

PERFORMANCE OF DISCONTINUOUS GALERKIN METHODS FOR ELLIPTIC PDE'S

PAUL CASTILLO *

Abstract. In this paper, we compare the performance of the main discontinuous Galerkin (DG) methods for elliptic partial differential equations on a model problem. Theoretical estimates of the condition number of the stiffness matrix are given for DG methods whose bilinear form is symmetric, which are shown to be sharp numerically. Then, the efficiency of the methods in the computation of both the potential and its gradient is tested on unstructured triangular meshes.

Key words. Finite elements, discontinuous Galerkin methods, conditioning of stiffness matrix

AMS subject classifications. 65N30

1. Introduction. In the last decade, several discontinuous Galerkin (DG) methods have been proposed for solving non-linear hyperbolic and convection dominated problems; see [8] for an introduction to the subject and [16] for an overview of the state of the art. They are preferred over standard continuous finite element methods because they provide high-order accurate approximations in contrast to traditional finite volume methods; they have a high degree of parallelism and since no inter-element continuity is imposed, polynomials of arbitrary degree can be used on different elements, making these methods suitable for *hp* refinement. Over the last years, there has been a tremendous interest in their application to problems where the diffusion is not negligible, and to pure elliptic problems.

Recently, Arnold, Brezzi, Cockburn and Marini [2] developed a unified framework in which theoretical stability analysis and optimal error estimates can be obtained for virtually all the existing DG methods. However, they do not discuss important issues that could be relevant to the practitioner. In this paper, which is part of my Ph.D. work [12], we complete the work presented in [2] by analyzing the methods from a practical point of view. The DG methods we consider are the following : the Babuška-Zlámal's penalty method [6], which is the simplest of all DG methods; the *Interior Penalty* (IP) method [18, 7, 24, 3, 4], which is one of the first symmetric DG methods for linear and non-linear parabolic problems with provable optimal error estimates; the steady state version of the so called *Local Discontinuous Galerkin* method (LDG) for purely elliptic problems [17, 13, 14], and, finally, a class of non-symmetric methods, called *Non-symmetric Interior Penalty Galerkin* (NIPG) [22, 23] which includes the method proposed by Baumann and Oden [5, 10].

We are interested in the quality and efficiency of the numerical approximation. We compare the above methods with respect to asymptotic behavior of the spectral condition number of the stiffness matrix, storage cost, rates of convergence and accuracy of the approximation of the potential and gradient. This comparison is carried out on a model elliptic problem with a smooth solution. The spectral condition number is analyzed numerically and theoretically for symmetric DG methods. The analysis strives in getting explicit expressions of the bounds in terms of the stabilization parameters of each method. For the non-symmetric methods, we perform a

*Scientific Computing Program, School of Mathematics, University of Minnesota, Vincent Hall, Minneapolis, MN 55455, castillo@math.umn.edu. The work of this author was partially supported by the National Science Foundation (Grant DMS-9807491) and by the University of Minnesota Supercomputer Institute.

numerical study of the spectral condition number as a function of the mesh size as well as of their stabilization parameters.

The organization of the paper is as follows. In section 2, we present the general formulation of a class of discontinuous Galerkin (DG) methods which contains all the DG methods considered in this paper. In section 3, we present a theoretical analysis of the spectral condition number for DG methods with symmetric bilinear forms. We also carry out a numerical study of the condition number in terms of the stabilization parameter of each method. In section 4, we describe the computational framework in which the comparisons are made. In section 5, we perform a comparison from a practical perspective. We analyze the storage cost and compute the rates of convergence of the LDG, IP and NIPG, as well as the accuracy of the potential and the gradients; unstructured meshes are used. Finally, we end in 6 with some concluding remarks.

2. General formulation of DG methods. We describe the formulation of a general discontinuous Galerkin method applied to the following elliptic model problem with Dirichlet boundary conditions.

$$-\Delta u = f, \quad \text{in } \Omega, \quad (2.1)$$

$$u = g, \quad \text{on } \partial\Omega, \quad (2.2)$$

where Ω is a bounded convex domain in \mathbb{R}^d .

We follow [13]; see also [2]. By introducing a new variable $\mathbf{q} = \nabla u$, we can rewrite our model problem as a system of the form

$$\mathbf{q} = \nabla u, \quad \text{in } \Omega, \quad (2.3)$$

$$-\nabla \cdot \mathbf{q} = f, \quad \text{in } \Omega, \quad (2.4)$$

$$u = g, \quad \text{on } \partial\Omega.$$

We must point out that the auxiliary variable \mathbf{q} can be eliminated from the equations which is usually not the case for classical mixed methods.

Let \mathcal{T}_h be a general triangulation of Ω . The weak formulation is obtained by multiplying equations (2.3) and (2.4) by smooth test functions \mathbf{r} and v , respectively on each element, T of \mathcal{T}_h . After integrating by parts we obtain the following weak formulation

$$\int_T \mathbf{q} \cdot \mathbf{r} = \oint_{\partial T} u \mathbf{r} \cdot \vec{n}_T - \int_T u \nabla \cdot \mathbf{r}, \quad (2.5)$$

$$\int_T \mathbf{q} \cdot \nabla v = \oint_{\partial T} v \mathbf{q} \cdot \vec{n}_T + \int_T f v, \quad (2.6)$$

where \vec{n}_T is the outward unit vector normal to the element T . Note that the above equations are well defined for any functions (u, \mathbf{q}) and (v, \mathbf{r}) in $\mathcal{V} \times \mathcal{M}$ where

$$\begin{aligned} \mathcal{V} &= \{u \in L^2(\Omega) : u|_T \in H^1(T), \forall T \in \mathcal{T}_h\}, \\ \mathcal{M} &= \{\mathbf{q} \in (L^2(\Omega))^d : \mathbf{q}|_T \in H^1(T)^d, \forall T \in \mathcal{T}_h\}. \end{aligned}$$

Next, we seek to approximate the exact solution (u, \mathbf{q}) with functions (u_h, \mathbf{q}_h) in the finite element space $\mathcal{V}_h \times \mathcal{M}_h \subset \mathcal{V} \times \mathcal{M}$, where

$$\begin{aligned} \mathcal{V}_h &= \{u \in L^2(\Omega) : u|_T \in \mathcal{P}_k(T), \forall T \in \mathcal{T}_h\}, \\ \mathcal{M}_h &= \{\mathbf{q} \in (L^2(\Omega))^d : \mathbf{q}|_T \in \mathcal{P}_k(T)^d, \forall T \in \mathcal{T}_h\}, \end{aligned}$$

and the *local* finite element space $\mathcal{P}_k(T)$ is the set of polynomials of degree at most k . Since for a given element T , the restrictions to T of u_h and of each of the components of \mathbf{q}_h belong to the *same* local space; this renders the coding of these methods considerably simpler than that of the standard mixed methods, especially for high-degree polynomial local spaces. The finite element solution (u_h, \mathbf{q}_h) is defined by using the aforementioned weak formulation by requiring that for all $T \in \mathcal{T}_h$, and for all $(v, \mathbf{r}) \in \mathcal{P}_k(T) \times \mathcal{P}_k(T)^d$ we have

$$\int_T \mathbf{q}_h \cdot \mathbf{r} = \oint_{\partial T} \widehat{u}_{h\{e,T\}} \mathbf{r} \cdot \vec{n}_T - \int_T u_h \nabla \cdot \mathbf{r}, \quad (2.7)$$

$$\int_T \mathbf{q}_h \cdot \nabla v = \oint_{\partial T} v \widehat{\mathbf{q}}_{h\{e,T\}} \cdot \vec{n}_T + \int_T f v, \quad (2.8)$$

where $\widehat{u}_{h\{e,T\}}$ and $\widehat{\mathbf{q}}_{h\{e,T\}}$ are the so called *numerical fluxes* which can be thought as being approximations to the traces of the function u and \mathbf{q} , respectively. We assume that these fluxes are local quantities in the sense that they depend only on the traces to the edge e of functions $u_{h\{T,K\}}$, $\mathbf{q}_{h\{T,K\}}$ and/or $\nabla u_{h\{T,K\}}$, where T and K are the elements sharing edge e . Moreover, we expect these functions to satisfy some basic properties such as *consistency*, that is, $\widehat{u}_{h\{e,T\}} = u$ and $\widehat{\mathbf{q}}_{h\{e,T\}} = \nabla u$, for a smooth function u , required in numerical methods for conservation laws and the so called *conservation* property, which can be formally defined as follows. Let T and K be the elements sharing edge e , a numerical flux $\widehat{\sigma}_{e,T}$ is conservative if

$$\widehat{\sigma}_{e,T} = \widehat{\sigma}_{e,K}.$$

By suitably choosing the numerical fluxes, we obtain the DG methods we are interested in, as shown in [2]. To define the fluxes we need to introduce some notation. Let e be an interior edge shared by elements T and K , we denote by \vec{n}_T and \vec{n}_K the outward unit normal vectors on e , relative to T and K respectively. For any function $v \in \mathcal{V}$, we define the jump, $[[v]]$, and the average, $\{\{v\}\}$, of v on an interior edge e , by

$$[[v]] = v|_K \vec{n}_K + v|_T \vec{n}_T \quad \text{and} \quad \{\{v\}\} = \frac{1}{2} (v|_K + v|_T).$$

For the boundary edges we simply define set $[[u]] = u|_T \vec{n}_T$ and $\{\{u\}\} = u|_T$. For any function $r \in \mathcal{M}$, the jump and the average are defined similarly,

$$[[\mathbf{r}]] = \mathbf{r}|_K \cdot \vec{n}_K + \mathbf{r}|_T \cdot \vec{n}_T \quad \text{and} \quad \{\{\mathbf{r}\}\} = \frac{1}{2} (\mathbf{r}|_K + \mathbf{r}|_T).$$

In Table 2.1, we show the definition of the numerical fluxes for the DG methods considered in this paper. Observe that, in Babuška-Zlámal's method, $\widehat{u}_{h\{e,T\}}$ is not conservative and $\widehat{\mathbf{q}}_{h\{e,T\}}$ is not consistent. Although optimal error estimates can still be obtained by using penalty terms of the order $O(h^{-(2p+1)})$, these type of methods are not suitable for practical computations, since the condition number of the stiffness matrix is proportional to $O(h^{-(2p+2)})$ which is especially bad when $p \gg 1$. In the IP method both numerical fluxes are consistent and conservative. The method is symmetric and achieves optimal rates of convergence for both the potential and the gradient, using, unlike Babuška-Zlámal's method, a penalization term independent of the approximation polynomial degree. However, the stabilization parameter η is mesh-dependent and must be chosen large enough to make the bilinear form coercive. As

for the IP, the LDG is a symmetric method and both numerical fluxes are consistent and conservative. Unlike the IP method, the LDG method is stable for $\eta > 0$. In the NIPG and Baumann-Oden's method, the numerical flux $\widehat{u}_{h\{e,T\}}$ is not conservative; this renders the bilinear form non-symmetric. Note that, Baumann-Oden's method does not require a penalty terms. However, this lack of stabilization is responsible for the suboptimality of the accuracy of the method. The NIPG method tries to fix this problem by including a penalty term which is of the order $O(h^{-\beta})$, $\beta = 1$ for the NIPG1 and $\beta = 3$ for NIPG3.

TABLE 2.1

Definition of the numerical fluxes for various DG methods, when using approximations of degree p

method	$\widehat{u}_{h\{e,T\}}$	$\widehat{\mathbf{q}}_{h\{e,T\}}$
Babuška-Zlámal	$u_{h _T}$	$-\frac{\eta}{h_e^{2p+1}}[[u_h]]$
IP	$\{\{u_h\}\}$	$\{\{\nabla u_h\}\} - \frac{\eta}{h_e}[[u_h]]$
LDG	$\{\{u_h\}\} + \beta_e \cdot [[u_h]]$	$\{\{\mathbf{q}_h\}\} - \beta_e \cdot [[\mathbf{q}_h]] - \frac{\eta}{h_e}[[u_h]]$
Baumann-Oden	$\{\{u_h\}\} + \vec{n}_T \cdot [u_h]$	$\{\{\nabla u_h\}\}$
NIPG1	$\{\{u_h\}\} + \vec{n}_T \cdot [u_h]$	$\{\{\nabla u_h\}\} - \frac{\eta}{h_e}[[u_h]]$
NIPG3	$\{\{u_h\}\} + \vec{n}_T \cdot [u_h]$	$\{\{\nabla u_h\}\} - \frac{\eta}{h_e^3}[[u_h]]$

3. Conditioning of the stiffness matrix. It is well known that, under certain conditions of the uniformity of the mesh, the spectral condition number of the stiffness matrix for the standard finite element method is of order $O(h^{-2})$ where h is the mesh size. In this section, we analyze the spectral condition number of the reduced stiffness matrix, i.e., the matrix obtained after the elimination of the auxiliary variable \mathbf{q}_h . We derive theoretical bounds for the DG methods with symmetric bilinear forms. The proof relies on a Poincaré-Friedrichs-type inequality; see Lemma 2.2 of Arnold [4]. For non-symmetric methods, namely, the method of Baumann-Oden [20] and the NIPG of Wheeler, [21, 22] and Süli *et al.* [23], a theoretical characterization of the spectrum is significantly difficult; this is why a numerical study of the spectral condition number is mandatory.

3.1. Preliminaries. For each element $T \in \mathcal{T}_h$, we denote by h_T the diameter of T and by ρ_T the diameter of the largest ball contained in T ; and we set $h = \max\{h_T : T \in \mathcal{T}_h\}$. We denote by \mathcal{E}_i the set of interior edges and by \mathcal{E} the set of all the edges, including the boundary edges. We assume that the triangulation is quasi regular, that is, there exists a positive constant σ such that

$$\forall T \in \mathcal{T}_h, \quad \frac{h_T}{\rho_T} \leq \sigma.$$

Then we have the following relation between the l_2 norm, $|\cdot|_\Omega$, and the L^2 norm $\|\cdot\|_{0,\Omega}^2$ in \mathcal{V}_h . There exist positive constant C^* and C^{**} that depend on d and σ , such that

for each $u \in \mathcal{V}_h$ we have

$$C^* h^d |u|_{\Omega}^2 \leq \|u\|_{0,\Omega}^2 \leq h^d C^{**} |u|_{\Omega}^2.$$

The spectral condition number $\kappa(A)$ of the stiffness matrix A can then be bounded by

$$\kappa(A) \leq R^* \frac{C_2}{C_1}, \quad (3.1)$$

where R^* is the ratio C^{**}/C^* and C_1, C_2 are positive constants, which can depend on the mesh size h , such that for all $u \in \mathcal{V}_h$, the bilinear form $\mathcal{A}(\cdot, \cdot)$ is bounded by

$$C_1 \|u\|_{0,\Omega}^2 \leq \mathcal{A}(u, u) \leq C_2 \|u\|_{0,\Omega}^2.$$

Thus, to obtain the bound (3.1) we prove the existence of the constants C_1 and C_2 . These are derived from the Poincaré-Friedrichs type of inequality Lemma 3.2, and the following standard inverse and trace inequalities.

LEMMA 3.1. *For any $u \in \mathcal{P}_k(T)$ there exist positive constants $C_1 = C_1(k, \sigma)$ and $C_2 = C_2(k, \sigma)$ such that*

$$\|\nabla u\|_{0,T}^2 \leq C_1 h_T^{-2} \|u\|_{0,T}^2, \quad (3.2)$$

$$\|u\|_{0,e}^2 \leq C_2 h_T^{-1} \|u\|_{0,T}^2. \quad (3.3)$$

As mentioned before, our results rely upon a Poincaré-Friedrichs type of inequality. This inequality will be crucial in proving the coercitivity of the bilinear form of several DG methods. Here, we want to slightly refine that result in order to obtain an explicit expression in terms of the parameters of a DG method.

LEMMA 3.2 (Lemma 2.2 in [4]). *Let u be a function in \mathcal{V} , for any positive number λ , there exist constants C_1 and C_2 such that*

$$\|u\|_{0,\Omega}^2 \leq C_\lambda \left(\|\nabla u\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{\lambda}{h_e} \llbracket u \rrbracket^2 \right),$$

where $C_\lambda = C_1 + \frac{\max\{1, h^2\}}{\lambda} C_2$; C_1 depends on Ω and C_2 depends on the minimum angle bound of the mesh.

Proof. We proceed as in [4]. Let $z \in H^2(\Omega) \cap H_0^1(\Omega)$ be the solution of the problem, $-\Delta z = u$ in Ω and $z = 0$ on $\partial\Omega$. It is clear that $(z, \nabla z)$ satisfies

$$\|u\|_{0,\Omega}^2 = \int_{\Omega} \nabla u \cdot \nabla z - \int_{\mathcal{E}} \llbracket u \rrbracket \nabla z,$$

then, after an application of Cauchy-Schwarz inequality we obtain

$$\|u\|_{0,\Omega}^2 \leq \left(\|\nabla u\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{\lambda}{h_e} \llbracket u \rrbracket^2 \right)^{1/2} \left(\|\nabla z\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{h_e}{\lambda} \left\| \frac{\partial z}{\partial n} \right\|_{0,e}^2 \right)^{1/2},$$

where λ is an arbitrary positive parameter. A simple computation gives

$$\|\nabla z\|_{0,\Omega}^2 \leq C_1 \|u\|_{0,\Omega}^2,$$

where $C_1 = C_1(\Omega)$ is the Poincaré constant. Moreover, using the following trace inequality; see [1, 4] :

$$\forall \phi \in H^2(T), \quad \left\| \frac{\partial \phi}{\partial n} \right\|_{0,e}^2 \leq C_2 \left(\frac{1}{h_e} |\phi|_{1,T}^2 + h_e |\phi|_{2,T}^2 \right),$$

where the constant $C_2 = C_2(\sigma)$ depends only on the minimal angle bound, we get

$$\|\nabla z\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{h_e}{\lambda} \left\| \frac{\partial z}{\partial n} \right\|_{0,e}^2 \leq C_\lambda \|u_h\|_{0,\Omega}^2,$$

where $C_\lambda = C_1 + \frac{\max\{1, h^2\}}{\lambda} C_2$, and C_1, C_2 have the above stated dependence. This completes the proof. \square

We are now ready to prove the estimates of the condition number for those methods with symmetric bilinear form. In subsection 3.2, we prove that for Babuška-Zlámal's penalty method, the condition number is of order $O(h^{-(2p+2)})$, where p is the approximation polynomial degree, showing that the method is not suitable for high order approximations. In subsections 3.3 and 3.4, we analyze the IP and the LDG methods, respectively. We show that both methods have a condition number that is of order $O(h^{-2})$, similar to the asymptotic behavior of the standard continuous finite element. Finally, we end this section with a numerical study of the condition number for the non-symmetric methods, subsection 3.5, where we show that the condition number is of order $O(h^{-2})$ for NIPG1 and Baumann-Oden's method or $O(h^{-4})$ for NIPG3.

3.2. Babuška-Zlámal's penalty method. To obtain the bilinear form we proceed as follows. First, observe that the auxiliary variable \mathbf{q}_h is equal to ∇u_h . Indeed, for any $u_h \in \mathcal{V}_h$, using integration by parts on the right term of equation (2.7) and the definition of $\widehat{\mathbf{u}}_{h\{e,T\}}$, we have

$$\forall \mathbf{r} \in \mathcal{M}_h, \quad \int_{\mathbf{T}} (\mathbf{q}_h - \nabla \mathbf{u}_h) \cdot \mathbf{r} = \oint_{\partial \mathbf{T}} (\widehat{\mathbf{u}}_{h\{e,T\}} - \mathbf{u}_h) \mathbf{r} \cdot \tilde{\mathbf{n}}_{\mathbf{T}} = \mathbf{0}.$$

Hence, $\mathbf{q}_h = \nabla u_h$. Inserting this expression in equation (2.8), and adding over all the elements, we obtain

$$\mathcal{A}(u_h, v_h) := \int_{\Omega} \nabla u_h \cdot \nabla v_h + \int_{\mathcal{E}} \frac{\eta}{h^{2p+1}} \llbracket u_h \rrbracket \llbracket v_h \rrbracket = \int_{\Omega} f v_h.$$

The corresponding energy norm $\|\cdot\|$ is given by

$$\|u_h\|^2 = \mathcal{A}(u_h, u_h) = \|\nabla u_h\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{\eta}{h^{2p+1}} \llbracket u_h \rrbracket^2. \quad (3.4)$$

The following result shows the dependence of the condition number of $\mathcal{A}(\cdot, \cdot)$ with the mesh size h .

THEOREM 3.3. *We have*

$$C_{\eta,h}^1 \|u_h\|_{0,\Omega}^2 \leq \mathcal{A}(u_h, u_h) \leq C_{\eta,h}^2 \|u_h\|_{0,\Omega}^2,$$

where $C_{\eta,h}^1 = 1 / \left(C_1^* + \frac{h^{2p}}{\eta} C_2^* \right)$ and $C_{\eta,h}^2 = \frac{1}{h^2} C_1 + \frac{\eta}{h^{2p+2}} C_2$ where C_1^*, C_2^*, C_1, C_2 are positive constants that depend on Ω, σ, k .

Proof. The lower bound constant $C_{\eta,h}^1$ can be obtained by using the Poincaré-Friedrichs inequality, Lemma 3.2, with $\lambda = \frac{\eta}{h^{2p}}$. We have

$$\|u\|_{0,\Omega}^2 \leq C_\lambda \left(\|\nabla u\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{\eta}{h^{2p+1}} \llbracket u \rrbracket^2 \right)$$

where $C_\lambda = C_1^* + \frac{h^{2p}}{\eta} C_2^*$, and $C_1^* = C(\Omega)$, $C_2^* = C(\sigma)$. Then set $C_{\eta,h}^1 = 1 / \left(C_1^* + \frac{h^{2p}}{\eta} C_2^* \right)$.

The upper bound constant $C_{\eta,h}^2$ can be easily obtained after an application of Lemmas (3.2) and (3.3) :

$$\mathcal{A}(u_h, u_h) \leq \left(\frac{C_1}{h^2} + \frac{\eta C_2}{h^{2p+2}} \right) \|u_h\|_{0,\Omega}^2,$$

where $C_1 = C(k, \sigma)$ and $C_2 = C(k, \sigma)$. Hence, let $C_{\eta,h}^2 = C_1/h^2 + \eta C_2/h^{2p+2}$. \square

Using (3.1), it is then easy to see that if A_h is the corresponding stiffness matrix, the spectral condition number $\kappa(A_h)$ is bounded by

$$\kappa(A_h) \leq R^* \frac{C_{\eta,h}^2}{C_{\eta,h}^1} = R^* \left(\frac{1}{h^2} C_1 + \frac{\eta}{h^{2p+2}} C_2 \right) \left(C_1^* + \frac{h^{2p}}{\eta} C_2^* \right) = O(h^{2p+2}).$$

3.3. The Interior Penalty method. We now consider the IP method. It can be shown that the energy norm $\|\cdot\|$ associated to this method is defined

$$\|u_h\|^2 = \mathcal{A}_h(u_h, u_h) = \|\nabla u_h\|_{0,\Omega}^2 - 2 \int_{\mathcal{E}} \llbracket u_h \rrbracket \{\!\{ \nabla u_h \}\!\} + \int_{\mathcal{E}} \frac{\eta}{h_e} \llbracket u_h \rrbracket^2.$$

THEOREM 3.4. *We have*

$$C_\eta^1 \|u_h\|_{0,\Omega}^2 \leq \mathcal{A}(u_h, u_h) \leq \frac{C_\eta^2}{h^2} \|u_h\|_{0,\Omega}^2,$$

where $C_\eta^2 = 1 / \left(C_1^* + \frac{C_2^*}{\eta - C_3} \right)$ and $C_\eta^1 = C_1 + C_2 \eta$, and the constants C_1, C_2, C_3 are generic constants that depend on Ω, σ, k .

This represents a major drawback for the IP method since the stabilization parameter η must be larger than the constant C_3 , which depends on the shape regularity of the mesh and the approximation polynomial degree and is difficult to find in practice.

Proof. Using the inequality, $ab \leq \frac{\epsilon}{2} a^2 + \frac{1}{2\epsilon} b^2$, for any $\epsilon > 0$, we have

$$\int_{\mathcal{E}} \llbracket u_h \rrbracket \{\!\{ \nabla u_h \}\!\} \leq \frac{\epsilon}{2} \int_{\mathcal{E}} \frac{1}{h} \llbracket u_h \rrbracket^2 + \frac{1}{2\epsilon} \int_{\mathcal{E}} h \|\{\!\{ \nabla u_h \}\!\}\|_{0,e}^2,$$

and by Lemma (3.2), we obtain

$$\int_{\mathcal{E}} \llbracket u_h \rrbracket \{\!\{ \nabla u_h \}\!\} \leq \frac{\epsilon}{2} \int_{\mathcal{E}} \frac{1}{h} \llbracket u_h \rrbracket^2 + \frac{C}{2\epsilon} \|\nabla u_h\|_{0,\Omega}^2, \quad (3.5)$$

where $C = C(k, \sigma)$ is a positive constant. Then,

$$\begin{aligned} \mathcal{A}(u_h, u_h) &\geq (1 - C/\epsilon) \|\nabla u_h\|_{0,\Omega}^2 + \int_{\mathcal{E}} \left(\frac{\eta - \epsilon}{h} \right) \llbracket u_h \rrbracket^2, \\ &\geq (1 - C/\epsilon) \left(\|\nabla u_h\|_{0,\Omega}^2 + \frac{\eta - \epsilon}{1 - C/\epsilon} \int_{\mathcal{E}} \frac{1}{h} \llbracket u_h \rrbracket^2 \right). \end{aligned}$$

The existence of the constant C_η^1 will follow by choosing η such that $C < \epsilon < \eta$. The lower bound is obtained by using Lemma 3.2, with $\lambda = (\eta - \epsilon) / (1 - C/\epsilon)$,

$$\mathcal{A}(u_h, u_h) \geq \frac{1 - C/\epsilon}{C_\lambda} \|u_h\|_{0,\Omega}^2.$$

Then $C_\eta^1 = \frac{1 - C/\epsilon}{C_\lambda}$, which is of the form $1 / (C_1^* + \frac{C_2^*}{\eta - \epsilon})$.

To obtain the upper bound, we proceed as follows :

$$\begin{aligned} \mathcal{A}_h(u_h, u_h) &\leq (1 + C) \|\nabla u_h\|_{0,\Omega}^2 + (1 + \eta) \int_\mathcal{E} \frac{1}{h} [[u_h]]^2, \quad \text{by 3.5, with } \epsilon = 1, \\ &\leq \frac{C_1}{h^2} \|u_h\|_{0,\Omega}^2 + \frac{C_2(1 + \eta)}{h^2} \|u_h\|_{0,\Omega}^2, \quad \text{by Lemmas 3.2 and 3.3,} \\ &\leq \frac{C_\eta^2}{h^2} \|u_h\|_{0,\Omega}^2, \end{aligned}$$

where $C_\eta^2 = C_1 + \eta C_2$, and $C_1 = C(k, \sigma)$, $C_2 = C(k, \sigma)$ are positive constants. \square

An upper bound for the spectral condition number of the stiffness matrix of the IP method is given by

$$\kappa(A_h) \leq R^* (C_1 + \eta C_2) \left(C_1^* + \frac{C_2^*}{\eta - C} \right) \frac{1}{h^2}, \quad (3.6)$$

where $C, C_1, C_2, C_1^*, C_2^*$ are positive constants.

In Fig 3.1, we show the condition number of the reduced stiffness matrix for the IP method as a function of the mesh size h . A given line represents the variation of the conditioning for a fixed polynomial degree. Using linear regression we find that the slopes of the lines are $-2.1360e + 00$ for $p = 1$, $-2.0603e + 00$ for $p = 2$ and $-2.0276e + 00$ for $p = 3$. Thus, the conditioning of the matrix behaves like $O(h^{-2})$ independently of the polynomial degree, showing that our analysis is sharp.

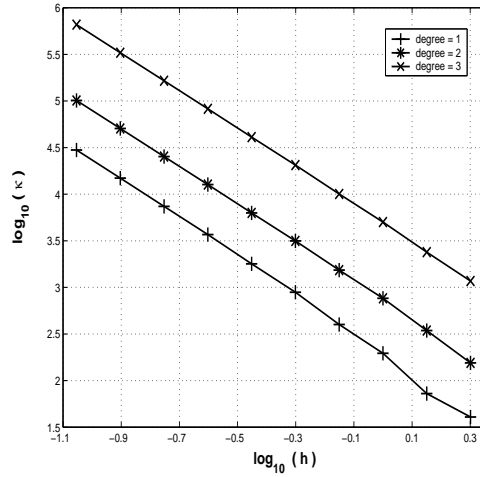


FIG. 3.1. Spectral condition number of the IP method as a function of the mesh size h

The upper bound (3.6) shows that there exists a mesh-dependent positive constant η_0 such that the spectral condition number $\kappa(\eta)$ has the following asymptotic behavior

$$\kappa(\eta) = \begin{cases} O(1/\eta - \eta_0), & \text{if } \eta - \eta_0 \ll 1, \\ O(\eta), & \text{if } \eta \gg 1. \end{cases}$$

In Fig 3.2, we show the function $\kappa(\eta)$ for linear and quadratic approximations. The plot is in logarithmic scale and has been shifted to the critical value $\eta_0 \approx 2.8666825$ for $p = 1$ and $\eta_0 \approx 6.9016989$ for $p = 2$. Observe that $\kappa(\eta)$ grows linearly as η goes to infinity and when η approaches to η_0 , that is $\log_{10}(\eta - \eta_0) \rightarrow -\infty$, the curve resembles a straight line with slope -1 , meaning that $\kappa(\eta) = O(1/\eta - \eta_0)$. This shows that our theoretical analysis is also sharp for the stabilization parameter η .

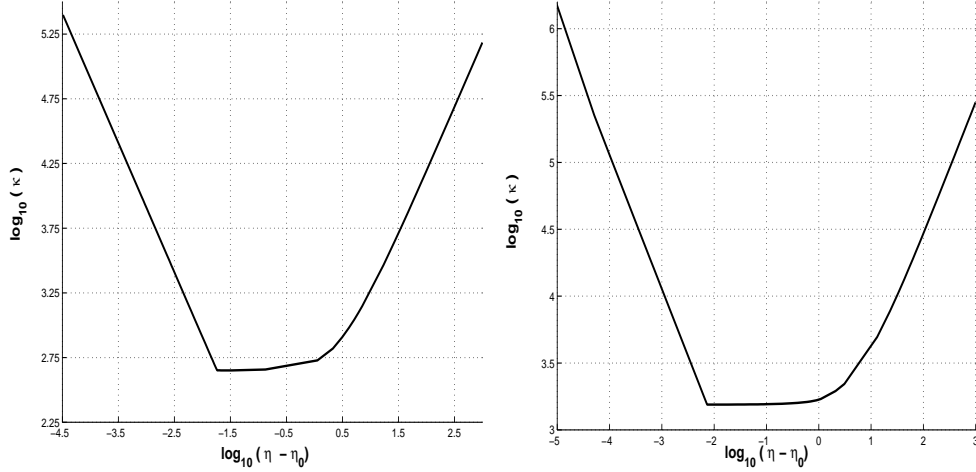


FIG. 3.2. Spectral condition number of the IP method, using linear (left) and quadratic (right) approximations as a function of the stabilization parameter η

3.4. The Local Discontinuous Galerkin method. The LDG method was introduced by Cockburn and Shu in [17] as a generalization of the DG method proposed by Bassi and Rebay [9] for the solution of the compressible Navier-Stokes equations. Recently, Castillo, Cockburn, Perugia and Schötzau [13] presented the first optimal *a priori* error estimates for the steady state version of the LDG on arbitrary meshes. Using the general formulation of a discontinuous Galerkin method, equations (2.7), (2.8), it can be shown that a general expression for the mesh-dependent energy norm $\|\cdot\|$, is given by

$$\|u_h\|^2 = \mathcal{A}_h(u_h, u_h) = \|\mathbf{q}_h\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{\eta}{h} \|[u_h]\|_{0,e}^2, \quad (3.7)$$

where the vector function \mathbf{q}_h belongs to \mathcal{M}_h and is defined by equation 2.7 in terms of u_h . In general, for any discontinuous Galerkin method defined by equations (2.7) and (2.8), we can view \mathbf{q}_h as an approximation of the gradient of the exact solution u . This approximation is of practical interest since it could lead to a more accurate approximation of the gradient or could even converge with higher rates of convergence than the standard piecewise gradient ∇u_h . Recently, Cockburn, Kanschat, Perugia and Schötzau [15] obtained a superconvergence result for the gradient on Cartesian grids, using \mathbf{q}_h as an approximation to ∇u . The following lemma establishes a connection between \mathbf{q}_h and ∇u_h .

LEMMA 3.5. For any function $u_h \in \mathcal{V}_h$, there exists a positive constant $C_\beta = C(\beta, k, \sigma)$ such that

$$\|\mathbf{q}_h - \nabla u_h\|_{0,\Omega}^2 \leq C_\beta \int_{\mathcal{E}_i} \frac{1}{h_e} \|[u_h]\|^2.$$

Proof. Since \mathbf{q}_h and ∇u_h belong to \mathcal{M}_h , we can take $\mathbf{r} = \mathbf{q}_h - \nabla u_h$, in equation (2.7). Then, after integrating by parts the right term of equation (2.7) and adding over all the elements, we obtain

$$\begin{aligned} \|\mathbf{q}_h - \nabla u_h\|_{0,\Omega}^2 &= \int_{\mathcal{E}_i} [(\widehat{u}_{h\{e,T\}} - u_h)(\mathbf{q}_h - \nabla u_h)] \\ &= \int_{\mathcal{E}_i} (\beta_e [\mathbf{q}_h - \nabla u_h] - \{\mathbf{q}_h - \nabla u_h\}) \cdot [u_h], \end{aligned}$$

since

$$[\mathbf{ur}] = \{\{u\}\}[\mathbf{r}] + [u]\{\{\mathbf{r}\}\}.$$

Using Cauchy-Schwarz inequality, we get

$$\|\mathbf{q}_h - \nabla u_h\|_{0,\Omega}^2 \leq \left(\int_{\mathcal{E}_i} h_e \|\beta_e [\mathbf{q}_h - \nabla u_h] - \{\mathbf{q}_h - \nabla u_h\}\|_{0,e}^2 \right)^{\frac{1}{2}} \left(\int_{\mathcal{E}_i} \frac{1}{h_e} [u_h]^2 \right)^{\frac{1}{2}}.$$

Finally, after applying the inverse inequality (3.3), we obtain the estimate

$$\|\mathbf{q}_h - \nabla u_h\|_{0,\Omega}^2 \leq C_\beta \int_{\mathcal{E}_i} \frac{1}{h_e} [u_h]^2.$$

where the constant C_β has the stated dependence. This completes the proof. \square

The following result shows that the spectral condition number of the stiffness matrix for the LDG method is bounded by $O(h^{-2})$.

THEOREM 3.6. *We have*

$$C_{\beta,\eta}^1 \|u_h\|_{0,\Omega}^2 \leq \mathcal{A}_h(u_h, u_h) \leq \frac{C_{\beta,\eta}^2}{h^2} \|u_h\|_{0,\Omega}^2,$$

where $C_{\beta,\eta}^1 = 1 / \left(C_\beta \max\{1, \frac{1}{\eta}\} \right)$ and $C_{\beta,\eta}^2 = C_1 + \eta C_2$, where C_β, C_1, C_2 are positive constants that depend on β, σ, k .

Proof. From the definition of the stabilizing term α , and the general expression of the energy norm \mathcal{A}_h , equation 3.7, we have

$$\mathcal{A}_h(u_h, u_h) = \|q_h\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{\eta}{h_e} [u_h]^2.$$

To obtain the upper bound, we proceed as follows :

$$\begin{aligned} \mathcal{A}_h(u_h, u_h) &\leq 2 \|\nabla u_h\|_{0,\Omega}^2 + 2 \|q_h - \nabla u_h\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{\eta}{h_e} [u_h]^2 \\ &\leq 2 \|\nabla u_h\|_{0,\Omega}^2 + C_\beta \int_{\mathcal{E}} \frac{1}{h_e} [u_h]^2 + \int_{\mathcal{E}} \frac{\eta}{h_e} [u_h]^2, \quad \text{by Lemma 3.5} \\ &\leq \frac{C_1}{h^2} \|u_h\|_{0,\Omega}^2 + (C_\beta + \eta) \frac{C_2}{h^2} \|u_h\|_{0,\Omega}^2, \quad \text{by Lemmas 3.2 and 3.3} \\ &\leq \frac{C_{\beta,\eta}^2}{h^2} \|u_h\|_{0,\Omega}^2, \end{aligned}$$

where the positive constant $C_{\beta,\eta}^2$ has the general form, $C_{\beta,\eta}^2 = C_1 + \eta C_2$. and the positive generic constant C_1 and C_2 have the following dependence $C_1 = C(k, \sigma, \beta)$ and $C_2 = C(k, \sigma)$.

To prove the lower bound we proceed as follows :

$$\begin{aligned} \|u\|_{0,\Omega}^2 &\leq C \left(\|\nabla u_h\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{1}{h_e} \llbracket u \rrbracket^2 \right), && \text{by Lemma 3.2 with } \lambda = 1 \\ &\leq C \left(2 \|q_h\|_{0,\Omega}^2 + 2 \|\nabla u_h - q_h\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{1}{h_e} \llbracket u \rrbracket^2 \right), \\ &\leq C \left(2 \|q_h\|_{0,\Omega}^2 + (1 + C_\beta) \int_{\mathcal{E}} \frac{1}{h_e} \llbracket u \rrbracket^2 \right), && \text{by Lemma 3.5} \\ &\leq C_2 \left(\|q_h\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{1}{h_e} \llbracket u \rrbracket^2 \right), \end{aligned}$$

where the positive constant $C_2 = \max\{2, 1 + C_\beta\}$. Then we have

$$\|u\|_{0,\Omega}^2 \leq C_2 \max\left\{1, \frac{1}{\eta}\right\} \left(\|q_h\|_{0,\Omega}^2 + \int_{\mathcal{E}} \frac{\eta}{h} \llbracket u_h \rrbracket^2 \right).$$

The lower bound constant follows by setting $C_{\beta,\eta}^1 = 1 / \left(C_\beta \max\left\{1, \frac{1}{\eta}\right\} \right)$. \square

The spectral condition number $\kappa(A_h)$, of the stiffness matrix A_h of the LDG method is bounded by

$$\kappa(A_h) \leq R^* \frac{C_{\beta,\eta}^2}{C_{\beta,\eta}^1} h^{-2} = R^* (C_1 + \eta C_2) \left(C_\beta \max\left\{1, \frac{1}{\eta}\right\} \right) h^{-2}. \quad (3.8)$$

In Fig 3.3, we show the condition number of the reduced stiffness matrix for the LDG as a function of the mesh size h . Each line represents the condition number using a different polynomial degree $p = 1, 2, 3$. Using linear regression we find that the slopes of the lines are $-2.1109e+00$ for $p = 1$, $-2.0334e+00$ for $p = 2$ and $-2.0207e+00$ for $p = 3$, thus the condition number varies like $O(h^{-2})$, independently of the polynomial degree as predicted by our theoretical analysis.

Note that, unlike the IP method, the LDG method is stable for any $\eta > 0$. Next we show that the theoretical upper bound obtained in (3.8) is sharp. In Fig 3.4, we plot the spectral condition number $\kappa(\eta)$ as a function of η on a logarithmic scale; we have used approximations of degree $p = 1, 2, 3$. Observe that the global asymptotic behavior is independent from the approximation polynomial degree. The stiffness matrix becomes ill-conditioned as η approaches to 0, meaning that the LDG method becomes less stable as expected. For $\eta \ll 1$, $\log \kappa(\eta)$ behaves like a straight line with slope -1 , that is $\kappa(\eta) = O(\frac{1}{\eta})$. Finally, for $\eta \gg 1$, the curve behaves like a straight line with slope 1, that is $\kappa(\eta) = O(\eta)$. Both asymptotic behaviors agree with our theoretical analysis.

3.5. The Non-symmetric Interior Penalty Galerkin methods . This class of DG methods includes the methods analyzed by Rivière, Wheeler and Girault [22], Baumann-Oden's method [10, 5, 11] and the discontinuous hp -finite element method of Houston, Schwab and Süli [23].

Unlike the IP method, Baumann-Oden's method is always well defined and weakly stable in the sense used in [2]. Unfortunately, this lack of stabilization has

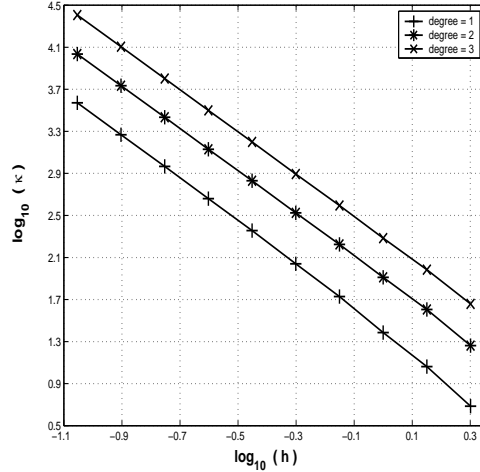


FIG. 3.3. Spectral condition number of the LDG method as a function of the mesh size h

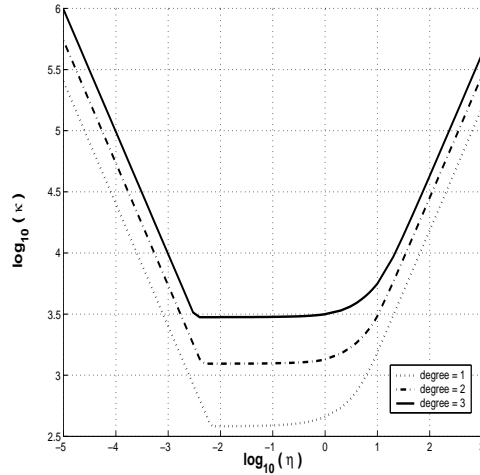


FIG. 3.4. Spectral condition number of the LDG method for approximations of degree $p = 1, 2, 3$.

a negative impact in the accuracy of the method since it has been observed that a loss in the order of convergence for polynomials of even degree occurs on meshes with quadrilateral elements [5]. In our experiments, using unstructured triangular meshes, we also observe this loss of accuracy; see section 5.3. The NIPG methods of Wheeler *et al.* [22] are a stabilization of Baumann-Oden's method. Here, we consider two cases; NIPG1 (i.e., $\beta = 1$, which corresponds to a stabilization term similar to that of the LDG and IP methods), and NIPG3 (i.e., $\beta = 3$, which is the minimal value that gives optimal rates of convergence in the L_2 norm for the potential. The analysis can be found in [22]).

In Table 3.1, we present the spectral condition number of the stiffness matrix for Baumann-Oden's method, NIPG1 and NIPG3 method, on a sequence of nested structured meshes. Since Baumann-Oden's method is stable for polynomial of degree greater or equal than 2, the column corresponding to $p = 1$ is left empty. We assume

the asymptotic behavior of the spectral condition number is of order $O(h^\alpha)$. In Table 3.2, we show a numerically estimate of α , which was obtained by linear regression. Thus, the spectral condition number of the NIPG1 and Baumann-Oden's method, is of order $O(h^{-2})$, and of order $O(h^{-4})$ for the NIPG3.

TABLE 3.1
Spectral condition number for a sequence of meshes

	$p = 1$	$p = 2$	$p = 3$
Baumann-Oden	—	7.8607e+01	2.0620e+02
	—	3.1155e+02	8.2943e+02
	—	1.2509e+03	3.3246e+03
	—	5.0113e+03	1.3318e+04
	—	2.0076e+04	5.3348e+04
NIPG1	8.0255e+02	1.7005e+03	2.6404e+03
	3.7107e+03	7.0816e+03	1.0785e+04
	1.5485e+04	2.8607e+04	4.3358e+04
	6.2639e+04	1.1471e+05	1.7365e+05
	2.5338e+05	4.6000e+05	6.9546e+05
NIPG3	1.3979e+03	2.9992e+03	4.8334e+03
	2.5128e+04	4.8332e+04	7.6398e+04
	4.1695e+05	7.7360e+05	1.2185e+06
	6.7348e+06	1.2375e+07	1.9481e+07
	1.0878e+08	1.9796e+08	3.1146e+08

TABLE 3.2
Numerical estimate of the order α of the conditioning number

	$p = 1$	$p = 2$	$p = 3$
Baumann-Oden	—	-2.0001e+00	-2.0036e+00
NIPG1	-2.0682e+00	-2.0177e+00	-2.0091e+00
NIPG3	-4.0562e+00	-4.0021e+00	-3.9946e+00

In Fig 3.5, we plot $\kappa(\eta)$ for the NIPG1 (left) and NIPG3 (right) methods for linear, quadratic and cubic approximations. Since the NIPG is a stabilization of Baumann-Oden's method, the spectral condition number should remain bounded as η goes to 0, for polynomials of degree greater or equal than 2. For linear approximations, Baumann-Oden's method is unstable, thus the condition number of the NIPG method should grow as η becomes smaller. Now, for large values of η , κ grows linearly with respect to η , for both values of β . This suggests an asymptotic behavior of the order

$O(\eta)$ for $\eta \gg 1$, similar to that of the IP and LDG methods.

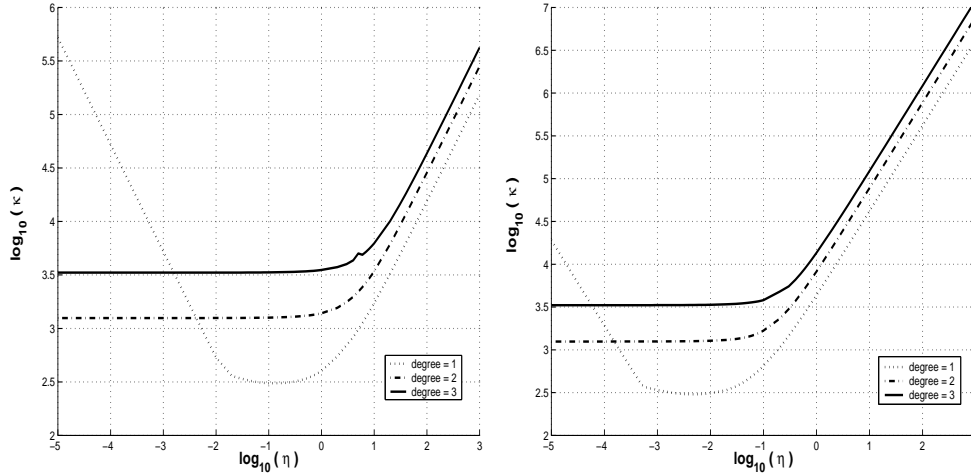


FIG. 3.5. Spectral condition number as a function of the stabilization parameter η for the NIPG method, NIPG1 $\beta = 1$ (left) and NIPG3 $\beta = 3$ (right)

4. Implementation issues. The implementation used for this paper is derived from the formulation of a general DG method, equations (2.7) and (2.8). This abstract formulation can be easily translated into a practical modular code by using an Object-Oriented Programming (OOP) paradigm. For this particular implementation we have used the C++ programming language. OOP provides a natural way for implementing a general class of DG methods in a single computational framework. By defining an abstract base class, we can define a general DG method, while keeping the implementation details out of the interfaces. In this abstract class, we define the common interfaces such as routines for the generation of the stiffness matrix and right hand side, routines for the computation of the gradients, and differential and edge flux operators. A particular method can now be implemented as an object derived from the abstract class. Since it inherits all the members (data and functions) from its base class, we avoid the redevelopment and testing of existing routines.

In order to have a uniform framework, we have used the same local polynomial space. Here, we have implemented the high-order orthogonal basis for triangular simplices proposed by Dubiner [19].

4.1. Stiffness matrix. The global linear system obtained from the general discontinuous formulation is a large sparse system which involves not only the potential but also the auxiliary variable \mathbf{q} as well. Moreover, the resulting system is indefinite. For these reasons it is preferable in practice to eliminate the auxiliary variable and solve only for the potential. We must point out that this elimination is possible only because of the particular choice of the numerical flux $\widehat{u}_{h\{e,T\}}$, which depends only on the variable u .

The assembly of the reduced stiffness matrix can be done either by computing a Schur complement from the global linear system or by exploiting the discontinuous formulation at the element level. The first approach is the classical procedure used in

mixed methods. First we obtain the matrix expression for equations (2.7) and (2.8) :

$$Mq_h + B_1u_h = f, \quad (4.1)$$

$$B_2q_h + Cu_h = g, \quad (4.2)$$

where the matrix M is a block diagonal mass matrix. The block structure of the global matrix is similar to that of mixed methods. Here, the stabilizing term is represented by the block matrix C . Since M is invertible, we can obtain an explicit expression for q_h from equation 4.1. Using this expression in the equation 4.2 we obtain the following linear system for the variable u_h

$$(C - B_2M^{-1}B_1)u_h = g - B_2M^{-1}f.$$

The second approach, which is the one we have used in our current implementation takes advantage of the block structure at the element level, making the assembly process more local. This approach is suitable for refinement since we only need to reassemble the block entries of those elements that have been marked for refinement and the elements in their stencil. The process is similar to the Schur complement approach, however, here, we consider the local expressions of the equations at each element. The stiffness matrix can then be stored using a block version of a standard scheme for storing sparse matrices such as Ellpack-Itpack scheme or compressed sparse row.

5. Comparison of methods. In this section we compare storage cost, condition number, rates of convergence and accuracy. To this end, we solve the model problem 2.1 with homogeneous boundary conditions in the convex domain, $\Omega = (-1, 1) \times (-1, 1)$. The right hand side is chosen such that the exact solution is the smooth function $u(x, y)$ given by

$$u(x, y) = 4(1 - x^2)(1 - y^2)e^{0.75(x+y)}.$$

We use restarted GMRES to solve the linear system. In order to obtain as much precision as possible, the stopping criterion is such that the relative residual norm is less than 10^{-13} .

5.1. Memory requirements. The size of the stencil provides an upper bound for the storage cost of the stiffness matrix and has a direct impact in parallel implementations. It is completely determined by the choice of the numerical flux $\widehat{q}_n|_{\{e, T\}}$. If this flux depends on the auxiliary variable q_h , an element interacts with its immediate neighbors and also with the neighbors of its neighbors, the maximum size of the stencil is 10 (in two dimensional domains). This is the case for LDG method. Otherwise, an element interacts only with its immediate neighbors, hence the possible maximum size for a stencil is 4, i.e. IP, Baumann-Oden and NIPG methods.

We measure the storage cost of the stiffness matrix in terms of the total number of its non-zero entries. Let s be the size of the stencil, d_k the dimension of the local polynomial space (we assume that the polynomial degree k is the same on each element) and n the total number of elements in the triangulation \mathcal{T}_h . Then, an upper bound for nnz is

$$nnz \leq nsd_k^2.$$

However, we must point out that this is a rough estimate and in practice the actual number will be below this bound. In Table 5.1 we show the ratio of cost relative to

the matrix with minimum storage for approximations of degree $p = 1, 2, 3$. Since the IP, NIPG and Baumann-Oden's have a compact stencil this ratio should be close to 1. However, for the LDG the ratio should be close to the theoretical bound $10/4 = 2.5$.

TABLE 5.1
Ratio.

	dofs	IP	LDG	BO	NIPG1	NIPG3
$p = 1$	16×3	1.000	1.571	—	1.000	1.041
	64×3	1.000	1.861	—	1.000	1.059
	256×3	1.000	2.007	—	1.000	1.068
	1024×3	1.000	2.081	—	1.000	1.073
	4096×3	1.000	2.117	—	1.000	1.075
$p = 2$	16×6	1.183	1.917	1.000	1.195	1.207
	64×6	1.185	2.288	1.000	1.202	1.219
	256×6	1.186	2.465	1.000	1.206	1.225
	1024×6	1.186	2.552	1.000	1.207	1.228
	4096×6	1.187	2.595	1.000	1.208	1.229
$p = 3$	16×10	1.090	1.792	1.000	1.106	1.110
	64×10	1.087	2.159	1.000	1.110	1.116
	256×10	1.086	2.336	1.000	1.112	1.119
	1024×10	1.085	2.423	1.000	1.113	1.120
	4096×10	1.085	2.466	1.000	1.114	1.121

5.2. Conditioning. In Table 5.2, we summarize the theoretical estimates obtained in section 3. We show the order of the condition number as well as the order of the penalization term of the method as function of the mesh size h and the polynomial degree p . We have also included the non-symmetric methods based on our numerical results. For small mesh size h , the penalization term dominates, thus for a penalization term of order $O(h^\alpha)$, the spectral condition number should be of order $O(h^{\alpha+1})$.

We consider the condition number as a function of the stabilization parameter η . In Fig 5.1, we have plotted the spectral condition number of the IP, LDG, NIPG1 and NIPG3 methods, for linear, quadratic and cubic approximations on a structured triangular mesh with 256 elements. For large values of η , the condition number of the IP, LDG and NIPG1 are asymptotically the same. This result is, to some extent, expected since their stabilization term is of the same order, i.e., $O(\frac{\eta}{h})$. This suggests that when no preconditioning is used, the IP method will take less CPU time than the LDG method, since performing a matrix-vector multiplication with the LDG is 2 to 2.5 times more expensive than with the IP. For the non-symmetric methods, the NIPG1 will outperform the NIPG3, since the condition number of the NIPG3 is significantly larger. However, we must keep in mind the loss of accuracy of the NIPG1 on polynomials of even degree.

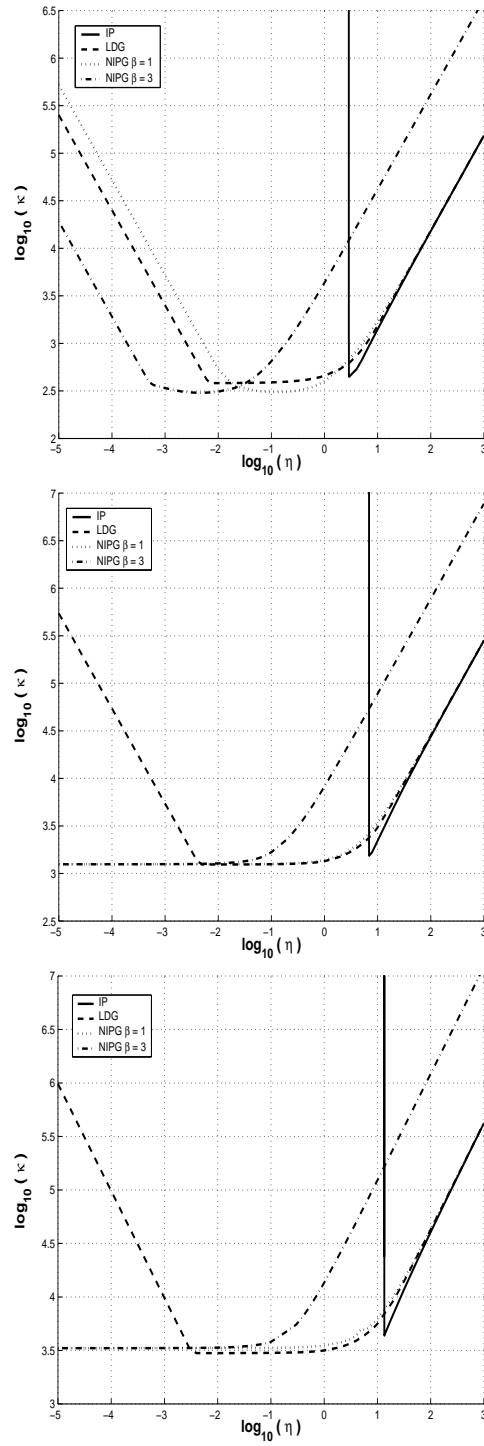


FIG. 5.1. Comparison of the spectral condition number $\kappa(\eta)$ as a function of the stabilization parameter η for linear (top), quadratic (middle) and cubic (bottom) approximations.

TABLE 5.2

Asymptotic behavior of the spectral condition number $\kappa(h)$ as a function of the mesh size, when using uniform approximations of degree p .

method	penalization	$\kappa(h)$
Babuška-Zlámal	$O(h^{-(2p+1)})$	$O(h^{-(2p+2)})$
IP	$O(h^{-1})$	$O(h^{-2})$
LDG	$O(h^{-1})$	$O(h^{-2})$
Baumann-Oden	no penalization	$O(h^{-2})$
NIPG1	$O(h^{-1})$	$O(h^{-2})$
NIPG3	$O(h^{-3})$	$O(h^{-4})$

5.3. Orders of convergence. We compare the orders of convergence for both the potential and the gradient. We used a sequence of structured triangular meshes $\{\mathcal{T}_i\}$, $i = 1, 2, \dots$, where the mesh-size parameter of \mathcal{T}_{i+1} is half the one of \mathcal{T}_i . If e_i denotes the error on mesh \mathcal{T}_i in the L^2 -norm then the numerical rate of convergence r_i , is defined as

$$r_i = \frac{\log_{10}(e_{i+1}/e_i)}{\log_{10}(0.5)}.$$

We consider two approximations of the gradient. The first, ∇u_h , is obtained by a direct application of the gradient operator to the computed solution u_h . The other approximation is derived from the general discontinuous formulation, using the definition of the auxiliary variable \mathbf{q}_h in equation 2.3. In Tables 5.3, 5.4 and 5.5, we present the numerical orders of convergence of the error in the potential u_h , gradient ∇u_h and gradient \mathbf{q}_h , respectively, for polynomials of degree 1 to 4. Observe that for the NIPG1 and Baumann-Oden's method, the rates of the potential are sub-optimal when polynomials of even degree are used. However, they are optimal for polynomials of odd degree. This behavior has also been reported by Baumann, Oden and Babuška, [5] on quadrilateral meshes. NIPG3 is the only non-symmetric method achieving optimal rates of convergence in both the potential and gradient. Both approximations of the gradient converge optimally for all the methods.

5.4. Effect of the stabilization parameter η . We carry out a numerical study of the effect of the stabilization parameter η on the quality of the approximation. Testing the accuracy of a method is a difficult if not an impossible task. Ideally, for a particular mesh, we would have to find the set of parameter which gives the minimal error. Here, we compare the ratio of the errors between the non-symmetric and the LDG method, as a function of the stabilization parameter η for linear, quadratic and cubic approximations. We have used three unstructured meshes of 248, 264 and 304 elements, which are shown in 5.2. In figures 5.3, 5.4 and 5.5, we show the ratio of the error in the potential, the gradient \mathbf{q}_h and piecewise gradient ∇u_h , respectively. It is clear that, when η is small, the LDG is more accurate than both variants of the

TABLE 5.3
Orders of convergence for the potential u_h .

method	p = 1	p = 2	p = 3	p = 4
IP	1.9004	2.9637	3.9283	4.9396
	1.9694	2.9809	3.9516	4.9593
	1.9895	2.9889	3.9724	4.9772
LDG	1.9027	2.9671	3.9397	4.9560
	1.9706	2.9830	3.9572	4.9672
	1.9900	2.9900	3.9749	4.9802
Baumann-Oden	—	2.8608	3.8608	4.7267
	—	2.4596	3.9390	4.2964
	—	2.1391	3.9781	4.0699
NIPG1	1.6906	2.8581	3.8524	4.7765
	1.8178	2.4510	3.9322	4.3458
	1.9067	2.1489	3.9755	4.0849
NIPG3	1.7900	2.7301	3.6508	5.0263
	1.9630	2.9338	3.9062	4.9425
	1.9924	2.9814	3.9678	4.9561

TABLE 5.4
Orders of convergence for q_h .

method	p = 1	p = 2	p = 3	p = 4
IP	0.9239	1.9495	2.9246	3.9179
	0.9733	1.9760	2.9594	3.9548
	0.9896	1.9890	2.9801	3.9778
LDG	0.9268	1.9550	2.9401	3.9442
	0.9749	1.9793	2.9667	3.9666
	0.9904	1.9905	2.9833	3.9827
Baumann-Oden	—	2.0745	3.1677	4.1756
	—	2.1127	3.1321	4.2221
	—	2.1150	3.0901	4.2111
NIPG1	1.1661	2.1794	3.2240	4.0711
	1.1599	2.1608	3.2056	4.0537
	1.1323	2.1403	3.1631	4.0734
NIPG3	0.8339	1.7886	2.6990	4.0246
	0.9718	1.9710	2.9510	3.9518
	0.9951	1.9957	2.9891	3.9882

NIPG, and consequently than Baumann-Oden's method. However, for large values of η all the methods, including the IP (not shown in the figures) have the same accuracy.

We have also compared \mathbf{q}_h and ∇u_h . We consider the ratio of the errors $R_{\nabla u}$ defined as follows

$$R_{\nabla u} = \frac{\|\nabla u - \nabla u_h\|_{\Omega}}{\|\nabla u - \mathbf{q}_h\|_{\Omega}}.$$

In Fig 5.6, we compare this ratio. For the IP method we only considered large values of η , for the LDG and NIPG we considered small values as well. In general, we

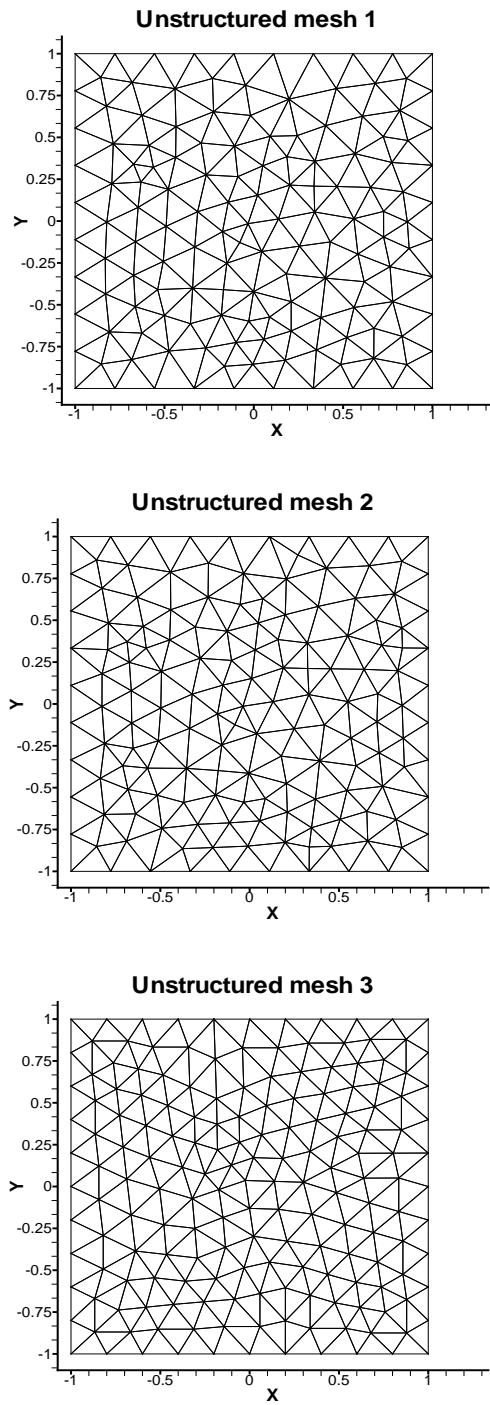


FIG. 5.2. *Unstructured meshes, 248, 264 and 304 elements.*

TABLE 5.5
Orders of convergence for ∇u_h .

method	p = 1	p = 2	p = 3	p = 4
IP	0.9458	1.9685	2.9806	3.9878
	0.9859	1.9879	2.9871	3.9878
	0.9960	1.9946	2.9926	3.9920
LDG	0.9462	1.9696	2.9816	3.9861
	0.9861	1.9885	2.9873	3.9871
	0.9961	1.9948	2.9926	3.9918
Baumann-Oden	—	2.0611	3.1524	4.2098
	—	2.0912	3.1114	4.1906
	—	2.0935	3.0698	4.1450
NIPG1	1.2214	2.1639	3.2161	4.0229
	1.1366	2.1557	3.1868	4.0420
	1.0780	2.1426	3.1369	4.0180
NIPG3	0.9212	1.9118	2.9647	4.0140
	0.9817	1.9733	2.9692	3.9636
	0.9956	1.9919	2.9884	3.9846

observe two regimes. For $\eta \ll 1$, q_h is more accurate for the LDG than for the NIPG methods. For $\eta \gg 1$ both approximations have the same accuracy. So in this case, the approximation obtained from the gradient operator is more efficient since it is completely local and offers the same accuracy than the auxiliary variable q_h .

6. Conclusions. In this paper, we present the first numerical comparison of DG methods for a model elliptic problem. We give a theoretical analysis of the behavior of the spectral condition number for methods with symmetric bilinear forms in terms of the parameters of the method, which has been shown to be sharp.

From our numerical experiments we can extract the following conclusions :

- The conditioning of the IP and LDG method is asymptotically of the same order as for the standard continuous case.
- The non-symmetric methods can only achieve optimal rates of convergence for the potential by using large penalty terms, hence increasing conditioning of the method from h^{-2} to h^{-4} , when β increases from 1 to 3. This can severely degrade the performance of the iterative method used for solving the linear system.
- Methods with symmetric discretizations, i.e, LDG, IP, have a better performance than those with non-symmetric discretizations. However, for large values of the stabilizing parameter all the methods have the same accuracy.
- The LDG requires at most 2.5 times the storage of the methods with compact stencil, such IP, Baumann-Oden and NIPG. This has a negative impact in the operation counts and parallel implementations.
- For larger values of η , the asymptotic behavior of the spectral condition number and accuracy of both symmetric methods, IP and LDG is the same. In this range of penalization parameter, since the LDG is a more expensive method, the IP is more efficient.
- The LDG method is more stable than the IP method since its stabilization parameter does not depend on the mesh nor on the approximation polynomial

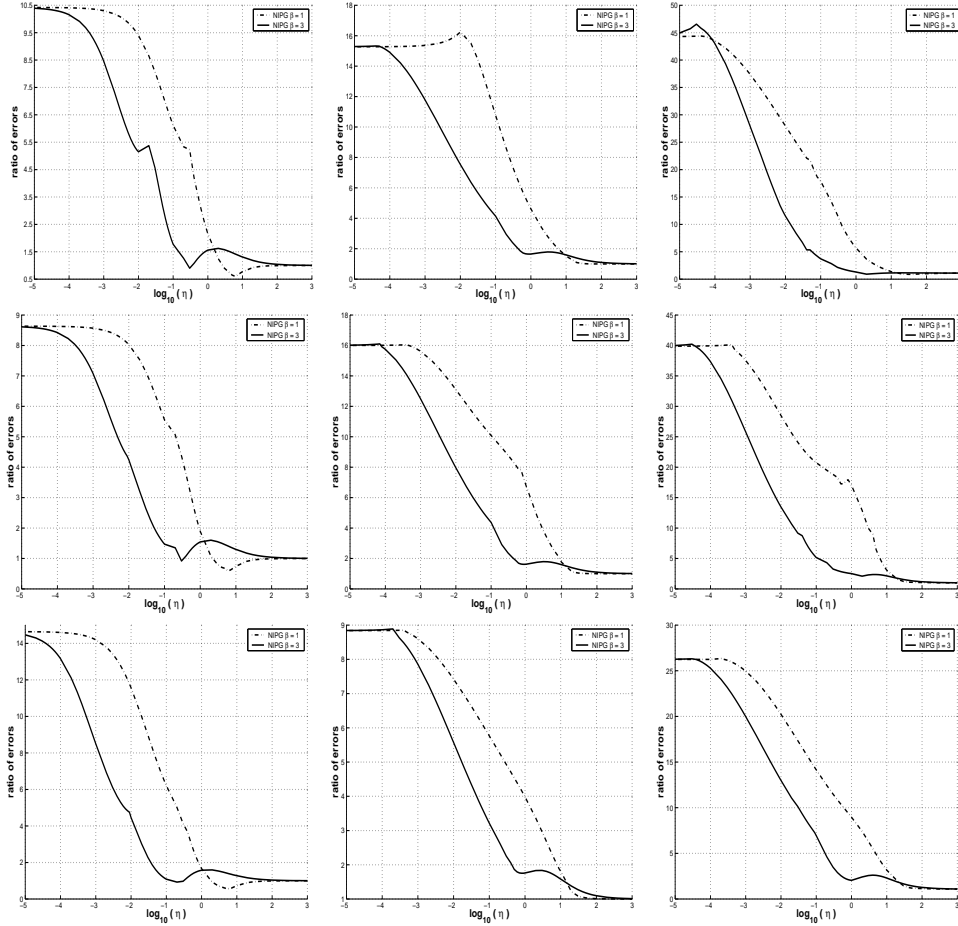


FIG. 5.3. Ratio of the errors for the potential NIPG : LDG. Row i corresponds to mesh i and column j to approximation of degree j .

degree.

- Although the LDG method becomes less stable as its stabilization parameter η approaches 0, the method is more accurate than the non-symmetric methods, like the Baumann-Oden's method.
- For large values of η , all the methods have similar accuracy in the potential and gradients \mathbf{q}_h and ∇u_h . However, the gradient \mathbf{q}_h of the IP and LDG methods is more accurate than the piecewise gradient ∇u_h for small values of η . For non-symmetric methods is the opposite.

7. Acknowledgements. The author would like to thank Prof. Bernardo Cockburn and Dr. Ilaria Perugia for insightful discussions.

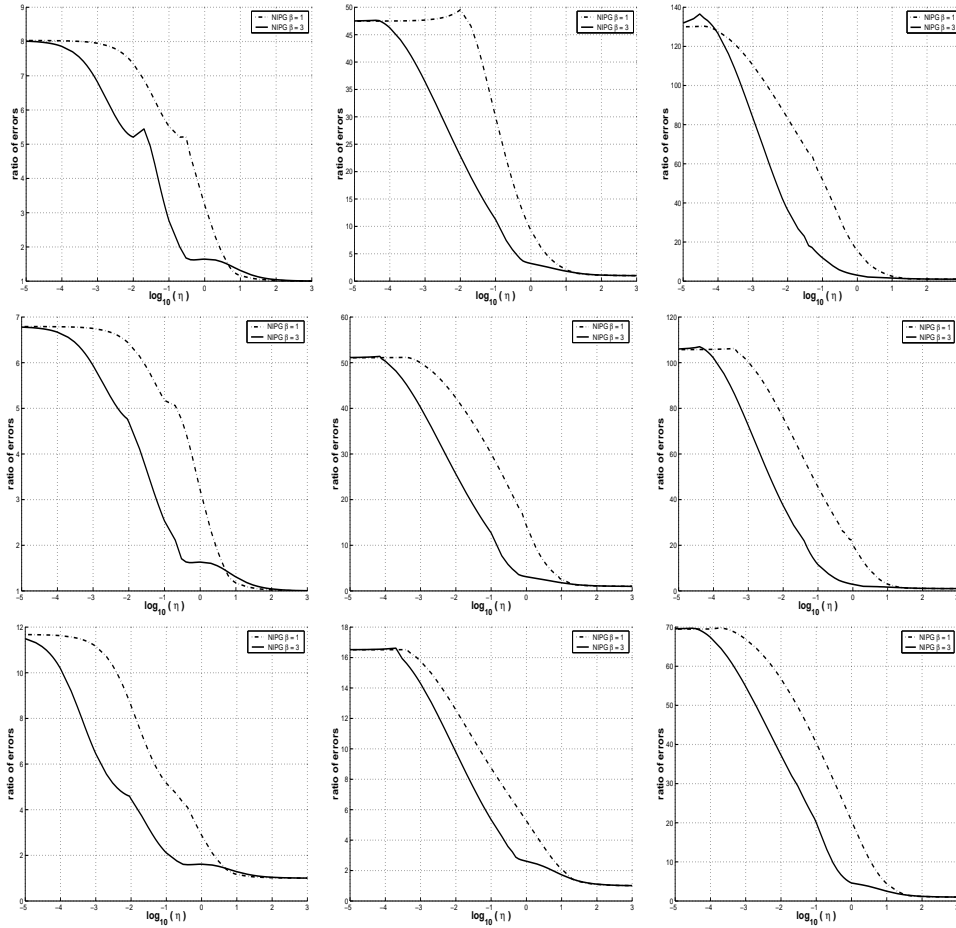


FIG. 5.4. Ratio of the errors for the gradient q NIPG : LDG. Row i corresponds to mesh i and column j to approximation of degree j .

REFERENCES

- [1] S. Agmon, *Lectures on elliptic boundary value problems*, Van Nostrand Mathematical Studies, Princeton, NJ, 1965.
- [2] D. Arnold, F. Brezzi, B. Cockburn, and D. Marini, *DG methods for elliptic problems*, Discontinuous Galerkin Methods: Theory, Computation and Applications (B. Cockburn, G.E. Karniadakis, and C.-W. Shu, eds.), Lecture Notes in Computational Science and Engineering, Springer Verlag, February 2000, pp. 89–101.
- [3] D.N. Arnold, *An interior penalty finite element method with discontinuous elements*, Ph.D. thesis, University of Chicago, Chicago, IL, 1979.
- [4] ———, *An interior penalty finite element method with discontinuous elements*, SIAM J. Num. Anal. **19** (1982), 742–760.
- [5] I. Babuška, C.E. Baumann, and J.T. Oden, *A discontinuous hp finite element method for diffusion problems: 1-D analysis*, Comput. Math. Appl. **37** (1999), 103–122.
- [6] I. Babuška and M. Zlámal, *Nonconforming elements in the finite element method with penalty*, SIAM J. Num. Anal. **10** (1973), 863–875.
- [7] G.A. Baker, *Finite element methods for elliptic equations using nonconforming elements*, Math. Comp. **31** (1977), 45–59.
- [8] T. Barth and H. Deconink (eds.), *Discontinuous Galerkin methods for convection-dominated problems*, Lecture Notes in Computational Science and Engineering, vol. 9, Springer Verlag,

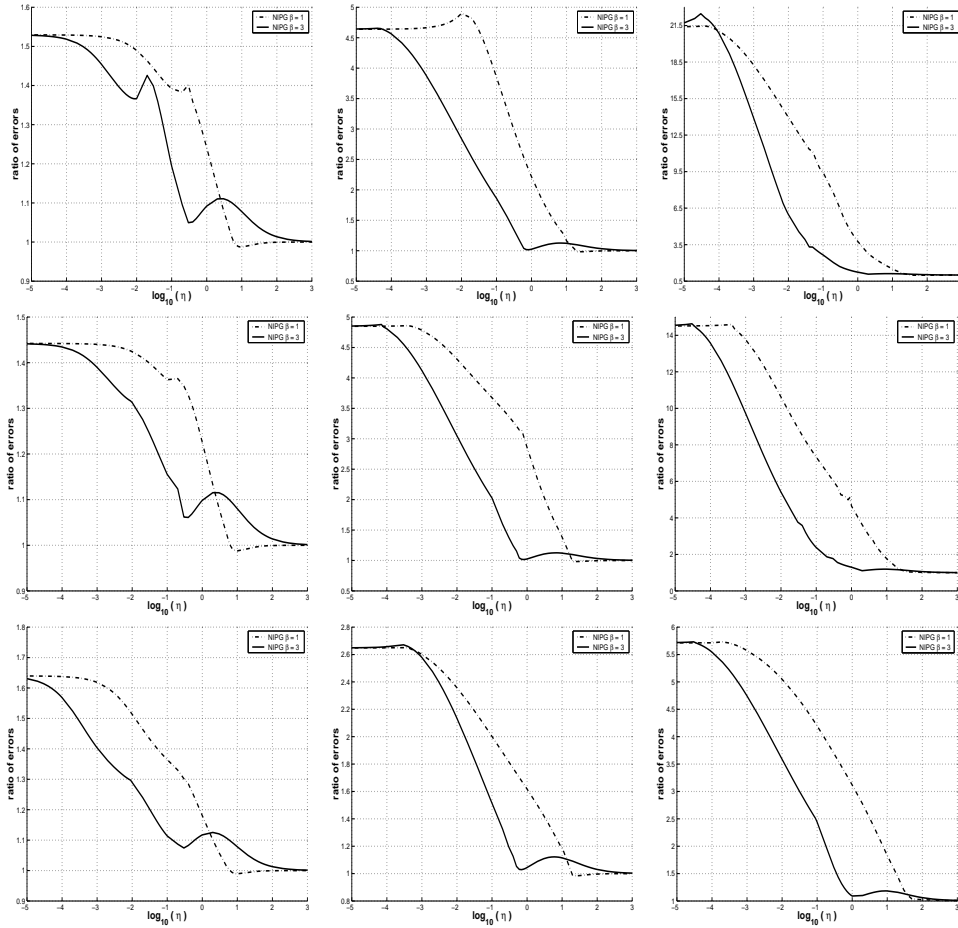


FIG. 5.5. Ratio of the errors for the gradient ∇u_h NIPG : LDG. Row i corresponds to mesh i and column j to approximation of degree j .

1999.

- [9] F. Bassi and S. Rebay, *A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations*, J. Comput. Phys. **131** (1997), 267–279.
- [10] C.E. Baumann, *An hp-adaptive discontinuous Galerkin method for computational fluid dynamics*, Ph.D. thesis, The University of Texas at Austin, 1997.
- [11] C.E. Baumann and J.T. Oden, *A discontinuous hp finite element method for convection-diffusion problems*, Comput. Methods Appl. Mech. Engrg., in press, special issue on Spectral, Spectral Element, and hp Methods in CFD, edited by G.E. Karniadakis, M. Ainsworth and C. Bernardi.
- [12] P. Castillo, *Local discontinuous Galerkin methods for convection-diffusion and elliptic problems*, Ph.D. thesis, University of Minnesota, Minneapolis, MN, 2001.
- [13] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau, *An a priori error analysis of the local discontinuous Galerkin method for elliptic problems*, SIAM J. Num. Anal. **38** (2000), no. 5, 1676–1706.
- [14] B. Cockburn and C. Dawson, *Some extensions of the local discontinuous Galerkin method for convection-diffusion equations in multidimensions*, Tech. Report Tech. Report 99-27, Texas Institute for Computational and Applied Mathematics, 1999.
- [15] B. Cockburn, G. Kanschat, I. Perugia, and D. Schötzau, *Superconvergence of the local discontinuous galerkin method for elliptic problems on cartesian grids*, SIAM J. Num. Anal., to

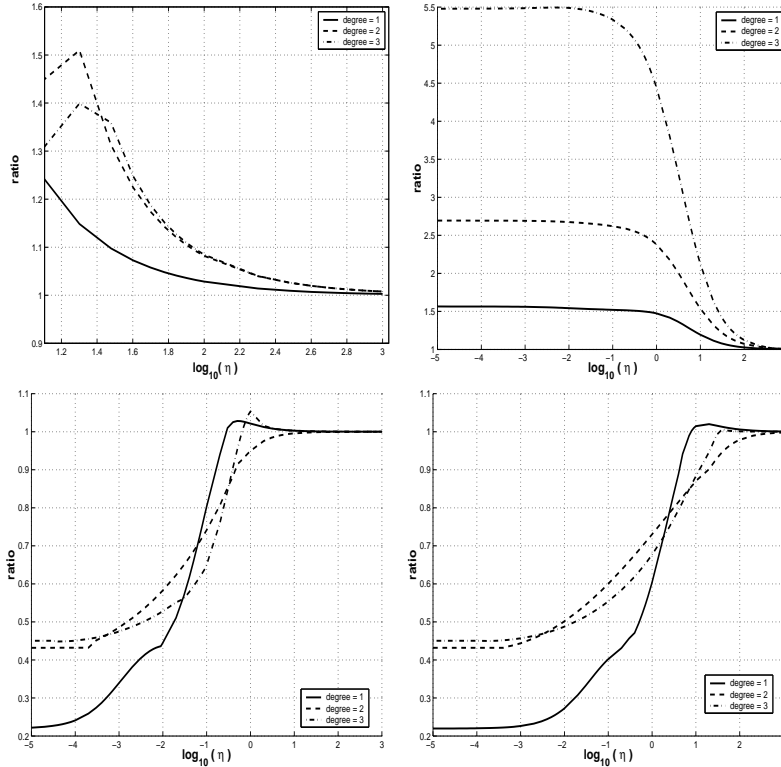


FIG. 5.6. Ratio $R_{\nabla u}$ for linear, quadratic and cubic approximations. Starting from top left in clockwise order, IP, LDG, NIPG1, NIPG3.

appear.

- [16] B. Cockburn, G.E. Karniadakis, and C.-W. Shu (eds.), *Discontinuous galerkin methods: Theory, computation and applications*, Lecture Notes in Computational Science and Engineering, vol. 11, Springer Verlag, February 2000.
- [17] B. Cockburn and C.W. Shu, *The local discontinuous Galerkin finite element method for convection-diffusion systems*, SIAM J. Num. Anal. **35** (1998), 2440–2463.
- [18] J. Douglas, Jr. and T. Dupont, *Interior penalty procedures for elliptic and parabolic Galerkin methods*, Lecture Notes in Physics, vol. 58, Springer-Verlag, Berlin, 1976.
- [19] M. Dubiner, *Spectral methods on triangles and other domains*, J. Sci. Comp. **6** (1991), 345–390.
- [20] J.T. Oden, Ivo Babuška, and C.E. Baumann, *A discontinuous hp finite element method for diffusion problems*, J. Comput. Phys. **146** (1998), 491–519.
- [21] B. Rivière and M.F. Wheeler, *Part I. Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems*, Tech. Report 99-09, TICAM, 1999.
- [22] B. Rivière, M.F. Wheeler, and V. Girault, *Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. Part I*, Tech. Report 99-09, TICAM, 1999.
- [23] E. Süli, C. Schwab, and P. Houston, *hp-DGFEM for partial differential equations with non-negative characteristic form*, Discontinuous Galerkin Methods: Theory, Computation and Applications (B. Cockburn, G.E. Karniadakis, and C.-W. Shu, eds.), Lecture Notes in Computational Science and Engineering, vol. 11, Springer Verlag, 2000, pp. 221–230.
- [24] M.F. Wheeler, *An elliptic collocation-finite element method with interior penalties*, SIAM J. Num. Anal. **15** (1978), 152–161.