AN EQUILIBRATED A POSTERIORI ERROR ESTIMATOR FOR
THE INTERIOR PENALTY DISCONTINUOUS GALERKIN
METHOD∗

D. BRAESS†, T. FRAUNHOLZ‡, AND R. H. W. HOPPE§

Abstract. Interior penalty discontinuous Galerkin (IPDG) methods for second order elliptic
boundary value problems have been derived from a mixed variational formulation of the problem.
Numerical flux functions across interelement boundaries play an important role in that theory. Resid-
ual type a posteriori error estimators for IPDG methods have been derived and analyzed by many
authors including the convergence analysis of the resulting adaptive schemes. Typically, the effectiv-
ity indices deteriorate with increasing polynomial order of the IPDG methods. The situation is more
favorable for a posteriori error estimators derived by means of the so-called hypercircle method.
Equilibrated fluxes are obtained by using a mixed method and an extension operator for Brezzi–
Douglas–Marini elements, and this can be done in the same way for all the discontinuous Galerkin
methods that fit into a known unified framework. The hypercircle method immediately provides the
reliability of the estimator, whereas its efficiency can be easily deduced from the efficiency of the
residual operators. In contrast to the residual-type estimators, the new estimators do not contain
unknown generic constants. Numerical results illustrate the performance of the suggested approach.

Key words. interior penalty discontinuous Galerkin method, a posteriori error estimation,
equilibration

AMS subject classifications. 65N30, 65N15, 65N50

DOI. 10.1137/130916540

1. Introduction. Residual-type error estimates were the first estimates that
have been studied in the a posteriori error analysis of discontinuous Galerkin (DG)
elements; see, e.g., [6, 20, 21, 22, 25]. They are more involved than the analogous
ones for conforming elements.

During the past decade, the hypercircle method also known as the Prager–Synge
theorem and the two-energies principle (cf. [7, section III.9], [9, 10]) and the method
of equilibrated fluxes (cf. [1, 2, 3, 4, 13, 14, 15, 16, 17, 18, 23, 24]) have attracted
a lot of interest. The advantage is that the error bounds do not contain (unknown)
generic constants. In this paper, we present a unified construction for all DG meth-
ods which are included in the general theory by Arnold et al. [5]. They present a
mixed variational approach that is equivalent to the primal variational formulation.
The specified numerical fluxes across the interelement boundaries are crucial in our
construction. As an example, we consider the interior penalty discontinuous Galerkin
(IPDG) method.

∗Received by the editors April 11, 2013; accepted for publication (in revised form) May 28, 2014;
†Faculty of Mathematics, Ruhr-University, D-44780 Bochum, Germany (Dietrich.Braess@rub.de).
‡Institut für Mathematik, Universität of Augsburg, D-86159 Augsburg, Germany (thomas.
fraunholz@math.uni-augsburg.de). This author was supported by the DFG within the Priority Pro-
gram SPP 1506.
§Department of Mathematics, University of Houston, Houston, TX 77204-3008 and Institut
für Mathematik, Universität of Augsburg, D-86159 Augsburg, Germany (rohop@math.uh.edu,
hoppe@math.uni-augsburg.de). This author was supported by NSF grants DMS-1115658, DMS-
1216857, by the DFG within the Priority Program SPP 1506, by the BMBF within the Collabora-
tive Research Projects “FROPT” and “MeFreSim,” and by the ESF within the Research Networking
Programme “OPTPDE.”
Moreover, \( \tau \in \text{length of } \) and

\[
2122 \quad \text{in a polygonal domain } \Omega \text{ consider the Poisson equation (2.1) for convenience, we}
\]

\[
-\Delta u = f \quad \text{in } \Omega,
\]

\[
u = 0 \quad \text{on } \Gamma,
\]

in a polygonal domain \( \Omega \subset \mathbb{R}^2 \) with homogeneous Dirichlet boundary conditions on \( \Gamma = \partial \Omega \). The extension to more general second order elliptic differential operators and boundary conditions can be accommodated.

Let \( T_h(\Omega) \) be a simplicial triangulation of the computational domain \( \Omega \). Given \( D \subset \Omega \), we denote by \( N_h(D) \) and \( E_h(D) \) the set of vertices and edges of \( T_h(\Omega) \) in \( D \), and we refer to \( P_k(D) \), \( k \in \mathbb{N} \), as the set of polynomials of degree \( \leq k \) on \( D \). Moreover, \( h_K, K \in T_h(\Omega) \), and \( h_E, E \in E_h(\Omega) \), stand for the diameter of \( K \) and the length of \( E \), respectively, and \( h := \max(h_K \mid K \in T_h(\Omega)) \). We consider the finite element approximation with the DG spaces

\[
\text{(2.2a)} \quad V_h := \{ v_h \in L^2(\Omega) \mid v_h|_K \in P_k(K), K \in T_h(\Omega) \},
\]

\[
\text{(2.2b)} \quad V_h := \{ \tau_h \in L^2(\Omega)^2 \mid \tau_h|_K \in P_k(K)^2, K \in T_h(\Omega) \}.
\]

For \( E \in E_h(\Omega) \), \( E = K_+ \cap K_- \), \( K_+ \in T_h(\Omega) \), and \( v_h \in V_h \), we denote the average and jump of \( v_h \) across \( E \) by \( \{ v_h \}_E \) and \( [v_h]_E \), i.e.,

\[
\{ v_h \}_E := \frac{1}{2} \left( v_h|_{E \cap K_+} + v_h|_{E \cap K_-} \right), \quad [v_h]_E := v_h|_{E \cap K_+} - v_h|_{E \cap K_-}.
\]
whereas for $E \in \mathcal{E}_h(\Gamma)$ we set

$$\{v_h\}_E := v_h|_E, \quad [v_h]_E := v_h|_E.$$  

We follow the general scheme of DG methods in the mixed formulation as in [5] rather than the equivalent primal variational formulations. The finite element approximation of the Poisson equation with homogeneous Dirichlet boundary conditions amounts to the computation of $(u_h, \sigma_h) \in \mathcal{V}_h \times \mathcal{V}_h$ such that for all $(v, \tau) \in \mathcal{V}_h \times \mathcal{V}_h$

$$\sum_{K \in T_h(\Omega)} \int_K \sigma_h \cdot \tau \, dx = \sum_{K \in T_h(\Omega)} \left( - \int_K u_h \div \tau \, dx + \int_{\partial K} \tilde{u}_{\partial K} \nu_K \cdot \tau \, ds \right),$$

$$\sum_{K \in T_h(\Omega)} \int_K \sigma_h \cdot \nabla v \, dx = \sum_{K \in T_h(\Omega)} \left( \int_K f v \, dx + \int_{\partial K} \tilde{\sigma}_{\partial K} \cdot \nu_K v \, ds \right),$$

where $\nu_K$ stands for the exterior normal unit vector on $\partial K$.

The definition of the DG method is completed by a rule for computing $\tilde{u}_{\partial K}$ and the numerical fluxes $\tilde{\sigma}_{\partial K}$, and each DG method is characterized by the associated definition. In particular, the IPDG method is obtained by the specification

$$\tilde{u}_{\partial K}|_E := \{u_h\}_E, \quad \tilde{\sigma}_{\partial K}|_E := \{\text{grad } u_h\}_E - \alpha h^{-1}_E [u_h]_E \nu, \quad E \in \mathcal{E}_h(\overline{\Omega}),$$

where $\alpha > 0$ is a penalty parameter, and $\alpha = 2.5(k+1)^2$ is considered as a convenient choice [19]. Here and below, for $E \in \mathcal{E}_h(\Omega), E = K_+ \cap K_-, K_+ \in T_h(\Omega)$, the unit normal vector $\nu$ is defined such that its product with a jump $[\cdot]|_E$ is independent of the orientation of $E$, i.e., $[u_h]_E \nu = u_h|_{K_+} \nu_{K_+} + u_h|_{K_-} \nu_{K_-}$.

For completeness, we recall the connection of the primal variational formulation and the mixed method for IPDG. Choosing $\tau = \nabla v$ in (2.3a), using the integration by parts formula

$$\int_K u_h \div \nabla v \, dx = - \int_K \nabla u_h \cdot \nabla v \, dx + \int_{\partial K} u_h \nu_K \cdot \nabla v \, ds,$$

and eliminating $\sigma_h$ from (2.3a), (2.3b) we obtain the primal variational formulation of the IPDG method: Find $u_h \in \mathcal{V}_h$ such that for all $v \in \mathcal{V}_h$

$$\sum_{K \in T_h(\Omega)} \int_K \nabla u_h \cdot \nabla v \, dx - \sum_{E \in \mathcal{E}_h(\overline{\Omega})} \int_E \left( \nu \cdot \{\nabla u_h\}_E [v]_E + [u_h]_E \nu \cdot \{\nabla v\}_E \right) \, ds$$

$$+ \alpha \sum_{E \in \mathcal{E}_h(\overline{\Omega})} h^{-1}_E \int_E [u_h]_E v_E \, ds = \sum_{K \in T_h(\Omega)} \int_K f v \, dx.$$  

Conversely, if $u_h \in \mathcal{V}_h$ solves (2.6), define $\sigma_h \in \mathcal{V}_h$ by

$$\sum_{K \in T_h(\Omega)} \int_K \sigma_h \cdot \tau \, dx = \sum_{K \in T_h(\Omega)} \left( - \int_K u_h \div \tau \, dx + \int_{\partial K} \tilde{u}_{\partial K} \nu_K \cdot \tau \, ds \right), \quad \tau \in \mathcal{V}_h.$$  

By setting $\tau = \nabla v, v \in \mathcal{V}_h$, and using the integration by parts formula (2.5) along with (2.6), it follows that the pair $(u_h, \sigma_h) \in \mathcal{V}_h \times \mathcal{V}_h$ satisfies (2.3a), (2.3b), (2.3a), (2.3b), and (2.4) are equivalent to (2.6).
3. An interpolation by BDM elements. The numerical fluxes \( \hat{\sigma} \) that live on the interelement boundaries will be extended to the elements by an interpolation. The finite element space for the fluxes is the BDM element, where \( \text{BDM}_k(K), k \in \mathbb{N}, \) is given by
\[
\text{BDM}_k(K) = P_k(K)^2, \quad \dim \text{BDM}_k(K) = (k+1)(k+2).
\]
We refer to \( \lambda^K_i, 1 \leq i \leq 3, \) as the barycentric coordinates of \( K \in \mathcal{T}_h(\Omega) \) and denote by \( b_K \) the element bubble function \( b_K := \lambda^K_1 \lambda^K_2 \lambda^K_3. \) By (3.41) in [11, p. 125] any \( q_K \in \text{BDM}_k(K) \) is uniquely determined by the following degrees of freedom (DOF):
\[
\begin{align*}
\int_E q_K \cdot \nu_K p_k ds, & \quad p_k \in P_k(E), \ E \in \mathcal{E}_h(\partial K), \\
\int_K q_K \cdot \text{grad} p_{k-1} dx, & \quad p_{k-1} \in P_{k-1}(K), \\
\int_K q_K \cdot \text{curl}(b_K p_{k-2}) dx, & \quad p_{k-2} \in P_{k-2}(K).
\end{align*}
\]
A standard scaling argument yields a bound of the \( L^2 \) norm when a BDM element is interpolated with these data.

**Lemma 3.1.** There exists a constant \( c \) that depends only on \( k \) and the shape regularity of \( \mathcal{T}_h \) such that for each \( q_K \in \text{BDM}_k(K) \)
\[
\int_K q_K^2(x) dx \leq c \left( h_K \int_{\partial K} (q_K \cdot \nu_K)^2 ds \right.
\]
\[
+ h_K^2 \max \left\{ \int_K (q_K \text{grad} p)^2 dx; \ p \in P_{k-1}, \ \max_{x \in K} |p(x)| \leq 1 \right\}
\]
\[
+ h_K^2 \max \left\{ \int_K (q_K \cdot \text{curl}(b_K p))^2 dx; \ p \in P_{k-2}, \ \max_{x \in K} |p(x)| \leq 1 \right\}.
\]

**Remark 3.2.** For \( k = 1, q_K \in \text{BDM}_1(K) \) is uniquely determined by the DOF on \( \partial K; \) cf. Figure 1.

**Remark 3.3.** A BDM element may be specified by \( \text{div} q_K \) instead of (3.2b). Therefore, the bound in Lemma 3.1 can be replaced by
\[
\int_K q_K^2(x) dx \leq c \left( h_K \int_{\partial K} (q_K \cdot \nu_K)^2 ds + h_K^2 \int_K (\text{div} q_K)^2 dx \right.
\]
\[
+ h_K^2 \max \left\{ \int_K (q_K \cdot \text{curl}(b_K p))^2 dx; \ p \in P_{k-2}, \ \max_{x \in K} |p(x)| \leq 1 \right\}.
\]
Moreover, we will refer to the following lemma.

**Lemma 3.4.** There exists a constant $c$ that depends only on $k$ and the shape regularity of $T_h$ such that for each $q_K \in BDM_k(K)$

$$\|q_K \cdot \nu_K\|_{0,\partial K} \leq c h_K^{-1/2} \|q_K\|_{0,K}.$$  

This inequality follows from the fact that

$$(3.3) \inf_{\|q_K \cdot \nu_K\|_{0,\partial K} = 1} \|q_K\|_{0,K} > 0.$$  

The constant $c$ depends on the degree $k$, since (3.3) is not true, if we take the infimum over all $H^1$ functions.

4. Application of the hypercircle method to nonconforming finite elements. The starting point is the theorem of Prager and Synge [7, 27] that is also called the two-energies principle. We restrict ourselves to the Poisson equation; the generalization to other elliptic problems can be found in [7, Chap. III, section 9].

**Theorem 4.1** (theorem of Prager and Synge, two-energies principle). Let $\sigma \in H(\text{div}, \Omega)$ and $v \in H^1_0(\Omega)$. Furthermore, let $u$ be the solution of (2.1). If $\sigma$ satisfies the equilibrium condition

$$(4.1) \text{div} \sigma + f = 0,$$

then,

$$|u - v|^2_{1,\Omega} + \|\text{grad} u - \sigma\|^2_{0,\Omega} = \|\text{grad} v - \sigma\|^2_{0,\Omega}.$$  

**Supplement.** Let $J(v) := \frac{1}{2} \|\text{grad} v\|^2_{0,\Omega} - \int_\Omega f v dx$ and $J^c(\sigma) := \frac{1}{2} \|\sigma\|^2_{0,\Omega}$ denote the (direct) energy and the complementary energy, respectively. If the assumptions above hold, then

$$|u - v|^2_{1,\Omega} + \|\text{grad} u - \sigma\|^2_{0,\Omega} = 2J(v) + 2J^c(\sigma).$$  

A proof is provided, e.g., in [7].

It is an advantage of the two-energies principle that the function $v$ in Theorem 4.1 may be an auxiliary function which does not arise from a variational problem. The piecewise gradient of a finite element function $v_h$ in the broken $H^1$ space will be denoted as $\text{grad}_h v_h$. We have $\text{grad}_h v_h \in L^2(\Omega)$.

**Corollary 4.2.** Let $u_h$ be the finite element solution of a nonconforming method in a broken $H^1$ space, e.g., a DG method. Assume that an auxiliary function $u_h^{\text{conf}} \in H^1(\Omega)$ that satisfies the Dirichlet boundary condition, is obtained by postprocessing. Moreover, let $\sigma_h^{eq} \in H(\text{div}, \Omega)$ be a flow that satisfies the equilibrium condition (4.1). Then we have the estimate

$$\|\text{grad} u - \text{grad}_h u_h\|_{0,\Omega} \leq \|\text{grad}_h u_h^{\text{conf}} - \sigma_h^{eq}\|_{0,\Omega} + \|\text{grad}_h u_h - \text{grad}_h u_h^{\text{conf}}\|_{0,\Omega}$$  

$$(4.2) \leq \|\text{grad}_h u_h - \sigma_h^{eq}\|_{0,\Omega} + 2\|\text{grad}_h u_h - \text{grad}_h u_h^{\text{conf}}\|_{0,\Omega}.$$  

Indeed, the two-energies principle yields the following bound for the auxiliary function $u_h^{\text{conf}}$:

$$\|\text{grad} u - \text{grad}_h u_h^{\text{conf}}\|_{0,\Omega} \leq \|\text{grad}_h u_h^{\text{conf}} - \sigma_h^{eq}\|_{0,\Omega}.$$  

By applying the triangle inequality twice, we obtain (4.2).
The corollary was implicitly used in [2, 8].

In actual computations, we have frequently an additional term due to data oscillations. We only have the equilibration for an approximate function \( \bar{f} \), i.e.,

\[
\text{div } \sigma + \bar{f} = 0
\]

and \( \int_K (\bar{f} - f) dx = 0, \ K \in \mathcal{T}_h(\Omega) \). The difference between the solutions of the Poisson equations for \( f \) and for \( \bar{f} \) will be estimated by the following lemma.

**Lemma 4.3.** Let \( \mathcal{T}_h(\Omega) \) be a triangulation of \( \Omega \), \( h_K \) denote the length of the longest side of \( K \in \mathcal{T}_h(\Omega) \). If \( g \in L^2(\Omega) \) satisfies

\[
\int_K g dx = 0, \ K \in \mathcal{T}_h(\Omega),
\]

then the solution of \( -\Delta v = g \) in \( \Omega \) with zero Dirichlet and/or Neumann boundary conditions satisfies

\[
\|\nabla v\|_{0,\Omega} \leq \frac{1}{\pi} \left( \sum_{K \in \mathcal{T}_h(\Omega)} h_K^2 \|g\|_{0,K}^2 \right)^{1/2}.
\]

A proof is given in [26]; see also [1]. Extra terms of the form (4.4) with \( g := f - \bar{f} \) will cope with the data oscillation.

**5. Equilibration.** Let \( \bar{f} \) be the \( L^2 \)-projection of \( f \) onto piecewise polynomials of degree \( k - 1 \), i.e.,

\[
\int_K \bar{f} v dx = \int_K f v dx, \ v \in P_{k-1}(K).
\]

We construct a flux \( \hat{\sigma}_K \in \text{BDM}_k(K) \) by the specifications

\[
\hat{\sigma}_K|_{\partial K} = \hat{\sigma}_{\partial K},
\]

\[
\int_K \hat{\sigma}_K \cdot \text{grad } p_{k-1} dx = \int_K \sigma_h \cdot \text{grad } p_{k-1} dx, \ p_{k-1} \in P_{k-1}(K),
\]

\[
\int_K \hat{\sigma}_K \cdot \text{curl}(b_K p_{k-2}) dx = \int_K \sigma_h \cdot \text{curl}(b_K p_{k-2}) dx, \ p_{k-2} \in P_{k-2}(K).
\]

The first equation corresponds to (3.2a) and shows that the flux is an extension of the numerical flux that is originally defined on the element boundaries. Now, it follows from (5.2b), Gauss’ theorem, (5.1), and the DG finite element equation (2.3b) that

\[
\int_K \text{div } \hat{\sigma}_K p_{k-1} dx = -\int_K \hat{\sigma}_K \cdot \text{grad } p_{k-1} dx + \int_{\partial K} \hat{\sigma}_K \cdot \nu_K p_{k-1} ds
\]

\[
= -\int_K \sigma_h \cdot \text{grad } p_{k-1} dx + \int_{\partial K} \hat{\sigma}_{\partial K} \cdot \nu_K p_{k-1} ds
\]

\[
= -\int_K f p_{k-1} dx = -\int_K \bar{f} p_{k-1} dx.
\]

Since \( \text{div } \hat{\sigma}_K \) and \( \bar{f} \) are contained in \( P_{k-1}(K) \), we readily deduce from (5.3) that

\[
\text{div } \hat{\sigma}_K + \bar{f} = 0.
\]
By setting $v = 1$ in (5.1) we see that $f - \bar{f}$ satisfies the assumption (4.3), and the effect of the latter can be bounded by Lemma 4.3. Let $\hat{\sigma}_h \in H(\text{div}, \Omega)$ be defined by $\hat{\sigma}_h|_K = \hat{\sigma}_K, K \in T_h(\Omega)$, and we have obtained an equilibrated flux up to data oscillations.

The last specification (5.2c) aims at the minimization of the error bound with respect to the known quantities. This is one difference from the equilibration procedure by Ern and Vohralík [18] who used Raviart–Thomas elements.

**Remark 5.1.** If $k = 1$, due to Remark 3.2, $\hat{\sigma}_K$ is uniquely defined by (5.2a), that is by data of the numerical flux on the edges.

Note that up to now we have not used the specification (2.4) that distinguishes the interior penalty method IPDG from the other DG elements.

**6. An approximation by conforming elements.** Corollary 4.2 shows that we require an approximation of $u_h$ by an $H^1$ function. We want to have a conforming element $u_{conf}^h$, and we will evaluate the norm

$$\|\text{grad } u_{conf}^h - \text{grad}_h u_h\|_{0,\Omega}$$

after computing the finite element approximations $u_h$ and $u_{conf}^h$. We follow the construction in [21] that was used in connection with residual-based error estimates. The estimates of (6.1) in [21] will be useful in the verification of the efficiency and Theorem 6.1. We note that similar constructions and estimates are found in several papers.

Let $N^L$ be the set of Lagrangian nodal points for the elements in $V_h$. Let $\kappa_i$ be the number of triangles that share the nodal point $x_i \in N^L$. We have $\kappa_i = 1$ if $x_i$ is contained in the interior of an element $K \in T_h(\Omega)$ or if $x_i$ is situated in the interior of an edge $E \in E_h(\Gamma)$, while $\kappa_i > 1$ if $x_i \in N^L \cap E_h(\Omega)$. The multiplicity $\kappa_i$ is bounded, since a minimal angle condition is assumed. The associated conforming element is now defined by its nodal values

$$u_{conf}^h(x_i) := \frac{1}{\kappa_i} \sum_{K \in T_h(\Omega), x_i \in K} u_h|_K(x_i).$$

The following estimate is provided by Theorem 2.2 in [21]:

$$\sum_{K \in T_h(\Omega)} \|\text{grad } u_{conf}^h - \text{grad}_h u_h\|_{0,K}^2 \leq c \sum_{E \in E_h(\Gamma)} h_E^{-1} \|[u_h]|_E\|_{0,E}^2.$$  (6.3)

The constant $c$ depends only on the degree $k$ of the finite elements and the shape regularity of the triangulation.

For sufficiently large penalty parameter, say $\alpha \geq \alpha_1$, the right-hand side of (6.3) can be bounded by Theorem 3.2(iv) in [22]:

$$\sum_{E \in E_h(\Omega)} h_E^{-1} \|[u_h]|_E\|_{0,E}^2 \leq c \left( \|\text{grad } u - \text{grad}_h u_h\|_{0,\Omega}^2 + \sum_{K \in T_h(\Omega)} h_K^2 \|f - \bar{f}\|_{0,K}^2 \right).$$  (6.4)

The last term on the right-hand side of (6.4) is added, since the analysis in [21] is done for zero data oscillation.
We note that $\alpha_1$ can be larger than the minimal stability estimate for the IPDG method. For sharp bounds on the jump seminorm we refer to [2, 3]. From (6.4) we eventually obtain

\begin{equation}
\| \nabla u_h^{\text{conf}} - \nabla u_h \|_{0, \Omega}^2 \\
\leq c \left( \| \nabla u - \nabla u_h \|_{0, \Omega}^2 + \sum_{K \in \mathcal{T}_h(\Omega)} h_K^2 \| f - \bar{f} \|_{0,K}^2 \right).
\end{equation}

This inequality will be used for the verification of the efficiency of the a posteriori error bound under consideration.

The inequality (6.5) is obtained from the efficiency of a residual a posteriori error estimate [21]. It gives rise to a comparison theorem for the solutions of two finite element methods in the spirit of the results in [8].

Theorem 6.1. Let $u_h^G$ be the solution of the Poisson equation by the conforming finite elements $V_h \cap H^1(\Omega)$ on the same triangulation. Then

\begin{equation}
\| \nabla (u - u_h^G) \|_{0, \Omega} \leq c \left( \| \nabla (u - u_h) \|_{0, \Omega} + \left( \sum_{K \in \mathcal{T}_h(\Omega)} h_K^2 \| f - \bar{f} \|_{0,K}^2 \right)^{1/2} \right).
\end{equation}

Proof. From the Galerkin orthogonality $(\nabla (u - u_h^G), \nabla v)_{0, \Omega} = 0$ for all $v \in V_h \cap H^1(\Omega)$ it follows that $\| \nabla (u - u_h^G) \|_{0, \Omega} \leq \| \nabla (u - u_h^{\text{conf}}) \|_{0, \Omega}$. Now we obtain from (6.5)

\begin{equation}
\| \nabla (u - u_h^G) \|_{0, \Omega} \leq \| \nabla (u - u_h^{\text{conf}}) \|_{0, \Omega} \\
\leq \| \nabla (u - u_h) \|_{0, \Omega} + \| \nabla (u_h - u_h^{\text{conf}}) \|_{0, \Omega} \\
\leq \| \nabla (u - u_h) \|_{0, \Omega} \\
+ c \left( \| \nabla (u - u_h) \|_{0, \Omega} + \left( \sum_{K \in \mathcal{T}_h(\Omega)} h_K^2 \| f - \bar{f} \|_{0,K}^2 \right)^{1/2} \right),
\end{equation}

and the proof is complete. □

We note that the comparison theorem was established independently of the equilibration.

7. The error bound and its efficiency. Let $\hat{\sigma}_h \in H(\text{div}, \Omega)$ with $\hat{\sigma}_h|_K \in BDM_k(K), K \in \mathcal{T}_h(\Omega)$, be the equilibrated flux constructed according to (5.2) and let $u_h^{\text{conf}} \in V_h \cap H^1(\Omega)$ be defined by the averaging procedure from the previous section. Recalling Corollary 4.2 we introduce the estimator

\begin{equation}
\eta_{\text{hyp}} := \eta_{\text{hyp}}^{(1)} + \eta_{\text{hyp}}^{(2)},
\end{equation}

\begin{equation}
\eta_{\text{hyp}}^{(\nu)} := \sum_{K \in \mathcal{T}_h(\Omega)} \eta_{K}^{(\nu)}, \quad 1 \leq \nu \leq 2,
\end{equation}

\begin{equation}
\eta_{K}^{(1)} := \| \nabla (u_h - \hat{\sigma}_h) \|_{0,K}, \quad \eta_{K}^{(2)} := 2 \| \nabla (u_h - u_h^{\text{conf}}) \|_{0,K}, \quad K \in \mathcal{T}_h(\Omega).
\end{equation}
By Corollary 4.2 and Lemma 4.3 we get the reliable a posteriori error estimate

\[(7.2) \| \text{grad} u - \text{grad}_h u_h \|_{0,h} \leq \eta_{\text{hyp}} + \frac{1}{\pi} \left( \sum_{K \in \mathcal{T}_h(\Omega)} h_K^2 \| f - \bar{f} \|_{0,K}^2 \right)^{1/2}. \]

From (6.5) it follows that the efficiency of the error bound (7.2) without the contribution of the data oscillation is guaranteed when we have appropriate bounds for \( \| \text{grad}_h u_h - \widehat{\sigma}_h \|_{0,\Omega} \). To this end, we will establish bounds for the terms in the triangle inequality \( \| \text{grad}_h u_h - \widehat{\sigma}_h \|_{0,\Omega} \leq \| \text{grad}_h u_h - \sigma_h \|_{0,\Omega} + \| \sigma_h - \widehat{\sigma}_h \|_{0,\Omega} \).

First, (2.3a) and Gauss’ theorem yield for \( \tau \in V_h \)

\[
\int_K (\sigma_h - \text{grad}_h u_h) \cdot \tau \, dx = \int_K \sigma_h \cdot \tau \, dx - \int_K \text{grad}_h u_h \cdot \tau \, dx
= -\int_K u_h \text{div} \tau \, dx + \int_{\partial K} \hat{u}_{\partial K} \nu_K \cdot \tau \, ds
+ \int_K u_h \text{div} \tau \, dx - \int_{\partial K} u_h \nu_K \cdot \tau \, ds
= \int_{\partial K} (\hat{u}_{\partial K} - u_h) \nu_K \cdot \tau \, ds.
\]

It follows from the specification of the internal penalty method (2.4) that \( \hat{u}_K - u_h = \frac{1}{2}[u_h]_E \) holds on \( E \subset \partial K \). We set \( \tau := \sigma_h - \text{grad}_h u_h \), and a standard scaling argument yields

\[(7.3) \| \sigma_h - \text{grad}_h u_h \|_{0,K} \leq c h_K^{-1/2} \| [u_h]_{\partial K} \|_{0,\partial K}. \]

After summing over all elements we obtain with (6.4) the required bound for the left-hand side of (7.3),

\[
\| \sigma_h - \text{grad}_h u_h \|_{0,\Omega} \leq c \left( \| \text{grad} u - \text{grad}_h u_h \|_{0,\Omega}^2 + \sum_{K \in \mathcal{T}_h(\Omega)} h_K^2 \| f - \bar{f} \|_{0,K}^2 \right)^{1/2}.
\]

Moreover, it follows from Lemma 3.4 and (7.3) that

\[(7.4) \| (\sigma_h - \text{grad}_h u_h) \cdot \nu_K \|_{0,\partial K} \leq c h_K^{-1} \| [u_h]_{\partial K} \|_{0,\partial K}. \]

Eventually we derive a bound for \( \widehat{\sigma}_h - \sigma_h \). Lemma 3.1 together with (5.2b) and (5.2c) yields

\[
\| \widehat{\sigma}_h - \sigma_h \|_{0,K} \leq c h_K^{1/2} \| (\sigma_h - \sigma_h) \cdot \nu_K \|_{0,\partial K}.
\]

Recalling the specification (2.4) for the IPDG method, we obtain on \( E \subset \partial K \)

\[
\widehat{\sigma}_h - \sigma_h = \sigma_h - \text{grad}_h u_h + (\text{grad}_h u_h - \sigma_h)
= \frac{1}{2}[\text{grad}_h u_h]_E - \alpha h_E^{-1} [u_h]_E \nu + (\text{grad}_h u_h - \sigma_h).
\]
Let $E = \partial K \cap \partial K'$. Theorem 3.2(ii) in [22] asserts that
\begin{align*}
\|\text{grad}_h u_h|_E \cdot \nu\|_{0,E} &\leq c h^{-1/2}_K \left( \|\text{grad} u - \text{grad}_h u_h\|_{0,\omega_E} + \left( \sum_{K \in T_h(\omega_E)} h^2_K \|f - \bar{f}\|_{0,K}^2 \right)^{1/2} \right),
\end{align*}
where $\omega_E := \bigcup\{K \in T_h(\Omega) | E \in E_h(K)\}$. The second term in (7.5) is already estimated in (6.4). The third term is reduced by (7.4) also to the second one, and we get
\begin{align*}
\sum_{K \in T_h(\Omega)} \|\hat{\sigma}_h - \sigma_h\|_{0,K}^2 
\leq c \left( \sum_{K \in T_h(\Omega)} \|\text{grad} u - \text{grad}_h u_h\|_{0,K}^2 + \left( \sum_{K \in T_h(\Omega)} h^2_K \|f - \bar{f}\|_{0,K}^2 \right)^{1/2} \right).
\end{align*}
By collecting all terms we obtain the efficiency of the a posteriori error estimate deduced from Corollary 4.2.

**Theorem 7.1.** Let $u_h$ and $\hat{\sigma}_h$ be the finite element solution of the IPDG method and the equilibrated flux, respectively. Further, assume that a conforming function $u_{h\text{ conf}}$ has been constructed as described in section 6. There is a constant $c$ that depends only on the degree $k$ and the shape regularity of the triangulation such that
\begin{align*}
\eta_{\text{hyp}} \leq c \left( \|\text{grad} u - \text{grad}_h u_h\|_{0,\Omega} + \left( \sum_{K \in T_h(\Omega)} h^2_K \|f - \bar{f}\|_{0,K}^2 \right)^{1/2} \right).
\end{align*}

**Remark 7.2.** There is the natural question whether the efficiency of the bound (7.2) can also be obtained in a unified way. The contributions from the hypercircle method are bounded by the differences $\hat{u} - u_h$ and $\hat{\sigma}_h - \sigma_h$ on the interelement boundaries. Table 3.2 in [5] in turn shows that these differences can be expressed in terms of the jumps of $u_h$ and $\nu \cdot \nabla u_h$ in many cases. The latter are found in the residual-based estimates; cf. [25]. If the efficiency of residual-based estimates is verified, we are done.

The treatment of $u_h - u_{h\text{ conf}}$ is independent of the origin of $u_h$.

**8. Numerical results.** In this section, we present a documentation of numerical results for two representative examples illustrating the performance of the suggested adaptive approach which consists of successive cycles of the steps
\begin{align*}
\text{SOLVE} \implies \text{ESTIMATE} \implies \text{MARK} \implies \text{REFINE}.
\end{align*}
In the step SOLVE we compute the solution of the IPDG approximation (2.3), whereas the second step ESTIMATE is devoted to the computation of the local components $\eta^{(1)}_K$ and $\eta^{(2)}_K$ of the error estimator $\eta_{\text{hyp}}$ (cf. (7.1)). We use the standard Dörfler marking in step MARK: Given some constant $0 < \theta \leq 1$, we choose a set $\mathcal{M} \subseteq T_h(\Omega)$ of elements $K \in T_h(\Omega)$ such that
\begin{align*}
\theta \eta_{\text{hyp}} \leq \sum_{K \in \mathcal{M}} \left( \eta^{(1)}_K + \eta^{(2)}_K \right).
\end{align*}
The final step REFINE takes care of the practical realization of the refinement process which is based on the newest vertex bisection \cite{12}.

**Example 1.** We consider the Laplace equation with inhomogeneous Dirichlet boundary conditions

\begin{align}
-\Delta u &= 0 \quad \text{in } \Omega, \\
  u &= g \quad \text{on } \partial \Omega
\end{align}

in the L-shaped domain $\Omega := (-1, +1)^2 \setminus ([0, +1] \times (-1, 0])$, where $g$ in (8.2b) is chosen such that

$$u(r, \varphi) = r^{2/3} \sin(2\varphi/3)$$

is the exact solution (in polar coordinates). The solution exhibits a singularity at the origin.

For $\theta = 0.3$ in the Dörfler marking, Figure 2 displays the adaptively refined meshes for polynomial degrees $1 \leq k \leq 4$. As can be expected, the adaptive algorithm refines in the vicinity of the origin with coarser meshes for increasing polynomial degree $k$. For adaptive refinement ($\theta = 1$), $\theta = 0.7$, and $\theta = 0.3$, Figure 3 (left) shows the decrease of the global discretization error $\| \text{grad } u - \text{grad}_h u_h \|_0, h$ as a function of the total number of DOF on a logarithmic scale for polynomial degree $k = 1$ (top left) to $k = 4$ (bottom left). The negative slope is indicated for each curve. We see that for $\theta = 0.3$ the optimal convergence rates are approached asymptotically. Figure 3 (right) displays the associated effectivity indices (ratio of the a posteriori error estimator and the global discretization error). In contrast to standard residual type a posteriori...
Example 1: $\lambda = 1, \alpha = 10.0$

Example 1: $\lambda = 2, \alpha = 22.5$

Example 1: $\lambda = 3, \alpha = 40.0$

Example 1: $\lambda = 4, \alpha = 62.5$

Fig. 3. Example 1: the discretization error $\|\nabla u - \nabla u_h\|_{0,h}$ as a function of the DOF on a logarithmic scale for various $\theta$ in the Dörfler marking (left) and the associated effectivity indices (right).
error estimators for IPDG approximations, the effectivity indices are slightly above 1 and do not significantly deteriorate with increasing polynomial degree $k$.

**Example 2.** We consider Poisson’s equation with homogeneous Dirichlet boundary conditions

\begin{align}
-\Delta u &= f \quad \text{in } \Omega, \\
u &= 0 \quad \text{on } \partial\Omega
\end{align}

in the unit square $\Omega = (0,1)^2$, where the right-hand side $f$ in (8.3a) is chosen such that

$$u(x,y) = x(1-x)y(1-y) \arctan(60(r-1)), \quad r^2 := (x-5/4)^2 + (y+1/4)^2$$

is the exact solution. The solution exhibits an interior layer along a circular segment inside the computational domain.

Figure 4 shows the adaptively refined meshes for polynomial degree $k = 1$ and $k = 4$ in the case of $\theta = 0.3$ in the Dörfler marking, whereas for uniform refinement ($\theta = 1$), $\theta = 0.7$ and $\theta = 0.3$. Figure 5 displays the global discretization error as a function of the DOF on a logarithmic scale (left) and the associated effectivity indices (right). We see that both for $\theta = 0.7$ and $\theta = 0.3$ the optimal convergence rates are achieved asymptotically and that the effectivity indices even slightly improve with increasing polynomial degree $k$.

### 9. Concluding remarks

The design of the a posteriori error bound is the same for all discontinuous Galerkin methods. There is no generic constant, and the proof of the reliability is much easier than that for residual-based estimators. In essence, it is focused on the terms which measure the nonconformity.

The proof of the efficiency is very similar to the analysis of residual-based error estimates, but there is one term less. The typical term

$$h\|\Delta u_h + f\|_{0,\Omega}$$

that models the negative norm $\|\Delta u_h + f\|_{-1}$ is not present, since implicitly a left inverse of the divergence operator is involved. The left inverse is constructed by a local procedure.

We recall that the analysis is based on the mixed formulation in [5], and it is known (see [8]) that the efficiency of the estimator is related to the quality of the mixed finite element method; cf. the comparison (7.6).
Fig. 5. Example 2: the discretization error $\|\nabla u - \nabla h u_h\|_{0,h}$ as a function of the DOF on a logarithmic scale for various $\theta$ in the Dörfler marking (left) and the associated effectivity indices (right).
REFERENCES

