ESAIM: M2AN 48 (2014) 753–764
DOI: 10.1051/m2an/2013119

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

CONVERGENCE ANALYSIS OF THE LOWEST ORDER WEAKLY PENALIZED ADAPTIVE DISCONTINUOUS GALERKIN METHODS

Thirupathi Gudi[1] and Johnny Guzmán[2]

Abstract. In this article, we prove convergence of the weakly penalized adaptive discontinuous Galerkin methods. Unlike other works, we derive the contraction property for various discontinuous Galerkin methods only assuming the stabilizing parameters are large enough to stabilize the method. A central idea in the analysis is to construct an auxiliary solution from the discontinuous Galerkin solution by a simple post processing. Based on the auxiliary solution, we define the adaptive algorithm which guides to the convergence of adaptive discontinuous Galerkin methods.

Mathematics Subject Classification. 65N30, 65N15.

Received September 11, 2012. Revised June 3, 2013
Published online April 1, 2014.

1. Introduction

The design of adaptive finite element methods based on reliable and efficient *a posteriori* error estimates has been the subject in the past [2, 6, 7, 13, 32]. The adaptive finite element method consists typically the following successive loops of the sequence

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}.$$

Convergence analysis of adaptive finite element methods has been initiated by Dörfler [24] who introduced an important marking strategy. Subsequently important theoretical developments have been made by many researchers. We refer to [20, 29, 30] for the work on conforming finite element methods, to [19, 21] for the results on mixed finite element methods, to [9, 18] for the work on nonconforming methods and finally to [12, 26, 27] for discontinuous Galerkin methods. On the other hand, the optimality of adaptive finite element method is derived in [11] for two dimensional problems and in [31] for high dimensional problems.

In this article, we focus on the low order adaptive discontinuous Galerkin (DG) methods. Karakashian and Pascal [27] were the first to prove contraction properties for the symmetric interior penalty Galerkin (SIPG) method. Therein, the authors have proved the contraction property for SIPG method under an interior node property. Subsequently the interior node property is relaxed independently in the works of [12, 26]. Moreover the quasi-optimal convergence rates are derived in [12]. However, the common issue with the three articles [12, 26, 27] is that the contraction property is derived assuming the penalty parameters are sufficiently large (*i.e.* larger than

Keywords and phrases. Contraction, adaptive finite element, discontinuous Galerkin.

[1] Department of Mathematics, Indian Institute of Science, 56002 Bangalore, India. gudi@math.iisc.ernet.in

[2] Division of Applied Mathematics, Brown University, Providence, 02912 RI, USA. Johnny_Guzman@brown.edu

what is needed for stability of the method). In this article, we prove contraction properties for various symmetric weakly penalized discontinuous Galerkin methods only assuming that the penalty parameters are large enough to guarantee stability of the method. For example, in the case of the LDG method the stabilizing parameters only have to be positive. This is achieved by a new marking strategy that uses an auxiliary solution obtained by post-processing the discontinuous Galerkin solution which turns out to be the Crouzeix–Raviart non-conforming approximation [23]. In fact, we borrow the marking strategies [9,18] that have been developed for non-conforming methods and show that these are enough to contract the error of the entire DG approximation. We are able to do this by using a result of Burman and Stamm [17] which shows that weakly penalized DG approximation can be written as the Crouzeix–Raviart approximation plus a discontinuous part with zero averages across interfaces. We show that the discontinuous part is controlled by data oscillations.

The weakly penalized method differ from classic DG methods in the fact that only the lower moments of the jumps are penalized on interfaces of the triangulation. For example, for piecewise linear elements and the SIPG method the penalty term looks like

$$\sum_{e\in\mathcal{E}_h}\frac{\alpha}{h_e}\int_e [\![w_h]\!][\![v_h]\!].$$

In contrast, in the weakly penalized case, one uses the term

$$\sum_{e\in\mathcal{E}_h}\frac{\alpha}{h_e}\int_e \Pi_e\big([\![w_h]\!]\big)\Pi_e\big([\![v_h]\!]\big)$$

where we use the average of the jump $\Pi_e\big([\![w_h]\!]\big)=\frac{1}{h_e}\int_e[\![w_h]\!]$. Note that this is equivalent to using the midpoint rule to evaluate the integrals $\int_e[\![w_h]\!][\![v_h]\!]$, so the weakly penalized method is cheaper to implement. Weak penalization has been used in weakly over penalized methods [14, 15]. Moreover, the weakly penalized DG method considered here was already analyzed by Burman and Stamm [17] and by Ayuso and Zikatonov [5]. Burman and Stamm [17] gave an a priori error analysis and their results are crucial in our analysis. Ayuso and Zikatanov [5] study the convergence of a multigrid algorithm for this method.

We consider the following model problem of finding $u\in H_0^1(\Omega)$ such that

$$a(u,v)=(f,v)\quad \forall v\in H_0^1(\Omega), \tag{1.1}$$

where

$$a(w,v)=(\nabla w,\nabla v)\qquad \forall w,\, v\in H_0^1(\Omega), \tag{1.2}$$

and $(\cdot,\cdot)$ denotes the $L^2(\Omega)$ inner product. We assume that $\Omega\subset\mathbb{R}^2$ is a bounded domain with polygonal boundary $\partial\Omega$ and $f\in L^2(\Omega)$.

The rest of the article is organized as follows. In Section 2, we introduce the notation and preliminary results. In Section 3, we recall the DG methods and corresponding stability results. In Section 4, we construct an auxiliary solution by averaging the DG solution and there in we derive some useful properties and results for the auxiliary solution. Section 5 is devoted to the convergence analysis of DG methods. Finally we conclude the article in Section 6.

2. Notation and preliminaries

The following notation will be used throughout the article:

$$
\begin{aligned}
\mathcal{T}_h &= \text{a face to face, shape regular simplicial triangulations of } \Omega \\
T &= \text{a triangle of } \mathcal{T}_h \qquad h_T = \text{diameter of } T \\
\mathcal{E}_h^i &= \text{set of all interior edges of } \mathcal{T}_h \\
\mathcal{E}_h^b &= \text{set of all boundary edges of } \mathcal{T}_h \\
\mathcal{E}_h &= \mathcal{E}_h^i \cup \mathcal{E}_h^b \\
\mathcal{M}_T &= \text{set of midpoints of the edges of } T \\
\mathcal{M}_h^i &= \text{set of all midpoints of the edges in } \mathcal{E}_h^i \\
\mathcal{M}_h^b &= \text{set of all midpoints of the edges in } \mathcal{E}_h^b \\
\mathcal{M}_h &= \mathcal{M}_h^i \cup \mathcal{M}_h^b \\
h_e &= \text{ length of the edge } e \in \mathcal{E}_h \\
\nabla_h &= \text{piecewise (element-wise) gradient} \\
\mathbb{P}_m(T) &= \text{space of polynomials of degree less than or equal to } m \geq 0 \text{ and defined on } T.
\end{aligned}
$$

The discontinuous finite element space is defined by

$$V_h = \{v_h \in L^2(\Omega) : v_h|_T \in P_1(T)\}.$$

In the analysis below, we need the following Crouzeix–Raviart nonconforming space [23]:

$$V_{CR} = \{v_h \in V_h : \int_e [\![v_h]\!]\,\mathrm{d}s = 0 \quad \forall e \in \mathcal{E}_h\},$$

and the following vector valued discrete space:

$$W_h = \{w_h \in L^2(\Omega)^2 : w_h|_T \in [P_0(T)]^2\}.$$

Define a broken Sobolev space

$$H^1(\Omega, \mathcal{T}_h) = \{v \in L^2(\Omega) :\ v_T = v|_T \in H^1(T) \quad \forall\ T \in \mathcal{T}_h\}.$$

For the DG methods, we require to define jump and mean of discontinuous functions. For any $e \in \mathcal{E}_h^i$, there are two triangles T_+ and T_- such that $e = \partial T_+ \cap \partial T_-$. Let n_- be the unit normal of e pointing from T_- to T_+, and $n_+ = -n_-$. (*cf.* Fig. 1). For any $v \in H^1(\Omega, \mathcal{T}_h)$, we define the jump and mean of v on e by

$$[\![v]\!] = v_- n_- + v_+ n_+, \text{ and } \{\!\{v\}\!\} = \frac{1}{2}(v_- + v_+), \text{ respectively,}$$

where $v_\pm = v\big|_{T_\pm}$. Similarly define for $w \in H^1(\Omega, \mathcal{T}_h)^2$ the jump and mean of w on $e \in \mathcal{E}_h^i$ by

$$[\![w]\!] = w_- \cdot n_- + w_+ \cdot n_+, \text{ and } \{\!\{w\}\!\} = \frac{1}{2}(w_- + w_+), \text{ respectively,}$$

where $w_\pm = w|_{T_\pm}$.

For any edge $e \in \mathcal{E}_h^b$, there is a triangle $T \in \mathcal{T}_h$ such that $e = \partial T \cap \partial \Omega$. Let n_e be the unit normal of e that points outside T. For any $v \in H^1(T)$, we set on $e \in \mathcal{E}_h^b$

$$[\![v]\!] = v n_e \text{ and } \{\!\{v\}\!\} = v,$$

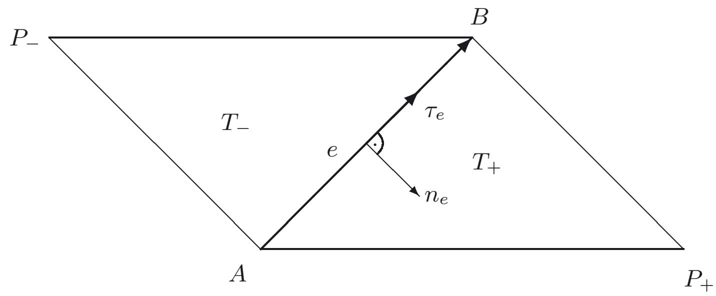

FIGURE 1. Two neighboring triangles T_- and T_+ that share the edge $e = \partial T_- \cap \partial T_+$ with initial node A and end node B and unit normal n_e. The orientation of $n_e = n_- = -n_+$ equals the outer normal of T_-, and hence, points into T_+.

and for $w \in H^1(T)^2$,

$$[\![w]\!] = w \cdot n_e, \text{ and } \{\!\{w\}\!\} = w.$$

The discontinuous Galerkin methods use a lifting operator $r : L^2(\mathcal{E}_h)^2 \to W_h$ defined by

$$\int_\Omega r(w) \cdot \tau \,\mathrm{d}x = -\sum_{e\in\mathcal{E}_h} \int_e w \cdot \{\!\{\tau\}\!\} \,\mathrm{d}s \quad \forall \tau \in W_h, \tag{2.1}$$

and a local analogue $r_e : L^2(e)^2 \to W_h$ defined by

$$\int_\Omega r_e(w) \cdot \tau \,\mathrm{d}x = -\int_e w \cdot \{\!\{\tau\}\!\} \,\mathrm{d}s \quad \forall \tau \in W_h. \tag{2.2}$$

Let $\Pi_e : L^2(e) \to R$ be the L^2-projection onto constants defined by

$$\Pi_e(v) = \frac{1}{h_e} \int_e v \,\mathrm{d}s. \tag{2.3}$$

3. DISCONTINUOUS GALERKIN METHODS

We consider four stable and symmetric weakly penalized discontinuous Galerkin Methods namely, the IP (or SIPG method) method [3, 25, 33], the LDG method [4, 22], the method by Brezzi *et al.* [16] and the method by Bassi *et al.* [8]. The original articles have considered the formulations using the penalty term with strong jumps, we replace them here with weak jumps. The concept of stabilizing the DG formulation by weak jumps was introduced in [14, 15].

The bilinear form for the IP method [3, 25, 33] is defined by

$$\begin{aligned}\mathcal{A}_h(w_h, v_h) = &\sum_{T\in\mathcal{T}_h} \int_T \nabla w_h \cdot \nabla v_h \,\mathrm{d}x - \sum_{e\in\mathcal{E}_h} \int_e \{\!\{\nabla w_h\}\!\} [\![v_h]\!] \,\mathrm{d}s \\ &- \sum_{e\in\mathcal{E}_h} \int_e \{\!\{\nabla v_h\}\!\} [\![w_h]\!] \,\mathrm{d}s + \sum_{e\in\mathcal{E}_h} \frac{\alpha}{h_e} \int_e \Pi_e\big([\![w_h]\!]\big) \Pi_e\big([\![v_h]\!]\big).\end{aligned} \tag{3.1}$$

where $\alpha > 0$ is the stabilizing parameter.

TABLE 1. Conditions for α.

Method	Condition on α
IP Method [3,25,33]	α satisfies (3.6)
LDG Method [4,22]	$\alpha > 0$
Brezzi *et al.* [16]	$\alpha > 0$
Bassi *et al.* [8]	$\alpha > 3$

The bilinear form for the LDG method [4,22] is defined by

$$\begin{aligned}\mathcal{A}_h(w_h,v_h) &= \sum_{T\in\mathcal{T}_h}\int_T \nabla w_h\cdot\nabla v_h\,\mathrm{d}x - \sum_{e\in\mathcal{E}_h}\int_e \{\!\{\nabla w_h\}\!\}[\![v_h]\!]\,\mathrm{d}s - \sum_{e\in\mathcal{E}_h}\int_e \{\!\{\nabla v_h\}\!\}[\![w_h]\!]\,\mathrm{d}s \\ &\quad + \int_\Omega r(\Pi_e([\![w_h]\!]))r(\Pi_e([\![v_h]\!])) + \sum_{e\in\mathcal{E}_h}\frac{\alpha}{h_e}\int_e \Pi_e\big([\![w_h]\!]\big)\Pi_e\big([\![v_h]\!]\big).\end{aligned} \tag{3.2}$$

The bilinear form for the Brezzi *et al.* method [16] is defined by

$$\begin{aligned}\mathcal{A}_h(w_h,v_h) &= \sum_{T\in\mathcal{T}_h}\int_T \nabla w_h\cdot\nabla v_h\,\mathrm{d}x - \sum_{e\in\mathcal{E}_h}\int_e \{\!\{\nabla w_h\}\!\}[\![v_h]\!]\,\mathrm{d}s - \sum_{e\in\mathcal{E}_h}\int_e \{\!\{\nabla v_h\}\!\}[\![w_h]\!]\,\mathrm{d}s \\ &\quad + \int_\Omega r(\Pi_e([\![w_h]\!]))r(\Pi_e([\![v_h]\!])) + \sum_{e\in\mathcal{E}_h}\alpha\int_\Omega r_e\big(\Pi_e([\![w_h]\!])\big)r_e\big(\Pi_e([\![v_h]\!])\big).\end{aligned} \tag{3.3}$$

The bilinear form for the Bassi *et al.* method [8] is defined by

$$\begin{aligned}\mathcal{A}_h(w_h,v_h) &= \sum_{T\in\mathcal{T}_h}\int_T \nabla w_h\cdot\nabla v_h\,\mathrm{d}x - \sum_{e\in\mathcal{E}_h}\int_e \{\!\{\nabla w_h\}\!\}[\![v_h]\!]\,\mathrm{d}s - \sum_{e\in\mathcal{E}_h}\int_e \{\!\{\nabla v_h\}\!\}[\![w_h]\!]\,\mathrm{d}s \\ &\quad + \sum_{e\in\mathcal{E}_h}\alpha\int_\Omega r_e\big(\Pi_e([\![w_h]\!])\big)r_e\big(\Pi_e([\![v_h]\!])\big).\end{aligned} \tag{3.4}$$

The DG method is to find $u_h \in V_h$ such that

$$\mathcal{A}_h(u_h,v_h) = (f,v_h) \qquad \forall v_h \in V_h, \tag{3.5}$$

where $\mathcal{A}_h$ is any of the bilinear form defined in (3.1)–(3.4).

It is proved in Lemma 1, [1] that the IP method is stable for any α satisfying

$$\alpha > 4\max_{T\in\mathcal{T}_h}\rho(S_T), \tag{3.6}$$

where $\rho(S_T)$ is the spectral radius of the local stiffness matrix $[S_T]_{mn} = (\nabla_h\lambda_m, \nabla_h\lambda_n)$, where λ_m's are barycentric coordinates of T. In Table 1, we present the condition on α for the above DG methods.

Define the mesh dependent norm

$$\|v\|_{1,h}^2 = \sum_{T\in\mathcal{T}_h}\int_T |\nabla v|^2\,\mathrm{d}x + \sum_{e\in\mathcal{E}_h}\Pi_e([\![v]\!])^2 \quad \forall v\in H^1(\Omega,\mathcal{T}_h). \tag{3.7}$$

The following lemma on the stability of the DG methods (3.5) is well-known [4].

Lemma 3.1. *Assume that α satisfies the conditions in Table 1. Then, it holds that*

$$C\|v_h\|_{1,h}^2 \le \mathcal{A}_h(v_h,v_h) \qquad \forall v_h\in V_h.$$

4. An auxiliary solution by post processing

Here we collect some results of the DG approximation that show how it is related to the Crouzeix–Raviart approximation. Lemmas 4.1, 4.3, 4.5, were proved by Burman and Stamm [17].

Let $u_h \in V_h$ be the solution of any of the DG methods (3.5). Define an auxiliary solution $u_h^* \in V_{CR}$ by the following:

$$u_h^*(m_e) := \begin{cases} \{\!\{u_h\}\!\}(m_e), \text{ if } m_e \in \mathcal{M}_h^i \\ 0, \qquad\qquad\quad \text{if } m_e \in \mathcal{M}_h^b. \end{cases} \tag{4.1}$$

In the following lemma, we establish an integral relation for u_h and u_h^*.

Lemma 4.1. *For any $v_h \in V_h$, it holds that*

$$\sum_{T\in\mathcal{T}_h} \int_T \nabla u_h \cdot \nabla v_h \,\mathrm{d}x - \sum_{e\in\mathcal{E}_h} \int_e \{\!\{\nabla v_h\}\!\} [\![u_h]\!] \,\mathrm{d}s = \sum_{T\in\mathcal{T}_h} \int_T \nabla u_h^* \cdot \nabla v_h \,\mathrm{d}s. \tag{4.2}$$

Proof. Using integration by parts, we find

$$\sum_{T\in\mathcal{T}_h} \int_T \nabla (u_h - u_h^*) \cdot \nabla v_h \,\mathrm{d}x = \sum_{e\in\mathcal{E}_h} \int_e \{\!\{\nabla v_h\}\!\} [\![u_h - u_h^*]\!] \,\mathrm{d}s + \sum_{e\in\mathcal{E}_h^i} \int_e [\![\nabla v_h]\!] \{\!\{u_h - u_h^*\}\!\} \,\mathrm{d}s.$$

The definition of u_h^* implies

$$\sum_{e\in\mathcal{E}_h^i} \int_e [\![\nabla v_h]\!] \{\!\{u_h - u_h^*\}\!\} \,\mathrm{d}s = 0,$$

$$\sum_{e\in\mathcal{E}_h} \int_e \{\!\{\nabla v_h\}\!\} [\![u_h^*]\!] \,\mathrm{d}s = 0.$$

This completes the proof. ◻

In the following lemma, we estimate the error between u_h and u_h^*.

Lemma 4.2. *It holds that*

$$\sum_{T\in\mathcal{T}_h} \Big(h_T^{-2} \|u_h - u_h^*\|_{L^2(T)}^2 + \|\nabla (u_h - u_h^*)\|_{L^2(T)}^2 \Big) \leq C \sum_{e\in\mathcal{E}_h} \mathit{\Pi}_e\big([\![u_h]\!]\big)^2$$

Proof. The proof is an easy consequence of the following estimate: for any $v_h \in P_1(T)$, it holds that

$$\|v_h\|_{L^2(T)}^2 \leq C|T| \sum_{m_e\in\mathcal{M}_T} v_h(m_e)^2. \tag{4.3}$$

The following identity is useful in our subsequent analysis. ◻

Lemma 4.3. *It holds that*

$$\mathcal{A}_h(u_h - u_h^*, u_h - u_h^*) = (f, u_h - u_h^*)$$

Proof. Using (3.5), we find

$$\mathcal{A}_h(u_h - u_h^*, u_h - u_h^*) - (f, u_h - u_h^*) = -\mathcal{A}_h(u_h^*, u_h - u_h^*)$$

It is remaining to show that $\mathcal{A}_h(u_h^*, u_h - u_h^*) = 0$. Using the fact that $\mathit{\Pi}_e([\![u_h^*]\!]) = 0$ for all $e \in \mathcal{E}_h$, integration by parts and (4.1), we obtain

$$\begin{aligned}\mathcal{A}_h(u_h^*, u_h - u_h^*) &= \sum_{T\in\mathcal{T}_h} \int_T \nabla u_h^* \cdot \nabla(u_h - u_h^*)\mathrm{d}x - \sum_{e\in\mathcal{E}_h} \int_e \{\!\{\nabla u_h^*\}\!\}[\![u_h - u_h^*]\!]\mathrm{d}s \\ &= \sum_{e\in\mathcal{E}_h^i} \int_e [\![\nabla u_h^*]\!]\{\!\{u_h - u_h^*\}\!\}\mathrm{d}s \\ &= 0.\end{aligned}$$

This completes the proof. □

In the following lemma, we estimate the error $u_h^* - u_h$ by volume residual. It is crucial that the $u_h^* - u_h$ is bounded only by an oscillation term for the contraction property to hold. Our bound is different from the bound given by Burman and Stamm [17].

Lemma 4.4. *Assume that α satisfies the conditions in Table 1. Then there exists some $C^* > 0$ such that*

$$\|u_h^* - u_h\|_{1,h}^2 \le C^* \sum_{T\in\mathcal{T}_h} h_T^2 \|f\|_{L^2(T)}^2.$$

Proof. Using Lemma 3.1, Lemma 4.3, Cauchy–Schwarz inequality, Lemma 4.2 and Young's inequality, we find

$$\begin{aligned}C\|u_h^* - u_h\|_{1,h} &\le \mathcal{A}_h(u_h^* - u_h, u_h^* - u_h) = (f, u_h^* - u_h) \\ &\le \left(\sum_{T\in\mathcal{T}_h} h_T^2\|f\|_{L^2(T)}^2\right)^{1/2} \left(\sum_{T\in\mathcal{T}_h} h_T^{-2}\|u_h^* - u_h\|_{L^2(T)}^2\right)^{1/2} \\ &\le C\left(\sum_{T\in\mathcal{T}_h} h_T^2\|f\|_{L^2(T)}^2\right)^{1/2} \left(\sum_{e\in\mathcal{E}_h} \mathit{\Pi}_e([\![u_h]\!])^2\right)^{1/2} \\ &\le \frac{C}{\epsilon}\left(\sum_{T\in\mathcal{T}_h} h_T^2\|f\|_{L^2(T)}^2\right) + \epsilon\|u_h^* - u_h\|_{1,h}^2.\end{aligned}$$

We complete the proof by choosing ϵ sufficiently small. □

Lemma 4.5. *The auxiliary solution u_h^* satisfies*

$$(\nabla_h u_h^*, \nabla_h v_h) = (f, v_h) \quad \forall v_h \in V_{CR},$$

i.e, u_h^* is the solution of the classical nonconforming method [23].

Proof. Using (3.5) for any $v_h \in V_{CR}$, we find

$$\sum_{T\in\mathcal{T}_h} \int_T \nabla u_h \cdot \nabla v_h \,\mathrm{d}x - \sum_{e\in\mathcal{E}_h} \int_e \{\!\{\nabla v_h\}\!\}[\![u_h]\!]\,\mathrm{d}s = (f, v_h).$$

Then using (4.2), we complete the proof. □

The following *a posteriori* error estimator is an easy consequence of Lemma 4.4 and the results in [18]:

Lemma 4.6. *Let u and u_h be the solutions of* (1.1) *and* (3.5). *Let u_h^* be the auxiliary solution defined in* (4.1). *Then it holds that*

$$\|\nabla_h(u-u_h)\|_{L^2(\Omega)} \le C\left(\sum_{T\in\mathcal{T}_h} h_T^2\|f\|^2_{L^2(T)} + \sum_{e\in\mathcal{E}_h^i}\int_e h_e [\![\partial u_h^*/\partial s]\!]^2\,\mathrm{d}s\right)^{1/2},$$

where $\partial/\partial s$ denotes the tangential derivative along the edge e.

Proof. First using triangle inequality

$$\|\nabla_h(u-u_h)\|_{L^2(\Omega)} \le \|\nabla_h(u-u_h^*)\|_{L^2(\Omega)} + \|\nabla_h(u_h^*-u_h)\|_{L^2(\Omega)},$$

and then using Lemma 4.4 and the results in [18], we complete the proof. □

5. Convergence of adaptive DG methods

Let $u_h \in V_h$ be the solution of any of the DG method (3.5) and let $u_h^* \in V_{CR}$ be the auxiliary solution constructed in (4.1). Using Lemma 4.5, recall that u_h^* is the solution of the Crouzeix–Raviart nonconforming method and Lemma 4.4 implies that there exists a positive constant $C^* > 0$ such that

$$\|u_h^* - u_h\|_h^2 + \sum_{e\in\mathcal{E}_h} \Pi_e([\![u_h]\!])^2 \le C^*\|hf\|^2,$$

where hereafter $\|\cdot\|_h = \|\nabla_h\cdot\|$ and

$$\|hf\|^2 = \sum_{T\in\mathcal{T}_h} h_T^2\|f\|^2_{L^2(T)}.$$

Let $\mathcal{T}_h$ be the conforming refinement of $\mathcal{T}_H$ obtained by refining the all the marked elements in $\mathcal{T}_H$ that are marked in the step MARK. The functions with h (resp. H) suffix corresponds to the mesh $\mathcal{T}_h$ (resp. $\mathcal{T}_H$). Below, we consider separately two different marking strategies that are introduced by Carstensen and Hoppe [18] and by Becker et al. [9] and prove the error reduction for both the algorithms separately.

Marking by Carstensen and Hoppe [18]:

Given the universal constants with $0 < \Theta, \rho_2 < 1$, the outcome of MARK is a set of edges $\mathcal{M} \subset \mathcal{E}_H$ such that

$$\Theta \sum_{e\in\mathcal{E}_H^i}\int_e H_e[\![\partial u_H^*/\partial s]\!]^2\,\mathrm{d}s \le \sum_{e\in\mathcal{M}}\int_e H_e[\![\partial u_H^*/\partial s]\!]^2\,\mathrm{d}s. \tag{5.1}$$

The refined regular triangulation $\mathcal{T}_h$ from REFINE generated by refining at least all the edges in $\mathcal{M}$ (and possibly further edges to avoid hanging nodes) with the new mesh-size $h < H$ is supposed to satisfy

$$\rho_2\|Hf\|^2_{L^2(\Omega)} \le \|hf\|^2_{L^2(\Omega)}. \tag{5.2}$$

The results in Carstensen and Hoppe ([18], Thm. 1.1), imply that there exists $0 < \rho_1 < 1$ and $C_1 > 0$ such that

$$\|u-u_h^*\|_h^2 \le \rho_1\|u-u_H^*\|_H^2 + C_1\|Hf\|^2, \tag{5.3}$$

$$\|hf\|^2 \le \rho_2\|Hf\|^2. \tag{5.4}$$

The results in (5.3)–(5.4) imply the following Q-linear convergence ([18], line 6 on page 254):

$$\left(\|u-u_h^*\|_h^2 + \beta\|hf\|^2\right) \le \max\{\rho_1, (1+\rho_2)/2\}\left(\|u-u_H^*\|_H^2 + \beta\|Hf\|^2\right)$$

where $\beta = 2C_1/(1-\rho_2)$.

In the following theorem, we derive the contraction property for the adaptive DG methods (3.5) using the marking strategy by Carstensen and Hoppe (5.1).

Theorem 5.1. *Let the marking be done by (5.1). Then there exists $\gamma > 0$ and $0 < \rho^* < 1$ such that*

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le \rho^* \left(\|u - u_H\|_H^2 + \gamma\|Hf\|^2\right).$$

Proof. Let $\epsilon > 0$. Using triangle inequality and Young's inequality, we find

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le (1+\epsilon)\|u - u_h^*\|_h^2 + (1+1/\epsilon)\|u_h - u_h^*\|_h^2 + \gamma\|hf\|^2.$$

Using Lemma 4.4, we obtain

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le (1+\epsilon)\|u - u_h^*\|_h^2 + \big(C^*(1+1/\epsilon) + \gamma\big)\|hf\|^2.$$

Using the error reduction for u_h^* (5.3), we find

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le (1+\epsilon)\rho_1\big(\|u - u_H^*\|_H^2 + C_1\|Hf\|^2\big) + \big(C^*(1+1/\epsilon) + \gamma\big)\|hf\|^2.$$

Again using triangle inequality, Young's inequality and lemma 4.4, we find

$$\begin{aligned}\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le{}& (1+\epsilon)\rho_1\big((1+\epsilon)\|u - u_H\|_H^2 + C^*(1+1/\epsilon)\|Hf\|^2 + C_1\|Hf\|^2\big)\\ &+ \big(C^*(1+1/\epsilon) + \gamma\big)\|hf\|^2.\end{aligned}$$

Therefore

$$\begin{aligned}\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le{}& (1+\epsilon)^2\rho_1\|u - u_H\|_H^2 + (1+\epsilon)\rho_1\big(C^*(1+1/\epsilon) + C_1\big)\|Hf\|^2\\ &+ \big(C^*(1+1/\epsilon) + \gamma\big)\|hf\|^2.\end{aligned}$$

We now use (5.4) and find

$$\begin{aligned}\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le{}& (1+\epsilon)^2\rho_1\|u - u_H\|_H^2 + (1+\epsilon)\rho_1\big(C^*(1+1/\epsilon) + C_1\big)\|Hf\|^2\\ &+ \big(C^*(1+1/\epsilon) + \gamma\big)\rho_2\|Hf\|^2.\end{aligned}$$

By simplifying

$$\begin{aligned}\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le{}& (1+\epsilon)^2\rho_1\|u - u_H\|_H^2\\ &+ \left[(1+\epsilon)\rho_1\big(C^*(1+1/\epsilon) + C_1\big) + \big(C^*(1+1/\epsilon)\big)\rho_2 + \gamma\rho_2\right]\|Hf\|^2.\end{aligned}$$

Choosing γ sufficiently large such that

$$\left[(1+\epsilon)\rho_1\big(C^*(1+1/\epsilon) + C_1\big) + \big(C^*(1+1/\epsilon)\big)\rho_2 + \gamma\rho_2\right] \le \gamma(1+\rho_2)/2.$$

That is by choosing γ such that

$$2\left[(1+\epsilon)\rho_1\big(C^*(1+1/\epsilon) + C_1\big) + \big(C^*(1+1/\epsilon)\big)\rho_2\right]/(1-\rho_2) \le \gamma,$$

we complete the proof by $\rho^* = \min\{(1+\epsilon)^2\rho_1, (1+\rho_2)/2\}$ and ϵ such that $(1+\epsilon)^2\rho_1 < 1$. □

Remark 5.2. The marking strategy by Carstensen and Hoppe in (5.1)–(5.2) is improved by Becker, Mao and Shi [9] so that the optimal rate of convergence can be derived. Below, we prove the contraction property for DG methods using Becker, Mao and Shi marking strategy.

Marking by Becker, Mao and Shi [9]: Choose the parameters $0 < \theta, \sigma < 1$ and γ. Then the out come of MARK step is the set of edges or elements according to the following:
If $\|Hf\|^2_{L^2(\Omega)} \le \gamma \sum_{e\in\mathcal{E}^i_H} \int_e H_e [\![\partial u^*_H/\partial s]\!]^2 \,\mathrm{d}s$, then mark a subset $\mathcal{M} \subset \mathcal{E}^i_h$ with minimal cardinality such that

$$\theta \sum_{e\in\mathcal{E}^i_H} \int_e H_e [\![\partial u^*_H/\partial s]\!]^2 \,\mathrm{d}s \le \sum_{e\in\mathcal{M}} \int_e H_e [\![\partial u^*_H/\partial s]\!]^2 \,\mathrm{d}s, \tag{5.5}$$

else find a subset $\mathcal{M}_1 \subset \mathcal{T}_H$ with minimal cardinality such that

$$\sigma \sum_{T\in\mathcal{T}_H} H_T^2 \|f\|^2_{L^2(T)} \le \sum_{T\in\mathcal{M}_1} H_T^2 \|f\|^2_{L^2(T)}. \tag{5.6}$$

The result by Becker, Mao and Shi [9] is that there exist $0 < \rho < 1$ and $\beta^* > 0$ such that

$$\|u - u_h^*\|_h^2 + \beta\|hf\|^2 \le \rho\big(\|u - u_H^*\|_H^2 + \beta\|Hf\|^2\big), \tag{5.7}$$

for all β such that $\beta \ge \beta^*$. Although, the result in Becker, Mao and Shi ([9], Thm. 4.1) states that this holds for a β sufficiently large a careful inspection of their proof shows that such an inequality holds for all $\beta > \beta^*$ with β^* sufficiently large [10].

Below, we prove the convergence of adaptive DG methods under the Becker, Mao and Shi marking (5.5)–(5.6).

Theorem 5.3. *Suppose that the marking is done by (5.5)–(5.6). Then there exists $\gamma > 0$ and $0 < \rho^* < 1$ such that*

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le \rho^* \left(\|u - u_H\|_H^2 + \gamma\|Hf\|^2\right).$$

Proof. Let $\epsilon > 0$. Using triangle inequality and Young's inequality, we find

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le (1+\epsilon)\|u - u_h^*\|_h^2 + (1 + 1/\epsilon)\|u_h - u_h^*\|_h^2 + \gamma\|hf\|^2.$$

Using Lemma 4.4, we obtain

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le (1+\epsilon)\|u - u_h^*\|_h^2 + \big(C^*(1 + 1/\epsilon) + \gamma\big)\|hf\|^2,$$

or equivalently

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le (1+\epsilon)\Big(\|u - u_h^*\|_h^2 + \big(C^*(1 + 1/\epsilon) + \gamma\big)/(1+\epsilon)\|hf\|^2\Big).$$

Assume that γ is sufficiently large such that

$$\big(C^*(1 + 1/\epsilon) + \gamma\big)/(1+\epsilon) =: \beta \ge \beta^*. \tag{5.8}$$

Then, using (5.7) we arrive at

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le (1+\epsilon)\rho\big(\|u - u_H^*\|_H^2 + \beta\|Hf\|^2\big).$$

Again using triangle inequality and Young's inequality and Lemma 4.4, we find

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le (1+\epsilon)\rho\left[(1+\epsilon)\|u - u_H\|_H^2 + C^*(1 + 1/\epsilon)\|Hf\|^2 + \beta\|Hf\|^2\right].$$

Therefore

$$\|u - u_h\|_h^2 + \gamma\|hf\|^2 \le (1+\epsilon)^2\rho\|u - u_H\|_H^2 + (1+\epsilon)\rho\big(C^*(1 + 1/\epsilon) + \beta\big)\|Hf\|^2.$$

First note that we can choose ϵ sufficiently small such that $(1+\epsilon)^2\rho < 1$. Then the proof will be completed if we can show that there is some $0 < \rho_2 < 1$ such that

$$(1+\epsilon)\rho\big(C^*(1+1/\epsilon)+\beta\big) \le \rho_2\gamma, \tag{5.9}$$

with $\rho^* = \min\{(1+\epsilon)^2\rho, \rho_2\}$.

Using (5.8) in (5.9),

$$(1+\epsilon)\rho C^*(1+1/\epsilon) + \frac{(1+\epsilon)\rho C^*(1+1/\epsilon)}{(1+\epsilon)} + \rho\gamma \le \rho_2\gamma,$$

equivalently

$$(1+\epsilon)\rho C^*(1+1/\epsilon) + \frac{(1+\epsilon)\rho C^*(1+1/\epsilon)}{(1+\epsilon)} \le (\rho_2-\rho)\gamma.$$

The proof is completed by choosing ρ_2 such that $\rho < \rho_2 < 1$ and γ sufficiently large. □

Remark 5.4. Using the triangle inequality and Lemma 4.4, we find

$$\begin{aligned}\|u-u_h\|_h &\le \|u-u_h^*\|_h + \|u_h^*-u_h\|_h\\ &\le \|u-u_h^*\|_h + C\|hf\|.\end{aligned}$$

Therefore the adaptive DG methods converge at least at the rate of adaptive nonconforming method. It was shown that the adaptive algorithm of Becker, Mao and Shi [9] has optimal rate of convergence for the nonconforming method.

6. Conclusions and future work

In this article, we have proved the contraction property for various symmetric discontinuous Galerkin (DG) methods. Unlike in the existing works for strongly penalized DG methods, we prove the convergence of weakly penalized adaptive DG methods without further assuming the stabilizing parameter is larger than what is required for stability. Although the analysis in this article is restricted to the lowest order case, we hope that similar ideas may be used in higher order cases. We remark that the convergence analysis of adaptive DG methods using strong jumps is still open when the stabilizing parameter is chosen just according to the stability of the method.

We would like to note that different markings can be used for which the contraction for the non-conforming method holds. For example, we can use the marking by Mao *et al.* [28].

In a future work we will consider the more general equation

$$\nabla(A\nabla u) + \mathbf{b}\cdot\nabla u + c\,u = f.$$

In this case the corresponding post-processed solution u_h^* will not be exactly the corresponding Crouzeix–Raviart solution. Instead, it satisfies a perturbed problem where the perturbation can be controlled by volume residuals.

Acknowledgements. We would like to thank Mark Ainsworth for many useful discussions. The work was done while the first author visited Brown University and Institute for Computational and Experimental Research in Mathematics (ICERM) with the help of the Indo-US Virtual Institute of Mathematical and Statistical Sciences. We would like to thank Govind Menon and Govindan Rangarajan for facilitating this visit. While the first author was partially supported by the DST-Fast Track scheme, the second author was partially supported by the NSF through grant number DMS-0914596.

References

[1] M. Ainsworth, A posteriori error estimation for discontinuous Galerkin finite element approximation. *SIAM J. Numer. Anal.* **39** (2007) 1777–1798.

[2] M. Ainsworth and J.T. Oden, A posteriori error estimation in finite element analysis. *Pure and Applied Mathematics*. Wiley-Interscience, John Wiley & Sons, New York (2000).

[3] D.N. Arnold, An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.* **19** (1982) 742–760.

[4] D.N. Arnold, F. Brezzi, B. Cockburn and L.D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.* **39** (2002) 1749–1779.

[5] B. Ayuso and L.L. Zikatanov, Uniformly convergent iterative methods for discontinuous Galerkin discretizations. *J. Sci. Comput.* **40** (2009) 4–36.

[6] I. Babuška and I. Strouboulis, *The Finite Element Method and its Reliability*. The Claredon Press, Oxford University Press (2001)

[7] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*. Birkhåuser Verlag, Basel (2003).

[8] F. Bassi, S. Rebay, G. Mariotti, S. Pedinotti, and M. Savini, A higher order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows, in *Proc. of 2nd European Conference on Turbomachinery, Fluid Dynamics and Thermodynamics*, edited by R. Decuypere and G. Dilbelius, Technologisch Instituut, Antewerpen, Belgium (1997) 99–108.

[9] R. Becker, S. Mao and Z.C. Shi, A convergent nonconforming adaptive finite element method with quasi-optimal complexity. *SIAM J. Numer. Anal.* **47** (2010) 4639–4659.

[10] R. Becker and S. Mao, Private Communication (2013).

[11] P. Binev, W. Dahmen, and R. DeVore, Adaptive finite element methods with convergence rates. *Numer. Math.* **97** (2004) 219–268.

[12] A. Bonito and R.H. Nochetto, Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method. *SIAM J. Numer. Anal.* **48** (2010) 734–771.

[13] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods, 3rd edn*. Springer-Verlag, New York (2008).

[14] S.C. Brenner and L. Owens, A weakly over-penalized non-symmetric interior penalty method. *J. Numer. Anal. Ind. Appl. Math.* **2** (2007) 35–48.

[15] S.C. Brenner, L. Owens and L.Y. Sung, A weakly over-penalized symmetric interior penalty method. *Electron. Trans. Numer. Anal.* **30** (2008) 107–127.

[16] F. Brezzi, G. Manzini, D. Marini, P. Pietra and A. Russo, Discontiuous Galerkin Approximations for Elliptic Problems. *Numer. Methods Partial Differ. Equ.* **16** (2000) 365–378.

[17] E. Burman and B. Stamm, Low order discontinuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.* **47** (2008) 508–533.

[18] C. Carstensen and R.H.W. Hoppe, Convergence analysis of an adaptive nonconforming finite element method. *Numer. Math.* **103** (2006) 251–266.

[19] C. Carstensen and R. Hoppe, Error reduction and convergence for an adaptive mixed finite element method. *Math. Comput.* **75** (2006) 1033–1042.

[20] J.M. Cascon, C. Kreuzer, R.H. Nochetto and K.G. Siebert, Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.* **46** (2008) 2524–2550.

[21] L. Chen, M. Holst and J. Xu, Convergence and optimality of adaptive mixed finite element methods. *Math. Comput.* **78** (2009) 35–53.

[22] B. Cockburn and C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.* **35** (1998) 2440–2463.

[23] M. Crouzeix and P.A. Raviart, Conforming and Nonconforming finite element methods for solving the stationary Stokes equations. *RAIRO Anal. Numer.* **7** (1973) 33–76.

[24] W. Dörfler, A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.* **33** (1996) 1106–1124.

[25] J. Douglas, Jr. and T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods. In vol. 58. *Lect. Notes Phys.* Springer-Verlag, Berlin (1976).

[26] R.H.W. Hoppe, G. Kanschat and T. Warburton, Convergence analysis ofan adaptive interior penalty discontinuous Galerkin method. *SIAM J. Numer. Anal.* **47** (2008/09) 534–550.

[27] O. A. Karakashian and F. Pascal, Convergence of adaptive discontinuous Galerkin approximations of second-order elliptic problems. *SIAM J. Numer. Anal.* **45** (2007) 641–665.

[28] S. Mao, X. Zhao and Z. Shi, Convergence of a standard adaptive nonconforming finite element method with optimal complexity. *Appl. Numer. Math.* **60** (2010) 673–688.

[29] P. Morin, R.H. Nochetto and K.G. Siebert, Data oscillation and convergence adaptive FEM. *SIAM J. Numer. Anal.* **38** (2000) 466–488.

[30] P. Morin, R.H. Nochetto and K.G. Siebert, Convergence of adaptive finite element methods. *SIAM Review* **44** (2002) 631–658.

[31] R. Stevenson, Optimality of a standard adaptive finite element method. *Found. Comput. Math.* **7** (2007) 245–269.

[32] R. Verfürth, *A Review of A Posteriori Error Estmation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, Chichester (1995).

[33] M.F. Wheeler, An elliptic collocation-finite-element method with interior penalties. *SIAM J. Numer. Anal.* **15** (1978) 152–161.