QUASI-OPTIMAL CONVERGENCE RATE OF AN ADAPTIVE DISCONTINUOUS GALERKIN METHOD

ANDREA BONITO* AND RICARDO H. NOCHETTO†

Abstract. We analyze an adaptive discontinuous finite element method (ADFEM) for symmetric second order linear elliptic operators. The method is formulated on nonconforming meshes made of simplices or quadrilaterals, with any polynomial degree and in any dimension ≥ 2. We prove that the ADFEM is a contraction for the sum of the energy error and the scaled error estimator, between two consecutive adaptive loops. We design a refinement procedure that maintains the level of nonconformity uniformly bounded, and prove that the approximation classes using continuous and discontinuous finite elements are equivalent. The geometric decay and the equivalence of classes are instrumental to derive optimal cardinality of ADFEM. We show that ADFEM (and AFEM on nonconforming meshes) yields a decay rate of energy error plus oscillation in terms of number of degrees of freedom as dictated by the best approximation for this combined nonlinear quantity.

Key words. Discontinuous Galerkin, interior penalty, nonconforming meshes, hanging nodes, adaptive algorithm, error reduction, convergence, optimal cardinality, approximation classes.

AMS subject classifications. 65N30, 65N50, 65N15, 65N12, 41A25

1. Introduction and Main Results. Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a polyhedral domain. We study the convergence of an interior penalty discontinuous Galerkin (IPDG) method for the diffusion problem:

$$-\mathrm{div}\,(A\nabla u) = f \quad \mathrm{in} \quad \Omega, \qquad u = 0 \quad \mathrm{on} \quad \partial\Omega; \tag{1.1}$$

precise conditions on $f : \Omega \to \mathbb{R}^d$ and $A : \Omega \to \mathbb{R}^{d \times d}$ are specified later.

Discontinuous Galerkin finite elements have received much attention due to their flexibility and ability to incorporate physical properties. We refer to [5] for a unified analysis of discontinuous Galerkin methods, [17] for a discussion of the inherent stability mechanism of those methods and [22] for comparison on their relative performances. For earlier work on interior penalty methods on elliptic problems we refer to Douglas and Dupont [28], Baker [9], Babŭska and Zlámal [7], Wheeler [51], and Arnold [4]. For further studies, we refer for instance to [18, 29, 34, 44, 35] and the references therein. However, fewer references are available for the a posteriori error analysis: see [1] for a review, [13, 47] regarding L^2-norm or function error estimation, and [33, 20, 36] regarding energy norm estimation. To the best of our knowledge, the only results concerning convergence of adaptive discontinuous finite element methods (ADFEM) have been obtained recently by Karakashian and Pascal [37] and Hoppe, Kanschat, and Warburton [31].

The ADFEM proposed here is based on a residual type estimator and consists of loops of the form

$$\mathrm{SOLVE} \ \to \mathrm{ESTIMATE} \to \mathrm{MARK} \ \to \mathrm{REFINE}. \tag{1.2}$$

Given a grid $\mathcal{T}_k$, SOLVE computes the solution U_k of the discrete problem

$$a_{\mathcal{T}_k}(U_k, V) = F_{\mathcal{T}_k}(V), \qquad \forall V \in \mathbb{V}(\mathcal{T}_k),$$

where $a_{\mathcal{T}_k}$ is the bilinear form over $\mathcal{T}_k$ associated with (1.1). The procedure ESTI-MATE computes the error indicators $\eta_{\mathcal{T}_k}(U_k, T)$, for all $T \in \mathcal{T}_k$. These indicators are used by the procedure MARK, which is the marking strategy originally introduced by Dörfler [27] and consisting of selecting a subset $\mathcal{M}_k$ of $\mathcal{T}_k$ satisfying

$$\eta_{\mathcal{T}_k}(U_k, \mathcal{M}_k) \geq \theta \, \eta_{\mathcal{T}_k}(U_k, \mathcal{T}_k),$$

where $0 < \theta < 1$ is given. Finally, REFINE subdivides each marked element $\rho \geq 1$ times and ensures that the level of nonconformity remains uniformly bounded.

The symmetric IPDG formulation considered by Arnold [4] reads: given a triangulation $\mathcal{T}$ of Ω, with skeleton Σ (set of interior interelement sides), seek $U \in \mathbb{V}(\mathcal{T})$ satisfying

$$a_{IP}(U, V) := (A\nabla U, \nabla V)_{\mathcal{T}} - \langle \{\!\{A\nabla U\}\!\}, [\![V]\!] \rangle_{\Sigma}$$
$$- \langle \{\!\{A\nabla V\}\!\}, [\![U]\!] \rangle_{\Sigma} + \gamma \langle h^{-1}[\![U]\!], [\![V]\!] \rangle_{\Sigma} = (f, V)_{\mathcal{T}}, \qquad \forall V \in \mathbb{V}(\mathcal{T}), \quad (1.3)$$

where γ is a positive penalty parameter and $\{\!\{\cdot\}\!\}, [\![\cdot]\!]$ stand for the average and jump operators; for precise definitions consult Section 2.1. As in [32], we employ a *lifting operator* [12, 5, 18, 46, 33] in Section 2.2 to replace $a_{IP}(\cdot, \cdot)$ by a different bilinear form $a_{\mathcal{T}}(\cdot, \cdot)$ and thereby obtain an equivalent discrete formulation to (1.3) which is also well defined for u under the minimal regularity $u \in H^1(\Omega)$ and $\mathrm{div}(Au) \in L^2(\Omega)$.

Karakashian and Pascal [37] were able to prove that if γ is sufficiently large then the following contraction property holds

$$a_{IP}(u - U_{k+1}, u - U_{k+1}) \leq \alpha a_{IP}(u - U_k, u - U_k), \qquad (1.4)$$

where $\alpha < 1$ and U_k, U_{k+1} are two consecutive solutions of the ADFEM on tetrahedral meshes. To obtain (1.4), the following assumptions were made in [37]
- f belongs to the finite element space and A is the identity matrix;
- the solution of (1.1) possesses the extra regularity $u \in H^2(\Omega)$;
- two successive subdivisions are not too far from each other;
- a relatively large amount of new nodes (12 vertices for piecewise linear elements in 2D) must be created by refining each marked element.

Hoppe et al. [31] improved upon [37]: first the refinement procedure consists of just one bisection per marked element, and second the right-hand side is in L^2. However,
- [31] assumes that the ensuing data oscillation contracts relative to itself between consecutive iterates, which is not guaranteed when marking only by the estimator;
- the technique used in [31] for error analysis is based on the Crouzeix-Raviart element, and thereby applies only to *conforming* meshes of $\Omega \subset \mathbb{R}^2$.

To enforce the aforementioned contraction of data oscillation one would need to mark also by oscillation. Unfortunately, this would lead to separate marking and, as discussed by Cascón et al [21], to the risk of getting sub-optimal meshes.

In this paper we extend and improve [31, 37] in several respects. Our novel contributions are to
- deal with general data $f : \Omega \to \mathbb{R}$, $A : \Omega \to \mathbb{R}^{d \times d}$, with regularity specified in Section 2.1.1, and non-convex bounded polyhedral domains of $\mathbb{R}^d$, $d \geq 2$. Since A may be discontinuous, this entails dealing with regularity $u \in H_0^1(\Omega)$ and $A\nabla u \in H(\mathrm{div})$ for the solution u of (1.1), which is minimal if $f \in L^2(\Omega)$;
- allow different types of nonconforming subdivisions such as tetrahedral or hexahedral meshes with hanging nodes and examine the complexity of REFINE with fixed level of non-conformity;

- refine each marked element using only one subdivision, either quad-refinement for hexahedral meshes, or red-refinement and bisection for tetrahedral meshes;
- prove a contraction property of the ADFEM, without further assumptions on REFINE, for the sum of energy error and scaled error estimator;
- show that the approximation classes consisting of continuous and discontinuous finite elements are equivalent;
- derive a quasi-optimal asymptotic rate of convergence for the ADFEM, which seems to be the first result of this type in the literature for dG methods;
- obtain a quasi-optimal asymptotic rate of convergence for the *continuous* Galerkin method on (hexahedral and tetrahedral) *nonconforming* meshes, as a by-product.

We would like to point out that [37, 31] consider mixed boundary conditions, but in the present work we restrict ourselves to homogeneous Dirichlet boundary condition to simplify the already technical presentation. Similarly, a reaction term of the type $c\, u$ with $0 \leq c \in L^\infty(\Omega)$ could have been added to the development as in [21] without changing the essence.

Our first Main Result concerns the reduction of the sum of energy error and scaled error estimator between two consecutive adaptive loops (1.2):

MAIN RESULT 1. *There exist a penalty threshold γ_m, a scaling factor $\delta > 0$, and a contraction constant $0 < \alpha < 1$ such that the scaled sum of energy error $a_{\mathcal{T}_k}(u - U_k, u - U_k)$ and estimator $\eta_k^2(U_k, \mathcal{T}_k)$ on mesh $\mathcal{T}_k$ contracts:*

$$
\begin{aligned}
a_{\mathcal{T}_{k+1}}(u - U_{k+1}, u - U_{k+1}) &+ \delta \eta_{k+1}^2(U_{k+1}, \mathcal{T}_{k+1}) \\
&\leq \alpha \left(a_{\mathcal{T}_k}(u - U_k, u - U_k) + \delta \eta_k^2(U_k, \mathcal{T}_k) \right).
\end{aligned}
\tag{1.5}
$$

A few comments are in order. The contraction property (1.5) is in the spirit of [21] in that it extends the ideas developed by Cascón et al for a continuous and conforming framework to dG (as well as cG) with *nonconforming* meshes. As a consequence of the coercivity property (Lemma 2.2) the adaptive finite element solution U_k converges in the broken energy norm, defined in (2.7), to the continuous solution of (1.1). The proof of (1.5), given in Section 4, does not rely on a lower bound for the error and, therefore, the *interior node property* as well as *marking by oscillation* can both be avoided; their significance for cG is pointed out in [42, 48, 21]. The estimator takes care of possible oscillations of $f \in L^2(\Omega)$ and A as well as possible increase of the error due to mesh refinement. A new major difficulty, compared to cG, is the presence of the negative power of the local meshsize h in the term $\left\| h^{-1/2} \llbracket U \rrbracket \right\|_\Sigma$, which appears in the upper bound (Lemma 3.1) and when comparing two successive solutions of the ADFEM (Corollary 4.3). To overcome this difficulty we need to

- resort to an idea of [2, 37] to control jumps by the estimator (Lemma 3.3);
- study the behavior of REFINE between consecutive refinements (Corollary 4.1).

The contraction property (1.5) is a crucial ingredient to prove *quasi-optimality* of the ADFEM. Let $\operatorname{osc}_{\mathcal{T}_k}(U_k, \mathcal{T}_k)$ be the oscillation over $\mathcal{T}_k$ (see Section 2.1) and let $\mathbb{A}_s$ be a suitable approximation class (see Section 5.1). Roughly speaking, (u, f, A) belongs to $\mathbb{A}_s$ if the following happens: given an arbitrary tolerance $\epsilon > 0$, it is possible to find a nonconforming refinement $\mathcal{T}$ of $\mathcal{T}_0$ with $O(\epsilon^{-s})$ DOFs so that the *total error* is bounded by ϵ, the latter being the sum of the broken energy error $|||u - U|||_{\mathcal{T}}$ and the oscillation $\operatorname{osc}_{\mathcal{T}}(U, \mathcal{T})$.

MAIN RESULT 2. *If $(u, f, A) \in \mathbb{A}_s$ with $s > 0$, then there holds*

$$
|||u - U_k|||_{\mathcal{T}_k} + \operatorname{osc}_{\mathcal{T}_k}(U_k, \mathcal{T}_k) \preceq \#\mathcal{T}_k^{-s}.
\tag{1.6}
$$

To the best of our knowledge no convergence rate for dG methods is yet known; we refer to [15, 48, 21] for results concerning cG methods. The main difficulty of the proof is again caused by the term $\left\|h^{-1/2}[\![U]\!]\right\|_\Sigma$. To see this, let ϵ be the error associated to $(U_k, \mathcal{T}_k)$ in step k of the ADFEM, and let $(U_*, \mathcal{T}_*)$ be the optimal solution-mesh pair for this tolerance ϵ. Since the subdivisions $\mathcal{T}_*$ and $\mathcal{T}_k$ are a priori unrelated, the negative power of mesh size prevents direct estimation of the jump terms as in the proof of the Main Result 1 (where at most ρ refinements separate two consecutive subdivisions). The cure to this issue resides on the following property of intrinsic interest and shown in Proposition 5.2:

> *the approximation classes for cG and dG methods are equivalent on noncon-forming partitions with bounded level of nonconformity,*

namely they exhibit the same asymptotic rate of convergence. Therefore, comparing with continuous optimal approximations, the effect of the jump term becomes manageable. The advantages of dG over cG are thus to be found in local conservation and stabilization properties rather than accuracy.

We finally emphasize that the Main Results 1 and 2 extend those of Cascón et al [21] for cG on conforming meshes to nonconforming meshes with hanging nodes.

The outline of this article is as follows. In Section 2 we start by setting the notations and definitions and conclude with the interior penalty method. In Section 3 we discuss each module of (1.2) and give the ingredients for the proof of the main results. Sections 4 and 5 are dedicated to the proofs of the main results. To unify the treatment of the various refinement strategies, we impose seven conditions below, as described in Table 1.1:

Condition	1	2	3	4	5	6	7
Page	5	12	17	18	24	25	27

TABLE 1.1
Page references of the Conditions required to obtain Main Results 1 and 2.

In Section 6, we verify that these conditions are valid for quad-refinement of hexahedral meshes, and bisection of simplicial meshes; red-refinement of simplicial meshes is somewhat in between the two and is not discussed in detail. In particular, note that Main Result 1 relies on Condition **3**, which restricts one step refinements, while Main Result 2 hinges also on Condition **4** (complexity of REFINE) and Condition **6** (overlay of two subdivisions). Finally, we conclude in Section 7 with a brief discussion of cG on nonconforming meshes with hanging nodes.

2. Interior Penalty Methods. Here we state the assumptions on data, discuss the notion of discontinuous finite elements on nonconforming meshes. The lifting operator [12, 5, 18, 46, 33] on nonconforming subdivisions is then introduced leading to the description of IPDG. We next define residual estimator and oscillation and conclude with an instrumental decomposition of the finite element space.

2.1. Notations and Preliminaries. In order to simplify the notation we write $f \preceq g$ whenever $f \le Cg$ with a constant C independent of the discretization parameters and the penalty parameter γ.

2.1.1. The Continuous Problem. Given an initial conforming partition $\mathcal{T}_0$ of Ω, we assume that $f \in L^2(\Omega)$ and

$$A \in \prod_{T \in \mathcal{T}_0} W^1_\infty(T)^{d \times d} \text{ is symmetric positive definite} \atop \text{with eigenvalues in } [a_m, a_M], \text{ where } 0 < a_m \le a_M < \infty, \tag{2.1}$$

the last condition meaning

$$a_m \left| \xi \right|^2 \le A\xi \cdot \xi \le a_M \left| \xi \right|^2, \quad \forall \xi \in \mathbb{R}^d, \ a.e. \text{ in } \Omega.$$

From now on, we shall not specify the dependency on a_m, a_M of the constants appearing below. The restriction to conforming initial subdivision is made in order to avoid technicalities, so nonconforming initial subdivisions can be treated similarly.

Invoking the Lax-Milgram Lemma, the above assumptions ensure that the weak formulation of (1.1), namely

$$u \in H_0^1(\Omega): \qquad \int_\Omega A\nabla u \cdot \nabla v = \int_\Omega fv, \qquad \forall v \in H_0^1(\Omega) \qquad (2.2)$$

possesses a unique solution. Note that no additional regularity on the solution of (1.1) is required: arbitrary polyhedral domains and discontinuous A are examined here.

2.1.2. Types of Subdivisions of Ω. The theoretical framework developed hereafter is not restricted to a special type of partitions of Ω. In fact, *nonconforming* partitions made of tetrahedrons (triangles when $d = 2$) or hexahedra (quadrilaterals when $d = 2$) are considered for $d \ge 2$. We now provide three relevant examples.

Quad Refinements. Elements are quadrilaterals (or hexahedra when $d = 3$), and refinement consists of subdividing them into 2^d children; see Fig. 2.1 for $d = 2$.
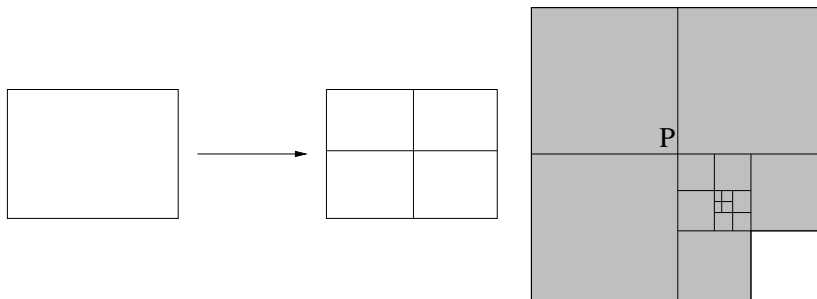

FIG. 2.1. *(Left) Quad refinement of a quadrilateral onto four children. (Right) Typical patch present in subdivisions generated by such refinement procedures allowing hanging nodes. The shaded elements depict the domain of influence of the proper node P; see Condition 7 in Section 6.*

Red Refinements. We only provide the description of this procedure when $d = 2$ because it is rather involved otherwise. A triangle is subdivided into four triangles by joining the midpoints of the element sides. Refer to Fig. 2.2 for an example of such a refinement together with a typical patch.

Bisections. This refinement procedure is studied in [15] for $d = 2$ and in [49] for $d > 2$; see also the survey [43]. Bisection splits a given element T into two children, regardless of d, upon connecting the midpoint of an edge with the vertices of T off such edge. Fig. 2.3 depicts the effect of one bisection along with a typical patch for $d = 2$. The specific choice of the refinement edge solely depends on the initial partition $\mathcal{T}_0$ and the bisection rules [49, condition (b) of Section 4].

From now on, we use the three dimensional denomination for elements of the partitions even when discussing the two dimensional case. Having the aforementioned examples in mind we assume the following behavior of the refinement procedure.

Condition 1 (Atomic Refinement). *The subdivision of an element into children is called atomic refinement and is dictated exclusively by the initial partition $\mathcal{T}_0$ and*
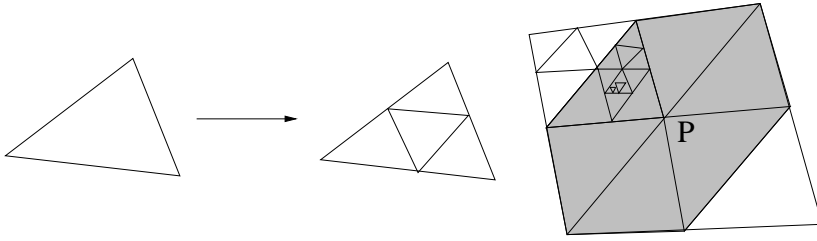
FIG. 2.2. *(Left) Red-refinement of a triangle onto four children. (Right) Typical patch present in subdivisions generated by such refinement procedures allowing hanging nodes. The shaded elements depict the domain of influence of the proper node P; see Condition **7** in Section 6.*
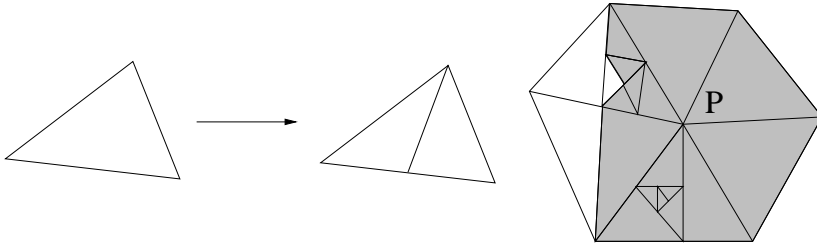

FIG. 2.3. *(Left) Bisection of one element onto two children. (Right) Typical patch present in subdivisions generated by such refinement procedures allowing hanging nodes. The shaded elements depict the domain of influence of the proper node P; see Condition **7** in Section 6.*

the refinement rules. If T' is a child of $T \in \mathcal{T}$ and $h_T = |T|^{1/d}$ is the meshsize, then there exist constants $\beta_m \leq \beta_M < 1$ only depending on the dimension d such that

$$\beta_m h_T \leq h_{T'} \leq \beta_M h_T. \tag{2.3}$$

Moreover, the shape regularity of any refinement $\mathcal{T}$ of $\mathcal{T}_0$ is determined by that of $\mathcal{T}_0$.

The above condition obviously holds for quad- and red-refinement since there is no ambiguity on how to subdivide an element and shape regularity is preserved. For bisection, instead, Stevenson's generalization [49] of the newest vertex rule studied by Binev et al [15] implies that the refinement edge only depends on the initial labeling of $\mathcal{T}_0$. Moreover, shape regularity is maintained [10, 38, 39, 41, 49, 50].

We now resort to the language of graph theory. Starting from an initial conforming mesh $\mathcal{T}_0$, we represent the refinement procedure by an infinite forest $\mathbb{F}$, called *master forest*, whose roots are the elements of $\mathcal{T}_0$. For instance, for bisections $\mathbb{F}$ is a binary tree [15, 21] regardless of d whereas for $d = 2$ quad and red refinements $\mathbb{F}$ is a quadtree. Let $\mathcal{F}_0$ be the initial forest corresponding to $\mathcal{T}_0$. If N_C is the number of children created by refining an element, a finite subset $\mathcal{F} \subset \mathbb{F}$ is called *forest* if $\mathcal{F}_0 \subset \mathcal{F}$ and

- all nodes in $\mathcal{F} \setminus \mathcal{F}_0$ have a predecessor;
- all nodes in $\mathcal{F}$ have N_C successors or none.

For bisections $N_C = 2$ whereas for quad refinements $N_C = 2^d$. The successors of an element T are called *children*.

Under Condition **1**, each forest $\mathcal{F}$ corresponds to a unique (possibly nonconforming) partitions $\mathcal{T}(\mathcal{F})$ of Ω by defining $\mathcal{T}(\mathcal{F})$ as the set of leaf nodes; the set of all such partitions is denoted by $\mathbb{T}$. We also denote by $\Sigma(\mathcal{T})$ the set of sides (faces) of a partition $\mathcal{T}(\mathcal{F})$, and use the abbreviation $\mathcal{T} := \mathcal{T}(\mathcal{F})$, $\Sigma := \Sigma(\mathcal{T})$ when no confusion is possible. The *level* (or generation) $\ell(T)$ of an element $T \in \mathcal{T}$ is the number of

atomic refinements needed to obtain T starting from an element of the initial mesh. Also we define the meshsize h to be the (piecewise constant) function

$$h|_K = |K|^{1/d'} \qquad \forall K \in \mathcal{T} \ (d' = d), \quad K \in \Sigma \ (d' = d-1). \tag{2.4}$$

If $\mathcal{F} \subset \mathcal{F}_*$ are two forests, we call $\mathcal{T}_* = \mathcal{T}(\mathcal{F}_*)$ a refinement of $\mathcal{T} = \mathcal{T}(\mathcal{F})$ and denote it as $\mathcal{T} \le \mathcal{T}_*$. We define $\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*} := \mathcal{T} \setminus (\mathcal{T}_* \cap \mathcal{T})$ to be the set of elements of $\mathcal{T}$ which have been refined to obtain $\mathcal{T}_*$.

We now discuss the overlay of two subdivisions. Given two forests $\mathcal{F}$, $\mathcal{F}_* \subset \mathbb{F}$ we define $\mathcal{F} \cup \mathcal{F}_* \subset \mathbb{F}$ to be the union of the nodes of $\mathcal{F}$ and $\mathcal{F}_*$. The resulting subdivision is called the *overlay* of $\mathcal{T}(\mathcal{F})$ and $\mathcal{T}(\mathcal{F}_*)$ and is denoted by

$$\mathcal{T}(\mathcal{F}) \oplus \mathcal{T}(\mathcal{F}_*) := \mathcal{T}(\mathcal{F} \cup \mathcal{F}_*).$$

Note that the above construction ensures that $\mathcal{T}(\mathcal{F}) \oplus \mathcal{T}(\mathcal{F}_*)$ is a refinement of both, $\mathcal{T}(\mathcal{F})$ and $\mathcal{T}(\mathcal{F}_*)$. Moreover following Lemma 3.7 (second part) in [21] we directly obtain that under Condition **1** and for $\mathcal{T}_A, \mathcal{T}_B$ two refinements of $\mathcal{T}_0$

$$\#(\mathcal{T}_A \oplus \mathcal{T}_B) \le \#\mathcal{T}_A + \#\mathcal{T}_B - \#\mathcal{T}_0. \tag{2.5}$$

A condition on the overlay of two refinements of $\mathcal{T}_0$, namely Condition **6**, is imposed in Section 5.3 when discussing the optimality of the algorithm.

2.1.3. Discontinuous Finite Elements and Energy Spaces. Let $n \ge 1$ be a given polynomial degree. The basic polynomial space $\mathbb{P}^n$ depends on the element type. If the master element $\widehat{K}$ is a d-simplex, then $\mathbb{P}^n(\widehat{K})$ indicates the set of polynomials of total degree $\le n$. If $\widehat{K}$ is a hypercube instead, then $\mathbb{P}^n(\widehat{K})$ stands for the set of polynomials of degree $\le n$ in each variable. Finally, iso-parametric mappings define the polynomial spaces in each element of $\mathcal{T}$ and side of Σ; see [16]. From now on, we will not specify the dependency on n of the constants appearing below.

We associate a discontinuous finite element space

$$\mathbb{V}(\mathcal{T}) := \prod_{T \in \mathcal{T}} \mathbb{P}^n(T)$$

to any partition $\mathcal{T}$ of Ω. To shorten notation, we define for any set of elements $\tau \subset \mathcal{T}$

$$(u, v)_\tau := \sum_{T \in \tau} (u, v)_T \qquad \forall u, v \in L^2(\tau).$$

Likewise, for any set of sides $\sigma \subset \Sigma$ we let

$$\langle \psi, \zeta \rangle_\sigma := \sum_{S \in \sigma} \langle \psi, \zeta \rangle_S \qquad \forall \psi, \zeta \in L^2(\sigma).$$

When no confusion is possible the same notation is used for functions in $L^2(\tau)^d$ or $L^2(\sigma)^d$. At the same time, we also define the corresponding "broken norms"

$$\|v\|_\tau := (v, v)_\tau^{1/2}, \qquad \|\psi\|_\sigma := \langle \psi, \psi \rangle_\sigma^{1/2}.$$

The broken H^1 space

$$\mathbb{E}(\mathcal{T}) := \prod_{T \in \mathcal{T}} H^1(T) \tag{2.6}$$

is referred to as the *energy space*. We now recall the mean and jump operators. We let $T^{\pm} \in \mathcal{T}$ satisfy $S = T^+ \cap T^- \in \Sigma$, and let $\nu_{T\pm}$ the outer unit normal to $T^{\pm}$. Given $v \in \mathbb{E}(\mathcal{T})$ and $\mathbf{w} \in \mathbb{E}(\mathcal{T})^d$, we define

$$\{\!\{v\}\!\}_S := \frac{1}{2}(v^+ + v^-), \qquad \{\!\{\mathbf{w}\}\!\}_S := \frac{1}{2}(\mathbf{w}^+ + \mathbf{w}^-),$$

and

$$[\![v]\!]_S := v^+\nu_{T+} + v^-\nu_{T-}, \qquad [\![\mathbf{w}]\!]_S := \mathbf{w}^+ \cdot \nu_{T+} + \mathbf{w}^- \cdot \nu_{T-}.$$

By convention, if $S \in \Sigma$ is a boundary side, then we set

$$\{\!\{v\}\!\}_S = v, \qquad \{\!\{\mathbf{w}\}\!\}_S = \mathbf{w}, \qquad [\![v]\!]_S = v\,\nu_\Gamma,$$

with ν_Γ denoting the unit outer normal on Γ. Note that we leave $[\![\mathbf{w}]\!]_S$ undefined.

Finally, the energy space $\mathbb{E}(\mathcal{T})$, is endowed with the mesh dependent *energy norm*

$$|||v|||_{\mathcal{T}}^2 := \left\|A^{1/2}\nabla v\right\|_{\mathcal{T}}^2 + \gamma \left\|h^{-1/2}\,[\![v]\!]\right\|_{\Sigma}^2, \qquad \forall v \in \mathbb{E}(\mathcal{T}). \tag{2.7}$$

2.2. The Discrete Formulation. We now introduce an alternate expression $a_{\mathcal{T}}(\cdot,\cdot)$ for the symmetric bilinear form $a_{IP}(\cdot,\cdot)$, which is well defined on the broken energy space $\mathbb{E}(\mathcal{T}) \times \mathbb{E}(\mathcal{T})$ of (2.6), and so on $H_0^1(\Omega) \times H_0^1(\Omega)$. To this end, we need the following lifting operator.

2.2.1. Lifting Operator. Let $L_{\mathcal{T}} : \mathbb{E}(\mathcal{T}) \to \mathbb{V}(\mathcal{T})^d$ be defined by [32]

$$\int_\Omega L_{\mathcal{T}}(v) \cdot AW := \langle [\![v]\!], \{\!\{AW\}\!\}\rangle_\Sigma, \qquad \forall W \in \mathbb{V}(\mathcal{T})^d. \tag{2.8}$$

Note that $L_{\mathcal{T}}$ is well defined on the whole broken energy space $\mathbb{E}(\mathcal{T})$, and that $L(v) = 0$ for all $v \in H_0^1(\Omega)$. We refer to [12, 5, 18, 46, 33] for more details on the lifting operator. In particular, we have the following stability result for $L_{\mathcal{T}}$, proved in [46] when A is the identity. The proof can be easily modified to incorporate the uniformly symmetric positive definite matrix A in definition (2.8), as done below for completeness.

LEMMA 2.1 (Stability of the Lifting Operator). *Assume that A satisfies* (2.1). *There exists a constant C_L only depending on the shape regularity of $\mathcal{T}$ such that*

$$\|L_{\mathcal{T}}(v)\|_{L^2(\Omega)} \le C_L \left\|h^{-1/2}\,[\![v]\!]\right\|_{\Sigma} \qquad \forall v \in \mathbb{E}(\mathcal{T}).$$

Proof. We first recall the inverse estimate $\|h^{1/2}W\|_{L^2(\partial T)} \preceq \|W\|_{L^2(T)}$ for all $W \in \mathbb{V}(\mathcal{T})^d$. In view of the positive definiteness assumptions (2.1) on A, this bound implies the global estimate $\|h^{1/2}\,\{\!\{AW\}\!\}\|_\Sigma \preceq \|A^{1/2}W\|_{L^2(\Omega)}$. Finally, it suffices to choose $W = L_{\mathcal{T}}(v)$ in (2.8), to obtain

$$\left\|A^{1/2}L_{\mathcal{T}}(v)\right\|_{\mathcal{T}}^2 = \int_\Omega L_{\mathcal{T}}(v) \cdot AL_{\mathcal{T}}(v) \preceq \left\|h^{-1/2}\,[\![v]\!]\right\|_\Sigma \left\|A^{1/2}L_{\mathcal{T}}(v)\right\|_{\mathcal{T}}.$$

This yields the desired result. $\square$

2.2.2. The Discrete Problem. Let $\gamma > 0$ be the penalty parameter, and let $a_{\mathcal{T}}(\cdot, \cdot) : \mathbb{E}(\mathcal{T}) \times \mathbb{E}(\mathcal{T}) \to \mathbb{R}$ and $F_{\mathcal{T}}(\cdot) : \mathbb{E}(\mathcal{T}) \to \mathbb{R}$ be defined by

$$a_{\mathcal{T}}(v, w) := (A\nabla v, \nabla w)_{\mathcal{T}} - (L_{\mathcal{T}}(w), A\nabla v)_{\mathcal{T}}$$
$$- (L_{\mathcal{T}}(v), A\nabla w)_{\mathcal{T}} + \gamma \left(h^{-1} \llbracket v \rrbracket, \llbracket w \rrbracket \right)_{\Sigma} \qquad (2.9)$$

and

$$F_{\mathcal{T}}(v) := (f, v)_{\mathcal{T}}, \qquad (2.10)$$

where the lifting operator $L_{\mathcal{T}} : \mathbb{E}(\mathcal{T}) \to \mathbb{V}(\mathcal{T})^d$ is given in (2.8). The discontinuous Galerkin interior penalty method for problem (2.2) reads: given a subdivision $\mathcal{T} \in \mathbb{T}$, $f \in L^2(\Omega)$, and A satisfying (2.1), seek

$$U \in \mathbb{V}(\mathcal{T}) : \quad a_{\mathcal{T}}(U, V) = F_{\mathcal{T}}(V), \qquad \forall V \in \mathbb{V}(\mathcal{T}). \qquad (2.11)$$

Note that the definition of the lifting operator (2.8) directly implies that

$$a_{\mathcal{T}}(V, W) = a_{IP}(V, W), \qquad \forall V, W \in \mathbb{V}(\mathcal{T}),$$

where the bilinear form $a_{IP}(\cdot, \cdot)$ is given in (1.3). In contrast to (1.3), the lifting operator $L_{\mathcal{T}} : \mathbb{E}(\mathcal{T}) \to \mathbb{V}(\mathcal{T})^d$ ensures that the bilinear form $a_{\mathcal{T}}(\cdot, \cdot)$ of (2.9) is well defined on $\mathbb{E}(\mathcal{T}) \times \mathbb{E}(\mathcal{T})$ and in particular on $H_0^1(\Omega) \times H_0^1(\Omega)$. Unfortunately, this is achieved at the expense of consistency, namely the solution $u \in H_0^1(\Omega)$ does not satisfy (2.11). However, since $L_{\mathcal{T}}(v) = 0$ for all $v \in H_0^1(\Omega)$, u still satisfies the *partial* consistency relation:

$$a_{\mathcal{T}}(u, v) = F_{\mathcal{T}}(v), \qquad \forall v \in H_0^1(\Omega). \qquad (2.12)$$

Moreover, elementwise integration by parts yields the error-residual relation

$$a_{\mathcal{T}}(u - U, v) = (R(U), v)_{\mathcal{T}} - \langle J(U), v \rangle_{\Sigma} + (L_{\mathcal{T}}(U), A\nabla v)_{\mathcal{T}}, \quad \forall v \in H_0^1(\Omega), \quad (2.13)$$

where the interior and jump residuals, $R(V)$ and $J(V)$ for any $V \in \mathbb{V}(\mathcal{T})$, are

$$R(V)|_T := f + \operatorname{div}(A\nabla V) \quad \forall T \in \mathcal{T}, \qquad J(V)|_S := \llbracket A\nabla V \rrbracket_S \quad \forall S \in \Sigma. \qquad (2.14)$$

We now review properties of the bilinear form that can be found in [45, Proposition 3.1], [33, Lemmas 4.3 and 4.4], and [8, Proposition 2.11]. We give a proof for completeness.

LEMMA 2.2 (Coercivity and Continuity). *Let $\mathcal{T}$ be a refinement of $\mathcal{T}_0$. There exist constants $\gamma_1, C_{cont}, C_{coer}$ such that for $\gamma \geq 1$ there holds*

$$a_{\mathcal{T}}(v, w) \leq C_{cont} |||v|||_{\mathcal{T}} |||w|||_{\mathcal{T}} \quad \forall v, w \in \mathbb{E}(\mathcal{T}), \qquad (2.15)$$

and for $\gamma \geq \gamma_1$ there holds

$$C_{coer} |||v|||_{\mathcal{T}}^2 \leq a_{\mathcal{T}}(v, v) \quad \forall v \in \mathbb{E}(\mathcal{T}). \qquad (2.16)$$

Proof. We focus on (2.16), since (2.15) readily follows upon using the Cauchy-Schwarz inequality for each term in $a_{\mathcal{T}}(v, w)$. For $v \in \mathbb{E}(\mathcal{T})$ we have

$$a_{\mathcal{T}}(v, v) = (A\nabla v, \nabla v)_{\mathcal{T}} - 2(L_{\mathcal{T}}(v), A\nabla v)_{\mathcal{T}} + \gamma \left(h^{-1} \llbracket v \rrbracket, \llbracket v \rrbracket \right)_{\Sigma}.$$

Lemma 2.1, in conjunction with (2.1) and Young inequality, yields

$$(L_{\mathcal{T}}(v), A\nabla v)_{\mathcal{T}} \leq \frac{1}{4} \left\| A^{1/2} \nabla v \right\|_{\mathcal{T}}^2 + C_L^2 a_M \left\| h^{-1/2} \, [\![v]\!] \right\|_{\Sigma}^2,$$

where C_L is the constant given in Lemma 2.1. Inserting this into $a_{\mathcal{T}}(v, v)$, setting $\gamma_1 = 4C_L^2 a_M$, and recalling the definition (2.7), we obtain (2.16) with $C_{\mathrm{coer}} = 1/2$. □

Continuity (2.15) and coercivity (2.16) of the bilinear form $a_{\mathcal{T}}(\cdot, \cdot)$ ensure the existence of a solution of problem (2.11) by invoking the Lax-Milgram Lemma.

REMARK 2.3 (Penalty Parameter). *The parameter γ in (2.9) must be chosen sufficiently large to ensure coercivity. Actually, we will assume later that γ is even bigger to prove the Main Results 1 and 2.*

2.2.3. Estimator and Oscillation. We use residual based estimators as the kernel of the ADFEM strategy. In view of (2.14), for any $T \in \mathcal{T}$ and $V \in \mathbb{V}(\mathcal{T})$, let

$$\eta_{\mathcal{T}}^2(V, T) := \|hR(V)\|_T^2 + \left\| h^{1/2} J(V) \right\|_{\partial T \cap \Omega}^2,$$

where h is defined in (2.4) and, for any subset τ of $\mathcal{T}$, let

$$\eta_{\mathcal{T}}^2(V, \tau) := \sum_{T \in \tau} \eta_{\mathcal{T}}^2(V, T); \tag{2.17}$$

note that the jump $\|h^{-1/2} [\![V]\!] \|_{\partial T \cap \Omega}$ is not part of $\eta_{\mathcal{T}}(V, T)$. A similar notion is useful for the matrix A

$$\eta_{\mathcal{T}}(A, \tau) := \max_{T \in \tau} \eta_{\mathcal{T}}(A, T), \tag{2.18}$$

where

$$\eta_{\mathcal{T}}^2(A, T) := \|h \, \mathrm{div} \, A\|_{L^\infty(T)}^2 + \|A\|_{L^\infty(T)}^2.$$

For $m \in \mathbb{N} \setminus \{0\}$, we recall that $\mathbb{P}^m(K)$ is the space of polynomials of degree $\leq m$ on d or $(d-1)$-simplices K or polynomials of degree $\leq m$ in each variable on d or $(d-1)$-quadrilaterals. Moreover, for $p \in [2, \infty]$ we let $P_m^p := Id - \Pi_m^p$, where Π_m^p is the L^p−best approximation operator onto $\mathbb{P}^m(K)$ where K depends on the context.

We define the *oscillation* for $V \in \mathbb{V}(\mathcal{T})$ and $T \in \mathcal{T}$ to be

$$\mathrm{osc}_{\mathcal{T}}^2(V, T) := \left\| h \, P_m^2 R(V) \right\|_{L^2(T)}^2 + \left\| h^{1/2} \, P_{m'}^2 J(V) \right\|_{L^2(\partial T \cap \Omega)}^2,$$

and

$$\mathrm{osc}_{\mathcal{T}}^2(V, \tau) := \sum_{T \in \tau} \mathrm{osc}_{\mathcal{T}}^2(V, T), \tag{2.19}$$

for any subset τ of $\mathcal{T}$. The specific choice of m and m' depends on the element type:
 • simplices: $m = 2n - 2, m' = 2n - 1$
 • quadrilaterals: $m = 2n - 1, m' = 2n$,
because $\nabla \mathbb{P}^n(K) \subset [\mathbb{P}^n(K)]^d$ for quadrilaterals instead of $\nabla \mathbb{P}^n(K) \subset [\mathbb{P}^{n-1}(K)]^d$. Moreover, both choices yield Lemmas 2.4 and 2.5 below. Note that due to the definition of the L^2-projection, the estimator dominates the oscillation, namely

$$\mathrm{osc}_{\mathcal{T}}(V, T) \leq \eta_{\mathcal{T}}(V, T), \tag{2.20}$$

for $V \in \mathbb{V}(\mathcal{T})$ and $T \in \mathcal{T}$. The oscillation counterpart related to A is given by

$$\text{osc}^2_{\mathcal{T}}(A, T) := \left\| h \, P^\infty_{n-1} \text{div} A \right\|^2_{L^\infty(T)} + \left\| P^\infty_n A \right\|^2_{L^\infty(T)}, \quad \forall T \in \mathcal{T}$$

and

$$\text{osc}_{\mathcal{T}}(A, \tau) := \max_{T \in \tau} \text{osc}_{\mathcal{T}}(A, T), \quad \forall \tau \subset \mathcal{T}.$$

Both $\eta_{\mathcal{T}}(V, \mathcal{T})$ and $\eta_{\mathcal{T}}(A, \mathcal{T})$ enjoy the following monotonicity property:

$$\eta_{\mathcal{T}_*}(V, \mathcal{T}_*) \leq \eta_{\mathcal{T}}(V, \mathcal{T}), \qquad \eta_{\mathcal{T}_*}(A, \mathcal{T}_*) \leq \eta_{\mathcal{T}}(A, \mathcal{T}), \qquad \forall \mathcal{T}_* \geq \mathcal{T}. \tag{2.21}$$

The latter is used to prove the following two lemmas, which are instrumental in deriving the main results. Even though the lemmas are stated for general meshes $\mathcal{T}$ and $\mathcal{T}_*$ made of tetrahedra or hexahedra, their original proofs in [21] for conforming simplices are still valid and are thus omitted.

LEMMA 2.4 (Estimator Reduction). *Assume that the refinement strategy satisfies Condition* **1**. *Let* $\mathcal{T}$, $\mathcal{T}_* \in \mathbb{T}$ *be two refinements of* $\mathcal{T}_0$ *with* $\mathcal{T} \leq \mathcal{T}_*$. *Let* $\mathcal{R} := \mathcal{R}_{\mathcal{T} \to \mathcal{T}_*}$ *be the set of elements of* $\mathcal{T}$ *refined to obtain* $\mathcal{T}_*$ *and* $\lambda := 1 - \beta_M$, *where* $0 < \beta_M < 1$ *is the constant appearing in Condition* **1**. *There exists a constant* C_{er} *only depending on the shape regularity of* $\mathcal{T}_0$ *and* $\eta_{\mathcal{T}_0}(A, \mathcal{T}_0)$ *such that for* $V_* \in \mathbb{V}(\mathcal{T}_*)$, $V \in \mathbb{V}(\mathcal{T})$ *and for all* $\zeta > 0$, *there holds*

$$\eta^2_{\mathcal{T}_*}(V_*, \mathcal{T}_*) \ \leq \ (1 + \zeta) \left(\eta^2_{\mathcal{T}}(V, \mathcal{T}) - \lambda \eta^2_{\mathcal{T}}(V, \mathcal{R}) \right) + (1 + \zeta^{-1}) C_{\text{er}} \left\| \nabla(V_* - V) \right\|^2_{\mathcal{T}_*}.$$

Proof. See [21, Corollary 3.4]. □

LEMMA 2.5 (Perturbation of Oscillation). *Let* $\mathcal{T}$, $\mathcal{T}_* \in \mathbb{T}$ *with* $\mathcal{T} \leq \mathcal{T}_*$. *Then there exists a constant* Λ *only depending on the shape regularity of* $\mathcal{T}_0$ *and* $\text{osc}_{\mathcal{T}_0}(A, \mathcal{T}_0)$, *such that for all* $V \in \mathbb{V}(\mathcal{T})$ *and* $V_* \in \mathbb{V}(\mathcal{T}_*)$ *there holds*

$$\text{osc}^2_{\mathcal{T}}(V, \mathcal{T} \cap \mathcal{T}_*) \leq 2\text{osc}^2_{\mathcal{T}_*}(V_*, \mathcal{T} \cap \mathcal{T}_*) + \Lambda \left\| A^{1/2} \nabla(V_* - V) \right\|^2_{\mathcal{T}_*}.$$

Proof. See [21, Corollary 3.5]. □

2.2.4. Space Decomposition and Continuous Approximation. It turns out that the decomposition of $\mathbb{V}(\mathcal{T})$ into its continuous and discontinuous components is useful; see Remark 2.7. Let $\mathbb{V}^0(\mathcal{T}) := \mathbb{V}(\mathcal{T}) \cap H^1_0(\Omega)$ be the underlying conforming finite element space. We introduce the following orthogonal decomposition of $\mathbb{V}(\mathcal{T})$

$$\mathbb{V}(\mathcal{T}) := \mathbb{V}^0(\mathcal{T}) \oplus \mathbb{V}^\perp(\mathcal{T}), \tag{2.22}$$

where $\mathbb{V}^\perp(\mathcal{T})$ is the orthogonal complement of $\mathbb{V}^0(\mathcal{T})$ in $\mathbb{V}(\mathcal{T})$ with respect to the (discrete) scalar product $a_{\mathcal{T}}(\cdot, \cdot)$.

REMARK 2.6 (Partial Galerkin Orthogonality). *Invoking the partial consistency relation* (2.12) *and using the above definition of* $\mathbb{V}^0(\mathcal{T})$ *we readily deduce*

$$a_{\mathcal{T}}(u - U, V^0) = 0, \qquad \forall V^0 \in \mathbb{V}^0(\mathcal{T}), \tag{2.23}$$

where $u \in H^1_0(\Omega)$ *and* $U \in \mathbb{V}(\mathcal{T})$ *are the solution to* (1.1) *and* (2.11) *respectively. This partial Galerkin orthogonality will be instrumental in our subsequent discussion.*

REMARK 2.7 (Continuous Galerkin Solution). *Let $U \in \mathbb{V}(\mathcal{T})$ be the dG solution of (2.11). In view of (2.22), we write $U = U^0 + U^\perp$ where $U^0 \in \mathbb{V}^0(\mathcal{T})$ and $U^\perp \in \mathbb{V}^\perp(\mathcal{T})$. We claim that U^0 is the corresponding cG solution, i.e.*

$$a_\mathcal{T}(U^0, V^0) = F_\mathcal{T}(V^0), \qquad \forall V^0 \in \mathbb{V}^0(\mathcal{T}). \tag{2.24}$$

Indeed, since $\mathbb{V}^0(\mathcal{T})$ and $\mathbb{V}^\perp(\mathcal{T})$ are orthogonal with respect to $a_\mathcal{T}(\cdot, \cdot)$, we obtain for all $V = V^0 + V^\perp \in \mathbb{V}(\mathcal{T})$

$$(f, V^0) + (f, V^\perp) = (f, V) = a_\mathcal{T}(U, V) = a_\mathcal{T}(U^0, V^0) + a_\mathcal{T}(U^\perp, V^\perp).$$

Hence, this implies (2.24).

We now describe local properties of an interpolation operators onto $\mathbb{V}^0(\mathcal{T})$. To this end, let $T \in \mathcal{T}$ and $\omega(T) \subset \mathcal{T}$ denotes a discrete set including T so that

$$\text{diam}\,(\omega(T)) \leq Ch_T, \tag{2.25}$$

with $C > 0$ a constant only depending on the shape regularity of $\mathcal{T}$. Let $\sigma(T) = \Sigma \cap \omega(T)$ be the skeleton of $\mathcal{T}$ within $\omega(T)$. For latter use, we also define for a set of element $\tau \subset \mathcal{T}$

$$\omega(\tau) := \bigcup_{T \in \tau} \omega(T). \tag{2.26}$$

Condition 2 (Interpolation Operator onto $\mathbb{V}^0(\mathcal{T})$). *There exists an interpolation operator $I_\mathcal{T} : \mathbb{E}(\mathcal{T}) \to \mathbb{V}^0(\mathcal{T})$ and a constant C_I only depending on the shape regularity of $\mathcal{T}$ such that for all $T \in \mathcal{T}$ the following inequalities hold:*

$$\|v - I_\mathcal{T} v\|_{L^2(T)} \leq C_I \|hDv\|_{L^2(\omega(T))} \quad \forall v \in H_0^1(\Omega); \tag{2.27}$$

and for $|a| = 0, 1$

$$\|D^a(V - I_\mathcal{T} V)\|_{L^2(T)} \leq C_I \left\| h^{\frac{1-2|a|}{2}} \llbracket V \rrbracket \right\|_{L^2(\sigma(T))} \quad \forall V \in \mathbb{V}(\mathcal{T}), \tag{2.28}$$

The operator $I_\mathcal{T}$ is defined locally and preserves $\mathbb{V}^0(\mathcal{T})$ locally.

REMARK 2.8 (Level of Non-Conformity). *It turns out that the constant C_I in Condition 2 depends on the relative size of the support of the basis functions in $\mathbb{V}(\mathcal{T})$ and $\mathbb{V}^0(\mathcal{T})$; see Section 6.*

LEMMA 2.9 (Estimate on the Nonconforming Component). *Let γ_1 be given by Lemma 2.2. If $V = V^0 + V^\perp \in \mathbb{V}(\mathcal{T})$ according to (2.22), and Condition 2 holds, then there exists a constant $C > 0$ only depending on shape regularity of $\mathcal{T}$ such that for all $\gamma \geq \max(1, \gamma_1)$ there holds*

$$\left\| |V^\perp| \right\|_\mathcal{T} \leq C\gamma^{1/2} \left\| h^{-1/2} \llbracket V^\perp \rrbracket \right\|_\Sigma = C\gamma^{1/2} \left\| h^{-1/2} \llbracket V \rrbracket \right\|_\Sigma. \tag{2.29}$$

Proof. From the definition (2.22) of $\mathbb{V}^\perp(\mathcal{T})$ we obtain

$$a_\mathcal{T}(V^\perp, V^\perp) = \inf_{W^0 \in \mathbb{V}^0(\mathcal{T})} a_\mathcal{T}(V - W^0, V - W^0).$$

Hence, choosing $W^0 = I_\mathcal{T} V$ given in Condition **2**, and invoking Lemma 2.2 and (2.28), yield the desired result. $\square$

3. The Adaptive Procedure. We analyze here an adaptive dG method (AD-FEM), which consists of loops of the following form with counter $k \geq 0$:

$$\text{SOLVE} \ \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \ \rightarrow \text{REFINE.} \tag{3.1}$$

Let $\{\mathcal{T}_k\}_{k \geq 0}$ and $\{\mathbb{V}_k\}_{k \geq 0} := \{\mathbb{V}(\mathcal{T}_k)\}_{k \geq 0}$ be the sequence of meshes and corresponding finite element spaces generated by iterating (3.1). We discuss each of these four procedures in the following subsections. To do so, we drop the subscript k whenever no confusion is possible.

3.1. The Module SOLVE. Given a mesh $\mathcal{T}$, SOLVE finds the solution $U \in \mathbb{V}$ of discrete linear problem (2.11). Due to the symmetry of the bilinear form $a_{\mathcal{T}}(\cdot, \cdot)$, conjugate gradient methods may be used to solve (2.11). Multilevel preconditioners are studied in [30, 19] but, for simplicity, we assume that U is computed *exactly*.

3.2. The Module ESTIMATE. The crucial ingredient of any ADFEM is the control of the error by the estimator, namely the so-called reliability. The following Lemma is proved in [36] for formulation (1.3) and in [32] for an hp version of it, both for simplices. Our proof applies to formulation (2.11).

LEMMA 3.1 (First Upper Bound). *Let $\mathcal{T}$ be any subdivision of Ω and let Condition 2 be valid. Let $u \in H_0^1(\Omega)$ and $U \in \mathbb{V}(\mathcal{T})$ be the corresponding solutions of (2.2) and (2.11), respectively, and $\gamma_1 \geq 1$ be as in Lemma 2.2. Then, for $\gamma \geq \gamma_1$ there holds*

$$a_{\mathcal{T}}(u - U, u - U) \preceq \eta_{\mathcal{T}}^2(U, \mathcal{T}) + \gamma \left\| h^{-1/2} \, [\![U]\!] \right\|_{\Sigma}^2, \tag{3.2}$$

where the constant hidden in "$\preceq$" depends on the shape regularity of $\mathcal{T}$.

Proof. To cope with the partial consistency (2.12) of $a_{\mathcal{T}}(\cdot, \cdot)$ we invoke the orthogonal decomposition (2.22) and write $U = U^0 + U^\perp$ with $U^0 \in \mathbb{V}^0(\mathcal{T})$, $U^\perp \in \mathbb{V}^\perp(\mathcal{T})$. The error $e := u - U \in \mathbb{E}(\mathcal{T})$ thus splits as $e = v - U^\perp$ with $v = u - U^0 \in H_0^1(\Omega)$ and

$$a_{\mathcal{T}}(e, e) = a_{\mathcal{T}}(e, v - I_{\mathcal{T}} v) - a_{\mathcal{T}}(e, U^\perp) \tag{3.3}$$

because of (2.23). In view of (2.13), the first term reads

$$a_{\mathcal{T}}(e, v - I_{\mathcal{T}} v) = (R(U), v - I_{\mathcal{T}} v)_{\mathcal{T}} - \langle J(U), v - I_{\mathcal{T}} v \rangle_{\Sigma \cap \Omega} + (L_{\mathcal{T}}(U), A \nabla(v - I_{\mathcal{T}} v))_{\mathcal{T}}.$$

Coupling Lemma 2.1 together with assumption (2.1) implies

$$(L_{\mathcal{T}}(U), A \nabla(v - I_{\mathcal{T}} v))_{\mathcal{T}} \preceq \left\| h^{-1/2} \, [\![U]\!] \right\|_{\Sigma} \left\| A^{1/2} \nabla(v - I_{\mathcal{T}} v) \right\|_{\mathcal{T}}.$$

Using the definition (2.17) of $\eta_{\mathcal{T}}(U, \mathcal{T})$, and Cauchy-Schwarz inequality, yields

$$a_{\mathcal{T}}(e, v - I_{\mathcal{T}} v) \preceq \left(\eta_{\mathcal{T}}(U, \mathcal{T}) + \left\| h^{-1/2} \, [\![U]\!] \right\|_{\Sigma} \right)$$
$$\times \left(\left\| A^{1/2} \nabla(v - I_{\mathcal{T}} v) \right\|_{\mathcal{T}} + \left\| h^{-1}(v - I_{\mathcal{T}} v) \right\|_{\mathcal{T}} + \left\| h^{-1/2}(v - I_{\mathcal{T}} v) \right\|_{\Sigma} \right).$$

Invoking a scaled trace inequality, assumption (2.1) on A and the approximation property (2.27) of the interpolation operator $I_{\mathcal{T}}$, implies

$$\left\| A^{1/2} \nabla(v - I_{\mathcal{T}} v) \right\|_{\mathcal{T}} + \left\| h^{-1}(v - I_{\mathcal{T}} v) \right\|_{\mathcal{T}} + \left\| h^{-1/2}(v - I_{\mathcal{T}} v) \right\|_{\Sigma} \preceq \left\| A^{1/2} \nabla v \right\|_{\mathcal{T}}.$$

Since $v = e + U^\perp$, coercivity (2.16) gives

$$\left\|A^{1/2}\nabla v\right\|_{\mathcal{T}} \preceq a_{\mathcal{T}}(e,e)^{1/2} + \left\|A^{1/2}\nabla U^\perp\right\|_{\mathcal{T}},$$

whence

$$a_{\mathcal{T}}(e, v - I_{\mathcal{T}}v) \preceq \left(\eta_{\mathcal{T}}(U,\mathcal{T}) + \left\|h^{-1/2}\,[\![U]\!]\right\|_{\Sigma}\right)\left(a_{\mathcal{T}}(e,e)^{1/2} + \left\|A^{1/2}\nabla U^\perp\right\|_{\mathcal{T}}\right).$$

Finally, if C denotes a generic constant independent of γ, then Young inequality and the definition (2.7) of the energy norm lead to

$$a_{\mathcal{T}}(e, v - I_{\mathcal{T}}v) \le \frac{1}{4}a(e,e) + \frac{1}{4}\left|\!\left|\!\left|\nabla U^\perp\right|\!\right|\!\right|_{\mathcal{T}}^2 + C\left(\eta_{\mathcal{T}}^2(U,\mathcal{T}) + \left\|h^{-1/2}\,[\![U]\!]\right\|_{\Sigma}^2\right).$$

Concerning the last term of (3.3), Cauchy-Schwarz inequality, together with continuity (2.15), coercivity (2.16) and Young inequality, implies

$$a_{\mathcal{T}}(e, U^\perp) \preceq a_{\mathcal{T}}(e,e)^{1/2}\left|\!\left|\!\left|U^\perp\right|\!\right|\!\right|_{\mathcal{T}} \le \frac{1}{4}a(e,e) + C\left|\!\left|\!\left|U^\perp\right|\!\right|\!\right|_{\mathcal{T}}^2.$$

Inserting this and the previous estimate into (3.3) leads to

$$a_{\mathcal{T}}(e,e) \preceq \eta_{\mathcal{T}}^2(U,\mathcal{T}) + \left\|h^{-1/2}\,[\![U]\!]\right\|_{\Sigma}^2 + \left|\!\left|\!\left|U^\perp\right|\!\right|\!\right|_{\mathcal{T}}^2,$$

Hence for $\gamma \ge 1$, the estimate (3.2) follows from (2.29). □

REMARK 3.2 (Influence on γ). *The upper bound obtained in [37] reads*

$$\|\nabla(u - U)\|_{\mathcal{T}}^2 \preceq \eta_{\mathcal{T}}^2(U,\mathcal{T}) + \gamma^2\left\|h^{-1/2}\,[\![U]\!]\right\|_{\Sigma}^2.$$

Compared to (3.2), we have been able to get a lower power of γ using the decomposition $U = U^0 + U^\perp$ and the interpolation operator of Condition 2.

The residual estimator $\eta_{\mathcal{T}}(U,\mathcal{T})$ is exactly the same as for cG [42, 40, 21]. The upper bound provided in Lemma 3.1 may suggest to add the jump term $\gamma^{1/2}\left\|h^{-1/2}\,[\![U]\!]\right\|_{\Sigma}$ to $\eta_{\mathcal{T}}(U,\mathcal{T})$. This would destroy the monotonicity property (2.21) of the estimator with respect to the meshsize, which is instrumental for Lemmas 2.4 and 2.5 and thus for the proof of the main results. However, the jump term $\gamma\left\|h^{-1/2}\,[\![U]\!]\right\|_{\Sigma}^2$ on the right hand side of (3.2) can be bounded by $\gamma^{-1}\eta_{\mathcal{T}}(U,\mathcal{T})^2$ for γ sufficiently large as stated in the following Lemma. We present an alternative proof to the original one in [37, eq. (3.20)]; see also [2] for a similar result in $2D$ using linear finite elements.

LEMMA 3.3 (Jump Control). *Assume that Condition 2 holds on any subdivision $\mathcal{T}$ of Ω. Let $U \in \mathbb{V}(\mathcal{T})$ be the solution of (2.11) and $\gamma_1 \ge 1$ be as in Lemma 2.2. There exists a constant $\gamma_2 \ge \gamma_1$ such that for $\gamma \ge \gamma_2$, there holds*

$$\gamma\left\|h^{-1/2}\,[\![U]\!]\right\|_{\Sigma} \preceq \eta_{\mathcal{T}}(U,\mathcal{T}).$$

The constant hidden in "$\preceq$" depends on the shape regularity of $\mathcal{T}$.

Proof. The coercivity property (2.16) of Lemma 2.2 implies the existence of constants C, γ_1 independent of γ and the meshsize h such that for $\gamma \ge \gamma_1$

$$\gamma\left\|h^{-1/2}\,[\![U]\!]\right\|_{\Sigma}^2 \le Ca_{\mathcal{T}}(U - V^0, U - V^0) \qquad \forall V^0 \in \mathbb{V}^0(\mathcal{T}) \qquad (3.4)$$

because $[\![V^0]\!] = 0$. On the other hand, since U solves (2.11) we have

$$a_{\mathcal{T}}(U - V^0, U - V^0) = F_{\mathcal{T}}(U - V^0) - a_{\mathcal{T}}(V^0, U - V^0). \qquad (3.5)$$

Let us concentrate for the moment on the second term of the right-hand side of (3.5). The penalty term containing γ on $a_{\mathcal{T}}(V^0, U - V^0)$ vanishes because $[\![V^0]\!] = 0$. Therefore, using that $L(V^0) \equiv 0$ on $\overline{\Omega}$ and rewriting $V^0 = U + (V^0 - U)$, we obtain

$$a_{\mathcal{T}}(V^0, U - V^0) = \left(A\nabla U, \nabla(U - V^0)\right)_{\mathcal{T}} - \left\|A^{1/2}\nabla(U - V^0)\right\|_{\mathcal{T}}^2 - \left(L_{\mathcal{T}}(U), A\nabla V^0\right)_{\mathcal{T}}.$$

Integration by parts, coupled with definition (2.8) of the lifting operator, leads to

$$\left(A\nabla U, \nabla(U - V^0)\right)_{\mathcal{T}} = \left(-\mathrm{div}(A\nabla U), U - V^0\right)_{\mathcal{T}}$$
$$+ \left\langle [\![A\nabla U]\!], \{\!\{U - V^0\}\!\}\right\rangle_{\Sigma \cap \Omega} + \left(L_{\mathcal{T}}(U - V^0), A\nabla U\right)_{\mathcal{T}},$$

whence, employing again that $L_{\mathcal{T}}(V^0) = 0$, we get

$$a_{\mathcal{T}}(V^0, U - V^0) = \left(-\mathrm{div}(A\nabla U), U - V^0\right)_{\mathcal{T}} + \left\langle [\![A\nabla U]\!], \{\!\{U - V^0\}\!\}\right\rangle_{\Sigma \cap \Omega}$$
$$- \left\|A^{1/2}\nabla(U - V^0)\right\|_{\mathcal{T}}^2 + \left(L_{\mathcal{T}}(U), A\nabla(U - V^0)\right)_{\mathcal{T}}.$$

Inserting this into (3.5), and using (2.17) and Cauchy-Schwarz inequality, yields

$$a_{\mathcal{T}}(U - V^0, U - V^0) \le \eta_{\mathcal{T}}(U, \mathcal{T})\left(\left\|h^{-1}(U - V^0)\right\|_{\mathcal{T}} + \left\|h^{-1/2}\{\!\{U - V^0\}\!\}\right\|_{\Sigma \cap \Omega}\right)$$
$$+ \left\|A^{1/2}\nabla(U - V^0)\right\|_{\mathcal{T}}^2 + \|L_{\mathcal{T}}(U)\|_{\mathcal{T}}\left\|A\nabla(U - V^0)\right\|_{\mathcal{T}}.$$

Therefore, choosing $V^0 = I_{\mathcal{T}}U$ with the interpolation operator $I_{\mathcal{T}} : \mathbb{E}(\mathcal{T}) \to \mathbb{V}^0(\mathcal{T})$ of Condition **2**, and invoking the trace inequality and Lemma 2.1, we arrive at

$$a_{\mathcal{T}}(U - V^0, U - V^0) \preceq \eta_{\mathcal{T}}(U, \mathcal{T})\left\|h^{-1/2}[\![U]\!]\right\|_{\Sigma} + \left\|h^{-1/2}[\![U]\!]\right\|_{\Sigma}^2,$$

where we have also used assumption (2.1) on A. Replacing this into (3.4) implies

$$(\gamma - C)\left\|h^{-1/2}[\![U]\!]\right\|_{\Sigma}^2 \preceq \eta_{\mathcal{T}}^2(U, \mathcal{T}).$$

This concludes the proof upon choosing $\gamma \ge \gamma_2 = 2C$. $\square$

As a direct consequence of the previous Lemma, we obtain a simpler upper bound and a quasi-localized upper bound.

COROLLARY 3.4 (Second Upper Bound). *Let Condition **2** (interpolation operator) hold and $\mathcal{T}$ be any subdivision of Ω. Let $u \in H_0^1(\Omega)$ and $U \in \mathbb{V}(\mathcal{T})$ be the corresponding solutions of (2.2) and (2.11), respectively. Let $\gamma_2 \ge 1$ be as in Lemma 3.3. Then, for $\gamma \ge \gamma_2$ there exists $C_U > 0$ depending on the shape regularity of $\mathcal{T}$ such that*

$$a_{\mathcal{T}}(u - U, u - U) \le C_U \eta_{\mathcal{T}}^2(U, \mathcal{T}). \qquad (3.6)$$

Proof. This is a direct consequence of Lemmas 3.1 and 3.3. $\square$

It is also possible to localize the error made by ADFEM, but in a sense weaker than [48, 21]; this is the subject of next Lemma. We recall that $\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*}$ stands for

the set of refined elements of $\mathcal{T}$ needed to obtain $\mathcal{T}_*$ (see Section 2.1.2), and let the enlarged refined set $\omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})$ defined by (2.26).

LEMMA 3.5 (Quasi-Localized Upper Bound). *Let Condition 2 hold and γ_2 be as in Lemma 3.3. Let $\mathcal{T}$ be any subdivision of Ω and $\mathcal{T}_*$ be a refinement of $\mathcal{T} \in \mathbb{T}$, and let $U \in \mathbb{V}(\mathcal{T})$ and $U_* \in \mathbb{V}(\mathcal{T}_*)$ be the corresponding solutions of (2.11). Then, there exists a constant C_{LU} only depending on the shape regularity of $\mathcal{T}$ so that*

$$\left|\left|\left|U_*^0 - U\right|\right|\right|_{\mathcal{T}}^2 \leq C_{LU} \left(\eta_{\mathcal{T}}^2(U, \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})) + \gamma^{-1} \eta_{\mathcal{T}}^2(U, \mathcal{T}) \right), \qquad \forall \gamma \geq \gamma_2,$$

where $U_* = U_*^0 + U_*^\perp$ is the orthogonal decomposition (2.22).

Proof. Write $U = U^0 + U^\perp$ with $U^0 \in \mathbb{V}^0(\mathcal{T})$, $U^\perp \in \mathbb{V}^\perp(\mathcal{T})$ according to (2.22). From Remark 2.7 we recall that $U_*^0 \in \mathbb{V}^0(\mathcal{T}_*)$ satisfies

$$a_{\mathcal{T}}(U_*^0, V^0) = (f, V^0), \qquad \forall V^0 \in \mathbb{V}^0(\mathcal{T}),$$

whence

$$a_{\mathcal{T}}(U_*^0 - U, V^0) = 0, \qquad \forall V^0 \in \mathbb{V}^0(\mathcal{T}).$$

Splitting $U_*^0 - U = E_*^0 - U^\perp + V^0$, with $E_*^0 = U_*^0 - U^0 - V^0$ and $V^0 \in \mathbb{V}^0(\mathcal{T})$, and using (2.16) for $\gamma \geq \gamma_1$ yields

$$\left|\left|\left|U_*^0 - U\right|\right|\right|_{\mathcal{T}}^2 \preceq a_{\mathcal{T}}(U_*^0 - U, E_*^0) - a_{\mathcal{T}}(U_*^0 - U, U^\perp).$$

The last term cannot be localized but its dependence on γ makes it still acceptable. In fact, combining (2.15) with Lemmas 2.9 and 3.3 gives

$$|a_{\mathcal{T}}(U_*^0 - U, U^\perp)| \preceq \gamma^{1/2} \left\|h^{-1/2} [\![U]\!]\right\|_{\Sigma(\mathcal{T})} \left|\left|\left|U_*^0 - U\right|\right|\right|_{\mathcal{T}}$$

$$\preceq \gamma^{-1/2} \eta_{\mathcal{T}}(U, \mathcal{T}) \left|\left|\left|U_*^0 - U\right|\right|\right|_{\mathcal{T}}. \tag{3.7}$$

In order to localize $a_{\mathcal{T}}(U_*^0 - U, E_*^0)$, we denote by Ω_i each connected component of $\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*}$ and choose $V^0 = I_{\mathcal{T}}(U_*^0 - U^0)$ where $I_{\mathcal{T}}$ is given in Condition **2**. Since $I_{\mathcal{T}}$ is locally a projection, the error E_*^0 vanishes outside the set $\omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})$ The error-residual relation (2.13), with u replaced by U_*^0, implies

$$a_{\mathcal{T}}(U_*^0 - U, E_*^0) = (R(U), E_*^0)_{\mathcal{T}} - \langle J(U), E_*^0 \rangle_{\Sigma} + (L_{\mathcal{T}}(U), A\nabla E_*^0)_{\mathcal{T}}.$$

For the last term, we invoke Lemmas 2.1 and 3.3 to obtain (recall that $\gamma > 1$)

$$(L_{\mathcal{T}}(U), A\nabla E_*^0)_{\mathcal{T}} \preceq \gamma^{-1/2} \eta_{\mathcal{T}}(U, \mathcal{T}) \left|\left|\left|U_*^0 - U\right|\right|\right|_{\mathcal{T}}.$$

So that, using (2.27), splitting $U_*^0 - U^0 = U_*^0 - U + U^\perp$, and recalling (2.29), we arrive at

$$a_{\mathcal{T}}(U_*^0 - U, E_*^0) \preceq \eta_{\mathcal{T}}(U, \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})) \left|\left|\left|U_*^0 - U\right|\right|\right|_{\mathcal{T}} + \gamma^{-1/2} \eta_{\mathcal{T}}(U, \mathcal{T}) \left|\left|\left|U_*^0 - U\right|\right|\right|_{\mathcal{T}}.$$

Finally, combining this with (3.7) gives the assertion. $\square$

The following lemma deals wth *efficiency* of the estimator. We only sketch its proof because one can find it in [36, Theorem 3.2] for simplices, except for the treatment of oscillation. For the latter we refer for instance to [42] and [3].

LEMMA 3.6 (Global Lower Bound). *Let $\mathcal{T}$ be any subdivision of Ω. Let $u \in H_0^1(\Omega)$ and $U \in \mathbb{V}(\mathcal{T})$ be the solutions of problems (1.1) and (2.11) respectively. Then, there exists a constant C_{GL} depending only on the shape regularity of $\mathcal{T}$ such that*

$$C_{GL}\eta_{\mathcal{T}}(U,\mathcal{T})^2 \le |||u-U|||_{\mathcal{T}}^2 + \operatorname{osc}_{\mathcal{T}}^2(U,\mathcal{T}). \tag{3.8}$$

Proof. We denote by ψ_T the bubble function with support in $T \in \mathcal{T}$. Since $v = \psi_T \Pi_{2n-2}^2 R(U) \in H_0^1(T)$, we easily deduce from (2.13) that

$$(R(U),v)_T = (A\nabla(u-U),\nabla v)_T,$$

whence

$$\|hR(U)\|_T^2 \preceq |||u-U|||_T^2 + \left\|hP_{2n-2}^2 R(U)\right\|_T^2.$$

Next, we let $S := T^+ \cap T^- \in \Sigma$ and assume that S is a full side of both T^+ and T^-. Let $\omega(S) = T^+ \cup T^-$ and ψ_S be the side bubble function with support in $\omega(S)$. Let $\Pi_{2n-1}^2 J$ be suitably extended to $\omega(S)$ and take $v = \psi_S \Pi_{2n-1}^2 J \in H_0^1(\omega(S))$ in (2.13). We readily find out the relation

$$\langle J(U),v\rangle_S = (R(U),v)_{\omega(S);\mathcal{T}} - (A\nabla(u-U),\nabla v)_{\omega(S);\mathcal{T}},$$

which implies

$$\left\|h^{1/2}J(U)\right\|_S^2 \preceq |||u-U|||_{\omega(S);\mathcal{T}}^2 + \|hR(U)\|_{\omega(S);\mathcal{T}}^2 + \left\|h^{1/2}P_{2n-1}^2 J(U)\right\|_S^2.$$

We now suppose that S is not a full side of T^+. We refine T^+ until S becomes a full side of an element $\widetilde{T}^+ \subset T^+$, which decreases the meshsize, and argue as before to reach the same inequality because $\widetilde{T}^+ \cup T^- \subset \omega(S)$.

Finally, invoking the finite overlapping property of sets $\omega(S)$, and adding over all elements $T \in \mathcal{T}$ the estimates for the interior and jump residuals yields (3.8). $\square$

3.3. The Module MARK. The element indicators $\{\eta_{\mathcal{T}}(U,T)\}_{T\in\mathcal{T}}$ are now used to choose which element to refine. Our marking strategy, originally proposed by Dörfler [27], consists of selecting a subset $\mathcal{M}$ of $\mathcal{T}$ satisfying

$$\eta_{\mathcal{T}}(U,\mathcal{M}) \ge \theta\eta_{\mathcal{T}}(U,\mathcal{T}), \tag{3.9}$$

where $0 < \theta < 1$ is a given parameter. We insist on marking exclusively by $\eta_{\mathcal{T}}(U,\mathcal{T})$, as in [21], because separate marking by $\eta_{\mathcal{T}}(U,\mathcal{T})$ and $\operatorname{osc}_{\mathcal{T}}(U,\mathcal{T})$ may be suboptimal for cG [11, 15, 23, 37, 40, 42]. We refer to [21, Section 6] and [43, Section 9.4] for a discussion of separate marking for cG. We observe that Hoppe et al [31] may need separate marking for dG to enforce a contraction of oscillation, which is avoided here.

3.4. The Module REFINE. The procedure $\mathcal{T}_* = \text{REFINE}(\mathcal{T},\mathcal{M})$ refines at least all the marked elements $\mathcal{M} \subset \mathcal{T}$, thereby creating a refinement $\mathcal{T}_* \ge \mathcal{T}$, and satisfies the following two conditions. The first one relates two successive subdivisions while the second condition restricts the global behavior of REFINE.

Condition 3 (Properties of REFINE). *The procedure $\mathcal{T}_* = REFINE(\mathcal{T},\mathcal{M})$ generates $\mathcal{T}_*$ from $\mathcal{T}$ by successive application of atomic refinements so that:*
  a. the atomic refinements satisfy Condition 1;

b. *all elements of $\mathcal{M}$ are refined at least ρ times, with $\rho \geq 1$ a given integer;*
c. *no element of $\mathcal{T}$ can undergo more than N atomic refinements, with $N \geq 1$ a universal integer, to give rise to elements of $\mathcal{T}_*$;*
d. *Condition 2 (interpolation operator) is valid on $\mathbb{V}(\mathcal{T}_*)$.*

If $T_* \in \mathcal{T}_*$ and $T \in \mathcal{T}$ is its ancestor, then Condition **3**(c) reads $\ell(T_*) - \ell(T) \leq N$.

Condition 4 (Complexity of REFINE). *Let $\mathcal{T}_0$ be an initial conforming subdivision. The REFINE procedure produces sequence of subdivisions $\{\mathcal{T}_k\}_{k \geq 0}$ such that*

$$\#\mathcal{T}_k - \mathcal{T}_0 \leq \Lambda_0 \sum_{j=0}^{k-1} \#\mathcal{M}_j,$$

where $\Lambda_0 > 0$ is a universal constant depending on $\mathcal{T}_0$, the dimension d, and the number of refinements $\rho \geq 1$.

Condition **4** is known to hold for refinement by bisections on conforming meshes: we refer to Binev, Dahmen and Devore [15] for $d = 2$ and Stevenson [49] for $d > 2$, as well as the survey [43, Section 4] for $d \geq 2$. The extension to the type of refinements described in Section 2.1.2 is provided in Section 6.

4. Contraction Property of the ADFEM. The aim of this section is to prove the Main Result 1. But this requires a few preliminary results.

4.1. Comparison Results. In order to relate two successive energy norms $a_{\mathcal{T}_k}(u - U_k, u - U_k)$, we start with a Lemma monitoring the decrease of side diameters.

COROLLARY 4.1 (Relation between Consecutive Sides). *Let $\mathcal{T}_* \geq \mathcal{T}$ be two successive subdivisions created by REFINE. If Condition 3 holds, then there exists a constant $C_* \geq 1$ only depending on shape regularity of $\mathcal{T}_0$, β_m in (2.3) and N in Condition 3 such that for all $S \in \Sigma$*

$$h_S \leq C_* h_{S_*}, \qquad \forall S_* \in \Sigma_*, \ S_* \subset S.$$

Proof. Let $S_* \in \Sigma_*$ be a side of $T_* \in \mathcal{T}_*$. Condition **3**(c) implies that $\ell(T_*) \leq \ell(T) + N$, which coupled with (2.3) yields $h_T \preceq h_{T_*}$. In view of Condition **3**(a) both $\mathcal{T}$ and $\mathcal{T}_*$ are shape regular with constant only depending on $\mathcal{T}_0$, whence

$$h_S \preceq h_T \preceq h_{T_*} \preceq h_{S_*}.$$

This ends the proof. □

LEMMA 4.2 (Mesh Perturbation). *Let $\mathcal{T}_* \geq \mathcal{T}$ be two successive refinements created by REFINE. If Condition 3 is valid, then there exists a constant C independent of γ and the level of refinement such that for all $0 < \epsilon < 1$ there holds*

$$a_{\mathcal{T}_*}(v,v) \leq (1+\epsilon)a_{\mathcal{T}}(v,v) + \frac{C}{\epsilon}\gamma \left\| h^{-1/2} \llbracket v \rrbracket \right\|_\Sigma^2 \quad \forall v \in \mathbb{E}(\mathcal{T}). \tag{4.1}$$

Proof. Since $(A\nabla v, \nabla v)_{\mathcal{T}} = (A\nabla v, \nabla v)_{\mathcal{T}_*}$ for $v \in \mathbb{E}(\mathcal{T})$, we readily have

$$a_{\mathcal{T}_*}(v,v) = a_{\mathcal{T}}(v,v) + 2\left(L_{\mathcal{T}}v, A\nabla v\right)_{\mathcal{T}} - 2\left(L_{\mathcal{T}_*}v, A\nabla v\right)_{\mathcal{T}_*}$$
$$- \gamma \left\| h^{-1/2} \llbracket v \rrbracket \right\|_\Sigma^2 + \gamma \left\| h^{-1/2} \llbracket v \rrbracket \right\|_\Sigma^2.$$

Using Lemma 2.1, coercivity (2.16) and Young inequality, we obtain

$$(L_{\mathcal{T}}v, A\nabla v)_{\mathcal{T}} \leq C \left\| h^{-1/2} \llbracket v \rrbracket \right\|_{\Sigma} a_{\mathcal{T}}(v,v)^{1/2} \leq \frac{\epsilon}{4} a_{\mathcal{T}}(v,v) + \frac{C}{\epsilon} \left\| h^{-1/2} \llbracket v \rrbracket \right\|_{\Sigma}^2,$$

with $C > 0$ independent of γ. A similar argument, but also employing Corollary 4.1 to relate the size of sides in Σ_* to those of Σ, leads to

$$(L_{\mathcal{T}_*}v, A\nabla v)_{\mathcal{T}_*} \leq \frac{\epsilon}{4} a_{\mathcal{T}}(v,v) + \frac{C}{\epsilon} \left\| h^{-1/2} \llbracket v \rrbracket \right\|_{\Sigma}^2.$$

We finally apply again Corollary 4.1, this time to write $\left\| h^{-1/2} \llbracket v \rrbracket \right\|_{\Sigma_*} \preceq \left\| h^{-1/2} \llbracket v \rrbracket \right\|_{\Sigma}$. Collecting these estimates gives (4.1). □

COROLLARY 4.3 (Comparison of Solutions). *Let $\mathcal{T}_* \geq \mathcal{T}$ be two successive refinements created by REFINE, and let $U_* \in \mathbb{V}(\mathcal{T}_*)$ and $U \in \mathbb{V}(\mathcal{T})$ be the corresponding discontinuous Galerkin solutions of (2.11). Let $u \in H_0^1(\Omega)$ be the solution of (2.2). If Condition 3 is valid and γ_2 is as in Lemma 3.3, then there exists a constant C_{comp}, independent of γ and the meshsize, such that for all $\gamma \geq \gamma_2$ and all $0 < \epsilon < 1$*

$$a_{\mathcal{T}_*}(u - U_*, u - U_*) \leq (1 + \epsilon) a_{\mathcal{T}}(u - U, u - U)$$
$$- \frac{C_{\mathrm{coer}}}{2} \|\nabla(U_* - U)\|_{\mathcal{T}_*}^2 + \frac{C_{\mathrm{comp}}}{\epsilon\gamma} \left(\eta_{\mathcal{T}}^2(U, \mathcal{T}) + \eta_{\mathcal{T}_*}^2(U_*, \mathcal{T}_*) \right).$$

Proof. We recall that Galerkin orthogonality does not hold in the whole energy space $\mathbb{E}(\mathcal{T})$. We thus invoke the partial orthogonality (2.23) to write

$$a_{\mathcal{T}_*}(u - U_*, u - U_*)$$
$$= a_{\mathcal{T}_*}(u - U_* - U^0 + U_*^0, u - U_* - U^0 + U_*^0) - a_{\mathcal{T}_*}(U_*^0 - U^0, U_*^0 - U^0).$$

Rewriting $U_* + U^0 - U_*^0 = U - U^\perp + U_*^\perp$, and employing the coercivity (2.16) and the continuity (2.15), it follows that

$$a_{\mathcal{T}_*}(u - U_*, u - U_*) \leq a_{\mathcal{T}_*}(u - U, u - U) - C_{\mathrm{coer}} \left\| \nabla(U_*^0 - U^0) \right\|_{\mathcal{T}_*}^2$$
$$+ C_{\mathrm{cont}} \left\| \left| U_*^\perp - U^\perp \right| \right\|_{\mathcal{T}_*}^2 \qquad (4.2)$$
$$+ 2C_{\mathrm{cont}}^{1/2} a_{\mathcal{T}_*}(u - U, u - U)^{1/2} \left\| \left| U^\perp - U_*^\perp \right| \right\|_{\mathcal{T}_*}.$$

Since $U^0 = U - U^\perp$, $U_*^0 = U_* - U_*^\perp$, the reversed triangle inequality yields

$$\left\| \nabla(U_*^0 - U^0) \right\|_{\mathcal{T}_*}^2 \geq \frac{1}{2} \left\| \nabla(U_* - U) \right\|_{\mathcal{T}_*}^2 - \left\| \nabla(U_*^\perp - U^\perp) \right\|_{\mathcal{T}_*}^2.$$

Plugging the above relation into (4.2), and employing Young's inequality, leads to

$$a_{\mathcal{T}_*}(u - U_*, u - U_*) \leq (1 + \epsilon) a_{\mathcal{T}_*}(u - U, u - U)$$
$$- \frac{C_{\mathrm{coer}}}{2} \|\nabla(U_* - U)\|_{\mathcal{T}_*}^2 + C\left(1 + \frac{1}{\epsilon}\right) \left\| \left| U_*^\perp - U^\perp \right| \right\|_{\mathcal{T}_*}^2,$$

where $0 < \epsilon < 1$. The triangle inequality, in conjunction with (2.29), gives

$$\left\| \left| U_*^\perp - U^\perp \right| \right\|_{\mathcal{T}_*}^2 \preceq \gamma \left\| h^{-1/2} \llbracket U_*^\perp \rrbracket \right\|_{\Sigma_*}^2 + \gamma \left\| h^{-1/2} \llbracket U^\perp \rrbracket \right\|_{\Sigma_*}^2.$$

Since the sides of Σ and Σ_* are comparable, according to Corollary 4.1, we can replace the last term by $\gamma \left\| h^{-1/2} \left[\!\left[U^\perp \right]\!\right] \right\|_\Sigma^2$. Invoking Lemma 3.3 with $\gamma \geq \gamma_2$ yields

$$\left\|\!\left|\!\left| U_*^\perp - U^\perp \right|\!\right|\!\right\|_{\mathcal{T}_*}^2 \preceq \gamma^{-1} \eta_{\mathcal{T}_*}^2(U_*, \mathcal{T}_*) + \gamma^{-1} \eta_{\mathcal{T}}^2(U, \mathcal{T}).$$

Finally, applying Lemma 4.2 with $v = u - U \in \mathbb{E}(\mathcal{T})$ concludes the proof. $\square$

Corollary 4.3 is the dG version of the orthogonality relation between two cG solutions (Pythagoras theorem). This is crucial to prove a contraction property [21].

4.2. Main Result 1: Contraction Property. We are now in a position to prove the Main Result 1 announced in the Introduction.

THEOREM 4.4 (Contraction Property). *Let REFINE satisfy Condition 3. Let $\mathcal{T}_0$ be the initial conforming subdivision and $\mathcal{T}_k \leq \mathcal{T}_{k+1}$ be two consecutive meshes obtained from $\mathcal{T}_0$ after $k \geq 0$ iterations of the adaptive algorithm (3.1). Let $U_k \in \mathbb{V}(\mathcal{T}_k)$, $U_{k+1} \in \mathbb{V}(\mathcal{T}_{k+1})$ be the corresponding finite element solutions of (2.11), and $u \in H_0^1(\Omega)$ be the solution of (2.2), with $f \in L^2(\Omega)$ and A satisfying (2.1).*

Then there exist three constants $\gamma_m \geq 0$, $\delta > 0$ and $\alpha < 1$ depending solely on the shape regularity of $\mathcal{T}_0$, the marking parameter $0 < \theta < 1$, the matrix A and the polynomial order n of the finite elements method, such that for $\gamma \geq \gamma_m$, there holds

$$a_{\mathcal{T}_{k+1}}(u - U_{k+1}, u - U_{k+1}) + \delta \eta_{\mathcal{T}_{k+1}}^2(U_{k+1}, \mathcal{T}_{k+1})$$
$$\leq \alpha \Big(a_{\mathcal{T}_k}(u - U_k, u - U_k) + \delta \eta_{\mathcal{T}_k}^2(U_k, \mathcal{T}_k) \Big).$$

Proof. To simplify the notation we write $E_k := u - U_k \in \mathbb{E}(\mathcal{T}_k)$ $a_k(\cdot, \cdot) := a_{\mathcal{T}_k}(\cdot, \cdot)$, and $\eta_k := \eta_{\mathcal{T}_k}(U_k, \mathcal{T}_k)$ for the k-th error, bilinear form and estimator.

Using Corollary 4.3 and Lemma 2.4 yields, for any $\zeta > 0$,

$$a_{k+1}(E_{k+1}, E_{k+1}) + \left(\delta - \frac{C_{\text{comp}}}{\epsilon \gamma} \right) \eta_{k+1}^2 \leq (1 + \epsilon) a_k(E_k, E_k)$$
$$+ \left(C_{\text{er}}(1 + \zeta^{-1})\delta - \frac{C_{\text{coer}}}{2} \right) \|\nabla(U_k - U_{k+1})\|_{\mathcal{T}_{k+1}}^2 \qquad (4.3)$$
$$+ (1 + \zeta)\delta \left(\eta_k^2 - \lambda \eta_{\mathcal{T}_k}^2(U_k, \mathcal{R}_k) \right) + \frac{C_{\text{comp}}}{\epsilon \gamma} \eta_k^2.$$

We choose the parameter δ to eliminate the term $\|\nabla(U_{k+1} - U_k)\|_{\mathcal{T}_{k+1}}$ in the right hand side of (4.3), namely,

$$C_{\text{er}}(1 + \zeta^{-1})\delta = \frac{C_{\text{coer}}}{2} \qquad \text{i.e.} \qquad (1 + \zeta)\delta = \frac{\zeta C_{\text{coer}}}{2C_{\text{er}}}. \qquad (4.4)$$

Moreover, Dörfler marking (3.9) implies that $-\eta_{\mathcal{T}_k}^2(U_k, \mathcal{R}_k) \leq -\theta^2 \eta_k^2$, and combined with (4.4) leads to

$$a_{k+1}(E_{k+1}, E_{k+1}) + \left(\delta - \frac{C_{\text{comp}}}{\epsilon \gamma} \right) \eta_{k+1}^2 \leq (1 + \epsilon) a_k(E_k, E_k) - \frac{(1 + \zeta)\delta \theta^2 \lambda}{2} \eta_k^2$$
$$+ \left((1 + \zeta)\delta - \frac{(1 + \zeta)\delta \theta^2 \lambda}{2} + \frac{C_{\text{comp}}}{\epsilon \gamma} \right) \eta_k^2.$$

The upper bound (3.6) together with (4.4), yields

$$a_{k+1}(E_{k+1}, E_{k+1}) + \delta \left(1 - \frac{C_{\text{comp}}}{\delta \epsilon \gamma} \right) \eta_{k+1}^2 \leq \Lambda_1 a_k(E_k, E_k) + \delta \left(\Lambda_2 + \frac{C_{\text{comp}}}{\delta \epsilon \gamma} \right) \eta_k^2,$$

with

$$\Lambda_1 := \left(1 + \epsilon - \frac{C_{\mathrm{coer}}}{4C_{\mathrm{U}}C_{\mathrm{er}}} \zeta \theta^2 \lambda \right) \quad \text{and} \quad \Lambda_2 := (1 + \zeta) \left(1 - \frac{\theta^2 \lambda}{2} \right).$$

We now choose $\epsilon = \frac{C_{\mathrm{coer}}}{8C_{\mathrm{U}}C_{\mathrm{er}}} \zeta \theta^2 \lambda$ and $\zeta > 0$ so that $\Lambda_2 = 1 - \frac{\theta^2 \lambda}{4}$, whence $\max(\Lambda_1, \Lambda_2) < 1$. Finally, on selecting $\gamma \geq \gamma_2$ sufficiently large so that

$$1 - \frac{C_{\mathrm{comp}}}{\delta \epsilon \gamma} \geq \Lambda_2 + \frac{C_{\mathrm{comp}}}{\delta \epsilon \gamma},$$

the assertion follows immediately. $\square$

REMARK 4.5 (Geometric Convergence). *Combining the coercivity* (2.16) *with the contraction property given by Theorem 4.4 yields* $|||u - U_k|||_{\mathcal{T}_k} \leq \Xi_0 \alpha^{\frac{k}{2}}$, *with constant* $\Xi_0^2 := |||u - U_0|||_{\mathcal{T}_0}^2 + \delta \eta_{\mathcal{T}_0}^2(U_0, \mathcal{T}_0)$. *Consequently, we get the geometric decay of jumps*

$$\left\| h^{-1/2} \llbracket U_k \rrbracket \right\|_{\Sigma_k} \leq \Xi_0 \alpha^{\frac{k}{2}}.$$

5. Quasi-Optimal Cardinality of the ADFEM. The aim of this section is to prove the Main Result 2, namely the asymptotic estimate for the total error

$$|||u - U_k|||_{\mathcal{T}_k} + \mathrm{osc}_{\mathcal{T}_k}(U_k, \mathcal{T}_k) \preceq \#\mathcal{T}_k^{-s},$$

provided $(u, f, A) \in \mathbb{A}_s$, the definition of approximation class $\mathbb{A}_s$ being the subject of next subsection. We prove in Proposition 5.2 that the class $\mathbb{A}_s$ is the same whether we employ continuous or discontinuous finite elements, provided that Condition **2** holds.

The key idea to relate a finite element solution-mesh pair $(U, \mathcal{T})$ with an optimal one $(U_*, \mathcal{T}_*)$, which achieves a prescribed total error reduction and induces a Dörfler marking, is due to Stevenson for cG and the Laplace operator [48]. We refer to Cascón et al [21], as well as the survey [43, Section 9], for an extension of cG to symmetric operators with variable coefficients such as (1.1). Two new difficulties arise in the discontinuous Galerkin context. Since the meshes $\mathcal{T}, \mathcal{T}_*$ are a priori unrelated, an overlay does not guarantee a fixed level of refinement and, thus, a way to compare the local meshsizes. This first difficulty results in the lack of control of jumps which contain negative powers of meshsize.

To overcome these difficulties, we resort to continuous finite elements and the equivalence of classes in Proposition 5.2. We prove an estimate for the number of marked elements in Lemma 5.5 and combine it with the contraction property of Theorem 4.4 to derive a quasi-optimal convergence rate in Theorem 5.7.

5.1. The Total Error. We recall that the decisions of ADFEM are based on the estimator $\eta_{\mathcal{T}}(U, \mathcal{T})$ via MARK. The global lower bound (3.8) and the control (2.20) of oscillation by $\eta_{\mathcal{T}}(U, \mathcal{T})$ give the equivalence of $\eta_{\mathcal{T}}(U, \mathcal{T})$ and the *total error*:

$$\eta_{\mathcal{T}}(U, \mathcal{T}) \approx |||u - U|||_{\mathcal{T}} + \mathrm{osc}_{\mathcal{T}}(U, \mathcal{T}).$$

Therefore, any decay rate delivered by ADFEM must hinge on the total error and any suitable definition of approximation classes should reflect this fact. Invoking the upper bound (3.6) and the coercivity (2.16) yields the equivalence

$$|||u - U|||_{\mathcal{T}} + \eta_{\mathcal{T}}(U, \mathcal{T}) \approx |||u - U|||_{\mathcal{T}} + \mathrm{osc}_{\mathcal{T}}(U, \mathcal{T}), \tag{5.1}$$

which, in view of the contraction property of Theorem 4.4, implies that ADFEM also reduces the total error. Moreover, the following lemma ensures that the dG solution is quasi-optimal with respect to the total error, and extends a similar result in [21] for cG; see also [43, Section 9].

LEMMA 5.1 (Quasi-Optimality of the Total Error). *Assume that Condition 2 holds. Let $u \in H_0^1(\Omega)$ and $U \in \mathbb{V}(\mathcal{T})$ be the solutions of problems (1.1) and (2.11) respectively, with $\mathcal{T} \in \mathbb{T}$. Then there exist constants C_{OPT} and γ_O only depending on the shape regularity of $\mathcal{T}$ and on the data such that for $\gamma \geq \gamma_O$ the following holds*

$$|||u - U|||_{\mathcal{T}}^2 + \operatorname{osc}_{\mathcal{T}}^2(U, \mathcal{T}) \leq C_{OPT} \inf_{V \in \mathbb{V}(\mathcal{T})} \left(|||u - V|||_{\mathcal{T}}^2 + \operatorname{osc}_{\mathcal{T}}^2(V, \mathcal{T}) \right). \qquad (5.2)$$

Proof. We use the decomposition (2.22), namely $U = U^0 + U^\perp$ and $V = V^0 + V^\perp$, where $U^0, V^0 \in \mathbb{V}^0(\mathcal{T})$ and $U^\perp, V^\perp \in \mathbb{V}^\perp(\mathcal{T})$. Let us start with the energy norm and invoke the coercivity (2.16) and partial consistency (2.23) to obtain

$$|||u - U|||_{\mathcal{T}}^2 \preceq a_{\mathcal{T}}(u - U, u - U) = a_{\mathcal{T}}(u - U, u - U^0) - a_{\mathcal{T}}(u - U, U^\perp)$$
$$\preceq a_{\mathcal{T}}(u - U, u - V) - a_{\mathcal{T}}(u - U, U^\perp) + a_{\mathcal{T}}(u - U, V^\perp).$$

Hence, the continuity (2.15) and coercivity (2.16) yield

$$|||u - U|||_{\mathcal{T}}^2 \preceq |||u - U|||_{\mathcal{T}} \left(|||u - V|||_{\mathcal{T}} + |||V^\perp|||_{\mathcal{T}} + |||U^\perp|||_{\mathcal{T}} \right).$$

The nonconforming parts $U^\perp$ and $V^\perp$ are controlled by their respective jumps, according to (2.29); thus $|||V^\perp|||_{\mathcal{T}} \preceq |||u - V|||_{\mathcal{T}}$. The jumps of U are in turn controlled by the estimator via Lemma 3.3 for $\gamma \geq \gamma_2$, whence $|||U^\perp|||_{\mathcal{T}} \preceq \gamma^{-1/2} \eta_{\mathcal{T}}(U, \mathcal{T})$ and the global lower bound (3.8) leads to

$$|||u - U|||_{\mathcal{T}}^2 \preceq |||u - V|||_{\mathcal{T}}^2 + \gamma^{-1} \left(|||u - U|||_{\mathcal{T}}^2 + \operatorname{osc}_{\mathcal{T}}^2(U, \mathcal{T}) \right); \qquad (5.3)$$

note the remaining factor γ^{-1} on the right-hand side. It remains to estimate the oscillation term in (5.2). Lemma 2.5 (for $\mathcal{T}_* \equiv \mathcal{T}$) and the triangle inequality imply

$$\operatorname{osc}_{\mathcal{T}}^2(U, \mathcal{T}) \preceq \operatorname{osc}_{\mathcal{T}}^2(V, \mathcal{T}) + |||u - V|||_{\mathcal{T}}^2 + |||u - U|||_{\mathcal{T}}^2.$$

Therefore, using (5.3) we arrive at

$$\operatorname{osc}_{\mathcal{T}}^2(U, \mathcal{T}) \preceq \operatorname{osc}_{\mathcal{T}}^2(V, \mathcal{T}) + |||u - V|||_{\mathcal{T}}^2 + \gamma^{-1} \left(|||u - U|||_{\mathcal{T}}^2 + \operatorname{osc}_{\mathcal{T}}^2(U, \mathcal{T}) \right).$$

Adding this to (5.3), and choosing $\gamma \geq \gamma_2$ sufficiently large, completes the proof. □

5.2. Approximation Classes. We now introduce the notion of approximation classes; see [21] and [43, Section 9] for cG. Let N be an integer bigger than $\#\mathcal{T}_0$ and $\mathbb{T}_N \subset \mathbb{T}$ be the set of subdivisions with a number of elements not exceeding N

$$\mathbb{T}_N := \{ \mathcal{T}(F) : \mathcal{F} \subset \mathbb{F}, \ \#\mathcal{T} \leq N \}.$$

The best approximation to the total error in the set $\mathbb{T}_N$ is given by the quantity

$$\sigma(N; u, f, A) := \inf_{\mathcal{T} \in \mathbb{T}_N} \inf_{V \in \mathbb{V}(\mathcal{T})} \left(|||u - V|||_{\mathcal{T}}^2 + \operatorname{osc}_{\mathcal{T}}^2(V, \mathcal{T}) \right)^{1/2},$$

which gives rise to the definition of approximation class $\mathbb{A}_s$ for $s > 0$

$$\mathbb{A}_s := \left\{ (u, f, A) : |u, f, A|_s := \sup_{N \geq \#\mathcal{T}_0} (N^s \sigma(N; u, f, A)) < \infty \right\}.$$

We point out the restriction $s \leq n/d$ since $\mathbb{A}_s$ contains only trivial functions for $s > n/d$; see for instance [26]. As it happens for cG, the class $\mathbb{A}_s$ for dG is not a typical approximation class for functions, as known from approximation theory, because of the nonlinear interaction between the data (f, A) and the solution U through the oscillation term $\mathrm{osc}_\mathcal{T}(U, \mathcal{T})$.

We also define the counterparts for continuous finite element approximations

$$\sigma^0(N; u, f, A) := \inf_{\mathcal{T} \in \mathbb{T}_N} \inf_{V^0 \in \mathbb{V}^0(\mathcal{T})} \left(|||u - V^0|||_\mathcal{T}^2 + \mathrm{osc}_\mathcal{T}^2(V^0, \mathcal{T}) \right)^{1/2},$$

and

$$\mathbb{A}_s^0 := \left\{ (u, f, A) : |u, f, A|_s^0 := \sup_{N \geq \#\mathcal{T}_0} (N^s \sigma^0(N; u, f, A)) < \infty \right\}.$$

We prove next the central result of this section, that both continuous and discontinuous finite elements deliver the same asymptotic rate over admissible subdivisions.

PROPOSITION 5.2 (Equivalence of Classes). *If Condition 2 holds, then for $0 < s \leq n/d$ we have $\mathbb{A}_s^0 \equiv \mathbb{A}_s$. Moreover, total errors are equivalent on the same mesh.*

Proof. The inclusion $\mathbb{A}_s^0 \subset \mathbb{A}_s$ is straightforward because for $(u, f, A) \in \mathbb{A}_s^0$

$$|u, f, A|_s \leq |u, f, A|_s^0.$$

We now prove $\mathbb{A}_s \subset \mathbb{A}_s^0$. To this end, let $(u, f, A) \in \mathbb{A}_s$ and, for $N > \#\mathcal{T}_0$, let $\mathcal{T}_* \in \mathbb{T}_N$, $V_* \in \mathbb{V}(\mathcal{T}_*)$ be so that

$$|||u - V_*|||_{\mathcal{T}_*} + \mathrm{osc}_{\mathcal{T}_*}(V_*, \mathcal{T}_*) = \inf_{\mathcal{T} \in \mathbb{T}_N} \inf_{V \in \mathbb{V}(\mathcal{T})} \left(|||u - V|||_\mathcal{T} + \mathrm{osc}_\mathcal{T}(V, \mathcal{T}) \right).$$

We prove that $(u, f, A) \in \mathbb{A}_s^0$ upon showing that $I_{\mathcal{T}_*} V_* \in \mathbb{V}^0(\mathcal{T}_*)$ satisfies

$$|||u - I_{\mathcal{T}_*} V_*|||_{\mathcal{T}_*} + \mathrm{osc}_{\mathcal{T}_*}(I_{\mathcal{T}_*} V_*, \mathcal{T}_*) \preceq N^{-s}, \tag{5.4}$$

where the interpolation operator $I_{\mathcal{T}_*}$ on $\mathbb{V}^0(\mathcal{T}_*)$ is given by Condition **2**. Using triangle inequality we obtain

$$|||u - I_{\mathcal{T}_*} V_*|||_{\mathcal{T}_*} \leq \left\| A^{1/2} \nabla(u - V_*) \right\|_{\mathcal{T}_*} + \left\| A^{1/2} \nabla(V_* - I_{\mathcal{T}_*} V_*) \right\|_{\mathcal{T}_*}.$$

Interpolation estimate (2.28) yields

$$\left\| A^{1/2} \nabla(V_* - I_{\mathcal{T}_*} V_*) \right\|_{\mathcal{T}_*} \preceq \left\| h^{-1/2} [\![V_*]\!] \right\|_{\Sigma_*} \leq |||u - V_*|||_{\mathcal{T}_*}, \tag{5.5}$$

whence

$$|||u - I_{\mathcal{T}_*} V_*|||_{\mathcal{T}_*} \preceq |||u - V_*|||_{\mathcal{T}_*}. \tag{5.6}$$

The oscillation term is treated similarly. In fact, Lemma 2.5 and (5.5) imply

$$\mathrm{osc}_{\mathcal{T}_*}(I_{\mathcal{T}_*} V_*, \mathcal{T}_*) \preceq \mathrm{osc}_{\mathcal{T}_*}(V_*, \mathcal{T}_*) + |||u - V_*|||_{\mathcal{T}_*}.$$

Coupling this estimate with (5.6), we end up with

$$|||u - I_{\mathcal{T}_*} V_*|||_{\mathcal{T}_*} + \mathrm{osc}_{\mathcal{T}_*}(I_{\mathcal{T}_*} V_*, \mathcal{T}_*) \preceq |||u - V_*|||_{\mathcal{T}_*} + \mathrm{osc}_{\mathcal{T}_*}(V_*, \mathcal{T}_*).$$

Since $(u, f, A) \in \mathbb{A}_s$, this implies (5.4) and so that $(u, f, A) \in \mathbb{A}_s^0$, as desired. □

5.3. Cardinality of $\mathcal{M}_k$. In order to derive a quasi-optimal decay of the total error $|||u - U_k|||_{\mathcal{T}_k} + \mathrm{osc}_{\mathcal{T}_k}(U_k, \mathcal{T}_k)$, we assume the following condition regarding the parameters (θ, γ) and the marking procedure MARK. Let

$$\gamma_* := \max\left(\frac{C_{LU}(2 + \Lambda)}{C_{GL}}, \gamma_m, \gamma_O\right), \quad \theta_*(\gamma) := \left(\frac{C_{GL} - C_{LU}\gamma^{-1}(2 + \Lambda)}{1 + C_{LU}(2 + \Lambda)}\right)^{1/2}$$

for $\gamma > \gamma_*$, where Λ is given by Lemma 2.5, γ_m by Theorem 4.4 and γ_O by Lemma 5.1. Notice that $\gamma > \gamma_*$ yields $\theta_* > 0$ and $C_{GL} < C_{LU}$ in turn implies that $\theta_* < 1$.

Condition 5 (Assumptions about MARK). *Let the set of marked elements $\mathcal{M}_k$ and marking parameter θ satisfy*

$$\mathcal{M}_k \text{ has minimal cardinality} \quad \text{and} \quad \theta \in (0, \theta_*).$$

In order to simplify the notation, we let $0 < \mu < 1/2$ be defined by

$$\mu(\gamma, \theta) := \frac{C_{GL} - C_{LU}\gamma^{-1}(2 + \Lambda)}{2C_{GL}}\left(1 - \frac{\theta^2}{\theta_*^2}\right), \quad \forall \gamma > \gamma_*, 0 < \theta < \theta_*.$$

REMARK 5.3 (Approximate Sort). *Condition **5** can be relaxed to*

$$\#\mathcal{M}_k \leq C\#M_k, \quad \text{and} \quad \theta \in (0, \theta_*),$$

where M_k is a set with minimal cardinality and C is a constant independent of k. This might be useful to improve upon the $O(\#\mathcal{T}_k \log(\#\mathcal{T}_k))$ operations required to perform an exact sort.

LEMMA 5.4 (Optimal Marking). *Let Condition **3** (properties of REFINE) hold. Let $\mathcal{T} \geq \mathcal{T}_0$ be a refinement of $\mathcal{T}_0$ obtained by algorithm (3.1), let $\mathcal{T}_* \geq \mathcal{T}$ be an admissible refinement of $\mathcal{T}$, and let $U \in \mathbb{V}(\mathcal{T}), U_* \in \mathbb{V}(\mathcal{T}_*)$ be the corresponding solutions of (2.11). In addition, assume that*

$$|||u - U_*^0|||_{\mathcal{T}_*}^2 + \mathrm{osc}_{\mathcal{T}_*}^2(U_*^0, \mathcal{T}_*) \leq \mu\left(|||u - U|||_{\mathcal{T}}^2 + \mathrm{osc}_{\mathcal{T}}^2(U, \mathcal{T})\right), \tag{5.7}$$

where $U_* := U_*^0 + U_*^\perp$ according to the decomposition (2.22). Then, for $\gamma > \gamma_*$ and $\theta \in (0, \theta_*(\gamma))$, the set $\omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})$, see (2.26), satisfies a Dörfler marking property

$$\eta_{\mathcal{T}}(U, \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})) \geq \theta\eta_{\mathcal{T}}(U, \mathcal{T}).$$

Proof. Since $0 < \mu < \frac{1}{2}$, the lower bound (3.8) in conjunction with (5.7) implies

$$(1 - 2\mu)C_{GL}\eta_{\mathcal{T}}^2(U, \mathcal{T}) \leq |||u - U|||_{\mathcal{T}}^2 - 2|||u - U_*^0|||_{\mathcal{T}_*}^2 + \mathrm{osc}_{\mathcal{T}}^2(U, \mathcal{T}) - 2\mathrm{osc}_{\mathcal{T}_*}^2(U_*^0, \mathcal{T}_*).$$

We estimate the terms involving the energy norm and the oscillation separately. For the former, the triangle inequality and Lemma 3.5 yield

$$|||u - U|||_{\mathcal{T}}^2 - 2|||u - U_*^0|||_{\mathcal{T}_*^0}^2 \leq 2|||U - U_*^0|||_{\mathcal{T}}^2$$
$$\leq 2C_{LU}\eta_{\mathcal{T}}^2(U, \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})) + 2C_{LU}\gamma^{-1}\eta_{\mathcal{T}}^2(U, \mathcal{T}).$$

For the latter we distinguish the case whether an element belongs to $\omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})$ or not

$$\mathrm{osc}_{\mathcal{T}}^2(U, \mathcal{T}) = \mathrm{osc}_{\mathcal{T}}^2(U, \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})) + \mathrm{osc}_{\mathcal{T}}^2(U, \mathcal{T} \setminus \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})).$$

Since the estimator dominates the oscillation elementwise (see (2.20)), we obtain

$$\operatorname{osc}^2_{\mathcal{T}}(U, \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})) \le \eta^2_{\mathcal{T}}(U, \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})),$$

while Lemma 2.5 gives

$$\operatorname{osc}^2_{\mathcal{T}}(U, \mathcal{T} \setminus \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})) - 2\operatorname{osc}^2_{\mathcal{T}_*}(U^0_*, \mathcal{T}_*) \le \Lambda \left\| A^{1/2} \nabla (U^0_* - U) \right\|^2_{\mathcal{T}}.$$

Collecting these estimates, and using the localized upper bound of Lemma 3.5, we get

$$(1 - 2\mu)C_{GL}\eta^2_{\mathcal{T}}(U, \mathcal{T}) \le (1 + C_{LU}(2 + \Lambda))\eta^2_{\mathcal{T}}(U, \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})) + C_{LU}\gamma^{-1}(2 + \Lambda)\eta^2_{\mathcal{T}}(U, \mathcal{T}),$$

whence, employing the definitions of θ_* and μ, results in

$$\eta^2_{\mathcal{T}}(U, \omega(\mathcal{R}_{\mathcal{T} \to \mathcal{T}_*})) \ge \frac{(1 - 2\mu)C_{GL} - C_{LU}\gamma^{-1}(2 + \Lambda)}{1 + C_{LU}(2 + \Lambda)}\eta^2_{\mathcal{T}}(U, \mathcal{T}) = \theta^2\eta^2_{\mathcal{T}}(U, \mathcal{T}).$$

This is the desired estimate. $\square$

In order to estimate the cardinality of $\mathcal{M}_k$ we need an additional condition on the overlays defined in Section 2.1.2.

Condition 6 (Mesh Overlay). *Let $\mathcal{T}_A$ and $\mathcal{T}_B$ be such that Condition 2 (interpolation operator) holds. Then, condition 2 holds on $\mathcal{T}_A \oplus \mathcal{T}_B$.*

LEMMA 5.5 (Cardinality of $\mathcal{M}_k$). *Let Conditions 3, 4, 5 and 6 be valid. Let $u \in H^1_0(\Omega)$ be the solution of (1.1) and $U_k \in \mathbb{V}(\mathcal{T}_k)$ be the k-th solution of (2.11) generated by (3.1). If $(u, f, A) \in \mathbb{A}_s$, then the following estimate holds for $\gamma > \gamma_*$*

$$\#\mathcal{M}_k \le C\mu(\gamma, \theta)^{-\frac{1}{2s}} |u, f, A|_s^{\frac{1}{s}} C_{OPT}^{\frac{1}{2s}} \left\{ |||u - U_k|||^2_{\mathcal{T}_k} + \operatorname{osc}^2_{\mathcal{T}_k}(U_k, \mathcal{T}_k) \right\}^{-\frac{1}{2s}},$$

where C_{OPT} is the constant in Lemma 5.1 and C is a constant independent of k.

Proof. Set

$$\epsilon^2 := \mu C_{OPT}^{-1} \left(|||u - U_k|||^2_{\mathcal{T}_k} + \operatorname{osc}^2_{\mathcal{T}_k}(U_k, \mathcal{T}_k) \right). \tag{5.8}$$

Since $(u, f, A) \in \mathbb{A}_s \equiv \mathbb{A}^0_s$ by Proposition 5.2, there exists $(V^0_\epsilon, \mathcal{T}_\epsilon)$ with $V^0_\epsilon \in \mathbb{V}^0(\mathcal{T}_\epsilon), \mathcal{T}_\epsilon \in \mathbb{T}$ so that

$$\#\mathcal{T}_\epsilon \preceq |u, f, A|_s^{1/s} \epsilon^{-1/s}, \qquad |||u - V^0_\epsilon|||^2_{\mathcal{T}_\epsilon} + \operatorname{osc}^2_{\mathcal{T}_\epsilon}(V^0_\epsilon, \mathcal{T}_\epsilon) \le \epsilon^2. \tag{5.9}$$

In order to relate $(U_k, \mathcal{T}_k)$ to $(V^0_\epsilon, \mathcal{T}_\epsilon)$, we overlay $\mathcal{T}_k$ and $\mathcal{T}_\epsilon$

$$\mathcal{T}_* := \mathcal{T}_k \oplus \mathcal{T}_\epsilon.$$

Let $U_* \in \mathbb{V}(\mathcal{T}_*)$ be the solution to (2.11) and use the orthogonal decomposition (2.22) to write $U_* = U^0_* + U^\perp_*$. In view of Remark 2.7, $U^0_* \in \mathbb{V}^0(\mathcal{T}_*)$ satisfies

$$a_{\mathcal{T}_*}(U^0_*, W^0) = (f, W^0)_{\mathcal{T}_*}, \qquad \forall W^0 \in \mathbb{V}^0(\mathcal{T}_*).$$

We proceed as in [21] to show that there is a reduction of the total error between U^0_* and V^0_ϵ. Since Condition 2 (interpolation operator) holds for $\mathcal{T}_*$, via Condition 6 (mesh overlay), we apply Lemma 5.1 for continuous solutions on $\mathcal{T}_* \ge \mathcal{T}_\epsilon$ to infer that

$$|||u - U^0_*|||^2_{\mathcal{T}_*} + \operatorname{osc}^2_{\mathcal{T}_*}(U^0_*, \mathcal{T}_*) \le C_{OPT} \left\{ |||u - V^0_\epsilon|||^2_{\mathcal{T}_\epsilon} + \operatorname{osc}^2_{\mathcal{T}_\epsilon}(V_\epsilon, \mathcal{T}_\epsilon) \right\}$$

$$\le C_{OPT}\epsilon^2 = \mu \left\{ |||u - U_k|||^2_{\mathcal{T}_k} + \operatorname{osc}^2_{\mathcal{T}_k}(U_k, \mathcal{T}_k) \right\}.$$

This, combined with Lemma 5.4, implies that $\omega(\mathcal{R}_{\mathcal{T}_k \to \mathcal{T}_*})$ verifies the Dörfler property

$$\eta_{\mathcal{T}_k}(U_k, \omega(\mathcal{R}_{\mathcal{T}_k \to \mathcal{T}_*})) \geq \theta \eta_{\mathcal{T}_k}(U_k, \mathcal{T}_k), \qquad \forall \theta \in (0, \theta_*).$$

Since MARK selects a minimal set $\mathcal{M}_k \subset \mathcal{T}_k$, according to Condition **5** (assumptions about MARK), we deduce

$$\#\mathcal{M}_k \leq \#\omega(\mathcal{R}_{\mathcal{T}_k \to \mathcal{T}_*}) \preceq \#\mathcal{R}_{\mathcal{T}_k \to \mathcal{T}_*},$$

where we used (2.25) and shape regularity of $\mathcal{T}$ (Condition **3**) to get the last estimate. Gathering the estimates above and using (2.5) (cardinality of mesh overlay) we obtain

$$\#\mathcal{M}_k \preceq \#\mathcal{R}_{\mathcal{T}_k \to \mathcal{T}_*} \leq \#\mathcal{T}_* - \#\mathcal{T}_k \leq \#\mathcal{T}_\epsilon.$$

Finally, combining this with (5.9) and (5.8), we reach the desired result. □

REMARK 5.6 (Exploiting the Equivalence of Classes). *The negative power of the local meshsize h, present in the energy norm $|||\cdot|||_{\mathcal{T}}$, is a source of difficulties when comparing discrete discontinuous functions on two different meshes. To by-pass this issue, we resort to the equivalence $\mathbb{A}_s \equiv \mathbb{A}_s^0$ and the continuous Ritz-Galerkin projection.*

5.4. Main Result 2: Quasi-Optimal Rate of Convergence. We are now ready to prove the second main result announced in the introduction.

THEOREM 5.7 (Quasi-Optimality). *Let Conditions **3**, **4**, **5** and **6** be valid. Let $\mathcal{T}_0$ be the initial conforming subdivision. Let $f \in L^2(\Omega)$, A satisfy (2.1) and $u \in H_0^1(\Omega)$ be the solution of (1.1). Let $\{\mathcal{T}_k, \mathbb{V}(\mathcal{T}_k), U_k\}_{k \geq 0}$ be the sequence of meshes, finite element spaces, and discrete solutions generated by ADFEM. Let γ_* and θ_* be given as in Section 5.3. If $\gamma > \gamma_*$, $\theta \in (0, \theta_*(\gamma))$ and $(u, f, A) \in \mathbb{A}_s$, then there holds*

$$\left\{ |||u - U_k|||_{\mathcal{T}_k}^2 + \mathrm{osc}_{\mathcal{T}_k}^2(U_k, \mathcal{T}_k) \right\}^{1/2} \leq C \# \mathcal{T}_k^{-s},$$

where C is a constant only depending on $\alpha, n, \mathcal{T}_0, A, \theta, s, \gamma$ and $|u, f, A|_s$.

Proof. Condition **4** and Lemma 5.5 yield

$$\#\mathcal{T}_k - \#\mathcal{T}_0 \preceq \sum_{j=0}^{k-1} \#\mathcal{M}_j \preceq \sum_{j=0}^{k-1} \left\{ |||u - U_j|||_{\mathcal{T}_j}^2 + \mathrm{osc}_{\mathcal{T}_j}^2(U_j, \mathcal{T}_j) \right\}^{-\frac{1}{2s}}$$

$$\preceq \sum_{j=0}^{k-1} \left\{ |||u - U_j|||_{\mathcal{T}_j}^2 + \eta_{\mathcal{T}_j}^2(U_j, \mathcal{T}_j) \right\}^{-\frac{1}{2s}},$$

where we have used (5.1). The contraction property of Theorem 4.4 implies

$$\#\mathcal{T}_k - \#\mathcal{T}_0 \preceq \left\{ |||u - U_k|||_{\mathcal{T}_k}^2 + \eta_{\mathcal{T}_k}^2(U_k, \mathcal{T}_k) \right\}^{-\frac{1}{2s}} \sum_{j=1}^{k-1} \alpha^{j/s}.$$

Since $\alpha < 1$ we obtain $\sum_{j=1}^{k-1} \alpha^{j/s} \leq \frac{1}{1-\alpha^{1/s}}$. Finally, invoking again (5.1) and noticing that $\#\mathcal{T}_k \leq C(\#\mathcal{T}_k - \#\mathcal{T}_0)$, with $C > 0$ independent of k, completes the proof. □

6. Nonconforming Hexahedral and Tetrahedral Meshes. In this section we briefly discuss under what additional assumptions, the standard refinement procedures described in Section 2.1.2 satisfy Conditions **3**, **4** and **6**, thereby yielding the quasi-optimal rate of convergence of Theorem 5.7. In particular, we shall see that we need to restrict the "level of incompatibility", which corresponds to the "level of noncomformity" for refinement by bisection on tetrahedra.

We consider a generalization of the K-meshes introduced by Babuška and Miller [6] to examine both hexahedral and tetrahedral meshes at once; see Section 6.1. We discuss in Section 6.2 a REFINE procedure ensuring that the above mesh restrictions are met. Finally, we verify Conditions **3**, **4**, **6** for hexahedral meshes in Section 6.3 and for tetrahedral meshes in 6.5; red refinement of tetrahedra is somewhat in between these two classes and is only mentioned in Section 6.4. As a by product, we obtain

convergence and optimal cardinality of continuous Galerkin adaptive methods on nonconforming meshes (see Section 7),

which does not seem to be available in the literature. To avoid technicalities, we restrict our discussion to polynomial degree one but the arguments extend to higher polynomial degree.

6.1. Domain of Influence and Admissible Meshes. We start by defining a basis of $\mathbb{V}^0(\mathcal{T})$, leading to the notion of domain of influence and admissible meshes.

Among the set of vertices $\mathcal{P}$ of $\mathcal{T}$, let $\mathcal{P}^0$ be the conforming ones and $N^0 = \#\mathcal{P}^0$. Let $P \in \mathcal{P}^0$ be a conforming node of $\mathcal{T}$. We construct a continuous piecewise linear basis function ψ_P over $\mathcal{T}$ recursively as in [6, Theorem 1.3.3]. We first rearrange the elements of $\mathcal{T}$ by level $\ell(T_1) \leq \ell(T_2) \leq ... \leq \ell(T_N)$. Suppose that ψ_P has been defined for each $T \in \mathcal{T}$ such that $l(T) < n$. If $T \in \mathcal{T}$ with $l(T') = n$ and Q is a vertex of T, we define ψ_P on T as follows:

- if Q is a conforming node, then set $\psi_P(Q) = 1$ if $Q = P$ and $\psi_P(Q) = 0$ otherwise;
- if Q is a hanging node, then Q belongs to an edge E contained in $T' \in \mathcal{T}$ with $\ell(T') < n$ and set $\psi_p(Q)|_T = \psi_p(Q)|_{T'}$.

Note that this definition is independent of the choice of T' since, by construction, ψ_P is continuous across interelements of lower level. Finally we observe that $\{\psi_P\}_{P \in \mathcal{P}^0}$ is a basis of $\mathbb{V}^0(\mathcal{T})$ because any $V \in \mathbb{V}^0(\mathcal{T})$ can be uniquely written as

$$V = \sum_{P \in \mathcal{P}^0} V(P)\psi_P. \tag{6.1}$$

The domain of influence of P is defined by $\mathcal{D}_\mathcal{T}(P) = \operatorname{supp}(\psi_P)$. Examples are depicted in Figs 2.3, 2.1 and 2.2. Moreover, we introduce the set

$$\omega(T) := \bigcup_{P:\ \mathcal{D}_\mathcal{T}(P) \cap T \neq \emptyset} \mathcal{D}_\mathcal{T}(P) \quad \forall T \in \mathcal{T}. \tag{6.2}$$

We say that a mesh $\mathcal{T}$ is *admissible*, if Condition **7** holds.

Condition 7 (Admissible Subdivision). *There exists a constant C uniform on the level of discretization of $\mathcal{T}$ such that for any conforming node $P \in \mathcal{P}^0$ of $\mathcal{T}$*

$$\operatorname{diam}(\mathcal{D}_\mathcal{T}(P)) \leq Ch_T, \qquad \forall T \in \mathcal{T},\ T \in \mathcal{D}_\mathcal{T}(P). \tag{6.3}$$

Condition **7** implies

$$\|\psi_P\|_{L^2(\mathcal{D}_\mathcal{T}(P))} \preceq h_T^{d/2} \quad \forall T \subset \mathcal{D}_\mathcal{T}(P), \tag{6.4}$$

which will be instrumental for contructing interpolation operators in Lemmas 6.6 and 6.9. For conforming meshes we find that $\mathcal{D}_{\mathcal{T}}(P) = \cup_{T \ni P} T$ so that (6.3) always holds with a constant only depending on the shape regularity of $\mathcal{T}$. The situation is strikingly different for nonconforming meshes: the number of hanging nodes per edge is obviously restricted by Condition **7** and the shape regularity of $\mathcal{T}$ (see Figs. 2.1-2.2). In addition, Condition **7** also restricts the maximum of level of incompatibility as illustrated in Figs. 2.1, 2.3 and 2.2: each side contained in the depicted patch has at most two hanging node and yet element levels differ up to 4. With the same pattern it is possible to reach any arbitrarily large level of nonconformity while keeping one hanging node per side for bisection and red refinements and two hanging nodes per side for quad refinement.

We now briefly discuss how to enforce Condition **7**. Let $T \in \mathcal{M}$ be an element of $\mathcal{T}$ marked for refinement by MARK. Each element $T \in \mathcal{M}$ has a tag $r(T)$, initially set to the desired number of refinements $\rho \geq 1$, and both T and $r(T)$ are updated as T is refined. Since subdivision of T may give rise to further refinement to keep the mesh admissible, the *local* procedure

$$[\mathcal{T}_*, \mathcal{M}_*] = \mathsf{MAKE\_ADMISSIBLE}(\mathcal{T}, \mathcal{M}, T)$$

subdivides T *once*, and perhaps other elements to enforce Condition **7** if necessary, and updates $(\mathcal{T}, \mathcal{M})$ to $(\mathcal{T}_*, \mathcal{M}_*)$; we give details in Sections 6.3 and 6.5 below. The mesh $\mathcal{T}_* \geq \mathcal{T}$ is the *smallest* refinement of $\mathcal{T}$ which is admissible, whereas $\mathcal{M}_*$ is obtained from $\mathcal{M}$ as follows: every time an element T is subdivided, T is removed from $\mathcal{M}$ and, provided that $r(T) \geq 2$, replaced by its children who inherit the tag $r(T) - 1$. This procedure must be iterated to perform ρ subdivisions.

6.2. Procedure REFINE and Condition 4. The aim of REFINE is to subdivide at least ρ times the elements in $\mathcal{M}$ while keeping the mesh admissible (see Condition **7**). It simply iterates MAKE_ADMISSIBLE until $\mathcal{M}$ is empty:

REFINE $(\mathcal{T}, \mathcal{M})$
for $T \in \mathcal{M}$ do
   $[\mathcal{T}, \mathcal{M}] = \mathsf{MAKE\_ADMISSIBLE}\ (\mathcal{T}, \mathcal{M}, T)$;
while $\mathcal{M} \neq \emptyset$
return$(\mathcal{T})$

To control the complexity of REFINE, namely the fulfillment of Condition **4**, we need to study MAKE_ADMISSIBLE more carefully. It turns out that the key properties to check are rather abstract and are valid regardless of the refinement strategy. This is formulated in the next lemma, whose original proof for simplices is due to Binev, Dahmen, and DeVore [15] for $d = 2$ and Stevenson [49] for $d > 2$. We refer to the survey [43, Section 4.5] for an extensive discussion of these properties.

LEMMA 6.1 (Condition **4**: Complexity of REFINE). *Let Condition **1** hold, $T \in \mathcal{M}$ and $T' \in \mathcal{T}_* \backslash \mathcal{T}$ be created by a call $[\mathcal{T}_*, \mathcal{M}_*] = MAKE\_ADMISSIBLE\ (\mathcal{T}, \mathcal{M}, T)$. If*

$$\ell(T') \leq \ell(T) + 1 \tag{6.5}$$

and there exists a geometric constant C, only depending on the shape regularity constant of $\mathcal{T}_0$ and the dimension d, so that the distance between T and T' satisfies

$$\mathrm{dist}(T, T') \leq C\beta_M^{\ell(T')}, \tag{6.6}$$

*with $0 < \beta_M < 1$ the constant in (2.3), then Condition **4** holds.*

6.3. Quad Refinement. No special initial labeling is necessary for hexahedral meshes. Condition **7** (admissible subdivision) can be written in more practical terms as follows. For any forest $\mathcal{F}$, we define $\mathcal{F}^1$ to be the biggest forest (in the sense of number of elements) such that $\mathcal{F}^1 \subset \mathcal{F}$ and $\mathcal{T}^1 := \mathcal{T}(\mathcal{F}^1)$ is a subdivision with at most one hanging node per side, hereafter called $1-mesh$. For any $T \in \mathcal{T}$, we also define its *1-compatible parent* $P_{\mathcal{T}}^1(T) \in \mathcal{T}^1$ to be such that $T \subset P_{\mathcal{T}}^1(T)$. The *level of incompatibility* of an element $T \in \mathcal{T}$ is then defined by

$$\ell^\perp(T) := \ell(T) - \ell(P_{\mathcal{T}}^1(T)). \tag{6.7}$$

We observe that $\ell^\perp(T) \geq 0$ and $\ell^\perp(T) = 0$ if and only if $T = P_{\mathcal{T}}^1(T) \in \mathcal{T}^1$ is an element of the biggest underlying $1-mesh$. Finally, the level of incompatibility of $\mathcal{T}$ is defined by

$$\ell^\perp(\mathcal{T}) := \max_{T \in \mathcal{T}} \ell^\perp(T).$$

Condition **7** (admissible subdivision) is equivalent to:

$$\textit{There exists a universal integer } L \geq 0 \textit{ such that } \ell^\perp(\mathcal{T}) \leq L. \tag{6.8}$$

This relation limits the local amount of elements not belonging to $\mathcal{T}^1$, as depicted in Figure 2.1, and is easier to check in practice. It is worth stressing that it is natural to have L uniformly bounded both algorithmically and theoretically.

REMARK 6.2 (Hanging Nodes vs Level of Incompatibility). *The number of hanging nodes per side is* unrelated *to the maximum level of incompatibility when $L \geq 2$. In fact, when $L \geq 2$, it is possible to construct a subdivision with an arbitrary large level of incompatibility. To see this, consider the example described in Figure 2.1 and continue to refine alternatively the lower left/upper right newly created children.*

The routine MAKE_ADMISSIBLE hinges on (6.8) and produces the smallest admissible refinement $\mathcal{T}_*$ of $\mathcal{T}$ so that T is refined at least once:

> MAKE_ADMISSIBLE $(\mathcal{T}, \mathcal{M}, T)$
> if $\ell^\perp(T) = L$ then
> $\mathcal{T}^1 = $ REFINE_COMPATIBLE $(\mathcal{T}^1, P_{\mathcal{T}}^1(T))$
> $\mathcal{T} = \mathcal{T} \oplus \mathcal{T}^1$ and update $\mathcal{M}$
> end
> if T has not been refined then
>   refine T and update $\mathcal{T}$ and $\mathcal{M}$
> end
> return $(\mathcal{T}, \mathcal{M})$

This procedure resorts to the routine REFINE_COMPATIBLE $(\mathcal{T}^1, P_{\mathcal{T}}^1(T))$ which acts recursively on $1-meshes$ $\mathcal{T}^1$, produces the smallest $1-mesh$ refinement $\mathcal{T}_*^1 \geq \mathcal{T}^1$, and refines $P_{\mathcal{T}}^1(T) \in \mathcal{T}^1$. Since REFINE_COMPATIBLE can only decrease the level of incompatibility, we find that the updated nonconforming mesh $\mathcal{T} \oplus \mathcal{T}^1$ satisfies $\ell^\perp(\mathcal{T} \oplus \mathcal{T}^1) \leq L$. Given $T \in \mathcal{T}^1$, we define its refinement patch to be

$$R(\mathcal{T}^1, T) := \left\{ T' \in \mathcal{T}^1 \mid T' \text{ and } T \text{ share an edge and } \ell(T') \leq \ell(T) \right\}.$$

The procedure REFINE_COMPATIBLE works solely on elements $T' \in R(\mathcal{T}^1, T)$ whose refinement is necessary to produce $1-meshes$, and reads as follows:

```
REFINE_COMPATIBLE (𝒯, T)
for T' ∈ R(𝒯, T) do
    if ℓ(T') = ℓ(T) − 1 then
        𝒯 = REFINE_COMPATIBLE (𝒯, T')
    end
end
if T has not been refined then
    subdivide T once and update 𝒯
end
return(𝒯)
```

We start by deducing properties of the REFINE_COMPATIBLE routine.

LEMMA 6.3 (Properties of REFINE_COMPATIBLE). *Let $\mathcal{T}$ be a $1 - mesh$ and $T \in \mathcal{T}$. Then* REFINE_COMPATIBLE $(\mathcal{T}, T)$ *terminates and produces the smallest 1-mesh refinement $\mathcal{T}_*$ of $\mathcal{T}$ such that T is refined and no element of $\mathcal{T}$ is refined more than once. Moreover, all newly created elements $T' \in \mathcal{T}_*$ satisfy* (6.5).

Proof. We first observe that the recursive nature of REFINE_COMPATIBLE, together with its conditional on $\ell(T') = \ell(T) - 1$, guarantees that this procedure creates the smallest $1-$mesh refinement $\mathcal{T}_*$ of $\mathcal{T}$ such that T is subdivided once.

Since a uniform refinement of $\mathcal{T}$ is a 1-mesh, and $\mathcal{T}_* \geq \mathcal{T}$ is minimal, we deduce that it must be a refinement of $\mathcal{T}_*$. This not only shows that REFINE terminates, but also that it subdivides elements of $\mathcal{T}$ at most once. This, in conjunction with the property $\ell(T') \leq \ell(T)$ for all elements $T' \in \mathcal{T}$ that are refined, yields (6.5) and completes the proof. □

We now show that the procedure REFINE implies Conditions **3**, **4**, **6**, and **7**. We recall that $\mathcal{T}_0$ is the initial conforming mesh, $L \geq 0$ is the constant in the admissibility condition (6.8), and $\rho \geq 1$ is the number of quad-refinement per marked element.

LEMMA 6.4 (Condition **7**: Admissible Subdivision). *Let $\mathcal{T}$ be a refinement of $\mathcal{T}_0$ satisfying* (6.8)*, and $\mathcal{M}$ be the set of elements of $\mathcal{T}$ to be refined. Then* REFINE $(\mathcal{T}, \mathcal{M})$ *terminates and produces the smallest refinement $\mathcal{T}_*$ of $\mathcal{T}$ fulfilling* (6.8). *The elements of $\mathcal{T}_*$ are obtained by at most ρ refinements of elements of $\mathcal{T}$, and the elements is $\mathcal{M}$ are refined exactly ρ times.*

Proof. We first note that MAKE_ADMISSIBLE only refines elements necessary to keep the mesh admissible, i.e. to satisfy condition (6.8). Therefore, in view of the update of $\mathcal{M}$ in MAKE_ADMISSIBLE, we see that $\mathcal{T}_* = $ REFINE $(\mathcal{T}, \mathcal{M})$ is the smallest refinement of $\mathcal{T}$ in which all the elements of $\mathcal{M}$ are refined ρ times. Moreover, since ρ uniform refinements of $\mathcal{T}$ preserve property (6.8), and so create an admissible partition, we deduce that it must be a refinement of $\mathcal{T}_*$. This not only shows that REFINE terminates, but also that it subdivides elements of $\mathcal{T}$ at most ρ times.

We now prove that $\mathcal{T}_*$ satisfies (6.8). If $\ell^\perp(T) = L$ for $T \in \mathcal{M}$, then a subdivision of T would violate (6.5). The call to REFINE_COMPATIBLE subdivides the 1-compatible parent $P^1_{\mathcal{T}}(T) \in \mathcal{T}^1$, thereby decreasing by at least 1 the level of incompatibility of all elements of $\mathcal{T}$ contained in $P^1_{\mathcal{T}}(T)$, and creates a 1-mesh refinement $\mathcal{T}^1_*$ of $\mathcal{T}^1$. This process is exemplified in Figure 6.1.

If T has not been refined after the first conditional of MAKE_ADMISSIBLE, then $\ell^\perp(T) < L$ because REFINE_COMPATIBLE reduced the level of incompatibility of T or the conditional was skipped altogether. Consequently, refining T increases $\ell^\perp(T)$ by 1, namely $\ell^\perp(T) \leq L$, whence (6.8) is verified. □

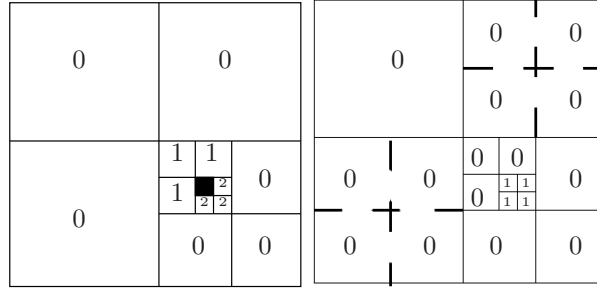LEMMA 6.5 (Condition **4**: Complexity of REFINE). *The procedure* REFINE *based*

Fig. 6.1. *Refinement procedure for* $d = 2$, $L = 2$ *and* $r = 1$. *The number on each element* T *indicates its level of incompatibility* $\ell^\perp(T)$. *Left: mesh* $\mathcal{T}$, *underlying* $1-mesh$ $\mathcal{T}^1$ *(bold lines), and shaded element to be refined with level of incompatibility* 2; *this forces a call to* REFINE_COMPATIBLE. *Right:* $1-mesh$ *refinement of* $\mathcal{T}^1$ *(dashed lines) via* REFINE_COMPATIBLE *to reduce the level of incompatibility of the shaded element to* 1. *Refinement of the shaded element leads now to an admissible mesh.*

on MAKE_ADMISSIBLE *satisfies*

$$\#\mathcal{T}_k - \mathcal{T}_0 \leq \Lambda_0 \sum_{j=0}^{k-1} \#\mathcal{M}_j,$$

with a constant Λ_0 *solely depending on* $\mathcal{T}_0$, L, *and* ρ.

Proof. Let $\mathcal{M} := \cup_{j=0}^{k-1} \mathcal{M}_j$ be the set of all marked elements to go from $\mathcal{T}_0$ to $\mathcal{T}_k$. The assertion follows from Lemma 6.1 provided we can show (6.5) and (6.6) (with $\beta_M = 1/2$) for every $T \in \mathcal{M}$ and newly created element T' by a call to MAKE_ADMISSIBLE. The proof of (6.6) is omitted since it follows strictly the proof by recursion given in [15] and [49] for triangles and tetrahedra; see also [43]. Hence, we focus on (6.5). If $\ell^\perp(T) = L$, then the newly created elements $T' \in \mathcal{T}_*^1 =$ REFINE_COMPATIBLE $(\mathcal{T}^1, P_\mathcal{T}^1(T))$ satisfy

$$\ell(T') \leq \ell(P_\mathcal{T}^1(T)) + 1 \leq \ell(T) - L + 1 \leq \ell(T) + 1,$$

according to Lemma 6.3. Consequently, the newly created elements of $\mathcal{T}_* = \mathcal{T} \oplus \mathcal{T}_*^1$ are some, but perhaps not all, of those $T' \in \mathcal{T}_*^1$ because $\mathcal{T}^1 \subset \mathcal{T}$.

It remains to examine what happens when $T \in \mathcal{M}$ has not been refined after the first conditional of MAKE_ADMISSIBLE. In this case, the refinement stays within T and its children T' satisfy $\ell(T') = \ell(T) + 1$.

This conclude the proof. □

LEMMA 6.6 (Condition **2**: Interpolation Operator). *Let* $\mathcal{T} \geq \mathcal{T}_0$ *be an admissible refinement of* $\mathcal{T}_0$. *Then, there exists an interpolation operator* $I_\mathcal{T} : \mathbb{E}(\mathcal{T}) \to \mathbb{V}^0(\mathcal{T})$ *and a constant* C *only depending on* $\mathcal{T}_0, L$ *and* ρ *such that* (2.27) *and* (2.28) *are satisfied. Moreover,* $I_\mathcal{T}$ *is locally a projection.*

Proof. Following Clemént [25], we construct $I_\mathcal{T} : \mathbb{E}(\mathcal{T}) \to \mathbb{V}^0(\mathcal{T})$ via local L^2-projections. We proceed in several steps.

1 Let $v \in \mathbb{E}(\mathcal{T})$. For each conforming node P let $\mathbb{V}^0(\mathcal{D}_\mathcal{T}(P)) := \Pi_{T \in \mathcal{D}_\mathcal{T}(P)} \mathbb{P}^n(T) \cap H_0^1(\Omega)$ and define $V_P \in \mathbb{V}^0(\mathcal{D}_\mathcal{T}(P))$ to be the solution of

$$\int_{\mathcal{D}_\mathcal{T}(P)} (v - V_P) W = 0, \qquad \forall W \in \mathbb{V}^0(\mathcal{D}_\mathcal{T}(P)), \tag{6.9}$$

where $\mathcal{D}_{\mathcal{T}}(P)$ denotes the domain of influence of P; see Section 6.1. The functions V_P evaluated at P give the nodal values of $I_{\mathcal{T}}v$, namely

$$I_{\mathcal{T}}v := \sum_{P \in \mathcal{P}^0} V_P(P)\psi_P. \qquad (6.10)$$

If $v \in \mathbb{V}^0(\mathcal{D}_{\mathcal{T}}(P))$, then $V_P(P) = v(P)$ and $I_{\mathcal{T}}$ is locally a projection.
[2] Definition (6.9) implies $\|V_P\|_{L^2(\mathcal{D}_{\mathcal{T}}(P))} \leq \|v\|_{L^2(\mathcal{D}_{\mathcal{T}}(P))}$, whence

$$\|I_{\mathcal{T}}v\|_{L^2(T)} \preceq \|v\|_{L^2(\omega(T))} \quad \forall T \in \mathcal{T}, \qquad (6.11)$$

where $\omega(T)$ is defined in (6.2). Since $I_{\mathcal{T}}$ reproduces constants exactly, namely

$$I_{\mathcal{T}}c|_T = c \quad \forall c \in \mathbb{P}^0(\omega(T)),$$

we have

$$\|v - I_{\mathcal{T}}v\|_{L^2(T)} = \|(v - c) - I_{\mathcal{T}}(v - c)\|_{L^2(T)} \preceq \|v - c\|_{L^2(\omega(T))} \quad \forall c \in \mathbb{P}^0(\omega(T)).$$

The Bramble-Hilbert lemma [16, Lemma 4.3.8], [24, Theorem 3.1.1] yields (2.27)

$$\|v - I_{\mathcal{T}}v\|_{L^2(T)} \preceq h_T \|\nabla v\|_{L^2(\omega(T))} \quad \forall v \in H_0^1(\Omega).$$

[3] We now consider $v \in \mathbb{V}(\mathcal{T})$ which may thus jump across interelement boundaries. We scale $\mathcal{D}_{\mathcal{T}}(P)$ to a reference domain with unit diameter and examine the seminorm $\|[\![v]\!]\|_{L^2(\Sigma \cap \mathcal{D}_{\mathcal{T}}(P))}$ on the space of discontinuous piecewise polynomials

$$\left\{ v \in \Pi_{T \in \mathcal{D}_{\mathcal{T}}(P)} \mathbb{P}^n(T) \mid V_P = 0 \right\},$$

where V_P is defined by (6.9) [11, 14]. If this seminorm vanishes then v is continuous and thus $v \in \mathbb{V}^0(\mathcal{D}_{\mathcal{T}}(P))$, whence $v = V_P = 0$ according to (6.9). This implies that the seminorm is equivalent to any norm because of the finite dimensional structure of the space. Consequently scaling gives

$$\|D^a(v - V_P)\|_{L^2(\mathcal{D}_{\mathcal{T}}(P);\mathcal{T})} \leq C \left\|h^{\frac{1}{2}-|a|} [\![v]\!]\right\|_{L^2(\Sigma \cap \mathcal{D}_{\mathcal{T}}(P))} \quad \forall |a| = 0, 1, \qquad (6.12)$$

where C is a constant depending on the integer L defined in (6.8).
[4] Finally, let $T \in \mathcal{T}$ and $P, Q \in \omega(T) \cap \mathcal{P}^0$. Then, in view of (6.12), we arrive at

$$\|V_P - V_Q\|_{L^2(T)} \preceq \left\|h^{1/2} [\![v]\!]\right\|_{L^2(\sigma(T))}, \qquad (6.13)$$

where $\sigma(T) = \Sigma \cap \omega(T)$. Consequently, (6.10) yields

$$v - I_{\mathcal{T}}v = v - \sum_{P \in \omega(T) \cap \mathcal{P}^0} V_P \psi_P = (v - V_Q) - \sum_{P \in \omega(T) \cap \mathcal{P}^0} (V_P - V_Q)\psi_P,$$

which, combined with (6.12) and (6.13), implies

$$\|v - I_{\mathcal{T}}v\|_{L^2(T)} \preceq \left\|h^{1/2} [\![v]\!]\right\|_{L^2(\sigma(T))}.$$

This is the desired estimate (2.28). The proof is thus complete. $\square$

Since Conditions **3**(a-c) are obvious (with $N_T = \rho$ according to Lemma 6.4) and Condition **3**(d) is a consequence of Lemma 6.6, it remains to show Condition **6**.

LEMMA 6.7 (Condition **6**: Overlay). *If $\mathcal{T}_A, \mathcal{T}_B \in \mathbb{T}$ are two refinements of $\mathcal{T}_0$ satisfying* (6.8)*, then the overlay $\mathcal{T} := \mathcal{T}_A \oplus \mathcal{T}_B$ satisfies* (6.8) *as well as Condition* **2**.

Proof. If $T \in \mathcal{T}$, then by construction T belongs to either $\mathcal{T}_A$ or $\mathcal{T}_B$, say $\mathcal{T}_A$; thus $\ell(T) - \ell(P^1_{\mathcal{T}_A}(T) \le L$. The underlying 1-meshes satisfy $\mathcal{T}^1_A \le \mathcal{T}^1$ because $\mathcal{T}_A \le \mathcal{T}$ and $\mathcal{T}^1$ is the largest 1-mesh, in the sense of trees, having $\mathcal{T}$ as a refinement. This implies $\ell(P^1_{\mathcal{T}}(T)) \ge \ell(P^1_{\mathcal{T}_A}(T))$, whence

$$\ell^\perp(T) = \ell(T) - \ell(P^1_{\mathcal{T}}(T)) \le \ell(T) - \ell(P^1_{\mathcal{T}_A}(T)) \le L.$$

Therefore, $\mathcal{T}$ is admissible and Condition **2** holds in view of Lemma 6.6. □

6.4. Red Refinement. As in Section 6.3, there is no need for an initial labeling and shape regularity of any sequence of meshes obtained using red refinement depends only on the initial mesh $\mathcal{T}_0$; this implies Condition **1**.

In contrast to quad-refinement, it is possible to create an arbitrary level of incompatibility with one hanging node per edge; to see this simply repeat the pattern in Figure 2.2. The definition of level of incompatibility $\ell^\perp(T)$ is again relative to the underlying 1-mesh, as in (6.7), and the practical Condition **7** becomes (6.8).

We observe that proof of Conditions **3**, **4**, **6**, and **7** are somewhat similar to those of Section 6.5 and are thus omitted.

6.5. Bisection Refinement. We assume that

> *the initial subdivision $\mathcal{T}_0$ is conforming and satisfies the labeling*
> *condition of Stevenson [49, condition (b) of Section 4].*   (6.14)

This ensures Condition **1** (atomic refinement); see [43, Section 4] for a discussion.

Condition **7** (admissible subdivision) can be written in more practical terms as follows. We resort to the construction of Section 6.3 for the quad refinement except that we now use an underlying conforming mesh instead of $1-$mesh. For any forest $\mathcal{F}$, $\mathcal{F}^0$ designates the biggest forest (in the sense of number of elements) such that $\mathcal{F}^0 \subset \mathcal{F}$ and $\mathcal{T}^0 := \mathcal{T}(\mathcal{F}^0)$ is conforming. For any $T \in \mathcal{T}$, we also define its *conforming parent* $P^0_{\mathcal{T}}(T) \in \mathcal{T}^0$ to be such that $T \subset P^0_{\mathcal{T}}(T)$. The *level of nonconformity* of an element $T \in \mathcal{T}$ is then given by

$$\ell^\perp(T) := \ell(T) - \ell(P^0_{\mathcal{T}}(T)).   \tag{6.15}$$

We observe that $\ell^\perp(T) \ge 0$ and $\ell^\perp(T) = 0$ if and only if $T = P^0_{\mathcal{T}}(T) \in \mathcal{T}^0$ is a conforming element. Finally, the level of nonconformity of $\mathcal{T}$ is defined by

$$\ell^\perp(\mathcal{T}) := \max_{T \in \mathcal{T}} \ell^\perp(T).$$

The concept of nonconformity is relative to 0-meshes, or conforming meshes, whereas that of incompatibility in Section 6.3 refers to 1-meshes. In any event, Condition **7** is again equivalent to:

> *There exists a universal integer $L \ge 0$ such that $\ell^\perp(\mathcal{T}) \le L$.*   (6.16)

This relation limits the local amount of nonconforming elements, as depicted in Figure 2.3, and is easier to check in practice. The routine MAKE_ADMISSIBLE hinges on it

and produces the smallest admissible refinement $\mathcal{T}_*$ of $\mathcal{T}$ by bisection (newest vertex bisection for $d = 2$) so that T is refined at least once:

> MAKE_ADMISSIBLE $(\mathcal{T}, \mathcal{M}, T)$
> if $\ell^\perp(T) = L$ then
> $\mathcal{T}^0 = \text{REFINE\_CONFORM} \ (\mathcal{T}^0, P_\mathcal{T}^0(T))$
> $\mathcal{T} = \mathcal{T} \oplus \mathcal{T}^0$
> end
> if T has not been refined then
>   bisect T and update $\mathcal{T}$ and $\mathcal{M}$
> end
> return $(\mathcal{T}, \mathcal{M})$

This procedure resorts to the routine $\text{REFINE\_CONFORM}(\mathcal{T}^0, P_\mathcal{T}^0(T))$, which proceeds recursively on conforming meshes $\mathcal{T}^0$ and produces the smallest conforming refinement $\mathcal{T}_*^0 \geq \mathcal{T}^0$ such that the level of nonconformity $\ell^\perp(\mathcal{T}_*) \leq L$ for all intermediate meshes and $P_\mathcal{T}^0(T) \in \mathcal{T}^0$ is refined. In [15, 49], REFINE_CONFORM has been proved to terminate; see also the survey [43].

We now show that this procedure implies Conditions **3**, **4**, **6**, and **7**.

LEMMA 6.8 (Condition **7**: Admissible Subdivision). *Let $\mathcal{T}_0$ satisfy* (6.14)*. Let $\mathcal{T}$ be a refinement of $\mathcal{T}_0$ satisfying* (6.16)*, and $\mathcal{M}$ be the set of elements of $\mathcal{T}$ to be refined. Then* REFINE $(\mathcal{T}, \mathcal{M})$ *terminates and produces the smallest refinement $\mathcal{T}_*$ of $\mathcal{T}$ such that* (6.16) *is still valid.*

Proof. The procedure REFINE_CONFORM acts on conforming meshes and terminates according to [49, Theorem 5.1] (since $\mathcal{T}_0$ satisfies (6.14)), and so does REFINE because the former is its only recursion.

The rest of the proof is identical to that of Lemma 6.4 and is thus omitted. □

LEMMA 6.9 (Condition **2**: Interpolation Operator). *Let $\mathcal{T}_0$ satisfy* (6.14)*. Then, for any admissible mesh $\mathcal{T} \geq \mathcal{T}_0$ there exists an interpolation operator $I_\mathcal{T} : \mathbb{E}(\mathcal{T}) \to \mathbb{V}^0(\mathcal{T})$ and a constant C only depending on $\mathcal{T}_0, L$ and ρ such that* (2.27) *and* (2.28) *are satisfied.*

Proof. Proceed exactly as in Lemma 6.6. □

Stevenson [49] proved that the initial labeling condition [49, condition (b) of Section 4] ensures that *any* uniform refinement (i.e. such that all the elements have the same level) is conforming. In particular, given $\mathcal{T}^0$ a conforming subdivision, if $\ell_{\max} := \max_{T \in \mathcal{T}^0} \ell(T)$, then refining recursively all the element of $\mathcal{T}^0$ until they all have a level equal to $\ell_{\max} + 1$ yields a conforming refinement of $\mathcal{T}^0$. Based on this idea, the next lemma provides an alternative way to construct a global refinement of any conforming subdivision, which turns out to yield Condition **3**(c). In the proof of the next lemma, we use the notion of proper t-subsimplices.

DEFINITION 6.10 (Proper Subsimplices). *If an element $T = $ conv hull $\{z_i\}_{i=0}^d$, then T' is a proper t-subsimplex of T if $T' = $ conv hull $\{z_j'\}_{j=0}^t$ where $t < d$ and $z_j' \in \{z_i\}_{i=0}^d$ for $0 \leq j \leq t$.*

The bisection rule, based on vertex order and type [49, 43], guarantees that each t-subsimplex of T is bisected exactly t-times whenever T is bisected d-times. In particular, edges of T, which are 1-subsimplices of T, are cut exactly once.

LEMMA 6.11 (Effect of d Uniform Refinements). *Let $\mathcal{T}_0$ satisfy* (6.14)*. If $\mathcal{T}$ is a conforming refinement of $\mathcal{T}_0$, then bisecting d times all the elements of $\mathcal{T}$ yields a conforming refinement $\mathcal{T}_*$ of $\mathcal{T}$.*

Proof. Let $T_1, T_2 \in \mathcal{T}$ be two adjacent elements sharing the t-subsimplex $K = $

$T_1 \cap T_2$ with $0 < t < d$ and let $\ell(T_1) \geq \ell(T_2)$. If suffices to show that after d bisections of T_1 and T_2, the t refinements of K induced by T_1 and T_2 are conforming.

We bisect T_2 further $\ell(T_1) - \ell(T_2)$ times, so that all the descendents of both T_1 and T_2 have the same level $\ell := \ell(T_1) + d$. Consequently, they belong to the uniform refinement $\mathcal{T}_\ell$ of $\mathcal{T}_0$ of level ℓ, which is conforming because $\mathcal{T}_0$ satisfies (6.14) [49, 43]. We now observe that T_2 cannot cause more than t bisections of K, because otherwise they would not be compatible with the t bisections induced by T_1 and they all belong to the same conforming mesh $\mathcal{T}_\ell$. This in turn shows that, after d refinements of T_1 and T_2, the t bisections of K are conforming as asserted. □

LEMMA 6.12 (Condition **3**: One-Step Refinement). *If $\mathcal{T}_0$ fulfills (6.14), then the REFINE procedure of Section 6.2 fulfills Condition **3** with $N = d\rho$.*

Proof. Condition **3**(a) follows from the initial labeling of $\mathcal{T}_0$ [10, 38, 39, 41, 49, 50], whereas the tagging of marked elements of Section 6.1 directly implies Condition **3**(b).

In order to prove Condition **3**(c), we construct an admissible refinement $\widetilde{\mathcal{T}}$ of $\mathcal{T}$ such that all the marked elements are refined at least ρ times and all the elements are refined at most ρd times. This implies that $\mathcal{T}_* \leq \widetilde{\mathcal{T}}$ because $\mathcal{T}_*$ is, by construction, the smallest refinement of $\mathcal{T}$ such that all the marked elements are refined ρ times; hence Condition **3**(c) follows with $N = \rho d$.

We construct $\widetilde{\mathcal{T}}$ in three steps. First we perform ρd successive uniform refinements of $\mathcal{T}^0$, the finest conforming mesh coarser than $\mathcal{T}$, and denote by $\widetilde{\mathcal{T}}^0$ the resulting subdivision. Lemma 6.11 ensures that $\widetilde{\mathcal{T}}^0$ is conforming. Second, we construct $\widetilde{\mathcal{T}}^\perp$ by refining ρd times all the elements of $\mathcal{T}^\perp = \mathcal{T} \setminus \mathcal{T}^0$. Finally, we set $\widetilde{\mathcal{T}} := \widetilde{\mathcal{T}}^0 \oplus \widetilde{\mathcal{T}}^\perp$ and claim that $\widetilde{\mathcal{T}}$ is admissible. To see this, we prove that $\ell^\perp(\widetilde{T}) \leq \ell^\perp(T)$ for all $\widetilde{T} \in \widetilde{\mathcal{T}}$ obtained by refinement of $T \in \mathcal{T}$. The latter is a simple consequence of the following observation: each $T \in \mathcal{T}$ is refined at most ρd times in steps 1 and 2, so that its children in the overlay $\widetilde{\mathcal{T}}$ arise from at most ρd refinements. Therefore,

$$\ell^\perp(\widetilde{T}) := \ell(\widetilde{T}) - \ell(P^0_{\widetilde{\mathcal{T}}}(\widetilde{T})) \leq \ell(T) + \rho d - \ell(P^0_{\mathcal{T}}(T)) - \rho d = \ell^\perp(T) \leq L,$$

which implies Condition **3**(c).

Finally, condition **3**(d) on the interpolation operator is a consequence of Lemma 6.9 coupled with the uniform bound on the level of nonconformity and shape regularity, both already proved to hold. This concludes the proof. □

LEMMA 6.13 (Condition **4**: Complexity of REFINE). *Let $\mathcal{T}_0$ satisfy (6.14). The procedure REFINE fulfills Condition **4** with a constant Λ_0 solely depending on $\mathcal{T}_0$, L, and ρ.*

Proof. Proceed as in Lemma 6.5. □

LEMMA 6.14 (Condition **6**: Overlay). *If $\mathcal{T}_A, \mathcal{T}_B \in \mathbb{T}$ are two refinements of $\mathcal{T}_0$ whose level of nonconformity is not bigger than $L \geq 0$, then the overlay $\mathcal{T} := \mathcal{T}_A \oplus \mathcal{T}_B$ has a level of nonconformity not bigger than L and Condition **2** holds on $\mathcal{T}$.*

Proof. Proceed as in Lemma 6.7. □

7. Continuous Galerkin Method on Nonconforming Meshes. The proof of quasi-optimality of dG hinges on comparing with the corresponding cG method. As a by-product, we thus obtain quasi-optimal cardinality of the cG method on nonconforming meshes for the refinement strategies of Section 6, thereby extending [21].

COROLLARY 7.1 (Quasi-Optimality of cG on Nonconforming Meshes). *Assume that the procedure REFINE consists of one of the strategies of Section 6, and that Condition **5** (assumptions on MARK) is valid. Let $\mathcal{T}_0$ be the initial conforming subdivision, A satisfy (2.1), and $u \in H_0^1(\Omega)$ be the solution of (1.1) with $f = -\mathrm{div}\,(A\nabla u) \in L^2(\Omega)$.*

Let $\left\{\mathcal{T}_k, \mathbb{V}^0(\mathcal{T}_k), U_k\right\}_{k \geq 0}$ *be the sequence of nonconforming meshes* $\mathcal{T}_k$, *conforming finite element spaces* $\mathbb{V}^0(\mathcal{T}_k) = \mathbb{V}(\mathcal{T}_k) \cap H_0^1(\Omega)$, *and discrete solutions* $U_k \in \mathbb{V}^0(\mathcal{T}_k)$ *generated by the continuous AFEM (cG). If* $\theta \in (0, \theta_*(\infty))$ *with* $\theta_*(\gamma)$ *given in Condition* **5**, *and* $(u, f, A) \in \mathbb{A}_s^0$, *then*

$$\left\{ |||u - U_k|||_{\mathcal{T}_k}^2 + \operatorname{osc}_{\mathcal{T}_k}^2(U_k, \mathcal{T}_k) \right\}^{1/2} \leq C \# \mathcal{T}_k^{-s},$$

where C *is a constant only depending on* $\alpha, n, \mathcal{T}_0, A, \theta, s$ *and* $|u, f, A|_s^0$.

Proof. This is a direct consequence of Proposition 5.2, Theorem 5.7, and the discussion of Section 6. □

We finally observe that the threshold $\theta_*(\infty)$ coincides with that derived in [21] for conforming bisection meshes.

REFERENCES

[1] M. Ainsworth. A synthesis of a posteriori error estimation techniques for conforming, nonconforming and discontinuous Galerkin finite element methods. In *Recent advances in adaptive computation*, volume 383 of *Contemp. Math.*, pages 1–14. Amer. Math. Soc., Providence, RI, 2005.

[2] M. Ainsworth. A posteriori error estimation for discontinuous Galerkin finite element approximation. *SIAM J. Numer. Anal.*, 45(4):1777–1798 (electronic), 2007.

[3] M. Ainsworth and J.T. Oden. *A Posteriori Error Estimation in Finite Element Analysis.* Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2000.

[4] D.N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19(4):742–760, 1982.

[5] 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(5):1749–1779 (electronic), 2001/02.

[6] I. Babuška and A. Miller. A feedback finite element method with a posteriori error estimation. I. The finite element method and some basic properties of the a posteriori error estimator. *Comput. Methods Appl. Mech. Engrg.*, 61(1):1–40, 1987.

[7] I. Babuška and M. Zlámal. Nonconforming elements in the finite element method with penalty. *SIAM J. Numer. Anal.*, 10:863–875, 1973.

[8] G. Baker, W.N. Jureidini, and O.A. Karakashian. Piecewise solenoidal vector fields and the stokes problem. *SIAM J. Numer. Anal.*, 27:1466–1485, 1990.

[9] G. A. Baker. Finite element methods for elliptic equations using nonconforming elements. *Math. Comp.*, 31(137):45–59, 1977.

[10] E. Bänsch. Local mesh refinement in 2 and 3 dimensions. *Impact Comput. Sci. Engrg.*, 3(3):181–191, 1991.

[11] E. Bänsch, P. Morin, and R.H. Nochetto. An adaptive Uzawa FEM for the Stokes problem: convergence without the inf-sup condition. *SIAM J. Numer. Anal.*, 40(4):1207–1229 (electronic), 2002.

[12] 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(2):267–279, 1997.

[13] R. Becker, P. Hansbo, and M. G. Larson. Energy norm a posteriori error estimation for discontinuous Galerkin methods. *Comput. Methods Appl. Mech. Engrg.*, 192(5-6):723–733, 2003.

[14] C. Bernardi and V. Girault. A local regularization operator for triangular and quadrilateral finite elements. *SIAM J. Numer. Anal.*, 35(5):1893–1916 (electronic), 1998.

[15] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

[16] S.C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[17] F. Brezzi, B. Cockburn, L. D. Marini, and E. Süli. Stabilization mechanisms in discontinuous Galerkin finite element methods. *Comput. Methods Appl. Mech. Engrg.*, 195(25-28):3293–3310, 2006.

[18] F. Brezzi, G. Manzini, D. Marini, P. Pietra, and A. Russo. Discontinuous Galerkin approximations for elliptic problems. *Numer. Methods Partial Differential Equations*, 16(4):365–378, 2000.

[19] K. Brix, M. Campos Pinto, and W. Dahmen. A multilevel preconditioner for the interior penalty discontinuous Galerkin method. *SIAM J. Numer. Anal.*, 46(5):2742–2768, 2008.

[20] R. Bustinza, G. N. Gatica, and B. Cockburn. An a posteriori error estimate for the local discontinuous Galerkin method applied to linear and nonlinear diffusion problems. *J. Sci. Comput.*, 22/23:147–185, 2005.

[21] J.M. Cascón, Ch. Kreuzer, R.H. Nochetto, and K.G. Siebert. Quasi-optimal rate of convergence of adaptive finite element methods. *SIAM Journal on Numerical Analysis*, 2008. to appear.

[22] P. Castillo. Performance of discontinuous Galerkin methods for elliptic PDEs. *SIAM J. Sci. Comput.*, 24(2):524–547 (electronic), 2002.

[23] Zh. Chen and J. Feng. An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems. *Math. Comp.*, 73(247):1167–1193 (electronic), 2004.

[24] Ph.G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[25] Ph. Clément. Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér., RAIRO Analyse Numérique*, 9(R-2):77–84, 1975.

[26] R.A. DeVore. Nonlinear approximation. In *Acta numerica, 1998*, volume 7 of *Acta Numer.*, pages 51–150. Cambridge Univ. Press, Cambridge, 1998.

[27] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.

[28] J. Douglas and T. Dupont. Interior penalty procedures for elliptic and parabolic Galerkin methods. In *Computing methods in applied sciences (Second Internat. Sympos., Versailles, 1975)*, pages 207–216. Lecture Notes in Phys., Vol. 58. Springer, Berlin, 1976.

[29] E. H. Georgoulis and E. Süli. Optimal error estimates for the *hp*-version interior penalty discontinuous Galerkin finite element method. *IMA J. Numer. Anal.*, 25(1):205–220, 2005.

[30] J. Gopalakrishnan and G. Kanschat. A multilevel discontinuous Galerkin method. *Numer. Math.*, 95(3):527–550, 2003.

[31] R. H. W. Hoppe, G. Kanschat, and T. Warburton. Convergence analysis of an adaptive interior penalty discontinuous Galerkin method. *SIAM J. Numer. Anal.*, 47(1):534–550, 2008/09.

[32] P. Houston, D. Schötzau, and T.P. Wihler. Mixed *hp*-discontinuous Galerkin finite element methods for the Stokes problem in polygons. In *Numerical Mathematics and Advanced Applications*, pages 493–501. Springer, Berlin, 2004.

[33] P. Houston, D. Schötzau, and T.P. Wihler. Energy norm a posteriori error estimation of *hp*-adaptive discontinuous Galerkin methods for elliptic problems. *Math. Models Methods Appl. Sci.*, 17(1):33–62, 2007.

[34] P. Houston, Ch. Schwab, and E. Süli. Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 39(6):2133–2163 (electronic), 2002.

[35] G. Kanschat and R. Rannacher. Local error analysis of the interior penalty discontinuous Galerkin method for second order elliptic problems. *J. Numer. Math.*, 10(4):249–274, 2002.

[36] O.A. Karakashian and F. Pascal. A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. *SIAM J. Numer. Anal.*, 41(6):2374–2399 (electronic), 2003.

[37] O.A. Karakashian and F. Pascal. Convergence of adaptive discontinuous Galerkin approximations of second-order elliptic problems. *SIAM J. Numer. Anal.*, 45(2):641–665, 2007.

[38] I. Kossaczký. A recursive approach to local mesh refinement in two and three dimensions. *J. Comput. Appl. Math.*, 55(3):275–288, 1994.

[39] J.M. Maubach. Local bisection refinement for *n*-simplicial grids generated by reflection. *SIAM J. Sci. Comput.*, 16(1):210–227, 1995.

[40] K. Mekchay and R.H. Nochetto. Convergence of adaptive finite element methods for general second order linear elliptic PDEs. *SIAM J. Numer. Anal.*, 43(5):1803–1827 (electronic), 2005.

[41] W.F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Trans. Math. Software*, 15(4):326–347 (1990), 1989.

[42] P. Morin, R.H. Nochetto, and K.G. Siebert. Convergence of adaptive finite element methods.

SIAM Rev., 44(4):631–658 (electronic) (2003), 2002. Revised reprint of "Data oscillation and convergence of adaptive FEM" [SIAM J. Numer. Anal. **38** (2000), no. 2, 466–488 (electronic); MR1770058 (2001g:65157)].

[43] R.H. Nochetto, K.G. Siebert, and A. Vesser. Theory of adaptive finite element methods: an introduction. In *Multiscale, Nonlinear and Adaptive Approximation*. Springer, 2009.

[44] J. T. Oden, I. Babuška, and C. E. Baumann. A discontinuous *hp* finite element method for diffusion problems. *J. Comput. Phys.*, 146(2):491–519, 1998.

[45] I. Perugia and D. Schötzau. An *hp*-analysis of the local discontinuous Galerkin method for diffusion problems. In *Proceedings of the Fifth International Conference on Spectral and High Order Methods (ICOSAHOM-01) (Uppsala)*, volume 17, pages 561–571, 2002.

[46] I. Perugia and D. Schötzau. The *hp*-local discontinuous Galerkin method for low-frequency time-harmonic Maxwell equations. *Math. Comp.*, 72(243):1179–1214 (electronic), 2003.

[47] B. Rivière and M. F. Wheeler. A posteriori error estimates for a discontinuous Galerkin method applied to elliptic problems. Log number: R74. *Comput. Math. Appl.*, 46(1):141–163, 2003. *p*-FEM2000: *p* and *hp* finite element methods—mathematics and engineering practice (St. Louis, MO).

[48] R. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007.

[49] R. Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77(261):227–241 (electronic), 2008.

[50] C. T. Traxler. An algorithm for adaptive mesh refinement in *n* dimensions. *Computing*, 59(2):115–137, 1997.

[51] M.F. Wheeler. An elliptic collocation-finite element method with interior penalties. *SIAM J. Numer. Anal.*, 15(1):152–161, 1978.