CONVERGENCE AND QUASI-OPTIMALITY OF AN ADAPTIVE
FINITE ELEMENT METHOD FOR CONTROLLING $L_2$ ERRORS

ALAN DEMLOW† AND ROB STEVENSON‡

Abstract. In this paper, a contraction property is proved for an adaptive finite element method
for controlling the global $L_2$ error. Furthermore, it is shown that the method converges in $L_2$ with
the best possible rate. The method that is analyzed is the standard adaptive method except that,
if necessary, additional refinements are made to keep the meshes sufficiently mildly graded. This
modification does not compromise the quasi-optimality of the resulting algorithm.

Key words. Adaptive finite element methods; convergence and quasi-optimality of adaptive
finite element methods.

AMS subject classification. 65N15, 65N30

1. Introduction and Results. Consider the elliptic model problem

$$
\begin{array}{ll}
-\Delta u = f \text{ in } \Omega, \\
u = 0 \text{ on } \partial \Omega,
\end{array}
$$

(1.1)

where $\Omega \subset \mathbb{R}^n$, $n \geq 2$, is a convex polyhedral domain and $f \in L_2(\Omega)$. The results
presented here can be generalized to other linear elliptic operators, but we present
the simplest case in order to focus on essential ideas.

In this work we prove contraction and quasi-optimality properties for an adaptive
finite element method (AFEM) for controlling errors in the $L_2(\Omega)$-norm. An AFEM
is an iterative feedback procedure of the form

$$
solve \rightarrow \text{estimate} \rightarrow \text{mark} \rightarrow \text{refine}.
$$

(1.2)

Such adaptive algorithms have for many years been a standard tool for efficiently approximating solutions to partial differential equations such as (1.1). The convergence
properties of AFEM have become the subject of intense theoretical study only in the
past few years, however. We refer to [7], [12], [11], [14], and [13] for an overview of
progress in basic convergence theory for AFEM for linear elliptic problems. Opti-
mal convergence rates were demonstrated in [19] and [4]. A common feature of most
theoretical results published to date concerning convergence and quasi-optimality of
AFEM is that the error estimated in the "estimate" step in (1.2) is the global energy
error. In particular, AFEM optimality results for error notions whose analysis require
nontrivial duality arguments have not to our knowledge appeared in the literature.

In order to describe our results, let $\{T_i\}_{i \geq 0}$ be a nested sequence of conforming,
uniformly shape regular partitions of $\Omega$ produced by the AFEM. In this work, we
consider simplicial partitions as they are produced by the newest vertex bisection
algorithm or by its generalization to more than two dimensions. Let $S_i \subset H^1_0(\Omega)$
be the standard Lagrange finite element space of some fixed degree $k$ on $T_i$, and let
$u_i \in S_i$ satisfy

$$
A(u_i, v_i) := \int_{\Omega} \nabla u_i \cdot \nabla v_i \, dx = \int_{\Omega} f v_i \, dx, \quad (v_i \in S_i).
$$

(1.3)
We also let $|||v||| := \sqrt{A(v,v)}$ and $|||v|||_D := (\int_D |\nabla v|^2 \, \text{d}x)^{1/2}$ denote the global and local energy (semi)norms over $D$, respectively.

Our first goal in this work is to prove that the AFEM for controlling the $L_2$-error $\|u - u_i\|_{L_2(\Omega)}$ is majorized linearly convergent. In order to obtain this result, we require that the sequence $(T_i)$ of partitions is sufficiently mildly graded. We will modify the “refine” routine so that, if necessary, it bisects additional simplices in order to preserve the mild grading of the partition. With this modification, and under the assumption that $\Omega$ is convex, in a sequence of estimates we prove that the quantities $\|u - u_i\|_{L_2(\Omega)}^2 + \text{osc}_i^2$, $|||h_i(u - u_i)|||_2^2 + \text{osc}_i^2$ and $\eta_i^2$ are all uniformly equivalent. Here $\text{osc}_i$ is the $L_2$-oscillation term, $h_i \in W^{1,\infty}(\Omega)$ is a regularized local mesh size function, and $\eta_i$ is the residual based $L_2$ a posteriori error estimator. A key ingredient of many of our proofs is that thanks to the mild grading of the partitions, the mesh size function can be designed so that $|||\nabla h_i|||_{L_\infty(\Omega)}$ is sufficiently small. Following an idea from [4], we then show that a suitably weighted sum of $|||h_i(u - u_i)|||_2^2$ and $\eta_i^2$ is contracted by the AFEM, which implies the convergence result.

The approach of treating the $L_2$-error as a weighted energy error and then employing techniques developed for proving convergence of AFEM for global energy errors was used earlier in [5] to prove convergence of an AFEM for controlling local energy errors. We note that the restriction that $|||\nabla h_i|||_{L_\infty(\Omega)}$ must be sufficiently small has previously appeared in the literature in connection with a priori and a posteriori estimates in $L_\infty$ (cf. [8], [9]) and a priori estimates in $L_2$ (cf. [3]).

Convergence of AFEM for controlling $L_2$-errors was also addressed in [13]. There are several substantial differences between our approach and the one taken in the latter work. [13] obtains convergence of AFEM for controlling weak norms under quite general assumptions on the marking strategy and norm of interest. While the assumptions are general, the convergence result obtained is correspondingly weak in that no estimate of the rate of convergence is obtained. In contrast, we require stronger assumptions; aside from the above-mentioned mesh restriction we also require a Dörfler-type marking strategy. The result is a much stronger convergence theory.

The second main theorem of this work states that for the sequence of partitions produced by the AFEM, $\|u - u_i\|_{L_2(\Omega)} + \text{osc}_i$ can be bounded by some absolute multiple of $(\#T_i)^{-s}$ for the best possible value of $s$. To arrive at this quasi-optimality result, we prove a localized a posteriori upper bound for the $L_2$-difference between Galerkin solutions on nested partitions. Furthermore, we show that, up to an oscillation term, the Galerkin solution is a near best approximation to $u$ from the finite element space in the $L_2$-norm. The latter result is of some interest independent of the current context, since it is known that Ritz projection is not stable in the $L_2$ norm (cf. [1]). Both of these results rely on the condition that $|||\nabla h_i|||_{L_\infty(\Omega)}$ is sufficiently small. Finally, we show that, even if additional bisections are made to preserve the mild grading of the partitions, at any stage in the AFEM the cardinality of the current partition can be bounded by the cardinality of the initial partition plus some absolute multiple of the number of all elements that so far were marked for refinement by the routine “mark”.

We finally comment on the practicality of the AFEM for which we prove optimality. Existing AFEM optimality results for global energy norms require that an essential user-supplied parameter in the “mark” step be sufficiently small. The $L_2$ AFEM for which we prove optimality requires two user-supplied parameters to be sufficiently small, one in the “mark” and the other in the “refine” step. Whereas in the energy case an upper bound for the parameter in terms of interpolation (Poincaré) constants can in principal be derived theoretically, this will be harder in the $L_2$-case.
as the corresponding parameters additionally depend on $H^2$ regularity constants.

The paper is organized as follows. Sections 2 and 3 contain a number of preliminaries and definitions. In Sect. 4 we establish relationships between the error notions used in the paper. Our $L_2$ AFEM is defined precisely in Sect. 5. Sections 6 and 7 contain proofs of convergence and optimality of our AFEM, respectively. Finally, a bisection routine which preserves a mild mesh grading is given in Appendix A along with results concerning existence and properties of mesh functions.

2. Preliminaries.

2.1. Partitions and mesh functions. Let $T_0$ be a conforming partition of $\Omega$ into (essentially) disjoint closed $n$-simplices. By fixing a local numbering of all vertices of all $T \in T_0$, all possible descendants $T$ of $T_0$ that can be created by newest vertex bisection are uniquely determined. Here, with newest vertex bisection, we mean either the refinement procedure as it was developed in two space dimensions, or its generalization to any space dimension. Details can be found in Appendix A.

The simplices in any of those partitions are uniformly shape regular, dependent only on the shape regularity parameters of $T_0$ and the dimension $n$.

Generally, a descendant of $T_0$ is non-conforming. Yet, with a suitable numbering of the vertices in the initial partition, any descendant can be refined to a conforming partition by inflating the total number of simplices by not more than some absolute multiple. Possibly after an initial refinement of the original initial partition, such a numbering always exists. Assuming such a numbering, we denote the set of all conforming descendants $T$ of $T_0$ by $T$. For $T, \tilde{T} \in T$, we will write $T \subset \tilde{T}$ when $\tilde{T}$ is a refinement of $T$ (or is equal to $T$).

For $T \in T \in T_0$, let $h_T = |T|^{1/n}$. Furthermore, let $\omega_T, \tilde{\omega}_T$ be the patches of elements (in $T$) sharing a vertex or a facet (edge, face, ...) with $T$. For $T \in T$, we construct the continuous piecewise linear "mesh-function" $h_T$ by defining, for any vertex $z$ of $T$, $h_T(z)$ as the average of the $h_{T'}$ over all $T' \in T$ with $T' \ni z$. For some constants $c_T$ and $C_T$, it satisfies

$$c_T h_T \leq h_T |T| \leq C_T h_T \quad (T \in T, T \in T). \tag{2.1}$$

By the uniform shape regularity, for another constant $\hat{C}_T$ we have

$$\|\nabla h_T\|_{L^\infty(\Omega)} \leq \hat{C}_T \quad (T \in T).$$

For a number of our results, we will need a mesh function $h_T$ that for some sufficiently small, but fixed constant $\mu > 0$ satisfies both

$$\|\nabla h_T\|_{L^\infty(\Omega)} \leq \mu, \tag{2.2}$$

and (2.1) for some absolute constants $c_T$ and $C_T$ that are independent of $\mu$. By an application of the mean value theorem, the existence of such a mesh function $h_T$ implies that $h_T \leq c_T^{-1}(c_T h_{T'} + \mu \text{dist}(T,T'))$ ($T, T' \in T$), and so in particular that $\max_{T \in T} h_T \leq c_T^{-1}(C_T \min_{T \in T} h_T + \mu \text{diam} \Omega)$. We conclude that the existence of such an $h_T$ imposes a restriction on the grading of the partition $T$ beyond that imposed by shape regularity alone.

In Appendix A, we show that given a parameter $\mu$, there is a class of sufficiently mildly graded partitions $T \in T$ such that mesh functions $h_T$ satisfying (2.2) and (2.1) independent of $\mu$ can be constructed. Given a $\mu > 0$, we denote the class of such partitions by $T_\mu$. The initial partition $T_0$, as any of its uniform refinements, has
no grading and is therefore contained in $T_\mu$ for any $\mu$. Each time that we apply a
mesh function $h_T$ for a $T \in T_\mu$, obviously we mean the mesh function that satisfies
(2.2) and (2.1), independent of $\mu$. These mesh functions are constructed such that,
additionally, they are pointwise non-increasing under refinements inside the class $T_\mu$.
Finally, it is shown that any $T \in T$ can be refined to a partition in $T_\mu$ by inflating
the number of simplices by not more than some fixed multiple, dependent on $\mu$.

The restriction that $\|\nabla h_T\|_{L_\infty(\Omega)}$ is sufficiently small appeared previously in [8],
where a priori $L_\infty$-estimates were proved under this assumption. There $\mu$ was al-
lowed to depend logarithmically on $\min_{T \in T_\mu} h_T$. In the introductory Chapter 0 of the
standard text [3], a priori $L_2$-estimates are also proved under the assumption (2.2)
in the 1-dimensional case; higher-dimensional versions of this result require a slightly
more involved technical development but may be obtained with a similar proof. To
our knowledge, (2.2) is the least restrictive assumption under which optimal a priori
bounds in the $L_2$ norm have appeared in the literature for general space dimension $n$.

Finally, it is pointed out in [8] that the graded meshes typically needed to resolve
corner singularities occurring in elliptic problems on corner domains satisfy (2.2) with
$\mu$ small. Thus although we modify the standard “refine” routine so that it bisects
additional simplices in order to preserve the mild grading of the partition if necessary,
it may be that additional bisections never have to be made.

2.2. Finite element spaces and interpolants. Given $T \in T$, let $S_T \subset H^1_0(\Omega)$
be a space of continuous Lagrange finite element functions which are piecewise poly-
nomials of some fixed degree $k$ on $T$. As $u$ always denotes the solution of the continuous
problem (1.1), the notation $u_T$ will be reserved for the solution of its Galerkin dis-
cretization $A(u_T, v_T) = \int_{\Omega} f v_T dx$ ($v_T \in S_T$). Note that $S_T \subset S_{\tilde{T}}$ whenever $T \subset \tilde{T}$.

We will employ two standard finite element (quasi)-interpolators onto $S_T$. The
first is the Lagrange interpolator, which we denote by $I_{L,T}$. We will also employ
the Scott-Zhang interpolator $I_{SZ,T}$, which is stable in $H^1(\Omega)$, uniformly in $T \subset T$
(cf. [18]). We do not list further properties of $I_{SZ,T}$ here, as its application in the
establishment of residual-type a posteriori upper bounds is now rather standard.

2.3. Regularity. Using that $f \in L_2(\Omega)$ and $\Omega$ is convex, the following $H^2$-
regularity result is well-known.

**Theorem 2.1.** The solution $u$ of (1.1) satisfies $u \in H^2(\Omega)$ with

$$\|u\|_{H^2(\Omega)} \leq C_{\text{reg}} \|f\|_{L_2(\Omega)}.$$ 

2.4. Constants. In what follows, we shall denote by $C$ and $C_i$ ($i = 1, 2, 3...$)
generic constants that may depend on the shape regularity parameters of $T_0$, the space
dimension $n$, the polynomial degree $k$, the constants $c_T$ and $C_T$ in (2.2), and the $H^2$-
regularity constant $C_{\text{reg}}$ from the theorem above. Other constants will be defined as
necessary. In addition, we will often write $a \lesssim b$ instead of $a \leq Cb$ with $C$ as above.
Obviously, with $a \lesssim b$ it is meant that $b \lesssim a$, and with $a \approx b$ that both $a \lesssim b$ and
$a \gtrsim b$.

3. Residual based $L_2$ a posteriori error estimator. Given $T \in T$, $v_T \in S_T$, 
and $T \in T$, we define the $L_2$-type error indicator $\eta(v_T, T)$ by

$$\eta(v_T, T)^2 = h_T^2 \|\nabla v_T\|^2_{L_2(T)} + h_T^2 \|\nabla v_T\|_{L_2(\partial T \setminus \partial \Omega)}^2.$$

(3.1)
We omit \( v_T \) in the above notation when \( v_T = u_T \), that is, \( \eta(T) = \eta(u_T, T) \). Next we define error estimators. If \( M \subset T \), we define

\[
\eta(v_T, M)^2 = \sum_{T \in M} \eta(v_T, T)^2,
\]

If \( v_T = u_T \), we shall omit the reference to \( v_T \) in our notation as above.

The following two results dealing with the reliability and efficiency of the estimator are well-known. Proofs can be found in [22]. The first result makes use of the fact that \( \Omega \) is convex.

**Proposition 3.1.** It holds that

\[
\| u - u_T \|_{L^2(\Omega)} \lesssim \eta(T) \quad (T \in \mathbb{T}).
\]

**Proposition 3.2.** There exists a constant \( C_1 > 0 \) such that

\[
\eta(v_T, T) \leq C_1 [\| u - v_T \|_{L^2(\Omega)} + \text{osc}(T)] \quad (T \in \mathbb{T}, \ v_T \in S_T).
\]

Here \( \text{osc}(M) \ (M \subset T) \), known as the data oscillation term, is defined by

\[
\text{osc}(M)^2 = \sum_{T \in M} h_T^4 \| f - f_T \|_{L^2(T)}^2,
\]

where \( f_T \) is the \( L^2(T) \)-projection of \( f |_T \) onto \( P_{k-1}(T) \).

**Remark 3.3.** Note that \( \text{osc}(M) \leq \eta(v_T, M) \ (M \subset T \in \mathbb{T}, \ v_T \in S_T) \).

### 4. Relations between several error notions

In this section we shall establish several relations between different error notions that will be used in the succeeding sections to prove convergence and quasi-optimality of AFEM in the \( L^2(\Omega) \)-norm.

Instead of directly proving that an AFEM for controlling \( \| u - u_i \|_{L^2(\Omega)} \) is contractive, we shall show that a total error notion based on \( \| h_i(u - u_i) \| \) contracts. The following proposition establishes that this error notion is meaningful so long as the mesh grading is sufficiently mild. It will be used that \( \Omega \) is convex.

**Proposition 4.1.** For sufficiently small \( \mu \), let \( T \in \mathbb{T}_\mu \). Then

\[
\| u - u_T \|_{L^2(\Omega)} \lesssim \| h_T (u - u_T) \|.
\]

**Proof.** We first employ a duality argument. Let \( e = u - u_T \), and let \( z \in H^2(\Omega) \cap H_0^1(\Omega) \) satisfy \( A(v, z) = (v, e) \) for all \( v \in H_0^1(\Omega) \). Employing Galerkin orthogonality, approximation properties, Theorem 2.1, \( h_T |_T \approx h_T \), and \( \| h_T \|_{L^\infty(\Omega)} \leq \mu \), we have for \( \epsilon > 0 \) that

\[
\begin{align*}
(\epsilon, e) &= A(e, z) = A(e, z - ISzT z) \lesssim \| h_T \nabla e \|_{L^2(\Omega)} \| h_T^{-1} \nabla (z - ISzT z) \|_{L^2(\Omega)} \\
&\lesssim \| h_T e \| + \| e \nabla h_T \|_{L^2(\Omega)} \| z \|_{H^2(\Omega)} \lesssim \frac{1}{\epsilon} \| h_T e \|^2 + (\epsilon + \mu) \| e \|^2_{L^2(\Omega)} \quad (4.1)
\end{align*}
\]

Taking \( \epsilon \) and \( \mu \) sufficiently small to kick back the last term completes the proof.

Next we prove that the a posteriori error estimator provides also an upper bound for \( \| h_T (u - u_T) \| \). Again, it is used that \( \Omega \) is convex.

**Proposition 4.2.** For any \( T \in \mathbb{T} \),

\[
\| h_T (u - u_T) \| \lesssim \eta(T).
\]
Proof. Letting $e = u - u_T$, after a short computation and inserting Proposition 3.1 while recalling that $\|\nabla h_T\|_{L_\infty(T)} \lesssim 1$, we find that
\[
A(h_T e, h_T e) = A(e, h_T^2 e) + \| e \nabla h_T \|_{L_2(T)}^2
\leq |A(e, h_T^2 e)| + \| \nabla h_T \|_{L_\infty(T)}^2 \| e \|_{L_2(T)}^2 (4.2)
\lesssim |A(e, h_T^2 e)| + \eta(T)^2.
\]
Using standard techniques for proving residual-type energy estimates along while recalling that $h_T |T \approx h_T$ and $\|\nabla h_T\|_{L_\infty(T)} \lesssim 1$ yields
\[
A(e, h_T^2 e) = A(e, h_T^2 e - ISZ, T(h_T^2 e))
\lesssim \sum_{T \in T}(h_T \| \Delta u_T + f \|_{L_2(T)} + h_T^{1/2} \| \nabla u_T \|_{L_2(\partial T)} \| \nabla (h_T^2 e) \|_{L_2(\omega_T)}
\lesssim \sum_{T \in T}(\eta(T)(\| \nabla (h_T e) \|_{L_2(\omega_T)} + \| \nabla h_T \|_{L_\infty(\omega_T)} \| e \|_{L_2(\omega_T)}
\lesssim \eta(T)(\| h_T e \| + \| e \|_{L_2(T)}).
\]
Inserting the result of Proposition 3.1 into (4.3), inserting the result into (4.2), and kicking back the term $\| h_T e \|$ above completes the proof. \( \square \)

Next we prove a stability result for the error indicator, cf. Proposition 3.3. of [4]. Lemma 4.3. It holds that
\[
|\eta(v, T) - \eta(w, T)| \lesssim \| h_T (v - w) \|_{L_2(\tilde{T})} + \| h_T \|_{L_\infty(\tilde{T})} \| v - w \|_{L_2(\tilde{T})}
\]
($T \in T \in T, v, w \in S_T$).

Proof. Recalling the definition (3.1), we use the triangle inequality to compute for $T \in T$ that
\[
\eta(v, T) \leq \eta(w, T) + (h_T^2 \| \Delta (v - w) \|_{L_2(T)} + h_T^3 \| \nabla (v - w) \|_{L_2(T)}))^{1/2}. (4.4)
\]
Employing an inverse inequality and $h_T |T \approx h_T$, we compute
\[
h_T^2 \| \Delta (v - w) \|_{L_2(T)} \lesssim h_T \| \nabla (v - w) \|_{L_2(T)}
\lesssim \| \nabla h_T (v - w) \|_{L_2(T)} + \| (v - w) \nabla h_T \|_{L_2(T)} (4.5)
\]
Employing the trace inequality $\| \nabla v \|_{L_2(\partial T)} \lesssim h_T^{1/2} \| \nabla v \|_{L_2(T)} + h_T^{1/2} \| v \|_{H^2(T)}$, we compute as above to obtain for the edge $e = T \cap T'$
\[
h_T^{3/2} \| \nabla (v - w) \|_{L_2(e)} \lesssim h_T^{3/2} (\| \nabla (v - w) \|_{L_2(e)} + \| \nabla (v - w) \|_{L_2(e)} + h_T^{1/2} \| v - w \|_{H^2(T)}
\lesssim h_T^{3/2} (h_T^{1/2} \| \nabla (v - w) \|_{L_2(T)} + h_T^{1/2} \| v - w \|_{H^2(T)})
\lesssim \| h_T (v - w) \|_{L_2(T \cup T')} + \| (v - w) \nabla h_T \|_{L_2(T \cup T')} (4.6)
\]
Above we have employed the convention that for $x$ lying in the edge $e = T \cap T'$, $\nabla (v - w)_T(x) = \lim_{y \in T, y \rightarrow x} \nabla (v - w)(y)$. Noting that $\| (v - w) \nabla h_T \|_{L_2(T \cup T')} \lesssim \| \nabla h_T \|_{L_\infty(\tilde{T})} \| v - w \|_{L_2(\tilde{T})}$, summing over the edges of $T$, and collecting (4.6) and (4.5) into (4.4) completes the proof. \( \square \)
The following localized upper bound will be used for proving quasi-optimality of the AFEM. It will be used that Ω is convex.

**Lemma 4.4.** There exists a constant $C_2 > 0$ such that for sufficiently small $μ$, $T ∈ T_μ$, and $T ⊂ T ∈ T$, \[ \| u_{\tilde{T}} - u_T \|_{L^2(Ω)} \leq C_2 \eta(\mathcal{R}_{T-\tilde{T}}). \] Here $\mathcal{R}_{T-\tilde{T}} ∈ T$ is the subset of elements that are refined in passing from $T$ to $\tilde{T}$.

**Proof.** Set $E = u_{\tilde{T}} - u_T$, so that $A(E, \chi) = 0$ for all $\chi ∈ S_T$. Arguing as in Lemma 4.1, it is easy to show that \[ \| E \|_{L^2(Ω)} \lesssim \| h_T \nabla E \|_{L^2(Ω)}. \] (4.7)

An elementary computation yields for $ε > 0$ \[ \| h_T \nabla E \|_{L^2(Ω)}^2 = (h_T \nabla E, h_T \nabla E) = A(E, h_T^2 E) - 2(h_T \nabla E, E \nabla h_T) \leq |A(E, h_T^2 E)| + ε \| h_T \nabla E \|_{L^2(Ω)}^2 + \frac{μ^2}{ε} \| E \|_{L^2(Ω)}^2. \] (4.8)

Taking $ε$ small enough to kick back the second term above and inserting the result into (4.7) yields \[ \| E \|_{L^2(Ω)}^2 \lesssim |A(E, h_T^2 E)| + μ^2 \| E \|_{L^2(Ω)}, \] so that for $μ$ sufficiently small we have \[ \| E \|_{L^2(Ω)}^2 \lesssim \| h_T \nabla E \|_{L^2(Ω)}^2 \lesssim |A(E, h_T^2 E)|. \] (4.10)

Let $I_{L, \tilde{T}} : C_0(Ω) → S_{\tilde{T}}$ and $I_{SZ, T}$ be as in §2.2. It is shown in [4, Lemma 3.6] that $I_{SZ, T}$ may be defined so that for $T ∈ T ∩ \tilde{T}$ (i.e., for $T ∈ T \setminus \mathcal{R}_{T-\tilde{T}}$), $(χ - I_{SZ, T} χ) T = 0$ for $χ ∈ S_{\tilde{T}}$. We assume that $I_{SZ, T}$ is so defined. Then

\[
A(E, h_T^2 E) = A(E, h_T^2 E - I_{L, \tilde{T}}(h_T^2 E)) + A(E, I_{L, \tilde{T}}(h_T^2 E) - I_{SZ, T}(I_{L, \tilde{T}}(h_T^2 E))).
\]

(4.11)

In order to bound the term “I” in (4.11), in essence we employ the classical superapproximation tool introduced in [15] and also use some modifications introduced in [6]. Restricted to $T ∈ T$, we have $h_T ∈ P_1$ and $E ∈ P_k$, and so $|D^2 h_T^2| \lesssim |\nabla h_T|^2$, $D^j h_T = 0$ for $j ≥ 3$, and $D^{k+1} E = 0$. Employing inverse inequalities and using that $h_T |_{T'} ≈ h_{T'}$ ($T' ∈ T$) and $\| \nabla h_T \|_{L^\infty(Ω)} \lesssim 1$, we thus compute

\[
\| \nabla (h_T^2 E - I_{L, \tilde{T}}(h_T^2 E)) \|_{L^2(T)} \lesssim h_T^k \| D^{k+1}(h_T^2 E) \|_{L^2(T)} \\
\lesssim h_T^k \| \nabla h_T^2 \|_{L^\infty(T)} \| D^k E \|_{L^2(T)} + \| D^2 h_T^2 \|_{L^\infty(T)} \| D^{k-1} E \|_{L^2(T)} \\
\lesssim h_T^k \| \nabla h_T \|_{L^\infty(T)} \| \nabla h_T \|_{L^\infty(T)} \| D^k E \|_{L^2(T)} + \| \nabla h_T \|_{L^\infty(T)} \| D^{k-1} E \|_{L^2(T)} \\
\lesssim \| \nabla h_T \|_{L^\infty(T)} \| \nabla h_T \|_{L^\infty(T)} \| E \|_{L^2(T)} \\
\lesssim \| h_T \|_{L^\infty(T)} \| \nabla h_T \|_{L^\infty(T)} \| E \|_{L^2(T)}.
\]

(4.12)
Using that \( \| \nabla h_T \|_{L_\infty(\Omega)} \leq \mu \), for any \( \epsilon > 0 \) it thus holds that

\[
|\mathcal{I}^*| \lesssim \sum_{T \in \mathcal{T}} \| h_T \nabla E \|_{L_2(T)} \| E \|_{L_2(T)} \lesssim \| h_T \nabla E \|_{L_2(\Omega)} \| E \|_{L_2(\Omega)}
\]

where

\[
\| h_T \nabla E \|_{L_2(\Omega)}^2 \lesssim \epsilon \| h_T \nabla E \|_{L_2(\Omega)}^2 + \frac{\mu^2}{4\epsilon} \| E \|_{L_2(\Omega)}^2.
\]

To bound the term \( \mathcal{I}^{**} \), we first note that since \( h_T^2 E \) is piecewise polynomial on \( \mathcal{T} \), the stability bound \( \| \nabla I_{L,T} (h_T^2 E) \|_{L_2(T)} \lesssim \| \nabla (h_T^2 E) \|_{L_2(T)} \) holds. Recall that the function \( I_{L,T} (h_T^2 E) - IS_Z,T (I_{L,T} (h_T^2 E)) \) is in \( S_\mathcal{T} \) and vanishes on all \( T \in \mathcal{T} \setminus R_{\mathcal{T} \rightarrow \mathcal{T}} \). We may then compute using Galerkin orthogonality and standard residual techniques that for \( \epsilon > 0 \),

\[
|\mathcal{I}^{**}| = |A(u - u_T, I_{L,T} (h_T^2 E) - IS_Z,T (I_{L,T} (h_T^2 E)))| \\
\lesssim \sum_{T \in R_{\mathcal{T} \rightarrow \mathcal{T}}} (h_T \| f + \Delta u_T \|_{L_2(T)} + h_T^{1/2} \| \nabla u_T \|_{L_2(\partial T)} \| \nabla I_{L,T} (h_T^2 E) \|_{L_2(\omega_T)}) \\
\lesssim \sum_{T \in R_{\mathcal{T} \rightarrow \mathcal{T}}} (h_T \| f + \Delta u_T \|_{L_2(T)} + h_T^{1/2} \| \nabla u_T \|_{L_2(\partial T)} \| \nabla (h_T^2 E) \|_{L_2(\omega_T)}) \\
\lesssim \sum_{T \in R_{\mathcal{T} \rightarrow \mathcal{T}}} \eta(T) (h_T \| \nabla E \|_{L_2(\omega_T)} + \mu \| E \|_{L_2(\omega_T)}) \\
\lesssim (1 + \frac{1}{\epsilon}) \eta(R_{\mathcal{T} \rightarrow \mathcal{T}})^2 + \epsilon \| h_T \nabla E \|_{L_2(\Omega)}^2 + \mu^2 \| E \|_{L_2(\Omega)}^2.
\]

Inserting (4.14) and (4.15) into (4.11) yields for \( \epsilon > 0 \),

\[
\| h_T \nabla E \|_{L_2(\Omega)}^2 \lesssim (1 + \frac{1}{\epsilon}) \eta(R_{\mathcal{T} \rightarrow \mathcal{T}})^2 + \epsilon \| h_T \nabla E \|_{L_2(\Omega)}^2 + \mu^2 \| E \|_{L_2(\Omega)}^2.
\]

Taking \( \epsilon \) small enough to kick back the second term above and then inserting the result into (4.10) yields

\[
\| E \|_{L_2(\Omega)}^2 \lesssim h_T \nabla E \|_{L_2(\Omega)}^2 \lesssim \eta(R_{\mathcal{T} \rightarrow \mathcal{T}})^2 + \mu^2 \| E \|_{L_2(\Omega)}^2.
\]

Taking \( \mu \) small enough to kick back the last term above completes the proof of the lemma. \( \Box \)

We next prove a \emph{quasi-orthogonality} result (cf. [11] for a similar estimate in the context of convergence of AFEM in the global energy norm for general second-order linear elliptic problems).

**Lemma 4.5.** For any \( \epsilon > 0 \), \( T \in \mathcal{T} \), and \( v_T \in S_\mathcal{T} \), it holds that

\[
\| h_T (u - u_T) \|_{L_\infty(\Omega)}^2 + \| h_T (u_T - v_T) \|_{L_\infty(\Omega)}^2 \lesssim (1 + \epsilon) \| h_T (u - v_T) \|_{L_\infty(\Omega)}^2 \\
\lesssim \epsilon^{-1} \| \nabla h_T \|_{L_\infty(\Omega)}^2 (\| u - u_T \|_{L_2(\Omega)}^2 + \| u - v_T \|_{L_2(\Omega)}^2).
\]

**Proof.** Writing \( \tilde{e}_T = u - v_T \) and \( e_T = u - u_T \), we calculate

\[
\| h_T e_T \|_{L_\infty(\Omega)}^2 = \| h_T \tilde{e}_T \|_{L_\infty(\Omega)}^2 - \| h_T (u_T - v_T) \|_{L_\infty(\Omega)}^2 - 2 A(h_T \tilde{e}_T, h_T (u_T - v_T))
\]
An elementary calculation yields
\[
|A(h_T e_T, h_T (u_T - v_T))| = |A(e_T, h_T^2 (u_T - v_T)) \\
+ (|∇ h_T|^2 e_T, u_T - v_T) - (e_T ∇ h_T, ∇ h_T e_T) \\
+ (e_T ∇ h_T, ∇ (h_T e_T))| \\
\leq \|e_T ∇ h_T\|_{L_2(Ω)} \|h_T e_T\| + \|e_T ∇ h_T\|_{L_2(Ω)} \|h_T e_T\| \\
+ \|e_T ∇ h_T\|_{L_2(Ω)} \|(u_T - v_T) ∇ h_T\|_{L_2(Ω)} \\
+ |A(e_T, h_T^2 (u_T - v_T))|.
\] (4.20)

Arguments analogous to those in (4.12) and (4.13) yield
\[
|A(e_T, h_T^2 (u_T - v_T))| = |A(e_T, (I - I_{L,T}) h_T^2 (u_T - v_T))| \\
\lesssim \|h_T ∇ e_T\|_{L_2(Ω)} \|∇ h_T\|_{L_2(Ω)} \|u_T - v_T\|_{L_2(Ω)} \\
\leq \left|\|h_T e_T\| + ∇ h_T\|_{L_2(Ω)} \|e_T\|_{L_2(Ω)} \|h_T\| \|u_T - v_T\|_{L_2(Ω)}\right|
\] (4.21)

Inserting this bound into (4.20) and applying Young’s inequality a few times yields for some constant \(C > 0\) and any \(δ > 0\) that
\[
|A(h_T e_T, h_T (u_T - v_T))| \leq 2 \left(\|h_T e_T\|^2 + \|h_T e_T\|^2 \right) \\
+ C(1 + \frac{1}{δ}) \|∇ h_T\|_{L_2(Ω)}^2 \|e_T\|^2_{L_2(Ω)} + \|e_T\|^2_{L_2(Ω)}.
\] (4.22)

Dividing through by \(1 - δ\) and selecting \(\frac{1 + δ}{1 - δ}\) as \(1 + ε\) completes the proof of Lemma 4.5.

A combination of the Propositions 4.2 and 3.2 shows that
\[
\|h_T (u - u_T)\| \lesssim \|u - u_T\|_{L_2(Ω)} + osc(T).
\]

In the following lemma it is shown that this inequality is even valid with \(u_T\) reading as any \(v_T \in S_T\).

**Lemma 4.6.** For \(T \in T\), it holds that
\[
\|h_T (u - v_T)\| \lesssim \|u - v_T\|_{L_2(Ω)} + osc(T) \quad (v_T \in S_T).
\]

**Proof.** Writing \(z_T = u - v_T\), from \(\|h_T\|_{L_2(Ω)} \lesssim 1\) we have for \(ε > 0\)
\[
\|h_T z_T\|^2 = A(h_T z_T, h_T z_T) \\
= (h_T ∇ z_T, ∇ (h_T z_T)) + (z_T ∇ h_T, ∇ (h_T z_T)) \\
\leq \frac{1}{2ε} \left(\|h_T ∇ z_T\|^2_{L_2(Ω)} + \|z_T\|^2_{L_2(Ω)} + ε\|h_T z_T\|^2\right)
\] (4.23)
or
\[
\|h_T z_T\|^2 \lesssim \|h_T ∇ z_T\|^2_{L_2(Ω)} + \|z_T\|^2_{L_2(Ω)}. 
\] (4.24)
Integrating by parts, we next compute
\[
\|h_T \nabla z_T\|_{L^2(\Omega)}^2 = (\nabla z_T, h_T^2 \nabla z_T)
\]
\[
= \sum_{T \in \mathcal{T}} \int_T -z_T \nabla \cdot (h_T^2 \nabla z_T) \, dx + \int_{\partial T} z_T h_T^2 \nabla z_T \cdot \vec{n} \, d\sigma
\]
\[
\leq \sum_{T \in \mathcal{T}} \|z_T\|_{L^2(T)} \|h_T\|_{L^\infty(T)} \|f + \Delta v_T\|_{L^2(T)} + 2 \|\nabla h_T\|_{L^\infty(T)} \|h_T \nabla z_T\|_{L^2(T)}
\]
\[
+ \|h_T^2 z_T\|_{L^2(\partial T \setminus \partial \Omega)} \|\nabla v_T\|_{L^2(\partial T \setminus \partial \Omega)}.
\]
(4.25)

Inserting the scaled trace inequality \(\|v\|_{L^2(\partial T)} \lesssim h_T^{-1/2}\|v\|_{L^2(T)} + h_T^{1/2}\|v\|_T\) into (4.25) yields for \(\epsilon > 0\)
\[
\|h_T^2 z_T\|_{L^2(\partial T)} \|\nabla v_T\|_{L^2(\partial T)}
\]
\[
\lesssim (h_T^{-1/2}\|h_T z_T\|_{L^2(T)} + h_T^{3/2}\|h_T z_T\|_T)\|\nabla v_T\|_{L^2(\partial T)}
\]
\[
\lesssim \|z\|_{L^2(T)} + \|h_T \nabla z_T\|_{L^2(T)} + \|\nabla h_T\|_{L^\infty(T)} \|z_T\|_{L^2(T)} h_T^{3/2}\|\nabla v_T\|_{L^2(\partial T)}
\]
\[
\lesssim \|z\|_{L^2(T)} + (1 + \frac{1}{\epsilon}) h_T^{3/2}\|\nabla v_T\|_{L^2(\partial T)} + \epsilon \|h_T \nabla z_T\|_{L^2(T)}^2.
\]
(4.26)

Inserting (4.26) into (4.25) and then employing Proposition 3.2 and \(h_T |T| \approx h_T\) yields for \(\epsilon > 0\)
\[
\|h_T \nabla z_T\|_{L^2(\Omega)}^2
\]
\[
\lesssim \sum_{T \in \mathcal{T}} (1 + \frac{1}{\epsilon}) (\|z_T\|_{L^2(T)}^2 + h_T^{3/2}\|\nabla v_T\|_{L^2(\partial T \setminus \partial \Omega)}^2) + h_T^{3/2}\|f + \Delta v_T\|_{L^2(T)}^2
\]
\[
+ \epsilon \|h_T \nabla z_T\|_{L^2(T)}^2
\]
\[
\lesssim (1 + \frac{1}{\epsilon}) (\|z_T\|_{L^2(\Omega)}^2 + \eta(v_T, T)^2) + \epsilon \|h_T \nabla z_T\|_{L^2(T)}^2
\]
\[
\lesssim (1 + \frac{1}{\epsilon}) (\|z_T\|_{L^2(\Omega)}^2 + \text{osc}(T)^2) + \epsilon \|h_T \nabla z_T\|_{L^2(T)}^2.
\]
(4.27)

Taking \(\epsilon\) small enough to kick back the last term above and inserting the result into (4.24) completes the proof. \(\square\)

We next prove that if \(T_i\) is sufficiently mildly graded, then \(\|u - u_i\|_{L^2(\Omega)}\) is bounded up to a constant by the best approximation to \(u\) lying in \(S_i\), as measured in \(L_2\), plus a data oscillation term.

**Corollary 4.7.** There exists a constant \(C_3 > 0\) such that for sufficiently small \(\mu\) and \(T \in \mathcal{T}_\mu\), it holds that
\[
\|u - u_T\|_{L^2(\Omega)} \leq C_3 \left( \inf_{v_T \in S_T} \|u - v_T\|_{L^2(\Omega)} + \text{osc}(T) \right).
\]
(4.28)

**Proof.** By using that \(\|\nabla h_T\|_{L^\infty(\Omega)} \leq \mu\), the application of Proposition 4.1, Lemma 4.5 for some fixed \(\epsilon\), and Lemma 4.6 shows that for any \(v_T \in S_T\)
\[
\|u - u_T\|_{L^2(\Omega)}^2 \lesssim \|h_T(u - u_T)\|_2^2
\]
\[
\lesssim \|h_T(u - v_T)\|_2^2 + \mu^2 (\|u - u_T\|_{L^2(\Omega)}^2 + \|u - v_T\|_{L^2(\Omega)}^2)
\]
\[
\lesssim (1 + \mu^2) \|u - v_T\|_{L^2(\Omega)}^2 + \text{osc}(T)^2 + \mu^2 \|u - u_T\|_{L^2(\Omega)}^2,
\]
By taking \( \mu \) sufficiently small, the proof is completed. \( \square \)

Remark 4.8. Corollary 4.7 is of some interest independent of the current context because the Galerkin approximation is not stable in \( L_2 \) even on quasi-uniform meshes, that is, (4.28) does not hold if the data oscillation term is removed. A simple onedimensional counterexample can be found in [1]. The counterexample to \( L_2 \)-stability given in [1] also shows that \( \| u - u_T \|_{L_2(\Omega)} \) does not bound the right hand side of (4.28) up to a constant. It does however trivially follow from Corollary 4.7 that \( \| u - u_T \|_{L_2(\Omega)} \) + osc(\( T \)) \( \approx \inf_{v_T \in S_T} \| u - v_T \|_{L_2(\Omega)} + \text{osc}(T) \) uniformly in all \( T \in T_\mu \) with \( \mu \) sufficiently small. That is, on sufficiently mildly graded meshes an analog to Céa’s Lemma holds for the total \( L_2 \) error \( \| u - u_T \|_{L_2(\Omega)} + \text{osc}(T) \).

5. Adaptive FEM. In this section we give details of our adaptive FEM. In particular, we give precise definitions of each module solve, estimate, mark, and refine of the generic adaptive iteration (1.2), with the goal of constructing an AFEM that produces by refinement a sequence of partitions \( (T_i)_{i \geq 0} \subset T_\mu \) for reducing the weighted energy error \( \| h_i (u - u_i) \|_{L_2(\Omega)} \|, \) and thus also the \( L_2 \)-error \( \| u - u_i \|_{L_2(\Omega)} \). Here, and in the following, \( h_i, u_i, S_i, \text{osc}_i, \eta_i \), and, for \( M \subset T_i \) \( \eta_i (M) \) denote \( h_{T_i}, u_{T_i}, S_{T_i}, \text{osc}(T_i), \eta(u_{T_i}, T_i), \) and \( \eta(u_{T_i}, M) \) respectively. The constant \( \mu \) will be assumed to be sufficiently small.

1. Module solve. Given the current partition \( T_i \), solve (1.3) for \( u_i \in S_i \). We assume that the finite element system is assembled and solved exactly.

2. Module estimate. In principle, the adaptive algorithm is terminated when \( C \eta(T_i) \leq \text{tol} \) for some prescribed tolerance \( \text{tol} \) and user-defined constant \( C \).

3. Module mark. We employ a Dörfler marking (cf. [7]). More precisely, we fix a parameter \( 0 < \theta < 1 \), and at each step of the algorithm choose the smallest set \( M_i \subset T_i \) so that

\[
\eta(M_i) \geq \theta \eta_i.
\] (5.1)

4. Module refine. Our results below assume that each marked element \( T_i \in M_i \) is bisected \( b \geq 1 \) times in passing from \( T_i \) to \( T_{i+1} \) and that generally additional elements are refined in the process in order to ensure that \( T_{i+1} \) is conforming and sufficiently mildly graded in the sense that it is in \( T_\mu \). In Appendix A, we show that, assuming the value of \( \mu \) is known, the standard newest vertex bisection algorithm or its generalization to more than two dimensions, as it has been implemented in the finite element toolbox ALBERTA ([17]), can be modified so that the sequence of adaptive meshes is indeed in \( T_\mu \). As we will see in Sect. 7, this modification does not compromise quasi-optimality of the resulting AFEM.

6. Convergence of the AFEM in \( L_2 \). In this section we prove a quasi-orthogonality property, an estimator reduction inequality, and finally an error contraction property. We employ the techniques of [4], which do not rely upon local a posteriori lower bounds as do previous proofs of convergence of adaptive FEM.

Proposition 6.1. For any \( \epsilon > 0 \), we have that

\[
\| h_{i+1} (u - u_{i+1}) \|^2 + \| h_{i+1} (u_{i+1} - u_i) \|^2 - (1 + \epsilon) \| h_i (u - u_i) \|^2 \\
\leq \epsilon^{-1} \mu^2 \| u - u_{i+1} \|_{L_2(\Omega)}^2 + \| u - u_i \|_{L_2(\Omega)}^2.
\] (6.1)
Proof. Writing \( e_i = u - u_i \), and using that \( T_i, T_{i+1} \in \mathcal{T}_\mu \) and \( h_{i+1} \leq h_i \), we have
\[
|||h_{i+1}e_i|||^2 \leq |||h_i\nabla e_i|||_{L^2(\Omega)} + |||\nabla h_{i+1}|||_{L^\infty(\Omega)} |||e_i|||_{L^2(\Omega)} \\
\leq ||h_i\nabla e_i|||_{L^2(\Omega)} + \mu ||e_i|||_{L^2(\Omega)} \leq ||h_i e_i|||^2 + 2\mu ||e_i|||_{L^2(\Omega)}.
\]
and so, for any \( \epsilon > 0 \),
\[
|||h_{i+1}e_i|||^2 \leq (1 + \epsilon)|||h_i e_i|||^2 + (1 + \frac{1}{\epsilon})4\mu^2 ||e_i|||^2_{L^2(\Omega)}.
\]
By substituting this result into the estimate of Lemma 4.5 with \( T = T_{i+1} \), and thus \( u_T = u_{i+1} \), and \( v_T = u_i \), we find (6.1) with \( 1 + \epsilon \) reading as \( (1 + \epsilon)^2 \), which is an equivalent statement. \[ \square \]

Next we establish an estimator reduction result. Our proof closely follows Corollary 3.4 of [4].

**Proposition 6.2.** With \( \lambda := 1 - 2^{-\frac{2b}{\theta}} \), for any \( \delta \in (0,1) \) we have
\[
\eta^2_{i+1} - (1 + \delta)(1 - \lambda \delta^2)\eta^2_i \leq \delta^{-1}|||h_{i+1}(u_{i+1} - u_i)|||^2 + \mu^2 ||u_{i+1} - u_i|||_{L^2(\Omega)}^2.
\]

**Proof.** We first apply Lemma 4.3 to \( u_i, u_{i+1} \in S_{i+1} \), square the result, and apply Young’s inequality to the resulting mixed terms to obtain for any \( \delta \in (0,1) \) and \( T \in T_{i+1} \)
\[
\eta(u_{i+1}, T)^2 - (1 + \delta)\eta(u_i, T)^2 \\
\leq \delta^{-1}|||h_{i+1}(u_{i+1} - u_i)|||^2_{L^2(\Omega)} + ||\nabla h_T|||_{L^\infty(\widetilde{\Omega})} ||u_{i+1} - u_i|||_{L^2(\widetilde{\Omega})}.
\]
Summing over \( T \in T_{i+1} \) and using the fact that no element is contained in more than \( n + 1 \) patches \( \widetilde{\omega}_T \), we have
\[
\eta^2_{i+1} - (1 + \delta)\eta(u_i, T_{i+1})^2 \leq \delta^{-1}|||h_{i+1}(u_{i+1} - u_i)|||^2_{L^2(\Omega)} + \mu^2 ||u_{i+1} - u_i|||_{L^2(\Omega)}^2. \tag{6.2}
\]

For \( T' \in T_i \), let \( T'_T = \{ T \in T_{i+1} : T \subset T' \} \). Note that for a marked element \( T' \in M_i \) and \( T \in T'_T \), \( h_T \leq 2^{-\frac{b}{\theta}} h_{T'} \) and \( ||\nabla u_i|||_{L^2(\Omega)} = 0 \) across interfaces of \( T \) lying in the interior of \( T' \), and so
\[
\eta(u_i, T'_T)^2 \leq 2^{-\frac{b}{\theta}} \eta(u_i, T')^2. \tag{6.3}
\]
Combining (6.3) with the trivially proved monotonicity property \( \eta(u_i, T_{i+1}) \leq \eta(u_i, T') \) for \( T' \in T_i \setminus M_i \), and summing over \( T \in T_{i+1} \) yields
\[
\eta(u_i, T_{i+1})^2 \leq \eta(u_{i+1}, T_i \setminus M_i)^2 + 2^{-\frac{b}{\theta}} \eta(u_i, M_i)^2 \leq \eta(u_i, T_i)^2 - \lambda \eta(u_i, M_i)^2 \leq (1 - \lambda \theta^2) \eta^2_i \tag{6.4}
\]
by (5.1). Combining (6.4) and (6.2) completes the proof. \[ \square \]

We finally establish a contraction property for the quantity \( |||h_i(u - u_i)|||^2 + \gamma \eta_i(T_i)^2 \) for some properly chosen constant \( \gamma > 0 \).

**Theorem 6.3.** There exist constants \( \gamma > 0 \) and \( \alpha \in (0,1) \) depending on a generic constant of the type defined in §2.4, the parameter \( \theta \) in (5.1), and the number of times \( b \) that each element in \( M_i \) is bisected such that for sufficiently small \( \mu \),
\[
|||h_{i+1}(u - u_{i+1})|||^2 + \gamma \eta^2_{i+1} \leq \alpha^2 (|||h_i(u - u_i)|||^2 + \gamma \eta^2_i). \tag{6.5}
\]
Proof. We will use the abbreviations $e_j = |||h_j(u - u_j)|||$ ($j \in \{i, i + 1\}$), and $E_i = |||h_{i+1}(u_i - u_{i+1})|||$.

From Propositions 4.2, 6.1, 3.1 and 6.2, we know that there exists constants $C_4, C_5, C_6 > 0$ such that for all $\epsilon, \delta > 0$,

$$e_i \leq C_4 \eta_i,$$  \hspace{1cm} (6.6)

$$e_{i+1}^2 \leq (1 + \epsilon) e_i^2 - E_i^2 + C_5 \epsilon^{-1} \mu^2 (\eta_i^2 + \eta_{i+1}^2),$$ \hspace{1cm} (6.7)

$$\eta_{i+1}^2 \leq (1 + \delta)(1 - \lambda \theta^2)\eta_i^2 + C_6 \delta^{-1} [E_i + \mu^2 (\eta_i^2 + \eta_{i+1}^2)]$$

$$\leq (1 + \delta)[(1 - \frac{2}{2C_4^2} \lambda \theta^2)\eta_i^2 - \frac{1}{2C_4^2} \lambda \theta^2 e_i^2] + C_6 \delta^{-1} [E_i + \mu^2 (\eta_i^2 + \eta_{i+1}^2)]$$ \hspace{1cm} (6.8)

where to arrive at (6.8) we used already (6.6).

Multiplying (6.8) with $\tilde{\gamma} = \tilde{\gamma}(\delta) = \delta C_6^{-1}$ and adding the result to (6.7) yields

$$e_{i+1}^2 + \tilde{\gamma} \eta_{i+1}^2 \leq (1 + \delta)(1 - \frac{2}{2C_4^2} \lambda \theta^2)\eta_i^2 + [(1 + \epsilon) - (1 + \delta)] \frac{\tilde{\gamma}}{2C_4^2} \lambda \theta^2 e_i^2 + q \mu^2 (\eta_i^2 + \eta_{i+1}^2),$$ \hspace{1cm} (6.9)

where $q = q(\epsilon) = C_5 \epsilon^{-1} + 1$. Now by fixing a sufficiently small $\delta$ and, subsequently, a sufficiently small $\epsilon$ such that

$$\tilde{\alpha}^2 := \max\{(1 + \delta)(1 - \frac{2}{2C_4^2} \lambda \theta^2), (1 + \epsilon) - (1 + \delta) \frac{\tilde{\gamma}}{2C_4^2} \lambda \theta^2\} < 1,$$

(6.9) implies that

$$e_{i+1}^2 + (\tilde{\gamma} - q \mu^2) \eta_{i+1}^2 \leq \tilde{\alpha}^2 e_i^2 + (\tilde{\alpha}^2 \tilde{\gamma} + q \mu^2) \eta_i^2.$$

Now by choosing $\mu$ sufficiently small such that $0 < \frac{\tilde{\alpha}^2 \tilde{\gamma} + q \mu^2}{\tilde{\gamma} - q \mu^2} \leq \tilde{\alpha}^2 := \frac{1 + \tilde{\alpha}^2}{2},$ the proof is completed for that $\alpha$ and $\gamma := \tilde{\gamma} - q \mu^2$. □

We finally show that $\|u - u_i\|_{L_2(\Omega)} \to 0$ with linear rate as $i \to \infty$.

**Corollary 6.4.** Assume that $\mu$ is sufficiently small as in Theorem 6.3. Then with $\alpha$ from that theorem, it holds that for $i \geq j$

$$\|u - u_i\|_{L_2(\Omega)} + \text{osc}_i \lesssim \alpha^{i-j}\|u - u_j\|_{L_2(\Omega)} + \text{osc}_j.$$

**Proof.** The proof follows from Theorem 6.3 and

$$\|u - u_i\|_{L_2(\Omega)} + \text{osc}_i \approx |||h_i(u - u)|||^2 + \gamma \eta_i^2,$$

the latter being a consequence of Propositions 3.1, 3.2 and 4.2, and Remark 3.3. □

**7. Quasi-optimality of AFEM in $L_2$.** For $s > 0$, we define the approximation class

$$\mathcal{A}^s = \{v \in H^1_0(\Omega) : -\Delta v \in L_2(\Omega),\}

|v|_{\mathcal{A}^s} := \sup_{N \in \mathbb{N}} \inf_{T \in \mathcal{T}_T} \inf_{\#T_0 \leq N} \left[ \inf_{\#T \leq N} \inf_{v_T \in S_T} \|u - v_T\|_{L_2(\Omega)} + \text{osc}(T) \right] < \infty.$$

Thus $\mathcal{A}^s$ contains all $v \in H^1_0(\Omega)$ with $-\Delta v \in L_2(\Omega)$, that, for some sequence $(T_i)_i$, can be approximated at rate $s$ in the $L_2$-norm by a sequence of functions from $(S_{T_i})_i$, and
for which \(-\Delta v\) can be approximated by a sequence of functions from \((\prod_{T \in T_i} \mathbb{P}_{k-1}(T))_i\) in the \(T_i\)-dependent weighted \(L_2\)-norm \(\sqrt{\sum_{T \in T_i} h_T^2 \|g\|^2_{L_2(T)}}\).

Note that in our case of \(\Omega\) being convex, \(v \in H^1_0(\Omega)\) with \(-\Delta v \in L^2(\Omega)\) is equivalent to \(v \in H^1_0(\Omega) \cap H^2(\Omega)\). Standard estimates show that for \(s \in [\frac{2}{n}, \frac{k+1}{n}]\), we have \(H^1_0 \cap H^m(\Omega) \subset \mathcal{A}^r\), where for \(v \in H^1_0 \cap H^m(\Omega)\) the rate \(s\) is already realized with uniform refinements. The class \(\mathcal{A}^r\), however, is much larger than \(H^1_0 \cap H^m(\Omega)\), which is the reason an AFEM is employed in the first place. What is in essence needed for \(v \in H^1_0 \cap H^2(\Omega)\) to be in \(\mathcal{A}^r\) is that its \(sn\)-th order partial derivatives are bounded in \(L_p(\Omega)\) for some \(p > (\frac{1}{2} + s)^{-1}\).

In this section, we are going to show that if, for whatever \(s > 0\), the solution \(u\) of (1.1) is in \(\mathcal{A}^r\), then for the sequence \((\mathcal{T}_0)_i \subset \mathcal{T}_n\) and corresponding Galerkin solutions \((u_i)_i\), produced by our AFEM, it holds that \(\|u - u_i\|_{L^2(\Omega)} + \text{osc}_i \lesssim \left(\#\mathcal{T}_i - \#\mathcal{T}_0\right)^{-s} \|u\|_{\mathcal{A}^r}\). Thus our AFEM realizes the best possible convergence rate.

For earlier results on quasi-optimality of AFEM with respect to the energy-norm and more details on approximation classes and references, we refer to [19, 4, 16].

The following lemma will be used to bound the number of marked elements in the Dörfler marking.

**Lemma 7.1.** With \(C_1\) and \(C_2\) being the constants from Proposition 3.2 and Lemma 4.4, let \(\theta < \frac{1}{C_1(C_2 + 1)}\). Let \(\mu\) be sufficiently small such that Lemma 4.4 is valid. Then for \((\mathcal{T}_n \ni \mathcal{T}_i) \subset \mathcal{T}_i\) with

\[
\|u - u_{\mathcal{T}}\|_{L^2(\Omega)} + \text{osc}(T) \leq [1 - \theta C_1(C_2 + 1)](\|u - u_i\|_{L^2(\Omega)} + \text{osc}_i),
\]

it holds that

\[
\eta(\mathcal{R}_{\mathcal{T}_i} \to \mathcal{T}) \geq \theta \eta_i.
\]

**Proof.** By adding the inequalities

\[
\|u - u_i\|_{L^2(\Omega)} \leq \|u - u_{\mathcal{T}}\|_{L^2(\Omega)} + \|u - u_{\mathcal{T}}\|_{L^2(\Omega)} + \text{osc}_i \leq \text{osc}(\mathcal{R}_{\mathcal{T}_i} \to \mathcal{T}) + \text{osc}(T)
\]

and employing (7.1), we infer that

\[
\theta(C_2 + 1) \eta_i \leq \theta C_1(C_2 + 1)(\|u - u_i\|_{L^2(\Omega)} + \text{osc}_i) \\
\leq \|u - u_i\|_{L^2(\Omega)} + \text{osc}_i - \|u - u_{\mathcal{T}}\|_{L^2(\Omega)} - \text{osc}(T) \\
\leq \|u_i - u_{\mathcal{T}}\|_{L^2(\Omega)} + \text{osc}(\mathcal{R}_{\mathcal{T}_i} \to \mathcal{T}) \\
\leq (C_2 + 1) \eta_i \eta(\mathcal{R}_{\mathcal{T}_i} \to \mathcal{T}),
\]

where the first and last inequality follow from applications of Proposition 3.2 and Lemma 4.4, respectively.

**Corollary 7.2.** For some \(s > 0\), let \(u \in \mathcal{A}^r\). Assume also that \(\mu\) is sufficiently small that Corollary 4.7 is valid. Then under the assumptions of Lemma 7.1, the collection of marked elements \(\mathcal{M}_i \subset \mathcal{T}_i\) defined by (5.1) satisfies

\[
\#\mathcal{M}_i \lesssim |u|_{\mathcal{A}^r}^{-s}(\|u - u_i\|_{L^2(\Omega)} + \text{osc}_i)^{-1/s}.
\]
Proof. With $C_3$ the constant from Corollary 4.7, by definition of $A^s$ there exists a partition $T' \in \mathbb{T}$ with

$$\#T' - \#T_0 \lesssim |u|_{A^s}^{1/s} \left( \frac{1 - \theta C_4(C_2 + 1)}{1 + C_3} \left( \|u - u_i\|_{L_2(\Omega)} + \text{osc}_i \right) \right)^{-1/s},$$

(7.2)

and a $v_T' \in S_{T'}$ with

$$\|u - v_T'\|_{L_2(\Omega)} + \text{osc}(T') \leq \frac{1 - \theta C_4(C_2 + 1)}{1 + C_3} \left( \|u - u_i\|_{L_2(\Omega)} + \text{osc}_i \right).$$

As is shown in Appendix A, $T'$ can be refined to a partition $T'' \in \mathbb{T}_\mu$ with $T'' - \#T_0 \lesssim \#T'' - \#T_0$, dependent on $\mu$. The smallest common refinement $T$ of $T_i$ and $T''$ is in $\mathbb{T}_\mu$ with $\#T - \#T_i \leq \#T'' - \#T_0$ (cf. [19] last lines of the proof of Lemma 5.2). Since $S_{T'} \subset S_T$, Corollary 4.7 shows that

$$\|u - u_T\|_{L_2(\Omega)} + \text{osc}(T) \leq (C_3 + 1) \left( \inf_{v_T' \in S_T} \|u - v_T'\|_{L_2(\Omega)} + \text{osc}(T) \right) \leq (1 - \theta C_4(C_2 + 1)) \left( \|u - u_i\|_{L_2(\Omega)} + \text{osc}_i \right),$$

and so $\eta(R_{T_i \rightarrow T}) \geq \eta_{T_i}$ by Lemma 7.1. Since $M_i$ is the smallest subset of $T_i$ with $\eta(M_i) \geq \eta_{T_i}$, we conclude that

$$\#M_i \leq \#R_{T_i \rightarrow T} \leq \#T - \#T_i \leq \#T'' - \#T_0 \lesssim \#T' - \#T_0,$$

so that the proof follows from (7.2). □

Finally, in the next theorem the quasi-optimality result is stated.

THEOREM 7.3. For some $s > 0$, let $u \in A^s$. Then, under the assumptions from Corollary 7.2, it holds that

$$\#T_i - \#T_0 \lesssim \left( \|u - u_{i-1}\|_{L_2(\Omega)} + \text{osc}_{i-1} \right)^{-1/s} |u|_{A^s}^{1/s}.$$

Proof. In Theorem A.5 it is shown that $\#T_i - \#T_0 \lesssim \sum_{j=0}^{i-1} M_j$. The proof follows by combining this result with Corollaries 7.2 and 6.4. □

Appendix A. Partitions and mesh functions. We specify the type of partitions that we consider, and derive some of their properties. When doing so we recall some results from [20], which generalize upon known results for newest vertex bisection in two dimensions. In addition, we introduce the concept of sufficiently mildly graded partitions and define suitable mesh-functions.

A.1. Bisection and uniform shape regularity. For $0 \leq p \leq n - 1$, a (closed) simplex spanned by $p + 1$ vertices of an $n$-simplex $T$ is called a hyperface of $T$. For $p = n - 1$, it will be called a facet. A partition $T$ of a domain $\Omega \subset \mathbb{R}^n$, i.e., a subdivision of $\Omega$ into (essentially) disjoint closed $n$-simplices, is called conforming when the intersection of any two different $T, T' \in T$ is either empty, or a hyperface of both simplices. Different simplices $T, T'$ that share a facet will be called neighbors.

Simplices are refined by means of bisection. In order to guarantee uniform shape regularity of all descendants, a cyclic choice of the refinement edges has to be made. To that end, given \{x_0, \ldots, x_n\} \subset \mathbb{R}^n, not on a joint $(n - 1)$-dimensional hyperplane, we distinguish between $n(n + 1)!$ tagged simplices given by all possible ordered
sequences \((x_0, x_1, \ldots, x_n)_\gamma\) and types \(\gamma \in \{0, \ldots, n-1\}\). Given a tagged simplex \(T = (x_0, x_1, \ldots, x_n)_\gamma\), its children are the tagged simplices
\[
(x_0, \frac{x_0 + x_n}{2}, x_1, \ldots, x_{\gamma}, x_{\gamma+1}, \ldots, x_{n-1})_{(\gamma+1) \mod n}
\]
and
\[
(x_n, \frac{x_0 + x_n}{2}, x_1, \ldots, x_{\gamma}, x_{n-1}, \ldots, x_{\gamma+1})_{(\gamma+1) \mod n},
\]
where the sequences \((x_{\gamma+1}, \ldots, x_{n-1})\) and \((x_1, \ldots, x_\gamma)\) should be read as being void for \(\gamma = n-1\) and \(\gamma = 0\), respectively. So these children are defined by bisecting the edge \(x_0 x_n\) of \(T\), i.e., by connecting its midpoint with the other vertices \(x_1, \ldots, x_{n-1}\), and by an appropriate ordering of their vertices, and by having type \((\gamma+1) \mod n\). See Figure A.1 for an illustration. This bisection process was introduced in [21], and

\[\text{Fig. A.1. Bisection of a tagged tetrahedron of type 0 with the next two level cuts indicated.}\]

in different notations, in [10]. The edge \(x_0 x_n\) is called the refinement edge of \(T\). In the \(n = 2\) case, the vertex opposite to this edge is known as the newest vertex.

Corresponding to a tagged simplex \(T = (x_0, \ldots, x_n)_\gamma\), we set
\[
T_R = (x_n, x_1, \ldots, x_{\gamma}, x_{n-1}, \ldots, x_{\gamma+1}, x_0)_\gamma
\]
which is the tagged simplex that has the same set of children as \(T\), and in this sense is equal to \(T\). So actually we distinguish between \(\frac{1}{2}n(n+1)!\) tagged simplices.

Given a fixed, conforming initial partition \(T_0\) of \(\bar{\Omega}\) into tagged simplices of some fixed type \(\gamma\), we exclusively consider partitions that can be created from \(T_0\) by recurrent bisectations of tagged simplices, for short, descendants of \(T_0\). Simplices from descendants of \(T_0\) are uniformly shape regular, dependent only on \(T_0\) and \(n\). In view of the refinement by bisection, this means that there exist constants \(d, D > 0\), dependent only on \(T_0\), such that
\[
d2^{-\ell(T)} \leq \text{vol}(T) = h_T^n, \quad \text{diam}(T) \leq D2^{-\ell(T)/n},
\]
Here \(\ell(T)\) denotes the level of a \(T\), being the number of bisections needed to create \(T\) from a simplex from \(T_0\).
A.2. Conforming partitions. For the application of an a posteriori error estimator, we need partitions that are conforming. The set of conforming descendants of $\mathcal{T}_0$ is denoted as $\mathcal{T}$. A partition generated from a conforming one by bisecting some marked simplices is generally non-conforming, so that additional bisections have to be made to restore conformity.

To bound the cardinality of the output partition of a call of AFEM, we would like that the total number of additional bisections needed to restore conformity after bisecting some marked simplices can be bounded by some absolute multiple of the total number of marked simplices inside such a call. To guarantee this property, we assume that the simplices from $\mathcal{T}_0$ are tagged in such a way that any two neighbors $T = (x_0, \ldots, x_n) \gamma$, $T' = (x'_0, \ldots, x'_n) \gamma$ from $\mathcal{T}_0$ match in the sense that if $x_0, x_n$ or $x'_0, x'_n$ is on $T \cap T'$, then either $T$ and $T'$ are reflected neighbors – meaning that the ordered sequence of vertices of either $T$ or $T_R$ coincides with that of $T'$ on all but one position–, or the pair of neighboring children of $T$ and $T'$ are reflected neighbors. See Figure A.2 for an illustration.

It is known, see [2] and the references therein, that for any conforming partition into triangles, i.e., for $n = 2$, there exists a local numbering of the vertices so that the matching condition is satisfied. For $n > 2$, any conforming partition of $n$-simplices can be refined, by inflating the number of simplices by not more than an absolute constant factor, into a conforming partition $\mathcal{T}_0$ that allows a local numbering of the vertices so that the matching condition is satisfied.

The matching condition on $\mathcal{T}_0$ is a necessary and sufficient condition so that any uniform refinement of $\mathcal{T}_0$ – i.e., a refinement of $\mathcal{T}_0$ in which all simplices have the same level– is conforming. For this reason, we always assume that $\mathcal{T}_0$ satisfies the matching condition.

Tagged neighbors are called compatibly divisible when they have the same refinement edge. For a descendant $T$ of $\mathcal{T}_0$, and $T \in \mathcal{T}$, we set

$$N(T, T) := \{\text{neighbors } T' \text{ of } T \text{ in } \mathcal{T} \text{ that contain the refinement edge of } T\}.$$ 

As a consequence of the matching condition, we have the following result.

**Corollary A.1.** For any descendant $T$ of $\mathcal{T}_0$, and $T \in \mathcal{T}$, it holds that
1. $|\ell(T') - \ell(T)| \leq 1$ for any neighbor $T' \in \mathcal{T}$ of $T$,
2. and for $T' \in N(T, T)$, either
   • $\ell(T') = \ell(T)$ and $T, T'$ are compatibly divisible, or
   • $\ell(T') = \ell(T) - 1$ and $T$ is compatibly divisible with one of both children of $T'$.

In view of the second part of this corollary, an algorithm $\text{bisect}(T, T')$ for finding the smallest $T \subset T' \in \mathcal{T}$ in which some selected simplex $T \in T$ is bisected may consist of the following 2 steps: Firstly, using recursive calls of $\text{bisect}(T, T''')$ for some
with \( \ell(T'') = \ell(T) - 1 \), construct a \( T \subset T' \in \mathcal{T} \) for which there exists a \( K \subset \mathcal{T} \) with \( T \in K \) and for which all \( T' \in K \) have the same refinement edge. Secondly, bisect all these \( T' \in K \) simultaneously, which does not create any “hanging nodes”.

### A.3. Graded partitions and the routine “bisect”.

For purposes of proving convergence and optimality of AFEM, we need partitions \( T \in \mathcal{T} \) whose (possible) grading is sufficiently mild. Given a constant \( G > 0 \), we call a descendant \( T \) of \( T_0 \) \( G \)-graded when for all \( T, T' \in \mathcal{T} \),

\[
\text{dist}(T, T') \leq G2^{-\ell(T)/n} \Rightarrow \ell(T') = \ell(T) - 1.
\]

Thus \( G \)-gradedness locally restricts the growth of the size of mesh elements as one moves away from a given element. If we want to bisect a simplex in such a partition, then generally additional bisections have to be made not only to guarantee conformity of the resulting partition, but also its \( G \)-gradedness. The following routine \( \text{bisect}[T,T] \) ensures both properties. The first part of this routine, whose steps were already described at the end of the previous subsection, can be found in [20], whereas its last, additional loop before the actual bisection of \( T \) in the last line ensures the \( G \)-gradedness of the output partition.

\[
\text{bisect}[T,T] \rightarrow T':
\]

\( \% \ T \in \mathcal{T} \text{ is } G \text{-graded, and } T \in \mathcal{T} \).

\[
K := \emptyset; \ F = \{T\}
\]

do \( F\text{new} := \emptyset \)

forall \( T' \in F \) do

forall \( T'' \in N(T, T') \) with \( T'' \notin F \cup K \) do

if \( T'' \) compatibly divisible with \( T' \)

then \( F\text{new} := F\text{new} \cup \{T''\} \)

else \( T := \text{bisect}[T,T''] \)

add to \( F\text{new} \) the child of \( T'' \) that is a neighbor of \( T' \)

endif

endfor

\( K := K \cup F \)

\( F := F\text{new} \)

until \( F = \emptyset \)

forall \( T''' \in \mathcal{T} \) with \( \text{dist}(K, T''' \} \leq G2^{-\ell(T)/n} \) and \( \ell(T''') = \ell(T) - 1 \) do

\( T := \text{bisect}[T,T'''] \)

endfor

Create \( T' \) from \( T \) by simultaneously bisecting all \( T' \in K \)

Properties of \( \text{bisect} \) are stated in the following two lemmas. Similar to [20, Theorem 5.1], the first lemma can be proven by induction to \( \ell(T) \).

**Proposition A.2.** \( T' := \text{bisect}[T,T] \) terminates, and \( T' \) is the smallest \( G \)-graded refinement of \( T \) in \( \mathcal{T} \) in which \( T \) has been bisected.

If \( T' \in T' \) is newly created by the call, then \( \ell(T') \leq \ell(T) + 1 \).

**Remark A.3.** Assuming that the datastructure allows us to find all neighbors of a simplex in \( T \in \mathcal{T} \) in \( O(1) \) operations, the number of operations needed for \( T' := \text{bisect}[T,T] \) is \( O(#T' - #T) \).
Proposition A.4. Any newly created $T'$ by the call \texttt{bisect}[T, T]$ satisfies

$$\text{dist}(T', T) \leq (D(1 + 2^{1/n}) + 2^{-1/n}G) \sum_{i=\ell(T')}^{\ell(T)} 2^{-i/n} \quad (\approx \text{diam}(T')).$$

Proof. For $\ell(T) = 0$, any newly created $T'$ is a child of a $\bar{T}$ that has its refinement edge on $\partial T$, so that $\text{dist}(T', T) = 0$. Note that in this case the sum over $i$ is empty since $\ell(T') = \ell(T) + 1$.

Assuming that the theorem holds for $\ell(T) = \ell - 1 \geq 0$, let us consider $T$ with $\ell(T) = \ell$. If $T'$ is created by bisection of any simplex from the set $K$, then $\text{dist}(T', T) = 0$ as in the $\ell(T) = 0$ case. If $T'$ is created by a recursive call \texttt{bisect}[T, T''], then either $\text{dist}(T'', T) = 0$ or

$$\text{dist}(T'', T) \leq \text{dist}(T''', K) + D 2^{-\ell(T'')} \leq (D + 2^{-1/n}G) 2^{-\ell(T'')/n},$$

where we used that for any $\bar{T} \in K$, $\ell(\bar{T}) = \ell(T)$ and thus $\text{diam}(\bar{T}) \leq D 2^{-\ell(T'')}$. Now by $\ell(T''') = \ell(T) - 1$, the induction hypothesis shows that

$$\text{dist}(T', T) \leq \text{dist}(T', T'') + \text{diam}(T'') + \text{dist}(T'', T)$$

$$\leq (D(1 + 2^{1/n}) + 2^{-1/n}G) \sum_{i=\ell(T'')}^{\ell(T''')} 2^{-i/n} + D 2^{1/n} 2^{-\ell(T'')} + (D + 2^{-1/n}G) 2^{-\ell(T'')/n}$$

$$= (D(1 + 2^{1/n}) + 2^{-1/n}G) \sum_{i=\ell(T')}^{\ell(T')} 2^{-i/n}.$$

Our AFEM is a loop of the following form:

$$\mathcal{T} := \mathcal{T}_0$$

\textbf{do} mark some set $\mathcal{M} \subset \mathcal{T}$ for bisection

\textbf{for} $T \in \mathcal{M}$ \textbf{do}

\textbf{if} $T \in \mathcal{T}$ \% i.e., if it has not already been bisected as a byproduct of a \% previous call of \texttt{bisect} in this \textbf{for}-loop

\textbf{then} $T := \text{bisect}[T, T]$

\textbf{endif}

\textbf{endfor}

\textbf{until} satisfied

Actually, we mark simplices possibly for some fixed number of multiple bisections. By scheduling such multiple bisections as a sequence of single ones, the algorithm can still be written in the above form.

The following theorem shows the important result that the difference between the cardinalities of the output and initial partition can be bounded by some absolute multiple of the total number of marked simplices. Besides the uniform shape regularity, the proof given in [20] relies only on the result of Proposition A.4 and the second statement from Proposition A.2. As we have shown, these results are still valid when considering conforming partitions that are additionally $G$-graded, and so is the theorem.

We emphasize that the proof from [20] is a harmless modification of the original, ingenious proof given in [2].
THEOREM A.5. Let $K$ be the total number of calls of `bisect` in the above AFEM loop, so that $K$ is no larger than the sum of the cardinalities of all sets of marked simplices. Then for the output partition $T$, it holds that $\#T - \#T_0 \lesssim K$, dependent only on the constants $d$, $D$, $G$, and $n$.

Any descendant $\tilde{T}$ of $T_0$, not necessarily conforming or $G$-graded, is constructed from $T_0$ by a sequence of bisections, where $\#(\tilde{T} \setminus T_0)$ is equal to the number of bisections. By replacing these bisections by calls of `bisect`, we infer the following consequence of Theorem A.5. It shows that there is no essential restriction in considering only $G$-graded partitions from $T$.

COROLLARY A.6. Any descendant $\tilde{T}$ of $T_0$ can be refined to a $G$-graded partition $T \in T$ with $\#T - \#T_0 \lesssim \#T - \#T_0$, dependent only on $d$, $D$, $G$, and $n$.

Remark A.7. For any descendant $T$ of $T_0$, it holds that $\#T - \#\tilde{T}_0 \leq \#(T \setminus T_0) \leq 2(\#T - \#T_0)$.

A.4. Mesh functions. In this subsection, we are going to construct mesh functions $h_T \in C(\tilde{\Omega}) \cap \prod_{T \in T_0} P_1(T)$ such that for any $\mu > 0$, $\|\nabla h_T\|_{L^\infty(\Omega)} \leq \mu$ for all $G$-graded $T \in T$ with $G \cdot \mu$ sufficiently large, and such that $h_T|_T \approx h_T$ uniformly in $\mu$, $T \in T$ and in all those $G$-graded $T \in T$. Moreover, these mesh functions will be pointwise non-increasing under conforming and $G$-graded refinements. The class of $G$-graded $T \in T$ for this value of $G$ can thus be taken as the class $T_\mu$ as introduced in Sect. 1. In view of Corollary A.6, this class is sufficiently large for our purposes.

For a $T \in T$, $z$ a vertex of $T$, and $p \in \mathbb{N}_0$, we define the patches

$$R_p(z, T) = \begin{cases} \cup_{\{T \in T: T \ni z\}} T & \text{when } p = 0, \\ \cup_{\{z': z' \ni z, z' \text{ are vertices of some } T \in T\}} R_{p-1}(z', T) & \text{when } p > 0. \end{cases}$$

PROPOSITION A.8. For $G \geq \frac{2}{(2p+1)} D 2^1/n$, a $G$-graded $T \in T$, and $z$ a vertex of $T$, the levels of any two simplices in $R_p(z, T)$ differ at most one.

Proof. Let $T \in R_p(z, T)$, and let $z'$ be the vertex of $T$ that is nearest to $z$. By definition of a $G$-graded partition, any $T'' \in T$ with $\text{dist}(T, T'') \leq G 2^{-\ell(T)/n}$ satisfies $\ell(T'') \geq \ell(T) - 1$, and so $\text{diam}(T'') \leq D 2^1/n 2^{-\ell(T)/n}$. As a consequence, for $\bar{p} \leq \frac{G 2^{-\ell(T)/n}}{D 2^1/n 2^{-\ell(T)/n}}$ and any $T'' \in R_p(z', T)$, it holds that $\ell(T'') \geq \ell(T) - 1$. Since $R_p(z, T) \subset R_{\bar{p}}(z', T)$ when $\bar{p} \geq 2p+1$, the proof is completed. \qed

Remark A.9. Note that a patch $R_p(z, T)$ is not necessarily simply connected. An example is given in Figure A.3.

Fig. A.3. An example of a non-simply connected $R_1(z, T)$.
For \( z \) a vertex of a conforming partition \( T \), let
\[
\ell(z) := \min_{\{T \in T : z \in T\}} \ell(T)
\]
be the level of \( z \) (in \( T \)).

**Lemma A.10.** For a \( G \)-graded \( T \in \mathcal{T} \), \( p \in (0, G2^{-2/n}] \), and \( z \) a vertex of \( T \), any two vertices in \( B(z, \rho 2^{-\ell(z)/n}) \cap T \) are connected by a path along the edges of \( T \) of length at most \( L \) edges, where \( L \) depends only on \( p \), \( n \) and the shape regularity parameters \( d \) and \( D \).

Furthermore, for any \( T \in \mathcal{T} \) with \( T \cap B(z, \rho 2^{-\ell(z)/n}) \neq \emptyset \), \( |\ell(T) - \ell(z)| \leq 1 \).

**Proof.** For all \( T \in \mathcal{T} \) with \( T \cap B(z, \rho 2^{-\ell(z)/n}) \neq \emptyset \), it holds that \( \ell(T) \leq \ell(z) + 1 \). Indeed, assume that such a \( T \) exists with \( \ell(T) \geq \ell(z) + 2 \). Then, using that the levels of neighboring simplices in \( T \) differ at most one (Corollary A.1(1)), there exists such a \( T \) with \( \ell(T) = \ell(z) + 2 \). However, the existence of such a \( T \) is in conflict with the \( G \)-gradedness of the partition since \( \text{dist}(T, z) \leq G2^{-\ell(z)+2/n} \). Finally, the \( G \)-gradedness of the partition also implies that \( \ell(T) \geq \ell(z) - 1 \) for any \( T \in \mathcal{T} \) with \( T \cap B(z, \rho 2^{-\ell(z)/n}) \neq \emptyset \).

We infer that the number of \( T \in \mathcal{T} \) with \( T \cap B(z, \rho 2^{-\ell(z)/n}) \neq \emptyset \) is uniformly bounded, dependent only on \( p \), \( n \), \( d \) and \( D \), and, consequently, that any 2 vertices in \( B(z, \rho 2^{-\ell(z)/n}) \cap T \) are connected by a path along the edges of \( T \) with a length that is uniformly bounded. \( \Box \)

**Proposition A.11.** It holds that
\[
\#R_p(z, T) \lesssim (p + 1)^n
\]
uniformly in \( p \in \mathbb{N}_0 \), all \( G \)-graded \( T \in \mathcal{T} \) with \( G \geq (2p + 1)D2^{1/n} \) and vertices \( z \) of \( T \), as well as
\[
\#R_p(z, T) \gtrsim (p + 1)^n
\]
when additionally \( p2^{-\ell(z)/n} \lesssim 1 \), dependent only on the domain \( \Omega \).

Finally, it holds that
\[
\frac{\#R_p(z, T)}{\sum_{q=0}^p \#R_q(z, T)} \lesssim \frac{1}{p + 1},
\]
uniformly in \( p \in \mathbb{N}_0 \), all \( G \)-graded partitions \( T \in \mathcal{T} \) with \( G \geq (2p + 1)D2^{1/n} \) and vertices \( z \) of \( T \).

**Proof.** From Lemma A.8, we know that \( \text{diam}(T) \approx 2^{-\ell(z)/n} \) uniformly in \( T \in R_p(z, T) \), and thus that \( R_p(z, T) \) is contained in a ball having radius not larger than some absolute multiple of \( (p + 1)2^{-\ell(z)/n} \). Because \( \text{vol}(T) \gtrsim 2^{-\ell(z)} \) for \( T \in R_p(z, T) \), we conclude that \( \#R_p(z, T) \lesssim (p + 1)^n \).

For any \( z \in \mathcal{T} \), there exists a cone \( C \) with vertex \( z \) and fixed height and opening, dependent only on \( \Omega \), such that \( C \subset \bar{\Omega} \). We are going to show that for any \( M \geq 1 \),
\[
C \cap B(z, M2^{-\ell(z)/n}) \cap T \subset R_p(z, T)
\]
whenever \( p \) is not less than some absolute multiple of \( M \) and \( G \geq (2p + 1)D2^{1/n} \). Since \( \text{vol}(B(z, M2^{-\ell(z)/n}) \cap C) \gtrsim M^n2^{-\ell(z)} \) whenever \( M2^{-\ell(z)/n} \lesssim 1 \), dependent only on the sizes of \( C \), and \( \text{vol}(T) \lesssim 2^{-\ell(z)} \) for \( T \in R_p(z, T) \), this will imply that \( \#R_p(z, T) \gtrsim (p + 1)^n \) whenever \( p2^{-\ell(z)/n} \lesssim 1 \).
Let \( \rho \) be a fixed, sufficiently large constant to be determined below, and let \( M \geq 1 \) be given. By taking \( p/M \) sufficiently large and requiring that \( G \geq (2p + 1)D2^{\ell/n} \), we may assume that \( G \geq 2^{2/n} \rho \). Let \( z' \in C \cap B(z, M2^{-\ell(z)/n}) \) be a vertex of \( T \) and \( M' := |z - z'|2^{\ell(z)/n} \).

If \( M' \leq \rho \), then, by taking \( p \geq L - 1 \), where \( L = L(\rho, m, d, D) \in \mathbb{N} \) is the constant from Lemma A.10, \( z' \) is a vertex in \( R_p(z, T) \) thanks to this lemma.

Assume instead that \( M' > \rho \). We now fix \( \rho \) so that \( \rho/D \) is sufficiently large to ensure that the quotient of the diameter of \( B(z, \rho2^{-\ell(z)/n}) \) and the maximal distance between any 2 neighboring vertices in \( B(z, \rho2^{-\ell(z)/n}) \cap T \) is in turn large enough to ensure the existence of a vertex \( z_1 \) of \( T \) in \( C \cap B(z, \rho2^{-\ell(z)/n}) \) with

\[
|z_1 - z'| \leq (M' - \frac{1}{2}\rho)2^{-\ell(z)/n}.
\]

The existence of such a \( \rho \) and \( z_1 \) is guaranteed by Lemma A.10, where it is shown that for any \( T \in T \) with \( T \cap B(z, \rho2^{-\ell(z)/n}) \neq \emptyset \), \( \ell(T) - \ell(z) \leq 1 \). Note that the required size of \( \rho \) depends only on the opening of the cone \( C \) and on \( D \).

With \( z_0 := z \), a repeated application of this argument shows that there exists a sequence of vertices \( z_1, z_2, \ldots \) of \( T \) with \( z_i \) in \( C \cap B(z_{i-1}, \rho2^{-\ell(z_{i-1})/n}) \) and

\[
|z_i - z'| \leq |z_{i-1} - z'| - \frac{1}{2}\rho2^{-\ell(z_{i-1})/n}
\]

whenever \( |z_{i-1} - z'| > \rho2^{-\ell(z_{i-1})/n} \), and with \( z_i = z' \) otherwise, see Figure A.4 for an illustration.

![Illustration with the proof of Proposition A.11.](image)

**Fig. A.4. Illustration with the proof of Proposition A.11.**

Let \( m \geq 2 \) be the smallest integer with \( M - \frac{1}{2}\rho - (m - 1)\frac{1}{2}2^{-\ell(z)} \rho \leq 0 \). Taking \( p \geq (m - 1)L \), we have for \( 1 \leq i \leq m - 1 \) that any \( T \in T \) with \( T \ni z_i \) is in \( R_p(z, T) \) by Lemma A.10. Taking \( G \geq (2p + 1)D2^{\ell/n} \) ensures that \( |\ell(z_{i-1}) - \ell(z)| \leq 1 \) by Proposition A.8. By substituting \( \ell(z_{i-1}) \leq \ell(z) + 1 \) in (A.1), we conclude that \( z_m = z' \). Requiring in addition that \( p \geq mL - 1 \) implies also \( z' = z_m \in R_p(z, T) \).

Since \( mL \leq M \) and thus \( p \geq mL - 1 \) is satisfied when \( p/M \) is sufficiently large, we have completed the proof that \( \#R_p(z, T) \geq (p + 1)^{n} \) whenever \( p2^{-\ell(z)/n} \leq 1 \).

The combination of the upper and lower bound for \( \#R_p(z, T) \) shows the last statement of this Proposition. The above arguments show that the lower bound \( \#R_p(z, T) \geq (p + 1)^{n} \) will only be violated when \( p \) is that large that \( R_p(z, T) \) (nearly)
fills the domain, i.e., that it is equal to \( T \). From that value of \( p \) on, \(#R_p(z, T)\) will not grow anymore, from which we infer the last statement also in that case. \( \square \)

For any \( T \in \mathbb{T} \), we define \( \bar{h}_{T,p} \in C(\bar{\Omega}) \cap \prod_{T \in \mathbb{T}} P_1(T) \) by

\[
\bar{h}_{T,p}(z) = \frac{\sum_{q=0}^{p} \sum_{T \in \mathbb{R}_q(z,T)} h_T}{\sum_{q=0}^{p} \#R_q(z, T)}
\]

for any vertex \( z \) of \( T \). The next proposition shows that \( \bar{h}_{T,p} \) is appropriately called a mesh function. The proof follows from Proposition A.8 and the uniform shape regularity.

**Proposition A.12.** Assume that \( T \in \mathbb{T} \) is \( G \)-graded with \( G \geq (2p + 1)D2^{1/n} \). Then

\[
\bar{h}_{T,p}|_T \approx h_T
\]

uniformly in \( p \in \mathbb{N}_0, G \), and \( T \in \mathbb{T} \).

Next we show that \( \bar{h}_{T,p} \) can be made arbitrarily flat by decreasing the grading of the partition, i.e., by increasing \( G \), and by increasing \( p \) proportionally.

**Proposition A.13.** Assume that \( T \in \mathbb{T} \) is \( G \)-graded with \( G \geq (2p + 1)D2^{1/n} \). Then

\[
|\bar{h}_{T,p}(z) - \bar{h}_{T,p}(z)| \lesssim \frac{1}{p+1} h_T,
\]

uniformly in \( p \in \mathbb{N}_0, G, T \in \mathbb{T}, \) and vertices \( z, z' \in T \). Thus \( \|\nabla \bar{h}_{T,p}\|_{L_\infty(\Omega)} \lesssim \frac{1}{p+1} \).

**Proof.** For \( z'' \in \{z, z'\} \), let

\[
n(z'') = \sum_{q=0}^{p} \sum_{T' \in \mathbb{R}_q(z'', T)} h_{T'}, \quad d(z'') = \sum_{q=0}^{p} \#R_q(z'', T).
\]

Letting \( R_{-1} (\cdot, \cdot) := \emptyset \), we have \( R_q(z'', T) \setminus (R_q(z, T) \cap R_q(z', T)) \subset R_q(z'', T) \setminus R_{q-1}(z'', T) \). Employing Proposition A.11, we compute

\[
\sum_{q=0}^{p} \#(R_q(z'', T) \setminus R_{q-1}(z'', T)) = \#R_p(z'', T) \lesssim \frac{1}{p+1} d(z'').
\]

Applying Proposition A.8 then yields \( |n(z) - n(z')| \lesssim \frac{h_T}{p+1} (d(z)+d(z')) \), \( |d(z) - d(z')| \lesssim \frac{1}{p+1} (d(z) + d(z')) \), and \( \frac{n(z'')}{d(z'')} \lesssim h_T. \) Assuming \( d(z') \geq d(z) \), we may complete the proof by writing

\[
\frac{n(z)}{d(z)} - \frac{n(z')}{d(z')} = \frac{n(z) d(z') - d(z) n(z')}{d(z) d(z')} + \frac{n(z) - n(z')}{d(z')}.
\]

\( \square \)

**Remark A.14.** Proposition A.12 is still valid if the more obvious definition \( \bar{h}_{T,p} = \sum_{T \in \mathbb{R}_p(z,T)} h_T / \#R_p(z, T) \) of the mesh function is used. In order to prove a result such as Proposition A.13 for this mesh function, however, it is necessary to show that, for \( G \geq (2p + 1)D2^{1/n} \), \( \#(R_p(z'', T) \setminus R_{p-1}(z'', T))/\#R_p(z'', T) \lesssim \frac{1}{p+1} \), or at least that it tends to zero for \( p \to \infty \). It turns out that such an estimate is difficult to establish, cf. also Remark A.9.
For some applications it is important to have a family of mesh functions that is non-increasing under (conforming, \( G \)-graded) refinements. For \( G \)-graded \( T \in \mathcal{T} \), we define \( h_{T,p} \in C(\bar{\Omega}) \cap \prod_{T \in \mathcal{T}} P_1(T) \) by
\[
h_{T,p}(z) = \min_{\{T' \in \mathcal{T} \text{ is } G-\text{graded}; T' \subset T\}} \tilde{h}_{T,p}(z),
\]
where \( z \) is any vertex of \( T \).

**Proposition A.15.** The mapping \( \mathcal{T} \ni T \mapsto h_{T,p} \) is pointwise non-increasing under \( G \)-graded refinements, and for some constants \( c_T, C_T, \tilde{C}_\mathcal{T} \), which are independent of \( p \), it holds that
\[
c_T h_T \leq h_{T,p} \leq C_T h_T, \quad \|\nabla h_{T,p}\|_{L_\infty(\Omega)} \leq \tilde{C}_\mathcal{T} \frac{1}{h_T}.
\]
for all \( G \)-graded \( T \in \mathcal{T} \) with \( G \geq (2p+1)D2^{1/n} \), and \( T \in \mathcal{T} \).

**Proof.** The first two statements follow by definition of \( h_{T,p} \) and Proposition A.12. For vertices \( z_1, z_2 \) of \( T \), let \( h_{T,p}(z_1) = h_{T,p}(z_2) = \tilde{h}_{T,p}(z_2) \). Then \( |h_{T,p}(z_1) - h_{T,p}(z_2)| \leq \|\nabla h_{T,p}\|_{L_\infty(\Omega)} |z_1 - z_2| \). Now an application of Proposition A.13 shows the second statement. \( \square \)

We conclude that given a \( \mu > 0 \), for \( p + 1 \geq \tilde{C}_\mathcal{T} \mu^{-1} \) and \( G \geq (2p+1)D2^{1/n} \), the collection of all \( G \)-graded \( T \in \mathcal{T} \) can be used as the class \( \mathcal{T}_\mu \).

**REFERENCES**


