Foundations of Computational Mathematics manuscript No.
(will be inserted by the editor)

Optimality of a standard adaptive finite element method*

Rob Stevenson

Department of Mathematics, Utrecht University, The Netherlands,
e-mail: stevenson@math.uu.nl

The date of receipt and acceptance will be inserted by the editor

Abstract In this paper, an adaptive finite element method is constructed for solving elliptic equations that has optimal computational complexity. Whenever for some $s > 0$, the solution can be approximated to accuracy $\mathcal{O}(n^{-s})$ in energy norm by a continuous piecewise linear function on some partition with n triangles, and one knows how to approximate the right-hand side in the dual norm with the same rate with piecewise constants, then the adaptive method produces approximations that converge with this rate, taking a number of operations that is of the order of the number of triangles in the output partition. The method is similar in spirit to that from [*SINUM*, 38 (2000), pp.466–488] by Morin, Nochetto, and Siebert, and so in particular it does not rely on a recurrent coarsening of the partitions. Although the Poisson equation in two dimensions with piecewise linear approximation is considered, it can be expected that the results generalize in several respects.

Key words Adaptive finite element method, optimal computational complexity, a posteriori error estimator, non-linear approximation
Mathematics Subject Classification (2000): 65N30, 65N50, 65N15, 65Y20, 41A25

1 Introduction

Adaptive finite element methods for solving elliptic boundary value problems have the potential to produce a sequence of approximations to the solution that converges with a rate that is optimal in view of the polynomial

* This work was supported by the Netherlands Organization for Scientific Research and by the European Community's Human Potential Programme under contract HPRN-CT-2002-00286.

order that is applied, also in the, common, situation that finite element approximations with respect to uniformly refined partitions exhibit a reduced rate due to a lacking (Sobolev) regularity of the solution. The basic idea of an adaptive finite element method is, given some finite element approximation, to create a refined partition by subdividing those elements where local error estimators indicate that the error is large, and then, on this refined partition, to compute the next approximation, after which the process can be repeated. Although, because of their success in practice, during the last 25 years the use of these adaptive methods became more and more widely spread, apart from results in the one-dimensional case by Babuška and Vogelius ([2]), their convergence was not shown before the work by Dörfler ([8]), that was later extended by Morin, Nochetto and Siebert ([10]).

Although these results meant a break through in the theoretical understanding of adaptive methods, they do not tell anything about the rate of convergence, and so, in particular, they do not show that adaptive methods are more effective than, or even competitive with non-adaptive ones in the situation that the solution has a lacking regularity.

Recently, in [4], Binev, Dahmen and DeVore developed an adaptive finite element method which they showed to be of optimal computational complexity. Whenever for some $s > 0$, the solution is in the approximation class $\mathcal{A}^s$, meaning that there exists a sequence of partitions of the domain into n elements such that the best finite element approximation with respect to this partition has an error in energy norm of order n^{-s}, then the adaptive method produces a sequence of approximations that converge with the same rate, where, moreover, the cost of computing such an approximation is of the order of the number of elements in the underlying partition. A combination of the (near) characterization of $\mathcal{A}^s$ in terms of Besov spaces from [5], and Besov regularity theorems from [7,6], indicate that under very mild conditions the value of s is indeed only restricted by the polynomial order. An additional condition was required on the right-hand side, the discussion of which we postpone to the end of this introduction.

The key to obtain the optimal computational complexity result was the addition of a so-called coarsening or derefinement routine to the method from [10], that has to be applied after each fixed number of iterations, as well as, in view of the cost, to replace the exact Galerkin solvers by inexact ones. Thanks to the linear convergence of the method from [10], and the fact that after this coarsening, the underlying partition can be shown to have, up to some constant factor, the smallest possible cardinality in relation to the current error, optimal computational complexity could be shown.

The result of [4] is of great theoretical importance, but the adaptive method seems not very practical. The implementation of the coarsening procedure is not trivial, whereas, moreover, numerical results indicate that coarsening is not needed for obtaining an optimal method. In this paper, we will give a proof of this fact. We construct an adaptive finite element method, that, except that we solve the Galerkin systems inexactly, is very similar to the one from [10], and show that it has optimal computational complexity.

As in [4,10], we restrict ourselves to the model case of the Poisson equation in two space dimensions, linear finite elements, and partitions that are created by newest vertex bisection. Our results, however, rely on three ingredients only, two dealing with residual based a posteriori error estimators (Theorem 4.1, and Theorem 4.3 originated from [10]), and one dealing with bounding the number of bisections needed to find the smallest conforming refinement of a partition (Theorem 3.2, originated from [4]). The two results on a posteriori error estimators extend to more general second order elliptic differential operators, to more space dimensions, and to higher order finite elements. It can be expected that also Theorem 3.2 extends to more space dimensions, which, however, has to be investigated.

To solve a boundary value problem on a computer, it is indispensable to be able to approximate the right-hand side by some finite representation within a given tolerance. As (implicitly) in [10,4], we use piecewise constant approximations, but, in particular for higher order elements, by a modification of the adaptive refinement routine, piecewise polynomial approximations of higher order can be applied as well. Our aforementioned result concerning optimal computational complexity is valid only under the additional assumption that if the solution $u \in \mathcal{A}^s$, then for any n we know how to approximate the right-hand side f by a piecewise constant function with respect to a partition of n elements such that the error in the dual norm is of order n^{-s}. For $s \in (0, \frac{1}{2}]$, which is the relevant range for piecewise linear elements, we conjecture that if $u \in \mathcal{A}^s$, then such approximations for the corresponding right-hand side exist, which, however, is something different than knowing how to construct them. For $f \in L_2(\Omega)$, however, the additional assumption is always satisfied, where for constructing the approximations of the right-hand side we may even rely on uniform refinements.

The adaptive methods from [10,4] apply only to $f \in L_2(\Omega)$. Our additional assumption on the right-hand side is weaker than that of [4], but for $f \in H^{-1}(\Omega)$ not in $L_2(\Omega)$, it has to be verified for the right-hand side at hand.

This paper is organized as follows. In Sect. 2, we define the boundary value problem. Sect. 3 deals with newest vertex bisection. In Sect. 4, we derive or recall properties of a residual based a posteriori error estimator. To expose the main idea underlying the proof of optimal computational complexity, in Sect. 5 we consider an adaptive finite element method in the ideal situation that the right-hand side is piecewise constant with respect to the initial partition. We show that this method produces approximations with respect to partitions that, up to some constant factor, have minimal cardinality in view of the error. Finally, in Sect. 6, we consider a general applicable adaptive finite element method, and prove its optimal computational complexity.

In this paper, in order to avoid the repeated use of generic but unspecified constants, by $C \lesssim D$ we mean that C can be bounded by a multiple of D, independently of parameters which C and D may depend on, in some cases with the exception of the initial partition P_0, and the parameter s when it tends to 0 or ∞. Obviously, $C \gtrsim D$ is defined as $D \lesssim C$, and $C \eqsim D$ as $C \lesssim D$ and $C \gtrsim D$.

2 Boundary value problem

Let Ω be a polygonal bounded domain in $\mathbb{R}^2$. We consider the Poisson problem in variational form: Given $f \in H^{-1}(\Omega)\,(= H_0^1(\Omega)')$, find $u \in H_0^1(\Omega)$ such that

$$a(u, w) := \int_\Omega \nabla u \cdot \nabla w = f(w), \quad (w \in H_0^1(\Omega)). \tag{2.1}$$

Defining $L : H_0^1(\Omega) \to H^{-1}(\Omega)$ by $(Lu)(w) = a(u, w)$, (2.1) can be rewritten as

$$Lu = f.$$

For $f \in L_2(\Omega)$, we interpret $f(w)$ as $\int_\Omega f w$.

We will measure the error of any approximation for u in the energy norm

$$|w|_{H^1(\Omega)} = a(w, w)^{\frac{1}{2}}, \quad (w \in H_0^1(\Omega)),$$

with dual norm $\|f\|_{H^{-1}(\Omega)} := \sup_{0 \neq w \in H_0^1(\Omega)} \frac{|f(w)|}{|w|_{H^1(\Omega)}}$, $(f \in H^{-1}(\Omega))$. Equipped with these norms, L is an isomorphism between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$.

3 Finite element approximation

Let P_0 be a *fixed partition* of Ω, i.e., a collection of closed, essentially disjoint triangles, also called elements, with $\overline{\Omega} = \cup_{\triangle \in P_0} \triangle$, and $\#P_0 \geq 2$. We assume that P_0 is *conforming*, meaning that for any pair of different $\triangle, \triangle' \in P_0$, the intersection $\triangle \cap \triangle'$ is either empty, or a common edge or vertex.

Throughout this paper, we consider exclusively partitions that are created from P_0 by the so called *newest vertex bisection* method ([9]). To each $\triangle \in P_0$, we assign one of its vertices $v(\triangle)$ as its newest vertex. We consider refinements of P_0 by subdividing one or more $\triangle \in P_0$ by connecting $v(\triangle)$ to the midpoint of the edge of $\triangle$ opposite to $v(\triangle)$. This midpoint is assigned to both newly created triangles as their newest vertex. By applying this refinement rule recursively, we obtain an infinite set of possible partitions P, each of them being the leaves of a tree $T(P)$, which is a subtree of the infinite binary tree, having as roots the triangles of P_0, that corresponds to recursive bisections of all triangles. The number of nodes in $T(P)$ is not larger than $2\#P$. The *generation* of $\triangle \in P$ is the number $\text{gen}(\triangle)$ of ancestors it has in $T(P)$. We will call two different $\triangle, \triangle' \in P$ *neighbours* when $\triangle \cap \triangle'$ is an edge of $\triangle$ or $\triangle'$.

Generally, a partition is *nonconforming* meaning that there is a $\triangle \in P$ that contains a vertex of a $\triangle' \in P$ interior to an edge. Such a $\triangle$ will be said to contain a *hanging vertex*. Usually, we will use the symbols P^c, $\tilde{P}^c$, etc. to denote conforming partitions. We will assume that in P_0 the newest vertices are assigned in such a way that for any neighbours $\triangle, \triangle' \in P_0$ it holds that if $\triangle \cap \triangle'$ is opposite to $v(\triangle)$, then it is opposite to $v(\triangle')$, which, as shown in [4, Lemma 2.1], is possible for any P_0.

Below we describe an algorithm to construct a conforming refinement of a partition.

MAKECONF$[P] \to P$
`for` $\triangle \in P$ `do`
  when $\triangle$ *contains a hanging vertex, put* $\triangle$ *into* M*.*
`end do`
`while` $M \neq \emptyset$ `do`
  extract a $\triangle$ *from* M*, and bisect it. For both children of* $\triangle$*, check if it contains a hanging vertex, and if so, put this triangle into* M*. Inspect whether the bisection did not create a hanging vertex in a neighbour of* $\triangle$ *that did not already contain such a vertex, and if so put this neighbour into* M*.*
`end do`

The above algorithm terminates and yields a conforming refinement ([3]). Since the only way to remove a hanging vertex in a triangle is to bisect this triangle, all bisections made in the algorithm are unavoidable, showing that **MAKECONF**$[P]$ produces the *smallest* conforming partition that is a refinement of P.

Proposition 3.1 *If each partition* P *is stored by means of the tree* $T(P)$*, together with additional information that, for any* $\triangle \in P$*, allows to find its neighbours of the same generation in* $\mathcal{O}(1)$ *operations, then the algorithm can be implemented in such a way that the number of arithmetic operations required by the call* $P^c =$ **MAKECONF**$[P]$ *is bounded by an absolute multiple of* $\#P^c$*.*

Proof As shown in [4], a consequence of our assumption on the assignment of the newest vertices in P_0 is that when $\triangle$ has a neighbour $\triangle'$ in P with $\text{gen}(\triangle') - \text{gen}(\triangle) > 1$, then $\triangle$ has a hanging vertex. Since, on the other hand, $\triangle$ has no hanging vertex on $\triangle \cap \triangle'$ when $\text{gen}(\triangle') - \text{gen}(\triangle) < 0$, we conclude that the verification of the statement inside the first loop can be implemented in $\mathcal{O}(1)$ operations. The same reasoning shows that each iteration of the second loop can be implemented in $\mathcal{O}(1)$ operations, meaning that the total number of arithmetic operations can be bounded by an absolute multiple of the sum of $\#P$ and the number of bisections that were made, which sum is equal to $\#P^c$. $\square$

Our adaptive finite element method will produce a sequence of increasingly refined partitions created by, alternately, some elementary refinements to reduce the error in the current approximate solution induced by an error estimator, or such refinements to reduce the error in the current approximation of the right-hand side, and calls of **MAKECONF** to restore conformity which is needed for the application of the error estimator. To show optimality, it will be essential to bound the cardinality of the final partition in terms of the number of the above elementary refinements, that is, to know that all intermediate applications of **MAKECONF** will not essentially inflate the number of triangles. The following theorem is a direct consequence of [4, Theorem 2.4].

Theorem 3.2 *With $P_0^c := P_0$, for $i = 1, 2, \ldots$, let P_i be a refinement of P_{i-1}^c, and $P_i^c :=$ **MAKECONF**$[P_i]$. Then*

$$\#P_n^c - \#P_0^c \lesssim \sum_{i=1}^{n} \#P_i - \#P_{i-1}^c,$$

only dependent on P_0.

For a partition P, let $\mathcal{S}_P \subset H_0^1(\Omega)$ denote the space of continuous, *piecewise linear* functions subordinate to P which vanish at $\partial\Omega$. The solution $u_P \in \mathcal{S}_P$ of

$$a(u_P, w_P) = f(w_P), \quad (w_P \in \mathcal{S}_P), \tag{3.1}$$

is called a *Galerkin approximation* of the solution u of (2.1). Defining $L_P : \mathcal{S}_P \to (\mathcal{S}_P)' \supset H^{-1}(\Omega)$ by $(L_P u_P)(w_P) = a(u_P, w_P)$, it is given by $L_P^{-1} f$.

For approximating the right-hand side, we will make use of the spaces of *piecewise constants* subordinate to P, denoted as $\mathcal{S}_P^0$.

4 A residual based a posteriori error estimator

For a conforming partition P^c, let V_{P^c} and E_{P^c} be the set of its interior vertices and edges, respectively. For each $e \in E_{P^c}$, let P_e^c be the set of the two $\triangle \in P^c$ that have e as their common edge. For $f \in L_2(\Omega)$, $w_{P^c} \in \mathcal{S}_{P^c}$, we set

$$\eta_e(P^c, f, w_{P^c}) := \operatorname{diam}(e)\|[\nabla w_{P^c}]_e \cdot \mathbf{n}_e\|_{L_2(e)}^2 + \sum_{\triangle \in P_e^c} \operatorname{diam}(\triangle)^2 \|f\|_{L_2(\triangle)}^2,$$

where $\mathbf{n}_e$ is a unit vector orthogonal to e, and $[\nabla w_{P^c}]_e$ denotes the jump of ∇w_{P^c} in the direction of $\mathbf{n}_e$. We set the error estimator

$$\mathcal{E}(P^c, f, w_{P^c}) := [\sum_{e \in E_{P^c}} \eta_e(P^c, f, w_{P^c})]^{\frac{1}{2}}.$$

The following theorem deals with an easy generalization of a well-known result on a posteriori error estimators (Theorem 4.2, cf. [13,1]). This generalization concerns the fact that the difference between two Galerkin solutions with respect to different partitions is estimated, instead of the error in one Galerkin solution. Since the estimate will be essential for our analysis, for completeness, we include a proof.

Theorem 4.1 *Let $f \in L_2(\Omega)$, P^c be a conforming partition and $\tilde{P}$ be a refinement of P^c. With*

$$\overline{F} = \overline{F}(P^c, \tilde{P}) := \{e \in E_{P^c} : \exists \triangle' \in P^c,\ \triangle' \notin \tilde{P} \text{ with } \triangle' \cap \cup_{\triangle \in P^c_e} \triangle \neq \emptyset\},$$

see Figure 1, for $u_{P^c} = L_{P^c}^{-1} f$ and $u_{\tilde{P}} = L_{\tilde{P}}^{-1} f$, we have

$$|u_{\tilde{P}} - u_{P^c}|_{H^1(\Omega)} \leq C_1 [\sum_{e \in \overline{F}} \eta_e(P^c, f, u_{P^c})]^{\frac{1}{2}},$$

for some absolute constant $C_1 > 0$. Note that $\#\overline{F} \lesssim \#\tilde{P} - \#P^c$.

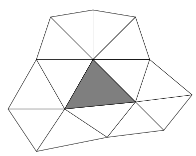

Fig. 1: Edges in $\overline{F}$ because of a refinement of the shaded triangle.

Proof Obviously, we have $|u_{\tilde{P}} - u_{P^c}|_{H^1(\Omega)} = \sup_{w_{\tilde{P}} \in \mathcal{S}_{\tilde{P}} \setminus \{0\}} \frac{|a(u_{\tilde{P}} - u_{P^c}, w_{\tilde{P}})|}{|w_{\tilde{P}}|_{H^1(\Omega)}}$, and $a(u_{\tilde{P}} - u_{P^c}, w_{P^c}) = 0$ for any $w_{P^c} \in \mathcal{S}_{P^c}$. For any $w_{\tilde{P}} \in \mathcal{S}_{\tilde{P}}$, $w_{P^c} \in \mathcal{S}_{P^c}$, integration by parts shows that

$$\begin{aligned}
&|a(u_{\tilde{P}} - u_{P^c}, w_{\tilde{P}})| \\
&\quad = |a(u_{\tilde{P}} - u_{P^c}, w_{\tilde{P}} - w_{P^c})| = |\int_\Omega f(w_{\tilde{P}} - w_{P^c}) - a(u_{P^c}, w_{\tilde{P}} - w_{P^c})| \\
&\quad = |\sum_{\triangle \in P^c} \{ \int_{\triangle \in P^c} f(w_{\tilde{P}} - w_{P^c}) - \int_{\partial \triangle} \partial_{\mathbf{n}} u_{P^c}(w_{\tilde{P}} - w_{P^c}) \}|, \\
&\quad \leq \sum_{\triangle \in P^c} \|f\|_{L_2(\triangle)} \|w_{\tilde{P}} - w_{P^c}\|_{L_2(\triangle)} && (4.1) \\
&\qquad\qquad + \sum_{e \in E_{P^c}} \|[\nabla u]_e \cdot \mathbf{n}_e\|_{L_2(e)} \|w_{\tilde{P}} - w_{P^c}\|_{L_2(e)}. && (4.2)
\end{aligned}$$

We select w_{P^c} to be the following interpolant of $w_{\tilde{P}}$. For any $v \in V_{P^c}$, choose a $\triangle_v \in P^c$ with $v \in \triangle_v$. As shown in [11, p. 17-18], there exists a $\omega(\triangle_v, v) \in L_\infty(\triangle_v)$ such that $\int_{\triangle_v} \omega(\triangle_v, v) p = p(v)$ for any $p \in P_1(\triangle_v)$, and $\|\omega(\triangle_v, v)\|_{L_\infty(\triangle_v)} \lesssim \text{meas}(\triangle_v)^{-1}$, independently of $\triangle_v$. We now define $w_{P^c} \in \mathcal{S}_{P^c}$ by $w_{P^c}(v) = \int_{\triangle_v} \omega(\triangle_v, v) w_{\tilde{P}}$, so that $|w_{P^c}(v)| \lesssim \mathrm{vol}(\triangle_v)^{-\frac{1}{2}} \|w_{\tilde{P}}\|_{L_2(\triangle_v)}$.

For each $\triangle \in P^c$, let $\Omega_\triangle = \cup_{\{\triangle' \in P^c : \triangle' \cap \triangle \neq \emptyset\}}$. By construction, we have

$$\|w_{\tilde{P}} - w_{P^c}\|_{L_2(\triangle)} \lesssim \|w_{\tilde{P}}\|_{L_2(\Omega_\triangle)}.$$

From a homogeneity argument, and either the fact that our interpolator reproduces any polynomial of first order degree, and, in particular, any constant, together with the Bramble-Hilbert lemma, or, in case one of the $\triangle'$ that form $\Omega_\triangle$ has an edge on $\partial\Omega$, the Poincaré-Friedrichs inequality, we infer that

$$\mathrm{diam}(\triangle)^{-1} \|w_{\tilde{P}} - w_{P^c}\|_{L_2(\triangle)} + |w_{\tilde{P}} - w_{P^c}|_{H^1(\triangle)} \lesssim |w_{\tilde{P}}|_{H^1(\Omega_\triangle)}. \quad (4.3)$$

For each $e \in E_{P^c}$ and both $\triangle \in P^c_e$, from the trace theorem and (4.3), we have

$$\begin{aligned}
\|w_{\tilde{P}} - & w_{P^c}\|_{L_2(e)} \\
&\lesssim \mathrm{diam}(e)^{-\frac{1}{2}} \|w_{\tilde{P}} - w_{P^c}\|_{L_2(\triangle)} + \mathrm{diam}(e)^{\frac{1}{2}} |w_{\tilde{P}} - w_{P^c}|_{H^1(\triangle)} \\
&\lesssim \mathrm{diam}(e)^{\frac{1}{2}} |w_{\tilde{P}}|_{H^1(\Omega_\triangle)} \qquad (4.4)
\end{aligned}$$

Noting that $w_{P^c}(v) = w_{\tilde{P}}(v)$ if all $\triangle \in P^c$ that contain v are also in $\tilde{P}$, we infer that terms in the sums (4.1) or (4.2) vanish for all $\triangle$ or e, respectively, that only share vertices with $\triangle' \in P^c \cap \tilde{P}$. By substituting (4.3) or (4.4) in these sums, applying the Cauchy-Schwarz inequality, and, since by assumption $\#P_0 \geq 2$, by using that any $\triangle \in P^c$ has an interior edge, the proof follows. □

By formally thinking of $H^1_0(\Omega)$ as being $\mathcal{S}_{\tilde{P}}$ with $\tilde{P}$ an infinite uniform refinement of P^c, the proof of Theorem 4.1 also yields the following well-known result (see, e.g., [13,1]).

Theorem 4.2 *For any $f \in L_2(\Omega)$ and any conforming partition P^c, with $u := L^{-1}f$ and $u_{P^c} := L^{-1}_{P^c} f$, we have*

$$|u - u_{P^c}|_{H^1(\Omega)} \leq C_1 \mathcal{E}(P^c, f, u_{P^c}).$$

Next we study whether the estimator also provides a lower bound for the error, or for the difference between two Galerkin solutions with respect to some conforming partition and a certain refinement. The following result was proven in [10, Lemma 4.2]. Note that here, and on more places, we restrict ourselves to piecewise constant right-hand sides. The case of having a general right-hand side will be discussed in §6.

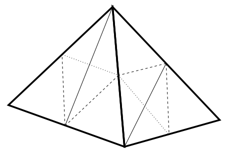

Fig. 2: A full refinement of both $\triangle \in P_e^c$. It creates vertices interior to both $\triangle \in P_e^c$ and e, which is essential to the proof of Theorem 4.3

Theorem 4.3 *Let P^c be a conforming partition and $\tilde{P}$ be a refinement of P^c. Let $w_{P^c} \in \mathcal{S}_{P^c}$, $f_{P^c} \in \mathcal{S}^0_{P^c}$, and let $u_{\tilde{P}} = L_{\tilde{P}}^{-1} f_{P^c}$ be the corresponding Galerkin solution. For an $e \in E_{P^c}$, assume that in $\tilde{P}$ both $\triangle \in P_e^c$ are replaced by at least 6 subtriangles by bisecting successively $\triangle$, both its children, and those two of its four grandchildren that have an edge in common with the common edge of both children of $\triangle$. As in [4], we will call such a refinement into 6 subtriangles a* full refinement*, see Figure 4.3. Then*

$$\sum_{\triangle \in P_e^c} |u_{\tilde{P}} - w_{P^c}|^2_{H^1(\triangle)} \gtrsim \eta_e(P^c, f_{P^c}, w_{P^c}).$$

As a corollary we obtain basically the converse of Theorem 4.1, assuming that the right-hand side is piecewise constant with respect to the current partition. In [10], it was demonstrated that such a result is not valid for general right-hand sides in $L_2(\Omega)$.

Corollary 4.4 *Let P^c be a conforming partition, and $f_{P^c} \in \mathcal{S}^0_{P^c}$. Let $\tilde{P}$ be a refinement of P^c such that for each $e \in \underline{F} \subset E_{P^c}$, both $\triangle \in P_e^c$ are refined by a full refinement. Then for $u_{\tilde{P}} = L_{\tilde{P}}^{-1} f_{P^c}$, and $w_{P^c} \in \mathcal{S}_{P^c}$, we have*

$$|u_{\tilde{P}} - w_{P^c}|_{H^1(\Omega)} \geq c_2 [\sum_{e \in \underline{F}} \eta_e(P^c, f_{P^c}, w_{P^c})]^{\frac{1}{2}},$$

for some absolute constant $c_2 > 0$. Note that $\#\tilde{P} - \#P^c \lesssim \#\underline{F}$.

Exploiting Galerkin orthogonality, from Corollary 4.4 one infers the following result.

Corollary 4.5 *Let P^c be a conforming partition, $f_{P^c} \in \mathcal{S}^0_{P^c}$. Then for $u = L^{-1} f_{P^c}$ and $w_{P^c} \in \mathcal{S}_{P^c}$, we have*

$$|u - w_{P^c}|_{H^1(\Omega)} \geq c_2 \mathcal{E}(P^c, f_{P^c}, w_{P^c}).$$

Finally in this section, we investigate the stability of the error estimator.

Proposition 4.6 *For a conforming partition P^c, $f \in L_2(\Omega)$, and $w_P, \tilde{w}_{P^c} \in \mathcal{S}_{P^c}$, we have*

$$c_2 |\mathcal{E}(P^c, f, w_{P^c}) - \mathcal{E}(P^c, f, \tilde{w}_{P^c})| \leq |w_{P^c} - \tilde{w}_{P^c}|_{H^1(\Omega)}.$$

Proof For $\tilde{f} \in L_2(\Omega)$, by two applications of the triangle inequality in the form $\big|\|\cdot\| - \|\cdot\|\big|^2 \leq \|\cdot - \cdot\|^2$, first for vectors and then for functions (cf. [12, Lemma 6.5], we have

$$|\mathcal{E}(P^c, f, w_{P^c}) - \mathcal{E}(P^c, \tilde{f}, \tilde{w}_{P^c})| \leq \mathcal{E}(P^c, f - \tilde{f}, w_{P^c} - \tilde{w}_{P^c}).$$

By substituting $\tilde{f} = f$, and by applying Corollary 4.5 the proof is completed. □

5 An idealized adaptive finite element method

For some *fixed*

$$\theta \in (0, 1],$$

we will make use of the following routine to determine a suitable adaptive refinement:

REFINE$[P^c, f_{P^c}, w_{P^c}] \to \tilde{P}$
% P^c is a conforming partition, $f_{P^c} \in \mathcal{S}^0_{P^c}$ and $w_{P^c} \in \mathcal{S}_{P^c}$.
Select, in $\mathcal{O}(\#P^c)$ operations, a set $\underline{F} \subset E_{P^c}$ with, up to some absolute factor, minimal cardinality *such that*

$$\sum_{e \in \underline{F}} \eta_e(P^c, f_{P^c}, w_{P^c}) \geq \theta^2\, \mathcal{E}(P^c, f_{P^c}, w_{P^c})^2. \tag{5.1}$$

Construct the partition $\tilde{P}$ from P^c by means of a full refinement of all $\triangle \in \{P^c_e : e \in \underline{F}\}$.

Selecting $\underline{F}$ that satisfies (5.1) with true minimal cardinality would require sorting all $e \in E_{P^c}$ by the values of $\eta_e = \eta_e(P^c, f_{P^c}, w_{P^c})$, which takes $\mathcal{O}(\#P^c \log(\#P^c))$ operations. Although it is unlikely that in applications the cost of this sorting, due to the log-factor, dominates the total cost, in order to give a full proof of our claim of optimal computational complexity, we recall a procedure with which this log-factor is avoided.

With $N := \#E_{P^c}$, we may discard all e with $\eta_e \leq (1-\theta^2)\mathcal{E}(P^c, f, w_{P^c})^2/N$. With $M := \max_{e \in E_{P^c}} \eta_e$, and q the smallest integer with $2^{-q-1}M \leq (1-\theta^2)\mathcal{E}(P^c, f, w_{P^c})^2/N$, we store the others in $q+1$ bins corresponding whether η_e is in $[M, \frac{1}{2}M)$, $[\frac{1}{2}M, \frac{1}{4}M), \ldots,$ or $[2^{-q}M, 2^{-q-1}M)$. We then build $\underline{F}$ by extracting edges from the bins, starting with the first bin, and when it got empty moving to the second bin and so on until (5.1) is satisfied. Let the resulting $\underline{F}$ now contains edges from the pth bin, but not from further bins. Then a minimal set $\tilde{\underline{F}}$ that satisfies (5.1) contains all edges from the bins up to the $(p-1)$th one. Since any two η_e for e in the pth bin differ at most a factor 2, we infer that the cardinality of the contribution from the pth bin to $\underline{F}$ is at most twice at large as that to $\tilde{\underline{F}}$, so that $\#\underline{F} \leq 2\#\tilde{\underline{F}}$. The number of operations and storage locations required by this procedure is $\mathcal{O}(q+\#P^c)$, with $q < \log_2(MN/[(1-\theta^2)\mathcal{E}(P^c,f,w_{P^c})^2]) \leq \log_2(N/(1-\theta^2)) \lesssim \log_2(\#P^c) < \#P^c$.

We define

$$\mathcal{A}^s = \{u \in H_0^1(\Omega) : |u|_{\mathcal{A}^s} := \sup_{n \geq \#P_0} n^s \inf_{\#P \leq n} \inf_{u_P \in \mathcal{S}_P} |u - u_P|_{H^1(\Omega)} < \infty\},$$

and equip it with norm $\|u\|_{\mathcal{A}^s} := |u|_{H^1(\Omega)} + |u|_{\mathcal{A}^s}$. So $\mathcal{A}^s$ is the class of functions that can be approximated by a continuous piecewise linear function with respect to a partition into n triangles created by newest vertex bisection with an error of order n^{-s}. An adaptive finite element method realizes optimal convergence rates if whenever $u \in \mathcal{A}^s$, it produces approximations with respect to partitions into n triangles with an error of order n^{-s}, and it has optimal computational complexity, if, in addition, it needs only $\mathcal{O}(n)$ arithmetical operations to produce such an approximation.

Although $\mathcal{A}^s$ is non-empty for any s, as it contains $\mathcal{S}_P$ for any partition, because we are approximating with piecewise linears in two dimensions, only for $s \leq \frac{1}{2}$ even for C^∞-functions membership in $\mathcal{A}^s$ is guaranteed, meaning that the classes for $s > \frac{1}{2}$ are less relevant. Classical estimates show that for $s \leq \frac{1}{2}$, $H^{1+2s}(\Omega) \cap H_0^1(\Omega) \subset \mathcal{A}^s$, where for $u \in H^{1+2s}(\Omega) \cap H_0^1(\Omega)$, the rate n^{-s} is already realized using uniform refinements. Obviously the class $\mathcal{A}^s$ contains many more functions, which is the reason to consider adaptive methods in the first place. A (near) characterization of $\mathcal{A}^s$ for $s \leq \frac{1}{2}$ in terms of Besov spaces can be found in [5].

To prove optimality of an adaptive algorithm based on the above routine **REFINE**, we will need that the constant θ satisfies

$$\theta \in \Big(0, \frac{c_2}{C_1}\Big),$$

which we will assume in the following. So the closer $\frac{c_2}{C_1}$ is to 1, i.e, the more 'efficient' is the, properly scaled, estimator, the larger is the fraction of the sum of the local estimators that can be used to induce refinements.

To express the main ideas, without being distracted by too many technical details, in the remainder of this section we consider the idealized situation that the right-hand side is piecewise constant with respect to any partition that we encounter, i.e., that it is piecewise constant with respect to the initial partition. Furthermore, we do not take the cost of the adaptive algorithm into account, and assume that the arising Galerkin systems are solved *exactly*.

The key to the proof that the adaptive algorithm produces a partition with, up to some constant factor, minimal cardinality is the following result.

Lemma 5.1 *Let $f \in \mathcal{S}^0_{P^c}$ such that, for some $s > 0$, $u := L^{-1}f \in \mathcal{A}^s$. Then for any conforming partition P^c, $u_{P^c} := L^{-1}_{P^c}f$, and $\tilde{P} :=$* **REFINE**$[P^c, f, u_{P^c}]$, *we have*

$$\#\tilde{P} - \#P^c \lesssim |u - u_{P^c}|_{H^1(\Omega)}^{-1/s} |u|_{\mathcal{A}^s}^{1/s},$$

only dependent on s when it tends to 0.

Proof Let $\lambda \in (0,1)$ be a constant with

$$\frac{c_2^2(1-\lambda^2)}{C_1^2} \geq \theta^2. \tag{5.2}$$

Suppose that $\breve{P}$ is a refinement of P^v such that $u_{\bar{P}} := L_{\bar{P}}^{-1} f$ satisfies

$$|u - u_{\bar{P}}|_{H^1(\Omega)} \leq \lambda |u - u_{P^c}|_{H^1(\Omega)}. \tag{5.3}$$

Then with $\overline{F} \subset E_{P^c}$ from Theorem 4.1, we have

$$\#\overline{F} \lesssim \#\breve{P} - \#P^c,$$

and

$$\begin{aligned} C_1^2 \sum_{e \in \overline{F}} \eta_e(P^c, f_{P^c}, u_{P^c}) &\geq |u_{\breve{P}} - u_{P^c}|^2_{H^1(\Omega)} \\ &= |u - u_{P^c}|^2_{H^1(\Omega)} - |u - u_{\breve{P}}|^2_{H^1(\Omega)} \\ &\geq (1-\lambda^2)|u - u_{P^c}|^2_{H^1(\Omega)} \\ &\geq (1-\lambda^2) c_2^2 \mathcal{E}(P^c, f, u_{P^c})^2, \end{aligned}$$

by Corollary 4.5, and so, by (5.2),

$$\sum_{e \in \overline{F}} \eta_e(P^c, f_{P^c}, u_{P^c}) \geq \theta^2 \mathcal{E}(P^c, f, u_{P^c})^2. \tag{5.4}$$

Since $\underline{F}$, determined in the call $\tilde{P} := \mathbf{REFINE}[P^c, f, u_{P^c}]$, is a set with, up to some absolute factor, *minimal* cardinality with the property (5.4), we infer that

$$\#\tilde{P} - \#P^c \lesssim \#\underline{F} \lesssim \#\overline{F} \lesssim \#\breve{P} - \#P^c.$$

Now let $\bar{P}$ be a smallest partition such that $u_{\bar{P}} := L_{\bar{P}}^{-1} f$ satisfies $|u - u_{\bar{P}}|_{H^1(\Omega)} \leq \lambda |u - u_{P^c}|_{H^1(\Omega)}$. Then

$$\#\bar{P} \leq \lambda^{-1/s} |u - u_{P^c}|^{-1/s}_{H^1(\Omega)} |u|^{1/s}_{\mathcal{A}^s}.$$

Taking $\breve{P}$ to be the smallest common refinement of P^c and $\bar{P}$, (5.3) is satisfied, and we conclude that

$$\#\tilde{P} - \#P^c \lesssim \#\breve{P} - \#P^c \leq \#\bar{P} \lesssim |u - u_{P^c}|^{-1/s}_{H^1(\Omega)} |u|^{1/s}_{\mathcal{A}^s}.$$

$\square$

We now consider the following adaptive algorithm:

SOLVE$[f,\varepsilon] \to [P_k^c, u_{P_k^c}]$
% For this preliminary version of the adaptive solver it is assumed
% that $f \in \mathcal{S}^0_{P_0}$
$P_0^c := P_0$; $u_{P_0^c} := L_{P_0^c}^{-1} f$; $k := 0$
`while` $C_1\mathcal{E}(P_k^c, f, u_{P_k^c}) \geq \varepsilon$ `do`
    $\tilde{P}_{k+1} :=$ **REFINE**$[P_k^c, f, u_{P_k^c}]$
    $P_{k+1}^c :=$ **MAKECONF**$[\tilde{P}_{k+1}]$
    $u_{P_{k+1}^c} := L_{P_{k+1}^c}^{-1} f$
    $k := k+1$
`end do`

Theorem 5.2 *Let $f \in \mathcal{S}^0_{P_0}$, then $[P^c, u_{P^c}] =$ **SOLVE**$[f,\varepsilon]$ terminates, and, with $u := L^{-1}f$, $|u - u_{P^c}|_{H^1(\Omega)} \leq \varepsilon$. If $u \in \mathcal{A}^s$, then $\#P^c - \#P_0 \lesssim \varepsilon^{-1/s}|u|_{\mathcal{A}^s}^{1/s}$, only dependent on P_0, and on s when it tends to 0 or ∞.*

Note that the given bound on $\#P^c$ as function of ε is, up to some constant factor, the *best* one can achieve in view of the assumption $u \in \mathcal{A}^s$.

Proof From

$$|u - u_{P_k^c}|^2_{H^1(\Omega)} = |u - u_{P_{k+1}^c}|^2_{H^1(\Omega)} + |u_{P_{k+1}^c} - u_{P_k^c}|^2_{H^1(\Omega)},$$

and, by Corollary 4.4, (5.1) and Theorem 4.2,

$$|u_{P_{k+1}^c} - u_{P_k^c}|_{H^1(\Omega)} \geq c_2\theta\mathcal{E}(P_k^c, f, u_{P_k^c}) \geq \frac{c_2\theta}{C_1}|u - u_{P_k^c}|_{H^1(\Omega)},$$

with $\mu := (1 - \frac{c_2^2\theta^2}{C_1^2})^{\frac{1}{2}} < 1$, we obtain

$$|u - u_{P_{k+1}^c}|_{H^1(\Omega)} \leq \mu|u - u_{P_k^c}|_{H^1(\Omega)}.$$

From Corollary 4.5 and again Theorem 4.2, we conclude the first two statements.

Let n be the value of k at termination. It is sufficient to consider $n \geq 1$, so that $C_1\mathcal{E}(P_{n-1}^c, f, u_{P_{n-1}^c}) > \varepsilon$. By Lemma 5.1, we have

$$\#\tilde{P}_{k+1} - \#P_k^c \lesssim |u - u_{P_k^c}|^{-1/s}_{H^1(\Omega)}|u|^{1/s}_{\mathcal{A}^s}, \quad 0 \leq k \leq n-1,$$

and so an application of Theorem 3.2 shows that

$$\#P_n^c - \#P_0 \lesssim |u|^{1/s}_{\mathcal{A}^s}\sum_{k=0}^{n-1}|u - u_{P_k^c}|^{-1/s}_{H^1(\Omega)}.$$

Using that $|u - u_{P_k^c}|^{-1/s}_{H^1(\Omega)} \leq \mu^{1/s}|u - u_{P_{k+1}^c}|^{-1/s}_{H^1(\Omega)}$, we end up with

$$\begin{aligned}\#P_n^c - \#P_0 &\lesssim |u|^{1/s}_{\mathcal{A}^s}|u - u_{P_{n-1}^c}|^{-1/s}_{H^1(\Omega)}\\ &\lesssim |u|^{1/s}_{\mathcal{A}^s}\mathcal{E}(P_{n-1}^c, f, u_{P_{n-1}^c})^{-1/s} \leq |u|^{1/s}_{\mathcal{A}^s}\varepsilon^{-1/s},\end{aligned}$$

where, for the last $\lesssim$ symbol, we applied Corollary 4.5. □

6 A practical adaptive finite element method

In this paper, we will deal in a somewhat different way than in [10] with the practical relevant situation that the right-hand side is *not* piecewise constant with respect to the initial partition. In [10], given P^c and w_{P^c}, and assuming that $f \in L_2(\Omega)$, the error estimator was applied to the triple (P^c, f, w_{P^c}). Instead of our Theorem 4.3, that can only be applied to piecewise constant right-hand sides, a more general version was presented giving a lower bound involving the term $\operatorname{osc}(f, P^c) := (\sum_{\triangle \in P^c} \operatorname{diam}(\triangle)^2 \|f - f_{P^c}\|^2_{L_2(\triangle)})^{\frac{1}{2}}$, with $f_{P^c} \in \mathcal{S}^0_{P^c}$ defined by $f_{P^c}|_\triangle = \int_\triangle f$ $(\triangle \in P^c)$, which term, called 'data oscillation', measures the difference between f and its the best piecewise constant approximation. A reduction of the error in the approximate solution because of an adaptive refinement could be shown when this data oscillation was small enough, and so the algorithm contained the possibility that additional refinements are made in order to reduce it. The Galerkin systems were set up using the right-hand side f, which, actually as with the evaluation of the error estimator, generally gives rise to quadrature errors.

In our approach, on any current partition P, we will *replace* f by a piecewise constant approximation f_P both to evaluate the error estimator, and to set up the Galerkin system. As in [10], the error of this approximation should be small enough, and our algorithm will include the possibility that additional refinements are made in order to reduce this error. With this approach, we do not have to deal with quadrature errors, and it will turn out that we have to control $f - f_P$ only in $H^{-1}(\Omega)$-norm, which is the natural norm to measure perturbations in the right-hand side. When $f \in L_2(\Omega)$, the $H^{-1}(\Omega)$-norm of the difference of f and its best approximation from $\mathcal{S}_P$ can be bounded by some absolute multiple of the data oscillation, and, what is more, we can consider $f \in H^{-1}(\Omega)$ not in $L_2(\Omega)$. On the other hand, one may argue that we have to *construct* the approximation f_P. Note, however, that when f is employed, implicitly a similar task is required for evaluating the error estimator or for setting up the right-hand side vector of a Galerkin system.

The following lemma generalizes upon Lemma 5.1, relaxing both the condition that the right-hand is piecewise constant with respect to the current partition and the assumption that we have the exact Galerkin solutions available, assuming that the deviations from that ideal situation are sufficiently small.

Lemma 6.1 *Let $\omega > 0$ be a constant with*

$$\frac{c_2}{C_1} - \frac{[C_1 c_2^{-1} + 1 + \sqrt{2}]\omega}{C_1} > \theta. \tag{6.1}$$

Then any conforming partition P^c, $f \in H^{-1}(\Omega)$, $u = L^{-1} f$, $f_{P^c} \in \mathcal{S}^0_{P^c}$, $w_{P^c} \in \mathcal{S}_{P^c}$, with

$$\|f - f_{P^c}\|_{H^{-1}(\Omega)} + |L^{-1}_{P^c} f_{P^c} - w_{P^c}|_{H^1(\Omega)} \leq \omega \mathcal{E}(P^c, f_{P^c}, w_{P^c}),$$

and $\tilde{P} := \mathbf{REFINE}[P^c, f_{P^c}, w_{P^c}]$, *we have*

$$\#\tilde{P} - \#P^c \lesssim |u - w_{P^c}|_{H^1(\Omega)}^{-1/s} |u|_{\mathcal{A}^s}^{1/s}$$

only dependent s *when it tends to* 0.

Proof We use the technique of the proof of Lemma 5.1, where f from that lemma should be read as f_{P^c}. We apply perturbation arguments to take into account that f and w_{P^c} are only approximations to f_{P^c} and $L_{P^c}^{-1} f_{P^c}$, respectively. Let $\lambda \in (0,1)$ be a constant with

$$\frac{c_2(1-2\lambda^2)^{\frac{1}{2}} - [C_1 c_2^{-1} + (1-2\lambda^2)^{\frac{1}{2}} + \sqrt{2}(1+\lambda)]\omega}{C_1} \geq \theta. \tag{6.2}$$

Let $\bar{P}$ be a smallest partition such that $u_{\bar{P}} := L_{\bar{P}}^{-1} f$ satisfies $|u - u_{\bar{P}}|_{H^1(\Omega)} \leq \lambda |u - w_{P^c}|_{H^1(\Omega)}$. Then

$$\#\bar{P} \leq \lambda^{-1/s} |u - w_{P^c}|_{H^1(\Omega)}^{-1/s} |u|_{\mathcal{A}^s}^{1/s}.$$

Let $\breve{P}$ be the smallest common refinement of P^c and $\bar{P}$, and let $u_{\breve{P}} := L_{\breve{P}}^{-1} f$, $\hat{u} := L^{-1} f_{P^c}$, $\hat{u}_{\breve{P}} := L_{\breve{P}}^{-1} f_{P^c}$, and $\hat{u}_{P^c} := L_{P^c}^{-1} f_{P^c}$.

We have

$$\begin{aligned}
|\hat{u} - \hat{u}_{\breve{P}}|_{H^1(\Omega)} &\leq |u - u_{\breve{P}}|_{H^1(\Omega)} + \|f - f_{P^c}\|_{H^{-1}(\Omega)} \\
&\leq \lambda |u - w_{P^c}|_{H^1(\Omega)} + \|f - f_{P^c}\|_{H^{-1}(\Omega)} \\
&\leq \lambda |\hat{u} - \hat{u}_{P^c}|_{H^1(\Omega)} + (1+\lambda)\|f - f_{P^c}\|_{H^{-1}(\Omega)} + \lambda |\hat{u}_{P^c} - w_{P^c}|_{H^1(\Omega)} \\
&\leq \lambda |\hat{u} - \hat{u}_{P^c}|_{H^1(\Omega)} + (1+\lambda)\omega \mathcal{E}(P^c, f_{P^c}, w_{P^c}) \\
&\leq \left[2\lambda^2 |\hat{u} - \hat{u}_{P^c}|_{H^1(\Omega)}^2 + 2(1+\lambda)^2 \omega^2 \mathcal{E}(P^c, f_{P^c}, w_{P^c})^2\right]^{\frac{1}{2}}.
\end{aligned}$$

With $\overline{F} = \overline{F}(P^c, \breve{P}) \subset E_{P^c}$ from Theorem 4.1, we obtain

$$\begin{aligned}
C_1^2 \sum_{e \in \overline{F}} \eta_e(P^c, f_{P^c}, \hat{u}_{P^c}) &\geq |\hat{u}_{\breve{P}} - \hat{u}_{P^c}|_{H^1(\Omega)}^2 \\
&= |\hat{u} - \hat{u}_{P^c}|_{H^1(\Omega)}^2 - |\hat{u} - \hat{u}_{\breve{P}}|_{H^1(\Omega)}^2 \\
&\geq (1-2\lambda^2)|\hat{u} - \hat{u}_{P^c}|_{H^1(\Omega)}^2 - 2(1+\lambda)^2\omega^2 \mathcal{E}(P^c, f_{P^c}, w_{P^c})^2 \\
&\geq (1-2\lambda^2) c_2^2 \mathcal{E}(P^c, f_{P^c}, \hat{u}_{P^c})^2 - 2(1+\lambda)^2 \omega^2 \mathcal{E}(P^c, f_{P^c}, w_{P^c})^2,
\end{aligned}$$

by Corollary 4.5, which can be applied because $f_{P^c} \in \mathcal{S}_{P^c}^0$. By two applications of $|\mathcal{E}(P^c, f_{P^c}, w_{P^c}) - \mathcal{E}(P^c, f_{P^c}, \hat{u}_{P^c})| \leq c_2^{-1} |w_{P^c} - \hat{u}_{P^c}|_{H^1(\Omega)} \leq c_2^{-1} \omega \mathcal{E}(P^c, f_{P^c}, w_{P^c})$ by Proposition 4.6 and the assumption made in the lemma, we have

$$\begin{aligned}
&c_2(1-2\lambda^2)^{\frac{1}{2}}\mathcal{E}(P^c,f_{P^c},w_{P^c})\\
&\leq c_2(1-2\lambda^2)^{\frac{1}{2}}\mathcal{E}(P^c,f_{P^c},\hat{u}_{P^c})+(1-2\lambda^2)^{\frac{1}{2}}\omega\mathcal{E}(P^c,f_{P^c},w_{P^c})\\
&\leq C_1\big[\sum_{e\in\overline{F}}\eta_e(P^c,f_{P^c},\hat{u}_{P^c})\big]^{\frac{1}{2}}+[(1-2\lambda^2)^{\frac{1}{2}}+\sqrt{2}(1+\lambda)]\omega\mathcal{E}(P^c,f_{P^c},w_{P^c})\\
&\leq C_1\big[\sum_{e\in\overline{F}}\eta_e(P^c,f_{P^c},w_{P^c})\big]^{\frac{1}{2}}\\
&\qquad\qquad+[C_1c_2^{-1}+(1-2\lambda^2)^{\frac{1}{2}}+\sqrt{2}(1+\lambda)]\omega\mathcal{E}(P^c,f_{P^c},w_{P^c}),
\end{aligned}$$

and so, by bringing the terms with $\mathcal{E}(P^c,f_{P^c},w_{P^c})$ to one side, from (6.2) we have

$$\theta^2\mathcal{E}(P^c,f_{P^c},w_{P^c})^2\leq\sum_{e\in\overline{F}}\eta_e(P^c,f_{P^c},w_{P^c}).$$

Since $\underline{F}$, determined in the call $\tilde{P}:=\mathbf{REFINE}[P^c,f,u_{P^c}]$, is a set with, up to some absolute factor, *minimal* cardinality with this property, we conclude that

$$\#\tilde{P}-\#P^c\lesssim\#\underline{F}\leq\#\overline{F}\lesssim\#\check{P}-\#P^c\leq\#\bar{P}\lesssim|u-w_{P^c}|_{H^1(\Omega)}^{-1/s}|u|_{\mathcal{A}^s}^{1/s}.$$

□

In the proof of Theorem 5.2, we saw that when P^c is a conforming partition, and $f\in\mathcal{S}^0_{P^c}$, then for $u_{P^c}:=L_{P^c}^{-1}f$, $\tilde{P}:=\mathbf{REFINE}[P^c,f,u_{P^c}]$ or a refinement of it, and $u_{\tilde{P}}:=L_{\tilde{P}}^{-1}f$, we have $|u-u_{\tilde{P}}|_{H^1(\Omega)}\leq(1-\frac{c_2^2\theta^2}{C_1^2})^{\frac{1}{2}}|u-u_{P^c}|_{H^1(\Omega)}$. In the next lemma, we show that such an error reduction is also valid when we have a general $f\in H^{-1}(\Omega)$, approximated by $f_{P^c}\in\mathcal{S}^0_{P^c}$ or $f_{\tilde{P}}\in H^{-1}(\Omega)$ on P^c or $\tilde{P}$, respectively, and when the resulting Galerkin systems are solved only inexactly, assuming that the deviation from the above ideal situation is sufficiently small. Actually, in view of a repeated application, we have in mind that $\tilde{P}$ is the result of a call of **MAKECONF** applied to the output of **REFINE**, and that $f_{\tilde{P}}\in\mathcal{S}^0_{\tilde{P}}$.

Lemma 6.2 *For all*

$$\mu\in\Big(\big[1-\frac{c_2^2\theta^2}{C_1^2}\big]^{\frac{1}{2}},1\Big),$$

there exists an $\omega=\omega(\mu,\theta,C_1,c_2)\in(0,c_2)$, *such that for any* $f\in H^{-1}(\Omega)$, *a conforming partition* P^c, $\tilde{P}=\mathbf{REFINE}[P^c,f_{P^c},w_{P^c}]$ *or a refinement of it,* $f_{P^c}\in\mathcal{S}^0_{P^c}$, $f_{\tilde{P}}\in H^{-1}(\Omega)$, $w_{P^c}\in\mathcal{S}_{P^c}$, *and* $w_{\tilde{P}}\in\mathcal{S}_{\tilde{P}}$, *with*

$$\left.\begin{array}{r}\|f-f_{P^c}\|_{H^{-1}(\Omega)}+|L_{P^c}^{-1}f_{P^c}-w_{P^c}|_{H^1(\Omega)}\\ \|f-f_{\tilde{P}}\|_{H^{-1}(\Omega)}+|L_{\tilde{P}}^{-1}f_{\tilde{P}}-w_{\tilde{P}}|_{H^1(\Omega)}\end{array}\right\}\leq\omega\mathcal{E}(P^c,f_{P^c},w_{P^c}),$$

we have

$$|u-w_{\tilde{P}}|_{H^1(\Omega)}\leq\mu|u-w_{P^c}|_{H^1(\Omega)}.$$

Proof Let $u := L^{-1}f$, $\hat{u} := L^{-1}f_{P^c}$, $\hat{u}_{\tilde{P}} := L_{\tilde{P}}^{-1}f_{P^c}$, and $\hat{u}_{P^c} := L_{P^c}^{-1}f_{P^c}$. From Theorem 4.2 and Proposition 4.6, we have

$$\begin{aligned}
|\hat{u} - w_{P^c}|_{H^1(\Omega)} &\leq |\hat{u} - \hat{u}_{P^c}|_{H^1(\Omega)} + |\hat{u}_{P^c} - w_{P^c}|_{H^1(\Omega)} \\
&\leq C_1\mathcal{E}(P^c, f_{P^c}, \hat{u}_{P^c}) + |\hat{u}_{P^c} - w_{P^c}|_{H^1(\Omega)} \\
&\leq C_1\mathcal{E}(P^c, f_{P^c}, w_{P^c}) + (1 + C_1 c_2^{-1})|\hat{u}_{P^c} - w_{P^c}|_{H^1(\Omega)} \\
&\leq [C_1 + \omega(1 + C_1 c_2^{-1})]\mathcal{E}(P^c, f_{P^c}, w_{P^c}),
\end{aligned}$$

and so, by Corollary 4.4,

$$|\hat{u}_{\tilde{P}} - w_{P^c}|_{H^1(\Omega)} \geq c_2\theta\mathcal{E}(P^c, f_{P^c}, w_{P^c}) \geq \frac{c_2\theta}{C_1 + \omega(1 + C_1 c_2^{-1})}|\hat{u} - w_{P^c}|_{H^1(\Omega)},$$

or

$$\begin{aligned}
|\hat{u} - \hat{u}_{\tilde{P}}|_{H^1(\Omega)} &= \left[|\hat{u} - w_{P^c}|^2_{H^1(\Omega)} - |\hat{u}_{\tilde{P}} - w_{P^c}|^2_{H^1(\Omega)}\right]^{\frac{1}{2}} \\
&\leq \left[1 - \left[\frac{c_2\theta}{C_1 + \omega(1 + C_1 c_2^{-1})}\right]^2\right]^{\frac{1}{2}}|\hat{u} - w_{P^c}|_{H^1(\Omega)}.
\end{aligned}$$

The proof is completed by the observations that

$$\begin{aligned}
|u - w_{\tilde{P}}|_{H^1(\Omega)} &\leq |u - \hat{u}|_{H^1(\Omega)} + |\hat{u} - \hat{u}_{\tilde{P}}|_{H^1(\Omega)} + |L_{\tilde{P}}^{-1}(f_{P^c} - f_{\tilde{P}})|_{H^1(\Omega)} \\
&\qquad + |L_{\tilde{P}}^{-1}f_{\tilde{P}} - w_{\tilde{P}}|_{H^1(\Omega)} \\
&\leq |\hat{u} - \hat{u}_{\tilde{P}}|_{H^1(\Omega)} + 2\|f - f_{P^c}\|_{H^{-1}(\Omega)} + \|f - f_{\tilde{P}}\|_{H^{-1}(\Omega)} \\
&\qquad + |L_{\tilde{P}}^{-1}f_{\tilde{P}} - w_{\tilde{P}}|_{H^1(\Omega)} \\
&\leq |\hat{u} - \hat{u}_{\tilde{P}}|_{H^1(\Omega)} + 3\omega\mathcal{E}(P^c, f_{P^c}, w_{P^c}), \\
|\hat{u} - w_{P^c}|_{H^1(\Omega)} &\leq |u - w_{P^c}|_{H^1(\Omega)} + \omega\mathcal{E}(P^c, f_{P^c}, w_{P^c}),
\end{aligned}$$

and

$$\begin{aligned}
|u - w_{P^c}|_{H^1(\Omega)} &\geq |\hat{u} - w_{P^c}|_{H^1(\Omega)} - \|f - f_{P^c}\|_{H^{-1}(\Omega)} \\
&\geq (c_2 - \omega)\mathcal{E}(P^c, f_{P^c}, w_{P^c}),
\end{aligned}$$

by Corollary 4.5. □

For solving the Galerkin systems approximately, we assume that we have an iterative solver of the following type available:

GALSOLVE$[P^c, f_{P^c}, u^{(0)}_{P^c}, \delta] \to \bar{u}_{P^c}$
% P^c is a conforming partition, $f_{P^c} \in (\mathcal{S}_{P^c})'$, and $u^{(0)}_{P^c} \in \mathcal{S}_{P^c}$, the
% latter being an initial approximation for an iterative solver.
% With $u_{P^c} := L_{P^c}^{-1}f_{P^c}$, the output $\bar{u}_{P^c} \in \mathcal{S}_{P^c}$ satisfies

$$|u_{P^c} - \bar{u}_{P^c}|_{H^1(\Omega)} \leq \delta.$$

% The call requires $\lesssim \max\{1, \log(\delta^{-1}|u_{P^c} - u_{P^c}^{(0)}|_{H^1(\Omega)})\}\#P^c$
% arithmetic operations.

Additive or multiplicative multigrid methods can be shown to to be of this type. A proof of this fact in our case of applying newest vertex bisection can be found in [14].

A second routine, called **RHS**, will be needed to find a piecewise constant approximation to the right-hand side f that is sufficiently accurate. Since this might not be possible with respect to the current partition, a call of **RHS** may result in a further refinement.

RHS$[P, f, \delta] \to [\tilde{P}, f_{\tilde{P}}]$
% P is a partition, $f \in H^{-1}(\Omega)$ and $\delta > 0$. The output consists of a
% $f_{\tilde{P}} \in \mathcal{S}^0_{\tilde{P}}$, where $\tilde{P}$ is P, or, if necessary, a refinement of it, such
% that $\|f - f_{\tilde{P}}\|_{H^{-1}(\Omega)} \leq \delta$.

Assuming that the solution $u \in \mathcal{A}^s$ for some $s > 0$, the cost of approximating the right-hand side f using a routine **RHS** will generally not dominate the other costs of our adaptive method only if there is some constant c_f such that for any $\delta > 0$ and any partition P, for $[\tilde{P}, f_{\tilde{P}}] := \mathbf{RHS}[P, f, \delta]$, it holds that

$$\#\tilde{P} - \#P \leq c_f^{1/s}\delta^{-1/s},$$

and the number of arithmetic operations required by the call is $\lesssim \#\tilde{P}$. We will call such a pair $(f, \mathbf{RHS})$ to be *s-optimal*. Obviously, given s, such a pair can only exist when $f \in \bar{\mathcal{A}}^s$, defined by

$$\bar{\mathcal{A}}^s = \{f \in H^{-1}(\Omega) : \sup_{n \geq \#P_0} n^s \inf_{\#P \leq n} \inf_{f_P \in \mathcal{S}^0_P} \|f - f_P\|_{H^{-1}(\Omega)} < \infty\}.$$

Classical estimates show that for $s \in (0, 1]$, $H^{2s-1}(\Omega) \subset \bar{\mathcal{A}}^s$, where for $f \in H^{2s-1}(\Omega)$ the rate n^{-s} is already realized by considering uniform refinements. For such f, the routine $[\tilde{P}, f_{\tilde{P}}] := \mathbf{RHS}[P, f, \delta]$ satisfying the assumption of s-optimality can be realized by taking $\tilde{P}$ to be the smallest common refinement of P and a uniform refinement $\hat{P}$ of P_0 with mesh-size $C_1\delta^{\frac{1}{2s}}$, and for $\triangle \in \tilde{P}$, taking $f_{\tilde{P}}|_\triangle = \int_\triangle f$, or an approximation of it within tolerance $C_2\delta^{\frac{1}{2s}}$, with C_1, C_2 being some suitable constants. Here, in addition, we need to assume that the evaluation of such $f_{\tilde{P}}|_\triangle$ requires not more than a constant number of operations, which can be satisfied assuming some piecewise smoothness of f.

Although for $f \in L_2(\Omega)$ above procedure realizes s-optimality for $s = \frac{1}{2}$, which covers the relevant range $s \in (0, \frac{1}{2}]$, based on $\|f - f_P\|_{H^{-1}(\Omega)} \lesssim \operatorname{osc}(f, P)$, a more efficient routine **RHS** might be obtained by running an adaptive algorithm for reducing $\operatorname{osc}(f, P)$, see [4, 10].

Obviously the class $\bar{\mathcal{A}}^s$ contains many more functionals f than those from $H^{2s-1}(\Omega)$. Yet, for $f \notin L_2(\Omega)$, the realization of the routine **RHS** has to depend on the functional at hand. In [12, Ex. 7.3], an example is given of a a pair $(f, \mathbf{RHS})$ that is s-optimal with $s = \frac{1}{2}$, where f is defined by the integral of its argument over a curve, which f is not in $L_2(\Omega)$.

We now have the ingredients in hand to define our adaptive finite element routine:

SOLVE$[f,\varepsilon] \to [P_k^c, u_{P_k^c}]$
% Let $\omega > 0$ *be a sufficiently small constant so that it satisfies* (6.1), *and,*
% for some $\mu \in ([1-\frac{c_2^2\theta^2}{C_1^2}]^{\frac{1}{2}}, 1)$*, so that* $\omega \le \omega(\mu,\theta,C_1,c_2)$ *as introduced*
% in Lemma 6.2.
% Let $\beta > 0$ *a constant not larger than* $[(2+C_1c_2^{-1})/2 + C_1/\omega]^{-1}$.
Select $\bar{\delta} \eqsim \|f\|_{H^{-1}(\Omega)}$*;* $P_0^c := P_0$*;* $w_{P_0^c} := 0$*;* $k := 0$*;* $\delta_0 := 2\bar{\delta}$
`do`
  `do` $\delta_k := \delta_k/2$
    $[P_k, f_{P_k}] :=$ **RHS**$[P_k, f, \delta_k/2]$
    $P_k^c :=$ **MAKECONF**$[P_k]$
    $w_{P_k^c} :=$ **GALSOLVE**$[P_k^c, f_{P_k}, w_{P_k^c}, \delta_k/2]$
    `if` $\eta_k := (2+C_1c_2^{-1})\delta_k/2 + C_1\mathcal{E}(P_k^c, f_{P_k}, w_{P_k^c}) \le \varepsilon$ `then stop`
    `end if`
  `until` $\delta_k \le \omega\mathcal{E}(P_k^c, f_{P_k}, w_{P_k^c})$.
  $P_{k+1} :=$ **REFINE**$[P_k^c, f_{P_k}, w_{P_k^c}]$
  $w_{P_{k+1}^c} := w_{P_k^c}$, $\delta_{k+1} := 2\beta\eta_k$, $k := k+1$
`end do`

The idea of this routine **SOLVE** is, preceding to a call of **REFINE**, to find P^c, $f_{P^c} \in \mathcal{S}_{P^c}^0$, and $w_{P^c} \in \mathcal{S}_{P^c}$, with

$$\|f - f_{P^c}\|_{H^{-1}(\Omega)} + |L_{P^c}^{-1} f_{P^c} - w_{P^c}|_{H^1(\Omega)} \le \omega\mathcal{E}(P^c, f_{P^c}, w_{P^c}), \qquad (6.3)$$

where ω satisfies the conditions of both Lemmas 6.1 and 6.2. Then by Lemmas 6.1, we have an optimal bound for the number of refinements that are made in **REFINE**, and, as we will see, because of the choice of the initial value $\delta_{k+1} = 2\beta\eta_k$, by an application of Lemma 6.2 the error in any following approximation for u produced in **SOLVE** is at least a factor μ smaller. Since both sides of (6.3) depend on P^c, f_{P^c}, and w_{P^c}, it is a priori not known how small the tolerances for f_{P^c} and w_{P^c} should be to satisfy it, explaining why the calls of **RHS** and **GALSOLVE** are put inside an inner-loop.

Thinking of the situation that $(f, \textbf{RHS})$ is $\tilde{s}$-optimal with $\tilde{s} > s$ for any s for which $u \in \mathcal{A}^s$, usually no refinements are made because of a call of **RHS**. Assuming this situation, and furthermore $\inf_{\bar{w}_{P_k^c} \in \mathcal{S}_{P_k^c}} |u - \bar{w}_{P_k^c}|_{H^1(\Omega)} \gtrsim \inf_{\bar{w}_{P_{k-1}^c} \in \mathcal{S}_{P_{k-1}^c}} |u - \bar{w}_{P_{k-1}^c}|_{H^1(\Omega)}$, that, however, although $\#P_k^c \lesssim \#P_{k-1}^c$, is not necessarily always valid, then by selecting β sufficiently small, one can show that this inner-loop terminates in the first iteration. In that case, the algorithm consists of a repetition of the sequence of calls of **REFINE**, **RHS**, **MAKECONF**, and **GALSOLVE**, which, except that we solve the Galerkin systems inexactly, is essentially the algorithm from [10].

Theorem 6.3 $[P^c, w_{P^c}] =$ **SOLVE**$[f,\varepsilon]$ *terminates, and, with* $u := L^{-1}f$, $|u - w_{P^c}|_{H^1(\Omega)} \le \varepsilon$*. If* $u \in \mathcal{A}^s$, $(f, \textbf{RHS})$ *is s-optimal, and* $\varepsilon \lesssim \|f\|_{H^{-1}}$*, then*

$\#P^c - \#P_0 \lesssim \varepsilon^{-1/s}(|u|_{\mathcal{A}^s}^{1/s} + c_f^{1/s})$, and the number of arithmetic operations and storage locations required by the call are bounded by some absolute multiple of $\varepsilon^{-1/s}(\|u\|_{\mathcal{A}^s}^{1/s} + c_f^{1/s})$. The constant factors involved in these bounds depend only on P_0, and on s when it tends to 0 or ∞.

Proof We are going to show that the sequence of approximations to u produced in **SOLVE** is majorized linearly convergent. We start with collecting some useful estimates. At evaluation of η_k, by Theorem 4.2 and Proposition 4.6, we have

$$
\begin{aligned}
|u - &w_{P_k^c}|_{H^1(\Omega)} \\
&\leq |u - L^{-1} f_{P_k}|_{H^1(\Omega)} + |(L^{-1} - L_{P_k^c}^{-1}) f_{P_k}|_{H^1(\Omega)} + |L_{P_k^c}^{-1} f_{P_k} - w_{P_k^c}|_{H^1(\Omega)} \\
&\leq \delta_k/2 + C_1 \mathcal{E}(P_k^c, f_{P_k}, L_{P_k^c}^{-1} f_{P_k}) + \delta_k/2 \\
&\leq ((2 + C_1 c_2^{-1})\delta_k/2 + C_1 \mathcal{E}(P_k^c, f_{P_k}, w_{P_k^c}) = \eta_k, \qquad (6.4)
\end{aligned}
$$

and, obviously,

$$
\delta_k \leq \frac{2}{2 + C_1 c_2^{-1}} \eta_k, \qquad (6.5)
$$

whereas when the subsequent until-clause fails, we know that

$$
\eta_k < 1 + C_1(c_2^{-1}/2 + \omega^{-1})\delta_k. \qquad (6.6)
$$

By Proposition 4.6 and Corollary 4.5, we have

$$
\begin{aligned}
\mathcal{E}(P_k^c, &f_{P_k}, w_{P_k^c}) \leq c_2^{-1}[|L_{P_k^c}^{-1} f_{P_k} - w_{P_k^c}|_{H^1(\Omega)} + |L^{-1} f_{P_k} - L_{P_k^c}^{-1} f_{P_k}|_{H^1(\Omega)}] \\
&c_2^{-1}[|L_{P_k^c}^{-1} f_{P_k} - w_{P_k^c}|_{H^1(\Omega)} + \|f - f_{P_k}\|_{H^{-1}(\Omega)} + |L^{-1} f - L_{P_k^c}^{-1} f|_{H^1(\Omega)}] \\
&\leq c_2^{-1}[\delta_k + |L^{-1} f - L_{P_k^c}^{-1} f|_{H^1(\Omega)}]. \qquad (6.7)
\end{aligned}
$$

So at the moment that the test in the until-clause is passed, because of $\delta_k \leq \omega \mathcal{E}(P_k^c, f_{P_k}, w_{P_k^c})$ and $c_2^{-1}\omega < 1$, by the assumption $\omega \leq \omega(\mu, \theta, C_1, c_2)$, we have

$$
\begin{aligned}
\eta_k &\leq ((2 + C_1 c_2^{-1})\omega/2 + C_1)\mathcal{E}(P_k^c, f_{P_k}, w_{P_k^c}) \qquad (6.8) \\
&\leq \tfrac{(2 + C_1 c_2^{-1})\omega/2 + C_1}{1 - c_2^{-1}\omega} \inf_{\bar{w}_{P_k^c} \in \mathcal{S}_{P_k^c}} |u - \bar{w}_{P_k^c}|_{H^1(\Omega)}, \qquad (6.9)
\end{aligned}
$$

in particular meaning that $\eta_k \lesssim |u - w_{P_k^c}|_{H^1(\Omega)}$ for the current $w_{P_k^c}$, and, by (6.4), that η_k can be bounded by some absolute multiple of any previously computed η_k.

At evaluation of $\mathbf{REFINE}[P_k^c, f_{P_k}, w_{P_k^c}]$, we know that $\|f - f_{P_k}\|_{H^{-1}(\Omega)} + |L_{P_k^c}^{-1} f_{P_k} - w_{P_k^c}|_{H^1(\Omega)} \leq \delta_k \leq \omega \mathcal{E}(P_k^c, f_{P_k}, w_{P_k^c})$. Because of (6.8) and the condition on β, the initial value of δ_{k+1} computed directly after **REFINE** satisfies $\delta_{k+1} \leq 2\omega \mathcal{E}(P_k^c, f_{P_k}, w_{P_k^c})$. Since furthermore $\omega \leq \omega(\mu, \theta, C_1, c_2)$, Lemma 6.2 shows that any newly computed $w_{P_{k+1}^c}$ in the next inner-loop satisfies

$$
|u - w_{P_{k+1}^c}|_{H^1(\Omega)} \leq \mu |u - w_{P_k^c}|_{H^1(\Omega)}. \qquad (6.10)
$$

Having above results, we claim that for any $\alpha < 1$, there exists a $K \in \mathbb{N}$ such that starting with some evaluation of η_k in **SOLVE**, within the K following evaluations its value is reduced by a factor α, where to prevent termination by the stop-statement we think of ε being 0. Indeed, let us fix some $\alpha < 1$, and consider some evaluation of η_k, say giving the value η. Then from (6.5), (6.6) and the geometric decrease of δ_k inside the inner-loop, we infer that within some fixed number of iterations of the same inner-loop the reduction with α is reached, unless the loop terminates earlier by the until-clause. In the latter case, after this termination, by the second noted consequence of (6.9), and (6.4), we have $|u - w_{P_k^c}|_{H^1(\Omega)} \lesssim \eta$. Because of (6.10), (6.9), and the definition of the initial value of δ_{k+1} directly after the call of **REFINE**, any subsequent inner-loop starts with a $\delta_k \lesssim \eta$, and so the previous reasoning shows that within a fixed number of iterations, it produces an $\eta_k \leq \alpha\eta$, again unless it terminates earlier by the until-clause. Finally, from (6.10) and (6.9), we infer that the number of inner-loops in which $\eta_k \leq \alpha\eta$ is not reached is uniformly bounded, proving our claim.

By (6.7), and the definition of the initial value of δ_0, the firstly computed $\eta_0 \lesssim \|f\|_{H^{-1}(\Omega)}$, and so the algorithm terminates with, by (6.4), $|u - w_{P_k^c}|_{H^1(\Omega)} \leq \varepsilon$, which completes the proof of the first two statements of the theorem.

Next, we are going to bound the cardinality of the output partition. We claim that at evaluation of the until-clause, we have

$$|u - w_{P_k^c}|_{H^1(\Omega)} \lesssim \delta_k. \tag{6.11}$$

Indeed, let $\bar{w}$ denote the previously computed approximation for u inside **SOLVE**. If the evaluation of the until-clause was the first one in this inner-loop, then either by $|u - \bar{w}|_{H^1(\Omega)} = \|f\|_{H^{-1}(\Omega)} \lesssim \bar{\delta}$ in case we are dealing with the first inner-loop, or by (6.4) and the initial value for δ_k, we have $|u - \bar{w}|_{H^1(\Omega)} \lesssim \delta_k$. In the other case, again because of (6.4) and the fact that apparently the previous evaluation of the until-clause failed, we have the same result. Now since $|u - w_{P_k^c}|_{H^1(\Omega)} \leq |u - L_{P_k^c}^{-1} f|_{H^1(\Omega)} + \delta_k$, and $L_{P_k^c}^{-1} f$ is the best approximation from $\mathcal{S}_{P_k^c}$ to u, our claim is shown.

By the assumption that $(f, \mathbf{RHS})$ is s-optimal, the number of refinements made by a call $\mathbf{RHS}[P_k, f, \delta_k/2]$ can be bounded by some absolute multiple of $\delta_k^{-1/s} c_f^{1/s}$. By the assumptions that $u \in \mathcal{A}^s$ and ω satisfies (6.1), and because at the moment of a call $\mathbf{REFINE}[P_k^c, f_{P_k}, w_{P_k^c}]$, it holds $\|f - f_{P_k}\|_{H^{-1}(\Omega)} + |L_{P_k^c}^{-1} f_{P^c} - w_{P_k^c}|_{H^1(\Omega)} \leq \omega \mathcal{E}(P_k^c, f_{P_k}, w_{P_k^c})$, Lemma 6.1 shows that the number of refinements made by this call can be bounded by some absolute multiple of $|u - w_{P_k^c}|_{H^1(\Omega)}^{-1/s} |u|_{\mathcal{A}^s}^{1/s}$.

In view of (6.11) and the geometric decrease of δ_k inside an inner-loop, we conclude that the total number of refinements by calls of **RHS** made in an inner-loop that terminates by the until-clause, and those made in the subsequent call of $\mathbf{REFINE}[P_k^c, f_{P_k^c}, w_{P_k^c}]$ can be bounded by some absolute multiple of $|u - w_{P^c}|_{H^1(\Omega)}^{-1/s} (|u|_{\mathcal{A}^s}^{1/s} + c_f^{1/s})$. For the case that the last inner-loop performed in **SOLVE** is not the first one, in view of (6.10), and the fact that the initial value of δ_k for this last inner-loop satisfies $\delta_k = \beta\eta_{k-1} \lesssim |u - w_{P^c_{k-1}}|_{H^1(\Omega)}$ by (6.9), we conclude that the total number of refinements by all calls of **RHS** and **REFINE** made in **SOLVE** can be bounded by some absolute multiple of $\delta_k^{-1/s}(|u|_{\mathcal{A}^s}^{1/s} + c_f^{1/s})$, where δ_k has its value at termination of **SOLVE**.

In case **SOLVE** terminates by the first evaluation of the test $\eta_k \leq \varepsilon$, then $\delta_k = \bar{\delta} \gtrsim \|f\|_{H^{-1}(\Omega)} \gtrsim \varepsilon$ follows by assumption. In the other case, because apparently the preceding test of this statement failed, we have either $\delta_k > \beta\varepsilon$ in case this test was evaluated in a preceding inner-loop, or $\delta_k > (2 + C_1(c_2^{-1} + 2\omega^{-1}))^{-1}\varepsilon$ when this test was evaluated in the same inner-loop, where we also used that the intermediate until-clause failed.

Since, by Theorem 3.2, all calls of **MAKECONF** increase the total number of refinements by not more than a constant factor, we conclude that for the output partition P^c,

$$\#P^c - \#P_0 \lesssim \varepsilon^{-1/s}(|u|_{\mathcal{A}^s}^{1/s} + c_f^{1/s}),$$

proving the third statement of the the theorem.

Finally, we have to bound the cost of the algorithm. The reasoning leading to (6.11) shows that at evaluation of $\delta_k := \delta_k/2$, $|u - w_{P^c_k}|_{H^1(\Omega)} \lesssim \delta_k$, so that after the calls of $\mathbf{RHS}[P_k, f, \delta_k/2]$ and $\mathbf{MAKECONF}[P_k]$ we have $|L^{-1}_{P^c_k} f_{P_k} - w_{P^c_k}|_{H^1(\Omega)} \lesssim \|f - f_{P_k}\|_{H^{-1}(\Omega)} + \delta_k \lesssim \delta_k$. We conclude that the call of $\mathbf{GALSOLVE}[P^c_k, f_{P_k}, w_{P^c_k}, \delta_k/2]$ requires $\mathcal{O}(\#P^c_k)$ operations, and so that a call of any of the subroutines **RHS**, **MAKECONF**, **GALSOLVE** or **REFINE** inside **SOLVE** requires a number of operations that is bounded by some absolute multiple of their (output) partition.

To prove the last statement, it is sufficient to consider ε of the form $\varepsilon_\ell := 2^{-\ell}\|f\|_{H^{-1}(\Omega)}$ ($\ell \in \mathbb{N}$). As we have seen, the firstly computed $\eta_0 \lesssim \|f\|_{H^{-1}(\Omega)}$. Since furthermore we showed for any $\alpha < 1$, there exists a K such that starting with some evaluation of η_k in **SOLVE**, within the K following evaluations its value is reduced by a factor α, we conclude that the call $\mathbf{SOLVE}[f, \varepsilon_0]$ terminates within some bounded number of evaluations of η_k, thus involving some bounded number of calls of the subroutines. Since the output partition of the call $\mathbf{SOLVE}[f, \varepsilon_0]$ satisfies $\#P^c - \#P_0 \lesssim \varepsilon_0^{-1/s}(|u|_{\mathcal{A}^s}^{1/s} + c_f^{1/s})$, and $\#P_0 \lesssim 1 = \varepsilon_0^{-1/s}\|f\|_{H^{-1}(\Omega)}^{1/s} = \varepsilon_0^{-1/s}|u|_{H^1(\Omega)}^{1/s} \leq \varepsilon_0^{-1/s}\|u\|_{\mathcal{A}^s}^{1/s}$, we conclude that the cost of this call can be bounded on some absolute multiple of $\varepsilon_0^{-1/s}(\|u\|_{\mathcal{A}^s}^{1/s} + c_f^{1/s})$.

As soon as $\eta_k \leq \varepsilon_\ell$ inside **SOLVE**, then within a bounded number of following evaluations of η_k, we have $\eta_k \leq \varepsilon_{\ell+1}$, involving a number of additional operations that can be bounded by a constant multiple of the cardinality of the output partition. Since this output partition satisfies $\#P^c - \#P_0 \lesssim \varepsilon_{\ell+1}^{-1/s}(|u|_{\mathcal{A}^s}^{1/s} + c_f^{1/s})$, and $\#P_0 \lesssim \|u\|_{\mathcal{A}^s}^{1/s}$, using induction we conclude that the cost of the call $\mathbf{SOLVE}[f, \varepsilon_{\ell+1}]$ can be bounded by some absolute multiple of $\varepsilon_{\ell+1}^{-1/s}(\|u\|_{\mathcal{A}^s}^{1/s} + c_f^{1/s})$, with which the proof of the theorem is completed. $\square$

References

1. M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis*. Pure and Applied Mathematics. Wiley-Interscience, New York, 2000.
2. I. Babuška and M. Vogelius. Feedback and adaptive finite element solution of one-dimensional boundary value problems. *Numer. Math.*, 44(1), 1984.
3. E. Bänsch. Local mesh refinement in 2 and 3 dimensions. *Impact Comput. Sci. Engrg.*, 3(3):181–191, 1991.
4. P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219 – 268, 2004.
5. P. Binev, W. Dahmen, R. DeVore, and P. Petruchev. Approximation classes for adaptive methods. *Serdica Math. J.*, 28:391–416, 2002.
6. S. Dahlke. Besov regularity for elliptic boundary value problems in polygonal domains. *Appl. Math. Lett.*, 12(6):31–36, 1999.
7. S. Dahlke and R. DeVore. Besov regularity for elliptic boundary value problems. *Comm. Partial Differential Equations*, 22(1 & 2):1–16, 1997.
8. W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996.
9. W. F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Trans. Math. Software*, 15(4):326–347, 1989.
10. P. Morin, R. Nochetto, and K. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488, 2000.
11. P. Oswald. *Multilevel finite element approximation: Theory and applications*. B.G. Teubner, Stuttgart, 1994.
12. R.P. Stevenson. An optimal adaptive finite element method. Technical Report 1271, Utrecht University, March 2003. To appear in *SIAM J. Numer. Anal.*
13. R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, Chichester, 1996.
14. H. Wu and Z. Chen. Uniform convergence of multigrid V-cycle on adaptively refined finite element meshes for second order elliptic problems. Research Report 2003-07, Institute of Computational Mathematics, Chinese Academy of Sciences, 2003.