A GOAL-ORIENTED ADAPTIVE FINITE ELEMENT METHOD WITH CONVERGENCE RATES – EXTENDED VERSION*

MARIO S. MOMMER † AND ROB STEVENSON ‡

Abstract. An adaptive finite element method is analyzed for approximating functionals of the solution of symmetric elliptic second order boundary value problems. We show that the method converges, and derive a favourable upper bound for its convergence rate and computational complexity. We illustrate our theoretical findings with numerical results.

AMS subject classifications. 65N30, 65N50, 65N15, 65Y20, 41A25

Key words. Adaptive finite element method, convergence rates, computational complexity, quantity of interest, a posteriori error estimators

1. Introduction.

Adaptive finite element methods (AFEMs) have become a standard tool for the numerical solution of partial differential equations. Although being successfully in use for more than 25 years, in more than one space dimension, even for the most simple case of symmetric elliptic equations of second order $a(u,v) = f(v)$ $(\forall v)$, their convergence was not demonstrated before the works of Dörfler ([Dör96]) and that of Morin, Nochetto and Siebert ([MNS00]). Convergence alone, however, does not show that the use of an AFEM for a solution that has singularities improves upon, or even competes with that of a non-adaptive FEM. Recently, after the derivation of such a result by Binev, Dahmen and DeVore ([BDD04]) for an AFEM extended with a so-called coarsening routine, in [Ste07] it was shown that standard AFEMs converge with the best possible rate in linear complexity.

The aforementioned works all deal with AFEMs in which the error is measured in the energy norm $\|\cdot\|_E := a(\cdot,\cdot)^{\frac{1}{2}}$. In many applications, however, one is not so much interested in the solution u as a whole, but rather in a (linear) *functional* $g(u)$ of the solution, often being referred to as a *quantity of interest*. With u_τ denoting the finite element approximation of u with respect to a partition τ, from $|g(u) - g(u_\tau)| \leq \|g\|_{E'}\|u - u_\tau\|_E$, obviously it follows that convergence of u_τ towards u with respect to $\|\cdot\|_E$ implies that of $g(u_\tau)$ towards $g(u)$ with at least the same rate. It is, however, generally observed that with adaptive methods especially designed for the approximation of this quantity of interest, known as *goal-oriented adaptive methods*, convergence of $g(u_\tau)$ towards $g(u)$ takes place at a higher rate. Examples of such methods can be found in the monographs [AO00, BR03, BS01], and in references cited there. So far these goal-oriented adaptive methods are usually not proven to converge. An exception is the method from [DKV06], however, in which adaptivity is purely driven by energy-norm minimalisation of the error in the *dual problem* $a(v,z) = g(v)$ $(\forall v)$. Another exception is the goal-oriented method from [MvSST06], which is proven to converge with a rate equal to what we will demonstrate (for piecewise linears), where, however, in [MvSST06] the strong assumption $u, z \in C^3(\overline{\Omega})$ was made.

*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.

†Interdisciplinary Center for Scientific Computing (IWR), Universität Heidelberg, Im Neuenheimer Feld 368, 69120 Heidelberg, Germany (mario.mommer@iwr.uni-heidelberg.de)

‡Korteweg - de Vries Institute for Mathematics, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands (R.P.Stevenson@uva.nl)

The starting point of our method is the well-known upper bound

$$|g(u) - g(u_\tau)| = |a(u - u_\tau, z - z_\tau)| \leq \|u - u_\tau\|_E \|z - z_\tau\|_E, \tag{1.1}$$

where z_τ is the finite element approximation with respect to τ of z. Having available an AFEM that is convergent with respect to the energy norm, in view of (1.1) an obvious approach would be to use it for finding partitions τ_p and τ_d such that the corresponding finite element approximations u_{τ_p} and z_{τ_d} have, say, both energy norm errors less than $\sqrt{\varepsilon}$. Indeed, then the product of the errors in primal and dual finite element approximations with respect to the smallest common refinement of τ_p and τ_d, and so the error in the approximation of the quantity of interest, is less than ε. This approach, however, would not benefit from the situation in which, quantitatively or qualitatively, either primal or dual solution is easier to approximate by finite element functions.

The alternative method we propose here works, in essence, as follows. On the k-th iteration, we start from a partition τ_k, and compute on it the solutions of the primal and dual problems. To advance the iteration, this partition is refined in such a way that the product $\|u - u_\tau\|_E \|z - z_\tau\|_E$ is reduced by a constant factor. To achieve this, we consider the effort needed to reduce each of $\|u - u_\tau\|_E$ and $\|z - z_\tau\|_E$ by the same constant factor, which we do by separately computing suitable refinement sets. The smallest of these sets is then applied to τ_k to obtain τ_{k+1}.

We can show that this method is convergent. In particular, we prove that if, for whatever $s, t > 0$, the solutions of the primal and dual problems can be approximated in energy norm to any accuracy $\delta > 0$ from partitions of cardinality $\mathcal{O}(\delta^{-1/s})$ or $\mathcal{O}(\delta^{-1/t})$, respectively, then given $\varepsilon > 0$, our method constructs a partition of cardinality $\mathcal{O}(\varepsilon^{-1/(s+t)})$ such that

$$|g(u) - g(u_\tau)| \leq \|u - u_\tau\|_E \|z - z_\tau\|_E \leq \varepsilon.$$

In view of the assumptions, this order of cardinality realizing $\|u - u_\tau\|_E \|z - z_\tau\|_E \leq \varepsilon$ is optimal. Moreover, by solving the arising linear systems only inexactly, we show that the overall cost of the algorithm is of order $\mathcal{O}(\varepsilon^{-1/(s+t)})$.

The convergence rate $s + t$ of our goal-oriented method is thus the sum of the rates s and t of the best approximations in energy norm for primal and dual problems. With the approach of approximating both primal and dual problem within tolerance $\sqrt{\varepsilon}$, the rate would be $2\min(s, t)$. Another alternative approach, namely, to solve each of the problems to an accuracy of $\varepsilon^{s/(s+t)}$ and $\varepsilon^{t/(s+t)}$, respectively, would also result in the rate $s + t$. This approach, however, is not feasible, since the values s and t are generally unknown. Our method converges at the rate $s + t$ without previous knowledge about the regularity of the solutions.

Concerning the value of s (and similarly t), when applying finite elements of order p, for s up to p/n, a rate s is guaranteed when the solution has "ns orders of smoothness" in $L_\tau(\Omega)$ for some $\tau > (\frac{1}{2} + s)^{-1}$ (instead of in $L_2(\Omega)$ required for non-adaptive approximation) (cf. [BDDP02]).

Our method is based on minimizing an *upper bound* for the error in the functional, which under circumstances can be crude. Actually, in all available goal-oriented adaptive methods the decision which elements have to be refined is based on some upper bound for the error. Different than for the error in energy norm, there exists no computable two-sided bound for the error in a functional of the solution. This leaves open the possibility that some bounds are "usually" sharper than others. An argument against the upper bound (1.1) brought up in [BR03] is that it is based on the application of a global Cauchy-Schwarz inequality, whereas the dual weighted residual method advocated there would better respect the local information.The contribution of the current paper is that we *prove* a rate that is generally observed with goal-oriented methods. When applying finite element spaces of equal order at primal and dual side, we neither expect (see Remark 5.1 for details), nor observe in our experiments, that on average our bound gets increasingly more pessimistic when the iteration proceeds.

This paper is organized as follows: In Sect. 2, we describe the model boundary value problem that we will consider. The finite element spaces and the refinement rules based on bisections of n-simplices are discussed in Sect. 3. In Sec. 4, we give results on residual based a posteriori energy error estimators. A cheaper estimator, consisting of face contributions only, applicable to lowest order elements is discussed in the appendix. In Sect. 5, we present our goal-oriented adaptive finite element method under the simplifying assumption that the right-hand sides of both primal and dual problem are piecewise polynomial with respect to the initial finite element partition. We derive the aforementioned bound on the cardinality of the output partition. In Sect. 6, the method is extended to general right-hand sides. By replacing the exact solutions of the arising linear systems by inexact ones, it is further shown that the required number of arithmetic operations and storage locations satisfies the same favourable bound as the cardinality of the output partition. Finally, in Sect. 7, we present numerical results obtained with the method. To apply our approach also to unbounded functionals, here we recall the use of extraction functionals, an approach introduced in [BS01].

In this paper, by $C \lesssim D$ we will mean that C can be bounded by a multiple of D, independently of parameters which C and D may depend on. Similarily, $C \gtrsim D$ is defined as $D \lesssim C$, and $C \eqsim D$ as $C \lesssim D$ and $C \gtrsim D$.

2. The model problem. Let $\Omega \subset \mathbb{R}^n$ be a polygonal domain. We consider the following model boundary value problem in variational form: Given $f \in H^{-1}(\Omega)$, find $u \in H^1_0(\Omega)$ such that

$$a(u,v) := \int_\Omega \boldsymbol{A}\nabla u \cdot \nabla v = f(v), \quad (v \in H^1_0(\Omega)), \tag{2.1}$$

where $\boldsymbol{A} \in L_\infty(\Omega)$ is a symmetric $n \times n$ matrix with $\operatorname{ess\,inf}_{x\in\Omega} \lambda_{\min}(\boldsymbol{A}(x)) > 0$. We assume that $\boldsymbol{A}$ is piecewise constant with respect to an initial finite element partition τ_0 of Ω specified below. To keep the exposition simple, we do not attempt to derive results that hold uniformly in the size of jumps of $\rho(\boldsymbol{A})$ over element interfaces, although, under some conditions, this is likely possible, cf. [Ste05]. For $f \in L_2(\Omega)$, we interpret $f(v)$ as $\int_\Omega fv$.

Given some $g \in H^{-1}(\Omega)$, we will be interested in $g(u)$. With $z \in H^1_0(\Omega)$ we will denote the solution of the dual problem

$$a(v,z) = g(v), \quad (v \in H^1_0(\Omega)). \tag{2.2}$$

We set the energy norm on $H^1_0(\Omega)$ and dual norm on $H^{-1}(\Omega)$ by

$$\|v\|_E = a(v,v)^{\frac{1}{2}} \text{ and } \|h\|_{E'} = \sup_{0\neq v\in H^1_0(\Omega)} \frac{|h(v)|}{\|v\|_E},$$

respectively.

3. Finite element spaces. Given an essentially disjoint subdivision τ of $\bar{\Omega}$ into (closed) n-simplices, called a partition, we will search approximations for u and z from the finite element space

$$\mathbb{V}_\tau := H_0^1(\Omega) \cap \prod_{T\in\tau} P_p(T),$$

where $0 < p \in \mathbb{N}$ is some fixed constant. For approximating the functionals f and g, we will make use of spaces

$$\mathbb{V}_\tau^* := \prod_{T\in\tau} P_{p-1}(T).$$

Although it is not a finite element space in the usual sense, we also use

$$\mathbb{W}_\tau^* := \prod_{T\in\tau} \{\boldsymbol{h} \in H(\mathrm{div}; T) : [\![\boldsymbol{h}\cdot\mathbf{n}]\!]_{\partial T} \in L_2(\partial T)\} \tag{3.1}$$

with $\mathbf{n}$ being a unit vector normal to ∂T, and $[\![\,]\!]_{\partial T}$ denoting the jump of its argument over ∂T in the direction of $\mathbf{n}$, defined to be zero on $\partial\Omega$. Obviously, $[\mathbb{V}_\tau^*]^n \subset \mathbb{W}_\tau^*$.

Below, we specify the type of (nested) partitions we will consider, and recall some results from [Ste08], generalizing upon known results for *newest vertex bisection* in two dimensions.

For $0 \le k \le n-1$, a (closed) simplex spanned by $k+1$ vertices of an n-simplex T is called a hyperface of T. For $k = n-1$, it will be called a *true hyperface.* A partition τ is called *conforming* when the intersection of any two different $T, T' \in \tau$ is either empty, or a hyperface of both simplices. Different simplices T, T' that share a true hyperface will be called *neighbours.* (Actually, when $\Omega \neq \mathrm{int}(\overline{\Omega})$ above definition of a conforming partition can be unnecessarily restrictive. We refer to [Ste08] for a discussion of this matter.)

Simplices will be refined by means of bisection. In order to guarantee uniform shape regularity of all descendants, a proper cyclic choice of the refinement edges should 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

$$(x_0, \tfrac{x_0+x_n}{2}, x_1, \ldots, x_\gamma, x_{\gamma+1}, \ldots, x_{n-1})_{(\gamma+1)\mathrm{mod}n}$$

and

$$(x_n, \tfrac{x_0+x_n}{2}, x_1, \ldots, x_\gamma, x_{n-1}, \ldots, x_{\gamma+1})_{(\gamma+1)\mathrm{mod}n},$$

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 $\overline{x_0 x_n}$ 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 $(\gamma+1)\mathrm{mod}n$. See Figure 3.1 for an illustration. This bisection process was introduced in [Tra97], and in different notations, in [Mau95]. The edge $\overline{x_0 x_n}$ 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

$$T_{\mathrm{R}} = (x_n, x_1, \ldots, x_\gamma, x_{n-1}, \ldots, x_{\gamma+1}, x_0)_\gamma,$$

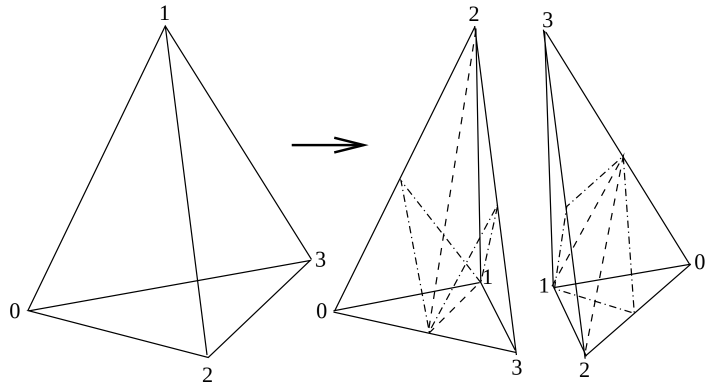

Fig. 3.1. *Bisection of a tagged tetrahedron of type* 0 *with the next two level cuts indicated.*

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 τ_0 of tagged simplices of some fixed type γ,

we will exclusively consider partitions that can be created from τ_0 *by recurrent bisections of tagged simplices*, for short, *descendants of* τ_0.

Simplices that can be created in this way are uniformly shape regular, only dependent on τ_0 and n. For the case that Ω might have slits, we assume that

$\partial\Omega$ *is the union of true hyperfaces of* $T \in \tau_0$.

We will assume that the simplices from τ_0 are tagged in a way such that any two neighbours $T = (x_0, \ldots, x_n)_\gamma$, $T' = (x'_0, \ldots, x'_n)_\gamma$ from P_0 *match* in the sense that if $\overline{x_0 x_n}$ or $\overline{x'_0 x'_n}$ is on $T \cap T'$, then *either* T and T' are reflected neighbours, 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 neighbouring children of T and T' are reflected neighbours. See Figure 3.2 for an illustration. It is known, see [BDD04] and the

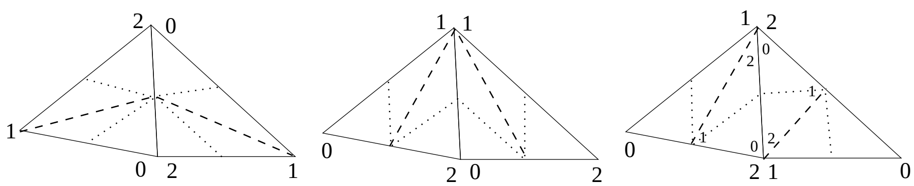

Fig. 3.2. *Matching neighbours for* $n = 2$, *and their level 1 and 2 descendants. The neighbours in the rightmost picture are not reflected neighbours, but the pair of their neighbouring children are.*

references therein, that for any conforming partition into triangles there exists a local numbering of the vertices so that the matching condition is satisfied. We do not now whether the corresponding statement holds in more space dimensions. Yet we showed that any conforming partition of n-simplices can be refined, inflating the number of simplices by not more than an absolute constant factor, into a conforming partition τ_0 that allows a local numbering of the vertices so that the matching condition is satisfied.

For applying a posteriori error estimators, we will need that the partitions τ underlying the approximation spaces are conforming. So in the following

τ, τ', $\hat{\tau}$ *etc. will always denote* conforming *partitions.*

Bisecting one or more simplices in a conforming partition τ generally results in a non-conforming partition ϱ. Conformity has to be restored by (recursively) bisecting any simplex $T \in \varrho$ that contains a vertex v of a $T' \in \varrho$ that does not coincides with any vertex of T (such a v is called a hanging vertex). This process, called completion, results into the smallest conforming refinement of ϱ.

Our adaptive method will be of the following form

`for` $j := 1$ `to` M
`do` *create some, possibly non-conforming refinement* ϱ_j *of* τ_{j-1}
complete ϱ_j *to its smallest conforming refinement* τ_j
`endfor`

As we will see, we will be able to bound $\sum_{j=1}^{M} \#\varrho_j - \#\tau_{j-1}$. Because of the additional bisections made in the completion steps, however, generally $\#\tau_M - \#\tau_0$ will be larger. The following crucial result, that relies on the matching condition in the initial partition, shows that these additional bisections inflate the total number of simplices by at most an absolute constant factor.

Theorem 3.1 (generalizes upon [BDD04, Theorem 2.4] for $n = 2$).

$$\#\tau_M - \#\tau_0 \lesssim \sum_{j=1}^{M} \#\varrho_j - \#\tau_{j-1},$$

only dependent on τ_0 and n, and in particular thus independently of M.

Remark 3.2. Note that this result in particular implies that any descendant ϱ of τ_0 has a conforming refinement τ with $\#\tau \lesssim \#\varrho$, only dependent on τ_0 and n. We end this section by introducing two more notations. For partitions τ', τ, we write $\tau' \supseteq \tau$ ($\tau' \supset \tau$) to denote that τ' is a (proper) refinement of τ. The smallest common refinement of τ and τ' will be denoted as $\tau \cup \tau'$.

4. A posteriori estimators for the energy error. Given a partition τ, and with u_τ denoting the solution in $\mathbb{V}_\tau$ of

$$a(u_\tau, v_\tau) = f(v_\tau), \quad (v_\tau \in \mathbb{V}_\tau), \tag{4.1}$$

in this section we discuss properties of the common residual based a posteriori error estimator for $\|u - u_\tau\|_E$. Since $a(\ ,\ )$ is symmetric, an analogous results will apply to $\|z - z_\tau\|_E$, with z_τ denoting the solution in $\mathbb{V}_\tau$ of

$$a(v_\tau, z_\tau) = g(v_\tau), \quad (v_\tau \in \mathbb{V}_\tau). \tag{4.2}$$

By formally viewing $H_0^1(\Omega)$ as $\mathbb{V}_\tau$ corresponding to the infinitely uniformly refined partition $\tau = \infty$, at some places we interpreted results derived for u_τ to hold for the solution u of (2.1) by substituting $\tau = \infty$.

For developing an adaptive finite element method that reduces the error in each iteration, it will be needed to approximate the right-hand side by discrete functions. Loosely speaking, in [MNS00] the error in this approximation is called *data oscillation.* Being on a partition τ, it will be allowed to use functions from $\mathbb{V}_\tau^* + \mathrm{div}[\mathbb{V}_\tau^*]^n$, where $\mathrm{div} := (-\nabla)' : L_2(\Omega)^n \to H^{-1}(\Omega)$. Depending on the right-hand side at hand, it might be more convenient to approximate it by functions from $\mathbb{V}_\tau^*$ or from $\mathrm{div}[\mathbb{V}_\tau^*]^n$, or by a combination of those. In view of this, we will write

$$f = f^1 + \mathrm{div}\boldsymbol{f}^2, \tag{4.3}$$

where $f^1 \in H^{-1}(\Omega)$ and $\boldsymbol{f}^2 \in L_2(\Omega)^n$ are going to be approximated by functions from $\mathbb{V}_\tau^*$ or from $\mathrm{div}[\mathbb{V}_\tau^*]^n$, respectively. Similarly, we write $g = g^1 + \mathrm{div}\boldsymbol{g}^2$.

Remark 4.1. Obviously, any $f \in H^{-1}(\Omega)$ can be written in the above form with vanishing $\boldsymbol{f}^2$. On the other hand, by taking $\boldsymbol{f}^2 = -\nabla w$ with $w \in H_0^1(\Omega)$ being the solution of $\int_\Omega \nabla w \cdot \nabla v = f(v)$ $(v \in H_0^1(\Omega))$, we see that we can equally well consider a vanishing f^1.

For $\bar{u}_\tau \in \mathbb{V}_\tau$, $\bar{f}^1 \in L_2(\Omega)$, and $\bar{\boldsymbol{f}}^2 \in \mathbb{W}_\tau^*$ (see (3.1)), where we have in mind approximations to u_τ, f^1, and $\boldsymbol{f}^2$, respectively, and $T \in \tau$, we set the local error indicator

$$\begin{aligned}\eta_T(\bar{f}^1, \bar{\boldsymbol{f}}^2, \bar{u}_\tau) := &\mathrm{diam}(T)^2 \|\bar{f}^1 + \nabla \cdot [\boldsymbol{A}\nabla \bar{u}_\tau + \bar{\boldsymbol{f}}^2]\|_{L_2(T)}^2 + \\ &\mathrm{diam}(T) \|[\![[\boldsymbol{A}\nabla \bar{u}_\tau + \bar{\boldsymbol{f}}^2] \cdot \mathbf{n}]\!]_{\partial T}\|_{L_2(\partial T)}^2.\end{aligned}$$

Note that the first term is the weighted local residual of the equation in strong form. We set the energy error estimator

$$\mathcal{E}(\tau, \bar{f}^1, \bar{\boldsymbol{f}}^2, \bar{u}_\tau) := [\sum_{T \in \tau} \eta_T(\bar{f}^1, \bar{\boldsymbol{f}}^2, \bar{u}_\tau)]^{\frac{1}{2}}.$$

The following Proposition 4.2 is a generalization of [Ste07, Theorem 4.1] valid for $\boldsymbol{A} = \mathrm{Id}$, $\boldsymbol{f}^2 = 0$, and polynomial degree $p = 1$. This result in turn was a generalization of [BMN02, Lemma 5.1(5.4)], see also [Ver96], in the sense that instead of $\|u - u_\tau\|_E$, the difference $\|u_{\tau'} - u_\tau\|_E$ for any $\tau' \supset \tau$ is estimated. Proposition 4.2 tells us that this difference can be bounded from above by the square root of the sum of the local error indicators corresponding to those simplices from τ that either are not in τ' since they were refined, or have non-empty intersection with such simplices. By taking $\tau' = \infty$, this result yields the known bound for $\|u - u_\tau\|_E$.

Proposition 4.2. *Let $\tau' \supset \tau$ be partitions, $f^1 \in L_2(\Omega)$, $\boldsymbol{f}^2 \in \mathbb{W}_\tau^*$ and*

$$G = G(\tau, \tau') := \{T \in \tau : T \cap \tilde{T} \neq \emptyset \textit{ for some } \tilde{T} \in \tau, \tilde{T} \notin \tau'\}.$$

Then we have

$$\|u_{\tau'} - u_\tau\|_E \leq C_1 \big[\sum_{T \in G} \eta_T(f^1, \boldsymbol{f}^2, u_\tau) \big]^{\frac{1}{2}},$$

for some absolute constant $C_1 > 0$. Note that $\#G \lesssim \#\tau' - \#\tau$.

In particular, by taking $\tau' = \infty$, we have

$$\|u - u_\tau\|_E \leq C_1 \mathcal{E}(\tau, f^1, \boldsymbol{f}^2, u_\tau). \tag{4.4}$$

Proof. We have $\|u_{\tau'} - u_\tau\|_E = \sup_{0 \neq v_{\tau'} \in \mathbb{V}_{\tau'}} \frac{|a(u_{\tau'} - u_\tau, v_{\tau'})|}{\|v_{\tau'}\|_E}$. For any $v_{\tau'} \in \mathbb{V}_{\tau'}$, $v_\tau \in \mathbb{V}_\tau$, we have

$$\begin{aligned}a(u_{\tau'} - u_\tau, v_{\tau'}) &= a(u_{\tau'} - u_\tau, v_{\tau'} - v_\tau) \\ &= \sum_T \int_T f^1(v_{\tau'} - v_\tau) - \boldsymbol{f}^2 \cdot \nabla(v_{\tau'} - v_\tau) - \boldsymbol{A}\nabla u_{\tau'} \cdot \nabla(v_{\tau'} - v_\tau) \\ &= \sum_T \big\{(f^1 + \nabla \cdot [\boldsymbol{A}\nabla u_\tau + \boldsymbol{f}^2])(v_{\tau'} - v_\tau) - \int_{\partial T} [\boldsymbol{A}\nabla u_\tau + \boldsymbol{f}^2] \cdot \mathbf{n}(v_{\tau'} - v_\tau)\big\},\end{aligned}$$

where the last line follows by integration by parts. By taking v_τ to be a suitable local quasi-interpolant of $v_{\tau'}$ as in [Ste07] (for $p > 1$, one may consult [KS08]) or, alternatively, a Clément type interpolator, and applying a Cauchy Schwarz inequality, one completes the proof. ▯

Remark 4.3. For the lowest order elements, i.e., $p = 1$, a statement similar to Proposition 4.2 is valid with error indicators consisting of the jump terms over the interfaces only. As a consequence, along the lines that we will follow for elements of general degree p, for $p = 1$ a cheaper goal-oriented adaptive finite element method can be developed that has similar properties. Details can be found in Appendix A.

Next we study whether the error estimator also provides a lower bound for $\|u - u_\tau\|_E$ and, when τ' is a sufficient refinement of τ, for $\|u_{\tau'} - u_\tau\|_E$. In order to derive such estimates, for the moment we further restrict the type of right-hand sides. The proof of the following proposition will be derived along the lines of the proof of [BMN02, Lemma 5.3] where the Stokes problem is considered (see also [MNS00, Lemma 4.2] for the case $p = 1$ and $\boldsymbol{f}^2 = 0$). For convenience of the reader we include it.

PROPOSITION 4.4. *Let $\tau \subset \tau'$ be partitions, and let $f^1 \in \mathbb{V}_\tau^*$, $\boldsymbol{f}^2 \in [\mathbb{V}_\tau^*]^n$, and $\bar{u}_\tau \in \mathbb{V}_\tau$.*

(a). *If $T \in \tau$ contains a vertex of τ' in its interior, then*

$$\operatorname{diam}(T)^2 \|f^1 + \nabla \cdot [\boldsymbol{A}\nabla \bar{u}_\tau + \boldsymbol{f}^2]\|_{L_2(T)}^2 \lesssim |u_{\tau'} - \bar{u}_\tau|_{H^1(T)}^2.$$

(b). *If a joint true hyperface e of $T_1, T_2 \in \tau$ contains a vertex of τ' in its interior, then*

$$\operatorname{diam}(e) \|[\![[\boldsymbol{A}\nabla \bar{u}_\tau + \boldsymbol{f}^2] \cdot \boldsymbol{n}]\!]_e\|_{L_2(e)}^2 \lesssim |u_{\tau'} - \bar{u}_\tau|_{H^1(T_1 \cup T_2)}^2 + \sum_{i=1}^{2} \operatorname{diam}(T_i)^2 \|f^1 + \nabla \cdot [\boldsymbol{A}\nabla \bar{u}_\tau + \boldsymbol{f}^2]\|_{L_2(T_i)}^2.$$

Proof. Let $\phi_T \in H_0^1(\Omega) \cap \prod_{T' \in \tau'} P_1(T')$ be the canonical nodal basis function associated to a vertex of τ' inside T. Writing $R_T = (f^1 + \nabla \cdot [\boldsymbol{A}\nabla \bar{u}_\tau + \boldsymbol{f}^2])|_T \in P_{d-1}(T)$, and $v_{\tau'} = R_T \phi_T \in \mathbb{V}_{\tau'}$, using that $\operatorname{supp} v_{\tau'} \subset T$, by integration by parts we get

$$\begin{aligned}\int_T R_T^2 \lesssim \int_T R_T^2 \phi_T = \int_T R_T v_{\tau'} &= (f_1 + \operatorname{div} \boldsymbol{f}^2)(v_{\tau'}) - \int_T \boldsymbol{A}\nabla \bar{u}_\tau \cdot \nabla v_\tau' \\ &= \int_T \boldsymbol{A}\nabla (u_{\tau'} - \bar{u}_\tau) \cdot \nabla v_{\tau'},\end{aligned}$$

and so by $|v_{\tau'}|_{H^1(T)} \lesssim \operatorname{diam}(T)^{-1} \|v_{\tau'}\|_{L_2(T)} \lesssim \operatorname{diam}(T)^{-1} \|R_T\|_{L_2(T)}$, we infer (a).

Let $\phi_e \in H_0^1(\Omega) \cap \prod_{T' \in \tau'} P_1(T')$ be the canonical nodal basis function associated to a vertex interior to e. Writing $J_e = [\![[\boldsymbol{A}\nabla \bar{u}_\tau + \boldsymbol{f}^2] \cdot \boldsymbol{n}]\!]_e \in P_{d-1}(e)$, let $\bar{J}_e \in P_{d-1}(T_1 \cup T_2)$ denote its extension constant in the direction normal to e, and let $v_{\tau'} = \bar{J}_e \phi_e \in \mathbb{V}_{\tau'}$. Using that $\operatorname{supp} v_\tau \subset T_1 \cup T_2$, by integration by parts we get

$$\int_e J_e^2 \lesssim \int_e J_e^2 \phi_e = \int_e J_e v_{\tau'} = \int_{T_1 \cup T_2} (\boldsymbol{A}\nabla \bar{u}_\tau + \boldsymbol{f}^2) \cdot \nabla v_{\tau'} + \int_{T_1 \cup T_2} \nabla \cdot (\boldsymbol{A}\nabla \bar{u}_\tau + \boldsymbol{f}^2) v_{\tau'}.$$

From

$$\int_{T_1 \cup T_2} \boldsymbol{f}^2 \cdot \nabla v_{\tau'} = -\operatorname{div} \boldsymbol{f}^2(v_{\tau'}) = -a(u_{\tau'}, v_{\tau'}) + \int_{T_1 \cup T_2} f^1 v_{\tau'},$$

we infer

$$\int_e J_e^2 \lesssim a(\bar{u}_\tau - u_{\tau'}, v_{\tau'}) + \int_{T_1\cup T_2} (f^1 + \nabla\cdot(\boldsymbol{A}\nabla\bar{u}_\tau + \boldsymbol{f}^2))v_{\tau'}$$
$$\lesssim \big[|\bar{u}_\tau - u_{\tau'}|_{H^1(T_1\cup T_2)}\operatorname{diam}(e)^{-1} + \sum_{i=1}^{2}\|R_{T_i}\|_{L_2(T_i)}\big]\|v_{\tau'}\|_{L_2(T_1\cup T_2)}.$$

Using that $\|v_{\tau'}\|_{L_2(T_1\cup T_2)} \eqsim \|\bar{J}_e\|_{L_2(T_1\cup T_2)} \eqsim \operatorname{diam}(e)^{\frac{1}{2}}\|J_e\|_{L_2(e)}$, we infer (b). ∎

In view of this last result, we will call a (possibly nonconforming) $\varrho\supset\tau$ *a full refinement with respect to* $T\in\tau$, when

T, and its neighbours in τ, as well as all true hyperfaces of T all contain a vertex of ϱ in their interiors.

As a direct consequence of Proposition 4.4 we have

Corollary 4.5. *Let τ be a partition, $f^1\in\mathbb{V}_\tau^*$, $\boldsymbol{f}^2\in[\mathbb{V}_\tau^*]^n$, and $\bar{u}_\tau\in\mathbb{V}_\tau$, and let $\tau'\supset\tau$ be a full refinement of τ with respect to all T from some $F\subset\tau$. Then*

$$c_2\big[\sum_{T\in F}\eta_T(f^1,\boldsymbol{f}^2,\bar{u}_\tau)\big]^{\frac{1}{2}} \le \|u_{\tau'}-\bar{u}_\tau\|_E, \tag{4.5}$$

for some absolute constant $c_2>0$. In particular, we have

$$c_2\mathcal{E}(\tau,f^1,\boldsymbol{f}^2,\bar{u}_\tau)\le\|u-\bar{u}_\tau\|_E. \tag{4.6}$$

Next, we investigate the stability of the energy error estimator.

Proposition 4.6. *Let τ be a partition, $f^1\in L_2(\Omega)$, $\boldsymbol{f}^2\in\mathbb{W}_\tau^*$, and $v_\tau,w_\tau\in\mathbb{V}_\tau$. Then*

$$c_2|\mathcal{E}(\tau,f^1,\boldsymbol{f}^2,v_\tau)-\mathcal{E}(\tau,f^1,\boldsymbol{f}^2,w_\tau)|\le\|v_\tau-w_\tau\|_E.$$

Proof. For $\tilde{f}^1\in L_2(\Omega)$ and $\tilde{\boldsymbol{f}}^2\in\mathbb{W}_\tau^*$ and $v_\tau,w_\tau\in\mathbb{V}_\tau$, by two applications of the triangle inequality in the form $\big|\|\cdot\|-\|\cdot\|\big|^2\le\|\cdot-\cdot\|^2$, first for vectors and then for functions, we have

$$|\mathcal{E}(\tau,f^1,\boldsymbol{f}^2,v_\tau)-\mathcal{E}(\tau,\tilde{f}^1,\tilde{\boldsymbol{f}}^2,w_\tau)|\le\mathcal{E}(\tau,f^1-\tilde{f}^1,\boldsymbol{f}^2-\tilde{\boldsymbol{f}}^2,v_\tau-w_\tau).$$

By substituting $\tilde{f}^1=f^1$ and $\tilde{\boldsymbol{f}}^2=\boldsymbol{f}^2$, and by applying (4.6) the proof is completed. ∎

5. An idealized goal-oriented adaptive finite element method. From (2.2), and $u-u_\tau\perp_{a(\,,\,)}\mathbb{V}_\tau\ni z_\tau$, we have

$$|g(u)-g(u_\tau)|=|a(u-u_\tau,z)|=|a(u-u_\tau,z-z_\tau)|\le\|u-u_\tau\|_E\|z-z_\tau\|_E. \tag{5.1}$$

We will develop an adaptive method for minimizing the right-hand side of this expression.

Remark 5.1. A question that naturally arises is whether there is something to be gained from using finite elements of different orders for the dual and the primal problem. Note that the derivation of (5.1) remains valid if the dual solution is computed in a lower order space, or for that matter, in any space that is a subspace of

$\mathbb{V}_\tau$. But this will result in a larger $\|z - z_\tau\|_E$, worsening our error estimate, without changing the actual error $|g(u) - g(u_\tau)|$.

And how about using a higher order space for the dual problem? In this case, (5.1) does not hold any longer. As $g(u) = f(z)$, we can approximate it by $f(z_\tau)$ with

$$|f(z) - f(z_\tau)| = |a(u, z - z_\tau)| = |a(u - u_\tau, z - z_\tau)| \leq \|u - u_\tau\|_E \|z - z_\tau\|_E. \quad (5.2)$$

Thus, as before, we obtain a worse error estimate as if we had used the same higher order space for the primal problem as well.

We conclude that with our approach there is no gain from using different orders, and, accordingly, will only consider here spaces of equal order.

Up to and including the following Lemma 5.3, we start with discussing a method for reducing $\|u - u_\tau\|_E$ or similarly $\|z - z_\tau\|_E$ separately. For some *fixed*

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

we will make use of the following routine to mark simplices for refinement:

MARK$[\tau, \bar{f}^1, \bar{\boldsymbol{f}}^2, \bar{u}_\tau] \to F$
% $\bar{f}^1 \in L_2(\Omega)$, $\bar{\boldsymbol{f}}^2 \in \mathbb{W}_\tau^*$, $\bar{u}_\tau \in \mathbb{V}_\tau$.
Select, in $\mathcal{O}(\#\tau)$ *operations, a set* $F \subset \tau$ *with, up to some absolute factor,* minimal cardinality *such that*

$$\sum_{T \in F} \eta_T(\bar{f}^1, \bar{\boldsymbol{f}}^2, \bar{u}_\tau) \geq \theta^2\, \mathcal{E}(\tau, \bar{f}^1, \bar{\boldsymbol{f}}^2, \bar{u}_\tau)^2. \quad (5.3)$$

Remark 5.2. Selecting F that satisfies (5.3) with truly minimal cardinality would require the sorting of all $\eta_T = \eta_T(\bar{f}^1, \bar{\boldsymbol{f}}^2, \bar{u}_\tau)$, which takes $\mathcal{O}(\#\tau \log(\#\tau))$ operations. The log-factor can be avoided by performing an approximate sorting based on binning that we recall here: With $N := \#\tau$, we may discard all $\eta_T \leq (1-\theta^2)\mathcal{E}(\tau, \bar{f}^1, \bar{\boldsymbol{f}}^2, \bar{u}_\tau)^2/N$. With $M := \max_{T \in \tau} \eta_T$, and q the smallest integer with $2^{-q-1}M \leq (1-\theta^2)\mathcal{E}(P^c, \bar{f}^1, \bar{\boldsymbol{f}}^2, w_{P^c})^2/N$, we store the others in $q+1$ bins corresponding whether η_T is in $[M, \frac{1}{2}M)$, $[\frac{1}{2}M, \frac{1}{4}M), \ldots$, or $[2^{-q}M, 2^{-q-1}M)$. Then we build F by extracting η_T from the bins, starting with the first bin, and when it got empty moving to the second bin and so on until (5.3) is satisfied. Let the resulting F now contains η_T from the ℓth bin, but not from further bins. Then a minimal set $\tilde{F}$ that satisfies (5.3) contains all η_T from the bins up to the $(\ell-1)$th one. Since any two η_T in the ℓth bin differ at most a factor 2, we infer that the cardinality of the contribution from the ℓth bin to F is at most twice at large as that to $\tilde{F}$, so that $\#F \leq 2\#\tilde{F}$. Assuming that each evaluation of η_T takes $\mathcal{O}(1)$ operations, the number of operations and storage locations required by this procedure is $\mathcal{O}(q + \#\tau)$, with $q < \log_2(MN/[(1-\theta^2)\mathcal{E}(\tau, \bar{f}^1, \bar{\boldsymbol{f}}^2, \bar{u}_\tau)^2]) \leq \log_2(N/(1-\theta^2)) \lesssim \log_2(\#\tau) < \#\tau$. The assumption on the cost of evaluating η_T is satisfied when $\bar{f}^1 \in \mathbb{V}_\tau^*$ and $\bar{\boldsymbol{f}}^2 \in [\mathbb{V}_\tau^*]^n$, as will be the case in our applications.

Having a set of marked elements F, the next step is to apply
REFINE$[\tau, F] \to \tau'$
% Determines the smallest $\tau' \supseteq \tau$ *which is a full refinement*
% with respect to all $T \in F$.

The cost of the call is $\mathcal{O}(\#\tau')$ operations.

Using the results on the a posteriori error estimator derived in the previous section, we have the following result:

LEMMA 5.3. *Let* $f^1 \in \mathbb{V}_\tau^*$, $\boldsymbol{f}^2 \in [\mathbb{V}_\tau^*]^n$. *Then for* $F = \mathbf{MARK}[\tau, f^1, \boldsymbol{f}^2, u_\tau]$, *and* $\tau' \supseteq \mathbf{REFINE}[\tau, F]$, *we have*

$$\|u - u_{\tau'}\|_E \le [1 - \tfrac{c_2^2\theta^2}{C_1^2}]^{\frac{1}{2}} \|u - u_\tau\|_E. \tag{5.4}$$

Furthermore

$$\#F \lesssim \#\hat{\tau} - \#\tau_0$$

for any partition $\hat{\tau}$ *for which*

$$\|u - u_{\hat{\tau}}\|_E \le [1 - \tfrac{C_1^2\theta^2}{c_2^2}]^{\frac{1}{2}} \|u - u_\tau\|_E.$$

Proof. Since this is a key result, for convenience of the reader we recall the arguments from [Ste07].

From

$$\|u - u_\tau\|_E^2 = \|u - u_{\tau'}\|_E^2 + \|u_{\tau'} - u_\tau\|_E^2,$$

and, by (4.5), (5.3) and (4.4),

$$\|u_{\tau'} - u_\tau\|_E \ge c_2\theta\mathcal{E}(\tau, f^1, \boldsymbol{f}^2, u_\tau) \ge \frac{c_2\theta}{C_1}\|u - u_\tau\|_E,$$

we conclude (5.4).

With $\hat{\tau}$ being a partition as in the statement of the theorem, let $\breve{\tau} = \tau \cup \hat{\tau}$. Then, as τ and $\hat{\tau}$, the partition $\breve{\tau}$ is a conforming descendant of τ_0, $\|u - u_{\breve{\tau}}\|_E \le \|u - u_{\hat{\tau}}\|_E$, and

$$\#\breve{\tau} - \#\tau \le \#\hat{\tau} - \#\tau_0.$$

To see the last statement, note that each simplex in $\breve{\tau}$ that is not in τ is in $\hat{\tau}$. Therefore, since $\tau \supset \tau_0$, the number of bisections needed to create $\breve{\tau}$ from τ, which number is equal to $\#\breve{\tau} - \#\tau$, is not larger than the number of bisections needed to create $\hat{\tau}$ from τ_0, which number is equal to $\#\hat{\tau} - \#\tau_0$.

With $G = G(\tau, \breve{\tau})$ from Proposition 4.2, we have

$$\begin{aligned} C_1^2 \sum_{T \in G} \eta_T(f^1, \boldsymbol{f}^2, u_\tau) &\ge \|u_{\breve{\tau}} - u_\tau\|_E^2 = \|u - u_\tau\|_E^2 - \|u - u_{\breve{\tau}}\|_E^2 \\ &\ge \tfrac{C_1^2\theta^2}{c_2^2}\|u - u_\tau\|_E^2 \ge C_1^2\theta^2\mathcal{E}(\tau, f^1, \boldsymbol{f}^2, u_\tau)^2, \end{aligned}$$

by (4.6). By construction of F, we conclude that

$$\#F \lesssim \#G \lesssim \#\breve{\tau} - \#\tau \le \#\hat{\tau} - \#\tau_0,$$

which completes the proof. ◻

The idea of the goal-oriented adaptive finite element method will be to mark sets of simplices for refinement corresponding to both primal and dual problems, and then to perform the actual refinement corresponding to that set of marked simplices that has the smallest cardinality. In order to assess the quality of the method, we first introduce the approximation classes $\mathcal{A}^s$.

For $s > 0$, we define

$$\mathcal{A}^s = \{u \in H_0^1(\Omega) : |u|_{\mathcal{A}^s} := \sup_{\varepsilon>0} \varepsilon \inf_{\{\tau : \|u-u_\tau\|_E \leq \varepsilon\}} [\#\tau - \#\tau_0]^s < \infty\}.$$

and equip it with norm $\|u\|_{\mathcal{A}^s} := \|u\|_E + |u|_{\mathcal{A}^s}$. So $\mathcal{A}^s$ is the class of functions that can be approximated within any given tolerance $\varepsilon > 0$ in $\| \, \|_E$ by a continuous piecewise polynomial of degree p on a partition τ with $\#\tau - \#\tau_0 \leq \varepsilon^{-1/s}|u|_{\mathcal{A}^s}^{1/s}$.

Remark 5.4. Although in the definition of $\mathcal{A}^s$ we only consider conforming descendants τ of τ_0, in view of Remark 3.2, we note that these approximation classes would remain the same if we would replace τ by any descendant ϱ of τ_0, conforming or not.

While the $\mathcal{A}^s$ contain $\mathbb{V}_\tau$ for any s, and thus are never empty, only the range $s \leq p/n$ is of interest, as even C^∞ functions are only guaranteed to belong to $\mathcal{A}^s$ for this range. Classical estimates show that for $s \leq p/n$, $H^{1+p}(\Omega) \cap H_0^1(\Omega) \subset \mathcal{A}^s$, where it is sufficient to consider uniform refinements. The class $\mathcal{A}^s$ is much larger than $H^{1+p}(\Omega) \cap H_0^1(\Omega)$, which is the reason to consider adaptive methods in the first place. A (near) characterization of $\mathcal{A}^s$ for $s \leq p/n$ in terms of Besov spaces can be found in [BDDP02] (although there the case $n = 2$ and $p = 1$ is considered, results easily generalize).

We now consider the following adaptive algorithm:

GOAFEM$[f^1, \boldsymbol{f}^2, g^1, \boldsymbol{g}^2, \varepsilon] \to [\tau_n, u_{\tau_n}, z_{\tau_n}]$
% For this preliminary version of the Goal-Oriented Adaptive Finite Element
% Method, it is assumed that $f^1, g^1 \in \mathbb{V}_{\tau_0}^*$ *and* $\boldsymbol{f}^2, \boldsymbol{g}^2 \in [\mathbb{V}_{\tau_0}^*]^n$.
$k := 0$
`while` $C_1\mathcal{E}(\tau_k, f^1, \boldsymbol{f}^2, u_{\tau_k}) \cdot C_1\mathcal{E}(\tau_k, g^1, \boldsymbol{g}^2, z_{\tau_k}) > \varepsilon$ `do`
    $F_p :=$ **MARK**$[\tau_k, f^1, \boldsymbol{f}^2, u_{\tau_k}]$
    $F_d :=$ **MARK**$[\tau_k, g^1, \boldsymbol{g}^2, z_{\tau_k}]$
    With F *being the smallest of* F_p *and* F_d, $\tau_{k+1} :=$ **REFINE**$[\tau_k, F]$
    $k := k + 1$
`end do`
$n{:=}k$

THEOREM 5.5. *Let* $f^1, g^1 \in \mathbb{V}_{\tau_0}^*$ *and* $\boldsymbol{f}^2, \boldsymbol{g}^2 \in [\mathbb{V}_{\tau_0}^*]^n$, *then* $[\tau_n, u_{\tau_n}, z_{\tau_n}] =$ **GOAFEM**$[f^1, \boldsymbol{f}^2, g^1, \boldsymbol{g}^2, \varepsilon]$ *terminates, and* $\|u - u_{\tau_n}\|_E \|z - z_{\tau_n}\|_E \leq \varepsilon$. *If* $u \in \mathcal{A}^s$ *and* $z \in \mathcal{A}^t$, *then*

$$\#\tau_n - \#\tau_0 \lesssim \varepsilon^{-1/(s+t)}(|u|_{\mathcal{A}^s}|z|_{\mathcal{A}^t})^{1/(s+t)},$$

only dependent on τ_0, *and on* s *or* t *when they tend to* 0 *or* ∞.

Remark 5.6. Only assuming that $u \in \mathcal{A}^s$ and $z \in \mathcal{A}^t$, given a partition τ, the generally smallest upper bound for the product of the errors in energy norm in primal and dual solution that can be expected is $[\#\tau - \#\tau_0]^{-s}|u|_{\mathcal{A}^s}[\#\tau - \#\tau_0]^{-t}|z|_{\mathcal{A}^t}$. Setting this expression equal to ε, one finds $\#\tau - \#\tau_0 = \varepsilon^{-1/(s+t)}(|u|_{\mathcal{A}^s}|z|_{\mathcal{A}^t})^{1/(s+t)}$. We conclude that the partition produced by **GOAFEM** is at most a constant factor larger than the generally smallest partition τ for which $\|u - u_\tau\|_E \|z - z_\tau\|_E$ is less than the prescribed tolerance.

Proof. Let $E_k := \|u - u_{\tau_k}\|_E \|z - z_{\tau_k}\|_E$. Then $E_{k+1} \leq [1 - \frac{c_2^2\theta^2}{C_1^2}]^{\frac{1}{2}} E_k$ by (5.4), and $c_2\mathcal{E}(\tau_k, f^1, \boldsymbol{f}^2, u_{\tau_k}) c_2\mathcal{E}(\tau_k, g^1, \boldsymbol{g}^2, z_{\tau_k}) \leq E_k$ by (4.6). So **GOAFEM**$[f^1, \boldsymbol{f}^2, g^1, \boldsymbol{g}^2, \varepsilon]$ terminates, with $E_n \leq C_1\mathcal{E}(\tau_n, f^1, \boldsymbol{f}^2, u_{\tau_n}) C_1\mathcal{E}(\tau_n, g^1, \boldsymbol{g}^2, z_{\tau_n}) \leq \varepsilon$ by (4.4).

With F_k being the set of marked cells inside the kth call of **REFINE**, Lemma 5.3 and the assumptions $u \in \mathcal{A}^s$, $z \in \mathcal{A}^t$ show that

$$\begin{aligned}
\#F_k &\leq \min\{[1-\tfrac{C_1^2\theta^2}{c_2^2}]^{-\frac{1}{2s}}\|u-u_{\tau_{k-1}}\|_E^{-1/s}|u|_{\mathcal{A}^s}^{1/s}, [1-\tfrac{C_1^2\theta^2}{c_2^2}]^{-\frac{1}{2t}}\|z-z_{\tau_{k-1}}\|_E^{-1/t}|z|_{\mathcal{A}^t}^{1/t}\} \\
&\lesssim \min\{\|u-u_{\tau_{k-1}}\|_E^{-1/s}|u|_{\mathcal{A}^s}^{1/s}, \|z-z_{\tau_{k-1}}\|_E^{-1/t}|z|_{\mathcal{A}^t}^{1/t}\} \\
&\leq \max_{\delta\eta \geq E_{k-1}} \min\{\delta^{-1/s}|u|_{\mathcal{A}^s}^{1/s}, \eta^{-1/t}|z|_{\mathcal{A}^t}^{1/t}\} = E_{k-1}^{-1/(s+t)}(|u|_{\mathcal{A}^s}|z|_{\mathcal{A}^t})^{1/(s+t)}.
\end{aligned}$$

The partition τ_k is smallest conforming refinement of the generally non-conforming ϱ_k, defined as the smallest refinement of τ_{k-1} which is a full refinement with respect to all $T \in F_k$. From Theorem 3.1, $\#\varrho_k - \#\tau_{k-1} \lesssim \#F_k$, the majorized linear convergence of $k \mapsto E_{k-1}$, and $E_{n-1} > \frac{c_2^2}{C_1^2}\varepsilon$, we conclude that

$$\#\tau_n - \#\tau_0 \lesssim \sum_{k=1}^{n} \#F_k \lesssim E_{n-1}^{-1/(s+t)}(|u|_{\mathcal{A}^s}|z|_{\mathcal{A}^t})^{1/(s+t)} \lesssim \varepsilon^{-1/(s+t)}(|u|_{\mathcal{A}^s}|z|_{\mathcal{A}^t})^{1/(s+t)}.$$

$\square$

6. A practical goal-oriented adaptive finite element method. So far, we assumed that $f = f^1 + \operatorname{div}\boldsymbol{f}^2$, $g = g^1 + \operatorname{div}\boldsymbol{g}^2$, with $f^1, g^1 \in \mathbb{V}_\tau^*$, $\boldsymbol{f}^2, \boldsymbol{g}^2 \in [\mathbb{V}_\tau^*]^n$ for any partition τ that we encountered, i.e., we assumed that $f^1, g^1 \in \mathbb{V}_{\tau_0}^*$, $\boldsymbol{f}^2, \boldsymbol{g}^2 \in [\mathbb{V}_{\tau_0}^*]^n$. From now on, given a partition τ, we will *approximate* $f, g \in H^{-1}(\Omega)$ by $f_{\tau'}^1 + \operatorname{div}\boldsymbol{f}_{\tau'}^2$, $g_{\tau'}^1 + \operatorname{div}\boldsymbol{g}_{\tau'}^2$, respectively, where $f_{\tau'}^1, g_{\tau'}^1 \in \mathbb{V}_{\tau'}^*$, $\boldsymbol{f}_{\tau'}^2, \boldsymbol{g}_{\tau'}^2 \in [\mathbb{V}_{\tau'}^*]^n$ and either $\tau' = \tau$, or, when it is needed to have a smaller approximation error, $\tau' \supset \tau$. We will set

$$f_{\tau'} := f_{\tau'}^1 + \operatorname{div}\boldsymbol{f}_{\tau'}^2, \quad g_{\tau'} := g_{\tau'}^1 + \operatorname{div}\boldsymbol{g}_{\tau'}^2.$$

To be able to distinguish between primal or dual solutions corresponding to different right-hand sides, we introduce operators $L : H_0^1(\Omega) \to H^{-1}(\Omega)$ by $(Lv)(w) = a(v,w)$ $(v, w \in H_0^1(\Omega))$, and $L_\tau : \mathbb{V}_\tau \to \mathbb{V}_\tau'$ by $(L_\tau v_\tau)(w_\tau) = a(v_\tau, w_\tau)$ $(v_\tau, w_\tau \in \mathbb{V}_\tau)$. The solutions u, z, u_τ, z_τ of (2.1), (2.2), (4.1), (4.2) can now be written as $L^{-1}f$, $(L')^{-1}g$, $L_\tau^{-1}f$, $(L_\tau')^{-1}g$, respectively. Since in our case $L' = L$ and $L_\tau' = L_\tau$, for notational convenience we will drop the prime. Note that $\|L\cdot\|_{E'} = \|\cdot\|_E$, $\|L_\tau^{-1}\|_{E' \to E} \leq 1$, and $\|(L^{-1} - L_\tau^{-1})\|_{E' \to E} \leq 1$.

Furthermore, in view of controlling the cost of our adaptive solver, from now on we will solve the arising Galerkin systems only approximately.

The following lemma generalizes upon Lemma 5.3, relaxing both the condition that the right-hand is in $\mathbb{V}_\tau^* + \operatorname{div}[\mathbb{V}_\tau^*]^n$ and the assumption that we have the exact Galerkin solution available, assuming that the deviations from that ideal situation are sufficiently small in a relative sense.

Lemma 6.1 ([Ste07, Lemmas 6.1, 6.2]). *There exist positive constants $\omega = \omega(\theta, C_1, c_2)$, and $\lambda = \lambda(\omega, C_1, c_2)$ such that for any $f \in H^{-1}(\Omega)$, partition τ, $f_\tau^1 \in \mathbb{V}_\tau^*$, $\boldsymbol{f}_\tau^2 \in [\mathbb{V}_\tau^*]^n$, $\bar{u}_\tau \in \mathbb{V}_\tau$ with*

$$\|f - f_\tau\|_{E'} + \|L_\tau^{-1}f_\tau - \bar{u}_\tau\|_E \leq \omega\mathcal{E}(\tau, f_\tau^1, \boldsymbol{f}_\tau^2, \bar{u}_\tau), \tag{6.1}$$

$F := \mathbf{MARK}[\tau, f_\tau^1, \boldsymbol{f}_\tau^2, \bar{u}_\tau]$ *satisfies*

$$\#F \lesssim \#\hat{\tau} - \#\tau_0$$

for any partition $\hat{\tau}$ for which

$$\|u - u_{\hat{\tau}}\|_E \leq \lambda\|u - \bar{u}_\tau\|_E.$$

Furthermore, given a

$$\mu \in \left(\left[1 - \tfrac{c_2^2\theta^2}{C_1^2}\right]^{\frac{1}{2}}, 1\right),$$

there exists an $\omega = \omega(\mu, \theta, C_1, c_2) > 0$*, such that if* (6.1) *is valid for this* ω*, and for* $\tau' \supseteq \mathbf{REFINE}[\tau, F]$*,* $f_{\tau'} \in H^{-1}(\Omega)$ *and* $\bar{u}_{\tau'} \in \mathbb{V}_{\tau'}$*,*

$$\|f - f_{\tau'}\|_{E'} + \|L_{\tau'}^{-1} f_{\tau'} - \bar{u}_{\tau'}\|_E \le \omega \mathcal{E}(\tau, f_\tau^1, \boldsymbol{f}_\tau^2, \bar{u}_\tau),$$

then

$$\|u - \bar{u}_{\tau'}\|_E \le \mu \|u - \bar{u}_\tau\|_E.$$

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

$\mathbf{GALSOLVE}[\tau, f_\tau, u_\tau^{(0)}, \delta] \to \bar{u}_\tau$
% $f_\tau \in (\mathbb{V}_\tau)'$*, and* $u_\tau^{(0)} \in \mathbb{V}_\tau$*, the latter being an initial approximation for an*
% iterative solver. The output $\bar{u}_\tau \in \mathbb{V}_\tau$ *satisfies*

$$\|L_\tau^{-1} f_\tau - \bar{u}_\tau\|_E \le \delta.$$

% The call requires $\lesssim \max\{1, \log(\delta^{-1}\|L_\tau^{-1} f_\tau - u_\tau^{(0)}\|_E)\}\#\tau$
% arithmetic operations.

Multigrid methods with local smoothing, or their additive variants (BPX) as preconditioners in Conjugate Gradients, are known to be of this type.

A routine called $\mathbf{RHS}_f$, and analogously $\mathbf{RHS}_g$, will be needed to find a sufficiently accurate approximation to the right-hand side f of the form $f_\tau^1 + \operatorname{div} \boldsymbol{f}_\tau^2$ with $f_\tau^1 \in \mathbb{V}_\tau^*$, $\boldsymbol{f}_\tau^2 \in [\mathbb{V}_\tau^*]^n$. Since this might not be possible with respect to the current partition, a call of $\mathbf{RHS}_f$ may result in a further refinement.

$\mathbf{RHS}_f[\tau, \delta] \to [\tau', f_{\tau'}^1, \boldsymbol{f}_{\tau'}^2]$
% $\delta > 0$*. The output consists of* $f_{\tau'}^1 \in \mathbb{V}_{\tau'}^*$*, and* $\boldsymbol{f}_{\tau'}^2 \in [\mathbb{V}_{\tau'}^*]^n$*, where* $\tau' = \tau$*, or,*
% if necessary, $\tau' \supset \tau$*, such that* $\|f - f_{\tau'}\|_{E'} \le \delta$*.*

Assuming that $u \in \mathcal{A}^s$ for some $s > 0$, the cost of approximating the right-hand side f using $\mathbf{RHS}_f$ 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 τ, for $[\tau', \cdot, \cdot] := \mathbf{RHS}_f[\tau, \delta]$, it holds that

$$\#\tau' - \#\tau \le c_f^{1/s} \delta^{-1/s},$$

and the number of arithmetic operations required by the call is $\lesssim \#\tau'$. We will call such a $\mathbf{RHS}_f$ to be *s-optimal* with constant c_f. Obviously, given s, such a routine can only exist when $f \in \bar{\mathcal{A}}^s$, defined by

$$\bar{\mathcal{A}}^s = \{f \in H^{-1}(\Omega) : \sup_{\varepsilon > 0} \varepsilon \inf_{\{\tau : \inf_{f_\tau^1 \in \mathbb{V}_\tau^*, \boldsymbol{f}_\tau^2 \in [\mathbb{V}_\tau^*]^n} \|f - f_\tau\|_{E'} \le \varepsilon\}} [\#\tau - \#\tau_0]^s < \infty\}.$$

On the one hand, $u \in \mathcal{A}^s$ implies that $f \in \bar{\mathcal{A}}^s$. Indeed, for any partition τ, let $\boldsymbol{f}_\tau^2 := -\boldsymbol{A}\nabla u_\tau$. Then $\boldsymbol{f}_\tau^2 \in [\mathbb{V}_\tau^*]^n$, and $\|f - \operatorname{div} \boldsymbol{f}_\tau^2\|_{E'} = \|u - u_\tau\|_E$. On the other hand, knowing that $f \in \bar{\mathcal{A}}^s$ is a different thing than knowing how to construct suitable approximations. If $s \in [\frac{1}{n}, \frac{p+1}{n}]$ and $f \in H^{sn-1}(\Omega)$, then the best approximations f_τ^1 to f from $\mathbb{V}_\tau^*$ with respect to $L_2(\Omega)$ using uniform refinements τ of τ_0 are known to converge with the required rate. For general $f \in \bar{\mathcal{A}}^s$, however, a realization of a suitable routine $\mathbf{RHS}_f$ has to depend on the functional f at hand.

Remark 6.2. When u and f are smooth, then $u \in \mathcal{A}^{p/n}$ and $f \in \bar{\mathcal{A}}^{(p+1)/n}$. Indeed, u is approximated by piecewise polynomials of degree p, and f by those of degree $p-1$ (apart from possible approximations from $\mathrm{div}[\mathbb{V}_\tau^*]^n$), whereas the errors are measured in $H_0^1(\Omega)$ or $H^{-1}(\Omega)$, respectively. Also for less smooth u and f, one can expect that usually $u \in \mathcal{A}^s$ and $f \in \bar{\mathcal{A}}^{s'}$ for some $s' > s$.

In our adaptive method, given some partition τ, for both computing the error estimator and setting up the Galerkin system, we will *replace* f by an approximation from $\mathbb{V}_{\tau'}^* + \mathrm{div}[\mathbb{V}_{\tau'}^*]^n$ where $\tau' \supseteq \tau$ (and similarly for g). This has the advantages that we can consider $f \notin L_2(\Omega) + \mathrm{div}\mathbb{W}_\tau^*$, for which thus the error estimator is not defined, and that we don't have to care about quadrature errors on various places in the algorithm.

Assuming $f \in L_2(\Omega) + \mathrm{div}\mathbb{W}_\tau^n$ for any τ, another option followed in [MNS00], is not to replace f by an approximation, but to check whether, on the current partition, the error in the best approximation for f from $\mathbb{V}_\tau^*(+\mathrm{div}[\mathbb{V}_\tau^*]^n)$, called *data oscillation*, is sufficiently small relative to the error in the current approximation to u, and, if not, to refine τ to achieve this. Convergence of this approach was shown, and it can be expected that by applying suitable quadrature and inexact Galerkin solves, optimal computational complexity can be shown as well. The observations at the beginning of this remark indicate that "usually", at least asymptotically, there will be no refinements needed to reduce the data oscillation. This explains why common adaptive methods that ignore data oscillation usually converge with optimal rates.

In addition to being s-optimal, we will have to assume that $\mathbf{RHS}_f$ is *linearly convergent*, with which we mean that for any $d \in (0,1)$, there exists a $D > 0$ such that for any $\delta > 0$, partitions τ and $\tau' \supseteq \hat{\tau}$ where $[\hat{\tau}, \cdot, \cdot] := \mathbf{RHS}_f[\tau, \delta]$, the output $[\tau'', \cdot, \cdot] := \mathbf{RHS}_f[\tau', d\delta]$ satisfies $\#\tau'' \leq D\#\tau'$.

Remark 6.3. Usually, a realization of $[\hat{\tau}, \cdot, \cdot] := \mathbf{RHS}_f[\tau, \delta]$ will be based on the selection of $\hat{\tau}$ such that an *upper bound* for the error is less than the prescribed tolerance. Since this upper bound will be an algebraically decreasing function of $\#\hat{\tau} - \#\tau_0$, linear convergence is obtained.

We now have the ingredients in hand to define our practical adaptive goal-oriented finite element routine **GOAFEM**. Compared to the idealized version from the previous section, we will have to deal with the fact that when solving the Galerkin systems only inexactly, and applying inexact right-hand sides, C_1 times the a posteriori error estimator $\mathcal{E}(\cdot)$ is not necessarily an upper bound for the energy norm of the error. We have to add corrections terms to obtain an upper bound. Furthermore, after applying **REFINE** at either primal or dual side, we have to specify a tolerance for the error in the new approximation of the right-hand side, and in that of the new approximate Galerkin solution. In order to know that a subsequent **REFINE** results in an error reduction, in view of Lemma 6.1 we would like to choose this tolerance smaller than ω times the new error estimator, that, however, is not known yet. Although we can expect that usually the new estimator is only some moderate factor less than the existing one, it cannot be excluded that the new estimator is arbitrarily small, e.g., when we happen to have reached a partition on which the solution can be exactly represented. In this case, an error reduction is immediate, and so we don't have to rely on **REFINE** to achieve it.

$\mathbf{GOAFEM}[f,g,\delta_p,\delta_d,\varepsilon] \to [\tau,\bar{u}_\tau,\bar{z}_\tau]$
% Let $\omega \in (0,c_2)$ be a constant not larger than the constants $\omega(\theta,C_1,c_2)$ and
% $\omega(\mu,\theta,C_1,c_2)$ for some $\mu \in ([1-\frac{c^2\theta^2}{C_1^2}]^{\frac{1}{2}},1)$ mentioned in Lemma 6.1.
% Let $0<\beta<[(\frac{2+3C_1c_2^{-1}}{2+C_1c_2^{-1}}+C_1c_2^{-1})(2+C_1(c_2^{-1}+2\omega^{-1}))]^{-1}$ be a constant.
$\tau := \tau_0$, $[\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}] := \mathbf{RHS}_f[\tau,\delta_p]$, $[\tau_d, g^1_{\tau_d}, \boldsymbol{g}^2_{\tau_d}] := \mathbf{RHS}_g[\tau,\delta_d]$
$\bar{u}_{\tau_p} := \bar{z}_{\tau_d} := 0$
`do`
  $\bar{u}_{\tau_p} := \mathbf{GALSOLVE}[\tau_p, f_{\tau_p}, \bar{u}_{\tau_p}, \delta_p]$
  $\bar{z}_{\tau_d} := \mathbf{GALSOLVE}[\tau_d, g_{\tau_d}, \bar{z}_{\tau_d}, \delta_d]$
  $\sigma_p := (2+C_1c_2^{-1})\delta_p + C_1\mathcal{E}(\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p})$
  $\sigma_d := (2+C_1c_2^{-1})\delta_d + C_1\mathcal{E}(\tau_d, g^1_{\tau_p}, \boldsymbol{g}^2_{\tau_p}, \bar{z}_{\tau_d})$
  `if` $\sigma_p\sigma_d \le \varepsilon$ `then` $\tau := \tau_p \cup \tau_d$, $\bar{u}_\tau := \bar{u}_{\tau_p}$, $\bar{z}_\tau := \bar{z}_{\tau_d}$ `stop endif`
  `if` $2\delta_p \le \omega\mathcal{E}(\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p})$ `then` $F_p := \mathbf{MARK}[\tau, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p}]$
  `else` $F_p := \emptyset$ `endif`
  `if` $2\delta_d \le \omega\mathcal{E}(\tau_d, g^1_{\tau_p}, \boldsymbol{g}^2_{\tau_p}, \bar{z}_{\tau_d})$ `then` $F_d := \mathbf{MARK}[\tau, g^1_{\tau_p}, \boldsymbol{g}^2_{\tau_p}, \bar{z}_{\tau_d}]$
  `else` $F_d := \emptyset$ `endif`
  `if` $\#\tau_p - \#\tau + \#F_p \le \#\tau_d - \#\tau + \#F_d$
  `then` $\tau := \mathbf{REFINE}[\tau_p, F_p]$, $\delta_p := \min(\delta_p, \beta\sigma_p)$
    $[\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}] := \mathbf{RHS}_f[\tau,\delta_p]$, $\tau_d := \tau \cup \tau_d$
  `else` $\tau := \mathbf{REFINE}[\tau_d, F_d]$, $\delta_d := \min(\delta_d, \beta\sigma_d)$
    $\tau_p := \tau \cup \tau_p$, $[\tau_d, g^1_{\tau_p}, \boldsymbol{g}^2_{\tau_p}] := \mathbf{RHS}_g[\tau,\delta_d]$
  `endif`
`enddo`

THEOREM 6.4. *$[\tau,\bar{u}_\tau,\bar{z}_\tau] = \mathbf{GOAFEM}[f,g,\underline{\delta}_p,\underline{\delta}_d,\varepsilon]$ terminates, and*

$$\|u-\bar{u}_\tau\|_E\|z-\bar{z}_\tau\|_E \le \varepsilon.$$

If $u \in \mathcal{A}^s$, $z \in \mathcal{A}^t$, $\mathbf{RHS}_f$ ($\mathbf{RHS}_g$) is s-optimal (t-optimal) with constant c_f (c_g), $\underline{\delta}_p > c_f$, and $\underline{\delta}_d > c_g$, then

$$\#\tau \lesssim \#\tau_0 + \varepsilon^{-1/(s+t)}\big[(|u|^{1/s}_{\mathcal{A}^s}+c_f^{1/s})^s(|z|^{1/t}_{\mathcal{A}^t}+c_g^{1/t})^t\big]^{1/(s+t)}.$$

If, additionally, $\|f\|_{E'} \lesssim \underline{\delta}_p$, $\|g\|_{E'} \lesssim \underline{\delta}_d$, and $\underline{\delta}_p\underline{\delta}_d \lesssim \|u-u_{\tau_0}\|_E\|z-z_{\tau_0}\|_E+\varepsilon$, then the number of arithmetic operations and storage locations required by the call are bounded by some absolute multiple of the same expression. The constant factors involved in these bounds may depend only on τ_0, and on s or t when they tends to 0 or ∞, and concerning the cost, on the constants involved in the additional assumptions.

Remark 6.5. The condition $\underline{\delta}_p > c_f$ implies that for a call $[\tau',\cdot,\cdot] = \mathbf{RHS}[\tau,\underline{\delta}_p]$, we have $\tau' = \tau$.

Proof. We start with collecting a few useful estimates. At evaluation of σ_p, by (4.4) and Proposition 4.6, we have

$$\begin{aligned}
\|u-\bar{u}_{\tau_p}\|_E &\le \|u-L^{-1}f_{\tau_p}\|_E + \|(L^{-1}-L^{-1}_{\tau_p})f_{\tau_p}\|_E + \|L^{-1}_{\tau_p}f_{\tau_p}-\bar{u}_{\tau_p}\|_E \\
&\le \delta_p + C_1\mathcal{E}(\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, L^{-1}_{\tau_p}f_{\tau_p}) + \|L^{-1}_{\tau_p}f_{\tau_p}-\bar{u}_{\tau_p}\|_E \\
&\le \delta_p + C_1\mathcal{E}(\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p}) + (C_1c_2^{-1}+1)\|L^{-1}_{\tau_p}f_{\tau_p}-\bar{u}_{\tau_p}\|_E \\
&\le (2+C_1c_2^{-1})\delta_p + C_1\mathcal{E}(\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p}) =: \sigma_p, \qquad (6.2)
\end{aligned}$$

and, by Corollary 4.5,

$$\begin{aligned}\mathcal{E}(\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p}) &\leq c_2^{-1}\|L^{-1}f_{\tau_p} - \bar{u}_{\tau_p}\|_E \\ &\leq c_2^{-1}[\|u - u_{\tau_p}\|_E + \|(L^{-1} - L^{-1}_{\tau_p})(f_{\tau_p} - f)\|_E + \|L^{-1}_{\tau_p} f_{\tau_p} - \bar{u}_{\tau_p}\|_E] \\ &\leq c_2^{-1}\|u - u_{\tau_p}\|_E + c_2^{-1} 2\delta_p. \end{aligned} \tag{6.3}$$

So if $2\delta_p \leq \omega\mathcal{E}(\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p})$, then $\mathcal{E}(\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p}) \leq [c_2 - \omega]^{-1}\|u - u_{\tau_p}\|_E$, and so

$$\sigma_p \leq D\|u - u_{\tau_p}\|_E \quad \text{ where } D := \tfrac{(1+\frac{1}{2}C_1 c_2^{-1})\omega + C_1}{c_2 - \omega}. \tag{6.4}$$

Now we are ready to show majorized linear convergence of $\sigma_p\sigma_d$. Consider any two instances $\sigma_p^{(A)}$ and $\sigma_p^{(B)}$ of σ_p, where $\sigma_p^{(A)}$ has been computed preceding to $\sigma_p^{(B)}$. With $\delta_p^{(A)}$, $\delta_p^{(B)}$ and $\tau_p^{(A)}$, $\tau_p^{(B)}$ being the corresponding tolerances and partitions, from (6.3), $\delta_p^{(B)} \leq \delta_p^{(A)}$, and $\tau_p^{(B)} \supseteq \tau_p^{(A)}$, and so $\|u - u_{\tau_p^{(B)}}\|_E \leq \|u - \bar{u}_{\tau_p^{(A)}}\|_E \leq \sigma_p^{(A)}$ by (6.2), we have

$$\begin{aligned}\sigma_p^{(B)} &= (2 + C_1 c_2^{-1})\delta_p^{(B)} + C_1\mathcal{E}(\tau_p^{(B)}, f^1_{\tau_p^{(B)}}, \boldsymbol{f}^2_{\tau_p^{(B)}}, \bar{u}_{\tau_p^{(B)}}) \\ &\leq (2 + 3C_1 c_2^{-1})\delta_p^{(A)} + C_1 c_2^{-1}\sigma_p^{(A)} \\ &\leq K\sigma_p^{(A)} \quad \text{ where } K := \tfrac{2+3C_1c_2^{-1}}{2+C_1c_2^{-1}} + C_1 c_2^{-1}. \end{aligned} \tag{6.5}$$

Let us denote with $\tau_p^{(i)}$, $\delta_p^{(i)}$, $f^1_{\tau_p^{(i)}}$, $\boldsymbol{f}^2_{\tau_p^{(i)}}$, $\bar{u}_{\tau_p^{(i)}}$, $\sigma_p^{(i)}$ the instances of τ_p, δ_p, $f^1_{\tau_p}$, $\boldsymbol{f}^2_{\tau_p}$, $\bar{u}_{\tau_p}$, σ_p at the moment of the ith call of **REFINE**$[\tau_p, F_p]$. If $2\delta_p^{(i)} > \omega\mathcal{E}(\tau_p^{(i)}, f^1_{\tau_p^{(i)}}, \boldsymbol{f}^2_{\tau_p^{(i)}}, \bar{u}_{\tau_p^{(i)}})$, then for any $k < i$,

$$\sigma_p^{(i)} < (2 + C_1(c_2^{-1} + 2\omega^{-1}))\delta_p^{(i)} \leq (2 + C_1(c_2^{-1} + 2\omega^{-1}))\beta\sigma_p^{(k)}.$$

If, for some $k \in I\!N_0$, $2\delta_p^{(j)} \leq \omega\mathcal{E}(\tau_p^{(j)}, f^1_{\tau_p^{(j)}}, \boldsymbol{f}^2_{\tau_p^{(j)}}, \bar{u}_{\tau_p^{(j)}})$ for $j = i, \ldots, i-k$, then by (6.4), Lemma 6.1, where we use that $\delta_p^{(j)} \leq \delta_p^{(j-1)}$, and (6.2),

$$\sigma_p^{(i)} \leq D\|u - \bar{u}_{\tau_p^{(i)}}\|_E \leq D\mu^k\|u - \bar{u}_{\tau_p^{(i-k)}}\|_E \leq D\mu^k\sigma_p^{(i-k)}.$$

Since $(2 + C_1(c_2^{-1} + 2\omega^{-1}))\beta < 1/K$ by definition of β, from (6.5) we conclude that for any $\alpha \in (0, 1)$ there exists an M such that $\sigma_p^{(i+M)} \leq \alpha\sigma_p^{(i)}$. Since all results derived so far are equally valid at the dual side, by taking $\alpha < 1/K$ we infer that by $2M$ iterations of the loop inside **GOAFEM**, the product $\sigma_p\sigma_d$ is reduced by a factor $\alpha K < 1$. Indeed, either σ_p or σ_p is reduced by a factor α, whereas the other cannot increase by a factor larger than K.

Next, we bound the cardinality of the output partition. If **GOAFEM** terminates as a result of the first evaluation of the test $\sigma_p\sigma_d \leq \varepsilon$, then by the assumptions that $\underline{\delta}_p > c_f$ and $\underline{\delta}_d > c_g$, the output partition $\tau_p \cup \tau_d = \tau_0$. In the following, we consider the case that initially $\sigma_p\sigma_d > \varepsilon$.

At evaluation of the test $\#\tau_p - \#\tau + \#F_p \leq \#\tau_d - \#\tau + \#F_d$, we have

$$\#\tau_p - \#\tau \leq (\beta K^{-1}\sigma_p)^{-1/s} c_f^{1/s}. \tag{6.6}$$

Indeed, the current $\#\tau_p - \#\tau$ is not larger than this difference at the moment of the most recent call of $\mathbf{RHS}_f[\tau, \delta_p]$. By the assumption of $\mathbf{RHS}_f$ being s-optimal, the latter difference was zero when at that time $\delta_p > c_f$. Otherwise, since $\underline{\delta}_p > c_f$ by assumption, this δ_p was equal to β times the minimum of all values attained by σ_p up to that moment. Using (6.5), and the fact that $\mathbf{RHS}_f$ is s-optimal with constant c_f, we end up with (6.6).

If, at evaluation of the test $\#\tau_p - \#\tau + \#F_p \leq \#\tau_d - \#\tau + \#F_d$, $F_p \neq \emptyset$, i.e., if in the preceding lines $2\delta_p \leq \omega\mathcal{E}(\tau_p, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p})$ and $F_p := \mathbf{MARK}[\tau, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p}]$, an application of Lemma 6.1 and the assumption that $u \in \mathcal{A}^s$ show that then

$$\#F_p \lesssim \|u - \bar{u}_{\tau_p}\|_E^{-1/s} |u|^{1/s}_{\mathcal{A}^s} \lesssim \sigma_p^{-1/s} |u|^{1/s}_{\mathcal{A}^s} \tag{6.7}$$

by (6.4).

Clearly, results analogous to (6.6) and (6.7) are valid at the dual side. Now with $\sigma_{p,j}$, $\sigma_{d,j}$ being the instances of σ_p, σ_d at the jth evaluation of the test $\#\tau_p - \#\tau + \#F_p \leq \#\tau_d - \#\tau + \#F_d$, with n being the last one, an application of Theorem 3.1 shows that for τ being the output of the call of **REFINE** following this last test, being thus the last call of **REFINE**, we have

$$\begin{aligned}
\#\tau - \#\tau_0 &\lesssim \sum_{j=1}^n \min\{\sigma_{p,j}^{-1/s}(|u|^{1/s}_{\mathcal{A}^s} + c_f^{1/s}), \sigma_{d,j}^{-1/t}(|z|^{1/t}_{\mathcal{A}^t} + c_g^{1/t})\} \\
&\leq \sum_{j=1}^n (\sigma_{p,j}\sigma_{d,j})^{-1/(s+t)} [(|u|^{1/s}_{\mathcal{A}^s} + c_f^{1/s})^s (|z|^{1/t}_{\mathcal{A}^t} + c_g^{1/t})^t]^{1/(s+t)} \\
&\lesssim \varepsilon^{-1/(s+t)} [(|u|^{1/s}_{\mathcal{A}^s} + c_f^{1/s})^s (|z|^{1/t}_{\mathcal{A}^t} + c_g^{1/t})^t]^{1/(s+t)},
\end{aligned} \tag{6.8}$$

by the majorized linear convergence of $(\sigma_{p,j}\sigma_{d,j})_j$ and $\sigma_{p,n}\sigma_{d,n} > \varepsilon$.

Suppose that this last call of **REFINE** took place at the primal side. Then the output partition of **GOAFEM** is $\tau_p \cup \tau_d$, where $[\tau_p, \cdot, \cdot] := \mathbf{RHS}_f[\tau, \delta_p]$ and $\tau_d := \tau \cup \tau_d$. As we have seen, if $\delta_p \leq c_f$, i.e., if possibly $\tau_p \supsetneq \tau$, then δ_p is larger than βK^{-1} times the current σ_p, that, by its definition, is larger than $2 + C_1 c_2^{-1}$ times the previous value of δ_p, denoted as $\delta_p^{(\mathrm{prev})}$. A call of $\mathbf{RHS}_f[\cdot, \delta_p^{(\mathrm{prev})}]$ has been made inside **GOAFEM**, and so $\tau \supseteq \tau'$ with $[\tau', \cdot, \cdot] := \mathbf{RHS}_f[\cdot, \delta_p^{(\mathrm{prev})}]$. The assumption of $\mathbf{RHS}_f$ being linearly convergent shows that $\#\tau_p \lesssim \#\tau$.

The current $\#\tau_d - \#\tau$ is not larger than this difference at the moment of the last call of $\mathbf{RHS}_g$, and so analogously we find that $\#\tau_d \lesssim \#\tau$. We conclude that

$$\#\tau_p \cup \tau_d \lesssim \#\tau \lesssim \#\tau_0 + \varepsilon^{-1/(s+t)} [(|u|^{1/s}_{\mathcal{A}^s} + c_f^{1/s})^s (|z|^{1/t}_{\mathcal{A}^t} + c_g^{1/t})^t]^{1/(s+t)}. \tag{6.9}$$

Finally, we have to bound the cost of the algorithm. At the moment of the first call of $\mathbf{GALSOLVE}[\tau_p, f_{\tau_p}, \bar{u}_{\tau_p}, \delta_p]$, we have

$$\|L^{-1}_{\tau_p} f_{\tau_p} - \bar{u}_{\tau_p}\|_E \leq \|f_{\tau_p} - f\|_{E'} + \|f\|_{E'} \leq \delta_p + \|f\|_{E'} \lesssim \delta_p,$$

by assumption. We now consider any further call. From (6.3), $\|u - u_{\tau_0}\|_E \leq \|f\|_{E'} \lesssim \underline{\delta}_p$ by assumption, and (6.5), we have that the current δ_p and σ_p at the moment of such a call satisfy $\sigma_p \lesssim \delta_p$. As a consequence, we have

$$\begin{aligned}
\|L^{-1}_{\tau_p} f_{\tau_p} - \bar{u}_{\tau_p}\|_E &\leq \|(L^{-1} - L^{-1}_{\tau_p}) f_{\tau_p}\|_E + \|L^{-1} f_{\tau_p} - \bar{u}_{\tau_p}\|_E \leq 2\|L^{-1} f_{\tau_p} - \bar{u}_{\tau_p}\|_E \\
&\leq 2[\|f - f_{\tau_p}\|_{E'} + \|u - \bar{u}_{\tau_p}\|_E] \leq 2\delta_p + 2\sigma_p \lesssim \delta_p.
\end{aligned}$$

By the assumption of **GALSOLVE** being an optimal iterative solver, we conclude that the cost of these calls are $\mathcal{O}(\#\tau_p)$.

The number of arithmetic operations needed for the calls $\mathbf{MARK}[\tau, f^1_{\tau_p}, \boldsymbol{f}^2_{\tau_p}, \bar{u}_{\tau_p}]$, $\tau := \mathbf{REFINE}[\tau_p, F_p]$, and $[\tau_p, \cdot, \cdot] := \mathbf{RHS}_f[\tau, \delta_p]$ are $\mathcal{O}(\#\tau)$, $\mathcal{O}(\#\tau)$, and $\mathcal{O}(\#\tau_p)$, respectively. Moreover, we know that $\#\tau_p \lesssim \#\tau$, and that $\#\tau - \#\tau_0$ as function of the iteration count is majorized by a linearly increasing sequence with upper bound (6.8). From the assumption that $\underline{\delta}_p\underline{\delta}_d \lesssim \|u - u_{\tau_0}\|_E\|z - z_{\tau_0}\|_E + \varepsilon$, the first $\sigma_p\sigma_d \lesssim \|u - u_{\tau_0}\|_E\|z - z_{\tau_0}\|_E + \varepsilon$, meaning that after some absolute constant number of iterations, either the current τ is unequal to τ_0 or the algorithm has terminated. Together, above observations show that the total cost is bounded by some absolute multiple of the right-hand side of (6.9). $\square$

Remark 6.6. The functions $\bar{u}_\tau$, $\bar{z}_\tau$ produced by **GOAFEM** are not the exact Galerkin approximations, and so $\|u - \bar{u}_\tau\|_E\|z - \bar{z}_\tau\|_E$ is not necessarily an upper bound for $|g(u) - g(\bar{u}_\tau)|$. Writing

$$g(u)-g(\bar{u}_\tau) = a(u-\bar{u}_\tau, z) = a(u-\bar{u}_\tau, z-z_\tau) = a(u-\bar{u}_\tau, z-\bar{z}_\tau) - a(u-\bar{u}_\tau, z_\tau-\bar{z}_\tau),$$

and using that $\|u - \bar{u}_\tau\|_E \le \sigma_p$, $\|z - \bar{z}_\tau\|_E \le \sigma_d$, $\|z_\tau - \bar{z}_\tau\| \le \delta_d \le (2 + C_1c_2^{-1})^{-1}\sigma_d$, and $\sigma_p\sigma_d \le \varepsilon$, we end up with $|g(u) - g(\bar{u}_\tau)| \le [1 + (2 + C_1c_2^{-1})^{-1}]\varepsilon$.

7. Numerical experiments. In this section we will consider the performance of the **GOAFEM** routine in practice. As many real-world problems require the evaluation of functionals that are unbounded on $H^1_0(\Omega)$, we will also consider such a problem. As **GOAFEM** can only handle bounded functionals, we need to do some additional work. Following [BS01], we will apply a so-called *extraction functional*, a technique that we recall below. An alternative approach would be to apply an regularized functional as suggested in [OR76, BR96]. This approach can be applied more generally since no Green's function is needed. On the other hand, it introduces an additional error that can only be controlled in terms of higher order derivatives of the solution beyond those that are needed for the functional to be well defined.

7.1. Extraction functionals. Let $\tilde{g}$ be some functional defined on the solution u of (2.1), but that is unbounded on $H^1_0(\Omega)$. With f being the right hand side of (2.1), we write $\tilde{g}(u) = g(u) + M(f)$, where $g \in H^{-1}(\Omega)$, and M a functional on f. Since u and f are related via an invertible operator, this is always possible, even for any $g \in H^{-1}(\Omega)$. Yet, we would like to do this under the additional constraint that $M(f)$ can be computed within any given tolerance at low cost. Basically, this additional condition requires that a Green's function for the differential operator is available.

We consider $\boldsymbol{A} = \mathrm{Id}$, i.e., the Poisson problem, on a two-dimensional domain Ω, and, for some $\bar{x} \in \Omega$, $\tilde{g} = \tilde{g}_{\bar{x}}$ given by

$$\tilde{g}_{\bar{x}}(u) = \frac{\partial u}{\partial x_1}(\bar{x}),$$

assuming that u is sufficiently smooth. With (r, θ) denoting polar coordinates centered at $\bar{x}$, we have $\triangle \frac{\log r}{2\pi} = \delta_{\bar{x}}$, and so $-\triangle \frac{\cos\theta}{2\pi r} = \tilde{g}_{\bar{x}}$, in the sense that for any smooth test function $\phi \in \mathcal{D}(\mathbb{R}^2)$, $-\int_{\mathbb{R}^2} \frac{\cos\theta}{2\pi r} \triangle\phi = \tilde{g}_{\bar{x}}(\phi)$. Generally, this formula cannot be applied with ϕ replaced by the solution u of (2.1). Indeed, in the general case this function has a non-vanishing normal derivative at the boundary of Ω, and therefore its zero extension is not sufficiently smooth. Therefore, with $w^{\bar{x}}_0 := \frac{\cos\theta}{2\pi r}$, $w^{\bar{x}}_1$ being a sufficiently smooth function equal to $w^{\bar{x}}_0$ outside some open $\Sigma \Subset \Omega$ that contains $\bar{x}$, and $w^{\bar{x}} := w_0^{\bar{x}} - w_1^{\bar{x}}$, for any $\phi \in \mathcal{D}(\mathbb{R}^2)$ we write

$$\begin{aligned}\tilde{g}_{\bar{x}}(\phi) &= -\int_{\mathbb{R}^2} w_1^{\bar{x}} \triangle\phi - \int_{\mathbb{R}^2} w^{\bar{x}} \triangle\phi \\ &= \int_{\mathbb{R}^2} \triangle(-w_1^{\bar{x}})\phi + \int_{\Omega} w^{\bar{x}}(-\triangle\phi) \\ &=: g_{\bar{x}}(\phi) + M_{\bar{x}}(-\triangle\phi).\end{aligned}$$

Clearly, $g_{\bar{x}}$ extends to a bounded functional on $L_1(\mathbb{R}^2)$, with $g_{\bar{x}}(v) = \int_\Omega \triangle(-w_1^{\bar{x}})v$ when $\operatorname{supp} v \subset \Omega$. In particular, $g_{\bar{x}}$ is bounded on $H_0^1(\Omega)$, which enables us to use **GOAFEM** to evaluate it. Moreover, since $\operatorname{supp} w^{\bar{x}} \Subset \Omega$, under some mild conditions above reformulation can be shown to be applicable to u. The details are as follows:

PROPOSITION 7.1. *If*
(a). $f \in L_2(\Omega)$,
(b). *u is continuously differentiable at $\bar{x}$,*
(c). *in a neighbourhood of $\bar{x}$, f is in L^p for some $p > 2$.*
then

$$\tilde{g}_{\bar{x}}(u) = g_{\bar{x}}(u) + M_{\bar{x}}(f)$$

Proof. Let $B(\bar{x};\varepsilon)$ be the ball centered at $\bar{x}$ with radius ε, small enough such that $B(\bar{x};\varepsilon) \Subset \Omega$. Since $u, w^{\bar{x}} \in H^1(\Omega \backslash B(\bar{x};\varepsilon))$, $\triangle u \in L_2(\Omega\backslash B(\bar{x};\varepsilon))$ by (a), $\triangle w^{\bar{x}} \in L_2(\Omega\backslash B(\bar{x};\varepsilon))$, and $\operatorname{supp} w^{\bar{x}} \Subset \Omega$, integration by parts shows that

$$\int_{\partial B(\bar{x};\varepsilon)} w^{\bar{x}} \frac{\partial u}{\partial \mathbf{n}} - u\frac{\partial w^{\bar{x}}}{\partial \mathbf{n}} = \int_{\Omega\backslash B(\bar{x};\varepsilon)} u\triangle w^{\bar{x}} - w^{\bar{x}}\triangle u, \tag{7.1}$$

where $\mathbf{n}$ is the outward pointing normal of $\partial B(\bar{x};\varepsilon)$.

We have $\lim_{\varepsilon\downarrow 0} \int_{\Omega\backslash B(\bar{x};\varepsilon)} u\triangle w^{\bar{x}} = -\lim_{\varepsilon\downarrow 0}\int_{\Omega\backslash B(\bar{x};\varepsilon)} u\triangle w_1^{\bar{x}} = g_{\bar{x}}(u)$.

Since $|\int_{B(\bar{x};\varepsilon)} w_0^{\bar{x}} f| \leq \|f\|_{L_p(B(\bar{x};\varepsilon))}\|w_0^{\bar{x}}\|_{L_q(B(\bar{x};\varepsilon))}$ $(\frac{1}{p}+\frac{1}{q}=1)$, and furthermore $\|w_0^{\bar{x}}\|_{L_q(B(\bar{x};\varepsilon))} = [\int_0^\varepsilon\int_0^{2\pi} |\frac{\cos\theta}{2\pi r}|^q r]^{1/q} \to 0$ when $\varepsilon \downarrow 0$ and $q < 2$, from (c) we conclude that $-\lim_{\varepsilon\downarrow 0}\int_{\Omega\backslash B(\bar{x};\varepsilon)} w^{\bar{x}}\triangle u = \int_\Omega w^{\bar{x}} f = M_{\bar{x}}(f)$.

The contributions of $w_1^{\bar{x}}$ to the left hand side of (7.1) vanish when $\varepsilon \downarrow 0$.

From $\int_{\partial B(\bar{x};\varepsilon)} w_0^{\bar{x}}\frac{\partial u}{\partial \mathbf{n}} = \int_0^{2\pi}(\cos\theta\,\frac{\partial u}{\partial x_1} + \sin\theta\,\frac{\partial u}{\partial x_2})\frac{\cos\theta}{2\pi\varepsilon}\varepsilon d\theta$, and (b), we infer that $\lim_{\varepsilon\downarrow 0}\int_{\partial B(\bar{x};\varepsilon)} w_0^{\bar{x}}\frac{\partial u}{\partial \mathbf{n}} = \frac{1}{2}\frac{\partial u}{\partial x_1}(\bar{x})$.

From

$$\begin{aligned}\int_{\partial B(\bar{x};\varepsilon)} u\frac{\partial w_0^{\bar{x}}}{\partial \mathbf{n}} &= \tfrac{-1}{2\pi\varepsilon}\int_0^{2\pi}\cos\theta\, u d\theta = \tfrac{1}{2\pi\varepsilon}\int_0^{2\pi}\sin\theta\,\tfrac{\partial u}{\partial\theta}d\theta \\ &= \tfrac{1}{2\pi}\int_0^{2\pi}\sin\theta\,(-\sin\theta\,\tfrac{\partial u}{\partial x_1} + \cos\theta\,\tfrac{\partial u}{\partial x_2})d\theta,\end{aligned}$$

and (b), we infer that $-\lim_{\varepsilon\downarrow 0}\int_{\partial B(\bar{x};\varepsilon)} u^{\bar{x}}\frac{\partial w_0^{\bar{x}}}{\partial \mathbf{n}} = \frac{1}{2}\frac{\partial u}{\partial x_1}(\bar{x})$. Together, above observations give the proof. ▯

7.2. Implementation. The implementation of the **GOAFEM** routine is, essentially, as described above, with the sole difference that we did not approximate the right-hand sides for setting up the Galerkin systems and computing the a posteriori error estimators, but instead used quadrature directly. This was possible, and in view of Remark 6.2 reasonable, because in our experiments, either the right-hand sides are very smooth, or are already in $\mathbb{V}^*_{\tau_0} + \operatorname{div}[\mathbb{V}^*_{\tau_0}]^n$.

For all experiments, we used $p = 2$, i.e., quadratic Lagrange elements.

The **GALSOLVE** routine we use solves the linear systems with the Conjugate Gradient method using the well-known Bramble-Pasciak-Xu (BPX) preconditioner.

All routines were implemented in Common Lisp, and run using the `SBCL` compiler and run-time environment. This allowed for a short development time and well instrumented code. With regards to efficiency, the only effort made in that direction consisted in making sure that the asymptotics were correct. While an efficient implementation would be possible with moderate effort (see [Neu03]), for our purposes convenience and correctness were the most important considerations.

As in one of the experiments we use the extraction functional for the partial derivative at a a point introduced above, we also have to solve a quadrature problem. For this we used the adaptive cubature routine `Cuhre` [BEG91] as implemented in the `Cuba` cubature package [Hah05].

7.3. Experiments. To test **GOAFEM**, we chose two distinct situations. For the first example, we want to compute a partial derivative at a point of a function given as the solution of a Poisson problem, thus illustrating the applicability of our method to this situation.

In our second example, we consider a problem in which the singularities of the solutions to the primal and dual problems are spatially separated.

Example 7.2. Let $\Omega = (0,1)^2$. We consider problem (2.1), choosing the right-hand side $f = 1$ (i.e., $f(v) = \int_\Omega v dx$). We will test the performance of **GOAFEM** on the task of computing

$$\frac{\partial u}{\partial x_1}(\bar{x}).$$

with $\bar{x} = (\frac{\pi}{7}, \frac{49}{100})$. The initial partition is as indicated in Figure 7.1, with $(\frac{1}{2}, \frac{1}{2})$ being

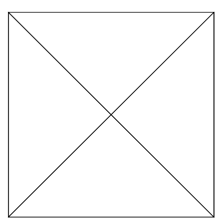

Fig. 7.1. *Initial partition τ_0 corresponding to Example 7.2*

the newest vertex of all 4 triangles.

Following the discussion from Subsect. 7.1, we take $w_1^{\bar{x}} = \psi w_0^{\bar{x}}$, and thus $w^{\bar{x}} = (1-\psi) w_0^{\bar{x}}$, with ψ being a sufficiently smooth function, 1 outside some neighboorhoud of $\bar{x}$ inside Ω, and 0 on some smaller neighbourhood of $\bar{x}$. Proposition 7.1 shows that $\frac{\partial u}{\partial x_1}(\bar{x}) = \int_\Omega u \triangle(-\psi w_0^{\bar{x}}) + \int_\Omega (1-\psi) w_0^{\bar{x}} f$. Writing (θ, r) for the polar coordinates around $\bar{x}$, we chose

$$\psi(\theta, r) := \int_0^r \psi^*(s) ds \,/\, \int_0^\infty \psi^*(s) ds, \tag{7.2}$$

with ψ^* a spline function of order 6, with support $[0.1, 0.45]$.

We evaluated $\int_\Omega (1-\psi) w_0^{\bar{x}} f$ using the adaptive quadrature routine `Cuhre`. To obtain precision of 10^{-12} it needed 216515 integrand evaluations. On current off-the-shelf hardware, it takes only a few seconds.

To approximate $\int_\Omega u\triangle(-\psi w_0^{\bar{x}})$ we used **GOAFEM**. Since the right-hand sides 1 and $\triangle(-\psi w_0^{\bar{x}})$ of primal and dual problem are smooth, their solutions are in $\mathcal{A}^{p/n} = \mathcal{A}^1$, so that the error in the functional is $\mathcal{O}([\#\tau - \#\tau_0]^{-2})$. We compared the results with those obtained with the corresponding non-goal oriented adaptive finite element method **AFEM** for minimizing the error in energy norm, which is obtained by applying refinements always because of the markings at primal side.

The solutions of primal and dual problem are in $H^{3-\varepsilon}(\Omega)$ for any $\varepsilon > 0$, but, because the right-hand sides do not vanish at the corners, they are not in $H^3(\Omega)$. Recalling that we use quadratic elements, as a consequence, (fully) optimal convergence rates with respect to $\|\,\|_E$ are not obtained using uniform refinements. On the other hand, since the (weak) singularities in primal and dual solution are solely caused by the shape of the domain, the same local refinements near the corners are appropriate for both primal and dual problem. Therefore, in view of (1.1), we may expect that also with **AFEM** the error in the functional is $\mathcal{O}([\#\tau - \#\tau_0]^{-2})$. On the other hand, since quantitatively the right-hand side, and so the solution of the dual problem are not that smooth, see Figure 7.2, we may hope that the application of **GOAFEM**

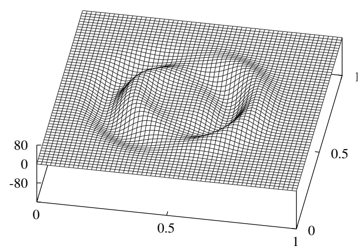

Fig. 7.2. *Right-hand side of the dual problem corresponding to Example 7.2*

yields quantitatively better results.

In Figure 7.3, we show errors in $\int_\Omega u\triangle(-\psi w_0^{\bar{x}})$ as function of $\#\tau - \#\tau_0$. The results confirm that for both **GOAFEM** and **AFEM**, these errors are $\mathcal{O}([\#\tau - \#\tau_0]^{-2})$, where on average for **GOAFEM** the errors are smaller. In 7.4, we show partitions produced by **GOAFEM** and **AFEM**. With **AFEM** local refinements are only made towards the corners, whereas with **GOAFEM** additional local refinements are made in areas where quantitatively the dual solution is non smooth due to oscillations in its right-hand side.

Example 7.3. As in Example 7.2, we consider Poisson's problem on the unit square. We now take as initial partition the one that is obtained from the partition from Figure 7.1 by 2 uniform refinements. We define the right-hand sides f and g of primal and dual problem by

$$f(v) = -\int_{T_f} \frac{\partial v}{\partial x_1}, \qquad g(v) = -\int_{T_g} \frac{\partial v}{\partial x_1}, \tag{7.3}$$

where T_f and T_g are the simplices $\{(0,0),(\frac{1}{2},0),(0,\frac{1}{2})\}$, and $\{(1,1),(\frac{1}{2},1),(1,\frac{1}{2})\}$, respectively, see Figure 7.5. That is, with χ_f being the characteristic function of T_f,

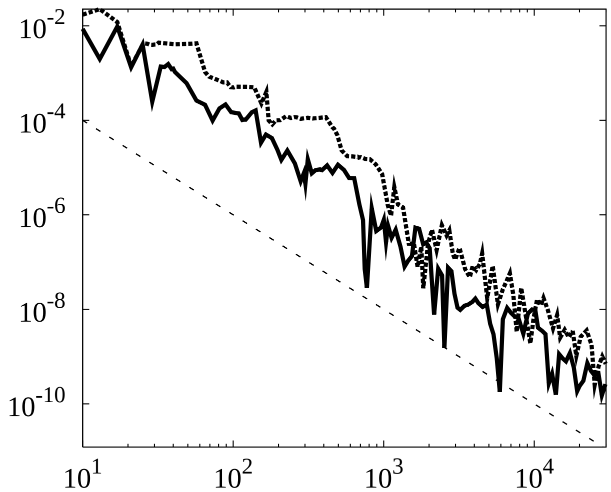

FIG. 7.3. *Error in the functional vs.* $\#\tau - \#\tau_0$ *using GOAFEM (solid) and AFEM (dashed) corresponding to Example 7.2, and a curve* $C[\#\tau - \#\tau_0]^{-2}$

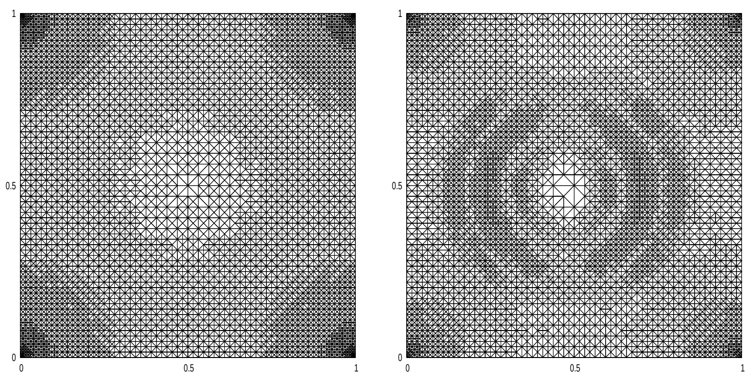

FIG. 7.4. *Partitions produced by* **AFEM** *and* **GOAFEM** *with nearly equal number of triangles for Example 7.2*

$f = \operatorname{div}[\chi_f \quad 0]^T$. So in view of (4.3), here we write f as $f^1 + \operatorname{div}\boldsymbol{f}^2$ with vanishing f^1, and benefit from the fact that $\boldsymbol{f}^2 \in [V^*_{\tau_0}]^2$. Similarly for g.

The primal solution has a singularity along the line connecting the points $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$, see Figure 7.6, and similarly the dual solution has one along the line connecting $(1, \frac{1}{2})$ and $(\frac{1}{2}, 1)$. Since the non-goal oriented adaptive finite element routine **AFEM** does not see the latter singularity, it behaves much worse than **GOAFEM** as appears from Figure 7.7. For **GOAFEM** we observe an error $\mathcal{O}([\#\tau - \#\tau_0]^{-2})$, which, since $p/n = 1$, is equal to the best possible rate predicted by Theorem 6.4. In Figure 7.8, we show partitions produced by **AFEM** and **GOAFEM**, respectively.

Appendix A. A cheaper adaptive algorithm for lowest order elements.

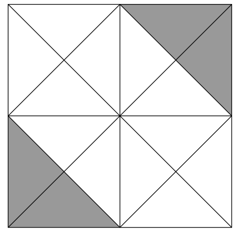

FIG. 7.5. *Initial partition τ_0 corresponding to Example 7.3, and T_f (left bottom), T_g (right top)*

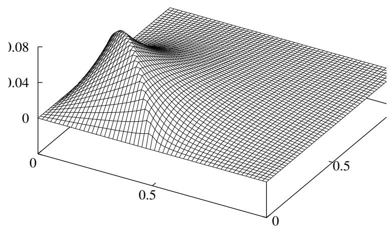

FIG. 7.6. *Primal solution corresponding to Example 7.3)*

Inspired by [CV99], for lowest order elements, i.e., for

$$\mathbb{V}_\tau = H_0^1(\Omega) \cap \prod_{T\in\tau} P_1(T), \text{ and } \mathbb{V}_\tau^* = \prod_{T\in\tau} P_0(T),$$

we construct a cheaper a posteriori error estimator for the energy norm, which will be used to construct a cheaper **GOAFEM**. We start with constructing a new biorthogonal projector.

For $k \in \mathbb{N}_0$ and $T \in \tau$, we define the ring $R_k(\tau,T)$ by $R_0(\tau,T) = T$, and, for $k > 0$, by $R_k(\tau,T) = \cup_{\{T'\in\tau: T'\cap R_{k-1}(\tau,T)\neq\emptyset\}} T'$. We set $N_T = \{\text{vertices of } T\}$, $\overline{N}_\tau = \cup_{T\in\tau} N_T$, and $N_\tau = \overline{N}_\tau \cap \Omega$. Under the assumption that for some absolute constant $\bar{k} \in \mathbb{N}_0$,

$$R_{\bar{k}}(\tau,T) \cap N_\tau \neq \emptyset, \quad (T \in \tau), \tag{A.1}$$

we will construct a projector P_τ onto $\mathbb{V}_\tau$ such that for any $p \in [1,\infty]$,

$$\|v - P_\tau v\|_{L_p(T)} \lesssim \inf_{v_\tau\in\mathbb{V}_\tau} \|v - v_\tau\|_{L_p(R_{\bar{k}+2}(\tau,T))} \quad (v \in L_p(\Omega),\, T \in \tau), \tag{A.2}$$

and

$$\|v - P_\tau v\|_{L_p(T)} + \operatorname{diam}(T)|v - P_\tau v|_{W_p^1(T)} \lesssim \operatorname{diam}(T)^m |v|_{W_p^m(R_{\bar{k}+3}(\tau,T))} \tag{A.3}$$

$(v \in \overset{\circ}{W}{}_p^1(\Omega) \cap W_p^m(\Omega),\, m \in \{1,2\},\, T \in \tau)$, *and* such that for its $L_2(\Omega)$-adjoint,

$$\|h - P_\tau^* h\|_{L_p(T)} \lesssim \inf_{s\in P_0(\Omega)} \|h - s\|_{L_p(R_{\bar{k}+2}(\tau,T))} \quad (h \in L_p(\Omega),\, T \in \tau). \tag{A.4}$$

Remark A.1. The assumption (A.1) is not always satisfied as illustrated in Figure A.1. On the other hand, since a sequence of refinements towards the boundary

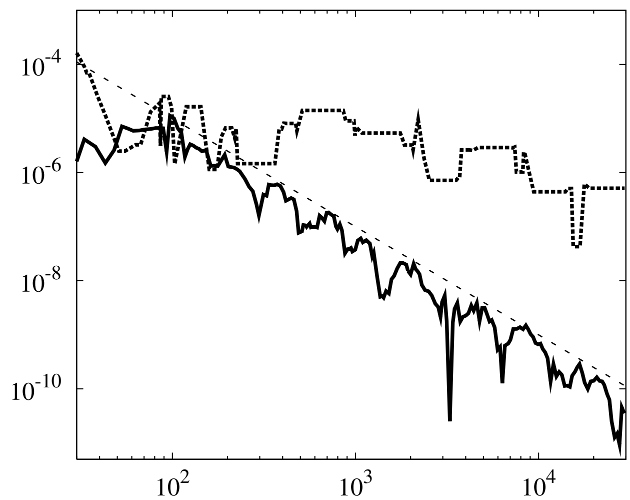

FIG. 7.7. *Error in the functional vs.* $\#\tau - \#\tau_0$ *using GOAFEM (solid) and AFEM (dashed) corresponding to example 7.3, and a curve* $C[\#\tau - \#\tau_0]^{-2}$

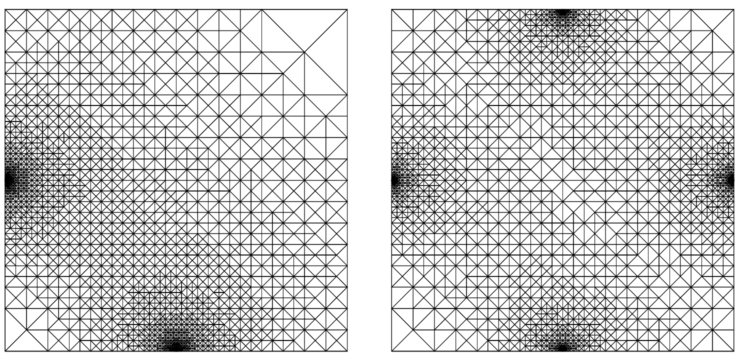

FIG. 7.8. *Partitions produced by* **AFEM** *and* **GOAFEM** *with nearly equal number of triangles for Example 7.3*

without creating interior vertices as in the right picture of Figure A.1 doesn't change $\mathbb{V}_\tau$, (A.1) is not an essential restriction.

We will modify the construction of a so-called quasi-interpolator from [Osw94]. Let $\{\lambda_a^{(T)} : a \in N_T\}$ denote the barycentric coordinates of T, i.e., $\lambda_a^{(T)} \in P_1(T)$ and $\lambda_a^{(T)}(b) = \delta_{ab}$ $(a, b \in N_T)$. For $a \in N_T$, we set

$$\mu_a^{(T)} = \tfrac{n+1}{\mathrm{vol}(T)}((n+1)\lambda_a^{(T)} - \sum_{a \neq b \in N_T} \lambda_b^{(T)}).$$

Then $\{\lambda_a^{(T)} : a \in N_T\}$ and $\{\mu_a^{(T)} : a \in N_T\}$ are $L_2(T)$-biorthogonal collections. Since $\|\lambda_a^{(T)}\|_{L_p(T)} \eqsim \mathrm{vol}(T)^{\frac{1}{p}-\frac{1}{p'}}\|\lambda_a^{(T)}\|_{L_{p'}(T)}$, and similarly for $\mu_a^{(T)}$, we have that $\|\lambda_a^{(T)}\|_{L_p(T)}\|\mu_a^{(T)}\|_{L_q(T)} \eqsim 1$ $(\frac{1}{p} + \frac{1}{q} = 1)$, independently of T.

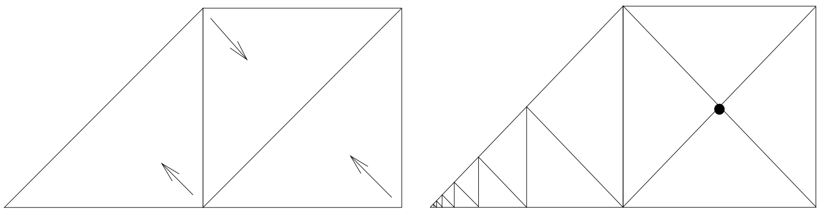

Fig. A.1. *Partition τ_0 with newest vertices indicated with arrows, and a family of descendants τ with $N_\tau = \{\bullet\}$*

For any $a \in \overline{N}_\tau$, we define ϕ_a, ψ_a by $\phi_a|_T = \begin{cases} \lambda_a^{(T)} & \text{when } a \in N_T \\ 0 & \text{otherwise} \end{cases}$, and $\psi_a|_T = \begin{cases} c_{aT}\mu_a^{(T)} & \text{when } a \in N_T \\ 0 & \text{otherwise} \end{cases}$, where the c_{aT} are uniformly bounded constants with $\sum_{\{T\in\tau: a\in N_T\}} c_{aT} = 1$. Thanks to the last property, the collections $\{\phi_a : a \in \overline{N}_\tau\}$ and $\{\psi_a : a \in \overline{N}_\tau\}$ are $L_2(\Omega)$-biorthogonal. In [Osw94], c_{aT} was chosen to be 1 for one $T \ni a$ and zero otherwise. In view of obtaining (A.4), we take $c_{aT} = \frac{\mathrm{vol}(T)}{\sum_{\{T'\in\tau: a\in N_{T'}\}} \mathrm{vol}(T')}$. Then for any $T \in \tau$,

$$\sum_{a\in\overline{N}_\tau} \langle 1, \phi_a\rangle \psi_a|_T = \sum_{a\in N_T} \sum_{\{T'\in\tau: a\in N_{T'}\}} \tfrac{\mathrm{vol}(T')}{n+1} c_{aT}\mu_a^{(T)} = \tfrac{\mathrm{vol}(T)}{n+1} \sum_{a\in N_T} \mu_a^{(T)} = 1,$$

i.e., this is the (unique) choice of the c_{aT} such that $P_0(\Omega) \subset \mathrm{span}\{\psi_a : a \in \overline{N}_\tau\}$.

Next, we remove degrees of freedom associated to $b \in \overline{N}_\tau \backslash N_\tau$. For each $b \in \overline{N}_\tau \backslash N_\tau$, we select an $a \in \cup_{\{T\in\tau: b\in N_\tau\}} R_{\bar{k}}(\tau, T) \cap N_\tau$, with $\bar{k}$ from assumption (A.1), and update

$$\psi_a \leftarrow \psi_a + \frac{\langle 1, \phi_b\rangle_{L_2(\Omega)}}{\langle 1, \phi_a\rangle_{L_2(\Omega)}} \psi_b.$$

Note that an ψ_a might be involved in more than 1 updates. Since

$$\langle 1, \phi_b\rangle_{L_2(\Omega)}\psi_b + \langle 1, \phi_a\rangle_{L_2(\Omega)}\psi_a = \langle 1, \phi_a\rangle_{L_2(\Omega)}\big[\psi_a + \frac{\langle 1, \phi_b\rangle_{L_2(\Omega)}}{\langle 1, \phi_a\rangle_{L_2(\Omega)}}\psi_b\big],$$

the resulting collection $\{\psi_a : a \in N_\tau\}$, thus with the generally redefined ψ_a, contains $P_0(\Omega)$ in its span, and is $L_2(\Omega)$-biorthogonal to $\{\phi_a : a \in N_\tau\}$, the latter being the standard nodal basis for $\mathbb{V}_\tau$. We set the projector $P_\tau v := \sum_{a\in N_\tau} \langle v, \psi_a\rangle_{L_2(\Omega)} \phi_a$, that has $L_2(\Omega)$-adjoint $P_\tau^* h := \sum_{a\in N_\tau} \langle h, \phi_a\rangle_{L_2(\Omega)} \psi_a$.

By construction,

$$\mathrm{supp}\,\phi_a \subset \bigcup_{\{T\in\tau: a\in N_T\}} R_0(\tau, T), \quad \mathrm{supp}\,\psi_a \subset \bigcup_{\{T\in\tau: a\in N_T\}} R_{\bar{k}+1}(\tau, T).$$

Furthermore, due to the fact that any two $T, T' \in \tau$ with $T \cap T' \neq \emptyset$ have uniformly comparable diameters, and all simplices from all partitions are uniformly shape regular, we have $\|\phi_a\|_{L_p(\Omega)} \eqsim \|\phi_a\|_{L_p(T)}$, $\|\psi_a\|_{L_p(\Omega)} \eqsim \|\psi_a\|_{L_p(T)}$ for any T in $\mathrm{supp}\,\phi_a$ or $\mathrm{supp}\,\psi_a$, respectively. We infer that for any $T \in \tau$,

$$\|P_\tau v\|_{L_p(T)} \leq \sum_{a\in N_T} \|v\|_{L_p(\mathrm{supp}\,\psi_a)} \|\psi_a\|_{L_q(\Omega)} \|\phi_a\|_{L_p(T)} \lesssim \|v\|_{L_p(R_{\bar{k}+2}(\tau, T))},$$

($v \in L_p(\Omega)$), which, since P_τ reproduces $\mathbb{V}_\tau$, implies (A.2). By a standard inverse inequality, we also have

$$\|P_\tau v\|_{L_p(T)} + \operatorname{diam}(T)|P_\tau v|_{W_p^1(T)} \lesssim \|v\|_{L_p(R_{\bar{k}+2}(\tau,T))} \quad (v \in L_p(\Omega)),$$

which, again by the reproduction of $\mathbb{V}_\tau \subset W_p^1(\Omega)$, gives

$$\begin{aligned}\|v-P_\tau v\|_{L_p(T)} &+ \operatorname{diam}(T)|v - P_\tau v|_{W_p^1(T)} \\ &\lesssim \inf_{v_\tau \in \mathbb{V}_\tau} [\|v - v_\tau\|_{L_p(R_{\bar{k}+2}(\tau,T))} + \operatorname{diam}(T)|v - v_\tau|_{W_p^1(T)}], \quad (v \in W_p^1(\Omega)).\end{aligned}$$

By now selecting v_τ to be the Scott-Zhang interpolation ([SZ90]) of $v \in \overset{\circ}{W}{}_p^1(\Omega) \cap W_p^m(\Omega)$ ($m = 1, 2$) in $C(\Omega) \cap \prod_{T\in\tau} P_1(T)$, that, by our assumption made in Sect.2 of $\partial\Omega$ being the union of true hyperfaces of $T \in \tau_0$, preserves the homogeneous boundary conditions, i.e., $v_\tau \in \mathbb{V}_\tau$, we end up with (A.3).

Analogously, we find that for any $T \in \tau$,

$$\|P_\tau^* h\|_{L_p(T)} \lesssim \sum_{\{a\in N_\tau : T\subset \cup_{\{T'\in\tau : a\in N_{T'}\}} R_{\bar{k}+1}(\tau,T')\}} \|v\|_{L_p(\operatorname{supp}\phi_a)} \lesssim \|v\|_{L_p(R_{\bar{k}+2}(\tau,T))}$$

($v \in L_p(\Omega)$), which, since P_τ^* reproduces $P_0(\Omega)$, gives (A.4).

Following an idea from [CV99], using the projection P_τ now we will show that for the lowest order elements, and a restricted class of right-hand sides, the a posteriori error estimator consisting of the face contributions only is already an upper bound for the energy norm error. In the following, let E_τ denote the set of interior true hyperfaces of τ.

Proposition A.2. *Let $\tau' \supset \tau$ be partitions, $f^1 \in \operatorname{Im} P_\tau^*$, $\boldsymbol{f}^2 \in [\mathbb{V}_\tau^*]^n$ and*

$$G = G(\tau, \tau') := \{e \in E_\tau : \forall T \in \tau,\, T \ni e,\, \exists T' \in R_{\bar{k}+2}(\tau,T) \text{ with } T' \notin \tau'\}.$$

Then, with for $e \in E_\tau$, $\eta_e(u_\tau, \boldsymbol{f}^2) := \operatorname{diam}(e)\|[\![[\boldsymbol{A}\nabla u_\tau + \boldsymbol{f}^2] \cdot \mathbf{n}]\!]_e\|^2_{L_2(e)}$, we have

$$\|u_{\tau'} - u_\tau\|_E \le \bar{C}_1 \big[\sum_{e\in G} \eta_e(u_\tau, \boldsymbol{f}^2)\big]^{\frac{1}{2}},$$

for some absolute constant $\bar{C}_1 > 0$. Note that $\#G \lesssim \#\tau' - \#\tau$, and that $G = E_\tau$ when $\tau' = \infty$, in which case thus $u_{\tau'} = u$.

Proof. We have $\|u_{\tau'} - u_\tau\|_E = \sup_{0\neq v_{\tau'}\in\mathbb{V}_{\tau'}} \frac{|a(u_{\tau'}-u_\tau, v_{\tau'})|}{\|v_{\tau'}\|_E}$, whereas for each $v_{\tau'} \in \mathbb{V}_{\tau'}$, $a(u_{\tau'} - u_\tau, v_{\tau'}) = a(u_{\tau'} - u_\tau, v_{\tau'} - v_\tau)$ for any $v_\tau \in \mathbb{V}_\tau$. Following an idea from [CV99], we select $v_\tau = P_\tau v_{\tau'}$. Then, since $f^1 \perp_{L_2(\Omega)} (I - P_\tau)v_{\tau'}$, and, on each $T \in \tau$, $\nabla \cdot [\boldsymbol{A}\nabla u_\tau + \boldsymbol{f}^2] = 0$, following the proof of Proposition 4.2 we find that

$$a(u_{\tau'} - u_\tau, v_{\tau'} - v_\tau) = -\sum_{e\in E_\tau} \int_e [\![(\boldsymbol{A}\nabla u_\tau + \boldsymbol{f}^2) \cdot \mathbf{n}]\!](v_{\tau'} - P_\tau v_{\tau'}).$$

Using the property (A.3) (and (A.2)) of the projector P_τ, the remainder of the proof can follow the lines indicated in the proof of Proposition 4.2. ◻

Although we are not going to apply it in this way, Proposition A.2 gives rise to the following upper bound for error in the energy norm for general $f \in H^{-1}(\Omega)$:

Corollary A.3. *For $f \in H^{-1}(\Omega)$, and any $\boldsymbol{f}^2_\tau \in [\mathbb{V}^*_\tau]^n$, we have*

$$\|u-u_\tau\|_E \leq \bar{C}_1\big[\sum_{e\in E_\tau}\eta_e(u_\tau,\boldsymbol{f}^2_\tau)\big]^{\frac{1}{2}}+(\bar{C}_1c_2^{-1}+1)\inf_{f^1_\tau\in \mathrm{Im}P^*_\tau}\|f-(f^1_\tau+\mathrm{div}\boldsymbol{f}^2_\tau)\|_{E'},$$

where, when $\tilde{f}:=f-\mathrm{div}\boldsymbol{f}^2_\tau \in L_2(\Omega)$, the second term can be bounded by

$$\|(I-P^*_\tau)\tilde{f}\|_{E'} \lesssim \Big\{\sum_{T\in\tau}\mathrm{diam}(T)^2\big\|\tilde{f}-\mathrm{vol}(R_{\bar{k}+2}(\tau,T))^{-1}\textstyle\int_{R_{\bar{k}+2}(\tau,T)}\tilde{f}\big\|^2_{L_2(R_{\bar{k}+2}(\tau,T))}\Big\}^{\frac{1}{2}}.$$

This result improves upon the result from [CV99] derived using a modified Clément interpolator, in the sense that for smooth $\tilde{f}$, in the last expression *all* terms are $\mathcal{O}(\mathrm{diam}(T)^2)$, also for T near the boundary and an $\tilde{f}$ that does not vanish on $\partial\Omega$.

Proof. By writing $u-u_\tau=(L^{-1}-L^{-1}_\tau)(f^1_\tau+\mathrm{div}\boldsymbol{f}^2_\tau)+(L^{-1}-L^{-1}_\tau)(f-(f^1_\tau+\mathrm{div}\boldsymbol{f}^2_\tau))$ for arbitrary $f^1_\tau\in\mathrm{Im}P^*_\tau$, and by applying Propositions A.2 and 4.6 to the first term, as well as $\|L^{-1}_\tau\|_{E'\to E}\leq 1$, and $\|(L^{-1}-L^{-1}_\tau)\|_{E'\to E}\leq 1$, the first claim follows.

The second claim follows from

$$\begin{aligned}\|\tilde{f}-P^*_\tau\tilde{f}\|_{E'} &= \sup_{0\neq v\in H^1_0(\Omega)}\frac{|\int_\Omega(\tilde{f}-P^*_\tau\tilde{f})(v-P_\tau v)|}{\|v\|_E}\\ &\leq \sup_{0\neq v\in H^1_0(\Omega)}\frac{\sum_{T\in\tau}\|\tilde{f}-P^*_\tau\tilde{f}\|_{L_2(T)}\|v-P_\tau v\|_{L_2(T)}}{\|v\|_E},\end{aligned}$$

and the application of (A.3) and (A.4). ☐

We apply Proposition A.2 as follows. Let us redefine $\mathbf{MARK}[\tau,\bar{\boldsymbol{f}}^2,\bar{u}_\tau]\to F$

% $\bar{\boldsymbol{f}}^2\in[\mathbb{V}^*_\tau]^n$, $\bar{u}_\tau\in\mathbb{V}_\tau$.
Select, in $\mathcal{O}(\#\tau)$ operations, a set $F\subset E_\tau$ with, up to some absolute factor, minimal cardinality *such that*

$$\sum_{e\in F}\eta_e(\bar{\boldsymbol{f}}^2,\bar{u}_\tau)\geq\theta^2\sum_{e\in E_\tau}\eta_e(\bar{\boldsymbol{f}}^2,\bar{u}_\tau),$$

$\mathbf{REFINE}[\tau,F]\to\tau'$
% Determines the smallest $\tau'\supseteq\tau$ for which all true
% hyperfaces $e\in F$ as well as all $T\in\tau$ adjacent to some $e\in F$ contain a
% vertex of τ' in their interiors,

$\mathbf{RHS}_f[\tau,\delta]\to[\tau',f^1_{\tau'},\boldsymbol{f}^2_{\tau'}]$
% $\delta>0$. The output consists of $f^1_{\tau'}\in\mathrm{Im}P^*_{\tau'}$, and $\boldsymbol{f}^2_{\tau'}\in[\mathbb{V}^*_{\tau'}]^n$, where $\tau'=\tau$, or,
% if necessary, $\tau'\supset\tau$, such that $\|f-f_{\tau'}\|_{E'}\leq\delta$,

and analogously, $\mathbf{RHS}_g[\tau,\delta]$. Then with C_1 reading as $\bar{C}_1$, and $\mathcal{E}(\tau_p,f^1_{\tau_p},\boldsymbol{f}^2_{\tau_p},\bar{u}_{\tau_p})$ as $\sum_{e\in E_{\tau_p}}\eta_e(\boldsymbol{f}^2_{\tau_p},\bar{u}_{\tau_p})$, and similarly at the dual side, all statements concerning the algorithm **GOAFEM** from Sect. 6 are still valid, whereas the evaluation of energy error estimators at primal and dual side has become cheaper.

Remark A.4. Thinking of approximations $f^1_\tau+\mathrm{div}\boldsymbol{f}^2_\tau$ for f with $\boldsymbol{f}^2_\tau=0$, thanks to (A.4) one can expect that the approximations produced by the new routine $\mathbf{RHS}_f$ are qualitatively as good as those by the old one. Since, because of (A.3), $P_\tau: H^1_0(\Omega)\to H^1_0(\Omega)$ is bounded, uniformly in τ, P^*_τ has a unique extension to a uniformly bounded operator $H^{-1}(\Omega) \to H^{-1}(\Omega)$, reading as $P_\tau^* f = \sum_{a \in N_\tau} f(\phi_a)\psi_a$. Therefore, a natural choice for $f^1_{\tau'}$ in $[\tau', f^1_{\tau'}] = \mathbf{RHS}_f[\tau, \delta]$ is $f^1_{\tau'} = P^*_{\tau'} f$, for simplicity ignoring quadrature errors. Since $(P^*_{\tau'} f)(v_{\tau'}) = f(v_{\tau'})$ for all $v_{\tau'} \in \mathbb{V}_{\tau'}$, instead of constructing an approximation to f, in this case a call $\mathbf{RHS}_f[\tau, \delta]$ merely consists of producing a $\tau' \supseteq \tau$ with $\|f - P^*_\tau f\|_{H^{-1}(\Omega)} \le \delta$, which for $f \in L_2(\Omega)$ can be done based on the second estimate from Corollary A.3. Obviously, similar remarks apply to approximations for g.

REFERENCES

[AO00] M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis*. Pure and Applied Mathematics. Wiley-Interscience, New York, 2000.

[BR96] R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: basic analysis and examples. *East-West J. Numer. Math.*, 4(4):237–264, 1996.

[BDD04] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219 – 268, 2004.

[BDDP02] P. Binev, W. Dahmen, R. DeVore, and P. Petruchev. Approximation classes for adaptive methods. *Serdica Math. J.*, 28:391–416, 2002.

[BEG91] J. Bernsten, T. O. Espelid, and A. Genz. An adaptive algorithm for the approximate calculation of multiple integrals. *ACM Transactions on Mathematical Software*, 17(4):437–451, 1991.

[BMN02] E. Bänsch, P. Morin, and R. Nochetto. An adaptive Uzawa FEM for the Stokes problem: Convergence without the inf-sup condition. *SIAM J. Numer. Anal.*, 40:1207–1229, 2002.

[BR03] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.

[BS01] I. Babuška and T. Strouboulis. *The finite element method and its reliability*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York, 2001.

[CV99] C. Carstensen and R. Verfürth. Edge residuals dominate a posteriori error estimates for lower order finite element methods. *SIAM J. Numer. Anal.*, 36(5):1571–1587, 1999.

[DKV06] W. Dahmen, A. Kunoth, and J. Vorloeper. Convergence of adaptive wavelet methods for goal-oriented error estimation. In *Numerical mathematics and advanced applications*, pages 39–61. Springer, Berlin, 2006.

[Dör96] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996.

[Hah05] T. Hahn. Cuba-a library for multidimensional numerical integration. *Computer Physics Communications*, 168(2):78–95, 2005.

[KS08] Y. Kondratyuk and R.P. Stevenson. An optimal adaptive finite element method for the Stokes problem. *SIAM J. Numer. Anal.*, 46(2):747–775, 2008.

[Mau95] J. Maubach. Local bisection refinement for n-simplicial grids generated by reflection. *SIAM J. Sci. Comput.*, 16(1):210–227, 1995.

[MNS00] P. Morin, R. Nochetto, and K. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488, 2000.

[MS08] M. Mommer and R.P. Stevenson. A goal-oriented adaptive finite element method with convergence rates – extended version. Preprint, Korteweg–De Vries Institute for Mathematics, University of Amsterdam, May 2008. Extended preprint version of current work. On http://staff.science.uva.nl/ rstevens/publ.html

[MvSST06] K.-S. Moon, E. von Schwerin, A. Szepessy, and R. Tempone. Convergence rates for an adaptive dual weighted residual finite element algorithm. *BIT*, 46(2):367–407, 2006.

[Neu03] N. Neuss. On using Common Lisp in scientific computing. In *Proceedings of the CISC 2002*. Springer-Verlag, 2003.

[OR76] J. T. Oden and J. N. Reddy. *An introduction to the mathematical theory of finite elements*. Wiley-Interscience [John Wiley & Sons], New York, 1976. Pure and Applied Mathematics.

[Osw94] P. Oswald. *Multilevel finite element approximation: Theory and applications*. B.G.

Teubner, Stuttgart, 1994.

[Ste05] R.P. Stevenson. An optimal adaptive finite element method. *SIAM J. Numer. Anal.*, 42(5):2188–2217, 2005.

[Ste07] R.P. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007.

[Ste08] R.P. Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77:227–241, 2008.

[SZ90] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[Tra97] C. T. Traxler. An algorithm for adaptive mesh refinement in n dimensions. *Computing*, 59(2):115–137, 1997.

[Ver96] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, Chichester, 1996.