ASC Report No. 21/2019

Instance-optimal goal-oriented adaptivity

M. Innerberger and D. Praetorius

Institute for Analysis and Scientific Computing
Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-00-1

Most recent ASC Reports

20/2019 *C.-M. Pfeiler and D. Praetorius*
Dörfler marking with minimal cardinality is a linear complexity problem

19/2019 *A. Jüngel, O. Leingang, and S. Wang*
Vanishing cross-diffusion limit in a Keller-Segel system with additional cross-diffusion

18/2019 *D. Praetorius, M. Ruggeri, and E.P. Stephan*
The saturation assumption yields optimal convergence of two-level adaptive BEM

17/2019 *G. Di Fratta and A. Fiorenza*
BMO-type seminorms from Escher-type tessellations

16/2019 *R. Pruckner, H. Woracek*
Limit behaviour of Nevanlinna functions

15/2019 *G. Di Fratta, C. Muratov, F. Rybakov, and V. Slastikov*
Variational principles of micromagnetics revisited

14/2019 *E. Davoli, G. Di Fratta*
Homogenization of chiral magnetic materials - A mathematical evidence of Dzyaloshinskii's predictions on helical structures

13/2019 *J.M. Melenk, A. Rieder*
On superconvergence of Runge-Kutta convuolution quadrature for the wave equation

12/2019 *J.M. Melenk, A. Rieder*
hp-FEM for the fractional heat equation

11/2019 *J.M. Melenk, S.A. Sauter, and C. Torres*
Wave number-explicit analysis for Galerkin discretizations of lossy Helmholtz problems

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: admin@asc.tuwien.ac.at
WWW: http://www.asc.tuwien.ac.at
FAX: +43-1-58801-10196

ISBN 978-3-902627-00-1

INSTANCE-OPTIMAL GOAL-ORIENTED ADAPTIVITY

MICHAEL INNERBERGER AND DIRK PRAETORIUS

ABSTRACT. We consider an adaptive finite element method with arbitrary but fixed polynomial degree $p \geq 1$, where adaptivity is driven by an edge-based residual error estimator. Based on the modified maximum criterion from [Diening et al, *Found. Comput. Math.* 16, 2016], we propose a goal-oriented adaptive algorithm and prove that it is instance optimal. Numerical experiments underline our theoretical findings.

1. INTRODUCTION

1.1. Rate optimality vs. instance optimality of AFEM. For an elliptic PDE with sought solution $u \in H^1_0(\Omega)$, the adaptive finite element method (AFEM) iterates the loop

$$\boxed{SOLVE} \longrightarrow \boxed{ESTIMATE} \longrightarrow \boxed{MARK} \longrightarrow \boxed{REFINE} \tag{1}$$

to successively adapt an initial mesh $\mathcal{T}_0$ to the possible singularities of the sought solution u. This leads to a sequence of meshes $\mathcal{T}_\ell$ and corresponding discrete solutions $u_\ell \in \mathcal{S}^p_0(\mathcal{T}_\ell) \subseteq H^1_0(\Omega)$ being $\mathcal{T}_\ell$-piecewise polynomials of degree $p \geq 1$. In the last decades, the mathematical understanding of AFEM has matured. We refer to [Dör96, MNS00, BDD04, Ste07, CKNS08, FFP14] for some milestones of the analysis of convergence with optimal algebraic rates, to [MSV08, Sie11] for a general theory on (plain) convergence, and to [CFPP14] for an abstract approach for optimal convergence rates.

The works [Ste07, CKNS08, FFP14, CFPP14] aim for rate optimality of adaptive algorithms, i.e., they prove that the usual adaptive loop (1), based on (quasi-) minimal Dörfler marking [Dör96] for the step *MARK*, leads to optimal decay of the *total error*

$$\mathrm{error}(\mathcal{T}_H) := \|\!| u - u_H |\!\| + \mathrm{osc}_H(\mathcal{T}_H), \tag{2}$$

being the sum of energy error $\|\!| u - u_H |\!\|$ plus data oscillation terms (or, equivalently, the error estimator, since $\eta_H \simeq \mathrm{error}(\mathcal{T}_H)$). Let $\#(\cdot)$ denote the number of elements of a finite set. Then, rate optimality reads as follows: For all $s > 0$, it holds that

$$\|u\|_{\mathbb{A}_s} < \infty \quad \Longrightarrow \quad \exists\, C > 0\, \forall\, \ell \in \mathbb{N}_0 : \quad \mathrm{error}(\mathcal{T}_\ell) \leq C\, (\#\mathcal{T}_\ell)^{-s}, \tag{3}$$

Date: August 1, 2019.

2010 *Mathematics Subject Classification.* 65N30, 41A25, 65N12, 65N50.

Key words and phrases. Adaptive finite element method, goal-oriented algorithm, quantity of interest, maximum marking strategy, convergence, instance optimality.

Acknowledgement: The authors acknowledge support through the Austrian Science Fund (FWF) through the doctoral school *Dissipation and dispersion in nonlinear PDEs* (grant W1245), the special research program *Taming complexity in PDE systems* (grant SFB F65), and the stand-alone project *Optimal adaptivity for BEM and FEM-BEM coupling* (grant P27005).

where $\|u\|_{\mathbb{A}_s} < \infty$ corresponds to certain approximation properties of u (which can be characterized in terms of Besov regularity [BDDP02, GM14, Gan17]).

Unlike this, the first work [BDD04] on AFEM with convergence rates implicitly proved even instance optimality of AFEM, i.e., the total error on an adaptive mesh is quasi-optimal with respect to all refinements of $\mathcal{T}_0$, which have essentially the same number of elements: It holds that

$$\exists C > 1 \, \forall \ell \in \mathbb{N}_0 \, \forall \mathcal{T}_H \in \operatorname{refine}(\mathcal{T}_0): \quad \big[C \, \#\mathcal{T}_H \le \#\mathcal{T}_\ell \implies \operatorname{error}(\mathcal{T}_\ell) \le C \operatorname{error}(\mathcal{T}_H)\big]. \tag{4}$$

The key argument for the proof of (4) in [BDD04] is an additional coarsening step in the adaptive loop (1). While [BDD04] also employed the Dörfler marking criterion [Dör96] for the step *MARK*, the work [DKS16] proposed a modified maximum criterion to single out edges for refinement. For P1-AFEM for the 2D Poisson problem, [DKS16] then proved instance optimality (4) of their adaptive strategy without resorting to an additional coarsening step. In [KS16], their analysis was extended to AFEM with non-conforming P1 elements for the Poisson problem and the Stokes system in 2D. We stress that any instance-optimal AFEM (4) is, in particular, rate optimal (3).

1.2. Goal-oriented adaptivity. While standard adaptivity aims to approximate the PDE solution $u \in H^1_0(\Omega)$ by some discrete approximation $u_\ell \in \mathcal{S}^p_0(\mathcal{T}_\ell)$ in the energy norm, a goal-oriented adaptive finite element method (GOAFEM) aims only to approximate $G(u)$ by $G(u_\ell)$, where $G : H^1_0(\Omega) \to \mathbb{R}$ is the so-called *goal functional* or *quantity of interest.*

In the present paper, we consider the following problem: Let $\Omega \subset \mathbb{R}^2$ be a polygonal Lipschitz domain, which is resolved by the initial mesh $\mathcal{T}_0$, where $\mathcal{T}_0$ is admissible in the sense of [BDD04, Ste08]. Given $f, g \in L^2(\Omega)$ and $\boldsymbol{f}, \boldsymbol{g} \in [L^2(\Omega)]^2$, the (linear) goal functional $G \in H^{-1}(\Omega)$ reads

$$G(u) := \int_\Omega gu - \boldsymbol{g} \cdot \nabla u \, \mathrm{d}x, \tag{5}$$

where $u \in H^1_0(\Omega)$ is the unique solution to

$$-\operatorname{div}(A\nabla u) = f + \operatorname{div} \boldsymbol{f} \quad \text{in } \Omega, \tag{6a}$$

$$u = 0 \quad \text{on } \Gamma := \partial\Omega. \tag{6b}$$

For technical reasons, we assume that the diffusion matrix $A \in [L^\infty(\Omega)]^{2\times 2}$ is $\mathcal{T}_0$-piecewise constant and $A|_T \in \mathbb{R}^{2\times 2}$ is symmetric and positive definite. Moreover, we assume that the restrictions $\boldsymbol{f}|_T, \boldsymbol{g}|_T$ are smooth for all $T \in \mathcal{T}_0$.

Convergence and rate-optimality of GOAFEM has been addressed in [MS09, BET11, HPZ15, HP16, FGH+16, FPZ16]. The key idea of the argument is to let $u^* \in H^1_0(\Omega)$ be the unique solution to the *dual problem*

$$-\operatorname{div}(A\nabla u^*) = g + \operatorname{div} \boldsymbol{g} \quad \text{in } \Omega, \tag{7a}$$

$$u^* = 0 \quad \text{on } \Gamma. \tag{7b}$$

Throughout, quantities associated with the dual problem are indexed by an asterisk. We note that the (symmetric) primal problem (6) and the dual problem (7) coincide up to the right-hand side. Let $|\!|\!| v |\!|\!| := \|A^{1/2}\nabla v\|_{L^2(\Omega)}$ be the problem-induced energy norm. For

FEM approximations $u \approx u_\ell \in \mathcal{S}^p_0(\mathcal{T}_\ell)$ and $u^* \approx u^*_\ell \in \mathcal{S}^p_0(\mathcal{T}_\ell)$, standard duality arguments (together with the Galerkin orthogonality) lead to

$$
\begin{aligned}
|G(u) - G(u_\ell)| &\le \| u - u_\ell \| \, \| u^* - u^*_\ell \| \\
&\le \big[\| u - u_\ell \| + \mathrm{osc}_\ell(\mathcal{T}_\ell) \big] \, \big[\| u^* - u^*_\ell \| + \mathrm{osc}^*_\ell(\mathcal{T}_\ell) \big];
\end{aligned} \tag{8}
$$

see, e.g., [MS09, FPZ16]. Therefore, GOAFEM aims to control and steer the product of the total errors

$$
\mathrm{error}(\mathcal{T}_H)\,\mathrm{error}^*(\mathcal{T}_H) := \big[\| u - u_H \| + \mathrm{osc}_H(\mathcal{T}_H) \big] \, \big[\| u^* - u^*_H \| + \mathrm{osc}^*_H(\mathcal{T}_H) \big]. \tag{9}
$$

While [BET11, HPZ15, HP16] focus on linear convergence of GOAFEM, the works [MS09, FGH$^+$16, FPZ16] also prove rate optimality. All works employ variants of the Dörfler marking criterion [Dör96]: The seminal work [MS09] employs (quasi-) minimal Dörfler marking separately for the primal and the dual problem, and then uses the smaller set for *MARK*. Instead, [BET11] proposes a (quasi-) minimal Dörfler marking for some combined estimator. Both strategies guarantee rate optimality for the product of the total errors

$$
\|u\|_{\mathbb{A}_s} + \|u^*\|_{\mathbb{A}_t} < \infty \implies \exists\, C > 0 \, \forall\, \ell \in \mathbb{N}_0 : \ \mathrm{error}(\mathcal{T}_\ell)\,\mathrm{error}^*(\mathcal{T}_\ell) \le C\,(\#\mathcal{T}_\ell)^{-(s+t)}, \tag{10}
$$

where the possible algebraic rate $s + t$ now depends on the approximability properties of the primal and dual problem; see [MS09, FGH$^+$16, FPZ16].

1.3. Instance-optimal GOAFEM. The new GOAFEM algorithm can briefly be outlined as follows: *SOLVE* computes the FEM solution $u_\ell \in \mathcal{S}^p_0(\mathcal{T}_\ell)$ to the primal problem (6) and $u^*_\ell \in \mathcal{S}^p_0(\mathcal{T}_\ell)$ to the dual problem (7). *ESTIMATE* computes the corresponding residual error estimators η_ℓ and η^*_ℓ. *MARK* employs the modified maximum strategy from [DKS16] to obtain two sets of marked edges, namely $\overline{\mathcal{M}}_\ell$ with respect to η_ℓ and $\overline{\mathcal{M}}^*_\ell$ with respect to η^*_ℓ. With $n := \min\{\#\overline{\mathcal{M}}_\ell, \#\overline{\mathcal{M}}^*_\ell\}$, we then define $\mathcal{M}_\ell := \underline{\mathcal{M}}_\ell \cup \underline{\mathcal{M}}^*_\ell$, where $\underline{\mathcal{M}}_\ell \subseteq \overline{\mathcal{M}}_\ell$ and $\underline{\mathcal{M}}^*_\ell \subseteq \overline{\mathcal{M}}^*_\ell$ are arbitrary up to $\#\underline{\mathcal{M}}_\ell = n = \#\underline{\mathcal{M}}^*_\ell$. Finally, *REFINE* employs 2D newest vertex bisection (NVB) to generate the coarsest mesh $\mathcal{T}_{\ell+1}$, where all edges in $\mathcal{M}_\ell$ have been bisected once.

The main result of the present work states that the proposed GOAFEM is instance optimal (4) with respect to the total-error product, i.e.,

$$
\begin{aligned}
&\exists\, C > 1 \, \forall\, \ell \in \mathbb{N}_0 \, \forall\, \mathcal{T}_H, \mathcal{T}_{H^*} \in \mathrm{refine}(\mathcal{T}_0) : \\
&\quad \big[\, C \max\{\#\mathcal{T}_H, \#\mathcal{T}_{H^*}\} \le \#\mathcal{T}_\ell \implies \mathrm{error}(\mathcal{T}_\ell)\,\mathrm{error}^*(\mathcal{T}_\ell) \le C\,\mathrm{error}(\mathcal{T}_H)\,\mathrm{error}^*(\mathcal{T}_{H^*}) \,\big].
\end{aligned} \tag{11}
$$

Again, we note that this implies, in particular, rate optimality (3). On a technical side, we note that the seminal work [DKS16] is restricted to the lowest-order FEM discretization $p = 1$, while the present analysis also allows higher (but fixed) polynomial degrees $p \ge 1$. In this sense, the present work provides also the technical tools to generalize the instance-optimal AFEM of [DKS16] from $p = 1$ to general but fixed $p \ge 1$.

1.4. Outline. The remainder of this paper is organized as follows: In Section 2, we give a precise formulation of the modules *SOLVE*, *ESTIMATE*, *MARK*, and *REFINE* of the adaptive loop (1). In particular, we state the modified maximum criterion (Algorithm 1) from [DKS16] as well as our extension to GOAFEM (Algorithm 2). Then, we thoroughly formulate our GOAFEM algorithm (Algorithm 6) and state our main result that the proposed GOAFEM is instance optimal (Theorem 7). Section 3 reviews the proof of the seminal work [DKS16] in an abstract framework. Section 4 collects the technical results to generalize the seminal work [DKS16] from lowest-order FEM $p = 1$ to arbitrary polynomial degree $p \geq 1$ (Theorem 4). Thereafter, Section 5 gives the proof of Theorem 7. Some numerical experiments in Section 6 conclude this work and empirically compare the instance-optimal GOAFEM algorithm from the present work with the rate-optimal GOAFEM strategies from [MS09, BET11, FPZ16].

1.5. General notation. In all results, the involved constants (as well as their dependencies) are stated explicitly. In proofs, however, we write $\lesssim$ to abbreviate $\leq$ up to a multiplicative constant which is clear from the context. Moreover, we write $\simeq$ if both estimates, $\lesssim$ and $\gtrsim$, hold.

2. Main result

Before stating our main result, we discuss the particular modules of the adaptive loop (1) and fix the necessary notation.

2.1. *REFINE*. A mesh $\mathcal{T}_H$ is a conforming triangulation of Ω into non-degenerate compact triangles $T \in \mathcal{T}_H$. The edges of $\mathcal{T}_H$ are denoted by $\mathcal{E}_H$. The set of interior edges of $\mathcal{T}_H$ is denoted by $\mathcal{E}_H^\Omega$, i.e., for each $E \in \mathcal{E}_H^\Omega$, there exist unique $T, T' \in \mathcal{T}_H$ such that $E = T \cap T'$. The set of vertices of $\mathcal{T}_H$ is denoted by $\mathcal{V}_H$. We define the patches

$$\mathcal{T}_H(\omega) := \big\{T \in \mathcal{T}_H \,:\, T \cap \omega \neq \emptyset\big\} \quad \text{for all } \omega \subset \overline{\Omega}. \tag{12}$$

For vertices $z \in \mathcal{V}_H$, we abbreviate $\mathcal{T}_H(z) := \mathcal{T}_H(\{z\}) = \big\{T \in \mathcal{T}_H \,:\, z \in T\big\}$. For neighbors $T, T' \in \mathcal{T}_H$, we also consider the reduced edge patch

$$\mathcal{T}_H^{\rm red}(E) := \big\{T \in \mathcal{T}_H \,:\, E \subset \partial T\big\} = \{T, T'\} \quad \text{for } E = T \cap T' \in \mathcal{E}_H. \tag{13}$$

Similarly, we define the area associated to a set of triangles $\mathcal{U}_H \subseteq \mathcal{T}_H$ by

$$\Omega(\mathcal{U}_H) := \bigcup_{T \in \mathcal{U}_H} T \subseteq \overline{\Omega} \quad \text{for all } \mathcal{U}_H \subseteq \mathcal{T}_H. \tag{14}$$

For mesh-refinement, we employ an edge-based variant of newest vertex bisection (NVB) [Ste08]. We suppose that the initial mesh $\mathcal{T}_0$ is admissible in the sense of [BDD04, Ste08]: For all neighbors $T, T' \in \mathcal{T}_0$, the joint edge $E = T \cap T' \in \mathcal{E}_0$ is the reference edge of T if and only if it is also the reference edge of T'. While this assumption is unnecessary for the NVB algorithm [KPP13], it provides additional structure which is crucial in the instance-optimality analysis of [DKS16].

For a mesh $\mathcal{T}_H$ and a set $\mathcal{M}_H \subseteq \mathcal{E}_H$ of marked edges, let $\mathcal{T}_h := \text{refine}(\mathcal{T}_H, \mathcal{M}_H)$ be the coarsest NVB refinement of $\mathcal{T}_H$ such that all edges $E \in \mathcal{M}_H$ have been bisected. Moreover, we write $\mathcal{T}_h \in \text{refine}(\mathcal{T}_H)$, if $\mathcal{T}_h$ can be obtained by finitely many steps of NVB refinement. Then, $\mathbb{T} := \text{refine}(\mathcal{T}_0)$ is the set of all possible NVB refinements of $\mathcal{T}_0$. We note that NVB leads to uniformly shape-regular meshes in the sense of

$$C_{\rm mesh} := \sup_{\mathcal{T}_H \in \mathbb{T}} \max_{T \in \mathcal{T}_H} \frac{\text{diam}(T)^2}{|T|} < \infty, \tag{15}$$

where $|T|$ is the area of a triangle T.

2.2. ***SOLVE.*** As usual, the primal problem (6) is understood in weak form. The Lax–Milgram lemma guarantees existence and uniqueness of $u \in H^1_0(\Omega)$ such that

$$a(u,v) := \int_\Omega A\nabla u \cdot \nabla v \,\mathrm{d}x = \int_\Omega fv - \boldsymbol{f}\cdot\nabla v \,\mathrm{d}x =: F(v) \quad \text{for all } v \in H^1_0(\Omega). \tag{16}$$

We define the energy norm $|\!|\!| v |\!|\!| := \|A^{1/2}\nabla v\|_{L^2(\Omega)} = a(v,v)^{1/2}$ and note that $|\!|\!| \cdot |\!|\!| \simeq |\cdot|_{H^1(\Omega)} := \|\nabla(\cdot)\|_{L^2(\Omega)} \simeq \|\cdot\|_{H^1(\Omega)}$ on $H^1_0(\Omega)$. Let $\mathcal{T}_H \in \mathbb{T}$ and $p \geq 1$. For the discretization of (16), define the space of $\mathcal{T}_H$-piecewise polynomials

$$\mathcal{P}^p(\mathcal{T}_H) := \big\{ v \in L^2(\Omega) \,:\, \forall T \in \mathcal{T}_H \quad v|_T \text{ is a polynomial of degree } \leq p \big\} \tag{17}$$

as well as the conforming FEM spaces

$$\mathcal{S}^p(\mathcal{T}_H) := \mathcal{P}^p(\mathcal{T}_H) \cap C(\Omega) = \mathcal{P}^p(\mathcal{T}_H) \cap H^1(\Omega) \quad \text{and} \quad \mathcal{S}^p_0(\mathcal{T}_H) := \mathcal{S}^p(\mathcal{T}_H) \cap H^1_0(\Omega). \tag{18}$$

Again, the Lax–Milgram lemma proves the existence and uniqueness of $u_H \in \mathcal{S}^p_0(\mathcal{T}_H)$ such that

$$a(u_H, v_H) = F(v_H) \quad \text{for all } v_H \in \mathcal{S}^p_0(\mathcal{T}_H). \tag{19}$$

The dual problem (7) is solved analogously with $F(\cdot)$ in (16) and (19) being replaced by $G(\cdot)$ from (5). The Lax–Milgram lemma guarantees existence and uniqueness of the dual solution $u^* \in H^1_0(\Omega)$ and its FEM approximation $u^*_H \in \mathcal{S}^p_0(\mathcal{T}_H)$.

2.3. ***ESTIMATE.*** For *a posteriori* error estimation, we employ an edge-based residual error estimator. Let $|E|$ be the length of an edge E. For the primal problem (6) with $F = f + \operatorname{div}\boldsymbol{f} \in H^{-1}(\Omega)$, we define

$$\begin{aligned} \eta_H(E)^2 := {} & |E| \, \| [\![(A\nabla u_H + \boldsymbol{f})\cdot\boldsymbol{\nu}]\!] \|^2_{L^2(E)} \\ & + \sum_{T \in \mathcal{T}^{\mathrm{red}}_H(E)} |T| \, \|f + \operatorname{div}(A\nabla u_H + \boldsymbol{f})\|^2_{L^2(T)} \quad \text{for all } E \in \mathcal{E}_H, \end{aligned} \tag{20}$$

where $\boldsymbol{\nu}$ is a normal vector on E and $[\![\cdot]\!]$ denotes the jump across E. For the dual problem (7) with $G = g + \operatorname{div}\boldsymbol{g} \in H^{-1}(\Omega)$, we define analogously

$$\begin{aligned} \eta^*_H(E)^2 := {} & |E| \, \| [\![(A\nabla u^*_H + \boldsymbol{g})\cdot\boldsymbol{\nu}]\!] \|^2_{L^2(E)} \\ & + \sum_{T \in \mathcal{T}^{\mathrm{red}}_H(E)} |T| \, \|g + \operatorname{div}(A\nabla u^*_H + \boldsymbol{g})\|^2_{L^2(T)} \quad \text{for all } E \in \mathcal{E}_H. \end{aligned} \tag{21}$$

With this notation, we define

$$\eta_H(\mathcal{U}_H) := \Big(\sum_{E \in \mathcal{U}_H} \eta_H(E)^2 \Big)^{1/2} \text{ and } \eta^*_H(\mathcal{U}_H) := \Big(\sum_{E \in \mathcal{U}_H} \eta^*_H(E)^2 \Big)^{1/2} \text{for all } \mathcal{U}_H \subseteq \mathcal{E}_H. \tag{22}$$

With the L^2-orthogonal projections $\Pi_E : L^2(E) \to \mathcal{P}^{p-1}(E)$ and $\Pi_T : L^2(T) \to \mathcal{P}^{p-2}(T)$, where $\mathcal{P}^{-1}(T) := \{0\}$, the data oscillations read

$$\operatorname{osc}_H(T)^2 := |T| \, \|(1-\Pi_T)(f + \operatorname{div}\boldsymbol{f})\|^2_{L^2(T)} + \sum_{\substack{E \in \mathcal{E}_H \\ E \subset \partial T}} |T|^{1/2} \, \|(1-\Pi_E)\, [\![\boldsymbol{f}\cdot\boldsymbol{\nu}]\!] \|^2_{L^2(E)}. \tag{23}$$

Note that for $p = 1$ the volume term simply reads $|T| \, \|f + \operatorname{div} \boldsymbol{f}\|_{L^2(T)}^2$. The oscillations $\operatorname{osc}_H^*(T)$ for the dual problem are defined analogously with g and $\boldsymbol{g}$ instead of f and $\boldsymbol{f}$, respectively. We note that

$$\operatorname{osc}_H(T) \lesssim \eta_H(E) \text{ and } \operatorname{osc}_H^*(T) \lesssim \eta_H^*(E) \quad \text{for all } T \in \mathcal{T}_H,\ E \in \mathcal{E}_H \text{ with } E \subseteq \partial T, \tag{24}$$

where the hidden constant depends only on $C_{\rm mesh}$ from (15). For a subset $\mathcal{U}_H \subseteq \mathcal{T}_H$, we define $\operatorname{osc}_H(\mathcal{U}_H)$ and $\operatorname{osc}_H^*(\mathcal{U}_H)$ analogously to (22). We note that

$$C_{\rm rel}^{-1} \, \lvert\!\lvert\!\lvert u - u_H \rvert\!\rvert\!\rvert \le \eta_H(\mathcal{E}_H) \le C_{\rm eff} \big[\lvert\!\lvert\!\lvert u - u_H \rvert\!\rvert\!\rvert + \operatorname{osc}_H(\mathcal{T}_H) \big], \tag{25}$$

$$C_{\rm rel}^{-1} \, \lvert\!\lvert\!\lvert u^* - u_H^* \rvert\!\rvert\!\rvert \le \eta_H^*(\mathcal{E}_H) \le C_{\rm eff} \big[\lvert\!\lvert\!\lvert u^* - u_H^* \rvert\!\rvert\!\rvert + \operatorname{osc}_H^*(\mathcal{T}_H) \big], \tag{26}$$

where the reliability constant $C_{\rm rel} > 0$ depends only on $C_{\rm mesh}$ from (15), while the efficiency constant $C_{\rm eff} > 0$ depends additionally on p.

2.4. *MARK.*

Let $\mathcal{T}_H \in \mathbb{T}$. We define the tail of an edge $E \in \mathcal{E}_H$ by

$$\operatorname{tail}_H(E) := \mathcal{E}_H \setminus \mathcal{E}_{H,E}, \quad \text{where} \quad \mathcal{T}_{H,E} := \operatorname{refine}(\mathcal{T}_H, \{E\}), \tag{27}$$

i.e., the tail consists of all edges, which have to be refined to ensure conformity of the triangulation if E is bisected. Moreover, we define

$$\operatorname{tail}_H(\mathcal{U}_H) := \bigcup_{E \in \mathcal{U}_H} \operatorname{tail}_H(E) \quad \text{for all } \mathcal{U}_H \subseteq \mathcal{E}_H. \tag{28}$$

With these definitions, we recall the following modified maximum criterion from [DKS16], which leads to an instance-optimal AFEM.

Algorithm 1 (Modified maximum criterion).
Input: *Edges* $\mathcal{E}_H$ *with indicators* $\mu_H := \big(\mu_H(E)\big)_{E \in \mathcal{E}_H}$, *marking parameter* $0 < \vartheta < 1$.
Output: *Set* $\mathcal{M}_H := \operatorname{markAFEM}(\mathcal{E}_H, \mu_H, \vartheta) \subseteq \mathcal{E}_H$ *of marked edges.*

$\mathcal{M}_H := \emptyset$ *and* $\mathcal{U} := \mathcal{E}_H$
$M := \max\big\{\mu_H(\operatorname{tail}_H(E)) \,:\, E \in \mathcal{E}_H\big\}$
while $\mathcal{U} \neq \emptyset$ ***do***
  pick $E \in \mathcal{U}$ *and update* $\mathcal{U} := \mathcal{U} \setminus \operatorname{tail}_H(E)$
  compute $m := \eta_H(\operatorname{tail}_H(E) \setminus \operatorname{tail}_H(\mathcal{M}_H))$
  if $m \ge \vartheta M$ ***then***
    $\mathcal{M}_H := \mathcal{M}_H \cup \{E\}$
  end if
end while

We refer to [DKS16] for a recursive implementation of Algorithm 1, which has linear costs. In case of GOAFEM, we need a different marking step that takes the primal as well as the dual problem into account. To this end, recall the Gauss brackets $\lfloor x \rfloor := \max\big\{n \in \mathbb{Z} \,:\, n \le x\big\}$ for $x \in \mathbb{R}$. We note that the following algorithm is slightly more general than the strategy outlined in Section 1.3 of the introduction (where $C_{\min} = 1$).

Algorithm 2 (Modified maximum criterion for GOAFEM).
Input: *Edges* $\mathcal{E}_H$*, indicators* $\eta_H := \big(\eta_H(E)\big)_{E\in\mathcal{E}_H}$ *and* $\eta_H^* := \big(\eta_H^*(E)\big)_{E\in\mathcal{E}_H}$*, marking parameters* $0 < \vartheta < 1$ *and* $C_{\min} > 0$*.*
Output: *Set* $\mathcal{M}_H := \mathrm{markGOAFEM}(\mathcal{E}_H, \eta_H, \eta_H^*, \vartheta, C_{\min}) \subseteq \mathcal{E}_H$ *of marked edges.*

1: *generate* $\overline{\mathcal{M}}_H := \mathrm{markAFEM}(\mathcal{E}_H, \eta_H, \vartheta)$ *by Algorithm 1*
2: *generate* $\overline{\mathcal{M}}_H^* := \mathrm{markAFEM}(\mathcal{E}_H, \eta_H^*, \vartheta)$ *by Algorithm 1*
3: *choose* $\mathcal{M}_{\min} := \arg\min\big\{\#\overline{\mathcal{M}}_H \,,\, \#\overline{\mathcal{M}}_H^*\big\}$ *and* $\mathcal{M}_{\max} := \big\{\overline{\mathcal{M}}_H \,,\, \overline{\mathcal{M}}_H^*\big\}\backslash\mathcal{M}_{\min}$
4: *define* $n := \min\big\{\#\mathcal{M}_{\max}, \max\{1, \lfloor C_{\min}\#\mathcal{M}_{\min}\rfloor\}\big\}$
5: *pick* $\mathcal{M}_{\max}' \subseteq \mathcal{M}_{\max}$ *with* $\#\mathcal{M}_{\max}' = n$
6: *choose* $\mathcal{M}_H = \mathcal{M}_{\min} \cup \mathcal{M}_{\max}'$

2.5. Instance-optimal AFEM. The work [DKS16] analyzes the following instance of the adaptive loop (1), which turns out to be instance-optimal; see Theorem 4.

Algorithm 3 (Instance-optimal AFEM).
Input: *Initial mesh* $\mathcal{T}_0$*, polynomial degree* $p \in \mathbb{N}$*, marking parameter* $0 < \vartheta < 1$*.*
Output: *Meshes* $\mathcal{T}_\ell$*, discrete solutions* u_ℓ*, and estimators* $\eta_\ell(\mathcal{E}_\ell)$ *for all* $\ell \in \mathbb{N}_0$*.*

1: ***for all*** $\ell = 0, 1, 2, \dots$ ***do***
2:     *compute FEM solution* $u_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell)$
3:     *compute indicators* $\eta_\ell = \big(\eta_\ell(E)\big)_{E\in\mathcal{E}_\ell}$
4:     *generate* $\mathcal{M}_\ell := \mathrm{markAFEM}(\mathcal{E}_\ell, \eta_\ell, \vartheta)$ *by Algorithm 1*
5:     *employ NVB to generate* $\mathcal{T}_{\ell+1} = \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$
6: ***end for***

For $p = 1$, the following theorem is the main result of [DKS16]. Our analysis below implies that the result remains true for arbitrary polynomial degrees $p \geq 1$.

Theorem 4. *Let the initial mesh* $\mathcal{T}_0$ *be admissible in the sense of* [BDD04]*. Let* $p \in \mathbb{N}$ *and* $0 < \vartheta < 1$*. Then, the AFEM Algorithm 3 for the primal problem* (6) *is instance optimal with respect to the total error, i.e.,*

$$(29)\quad \exists\, C > 1\, \forall\, \ell \in \mathbb{N}_0\, \forall\, \mathcal{T}_H \in \mathrm{refine}(\mathcal{T}_0):$$
$$\Big(C\,\#(\mathcal{T}_H\backslash\mathcal{T}_0) \leq \#(\mathcal{T}_\ell\backslash\mathcal{T}_0) \implies |\!|\!|\, u - u_\ell\, |\!|\!|^2 + \mathrm{osc}_\ell(\mathcal{T}_\ell)^2 \leq C\big[|\!|\!|\, u - u_H\, |\!|\!|^2 + \mathrm{osc}_H(\mathcal{T}_H)^2\big]\Big).$$

The constant C *depends only on* ϑ*,* p*,* C_{mesh}*, and the data* A*,* f*,* $\boldsymbol{f}$*.*

Remark 5. *We note that elementary calculation shows that, for all* $\mathcal{T}_H \in \mathbb{T}\backslash\{\mathcal{T}_0\}$*,*

$$\#\mathcal{T}_H - \#\mathcal{T}_0 \leq \#(\mathcal{T}_H\backslash\mathcal{T}_0) \leq \#\mathcal{T}_H \leq (\#\mathcal{T}_0)\,(\#\mathcal{T}_H - \#\mathcal{T}_0 + 1) \leq 2\,(\#\mathcal{T}_0)\,(\#\mathcal{T}_H - \#\mathcal{T}_0);$$

see, e.g., [BHP17, Lemma 22]*. Hence,* $\#(\mathcal{T}_\ell\backslash\mathcal{T}_0)$ *in* (29) *can, in fact, be replaced by* $\#\mathcal{T}_\ell$ *(at the cost that the constant* C *in* (4) *will additionally depend on* $\#\mathcal{T}_0$*). Therefore, the statement of Theorem 4 is equivalent to the introductory statement of instance optimality* (4) *in Section 1.1.*

2.6. Instance-optimal GOAFEM. As outlined in the introduction, the main idea behind GOAFEM is the duality-based estimate

$$|\,G(u) - G(u_\ell)\,| \;=|\,a(u-u_\ell, u^*)\,| \;=|\,a(u-u_\ell, u^*-u_\ell^*)\,| \le \;|\!|\!|\, u-u_\ell \,|\!|\!|\,|\!|\!|\, u^*-u_\ell^* \,|\!|\!|,$$

The formal statement of our GOAFEM algorithm reads as follows:

Algorithm 6 (Instance-optimal GOAFEM).
Input: *Initial mesh $\mathcal{T}_0$, polynomial degree $p \in \mathbb{N}$, marking parameters $0 < \vartheta < 1$ and $C_{\min} > 0$.*
Output: *Meshes $\mathcal{T}_\ell$, discrete solutions u_ℓ, u_ℓ^*, estimators $\eta_\ell(\mathcal{E}_\ell), \eta_\ell^*(\mathcal{E}_\ell)$ and goal quantities $G(u_\ell)$ for all $\ell \in \mathbb{N}_0$.*

1: ***for all*** $\ell = 0, 1, 2, \ldots$ ***do***
2: *compute FEM solutions* $u_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell)$ *and* $u_\ell^* \in \mathcal{S}_0^p(\mathcal{T}_\ell)$
3: *compute indicators* $\eta_\ell = \big(\eta_\ell(E)\big)_{E\in\mathcal{E}_\ell}$ *and* $\eta_\ell^* = \big(\eta_\ell^*(E)\big)_{E\in\mathcal{E}_\ell}$
4: *generate* $\mathcal{M}_\ell := \mathrm{markGOAFEM}(\mathcal{E}_\ell, \eta_\ell, \eta_\ell^*, \vartheta, C_{\min})$ *by Algorithm 2*
5: *employ NVB to generate* $\mathcal{T}_{\ell+1} = \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$
6: ***end for***

The following theorem is the main result of this work. We stress that the theorem involves the adaptively generated mesh $\mathcal{T}_\ell$ for the primal and the dual error and compares it with arbitrary meshes $\mathcal{T}_H$ and $\mathcal{T}_{H^*}$, where $\mathcal{T}_H$ is used for the primal error and $\mathcal{T}_{H^*}$ is used for the dual error.

Theorem 7. *Let the initial mesh $\mathcal{T}_0$ be admissible in the sense of* [BDD04]. *Let $p \in \mathbb{N}$ and $0 < \vartheta < 1$ as well as $C_{\min} > 0$. Let $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ be the sequence of meshes generated by Algorithm 6. Then, the AFEM Algorithm 3 is instance optimal with respect to the product of total errors, i.e.,* $\exists\, C > 1\, \forall\, \ell \in \mathbb{N}_0\, \forall\, \mathcal{T}_H, \mathcal{T}_{H^*} \in \mathrm{refine}(\mathcal{T}_0)$:

$$\begin{aligned}&\Big(C\, \max\{\#(\mathcal{T}_H\backslash\mathcal{T}_0), \#(\mathcal{T}_{H^*}\backslash\mathcal{T}_0)\} \le \#(\mathcal{T}_\ell\backslash\mathcal{T}_0)\\ &\qquad\Longrightarrow \big[\,|\!|\!|\, u-u_\ell \,|\!|\!|^2 + \mathrm{osc}_\ell(\mathcal{T}_\ell)^2\,\big]\big[\,|\!|\!|\, u^*-u_\ell^* \,|\!|\!|^2 + \mathrm{osc}_\ell^*(\mathcal{T}_\ell)^2\,\big]\\ &\qquad\quad\le C\big[\,|\!|\!|\, u-u_H \,|\!|\!|^2 + \mathrm{osc}_H(\mathcal{T}_\ell)^2\,\big]\big[\,|\!|\!|\, u^*-u_{H^*}^* \,|\!|\!|^2 + \mathrm{osc}_{H^*}^*(\mathcal{T}_\ell)^2\,\big]\Big).\end{aligned} \tag{30}$$

The constant C depends only on ϑ, p, C_{mesh} $C_{\min}$, and the data A, f, $\boldsymbol{f}$, g, $\boldsymbol{g}$.

Remark 8. *Note that the natural statement of instance-optimality for GOAFEM in the sense of* (4) *and* (9) *would be:* $\exists\, C > 1\, \forall\, \ell \in \mathbb{N}_0\, \forall\, \mathcal{T}_H \in \mathrm{refine}(\mathcal{T}_0)$ *such that*

$$\begin{aligned}&\Big(C\#(\mathcal{T}_H\backslash\mathcal{T}_0) \le \#(\mathcal{T}_\ell\backslash\mathcal{T}_0)\\ &\qquad\Longrightarrow \big[\,|\!|\!|\, u-u_\ell \,|\!|\!|^2 + \mathrm{osc}_\ell(\mathcal{T}_\ell)^2\,\big]\big[\,|\!|\!|\, u^*-u_\ell^* \,|\!|\!|^2 + \mathrm{osc}_\ell^*(\mathcal{T}_\ell)^2\,\big]\\ &\qquad\quad\le C\big[\,|\!|\!|\, u-u_H \,|\!|\!|^2 + \mathrm{osc}_H(\mathcal{T}_\ell)^2\,\big]\big[\,|\!|\!|\, u^*-u_H^* \,|\!|\!|^2 + \mathrm{osc}_H^*(\mathcal{T}_\ell)^2\,\big]\Big).\end{aligned}$$

Our Theorem 7, however, is stronger. There, the mesh for the right-hand side can be chosen for both factors independently.

---

3. Abstract result on instance optimality

This section aims to review the proof of [DKS16, Theorem 7.3] in an abstract framework. For arbitrary $m \in \mathbb{N}$, the tuple $(\mathcal{T}_H, \mathcal{T}_h; \mathcal{T}_1, \dots, \mathcal{T}_m) \in \mathbb{T}^{m+2}$ is a *diamond*, if

- $\mathcal{T}_j \in \mathbb{T}$ are meshes for all $j = 1, \dots, m$
- with finest common coarsening $\mathcal{T}_H \in \mathbb{T}$ and coarsest common refinement $\mathcal{T}_h \in \mathbb{T}$
- such that the areas $\Omega(\mathcal{T}_j \backslash \mathcal{T}_h)$ are pairwise disjoint for all $j = 1, \dots, m$.

Note that $\mathcal{T}_H, \mathcal{T}_h \in \mathbb{T}$ exist (and are unique), since newest vertex bisection is a a binary refinement rule, where the order of refinements does not matter. This allows to write

$$\mathcal{T}_H = \bigcup_{j=1}^m \big\{ T \in \mathcal{T}_j \, : \, \forall k \in \{1, \dots, m\} \, \forall T' \in \mathcal{T}_k \quad \big(T \subseteq T' \implies T = T' \big) \big\},$$
$$\mathcal{T}_h = \bigcup_{j=1}^m \big\{ T \in \mathcal{T}_j \, : \, \forall k \in \{1, \dots, m\} \, \forall T' \in \mathcal{T}_k \quad \big(T' \subseteq T \implies T = T' \big) \big\}.$$

Moreover, according to [DKS16], this also implies that

$$\mathcal{T}_H \setminus \mathcal{T}_h = \bigcup_{j=1}^m \big(\mathcal{T}_j \setminus \mathcal{T}_h \big). \tag{31}$$

Diamonds are a means to couple the lattice structure of $\mathbb{T}$ with an abstract energy

$$\mathbb{E} : \mathbb{T} \to \mathbb{R}_{\geq 0}. \tag{32}$$

Only energies that are compatible with this structure are suitable to prove instance optimality. This is encoded in the following properties (A1)–(A5), where C_{mark}, C_{diam}, C_{est}, $C_{\mathrm{err}} > 0$ are generic constants, η_h is a computable edge-based estimator, and `mark` is an abstract marking strategy:

(A1) Marking criterion: For all meshes $\mathcal{T}_H \in \mathbb{T}$ with edges $\mathcal{E}_H$, the marking strategy guarantees that the marked edges $\mathcal{M}_H := \mathtt{mark}(\mathcal{E}_H, (\eta_H(E))_{E \in \mathcal{E}_H})$ satisfy that

$$\mathcal{M}_H \neq \emptyset \quad \text{and} \quad \eta_H^2(\mathrm{tail}_H(\mathcal{M}_H)) \geq C_{\mathrm{mark}} \, (\#\mathcal{M}_H) \, \max_{E \in \mathcal{E}_H} \eta_H^2(\mathrm{tail}_H(E)).$$

(A2) Monotonicity of energy: For all $\mathcal{T}_H \in \mathbb{T}$ and all $\mathcal{T}_h \in$ refine$(\mathcal{T}_H)$, it holds that

$$0 \leq \mathbb{E}(\mathcal{T}_h) \leq \mathbb{E}(\mathcal{T}_H).$$

(A3) Diamond estimate: For all diamonds $(\mathcal{T}_H, \mathcal{T}_h; \mathcal{T}_1, \dots, \mathcal{T}_m) \in \mathbb{T}^{m+2}$, it holds that

$$C_{\mathrm{diam}}^{-1} \big[\mathbb{E}(\mathcal{T}_H) - \mathbb{E}(\mathcal{T}_h) \big] \leq \sum_{j=1}^m \big[\mathbb{E}(\mathcal{T}_j) - \mathbb{E}(\mathcal{T}_h) \big] \leq C_{\mathrm{diam}} \big[\mathbb{E}(\mathcal{T}_H) - \mathbb{E}(\mathcal{T}_h) \big].$$

(A4) Local energy estimates for the estimator: For all $\mathcal{T}_H \in \mathbb{T}$ and $\mathcal{T}_h \in$ refine$(\mathcal{T}_H)$, it holds that

$$C_{\mathrm{est}}^{-1} \big[\eta_H(\mathcal{E}_H \setminus \mathcal{E}_h) \big] \leq \mathbb{E}(\mathcal{T}_H) - \mathbb{E}(\mathcal{T}_h) \leq C_{\mathrm{est}} \big[\eta_H(\mathcal{E}_H \setminus \mathcal{E}_h) \big].$$

(A5) Equivalence of energy and total error: For all $\mathcal{T}_H \in \mathbb{T}$, it holds that

$$C_{\mathrm{err}}^{-1} \, \mathbb{E}(\mathcal{T}_H) \leq |||\, u - u_H \,|||^2 + \mathrm{osc}_H(\mathcal{T}_H)^2 \leq C_{\mathrm{err}} \, \mathbb{E}(\mathcal{T}_H)$$

As can be seen from the proof of [DKS16, Theorem 7.3], the conditions (A1)–(A4) are sufficient for an AFEM to be instance optimal with respect to $\mathbb{E}$. Moreover, condition (A5) allows to derive instance optimality with respect to the total error. We formulate this as a theorem, but refer to [DKS16] for the proof.

Theorem 9. *Consider an AFEM loop as given by* (1)*, which satisfies the conditions* (A1)–(A4)*. Then, the AFEM is instance optimal with respect to the energy, i.e.,*

$$\exists\, C>1\,\forall\, \ell\in\mathbb{N}_0\,\forall\, \mathcal{T}_H\in\mathrm{refine}(\mathcal{T}_0):\ \big[C\,\#(\mathcal{T}_H\backslash\mathcal{T}_0)\leq\#(\mathcal{T}_\ell\backslash\mathcal{T}_0)\implies\mathbb{E}(\mathcal{T}_\ell)\leq\mathbb{E}(\mathcal{T}_H)\big]. \tag{33}$$

If (A5) *is satisfied in addition, then the AFEM is instance optimal in the sense of* (4)*.* □

Remark 10. *We note that the proof of Theorem 9 (resp.* [DKS16, Theorem 7.3]*) is currently tailored to 2D newest vertex bisection, for which structural properties are exploited (so-called* populations*). Besides this, the proof only relies on the given properties* (A1)–(A4)*, as was already observed in* [KS16]*.*

4. Auxiliary results

In this section, we recall how the properties (A1)–(A4) from Section 3 are proved for our model problem from Section 2. To that end, we define the energy (32) corresponding to a mesh $\mathcal{T}_H\in\mathbb{T}$ by

$$\mathbb{E}(\mathcal{T}_H):=\frac{1}{2}\,a(u_H,u_H)-F(u_H)-\Big[\frac{1}{2}\,a(u,u)-F(u)\Big]+\mathrm{osc}_H(\mathcal{T}_H)^2. \tag{34}$$

In particular, we generalize the analysis of [DKS16] from lowest-order FEM $p=1$ to arbitrary fixed order $p\geq 1$.

Remark 11. *Our definition follows* [DKS16]*, but is shifted to ensure* $\mathbb{E}(\mathcal{T}_H)\geq 0$ *for all* $\mathcal{T}_H\in\mathbb{T}$*. This is important, since the GOAFEM analysis involves energy products.*

4.1. Verification of (A1): Marking criterion.

In [DKS16, Proposition 5.1], it is shown that Algorithm 1 satisfies the marking criterion (A1) with $C_{\mathrm{mark}}=\vartheta$. We state the following proposition, which is a straightforward generalization of this result and follows from the same arguments.

Proposition 12. *Let* $\overline{\mathcal{M}}_H\subseteq\mathcal{E}_H$ *be the set of edges marked by Algorithm 1 for* $0<\vartheta<1$*. Then, any subset* $\mathcal{M}_H\subseteq\overline{\mathcal{M}}_H$ *with* $\mathcal{M}_H\neq\emptyset$ *satisfies* (A1) *with* $C_{\mathrm{mark}}=\vartheta$*.* □

4.2. Verification of (A2): Monotonicity of energy.

Recall that $u\in H^1_0(\Omega)$ solves the variational formulation (16) if and only if it minimizes the Dirichlet energy, i.e.,

$$\frac{1}{2}\,a(u,u)-F(u)=\inf_{v\in H^1_0(\Omega)}\Big[\frac{1}{2}\,a(v,v)-F(v)\Big]. \tag{35}$$

The same holds (with $H^1_0(\Omega)$ being replaced by $\mathcal{S}^p_0(\Omega)$) for the Galerkin formulation (19). By definition (23) of the data oscillations, this proves $\mathbb{E}(\mathcal{T}_H)\geq 0$. Moreover, from nestedness $\mathcal{S}^p(\mathcal{T}_H)\subseteq\mathcal{S}^p(\mathcal{T}_h)$, we obtain the monotonicity $\mathbb{E}(\mathcal{T}_h)\leq\mathbb{E}(\mathcal{T}_H)$ for all $\mathcal{T}_H\in\mathbb{T}$ and $\mathcal{T}_h\in\mathrm{refine}(\mathcal{T}_H)$.

4.3. Verification of (A5): Equivalence of energy and total error. It is well-known from variational calculus that

$$\mathbb{E}(\mathcal{T}_H) = \frac{1}{2} \| u - u_H \|^2 + \mathrm{osc}_H(\mathcal{T}_H)^2. \tag{36}$$

This proves (A5) even with constant $C_{\mathrm{err}} = 2$. Moreover, for $\mathcal{T}_h \in \mathrm{refine}(\mathcal{T}_H)$, the Galerkin orthogonality proves the identity

$$\mathbb{E}(\mathcal{T}_H) - \mathbb{E}(\mathcal{T}_h) = \frac{1}{2} \| u_H - u_h \|^2 + \mathrm{osc}_H(\mathcal{T}_H)^2 - \mathrm{osc}_h(\mathcal{T}_h)^2, \tag{37}$$

which will be exploited below.

4.4. Scott–Zhang projector. The key ingredient to prove (A3)–(A4) is a slight variant [DKS16] of the Scott–Zhang projector from [SZ90]: Suppose $\mathcal{T}_H \in \mathbb{T}$ and $\mathcal{T}_h \in \mathrm{refine}(\mathcal{T}_H)$. Let $\mathcal{L}_H$ denote the set of Lagrange nodes of of $\mathcal{S}^p_0(\mathcal{T}_H)$. For each $z \in \mathcal{L}_H$, choose a simplex $\sigma_{H,z} \in \mathcal{T}_H \cup \mathcal{E}_H$ subject to the following constraints:

(a) If $z \in T \in \mathcal{T}_H$ lies in the interior of T, choose $\sigma_{H,z} = T$.
(b) If $z \in E \in \mathcal{E}_H$ lies in the interior of E, choose $\sigma_{H,z} = E$.
(c) If $z \in \mathcal{V}_H$ with $z \in \Omega(\mathcal{T}_H \cap \mathcal{T}_h)$ (resp. $z \in \Omega(\mathcal{T}_H \setminus \mathcal{T}_h)$), choose $\sigma_{H,z} = E \in \mathcal{E}_H$ with $E \subseteq \Omega(\mathcal{T}_H \cap \mathcal{T}_h)$ (resp. $E \subseteq \Omega(\mathcal{T}_H \setminus \mathcal{T}_h)$).

For a Lagrange point $z \in \mathcal{L}_H$, let $\phi_{H,z} \in \mathcal{S}^p(\mathcal{T}_h)$ be the corresponding nodal basis function, i.e., it holds that $\phi_{H,z}(z') = \delta_{zz'}$ for all $z' \in \mathcal{L}_H$. Moreover, let $\psi_{H,z} \in \mathcal{P}^p(\sigma_{H,z})$ be the corresponding dual basis function with respect to $L^2(\sigma_{H,z})$, i.e., it holds that

$$\int_{\sigma_{H,z}} \psi_{H,z} \phi_{z'} \,\mathrm{d}x = \delta_{zz'} \quad \text{for all } z, z' \in \mathcal{L}_H. \tag{38}$$

Then, we consider the Scott–Zhang projector $\mathcal{Q}_{H,h} : H^1(\Omega) \to \mathcal{S}^p(\mathcal{T}_H)$ defined by

$$\mathcal{Q}_{H,h} v := \sum_{z \in \mathcal{L}_H} \phi_{H,z} \int_{\sigma_{H,z}} \psi_{H,z} v \,\mathrm{d}x \quad \text{for all } v \in H^1(\Omega). \tag{39}$$

The following proposition collects the relevant properties of $\mathcal{Q}_{H,h}$. We note that the definition guarantees that, for $v_h \in \mathcal{S}^p(\mathcal{T}_h)$, the restriction of $\mathcal{Q}_{H,h} v_h$ to $\Omega(\mathcal{T}_H \cap \mathcal{T}_h)$ (resp. $\Omega(\mathcal{T}_H \setminus \mathcal{T}_h)$) depends only on v_h restricted to $\Omega(\mathcal{T}_H \cap \mathcal{T}_h)$ (resp. $\Omega(\mathcal{T}_H \setminus \mathcal{T}_h)$). This is enforced by the choice (c) of $\sigma_{H,z}$.

Proposition 13. *Let $\mathcal{T}_H \in \mathbb{T}$ and $\mathcal{T}_h \in \mathrm{refine}(\mathcal{T}_H)$. Let $\mathcal{U} \in \{\mathcal{T}_H \cap \mathcal{T}_h\,,\, \mathcal{T}_H \setminus \mathcal{T}_h\}$. Then, there hold the following assertions* (i)–(vi)*, where $C_{\mathrm{sz}} > 0$ depends only on C_{mesh} and p:*

(i) $|\mathcal{Q}_{H,h} v|_{H^1(T)} \le C_{\mathrm{sz}} |v|_{H^1(\mathcal{T}_H(T))}$ *for all $T \in \mathcal{T}_H$ and $v \in H^1(\Omega)$.*
(ii) $\|(1 - \mathcal{Q}_{H,h}) v\|_{L^2(T)} \le C_{\mathrm{sz}} h_T |v|_{H^1(\mathcal{T}_H(T))}$ *for all $T \in \mathcal{T}_H$ and $v \in H^1(\Omega)$.*
(iii) $\|(1 - \mathcal{Q}_{H,h}) v\|_{L^2(E)} \le C_{\mathrm{sz}} h_E^{1/2} |v|_{H^1(\mathcal{T}_H(E))}$ *for all $E \in \mathcal{E}_H$ and $v \in H^1(\Omega)$.*
(iv) $(\mathcal{Q}_{H,h} v_h)|_{\Omega(\mathcal{U})}$ *depends only on $v_h|_{\Omega(\mathcal{U})}$ for all $v_h \in \mathcal{S}^p(\mathcal{T}_h)$.*
(v) $\big(\mathcal{Q}_{H,h} v_h - v_h\big)|_T = 0$ *for all $T \in \mathcal{T}_H \cap \mathcal{T}_h$ and all $v_h \in \mathcal{S}^p(\mathcal{T}_h)$.*
(vi) $\big(\mathcal{Q}_{H,h} v_h\big)|_\Gamma = 0$ *for all $v_h \in \mathcal{S}^p_0(\mathcal{T}_h)$.*

Proof. The claims (i)–(iii) are proved in [SZ90]. For (iv)–(vi), we refer to [DKS16, Lemma 3.5] (which directly transfers from $p = 1$ to $p \ge 1$). □

4.5. Verification of (A3): Diamond estimate. In order to prove the diamond estimate (A3), we employ the Scott–Zhang projector from the previous section. The following lemma is proved in [DKS16, Theorem 3.7] (which directly transfers from $p = 1$ to $p \geq 1$).

Lemma 14. *Let $(\mathcal{T}_H, \mathcal{T}_h; \mathcal{T}_1, \dots, \mathcal{T}_m) \in \mathbb{T}^{m+2}$ be a diamond and $p \in \mathbb{N}$. Then, the Scott–Zhang projectors $\mathcal{Q}_{i,h}$ commute pairwise and the projection*

$$\mathcal{Q}_{\circ,h} := \mathcal{Q}_{1,h} \circ \dots \circ \mathcal{Q}_{m,h} : \mathcal{S}^p_0(\mathcal{T}_h) \to \mathcal{S}^p_0(\mathcal{T}_H) \tag{40}$$

is well-defined and satisfies that

$$|\mathcal{Q}_{\circ,h} v_h|_{H^1(\Omega)} \leq C |v_h|_{H^1(\Omega)} \quad \text{for all } v_h \mathcal{S}^p_0(\mathcal{T}_h), \tag{41}$$

where $C > 0$ depends only on C_{mesh} and p. Moreover, with $\Omega_i := \Omega(\mathcal{T}_i \setminus \mathcal{T}_h)$ for $i = 1, \dots, m$, it holds that

$$\mathcal{Q}_{\circ,h} v_h = \begin{cases} \mathcal{Q}_{i,h} v_h & \text{on } \Omega_i, \\ v_h & \text{on } \Omega \setminus \bigcup_{i=1}^m \Omega_i. \end{cases} \tag{42}$$

□

With this auxiliary result, we can prove the diamond estimate (A3).

Proposition 15. *The diamond estimate* (A3) *holds with a constant $C_{\mathrm{diam}} > 0$ depending only on C_{mesh}, p, and A.*

Sketch of proof. The proof is split into three steps.

Step 1. From the best approximation property of FEM solutions with respect to the energy norm and the stability of Scott–Zhang operators, we infer that

$$|\!|\!| u_h - u_i |\!|\!| \simeq |\!|\!| u_h - \mathcal{Q}_{i,h} u_h |\!|\!| \quad \text{for all } i = 1, \dots, m;$$

see also [DKS16, Lemma 3.4]. This equivalence holds also for u_H and $\mathcal{Q}_{\circ,h}$ instead of u_i and $\mathcal{Q}_{i,h}$, respectively. Together with (42) and Proposition 13(v), we obtain that

$$|\!|\!| u_h - u_H |\!|\!|^2 \simeq |\!|\!| u_h - \mathcal{Q}_{\circ,h} u_h |\!|\!|^2 \overset{(42)}{=} \sum_{i=1}^m |\!|\!| u_h - \mathcal{Q}_{i,h} u_h |\!|\!|^2 \simeq \sum_{i=1}^m |\!|\!| u_h - u_i |\!|\!|^2. \tag{43}$$

Step 2. Let $\mathcal{T}_\bullet \in \mathbb{T}$ and $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$. Then, newest vertex bisection guarantees that

$$|T'| \leq \tfrac{1}{2}|T| \quad \text{for all } T \in \mathcal{T}_\bullet \setminus \mathcal{T}_\circ \text{ and } T' \in \mathcal{T}_\circ \setminus \mathcal{T}_\bullet \text{ with } T' \subset T.$$

Together with the fact that $[\![\boldsymbol{f} \cdot \boldsymbol{\nu}]\!]$ vanishes on all newly created edges, since $\operatorname{div} \boldsymbol{f} \in L^2(T)$ for every $T \in \mathcal{T}_0$, this shows the equivalence

$$\big(1 - \tfrac{1}{\sqrt{2}}\big) \mathrm{osc}_\bullet(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)^2 \leq \mathrm{osc}_\bullet(\mathcal{T}_\bullet)^2 - \mathrm{osc}_\circ(\mathcal{T}_\circ)^2 \leq \mathrm{osc}_\bullet(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)^2. \tag{44}$$

Step 3. We use (44) on the meshes $\mathcal{T}_h, \mathcal{T}_i \in \mathrm{refine}(\mathcal{T}_H)$. This yields that

$$\begin{aligned}\mathrm{osc}_H(\mathcal{T}_H)^2 - \mathrm{osc}_h(\mathcal{T}_h)^2 &\overset{(44)}{\simeq} \mathrm{osc}_H(\mathcal{T}_H \setminus \mathcal{T}_h,)^2 = \mathrm{osc}_H\Big(\bigcup_{i=1}^m(\mathcal{T}_i \setminus \mathcal{T}_h)\Big)^2 \\ &= \sum_{i=1}^m \mathrm{osc}_H(\mathcal{T}_i \setminus \mathcal{T}_h)^2 \overset{(44)}{\simeq} \sum_{i=1}^m \big[\,\mathrm{osc}_i(\mathcal{T}_i)^2 - \mathrm{osc}_h(\mathcal{T}_h)^2\,\big].\end{aligned}$$

Together with (43), we see that

$$\begin{aligned}\mathbb{E}(\mathcal{T}_H) - \mathbb{E}(\mathcal{T}_h) &\overset{(37)}{=} \frac{1}{2}\,|\!|\!|\, u_H - u_h \,|\!|\!|^2 + \mathrm{osc}_H(\mathcal{T}_H)^2 - \mathrm{osc}_h(\mathcal{T}_h)^2 \\ &\overset{(43)}{\simeq} \sum_{i=1}^m \Big[\,\frac{1}{2}|\!|\!|\, u_i - u_h \,|\!|\!|^2 + \mathrm{osc}_i(\mathcal{T}_i) - \mathrm{osc}_h(\mathcal{T}_h)^2\,\Big] \overset{(37)}{=} \sum_{i=1}^m \big[\,\mathbb{E}(\mathcal{T}_i) - \mathbb{E}(\mathcal{T}_h)\,\big].\end{aligned}$$

This concludes the proof. $\square$

4.6. Verification of (A4): Local energy estimates for the estimator. Since we have already verified (37), it suffices to show the discrete local estimates (A4) for the total error.

4.6.1. *Discrete reliability.* We note that

$$\mathrm{osc}_H(\mathcal{T}_H)^2 - \mathrm{osc}_h(\mathcal{T}_h)^2 \overset{(44)}{\simeq} \mathrm{osc}_H(\mathcal{T}_H \setminus \mathcal{T}_h)^2 \overset{(24)}{\leq} \eta_H(\mathcal{E}_H \setminus \mathcal{E}_h)^2.$$

Thus, the next proposition shows the upper bound in (A4).

Proposition 16. *Let $\mathcal{T}_H \in \mathbb{T}$ and $\mathcal{T}_h \in \mathrm{refine}(\mathcal{T}_H)$. Let $p \in \mathbb{N}$. Then, it holds that*

$$|\!|\!|\, u_H - u_h \,|\!|\!|^2 \leq C_{\mathrm{drel}}\,\eta_H(\mathcal{E}_H \setminus \mathcal{E}_h)^2. \tag{45}$$

The constant $C_{\mathrm{drel}} > 0$ depends only on C_{mesh}, p, and A.

Sketch of proof. Recall the Galerkin orthogonality

$$\int_\Omega A\nabla(u_h - u_H)\cdot\nabla v_H \,\mathrm{d}x = 0 \qquad \text{for all } v_H \in \mathcal{S}^p_0(\mathcal{T}_H). \tag{46}$$

Therefore, we can insert $\mathcal{Q}_{H,h}(u_h - u_H) \in \mathcal{S}^p_0(\mathcal{T}_H)$ into the bilinear form $a(\cdot,\cdot)$. With $v_h := (1 - \mathcal{Q}_{H,h})(u_h - u_H) \in \mathcal{S}^p_0(\mathcal{T}_h)$, this yields that

$$|\!|\!|\, u_h - u_H \,|\!|\!|^2 \overset{(46)}{=} a(u_h - u_H, (1 - \mathcal{Q}_{H,h})(u_h - u_H)) = a(u_h, v_h) - \int_\Omega A\nabla u_H \cdot \nabla v_h \,\mathrm{d}x.$$

Using $\mathcal{T}_H$-elementwise integration by parts, we see that

$$|\!|\!|\, u_h - u_H \,|\!|\!|^2 = \sum_{T\in\mathcal{T}_H\setminus\mathcal{T}_h} \int_T \big(f + \mathrm{div}\,\boldsymbol{f} + \mathrm{div}(A\nabla u_H)\big)v_h \,\mathrm{d}x + \sum_{E\in\mathcal{E}^\Omega_H\setminus\mathcal{E}^\Omega_h} \int_E [\![(A\nabla u_H + \boldsymbol{f})\cdot\boldsymbol{\nu}]\!]\, v_h \,\mathrm{d}s.$$

Standard estimates conclude (45). $\square$

4.6.2. ***Discrete efficiency.*** The following proposition is proved along the lines of [FLOP10, Proposition 2] and adapts Verfürth's bubble function technique with cleverly chosen bubble functions. We note that the idea goes back to the seminal works [Dör96, MNS00]. This result extends [DKS16, Lemma 4.3] to polynomial degrees $p \geq 1$.

Proposition 17. *Let $\mathcal{T}_H \in \mathbb{T}$ and $\mathcal{T}_h \in \mathrm{refine}(\mathcal{T}_H)$. Let $p \in \mathbb{N}$. For $T \in \mathcal{T}_H \setminus \mathcal{T}_h$ and $E \in \mathcal{E}_H \setminus \mathcal{E}_h$ there hold the estimates*

$$|E| \, \| [\![(A\nabla u_H + \boldsymbol{f}) \cdot \boldsymbol{\nu}]\!] \|_{L^2(E)}^2 \lesssim \|A^{1/2}\nabla(u_H - u_h)\|_{L^2(\mathcal{T}_H^{\mathrm{red}}(E))}^2 + \mathrm{osc}_H(\mathcal{T}_H^{\mathrm{red}}(E))^2 + \sum_{T \in \mathcal{T}_H^{\mathrm{red}}(E)} |T| \, \|f + \mathrm{div}(A\nabla u_H + \boldsymbol{f})\|_{L^2(T)}^2, \tag{47}$$

$$|T| \, \|f + \mathrm{div}(A\nabla u_H + \boldsymbol{f})\|_{L^2(T)}^2 \lesssim \|A^{1/2}\nabla(u_H - u_h)\|_{L^2(T)}^2 + \mathrm{osc}_H(T)^2. \tag{48}$$

Together, there exists a constant $C_{\mathrm{deff}} > 0$ such that there holds discrete local efficiency

$$\eta_H^2(\mathcal{E}_H \setminus \mathcal{E}_h) \leq C_{\mathrm{deff}} \big[\, \| u_H - u_h \|^2 + \mathrm{osc}_H(\mathcal{T}_H \setminus \mathcal{T}_h)^2 \, \big]. \tag{49}$$

The constant C_{deff} depends only on C_{mesh}, A, and p.

Sketch of proof. We use ideas from [EGP19, Lemma 11] and employ Verfürth's bubble function technique with discrete, conforming bubbles. To this end, let $\mathcal{T}_H \in \mathbb{T}$ and $\mathcal{T}_h \in \mathrm{refine}(\mathcal{T}_H)$. For $z \in \mathcal{V}_H$, let $\phi_{H,z} \in \mathcal{S}^1(\mathcal{T}_H)$ be the piece-wise affine hat function. An element $T \in \mathcal{T}_H \setminus \mathcal{T}_h$ has at least one edge $E \subset \partial T$ with $E \in \mathcal{E}_H \setminus \mathcal{E}_h$. We denote the midpoint of this edge by $z' \in \mathcal{V}_h$ and the vertex opposite to E by $z \in \mathcal{V}_H$. We then define the corresponding edge and element bubble functions as

$$b_E := \phi_{h,z'} \in \mathcal{S}^1(\mathcal{T}_h) \quad \text{and} \quad b_T := \phi_{H,z}\phi_{h,z'} \in \mathcal{S}_0^2(\mathcal{T}_h), \tag{50}$$

respectively. The estimates (47) for $p \geq 1$ and (48) for $p \geq 2$ follow directly from the bubble function technique with b_E and b_T, respectively. For (48) with $p = 1$, it holds that

$$|T| \, \|f + \mathrm{div}(A\nabla u_H + \boldsymbol{f})\|_{L^2(T)}^2 = |T| \, \|f + \mathrm{div}\, \boldsymbol{f}\|_{L^2(T)}^2 \leq \mathrm{osc}_H(T)^2.$$

Combining (47)–(48), we conclude (49). □

4.7. Proof of Theorem 4. In the last sections, we have verified (A1)–(A5) for the primal problem (6). From Theorem 9, we thus infer instance optimality (Theorem 4) of Algorithm 3. Clearly, the same results hold for the dual problem (7), which differs from (6) only through the right-hand side G instead of F.

5. Proof of Theorem 7

The key observation for the proof of Theorem 7 is that the proposed GOAFEM (Algorithm 6) is simultaneously instance optimal for both, the primal and the dual error estimate. Since the properties (A2)–(A5) are already verified for primal and dual problem (see Section 4), it only remains to show that the marking strategy of GOAFEM (Algorithm 2) satisfies (A1) for both, the primal and the dual error estimator.

Lemma 18. *For $\mathcal{T}_H \in \mathbb{T}$, let $\mathcal{M}_H \subseteq \mathcal{E}_H$ be set of marked edges generated by Algorithm 2. Then, it holds that $\mathcal{M}_H \neq 0$ as well as*

$$\eta_H(\mathrm{tail}_H(\mathcal{M}_H)) \geq C\,\#\mathcal{M}_H \max_{E\in\mathcal{E}_H} \eta_H(\mathrm{tail}_H(E)), \tag{51}$$

$$\eta_H^*(\mathrm{tail}_H(\mathcal{M}_H)) \geq C\,\#\mathcal{M}_H \max_{E\in\mathcal{E}_H} \eta_H^*(\mathrm{tail}_H(E)), \tag{52}$$

where $C > 0$ depends only on $C_{\min} > 0$.

Proof. According to [DKS16, Proposition 5.1], Algorithm 1 guarantees that

$$\overline{\mathcal{M}}_H \neq \emptyset \qquad \text{with} \qquad \eta_H(\mathrm{tail}_H(\overline{\mathcal{M}}_H)) \geq C_{\mathrm{mark}}\,\#\overline{\mathcal{M}}_H \max_{E\in\mathcal{E}_H} \eta_H(\mathrm{tail}_H(E)),$$

$$\overline{\mathcal{M}}_H^* \neq \emptyset \qquad \text{with} \qquad \eta_H^*(\mathrm{tail}_H(\overline{\mathcal{M}}_H^*)) \geq C_{\mathrm{mark}}\,\#\overline{\mathcal{M}}_H^* \max_{E\in\mathcal{E}_H} \eta_H^*(\mathrm{tail}_H(E)).$$

Without loss of generality, we may assume that $\mathcal{M}_{\min} = \overline{\mathcal{M}}_H$ and $\mathcal{M}_{\max} = \overline{\mathcal{M}}_H^*$. Recall that $\#\mathcal{M}_{\min} \leq \#\mathcal{M}_{\max}$ and $\mathcal{M}_H := \mathcal{M}_{\min} \cup \mathcal{M}'_{\max} \neq \emptyset$, where

$$\mathcal{M}'_{\max} \subseteq \mathcal{M}_{\max} \quad \text{with} \quad \#\mathcal{M}'_{\max} = n := \min\big\{\#\mathcal{M}_{\max}, \max\{1, \lfloor C_{\min}\#\mathcal{M}_{\min}\rfloor\}\big\}.$$

First, note that $\mathcal{M}_{\max} \supseteq \mathcal{M}'_{\max} \neq \emptyset$. With Proposition 12, it follows that

$$\eta_H(\mathrm{tail}_H(\mathcal{M}_H)) \geq \eta_H(\mathrm{tail}_H(\mathcal{M}_{\min})) \geq C_{\mathrm{mark}}\,\#\mathcal{M}_{\min} \max_{E\in\mathcal{E}_H} \eta_H(\mathrm{tail}_H(E)),$$

$$\eta_H^*(\mathrm{tail}_H(\mathcal{M}_H)) \geq \eta_H^*(\mathrm{tail}_H(\mathcal{M}'_{\max})) \geq C_{\mathrm{mark}}\,\#\mathcal{M}'_{\max} \max_{E\in\mathcal{E}_H} \eta_H^*(\mathrm{tail}_H(E)),$$

In view of (51)–(52), it only remains to prove that $\#\mathcal{M}_H \lesssim \#\mathcal{M}_{\min} \lesssim \#\mathcal{M}'_{\max}$, where the hidden constants depend only on $C_{\min}$. First, note that

$$\begin{aligned}\#\mathcal{M}_H &\leq \#\mathcal{M}_{\min} + \#\mathcal{M}'_{\max} \\ &\leq \#\mathcal{M}_{\min} + \max\{1, \lfloor C_{\min}\#\mathcal{M}_{\min}\rfloor\} \leq \max\{2, (1 + C_{\min})\}\,\#\mathcal{M}_{\min}.\end{aligned}$$

This already guarantees that

$$\frac{1}{\max\{2, (1 + C_{\min})\}}\,\#\mathcal{M}_H \leq \#\mathcal{M}_{\min} \leq \#\mathcal{M}_H.$$

To estimate $\#\mathcal{M}'_{\max}$, we consider two cases:

Case 1. If $n = \#\mathcal{M}_{\max} \leq \max\{1, \lfloor C_{\min}\#\mathcal{M}_{\min}\rfloor\}$, then $\mathcal{M}_{\max} = \mathcal{M}'_{\max}$. Therefore,

$$\frac{1}{\max\{2, (1 + C_{\min})\}}\,\#\mathcal{M}_H \leq \#\mathcal{M}_{\min} \leq \#\mathcal{M}_{\max} \leq \#\mathcal{M}_H.$$

Case 2. If $n = \max\{1, \lfloor C_{\min}\#\mathcal{M}_{\min}\rfloor\} < \#\mathcal{M}_{\max}$, then $\#\mathcal{M}'_{\max} \geq 1$ leads to

$$\#\mathcal{M}'_{\max} \geq \lfloor C_{\min}\#\mathcal{M}_{\min}\rfloor \geq C_{\min}\#\mathcal{M}_{\min} - 1 \geq C_{\min}\#\mathcal{M}_{\min} - \#\mathcal{M}'_{\max}.$$

Hence, we see that

$$\frac{1}{\max\{2, (1 + C_{\min})\}}\,\#\mathcal{M}_H \leq \#\mathcal{M}_{\min} \leq \frac{2}{C_{\min}}\,\#\mathcal{M}'_{\max} \leq \frac{2}{C_{\min}}\,\mathcal{M}_H.$$

In any case, we see that $\mathcal{M}_H \simeq \#\mathcal{M}_{\min} \simeq \mathcal{M}'_{\max}$. This concludes the proof. □

Proof of Theorem 7. Section 4 shows that (A2)–(A5) are satisfied for the primal and the dual problem. Lemma 18 shows that $\mathcal{M}_H$ from Algorithm 2 satisfies (A1) simultaneously for both estimators. Theorem 9 thus implies instance optimality of Algorithm 6 for the primal and dual energy. In explicit terms, there exists $C, C^* > 0$ such that for all $\ell \in \mathbb{N}_0$

$$\forall\, \mathcal{T}_H \in \mathrm{refine}(\mathcal{T}_0): \quad \big[\, C\, \#(\mathcal{T}_H \backslash \mathcal{T}_0) \le \#(\mathcal{T}_\ell \backslash \mathcal{T}_0) \implies \mathbb{E}(\mathcal{T}_\ell) \le \mathbb{E}(\mathcal{T}_H)\,\big],$$
$$\forall\, \mathcal{T}_{H^*} \in \mathrm{refine}(\mathcal{T}_0): \quad \big[\, C^*\, \#(\mathcal{T}_{H^*} \backslash \mathcal{T}_0) \le \#(\mathcal{T}_\ell \backslash \mathcal{T}_0) \implies \mathbb{E}^*(\mathcal{T}_\ell) \le \mathbb{E}^*(\mathcal{T}_{H^*})\,\big],$$

where $\mathbb{E}(\cdot)$ denotes the energy (34) for the primal problem, and $\mathbb{E}^*(\cdot)$ denotes the energy for the dual problem. Obviously, this leads to

$$\begin{aligned}\forall\, \mathcal{T}_H, \mathcal{T}_{H^*} \in \mathrm{refine}(\mathcal{T}_0): \quad \big[\, &\max\{C, C^*\}\, \max\{\#(\mathcal{T}_H \backslash \mathcal{T}_0)\,,\, \#(\mathcal{T}_{H^*} \backslash \mathcal{T}_0)\} \le \#(\mathcal{T}_\ell \backslash \mathcal{T}_0) \\ &\implies \mathbb{E}(\mathcal{T}_\ell)\, \mathbb{E}^*(\mathcal{T}_\ell) \le \mathbb{E}(\mathcal{T}_H)\, \mathbb{E}^*(\mathcal{T}_{H^*})\,\big],\end{aligned}$$

Using the equivalence (A5) of energy and total error (for primal and dual problem), we conclude the proof. □

6. Numerical experiments

We conclude this work with some numerical experiments performed in MATLAB, where our implementation builds on the codes provided in [FPW11] for $p = 1$ and [FFP19] for $p = 2$. For the modified maximum criterion (Algorithm 1), we have implemented a recursive variant proposed in [DKS16].

6.1. Adaptive FEM with Z-shaped domain.

We consider the problem

$$-\Delta u = 1 \quad \text{in } \Omega := (-1,1)^2 \setminus \mathrm{conv}\{(0,0), (-1,0), (-1,-1)\}, \tag{53a}$$
$$u = 0 \quad \text{on } \Gamma := \partial\Omega, \tag{53b}$$

where $\mathrm{conv}(\cdot)$ denotes the convex hull and Ω is the Z-shaped domain from Figure 2. This problem is solved with the instance-optimal algorithm from [DKS16], i.e., Algorithm 3. Moreover, we compare the results with a rate-optimal algorithm, which builds on an edge-based Dörfler marking criterion [Dör96]: Find a subset $\mathcal{M}_H \subseteq \mathcal{E}_H$ with minimal cardinality such that

$$\theta\, \mu_H(\mathcal{E}_H)^2 \le \mu_H(\mathcal{M}_H)^2 \tag{54}$$

for an edge-based error estimator $\mu_H : \mathcal{E}_H \to \mathbb{R}$. Note that uniform refinement corresponds to $\theta = 1$ in (54) but $\vartheta = 0$ in Algorithm 1. Therefore, we set $\theta = 1 - \vartheta$ in the following to account for the different interpretations of the marking parameters. We note that both adaptive strategies only differ by the marking strategy. Throughout, we consider $\vartheta = 0.5$ and $C_{\min} = 1$.

In Figure 1, we visualize the edge-based residual error estimator η_ℓ and the energy error $|\!|\!|\, u - u_\ell \,|\!|\!|$. Since the exact solution u is unknown, we extrapolate $|\!|\!|\, u \,|\!|\!|$ from the computed values $|\!|\!|\, u_\ell \,|\!|\!|$ and use the Galerkin orthogonality (46) to obtain

$$|\!|\!|\, u - u_\ell \,|\!|\!|^2 = |\!|\!|\, u \,|\!|\!|^2 - |\!|\!|\, u_\ell \,|\!|\!|^2.$$

Since [DKS16] does not provide any numerical experiments, we also give qualitative plots of the resulting meshes in Figure 2. Both strategies mark edges near the re-entrant corner.

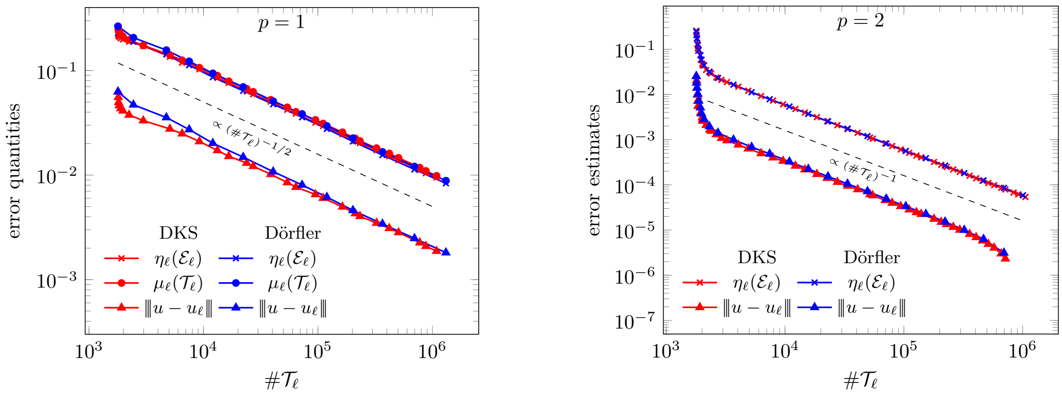
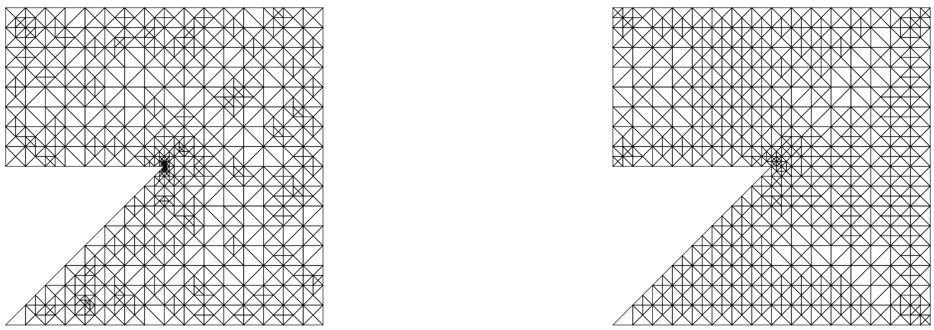


FIGURE 1. *Error estimators and energy error for problem* (53) *with FEM of order $p = 1$ (left) and $p = 2$ (right). We compare the modified maximum criterion (DKS) with the Dörfler marking* (54) *for $\vartheta = 0.5$.*



(A) *Modified maximum marking, $\#\mathcal{T}_{10} = 1204$.* (B) *Dörfler marking* (54)*, $\#\mathcal{T}_7 = 1196$.*

FIGURE 2. *Comparison of meshes generated by Algorithm 3 with $p = 1$, $\vartheta = 0.5$, and different marking strategies for problem* (53).

While for the Dörfler marking criterion (54) edges in the interior are refined mostly by mesh closure, the modified maximum criterion (Algorithm 1) also marks edges in the interior that have long tails.

6.2. Goal-oriented AFEM. The following numerical example empirically shows how the proposed goal-oriented adaptivity (Algorithm 6) handles possible singularities. We consider a problem proposed in [MS09], where the primal problem reads

$$-\Delta u = \operatorname{div} \boldsymbol{f} \qquad \text{in } \Omega := (0,1)^2, \tag{55a}$$

$$u = 0 \qquad \text{on } \Gamma := \partial(0,1)^2, \tag{55b}$$

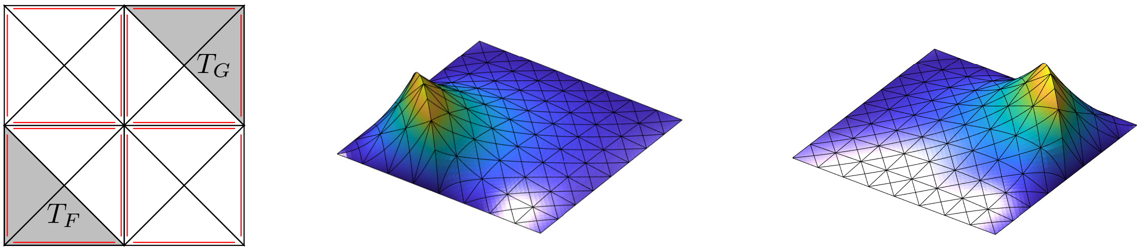


FIGURE 3. *Initial mesh (left), and qualitative plot of primal solution (middle) and negative dual solution (right) of* (55). *In the initial mesh, the sets T_F and T_G are highlighted in grey and the reference edge of each triangle is highlighted by a red line.*

with

$$\boldsymbol{f}(x) = \begin{cases} (1,0)^\top & \text{if } x \in T_F := \big\{x \in \Omega \,:\, x_1 + x_2 \le 1/2\big\}, \\ (0,0)^\top & \text{else.} \end{cases}$$

With $T_G := \big\{x \in \Omega \,:\, x_1 + x_2 \ge 3/2\big\}$ and

$$g = 0 \qquad \text{and} \qquad \boldsymbol{g}(x) = \begin{cases} (1,0)^\top & \text{if } x \in T_G, \\ (0,0)^\top & \text{else,} \end{cases}$$

the goal functional from (5) takes the form

$$G(v) = \int_{T_G} \frac{\partial v}{\partial x_1} \,\mathrm{d}x \qquad \text{for } v \in H^1_0(\Omega).$$

The initial triangulation $\mathcal{T}_0$ with the subsets T_F and T_G, together with approximations to the primal and dual solution can be seen in Figure 3.

In addition to Algorithm 6, we investigate the rate-optimal algorithms from [MS09, FPZ16, BET11]. These build on the Dörfler marking criterion (54). For the convenience of the reader, we briefly outline these marking strategies:

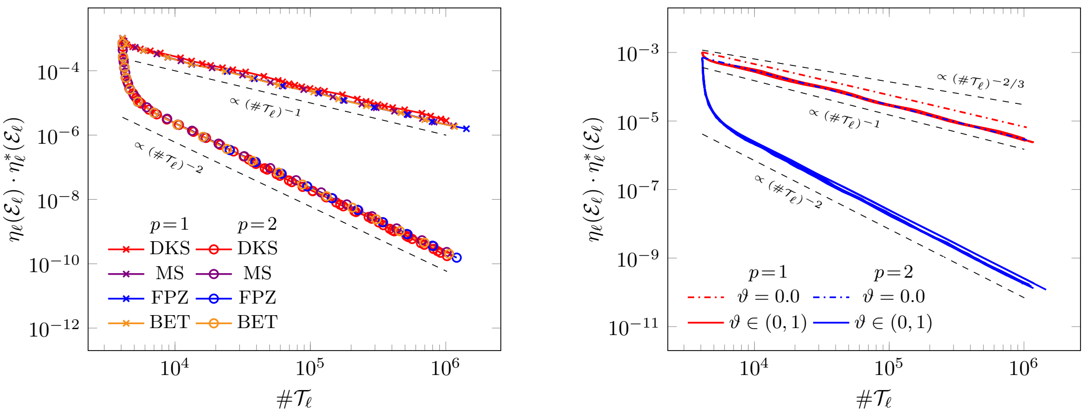
- In [MS09], the Dörfler criterion (54) is employed separately for $\mu_H = \eta_H$ as well as $\mu_H = \eta_H^*$ and thus provides sets $\overline{\mathcal{M}}_H, \overline{\mathcal{M}}_H^* \subseteq \mathcal{T}_H$. Then $\mathcal{M}_H$ is chosen as the smaller set $\mathcal{M}_H := \arg\min\{\overline{\mathcal{M}}_H, \overline{\mathcal{M}}_H^*\}$.
- In [FPZ16], one proceeds analogously, but chooses $\mathcal{M}_H := \underline{\mathcal{M}}_H \cup \underline{\mathcal{M}}_H^*$, where $\underline{\mathcal{M}}_H \subseteq \overline{\mathcal{M}}_H$ and $\underline{\mathcal{M}}_H^* \subseteq \overline{\mathcal{M}}_H^*$ satisfy $\#\underline{\mathcal{M}}_H = \#\underline{\mathcal{M}}_H^* = \min\{\#\overline{\mathcal{M}}_H, \#\overline{\mathcal{M}}_H^*\}$.
- For [BET11], the Dörfler criterion (54) is employed for

$$\mu_H(E)^2 := \eta_H(E)^2 \eta_H^*(\mathcal{E}_H)^2 + \eta_H(\mathcal{E}_H)^2 \eta_H^*(E)^2.$$

In Figure 4, we compare the products of error estimates for the primal and dual problem for the above marking strategies. For $p = 1, 2$, this product decays with optimal rate $(\#\mathcal{T}_\ell)^{-p}$. Furthermore, Figure 5 gives qualitative plots of the local mesh-size.

REFERENCES

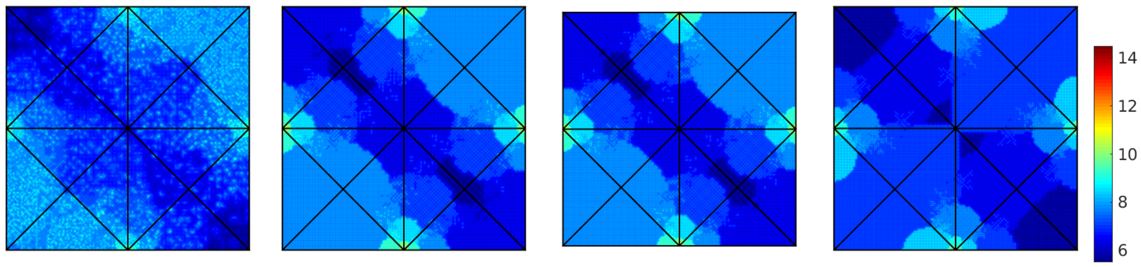
[BDD04] Peter Binev, Wolfgang Dahmen, and Ronald DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97:219–268, 2004.



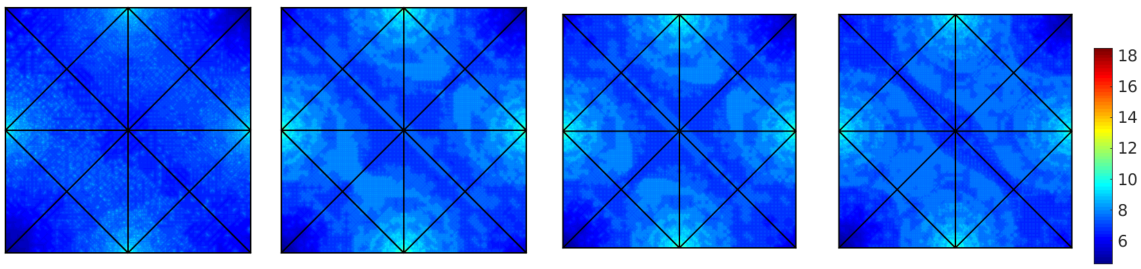
(A) *Different marking strategies in Algorithm 6, where we compare Algorithm 2 (DKS) and* [MS09, FPZ16, BET11] *with marking parameter* $\vartheta = \theta = 0.5$.

(B) *Comparison of estimates for different parameters of* ϑ. *For the adaptive case* $\vartheta \in (0,1)$ *we chose* $i/10$ *with* $i = 1,\dots,9$ *as representative parameters.*

FIGURE 4. *Product of error estimates for the solution of problem* (55) *with Algorithm 6 and FEM of order* $p = 1$ *and* $p = 2$.



(A) $\#\mathcal{T}_{22} = 53997$. (B) $\#\mathcal{T}_{18} = 52596$. (C) $\#\mathcal{T}_{9} = 52078$. (D) $\#\mathcal{T}_{10} = 80142$.



(E) $\#\mathcal{T}_{33} = 51619$. (F) $\#\mathcal{T}_{24} = 55118$. (G) $\#\mathcal{T}_{12} = 54164$. (H) $\#\mathcal{T}_{13} = 71080$.

FIGURE 5. *Comparison of meshes resulting from the recursive marking strategy from* [DKS16], [MS09], [FPZ16], *and* [BET11] *(left to right) with polynomial degree* $p = 1$ *(top) and* $p = 2$ *(bottom), and marking parameter* $\vartheta = 0.5$. *The colors show the value of* $\log_2(1/|T|)$ *for every element* T, *which corresponds to the element's level.*

[BDDP02] Peter Binev, Wolfgang Dahmen, Ronald DeVore, and Pencho Petrushev. Approximation classes for adaptive methods. *Serdica Math. J.*, 28(4):391–416, 2002. Dedicated to the memory of Vassil Popov on the occasion of his 60th birthday.

[BET11] Roland Becker, Elodie Estecahandy, and David Trujillo. Weighted marking for goal-oriented adaptive finite element methods. *SIAM J. Numer. Anal.*, 49(6):2451–2469, 2011.

[BHP17] Alex Bespalov, Alexander Haberl, and Dirk Praetorius. Adaptive FEM with coarse initial mesh guarantees optimal convergence rates for compactly perturbed elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 317:318–340, 2017.

[CFPP14] Carsten Carstensen, Michael Feischl, Markus Page, and Dirk Praetorius. Axioms of adaptivity. *Comput. Math. Appl.*, 67(6):1195–1253, 2014.

[CKNS08] J. Manuel Cascon, Christian Kreuzer, Ricardo H. Nochetto, and Kunibert G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.

[DKS16] Lars Diening, Christian Kreuzer, and Rob Stevenson. Instance optimality of the adaptive maximum strategy. *Found. Comput. Math.*, 16(1):33–68, 2016.

[Dör96] Willy Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.

[EGP19] Christoph Erath, Gregor Gantner, and Dirk Praetorius. Optimal convergence behavior of adaptive FEM driven by simple $(h - h/2)$-type error estimators. *Comput. Math. Appl., in print*, 2019.

[FFP14] Michael Feischl, Thomas Führer, and Dirk Praetorius. Adaptive FEM with optimal convergence rates for a certain class of nonsymmetric and possibly nonlinear problems. *SIAM J. Numer. Anal.*, 52(2):601–625, 2014.

[FFP19] Thomas Führer, Stefan A. Funken, and Dirk Praetorius. Adaptive isoparametric P2-FEM: Analysis and efficient Matlab implementation. In preparation, 2019.

[FGH$^+$16] Michael Feischl, Gregor Gantner, Alexander Haberl, Dirk Praetorius, and Thomas Führer. Adaptive boundary element methods for optimal convergence of point errors. *Numer. Math.*, 132(3):541–567, 2016.

[FLOP10] Samuel Ferraz-Leite, Christoph Ortner, and Dirk Praetorius. Convergence of simple adaptive galerkin schemes based on $h - h/2$ error estimators. *Numer. Math.*, 116:291–316, 2010.

[FPW11] Stefan Funken, Dirk Praetorius, and Philip Wissgott. Efficient implementation of adaptive P1-FEM in Matlab. *Comput. Methods Appl. Math.*, 11(4):460–490, 2011.

[FPZ16] Michael Feischl, Dirk Praetorius, and Kristoffer G. van der Zee. An abstract analysis of optimal goal-oriented adaptivity. *SIAM J. Numer. Anal.*, 54(3):1423–1448, 2016.

[Gan17] Tsogtgerel Gantumur. Convergence rates of adaptive methods, Besov spaces, and multilevel approximation. *Found. Comput. Math.*, 17(4):917–956, 2017.

[GM14] Fernando D. Gaspoz and Pedro Morin. Approximation classes for adaptive higher order finite element approximation. *Math. Comp.*, 83(289):2127–2160, 2014.

[HP16] Michael Holst and Sara Pollock. Convergence of goal-oriented adaptive finite element methods for nonsymmetric problems. *Numer. Methods Partial Differential Equations*, 32(2):479–509, 2016.

[HPZ15] Michael Holst, Sara Pollock, and Yunrong Zhu. Convergence of goal-oriented adaptive finite element methods for semilinear problems. *Comput. Vis. Sci.*, 17(1):43–63, 2015.

[KPP13] Michael Karkulik, David Pavlicek, and Dirk Praetorius. On 2D newest vertex bisection: optimality of mesh-closure and H^1-stability of L_2-projection. *Constr. Approx.*, 38(2):213–234, 2013.

[KS16] Christian Kreuzer and Mira Schedensack. Instance optimal Crouzeix-Raviart adaptive finite element methods for the Poisson and Stokes problems. *IMA J. Numer. Anal.*, 36(2):593–617, 2016.

[MNS00] Pedro Morin, Ricardo H. Nochetto, and Kunibert G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488, 2000.

[MS09] Mario S. Mommer and Rob Stevenson. A goal-oriented adaptive finite element method with convergence rates. *SIAM J. Numer. Anal.*, 47(2):861–886, 2009.

[MSV08] Pedro Morin, Kunibert G. Siebert, and Andreas Veeser. A basic convergence result for conforming adaptive finite elements. *Math. Models Methods Appl. Sci.*, 18(5):707–737, 2008.

[Sie11] Kunibert G. Siebert. A convergence proof for adaptive finite elements without lower bound. *IMA J. Numer. Anal.*, 31(3):947–970, 2011.

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

[Ste08] Rob Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77(261):227–241, 2008.

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

TU Wien, Institute for Analysis and Scientific Computing, Wiedner Hauptstr. 8–10/E101/4, 1040 Wien, Austria
E-mail address: `Michael.Innerberger@asc.tuwien.ac.at`

TU Wien, Institute for Analysis and Scientific Computing, Wiedner Hauptstr. 8–10/E101/4, 1040 Wien, Austria
E-mail address: `Dirk.Preatorius@asc.tuwien.ac.at`