SIAM J. NUMER. ANAL.
Vol. 52, No. 2, pp. 601–625

ADAPTIVE FEM WITH OPTIMAL CONVERGENCE RATES FOR A CERTAIN CLASS OF NONSYMMETRIC AND POSSIBLY NONLINEAR PROBLEMS*

M. FEISCHL†, T. FÜHRER†, AND D. PRAETORIUS†

Abstract. We analyze adaptive mesh-refining algorithms for conforming finite element discretizations of certain nonlinear second-order partial differential equations. We allow continuous polynomials of arbitrary but fixed polynomial order. The adaptivity is driven by the residual error estimator. We prove convergence even with optimal algebraic convergence rates. In particular, our analysis covers general linear second-order elliptic operators. Unlike prior works for linear nonsymmetric operators, our analysis avoids the interior node property for the refinement, and the differential operator has to satisfy a Gårding inequality only. If the differential operator is uniformly elliptic, no additional assumption on the initial mesh is posed.

Key words. adaptive algorithm, convergence, optimal cardinality, nonlinear, nonsymmetric

AMS subject classifications. 65N30, 65N50

DOI. 10.1137/120897225

1. Introduction.

Let Ω be a bounded polyhedral Lipschitz domain in $\mathbb{R}^d$, $d \geq 2$. We consider a homogeneous Dirichlet boundary value problem for a certain nonlinear second-order elliptic partial differential equation

$$\mathcal{L}u(x) := -\mathrm{div}\big(\boldsymbol{A}(x, \nabla u)\big) + g(x, u, \nabla u) = f(x) \quad \text{in } \Omega, \tag{1.1a}$$

$$u = 0 \quad \text{on } \Gamma := \partial\Omega. \tag{1.1b}$$

The differential operator $\mathcal{L} = \mathcal{A}+\mathcal{K}$ is split into a principal part $\mathcal{A}u = -\mathrm{div}\big(\boldsymbol{A}(\cdot, \nabla u)\big)$ and a compact perturbation $\mathcal{K}u = g(\cdot, u, \nabla u)$. The analysis restricts to strongly monotone operators $\mathcal{L}$ which consist of twice Frechet-differentiable parts $\mathcal{A}$, $\mathcal{K}$ with symmetric first derivative $D\mathcal{A}$ of the diffusion term. (See section 6.5 for the precise regularity assumptions.) This particularly includes the case of general linear second-order elliptic operators

$$\mathcal{L}u := -\mathrm{div}(\boldsymbol{A}\nabla u) + \boldsymbol{b} \cdot \nabla u + cu. \tag{1.2}$$

We consider a common adaptive mesh-refining algorithm which iterates the following loop:

$$\boxed{\texttt{solve}} \longrightarrow \boxed{\texttt{estimate}} \longrightarrow \boxed{\texttt{mark}} \longrightarrow \boxed{\texttt{refine}} \tag{1.3}$$

The module `solve` computes a piecewise polynomial finite element approximation U_ℓ of u with respect to a given mesh $\mathcal{T}_\ell$. For `estimate`, we use a residual error estimator;

*Received by the editors November 1, 2012; accepted for publication (in revised form) January 8, 2014; published electronically March 11, 2014. This work was supported through the FWF project Adaptive Boundary Element Method, funded by the Austrian Science Fund (FWF) under grant P21732; see http://www.asc.tuwien.ac.at/abem.

http://www.siam.org/journals/sinum/52-2/89722.html

†Institute for Analysis and Scientific Computing, Vienna University of Technology, A-1040 Wien, Austria (michael.feischl@tuwien.ac.at, thomas.fuehrer@tuwien.ac.at, dirk.praetorius@tuwien.ac.at). The first and third authors were supported by the FWF doctoral program "Dissipation and Dispersion in Nonlinear PDEs" under grant W1245; see http://npde.tuwien.ac.at/.

see, e.g., [1, 31]. Next, the Dörfler marking criterion [14] is used to single out elements for refinement. Finally, `refine` leads to a locally refined and improved mesh $\mathcal{T}_{\ell+1}$ by means of the newest vertex bisection algorithm (NVB).

So far, available results on convergence and quasi-optimality of adaptive finite element methods (AFEM) from the literature essentially have dealt with the linear, symmetric, and elliptic case (1.2) with $\boldsymbol{b} = 0$ and $c \geq 0$; see, e.g., [7, 9, 11, 14, 20, 29] and the references therein. As far as the linear and nonsymmetric case $\boldsymbol{b} \neq 0$ is concerned, we are only aware of the works [12, 21] which, however, considered the special situation $\operatorname{div} \boldsymbol{b} = 0$ and $c \geq 0$. Moreover, their analysis requires the interior node property for the refinement at least after a fixed number of steps, which has been introduced in [23] to guarantee a discrete lower bound for the error. Finally, the proofs of convergence and quasi-optimality in [12, 21] assume the initial mesh $\mathcal{T}_0$ to be sufficiently fine, although the assumption $\operatorname{div} \boldsymbol{b} = 0$ already ensures ellipticity of the associated bilinear form $b(\cdot,\cdot)$ in the weak formulation of (1.1), i.e., the operator $\mathcal{L}$ in (1.2) is uniformly elliptic. All this is different from the present work, and the advances over the state of the art (see, e.g., [11, 12, 20]), are fourfold:

- (i) In the linear case (1.2), our assumptions on the data $\boldsymbol{A} = \boldsymbol{A}(x)$, $\boldsymbol{b} = \boldsymbol{b}(x)$, and $c = c(x)$ only ensure that the bilinear form $b(\cdot,\cdot)$ of the weak formulation of (1.1) is continuous and satisfies a Gårding inequality on $H^1_0(\Omega)$.
- (ii) As for the symmetric case [11], we only rely on the standard NVB, and the interior node property is avoided.
- (iii) If $b(\cdot,\cdot)$ is elliptic, we avoid any assumption on the initial mesh $\mathcal{T}_0$. If $b(\cdot,\cdot)$ satisfies a Gårding inequality, we require the same assumption on the initial mesh as [12, 21] to ensure well-posedness of the finite element formulations.
- (iv) This work complements [16] beyond energy minimization and scalar coefficients. To the best of the authors' knowledge and besides [6] for the particular p-Laplace problem, this work and [16] provide the first quasi-optimality results for a class of nonlinear problems. Contraction of AFEM within an abstract framework for nonlinear problems but without convergence rates is shown in [18].

From a technical point of view, our analytical argument works as follows and is illustrated for the linear operator $\mathcal{L}$ from (1.2) with induced bilinear form $b(\cdot,\cdot)$. First, the estimator reduction

$$\eta_{\ell+1}^2 \leq q\,\eta_\ell^2 + C\,|||U_{\ell+1} - U_\ell|||^2 \tag{1.4}$$

together with a Céa-type quasi-optimality already implies convergence $U_\ell \to u$ as $\ell \to \infty$ (Proposition 3.3); see also [3] for this *estimator reduction principle*. Here, $0 < q < 1$ and $C > 0$ are generic constants, and $|||\cdot|||$ denotes the energy quasi norm induced by $b(\cdot,\cdot)$. Second, the novel contribution in our analysis is that this additional knowledge allows us to prove a quasi-Pythagoras theorem

$$|||U_{\ell+1} - U_\ell|||^2 + |||u - U_{\ell+1}|||^2 \leq \frac{1}{1-\varepsilon}\,|||u - U_\ell|||^2 \tag{1.5}$$

for all $\varepsilon > 0$ and $\ell \geq \ell_0(\varepsilon)$ sufficiently large (Proposition 3.6) which unlike [12, 21] avoids any additional assumption on the mesh-size of $\mathcal{T}_\ell$. With estimator reduction (1.4) and quasi-orthogonality (1.5) at hand, we next observe R-linear convergence

$$\eta_{\ell+k} \leq C q^k \eta_\ell \quad \text{for all } \ell, k \in \mathbb{N} \tag{1.6}$$

of the error estimator (Theorem 4.1) with further generic constants $C > 0$ and $0 < q < 1$. Finally, the R-linear convergence (1.6) suffices to follow the paths of [29, 11] to prove even quasi-optimal convergence rates in the sense of

$$(u, f) \in \mathbb{A}_s \quad \Longleftrightarrow \quad \eta_\ell \le C\,(\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-s} \quad \text{for all } \ell \in \mathbb{N}, \tag{1.7}$$

i.e., each theoretically possible convergence order $\mathcal{O}(N^{-s})$ for the error estimator will asymptotically be achieved by AFEM. The approximation class $\mathbb{A}_s$ involved in (1.7) is defined in section 5. By means of reliability and efficiency of the error estimator η_ℓ used, this quasi-optimality result can equivalently be stated in terms of error plus oscillations, as is done in [11, 12, 20, 29]. As was first observed in [2], this approach uses the lower bounds for the error only in the characterization of the approximation class in Lemma 5.1. On the one hand, this allows us to show that the upper bound $0 < \theta < \theta_\star$ of the range of optimal marking parameters does not depend on any efficiency estimate. Moreover, optimal convergence of the estimator (but not of the total error; see Remark 5.2 for further discussion) can be shown even without the efficiency estimate.

For the nonlinear problem (1.1), we observe that estimator reduction (1.4), R-linear convergence (1.6), and quasi-optimality (1.7) do not hinge on linearity of $\mathcal{L}$. We thus bootstrap the arguments developed for the linear case to prove a quasi-Pythagoras theorem (1.5) for nonlinear $\mathcal{L}$ (Proposition 6.11) and may derive convergence of AFEM with quasi-optimal algebraic rates.

The remainder of this paper is organized as follows. For the sake of a clear presentation, we first consider the linear case (1.2) with elliptic bilinear form $b(\cdot,\cdot)$ corresponding to the weak formulation of (1.1). This case already includes the main ideas of how to cope with compact perturbations. In section 2, we explicitly state the assumptions on the differential operator $\mathcal{L}$ from (1.2), recall the continuous and discrete variational formulation of (1.1), and give the necessary details on the four modules of (1.3). Section 3 then provides the estimator reduction (1.4), which follows as in [11], and the quasi-Galerkin orthogonality (1.5), which relies on the convergence of AFEM and compactness arguments. The short section 4 proves R-linear convergence (1.6) of the error estimator by use of (1.4)–(1.5). We stress that, so far, the analysis hinges on neither the precise mesh-refinement used nor on the adaptivity parameter chosen. By use of intrinsic properties of NVB, we then prove quasi-optimal convergence rates (1.7) in section 5. The final section 6 is concerned with extensions of our analysis. Among other topics, we discuss boundary conditions other than (1.1b) as well as changes of our analysis if the bilinear form $b(\cdot,\cdot)$ satisfies only a Gårding inequality. Subsection 6.5 bootstraps the arguments of the previous sections and incorporates the nonlinear case (1.1a) into the analysis.

In all statements, the constants involved and their dependencies are explicitly stated. In proofs, however, we use the symbol $\lesssim$ to abbreviate $\le$ up to a multiplicative constant. Moreover, $\simeq$ abbreviates that both estimates $\lesssim$ and $\gtrsim$ hold.

2. Model problem and adaptive algorithm.

This section is devoted to stating the model problem (1.1) with linear differential operator (1.2) in weak form and collects all the ingredients needed to formulate the adaptive algorithm. The presented problem is not the most general case on which the developed theory can be applied, but it allows for a rather simple presentation and illustrates the main difficulties of the problem. We refer to section 6 for possible extensions and generalizations.

2.1. Variational formulation.

For a given right-hand-side $f \in L^2(\Omega)$, we consider the elliptic boundary value problem (1.1) with linear operator $\mathcal{L}$ from (1.2).

For the weak formulation, the error estimator, and to prove optimal convergence rates, we require some regularity assumptions on the coefficients. We assume that $\boldsymbol{A} = \boldsymbol{A}(x) \in \mathbb{R}^{d\times d}$ with $\boldsymbol{A} \in \big(W_1^\infty(\Omega)\big)^{d\times d}$ is a symmetric matrix, $\boldsymbol{b} = \boldsymbol{b}(x) \in \mathbb{R}^d$ with $\boldsymbol{b} \in \big(L^\infty(\Omega)\big)^d$ is a vector, and $c = c(x) \in \mathbb{R}$ with $c \in L^\infty(\Omega)$ is a scalar. Here, $W_1^\infty(\Omega) := \big\{a \in L^\infty(\Omega) : \nabla a \in \big(L^\infty(\Omega)\big)^d \text{ in the weak sense}\big\}$ coincides with the space of Lipschitz continuous functions. This allows us to write the weak formulation of (1.1): Find $u \in H_0^1(\Omega) := \big\{v \in H^1(\Omega) : v|_\Gamma = 0 \text{ in the sense of traces}\big\}$ such that

$$b(u,v) := \int_\Omega \boldsymbol{A}\nabla u \cdot \nabla v + \boldsymbol{b}\cdot\nabla u\, v + cuv\,dx = \int_\Omega f v\,dx \quad \text{for all } v \in H_0^1(\Omega). \tag{2.1}$$

The bilinear form $b(\cdot,\cdot)$ is well-defined and bounded with

$$|b(u,v)| \le C_{\rm cont}\|\nabla u\|_{L^2(\Omega)}\|\nabla v\|_{L^2(\Omega)} \quad \text{for all } u,v \in H_0^1(\Omega), \tag{2.2}$$

where the constant $C_{\rm cont} := C_\Omega\big(\|\boldsymbol{A}\|_{L^\infty(\Omega)} + \|\boldsymbol{b}\|_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)}\big)$ depends only on the coefficients of $\mathcal{L}$ as well as the Poincaré constant $C_\Omega > 0$ of Ω. Additionally, we assume that the coefficients ensure that $b(\cdot,\cdot)$ is elliptic, i.e.,

$$b(u,u) \ge C_{\rm ell}\|\nabla u\|_{L^2(\Omega)}^2 \quad \text{for all } u \in H_0^1(\Omega) \tag{2.3}$$

for some constant $C_{\rm ell} > 0$ which may also depend on $C_\Omega > 0$; see section 6 if $b(\cdot,\cdot)$ satisfies only a Gårding inequality.

Now, the Lax–Milgram lemma guarantees unique solvability of (2.2) for all $f \in L^2(\Omega)$ and proves continuous dependence $\|\nabla u\|_{L^2(\Omega)} \lesssim \|f\|_{H^{-1}(\Omega)} \lesssim \|f\|_{L^2(\Omega)}$. (The hidden constant in the first estimate depends on $C_{\rm cont}$ and $C_{\rm ell}$, whereas the constant in the second estimate depends only on Ω.) Here, $H^{-1}(\Omega) := H_0^1(\Omega)^\star$ denotes the dual space of $H_0^1(\Omega)$, and duality is understood with respect to the extended L^2-scalar product, i.e.,

$$\|f\|_{H^{-1}(\Omega)} := \sup_{v\in H_0^1(\Omega)\setminus\{0\}} \frac{\int_\Omega f v\,dx}{\|\nabla v\|_{L^2(\Omega)}}.$$

Moreover, the bilinear form $b(\cdot,\cdot)$ defines a *quasi*-norm $|||\cdot||| := b(\cdot,\cdot)^{1/2}$, i.e., $|||\cdot|||$ is definite and homogeneous, but satisfies the triangle inequality only up to some multiplicative constant. Due to ellipticity and continuity of $b(\cdot,\cdot)$, it holds that

$$C_{\rm norm}^{-1}\|\nabla v\|_{L^2(\Omega)} \le |||v||| \le C_{\rm norm}\|\nabla v\|_{L^2(\Omega)} \quad \text{for all } v \in H_0^1(\Omega) \tag{2.4}$$

for a constant $C_{\rm norm} = \max\{C_{\rm cont}^{1/2}, C_{\rm ell}^{-1/2}\} > 0$.

2.2. Discrete formulation. For any regular triangulation $\mathcal{T}_\ell$ of Ω (see section 2.5 below) and $p \ge 1$, we consider the piecewise polynomials

$$\mathcal{P}^p(\mathcal{T}_\ell) := \big\{V_\ell \in L^2(\Omega) : \text{for all } T \in \mathcal{T}_\ell,\ V_\ell|_T \text{ is a polynomial of degree at most } p\big\}$$

as well as the conforming ansatz and test space

$$\mathcal{S}_0^p(\mathcal{T}_\ell) := \mathcal{P}^p(\mathcal{T}_\ell) \cap H_0^1(\Omega) \subset \mathcal{C}(\overline{\Omega}).$$

Now, the discrete formulation of (2.2) reads as follows: Find $U_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell)$ such that

$$b(U_\ell, V_\ell) = \int_\Omega f\,V_\ell\,dx \quad \text{for all } V_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell). \tag{2.5}$$

As in the continuous case (2.2), existence and uniqueness of U_ℓ follow from the Lax–Milgram lemma. Moreover, there holds the Céa lemma

$$\|\nabla(u - U_\ell)\|_{L^2(\Omega)} \le \frac{C_{\rm cont}}{C_{\rm ell}} \min_{V_\ell \in \mathcal{S}^p_0(\mathcal{T}_\ell)} \|\nabla(u - V_\ell)\|_{L^2(\Omega)}. \tag{2.6}$$

2.3. Error estimator. We use the standard weighted-residual error estimator with the local contributions

$$\eta_\ell(T)^2 := |T|^{2/d}\|\mathcal{L}|_T U_\ell - f\|^2_{L^2(T)} + |T|^{1/d}\|[\boldsymbol{A}\nabla U_\ell \cdot n]\|^2_{L^2(\partial T\cap\Omega)}$$

for all $T \in \mathcal{T}_\ell$ and all $\ell \in \mathbb{N}$. Here, $|T|$ is the d-dimensional volume of $T \in \mathcal{T}_\ell$, and $[\boldsymbol{A}\nabla U_\ell \cdot n]|_E := \big(\boldsymbol{A}\nabla U_\ell|_{T_1}\big)\cdot n_{T_1} + \big(\boldsymbol{A}\nabla U_\ell|_{T_2}\big)\cdot n_{T_2}$ denotes the conormal jump over the facet $E := T_1 \cap T_2$ for all $T_1, T_2 \in \mathcal{T}_\ell$, where n_{T_1}, n_{T_2} denote the outward pointing normal units on the respective element boundaries. Note that due to the regularity assumptions on the coefficients, there holds $\mathcal{L}|_T U_\ell \in L^2(T)$ for all $T \in \mathcal{T}_\ell$. The error estimator η_ℓ is defined as the ℓ_2-sum of the elementwise contributions

$$\eta_\ell^2 := \sum_{T\in\mathcal{T}_\ell} \eta_\ell(T)^2.$$

As shown in, e.g., [1, 31], the error estimator is reliable, i.e., for all regular triangulations $\mathcal{T}_\ell$ and corresponding solutions U_ℓ of (2.5), it holds that

$$\|\nabla(u - U_\ell)\|_{L^2(\Omega)} \le C_{\rm rel}\eta_\ell \tag{2.7}$$

for a constant $C_{\rm rel} > 0$. Moreover, η_ℓ is also efficient, i.e.,

$$C_{\rm eff}^{-1}\eta_\ell \le \|\nabla(u - U_\ell)\|_{L^2(\Omega)} + {\rm osc}_\ell(U_\ell) \tag{2.8a}$$

for a constant $C_{\rm eff} > 0$ and oscillation terms

$$\begin{aligned}{\rm osc}_\ell(U_\ell)^2 := \sum_{T\in\mathcal{T}_\ell} \Big(&|T|^{2/d}\|(1-\Pi_\ell^{2p-2})(\mathcal{L}|_T U_\ell - f)\|^2_{L^2(T)} \\ &+ |T|^{1/d}\|(1-\Pi_\ell^{2p-1})[\boldsymbol{A}\nabla U_\ell\cdot n]\|^2_{L^2(\partial T\cap\Omega)}\Big),\end{aligned} \tag{2.8b}$$

where Π_ℓ^q denotes the L^2-orthogonal projection onto piecewise polynomials of degree q on Ω, respectively, the skeleton of $\mathcal{T}_\ell$. The constants $C_{\rm rel}, C_{\rm eff} > 0$ depend only on γ-shape regularity of $\mathcal{T}_\ell$ (see section 2.5 below), the polynomial degree $p \ge 1$, and Ω. We stress that unlike [11, 12, 20], efficiency (2.8) is not used throughout our analysis.

2.4. Adaptive algorithm. Now, we are in position to formulate the adaptive algorithm (1.3) in detail.

ALGORITHM 2.1. INPUT: *Initial triangulation* $\mathcal{T}_0$ *and adaptivity parameter* $0 < \theta \le 1$.

Loop: *For* $\ell = 0, 1, 2, \dots$ *do* (i)–(iv)

(i) *Compute discrete solution* U_ℓ *of* (2.5).

(ii) *Compute refinement indicators* $\eta_\ell(T)$ *for all* $T \in \mathcal{T}_\ell$.

(iii) *Determine set* $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ *of minimal cardinality such that*

$$\theta\,\eta_\ell^2 \le \sum_{T\in\mathcal{M}_\ell} \eta_\ell(T)^2. \tag{2.9}$$

(iv) *Refine (at least) the marked elements* $T \in \mathcal{M}_\ell$ *to obtain the triangulation* $\mathcal{T}_{\ell+1}$.

OUTPUT: *Approximate solutions* U_ℓ *and error estimators* η_ℓ *for all* $\ell \in \mathbb{N}$.

2.5. Mesh-refinement. Given an initial mesh $\mathcal{T}_0$ which is regular in the sense of Ciarlet, we construct the subsequent meshes $\mathcal{T}_\ell$ by local refinement with the NVB for simplicial meshes in $\mathbb{R}^d$, $d \geq 2$; see, e.g., [31, Chapter 4], respectively, [30]. Consequently, the set of meshes which can be obtained reads

$$\mathbb{T} := \big\{\mathcal{T}_\ell \,:\, \mathcal{T}_\ell \text{ is a refinement of } \mathcal{T}_0\big\}. \tag{2.10}$$

The finite subset of meshes with at most $N \in \mathbb{N}$ elements more than the initial mesh is defined as

$$\mathbb{T}_N := \big\{\mathcal{T}_\ell \in \mathbb{T} \,:\, \#\mathcal{T}_\ell - \#\mathcal{T}_0 \leq N\big\}.$$

The meshes $\mathcal{T}_\ell \in \mathbb{T}$ are regular in the sense of Ciarlet and γ-shape regular in the sense of

$$\gamma^{-1}\,|T|^{1/d} \leq \operatorname{diam}(T) \leq \gamma\,|T|^{1/d} \tag{2.11}$$

for some $\gamma \geq 1$ which depends only on $\mathcal{T}_0$. A refined element $T \in \mathcal{T}_\ell$ is split into at least two sons, i.e., we have

$$\#(\mathcal{T}_\ell \setminus \mathcal{T}_\star) \leq \#\mathcal{T}_\star - \#\mathcal{T}_\ell \tag{2.12}$$

for all refinements $\mathcal{T}_\star \in \mathbb{T}$ of $\mathcal{T}_\ell \in \mathbb{T}$. As a key property for the optimality proof, the crucial closure estimate, for the meshes generated by Algorithm 2.1, is satisfied by

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 \leq C_{\mathrm{mesh}} \sum_{j=0}^{\ell-1} \#\mathcal{M}_j \quad \text{for all } \ell \in \mathbb{N} \tag{2.13}$$

with some constant $C_{\mathrm{mesh}} > 0$ which depends only on $\mathcal{T}_0$. For $d \geq 3$, $\mathcal{T}_0$ has to satisfy a certain condition on the reference edges (cf. [7, 30]), while this assumption can be dropped for $d = 2$; see the recent work [19]. Finally, for two meshes $\mathcal{T}_\ell, \mathcal{T}_\star \in \mathbb{T}$ there is a coarsest common refinement $\mathcal{T}_\ell \oplus \mathcal{T}_\star \in \mathbb{T}$ which satisfies

$$\#(\mathcal{T}_\ell \oplus \mathcal{T}_\star) \leq \#\mathcal{T} + \#\mathcal{T}' - \#\mathcal{T}_0; \tag{2.14}$$

see [11, 29]. We stress that NVB is a binary refinement rule, and the coarsest common refinement $\mathcal{T}_\ell \oplus \mathcal{T}_\star$ is just the overlay of both meshes.

3. Convergence and quasi-orthogonality. The aim of this section is to prove convergence, without relying on symmetry properties of $\mathcal{L}$, which can be done by use of the concept of estimator reduction [3]. To that end, we define the subspace $\mathcal{S}^p_0(\mathcal{T}_\infty)$ of $H^1_0(\Omega)$ which is *theoretically* affected by Algorithm 2.1 as

$$\mathcal{S}^p_0(\mathcal{T}_\infty) := \overline{\bigcup_{\ell \in \mathbb{N}} \mathcal{S}^p_0(\mathcal{T}_\ell)}, \tag{3.1}$$

where the closure is taken with respect to the H^1-norm. With convergence $U_\ell \to u$ and hence $u \in \mathcal{S}^p_0(\mathcal{T}_\infty)$ at hand, we are then able to prove a novel quasi-Galerkin orthogonality estimate (3.7), which is sufficient to prove linear convergence (4.1) as well as optimal convergence rates (5.5).

3.1. Convergence. The following result is proved in [11] for symmetric $\mathcal{L}$ and shows that the error estimator η_ℓ is contractive up to a certain perturbation.

LEMMA 3.1. *There exist constants* $0 < q_{\rm est} < 1$ *and* $C_{\rm est} > 0$ *such that there holds*

$$\eta_{\ell+1}^2 \le q_{\rm est}\eta_\ell^2 + C_{\rm est}\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2 \quad \textit{for all } \ell \in \mathbb{N}. \tag{3.2}$$

The constants $q_{\rm est}$ *and* $C_{\rm est}$ *depend only on* θ, γ*-shape regularity of* $\mathcal{T}_{\ell+1}$*, the polynomial degree* $p \in \mathbb{N}$*, and* Ω.

Proof. The proof follows verbatim the proof of [11, Corollary 3.4]. Therefore, we give a rough sketch only. The application of Young's inequality $2ab \le a^2 + b^2$ proves for $\delta > 0$

$$\begin{aligned}\eta_{\ell+1}^2 \le (1+\delta) \sum_{T'\in\mathcal{T}_{\ell+1}} \Big(|T'|^{2/d}\|\mathcal{L}|_{T'}U_\ell - f\|_{L^2(T')}^2 + |T'|^{1/d}\|[\boldsymbol{A}\nabla U_\ell\cdot n]\|_{L^2(\partial T'\cap\Omega)}^2\Big)\\ + (1+\delta^{-1}) \sum_{T'\in\mathcal{T}_{\ell+1}} \Big(|T'|^{2/d}\|\mathcal{L}|_{T'}(U_{\ell+1}-U_\ell)\|_{L^2(T')}^2\\ + |T'|^{1/d}\|[\boldsymbol{A}\nabla(U_{\ell+1}-U_\ell)\cdot n]\|_{L^2(\partial T'\cap\Omega)}^2\Big).\end{aligned}$$

By use of the regularity assumption on the coefficients and standard inverse estimates as well as the Poincaré inequality, we obtain

$$\begin{aligned}\eta_{\ell+1}^2 &\le (1+\delta) \sum_{T'\in\mathcal{T}_{\ell+1}} \Big(|T'|^{2/d}\|\mathcal{L}|_{T'}U_\ell - f\|_{L^2(T')}^2 + |T'|^{1/d}\|[\boldsymbol{A}\nabla U_\ell\cdot n]\|_{L^2(\partial T'\cap\Omega)}^2\Big)\\ &\le (1+\delta) \sum_{T'\in\mathcal{T}_{\ell+1}} \Big(|T'|^{2/d}\|\mathcal{L}|_{T'}U_\ell - f\|_{L^2(T')}^2 + |T'|^{1/d}\|[\boldsymbol{A}\nabla U_\ell\cdot n]\|_{L^2(\partial T'\cap\Omega)}^2\Big)\\ &\quad + (1+\delta^{-1})C_{\rm stab}\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2.\end{aligned} \tag{3.3}$$

The constant $C_{\rm stab} > 0$ depends only on the γ-shape regularity of $\mathcal{T}_{\ell+1}$, the norms $\|\boldsymbol{A}\|_{W_1^\infty(\Omega)}^2, \|\boldsymbol{b}\|_{L^\infty(\Omega)}^2, \|c\|_{L^\infty(\Omega)}^2$, and the polynomial degree $p \in \mathbb{N}$. Next, the sum is split into two sums over $T' \in \mathcal{T}_\ell \cap \mathcal{T}_{\ell+1}$ and $T' \in \mathcal{T}_{\ell+1} \setminus \mathcal{T}_\ell$. We use the reduction of the element size $|T'| \le |T|/2$ for $T' \subset T$ being a son of a refined element $T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$. Since $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$, one ends up with

$$\begin{aligned}\eta_{\ell+1}^2 &\le (1+\delta)\Big(2^{-1/d}\sum_{T\in\mathcal{T}_\ell\setminus\mathcal{T}_{\ell+1}} \eta_\ell(T)^2 + \sum_{T\in\mathcal{T}_\ell\cap\mathcal{T}_{\ell+1}} \eta_\ell(T)^2\Big)\\ &\quad + (1+\delta^{-1})C_{\rm stab}\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2\\ &\le (1+\delta)\Big(2^{-1/d}\sum_{T\in\mathcal{M}_\ell} \eta_\ell(T)^2 + \sum_{T\in\mathcal{T}_\ell\setminus\mathcal{M}_\ell} \eta_\ell(T)^2\Big)\\ &\quad + (1+\delta^{-1})C_{\rm stab}\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2\\ &\le (1+\delta)\Big((2^{-1/d}-1)\sum_{T\in\mathcal{M}_\ell} \eta_\ell(T)^2 + \eta_\ell^2\Big) + (1+\delta^{-1})C_{\rm stab}\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2.\end{aligned}$$

Finally, Dörfler marking (2.9) proves (3.2) with

$$q_{\rm est} = \big(1-\theta(1-2^{-1/d}\big)(1+\delta) \in (0,1) \quad \text{and} \quad C_{\rm est} = (1+\delta^{-1})C_{\rm stab}$$

for $\delta > 0$ sufficiently small. ∎

Adaptive algorithms of the type of Algorithm 2.1 with nested ansatz spaces $\mathcal{S}_0^p(\mathcal{T}_\ell) \subseteq \mathcal{S}_0^p(\mathcal{T}_{\ell+1})$ have in common that there holds a priori convergence. This has already been observed in the early work [4] and later was used in [24] to prove a general plain convergence result for AFEM.

LEMMA 3.2. *The sequence of Galerkin approximations U_ℓ of Algorithm* 2.1 *is convergent in $H_0^1(\Omega)$, i.e., there exists $u_\infty \in \mathcal{S}_0^p(\mathcal{T}_\infty)$ with*

$$U_\ell \to u_\infty \quad \text{as } \ell \to \infty. \tag{3.4}$$

Proof. The space $\mathcal{S}_0^p(\mathcal{T}_\infty)$ is a closed subspace of $H_0^1(\Omega)$ and therefore the Lax–Milgram lemma guarantees existence and uniqueness of a solution $u_\infty \in \mathcal{S}_0^p(\mathcal{T}_\infty)$ of (2.5) with test space $\mathcal{S}_0^p(\mathcal{T}_\infty)$ instead of $\mathcal{S}_0^p(\mathcal{T}_\ell)$. The Galerkin approximations U_ℓ are also Galerkin approximations of u_∞, since $\mathcal{S}_0^p(\mathcal{T}_\ell) \subseteq \mathcal{S}_0^p(\mathcal{T}_\infty)$ for all $\ell \in \mathbb{N}$. Therefore, the Céa lemma shows

$$\|\nabla(u_\infty - U_\ell)\|_{L^2(\Omega)} \lesssim \min_{V_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell)} \|\nabla(u_\infty - V_\ell)\|_{L^2(\Omega)} \to 0$$

as $\ell \to \infty$. ∎

The combination of estimator reduction (3.2) and a priori convergence (3.4) yields convergence of Algorithm 2.1.

PROPOSITION 3.3. *Algorithm* 2.1 *is convergent in $H_0^1(\Omega)$, i.e.,*

$$U_\ell \to u \in H_0^1(\Omega) \quad \text{as } \ell \to \infty. \tag{3.5}$$

In particular, this implies $u = u_\infty \in \mathcal{S}_0^p(\mathcal{T}_\infty)$.

Proof. According to Lemma 3.2, the estimator reduction (3.2) of Lemma 3.1 takes the form

$$\eta_{\ell+1}^2 \le q_{\rm est}\eta_\ell^2 + \alpha_\ell$$

with $\alpha_\ell \ge 0$ and $\lim_{\ell\to\infty} \alpha_\ell = 0$. From this, elementary calculus proves $\lim_{\ell\to\infty} \eta_\ell = 0$; see, e.g., [3]. Finally, reliability (2.7) of η_ℓ concludes the proof. ∎

3.2. Quasi-Galerkin orthogonality. The standard proof of the Pythagoras theorem $|||u - U_{\ell+1}|||^2 + |||U_{\ell+1} - U_\ell|||^2 = |||u - U_\ell|||^2$ relies on Galerkin orthogonality and symmetry of $b(\cdot,\cdot)$. The following lemmata provide a workaround for our case of a nonsymmetric bilinear form $b(\cdot,\cdot)$. We stress that the quasi-orthogonality proof makes explicit use of the fact that we already have convergence $U_\ell \to u$ in $H_0^1(\Omega)$ and $u \in \mathcal{S}_0^p(\mathcal{T}_\infty)$.

LEMMA 3.4. *The operators $\mathcal{A}, \mathcal{K} : H_0^1(\Omega) \to H^{-1}(\Omega)$ are bounded. Moreover, $\mathcal{A}$ is symmetric and $\mathcal{K}$ is compact.*

Proof. The symmetry of $\mathcal{A}$ is obvious, and both operators $\mathcal{A}$ and $\mathcal{K}$ are also bounded, i.e.,

$$\begin{aligned}
&\|\mathcal{A}v\|_{H^{-1}(\Omega)} \le \|\boldsymbol{A}\|_{L^\infty(\Omega)}\|\nabla v\|_{L^2(\Omega)},\\
&\|\mathcal{K}v\|_{H^{-1}(\Omega)} \le \|\mathcal{K}v\|_{L^2(\Omega)} \le (\|\boldsymbol{b}\|_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)})\|\nabla v\|_{L^2(\Omega)},
\end{aligned}$$

for all $v \in H_0^1(\Omega)$. This implies that $\widetilde{\mathcal{K}} : H_0^1(\Omega) \to L^2(\Omega)$, $\widetilde{\mathcal{K}}v := \mathcal{K}v$ is well-defined and bounded. It remains to prove that $\mathcal{K}$ is compact. The Rellich compactness theorem shows that the embedding $\iota : H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ is a compact operator. Therefore, according to Schauder's theorem (see, e.g., [25, Theorem 4.19]), the adjoint operator

$\iota^\star : L^2(\Omega) \to H^{-1}(\Omega)$ is also compact. Obviously, $\iota^\star : L^2(\Omega) \to H^{-1}(\Omega)$ coincides with the natural embedding, and we may write

$$\mathcal{K} = \iota^\star \circ \widetilde{\mathcal{K}} : H_0^1(\Omega) \to L^2(\Omega) \to H^{-1}(\Omega).$$

Therefore, $\mathcal{K}$ is the composition of a bounded operator and a compact operator and hence compact. This concludes the proof. ∎

LEMMA 3.5. *The sequences $(e_\ell)_{\ell\in\mathbb{N}}$ and $(E_\ell)_{\ell\in\mathbb{N}}$ defined by*

$$e_\ell := \begin{cases} \frac{u-U_\ell}{\|\nabla(u-U_\ell)\|_{L^2(\Omega)}} & \text{for } u \neq U_\ell, \\ 0 & \text{else} \end{cases} \quad \text{and}$$
$$E_\ell := \begin{cases} \frac{U_{\ell+1}-U_\ell}{\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}} & \text{for } U_{\ell+1} \neq U_\ell, \\ 0 & \text{else} \end{cases}$$

converge to zero, weakly in $H_0^1(\Omega)$, i.e.,

$$\lim_{\ell\to\infty} \langle w\,,\, e_\ell\rangle = 0 = \lim_{\ell\to\infty} \langle w\,,\, E_\ell\rangle \quad \text{for all } w \in H^{-1}(\Omega), \tag{3.6}$$

where $\langle\cdot\,,\,\cdot\rangle$ denotes the extended $L^2(\Omega)$-scalar product.

Proof. We prove weak convergence of e_ℓ to zero. The weak convergence of E_ℓ follows with the same arguments. Let (e_{ℓ_j}) be a subsequence of (e_ℓ). Due to boundedness $\|\nabla e_{\ell_j}\|_{L^2(\Omega)} \le 1$ for all $j \in \mathbb{N}$, we may extract a weakly convergent subsequence (e_{ℓ_k}) of (e_{ℓ_j}) with

$$e_{\ell_k} \rightharpoonup w \in H_0^1(\Omega).$$

First, note that $u, U_\ell \in \mathcal{S}_0^p(\mathcal{T}_\infty)$ implies $e_\ell \in \mathcal{S}_0^p(\mathcal{T}_\infty)$ and hence $w \in \mathcal{S}_0^p(\mathcal{T}_\infty)$. Second, for all $\ell_k \ge \ell$ with $e_{\ell_k} \neq 0$ and all $V_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell)$, it holds that

$$b(e_{\ell_k}, V_\ell) = \|\nabla(u - U_{\ell_k})\|_{L^2(\Omega)}^{-1} b(u - U_{\ell_k}, V_\ell) = 0.$$

For any $\ell \in \mathbb{N}$, $V_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell)$, and $\varepsilon > 0$, there exists $k_0 \in \mathbb{N}$ such that for all $k \ge k_0$, it holds that

$$|b(w, V_\ell)| = |\langle w\,,\, \mathcal{L}^\star V_\ell\rangle| \le \varepsilon + |\langle e_{\ell_k}\,,\, \mathcal{L}^\star V_\ell\rangle| = \varepsilon + |b(e_{\ell_k}, V_\ell)| = \varepsilon,$$

since k_0 is chosen large enough such that $\ell_k \ge \ell$. Therefore

$$b(w, V_\ell) = 0 \quad \text{for all } V_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell) \text{ and } \ell \in \mathbb{N}.$$

Due to definiteness of $b(\cdot,\cdot)$ and $w \in \mathcal{S}_0^p(\mathcal{T}_\infty) := \overline{\bigcup_{\ell\in\mathbb{N}} \mathcal{S}_0^p(\mathcal{T}_\ell)}$, this implies $w = 0$. Altogether, we have now shown that each subsequence of e_ℓ has a subsequence which converges weakly to zero. This immediately implies weak convergence $e_\ell \rightharpoonup 0$ as $\ell \to \infty$. ∎

The previous lemma shows that although $(E_\ell)_{\ell\in\mathbb{N}}$ is no orthonormal sequence, it shares the property of weak convergence to zero with orthonormal systems. Note that our proof already used convergence $U_\ell \to u$ as $\ell \to \infty$ in the sense that we required $u - U_\ell \in \mathcal{S}_0^p(\mathcal{T}_\infty)$. This suffices to prove the following quasi-Pythagoras theorem.

PROPOSITION 3.6. *For any $0 < \varepsilon < 1$, there exists $\ell_0 \in \mathbb{N}$ such that*

$$|\!|\!|U_{\ell+1} - U_\ell|\!|\!|^2 \le \frac{1}{1-\varepsilon}\, |\!|\!|u - U_\ell|\!|\!|^2 - |\!|\!|u - U_{\ell+1}|\!|\!|^2 \tag{3.7}$$

for all $\ell \ge \ell_0$.

Remark 3.7. As in [12, Theorem 5.1], the quasi-orthogonality (3.7) is an asymptotic statement. The advantage here is that (3.7) is automatically guaranteed after ℓ_0 steps of Algorithm 2.1. In contrast to that, [12, Assumption 4.3] used to prove [12, Theorem 5.1], includes a mesh-size condition of the form $|T|^{1/d} \le h_{\max} \ll 1$ for all $T \in \mathcal{T}_\ell$ which is not necessarily enforced by Algorithm 2.1, unless the initial mesh is already sufficiently fine. Moreover, $h_{\max}$ is unknown in general and depends on the regularity of the dual problem.

Proof. Lemma 3.5 shows that $e_\ell, E_\ell \rightharpoonup 0$ as $\ell \to \infty$. Due to Lemma 3.4, $\mathcal{K}$ is compact. Therefore, we have strong convergence $\mathcal{K}e_\ell, \mathcal{K}E_\ell \to 0$ in $H^{-1}(\Omega)$ as $\ell \to \infty$. This shows

$$\begin{aligned}\langle \mathcal{K}(u-U_{\ell+1})\,,\,U_{\ell+1}-U_\ell\rangle &= \langle \mathcal{K}e_{\ell+1}\,,\,U_{\ell+1}-U_\ell\rangle \|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}\\ &\le \|\mathcal{K}e_{\ell+1}\|_{H^{-1}(\Omega)}\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}\end{aligned}$$

as well as

$$\begin{aligned}\langle \mathcal{K}(U_{\ell+1}-U_\ell)\,,\,u-U_{\ell+1}\rangle &= \langle \mathcal{K}E_\ell\,,\,u-U_{\ell+1}\rangle \|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}\\ &\le \|\mathcal{K}E_\ell\|_{H^{-1}(\Omega)}\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}.\end{aligned}$$

For any $\delta > 0$, this may be employed to obtain some $\ell_0 \in \mathbb{N}$ such that for all $\ell \ge \ell_0$, it holds that

$$\begin{aligned}&|\langle \mathcal{K}(U_{\ell+1}-U_\ell)\,,\,u-U_{\ell+1}\rangle| + |\langle \mathcal{K}(u-U_{\ell+1})\,,\,U_{\ell+1}-U_\ell\rangle|\\ &\quad\le \delta\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}.\end{aligned}$$

Together with Galerkin orthogonality

$$0 = b(u-U_{\ell+1}, V_{\ell+1}) = \langle \mathcal{L}(u-U_{\ell+1})\,,\,V_{\ell+1}\rangle \quad \text{for all } V_{\ell+1} \in \mathcal{S}^p_0(\mathcal{T}_{\ell+1}), \tag{3.8}$$

we estimate

$$\begin{aligned}&|\langle \mathcal{L}(U_{\ell+1}-U_\ell)\,,\,u-U_{\ell+1}\rangle|\\ &\quad= |\langle \mathcal{A}(u-U_{\ell+1})\,,\,U_{\ell+1}-U_\ell\rangle + \langle \mathcal{K}(U_{\ell+1}-U_\ell)\,,\,u-U_{\ell+1}\rangle|\\ &\quad\le |\langle \mathcal{L}(u-U_{\ell+1})\,,\,U_{\ell+1}-U_\ell\rangle| + |\langle \mathcal{K}(U_{\ell+1}-U_\ell)\,,\,u-U_{\ell+1}\rangle|\\ &\qquad+ |\langle \mathcal{K}(u-U_{\ell+1})\,,\,U_{\ell+1}-U_\ell\rangle|\\ &\quad\le \delta\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}.\end{aligned} \tag{3.9}$$

The definition of $|\!|\!|\cdot|\!|\!|$ and Galerkin orthogonality (3.8) yield

$$|\!|\!|u-U_{\ell+1}|\!|\!|^2 + |\!|\!|U_{\ell+1}-U_\ell|\!|\!|^2 + \langle \mathcal{L}(U_{\ell+1}-U_\ell)\,,\,u-U_{\ell+1}\rangle = |\!|\!|u-U_\ell|\!|\!|^2,$$

whence

$$|\!|\!|U_{\ell+1}-U_\ell|\!|\!|^2 \le |\!|\!|u-U_\ell|\!|\!|^2 - |\!|\!|u-U_{\ell+1}|\!|\!|^2 + \delta C_{\rm norm}^2 |\!|\!|u-U_{\ell+1}|\!|\!|\,|\!|\!|U_{\ell+1}-U_\ell|\!|\!|.$$

The application of Young's inequality $2ab \le a^2 + b^2$ and the choice $\varepsilon = \delta C_{\rm norm}^2/2$ conclude the proof. ∎

4. Contraction. The quasi-Pythagoras theorem (3.7) from Proposition 3.6 allows us to prove R-linear convergence of the error estimator η_ℓ. Compared with the analysis of the symmetric case [11], this is a weaker result. However, R-linear convergence is still sufficient to prove quasi-optimal convergence rates in section 5.

THEOREM 4.1. *There exist constants* $0 < q_{\rm conv} < 1$ *and* $C_{\rm conv} > 0$ *such that for all* $\ell, k \in \mathbb{N}$*, there holds*

$$\eta_{\ell+k}^2 \le C_{\rm conv} q_{\rm conv}^k \, \eta_\ell^2. \tag{4.1}$$

The constants $q_{\rm conv}$ *and* $C_{\rm conv}$ *depend only on* $q_{\rm est}$*,* $C_{\rm est}$*,* $C_{\rm norm}$*,* $C_{\rm rel}$*, and hence implicitly on* Ω*, the used mesh-refinement strategy, the adaptivity parameter* θ*, as well as on the polynomial degree* p*.*

Remark 4.2. (i) The R-linear convergence (4.1) is sufficient to prove optimal convergence rates in Theorem 5.3 below. This generalizes the proofs of [12, 11] in the following sense. There, contraction of an appropriate sum of error and estimator (the so-called quasi error [11]) or of error and oscillations (the so-called total error [12]) is proved. Either sum is equivalent to the error estimator (up to reliability and efficiency constants), and hence R-linear convergence (4.1) of the estimator follows.

(ii) Moreover, R-linear convergence seems to be a necessary relaxation of the analysis to cover strongly nonsymmetric problems. The possible preasymptotic range, as observed for example in convection dominated problems, is reflected in the constant $C_{\rm conv}$ which is unknown in general. However, the result states that the adaptive algorithm is capable of overcoming the preasymptotic range at some point and reveals the optimal convergence rates as shown in Theorem 5.3 below.

Proof. We employ the estimator reduction (3.2) and reliability (2.7) to obtain for $N \ge \ell + 1$ and $\alpha < 1 - q_{\rm est}$

$$\begin{aligned}\sum_{k=\ell+1}^{N} \eta_k^2 &\le \sum_{k=\ell+1}^{N} \big(q_{\rm est}\eta_{k-1}^2 + C_{\rm est}\|\nabla(U_k - U_{k-1})\|_{L^2(\Omega)}^2 \big) \\ &\le \sum_{k=\ell+1}^{N} \Big((q_{\rm est} + \alpha)\eta_{k-1}^2 + C_{\rm est}\big(\|\nabla(U_k - U_{k-1})\|_{L^2(\Omega)}^2 \\ &\quad - \alpha C_{\rm rel}^{-2} C_{\rm est}^{-1} \|\nabla(u - U_{k-1})\|_{L^2(\Omega)}^2\big)\Big).\end{aligned}$$

Rearranging the terms in the above estimate, we end up with

$$\begin{aligned}(1 - q_{\rm est} - \alpha) \sum_{k=\ell+1}^{N} \eta_k^2 &\le (q_{\rm est} + \alpha)\eta_\ell^2 \\ &\quad + C_{\rm est} C_{\rm norm}^2 \sum_{k=\ell+1}^{N} \big(|||U_k - U_{k-1}|||^2 - \delta |||u - U_{k-1}|||^2 \big),\end{aligned}$$

where $\delta = \alpha C_{\rm rel}^{-2} C_{\rm est}^{-1} C_{\rm norm}^{-4}$. Next, we aim at proving that the sum on the right-hand side is bounded above by η_ℓ^2 for all $N \in \mathbb{N}$. To that end, we employ Proposition 3.6

with $\varepsilon > 0$ such that $1/(1-\varepsilon) \le 1+\delta$. This gives a number $\ell_0 \in \mathbb{N}$ such that for all $N > \ell \ge \ell_0$, we may estimate

$$
\begin{aligned}
(4.2)\qquad &\sum_{k=\ell+1}^{N} \big(|||U_k - U_{k-1}|||^2 - \delta |||u - U_{k-1}|||^2 \big) \\
&\le \sum_{k=\ell+1}^{N} \left(\left(\frac{1}{1-\varepsilon} - \delta \right) |||u - U_{k-1}|||^2 - |||u - U_k|||^2 \right) \\
&\le \sum_{k=\ell+1}^{N} \big(|||u - U_{k-1}|||^2 - |||u - U_k|||^2 \big) \\
&\le |||u - U_\ell|||^2 \le C_{\rm norm}^2 C_{\rm rel}^2 \eta_\ell^2 .
\end{aligned}
$$

For all $\ell < \ell_0$, we first observe that $|||u - U_\ell||| = 0$ implies $|||U_k - U_{k-1}||| = 0$ for all $k \ge \ell + 1$, since $U_k = u = U_{k-1}$. Therefore, we obtain with the convention $\infty \cdot 0 = 0$

$$C_{\rm sup} := \sup_{\ell \in \{1,\dots,\ell_0\}} \left(|||u - U_\ell|||^{-2} \sum_{k=\ell+1}^{\ell_0} |||U_k - U_{k-1}|||^2 \right) < \infty.$$

In combination with (4.2), we thus see for all $\ell \in \mathbb{N}$ and all $N > \ell$

$$\sum_{k=\ell+1}^{N} \big(|||U_k - U_{k-1}|||^2 - \delta |||u - U_{k-1}|||^2 \big) \le (1 + C_{\rm sup}) C_{\rm norm}^2 C_{\rm rel}^2 \eta_\ell^2 .$$

Plugging everything together, we have so far shown

$$(4.3)\qquad \sum_{k=\ell+1}^{\infty} \eta_k^2 \le C \eta_\ell^2 \quad \text{for all } \ell \in \mathbb{N}$$

for some constant $C > 0$ which depends only on $q_{\rm est}$, $C_{\rm est}$, $C_{\rm norm}$, and $C_{\rm rel}$. Therefore, we get

$$(1 + C^{-1}) \sum_{k=\ell+1}^{\infty} \eta_k^2 \le \sum_{k=\ell+1}^{\infty} \eta_k^2 + \eta_\ell^2 = \sum_{k=\ell}^{\infty} \eta_k^2 ,$$

and hence by induction

$$\eta_{\ell+j}^2 \le \sum_{k=\ell+j}^{\infty} \eta_k^2 \le (1 + C^{-1})^{-j} \sum_{k=\ell}^{\infty} \eta_k^2 \le (1 + C)(1 + C^{-1})^{-j} \eta_\ell^2 \quad \text{for all } \ell, j \in \mathbb{N}_0 .$$

This concludes the proof with $q_{\rm conv} = 1/(1 + C^{-1})$ and $C_{\rm conv} = (1 + C)$. ∎

Remark 4.3. Note that the R-linear convergence of Theorem 4.1 holds for arbitrary adaptivity parameters $0 < \theta < 1$. Moreover, the result is independent of NVB in the sense that the proof only requires that the meshes $\mathcal{T}_\ell$ are uniformly γ-shape regular and that $|T'| \le q|T|$ for some $0 < q < 1$ and all sons $T' \subset T$ of refined elements $T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$. These properties hold for each feasible mesh-refinement strategy and for NVB with $q = 1/2$. Finally, the minimal cardinality of the set $\mathcal{M}_\ell$ of marked elements has not been used yet. Instead, Theorem 4.1 holds as long as the set $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ satisfies the Dörfler marking (2.9) and, in particular, for $\mathcal{M}_\ell = \mathcal{T}_\ell$.

Remark 4.4. Note that the proof of Theorem 4.1 uses neither linearity nor uniform ellipticity of $\mathcal{L}$. Instead, we only require reliability (2.7), estimator reduction (3.2), quasi-Galerkin orthogonality (3.7), and equivalence (2.4) of the norm $\|\nabla(\cdot)\|_{L^2(\Omega)}$ and the energy quasi-norm $|\!|\!|\cdot|\!|\!|$ on $H^1_0(\Omega)$. With these ingredients, our analysis is thus also capable of covering certain nonlinear problems, as discussed in section 6.5.

5. Optimal convergence rates. With Theorem 4.1 at hand, we are in the position to prove quasi-optimal convergence rates for the sequence of Galerkin solutions obtained from Algorithm 2.1. First, however, we have to clarify what is the best possible convergence rate that can be aimed at. To that end, we follow, e.g., [11] and define the approximation class $\mathbb{A}_s$ by

$$(u,f)\in\mathbb{A}_s \quad\overset{\text{def}}{\Longleftrightarrow}\quad \|(u,f)\|_{\mathbb{A}_s} := \sup_{N\in\mathbb{N}} N^s\sigma(N;u,f)<\infty \tag{5.1a}$$

for all $s>0$, where

$$\sigma(N;u,f) := \inf_{\mathcal{T}_\star\in\mathbb{T}_N}\ \inf_{V_\star\in\mathcal{S}^p_0(\mathcal{T}_\star)} \big(\|\nabla(u-V_\star)\|^2_{L^2(\Omega)}+\mathrm{osc}_\star(V_\star)^2\big)^{1/2} \tag{5.1b}$$

and $\mathrm{osc}_\star$ is the oscillation term from (2.8) corresponding to the mesh $\mathcal{T}_\star$. We refer to [8, 17] for a characterization of approximation classes in terms of Besov regularity. However, in this work, we follow [2] and use an equivalent definition of $\mathbb{A}_s$, which involves the error estimator η_ℓ only. This equivalence is part of the next lemma, which is also implicitly contained in [11, Lemma 5.2].

LEMMA 5.1. *There exists a constant $C_1>0$ such that for all $\mathcal{T}_\star\in\mathbb{T}$ there holds*

$$C_1^{-1}\eta_\star^2 \le \inf_{V_\star\in\mathcal{S}^p_0(\mathcal{T}_\star)}\big(\|\nabla(u-V_\star)\|^2_{L^2(\Omega)}+\mathrm{osc}_\star(V_\star)^2\big)\le C_1\eta_\star^2, \tag{5.2}$$

where $\eta_\star$ denotes the error estimator with respect to $\mathcal{T}_\star$. Hence, $\mathbb{A}_s$ from (5.1) can equivalently be characterized as

$$(u,f)\in\mathbb{A}_s \quad\Longleftrightarrow\quad \sup_{N\in\mathbb{N}}\inf_{\mathcal{T}_\star\in\mathbb{T}_N} N^s\eta_\star<\infty \tag{5.3}$$

for all $s>0$. The constant C_1 depends only on $C_{\mathrm{cont}}, C_{\mathrm{ell}}$, the γ-shape regularity of $\mathcal{T}_\star$, and the polynomial degree $p\in\mathbb{N}$.

Proof. First, we prove (5.2). To that end, we observe $\|\nabla(u-U_\star)\|^2_{L^2(\Omega)}+\mathrm{osc}_\star(U_\star)^2\simeq\eta_\star^2$, which follows from reliability (2.7), efficiency (2.8), and $\mathrm{osc}_\star(U_\star)\le\eta_\star$. Moreover, the lower bound

$$\inf_{V_\star\in\mathcal{S}^p_0(\mathcal{T}_\star)}\big(\|\nabla(u-V_\star)\|^2_{L^2(\Omega)}+\mathrm{osc}_\star(V_\star)^2\big)\le\|\nabla(u-U_\star)\|^2_{L^2(\Omega)}+\mathrm{osc}_\star(U_\star)^2 \tag{5.4}$$

holds since $U_\star\in\mathcal{S}^p_0(\mathcal{T}_\star)$. To prove the converse estimate in (5.4), we argue as in Lemma 3.1 and use a standard inverse estimate as well as the Poincaré inequality, to see

$$
\begin{aligned}
\operatorname{osc}_\star(U_\star)^2 &= \sum_{T\in\mathcal{T}_\star} |T|^{2/d} \|(1-\Pi_\star^{p-1})(\mathcal{L}|_T U_\star - f)\|_{L^2(T)}^2 \\
&\lesssim \sum_{T\in\mathcal{T}_\star} |T|^{2/d} \|(1-\Pi_\star^{p-1})\mathcal{L}|_T (U_\star - V_\star)\|_{L^2(T)}^2 + \operatorname{osc}_\star(V_\star)^2 \\
&\lesssim \big(\|\boldsymbol{A}\|_{W_1^\infty(\Omega)}^2 + \|\boldsymbol{b}\|_{L^\infty(\Omega)}^2 + \|c\|_{L^\infty(\Omega)}^2\big) \sum_{T\in\mathcal{T}_\star} |T|^{2/d} \|U_\star - V_\star\|_{H^2(T)}^2 \\
&\quad + \operatorname{osc}_\star(V_\star)^2 \\
&\lesssim \|\nabla(U_\star - V_\star)\|_{L^2(\Omega)}^2 + \operatorname{osc}_\star(V_\star)^2.
\end{aligned}
$$

Finally, by use of the Céa lemma, we end up with

$$
\begin{aligned}
\operatorname{osc}_\star(U_\star)^2 &\lesssim \|\nabla(u - U_\star)\|_{L^2(\Omega)}^2 + \|\nabla(u - V_\star)\|_{L^2(\Omega)}^2 + \operatorname{osc}_\star(V_\star)^2 \\
&\lesssim \|\nabla(u - V_\star)\|_{L^2(\Omega)}^2 + \operatorname{osc}_\star(V_\star)^2.
\end{aligned}
$$

The combination of the last three estimates proves (5.2). The characterization (5.3) follows with (5.2) and the definition of $\sigma(N; u, f)$ in (5.1b). ∎

Remark 5.2. Note that efficiency (2.8) is only used in the foregoing characterization of the approximation class $\mathbb{A}_s$. In our opinion, this approach allows for a clearer presentation of the proof of the following quasi-optimality theorem and, in particular, we shall see that unlike the analysis of [11, 12, 20, 29], the upper bound $\theta_\star$ for optimal adaptivity parameters $0 < \theta < \theta_\star$ does not depend on the efficiency constant $C_{\rm eff}$. Moreover, if one defines the approximation class $\mathbb{A}_s$ by (5.3) instead of (5.1), one may fully avoid any efficiency estimate. The overall result then states that the algorithm leads to the best possible convergence rate for the estimator used. Such an observation might be of independent interest for further investigations of nonlinear problems.

The following result is the main theorem of this section.

THEOREM 5.3. *Define $\theta_\star := (1 + C_{\rm stab} C_{\rm dRel})^{-1}$ with the constants $C_{\rm dRel} > 0$ from Lemma* 5.4 *and $C_{\rm stab} > 0$ from the proof of Lemma* 3.1*. Then, for all adaptivity parameters $0 < \theta < \theta_\star$ and all $s > 0$, there exists a constant $C_{\rm opt} > 0$ such that*

$$
(5.5) \qquad (u, f) \in \mathbb{A}_s \iff \eta_\ell \le C_{\rm opt} \|(u,f)\|_{\mathbb{A}_s} (\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-s} \quad \text{for all } \ell \in \mathbb{N}.
$$

The constant $C_{\rm opt}$ depends only on θ, s, $q_{\rm conv}$, $C_{\rm conv}$, $C_{\rm eff}$, and $C_{\rm mesh}$, and the proof relies on the properties (2.11)–(2.14) *of NVB.*

For the proof of the quasi-optimality theorem, we need a refined reliability property of the error estimator η_ℓ.

LEMMA 5.4 (discrete reliability). *There exists a constant $C_{\rm dRel} > 0$ such that for all refinements $\mathcal{T}_\star \in \mathbb{T}$ of a triangulation $\mathcal{T}_\ell \in \mathbb{T}$, it holds that*

$$
(5.6) \qquad \|\nabla(U_\star - U_\ell)\|_{L^2(\Omega)}^2 \le C_{\rm dRel} \sum_{T\in\mathcal{T}_\ell\setminus\mathcal{T}_\star} \eta_\ell(T)^2.
$$

The constant $C_{\rm dRel}$ depends only on the γ-shape regularity of $\mathcal{T}_0$, the polynomial degree $p \in \mathbb{N}$, and Ω.

Proof. The statement is proved for $\boldsymbol{b} = 0$ and $c \ge 0$ in [11, Lemma 3.6]. The proof for the present case follows verbatim. ∎

So far, we have observed that Dörfler marking (2.9) implies contraction of η_ℓ (Proposition 4.1). Now we prove, in some sense, the converse. We follow the concept of proof of [2] and stress that unlike the corresponding results in, e.g., [11, 12, 20, 29] our proof does not use efficiency (2.8) of η_ℓ.

LEMMA 5.5 (optimality of Dörfler marking). *Suppose* $0 < \theta < \theta_\star := (1 + C_{\rm stab}C_{\rm dRel})^{-1}$. *Then, there exists* $0 < q_{\rm D} < 1$ *such that for all refinements* $\mathcal{T}_\star \in \mathbb{T}$ *of a triangulation* $\mathcal{T}_\ell \in \mathbb{T}$ *the following statement is true:*

$$\eta_\star^2 \le q_{\rm D}\eta_\ell^2 \quad\Longrightarrow\quad \theta\eta_\ell^2 \le \sum_{T\in\mathcal{T}_\ell\setminus\mathcal{T}_\star} \eta_\ell(T)^2. \tag{5.7}$$

Proof. Analogously to (3.3), we estimate for $\delta > 0$

$$\begin{aligned}\eta_\ell^2 &= \sum_{T\in\mathcal{T}_\ell\setminus\mathcal{T}_\star}\eta_\ell(T)^2 + \sum_{T\in\mathcal{T}_\ell\cap\mathcal{T}_\star}\eta_\ell(T)^2\\ &\le \sum_{T\in\mathcal{T}_\ell\setminus\mathcal{T}_\star}\eta_\ell(T)^2 + (1+\delta^{-1})\sum_{T\in\mathcal{T}_\ell\cap\mathcal{T}_\star}\eta_\star(T)^2 + (1+\delta)C_{\rm stab}\|\nabla(U_\star - U_\ell)\|_{L^2(\Omega)}^2\\ &\le \sum_{T\in\mathcal{T}_\ell\setminus\mathcal{T}_\star}\eta_\ell(T)^2 + (1+\delta^{-1})q_{\rm D}\eta_\ell^2 + (1+\delta)C_{\rm stab}\|\nabla(U_\star - U_\ell)\|_{L^2(\Omega)}^2.\end{aligned}\tag{5.8}$$

Rearranging the terms and employing the discrete reliability (2.7), we end up with

$$\frac{1-(1+\delta^{-1})q_{\rm D}}{1+(1+\delta)C_{\rm stab}C_{\rm dRel}}\,\eta_\ell^2 \le \sum_{T\in\mathcal{T}_\ell\setminus\mathcal{T}_\star}\eta_\ell(T)^2.$$

According to $\theta < (1 + C_{\rm stab}C_{\rm dRel})^{-1}$, we may finally choose $\delta > 0$ and $0 < q_{\rm D} < 1$ sufficiently small to ensure

$$\theta \le \frac{1-(1+\delta)q_{\rm D}}{1+(1+\delta^{-1})C_{\rm stab}C_{\rm dRel}} < \frac{1}{1+C_{\rm stab}C_{\rm dRel}}.$$

This concludes the proof. ∎

Now, we are in the position to prove Theorem 5.3. We stress that the concept of proof goes back to [29] and has been adopted by [11] and all succeeding works. We put emphasis on the fact that, first, efficiency (2.8) of η_ℓ is not used explicitly and that, second, R-linear convergence (4.1) instead of plain contraction in each step of the adaptive loop is sufficient.

Proof of Theorem 5.3. Let $\lambda > 0$ denote a free parameter, which is fixed later on. The definition of the approximation class $\mathbb{A}_s$ allows for given $\varepsilon^2 := \lambda\eta_\ell^2 > 0$ to choose a mesh $\mathcal{T}_\varepsilon \in \mathbb{T}$ such that

$$\eta_\varepsilon \le \varepsilon \quad\text{and}\quad \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \lesssim \|(u,f)\|_{\mathbb{A}_s}^{1/s}\varepsilon^{-1/s}.$$

Now, consider the overlay $\mathcal{T}_\star := \mathcal{T}_\varepsilon \oplus \mathcal{T}_\ell$ and argue similarly to (3.3) to see

$$\eta_\star^2 \lesssim \eta_\varepsilon^2 + \|\nabla(U_\star - U_\varepsilon)\|_{L^2(\Omega)}^2 \lesssim \eta_\varepsilon^2 \le \lambda\eta_\ell^2,$$

where we used the definition of $\varepsilon > 0$. We choose $\lambda > 0$ sufficiently small such that Lemma 5.5 is applicable and conclude that $\mathcal{T}_\ell \setminus \mathcal{T}_\star$ satisfies the Dörfler marking (2.9). By definition of step (iii) of Algorithm 2.1, the set $\mathcal{M}_\ell$ of marked elements is a set of minimal cardinality which satisfies the Dörfler marking. Therefore, we obtain by use of (2.12) and (2.14)

$$\begin{aligned}\#\mathcal{M}_\ell \le \#(\mathcal{T}_\ell \setminus \mathcal{T}_\star) \le \#\mathcal{T}_\star - \#\mathcal{T}_\ell \le \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 &\lesssim \|(u,f)\|_{\mathbb{A}_s}^{1/s}\varepsilon^{-1/s} \\ &\lesssim \|(u,f)\|_{\mathbb{A}_s}^{1/s}\eta_\ell^{-1/s}\end{aligned} \tag{5.9}$$

for all $\ell \in \mathbb{N}$. Finally, the closure estimate (2.13) and the contraction (4.1) of Proposition 4.1 yield

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 \lesssim \sum_{j=0}^{\ell-1} \#\mathcal{M}_\ell \lesssim \|(u,f)\|_{\mathbb{A}_s}^{1/s} \sum_{j=0}^{\ell-1} \eta_j^{-1/s} \lesssim \|(u,f)\|_{\mathbb{A}_s}^{1/s} \eta_\ell^{-1/s} \sum_{j=0}^{\ell-1} q_{\rm conv}^{(\ell-j)/s}.$$

Exploiting the convergence of the geometric series, we end up with

$$\eta_\ell^2 \lesssim \|(u,f)\|_{\mathbb{A}_s}(\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-s} \quad \text{for all } \ell \in \mathbb{N}.$$

Altogether, this proves that each theoretically possible convergence rate for the estimator is, in fact, asymptotically achieved by the adaptive algorithm. The converse implication in (5.5) is obvious. This concludes the proof. ∎

Remark 5.6. We stress that the overall proof of Theorem 5.3 depends only on properties (2.11)–(2.14) of NVB, R-linear convergence (4.1) of the estimator used, and the discrete reliability (5.6). In particular, there is no explicit use of the properties of the differential operator $\mathcal{L}$, i.e., neither linearity nor uniform ellipticity is required. However, the proofs of Proposition 3.6 and Theorem 4.1 require the equivalences

$$|||u - U_\ell||| \simeq \|\nabla(u - U_\ell)\|_{L^2(\Omega)} \quad \text{and} \quad |||U_{\ell+1} - U_\ell||| \simeq \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}$$

for all sufficiently large $\ell \in \mathbb{N}$, i.e., $\ell \ge \ell_0$.

6. Extensions. In this section, we want to discuss some possible extensions of our analysis.

6.1. Minimal cardinality of marked elements. The choice of the set of marked elements $\mathcal{M}_\ell$ in step (iii) of Algorithm 2.1 to be a set of minimal cardinality which satisfies the Dörfler marking (2.9) requires sorting the set $\{\eta_\ell(T) : T \in \mathcal{T}_\ell\}$, which takes at least $\mathcal{O}(\#\mathcal{T}_\ell \log(\#\mathcal{T}_\ell))$ operations. In comparison to $\mathcal{O}(\#\mathcal{T}_\ell)$ operations for iterative solvers on sparse matrices, marking becomes the bottleneck of Algorithm 2.1. To overcome this problem, we may allow the set $\mathcal{M}_\ell$ to be of *almost* minimal cardinality in the sense of

$$\#\mathcal{M}_\ell \le C\#\widetilde{\mathcal{M}}_\ell \quad \text{for all } \ell \in \mathbb{N}, \tag{6.1}$$

where $\widetilde{\mathcal{M}}_\ell$ is a set of minimal cardinality which satisfies Dörfler marking and $C > 0$ is an arbitrary but fixed constant. All the proofs hold true up to an additional factor C, which is involved in (5.9). The relaxation (6.1) allows us to apply an inexact sorting algorithm based on binning of the data (see, e.g., [22]) which performs in $\mathcal{O}(\#\mathcal{T}_\ell)$ operations.

6.2. Other mesh-refinement strategies. Instead of the simple NVB, one can consider other mesh-refinement strategies which satisfy (2.12)–(2.14), since no other property of the mesh-refinement strategy is used throughout this paper. In particular, one could use up to m NVBs per marked element, where $m \in \mathbb{N}$ is a fixed number; cf., e.g., [20]. This includes the strategy proposed in [12] which uses additional bisections every nth step to ensure the interior node property and hence to obtain a discrete lower bound on the error. Moreover, one can relax the regularity of the triangulations used and allow a fixed number of hanging nodes in each triangle $T \in \mathcal{T}_\ell$ [9].

6.3. Inhomogeneous Dirichlet data. Let $\mathcal{S}^p(\mathcal{T}_\ell) := \mathcal{P}^p(\mathcal{T}_\ell) \cap H^1(\Omega)$ with discrete trace space $\mathcal{S}^p(\mathcal{T}_\ell|_\Gamma) := \{V_\ell|_\Gamma \,:\, V_\ell \in \mathcal{S}^p(\mathcal{T}_\ell)\}$. We consider inhomogeneous Dirichlet data $g \in H^{1/2}(\Gamma)$ and an $H^{1/2}$-stable projection $P_\ell : H^{1/2}(\Gamma) \to \mathcal{S}^p(\mathcal{T}_\ell|_\Gamma)$, for instance, the Scott–Zhang projection [28] for $p \geq 1$ or the L^2-projection for $p = 1$. (See [19] for H^1-stability on NVB refined meshes.) The continuous problem we want to solve now reads as follows: Find $u \in H^1(\Omega)$ with $u|_{\partial\Omega} = g$ such that

$$\langle \mathcal{L}u\,,\, v\rangle = b(u,v) = \int_\Omega f v\,dx \quad \text{for all } v \in H^1_0(\Omega). \tag{6.2}$$

The corresponding discrete formulation reads as follows: Find $U_\ell \in \mathcal{S}^p(\mathcal{T}_\ell)$ with $U_\ell|_\Gamma = P_\ell g$ such that

$$b(U_\ell, V_\ell) = \int_\Omega f V_\ell\,dx \quad \text{for all } V_\ell \in \mathcal{S}^p_0(\mathcal{T}_\ell). \tag{6.3}$$

Well-posedness of (6.2)–(6.3) is well-known and discussed, e.g., in [2, 5, 26]. The approximation error which is introduced via $g \approx P_\ell g$ results in an additional error quantity. We assume regularity $g \in H^1(\Gamma)$ and define the Dirichlet data oscillations

$$\mathrm{osc}_{g,\ell} := \sum_{E \in \mathcal{T}_\ell|_\Gamma} \mathrm{diam}(E)\|\nabla_\Gamma(1 - P_\ell)g\|^2_{L^2(E)},$$

where $\nabla_\Gamma(\,\cdot\,)$ denotes the surface gradient on $\Gamma = \partial\Omega$.

Since the ansatz spaces are no longer nested, i.e., $U_{\ell+1} - U_\ell \notin \mathcal{S}^p_0(\mathcal{T}_\ell)$, we have to rely on a modified marking strategy proposed in [29]. We replace the Dörfler marking (2.9) by the following separate marking strategy with adaptivity parameters $0 < \theta, \vartheta < 1$:

- If $\mathrm{osc}^2_{g,\ell} \leq \vartheta\eta^2_\ell$, determine $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ as a set of minimal cardinality which satisfies (2.9).
- If $\mathrm{osc}^2_{g,\ell} > \vartheta\eta^2_\ell$, determine $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ as a set of minimal cardinality which satisfies

$$\theta\,\mathrm{osc}^2_{g,\ell} \leq \sum_{T \in \mathcal{M}_\ell} \mathrm{osc}_{g,\ell}(T)^2. \tag{6.4}$$

Now, the analysis of [2] can easily be transferred to the present problem as well, where η_ℓ in (3.2), (4.1), and (5.3)–(5.5) is replaced by $\rho_\ell := \eta_\ell + \mathrm{osc}_{g,\ell}$. For usual choices of P_ℓ as above, one obtains convergence of AFEM by means of the estimator reduction principle [2, Theorem 4]. Moreover, for arbitrary P_ℓ and sufficiently small marking parameters $0 < \vartheta, \theta < 1$, we obtain the optimality result of Theorem 5.3; cf. [2, Theorem 6].

For $d = 2$, one may even use nodal interpolation to discretize the inhomogeneous Dirichlet data. Then, the combined Dörfler marking (2.9) for $\rho_\ell := \eta_\ell + \mathrm{osc}_{g,\ell}$ instead of η_ℓ yields the contraction result of Theorem 4.1. Moreover, for sufficiently small $0 < \theta < 1$, Theorem 5.3 remains valid. We refer to [15] in case of symmetric $\mathcal{L} = -\Delta$ and stress that the analysis can easily be transferred to the present setting.

6.4. Coercive but not uniformly elliptic bilinear forms. Assume that instead of ellipticity (2.3), there holds a Gårding inequality

$$b(u,u) + C_{\text{gård}}\|u\|^2_{L^2(\Omega)} \geq \rho_{\text{gård}}\|\nabla u\|^2_{L^2(\Omega)} \quad \text{for all } u \in H^1(\Omega) \tag{6.5}$$

with constants $0 < \rho_{\text{gård}} < 1$ and $C_{\text{gård}} > 0$. We have to assume that $b(\cdot,\cdot)$ is definite on the continuous level, i.e., for all $v \in H^1_0(\Omega)$, it holds that

$$b(v,w) = 0 \quad \text{for all } w \in H^1_0(\Omega) \quad \Longrightarrow \quad v = 0, \tag{6.6a}$$

$$b(v_\infty, w_\infty) = 0 \quad \text{for all } w_\infty \in \mathcal{S}^p_0(\mathcal{T}_\infty) \quad \Longrightarrow \quad v_\infty = 0. \tag{6.6b}$$

This together with Fredholm's alternative already guarantees the unique solvability of (2.2) and (2.5) with test and ansatz space $\mathcal{S}^p_0(\mathcal{T}_\infty)$ instead of $\mathcal{S}^p_0(\mathcal{T}_\ell)$.

Remark 6.1. Usually, the conditions (6.6) are guaranteed under the assumption that the mesh-size of the initial mesh $\mathcal{T}_0$ is sufficiently small and that the solution $w \in H^1_0(\Omega)$ of the dual problem

$$b(v,w) = \int_\Omega fv\,dx \quad \text{for all } v \in H^1_0(\Omega)$$

satisfies some regularity estimate

$$\|w\|_{H^{1+s}(\Omega)} \lesssim \|f\|_{L^2(\Omega)} \quad \text{for some } s > 0;$$

see, e.g., [10, Theorem 5.7.6].

Now, we may apply [27, Theorem 4.2.9] to obtain the following result.

LEMMA 6.2. *There exists an index $\ell_0 \in \mathbb{N}$ such that for all $\ell \geq \ell_0$ the discrete formulation* (2.5) *is uniquely solvable, and it holds that*

$$\|\nabla(u_\infty - U_\ell)\|_{L^2(\Omega)} \leq C_{\text{Céa}} \min_{V_\ell \in \mathcal{S}^p_0(\mathcal{T}_\ell)} \|\nabla(u_\infty - V_\ell)\|_{L^2(\Omega)}, \tag{6.7}$$

where $u_\infty \in \mathcal{S}^p_0(\mathcal{T}_\infty)$ denotes the unique solution of (2.5) *with $\mathcal{S}^p_0(\mathcal{T}_\infty)$ instead of $\mathcal{S}^p_0(\mathcal{T}_\ell)$.*

Proof. Since (6.5) states that $b(u,v) + C_{\text{gård}}\langle u\,,\,v\rangle_{L^2(\Omega)}$ is elliptic and $\langle\cdot\,,\,\cdot\rangle_{L^2(\Omega)}$ is a compact perturbation, we apply [27, Theorem 4.2.9] on the Hilbert space $\mathcal{S}^p_0(\mathcal{T}_\infty)$ and the dense sequence of subspaces $\mathcal{S}^p_0(\mathcal{T}_\ell)$ for $\ell \to \infty$. ∎

The above lemma allows us to prove a priori convergence from Lemma 3.2 and consequently convergence $U_\ell \to u$ in $H^1_0(\Omega)$ as well as $u \in \mathcal{S}^p_0(\mathcal{T}_\infty)$. Moreover, Lemma 3.5 still holds true, since we assumed definiteness of $b(\cdot,\cdot)$ on $\mathcal{S}^p_0(\mathcal{T}_\infty)$ in (6.6b).

LEMMA 6.3. *There exists an index $\ell_1 \in \mathbb{N}$ such that for all $\ell \geq \ell_1$ there holds*

$$\|\nabla(u - U_\ell)\|_{L^2(\Omega)} \leq C_{\text{norm}}|||u - U_\ell||| \quad \textit{and}$$
$$\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)} \leq C_{\text{norm}}|||U_{\ell+1} - U_\ell|||.$$

Proof. With (6.5) and $b(\cdot,\cdot) = |||\cdot|||^2$, we may estimate

$$\begin{aligned}\|\nabla(u - U_\ell)\|^2_{L^2(\Omega)} &\lesssim |||u - U_\ell|||^2 + \|u - U_\ell\|^2_{L^2(\Omega)}\\ &= |||u - U_\ell|||^2 + \|e_\ell\|^2_{L^2(\Omega)}\|\nabla(u - U_\ell)\|^2_{L^2(\Omega)}.\end{aligned}$$

Lemma 3.5 shows weak convergence $e_\ell \rightharpoonup 0$ in $H^1_0(\Omega)$. The Rellich compactness theorem thus implies strong convergence $e_\ell \to 0$ in $L^2(\Omega)$. Therefore, there exists an index $\ell_1 \in \mathbb{N}$ such that there holds

$$\|\nabla(u - U_\ell)\|^2_{L^2(\Omega)} \lesssim |||u - U_\ell|||^2 \quad \text{for all } \ell \geq \ell_1.$$

The statement for $U_{\ell+1} - U_\ell$ follows analogously. ∎

Lemma 3.5 together with Lemma 6.3 allows to prove the quasi-Galerkin orthogonality of Proposition 3.6 and consequently also the R-linear convergence of Theorem 4.1. Therefore, all the results from section 5 hold and, in particular, we obtain the optimality result of Theorem 5.3.

6.5. Nonlinear operators $\mathcal{L}$. We consider following *nonlinear* operator

$$\mathcal{L}u(x) := -\mathrm{div}\boldsymbol{A}(x, \nabla u(x)) + g(x, u(x), \nabla u(x))$$

for functions $\boldsymbol{A} : \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$. We assume that $\boldsymbol{A}(\cdot, \nabla u), g(\cdot, u, \nabla u) \in L^2(\Omega)$ for all $u \in H^1_0(\Omega)$. Then, the weak formulation of (1.1) reads as follows: Find $u \in H^1_0(\Omega)$ such that

$$\langle \mathcal{L}u\,,\, v\rangle = \int_\Omega \boldsymbol{A}(x, \nabla u(x)) \cdot \nabla v(x) + g(x, u(x), \nabla u(x))v(x)\, dx = \int_\Omega f v\, dx \tag{6.8}$$

for all $v \in H^1_0(\Omega)$. We define two auxiliary operators $\mathcal{A}, \mathcal{K} :\ H^1_0(\Omega) \to H^{-1}(\Omega)$ as

$$\mathcal{A}v := -\mathrm{div}\boldsymbol{A}(\cdot, \nabla v) \quad \text{and} \quad \mathcal{K}v := g(\cdot, v, \nabla v) \quad \text{for all } v \in H^1_0(\Omega).$$

We formally define the residual error estimator for a mesh $\mathcal{T}_\ell$

$$\eta_\ell^2 := \sum_{T\in\mathcal{T}_\ell} \big(|T|^{2/d}\|\mathcal{L}|_T U_\ell - f\|^2_{L^2(T)} + |T|^{1/d}\|[\boldsymbol{A}(\cdot, \nabla U_\ell)\cdot n]\|^2_{L^2(\partial T\cap\Omega)}\big). \tag{6.9}$$

The solvability and uniqueness of (6.8) as well as the regularity assumptions needed such that (6.9) is well-defined are part of the subsequent sections.

6.5.1. Regularity assumptions. We consider the frame of strongly monotone operators and require the following regularity assumptions on $\mathcal{L}$:

$$\|\mathcal{A}\nabla w - \mathcal{A}\nabla v\|_{H^{-1}(\Omega)} \le C_{\mathrm{lip}}\|\nabla(w - v)\|_{L^2(\Omega)}, \tag{6.10a}$$

$$\|\mathcal{K}w - \mathcal{K}v\|_{L^2(\Omega)} \le C_{\mathrm{lip}}\|\nabla(w - v)\|_{L^2(\Omega)} \tag{6.10b}$$

for all $w, v \in H^1_0(\Omega)$ and some constant $C_{\mathrm{lip}} > 0$ as well as

$$\langle \mathcal{L}w - \mathcal{L}v\,,\, w - v\rangle \ge C_{\mathrm{mon}}\|\nabla(w - v)\|^2_{L^2(\Omega)} \tag{6.11}$$

for all $w, v \in H^1_0(\Omega)$ and some constant $C_{\mathrm{mon}} > 0$. These assumptions, in particular, allow us to apply the main theorem on strongly monotone operators [32, Theorem 26.A] and to obtain the unique solvability of (6.8) as well as of (2.5). Additionally, (6.10)–(6.11) guarantee that the norms of the residual and the error are equivalent, i.e.,

$$\|\mathcal{L}u - \mathcal{L}U_\ell\|_{H^{-1}(\Omega)} \simeq \|\nabla(u - U_\ell)\|_{L^2(\Omega)} \quad \text{for all } \ell \in \mathbb{N}. \tag{6.12}$$

We also obtain the Céa lemma (2.6) with the constant $2C_{\mathrm{lip}}/C_{\mathrm{mon}}$.

Moreover, we require that (6.9) is well-defined and that there holds the estimator reduction (3.2) from Lemma 3.1. For possible nonlinearities $\boldsymbol{A}$ which allow for (3.2), we refer to Lemma 6.6 below.

We assume that $\mathcal{L} :\ H^1_0(\Omega) \to H^{-1}(\Omega)$ as well as $\mathcal{A} :\ H^1_0(\Omega) \to H^{-1}(\Omega)$ are twice Fréchet differentiable, i.e., there exist

$$\begin{aligned} D\mathcal{L}, D\mathcal{A} &: H^1_0(\Omega) \to L(H^1_0(\Omega), H^{-1}(\Omega)),\\ D^2\mathcal{L}, D^2\mathcal{A} &: H^1_0(\Omega) \to L\big(H^1_0(\Omega), L(H^1_0(\Omega), H^{-1}(\Omega))\big). \end{aligned} \tag{6.13}$$

The second derivative should be bounded locally around the solution u of (6.8), i.e., there exists $\varepsilon_{\ell oc} > 0$ with

$$C_{\ell oc} := \sup_{\|\nabla(u-v)\|_{L^2(\Omega)} < \varepsilon_{\ell oc}} \Big(\|D^2\mathcal{L}(v)\|_{L\left(H_0^1(\Omega), L(H_0^1(\Omega), H^{-1}(\Omega))\right)} + \|D^2\mathcal{A}(v)\|_{L\left(H_0^1(\Omega), L(H_0^1(\Omega), H^{-1}(\Omega))\right)} \Big) < \infty. \tag{6.14}$$

Finally, we assume that $D\mathcal{A}(v) : H_0^1(\Omega) \to H^{-1}(\Omega)$ is symmetric for all $v \in H_0^1(\Omega)$, i.e., for all $w_1, w_2 \in H_0^1(\Omega)$ it holds that

$$\langle D\mathcal{A}(v)(w_1)\,,\, w_2\rangle = \langle D\mathcal{A}(v)(w_2)\,,\, w_1\rangle. \tag{6.15}$$

Remark 6.4. Note that if $\boldsymbol{A} : \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ are twice differentiable, and if $J_y\boldsymbol{A}(x,y) \in \mathbb{R}^{d\times d}$ additionally is a symmetric matrix, then $\mathcal{L}$ and $\mathcal{A}$ satisfy (6.13) as well as (6.14). Moreover, $D\mathcal{A}(v)$ is symmetric for all $v \in H_0^1(\Omega)$, since there holds for $w \in H_0^1(\Omega)$

$$D\mathcal{A}(v)(w) = \operatorname{div}_x\Big(\big(J_y\boldsymbol{A}(x, \nabla v(x))\big)\big(\nabla_x w(x)\big)\Big),$$

where $J_y\boldsymbol{A}(x,y)$ denotes the Jacobian of $\boldsymbol{A}$ with respect to y.

Example 6.5. We stress that the symmetry assumption (6.15) posed on $D\mathcal{A}$ covers in particular the operator class from [16], where

$$\boldsymbol{A}(x,y) = \alpha(x, |y|^2)y$$

for some function $\alpha : \Omega\times\mathbb{R} \to \mathbb{R}$ with continuous derivative $t \mapsto \partial_t\alpha(x,t)$. In contrast to [16], where $\alpha(x,\cdot) \in C^1(\mathbb{R})$ is sufficient, the analysis here covers a wider class of operators, but requires $\alpha(x,\cdot) \in C^2(\mathbb{R})$ to guarantee (6.14).

LEMMA 6.6. *Sufficient regularity assumptions in addition to* (6.10b) *and* (6.11) *to guarantee that the error estimator* (6.9) *is well-defined and satisfies the estimator reduction* (3.2) *are, for instance, either of the following conditions* (i) *and* (ii)*:*

(i) $\boldsymbol{A}(\cdot,\cdot) : \Omega\times\mathbb{R}^d \to \mathbb{R}^d$ *is Lipschitz continuous and there exists a constant* $C_3 > 0$ *such that for all* $\ell \in \mathbb{N}$ *and all* $V_\ell, W_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell)$ *there holds* $\operatorname{div}\boldsymbol{A}(\cdot, V_\ell(\cdot)) \in L^2(\Omega)$ *as well as*

$$\|\operatorname{div}|_T\big(\boldsymbol{A}(\cdot, V_\ell(\cdot)) - \boldsymbol{A}(\cdot, W_\ell(\cdot))\big)\|_{L^2(T)} \le C_3\|V_\ell - W_\ell\|_{H^2(T)} \quad \textit{for all } T \in \mathcal{T}_\ell. \tag{6.16}$$

(ii) *There holds* $p = 1$ *(lowest-order case) as well as*

$$\boldsymbol{A}(x,y) = \boldsymbol{A}(y) \quad \textit{for all } x \in \Omega,\ y \in \mathbb{R}^d,$$

and additionally $\boldsymbol{A}(\cdot) : \mathbb{R}^d \to \mathbb{R}^d$ *is Lipschitz continuous.*

Proof. The jump terms in (6.9) are well-defined in both cases (i) and (ii) since $\boldsymbol{A}(\cdot, \nabla U_\ell(\cdot))$ is a piecewise Lipschitz continuous function. Moreover, this shows that $\operatorname{div}\boldsymbol{A}(\cdot, \nabla U_\ell(\cdot)) \in L^\infty(T) \subset L^2(T)$ for all $T \in \mathcal{T}_\ell$. Therefore, (6.9) is well-defined.

Given $T_+, T_- \in \mathcal{T}_\ell$ as well as $W_\ell, V_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell)$, the Lipschitz continuity also proves the following pointwise estimate for all $x \in T_+ \cap T_-$:

$$\begin{aligned}
&|[(\boldsymbol{A}(x, \nabla W_\ell(x)) - \boldsymbol{A}(x, \nabla V_\ell(x))) \cdot n]| \\
&\quad \lesssim \Big| \big(\boldsymbol{A}(x, (\nabla W_\ell)|_{T_+}(x)) - \boldsymbol{A}(x, (\nabla V_\ell)|_{T_+}(x))\big) \cdot n|_{T_+} \\
&\qquad + \big(\boldsymbol{A}(x, (\nabla W_\ell)|_{T_-}(x)) - \boldsymbol{A}(x, (\nabla V_\ell)|_{T_-}(x))\big) \cdot n|_{T_-} \Big| \\
&\quad \lesssim \Big| (\nabla W_\ell)|_{T_+}(x) - (\nabla V_\ell(x))|_{T_+} \Big| + \Big| (\nabla W_\ell)|_{T_-}(x) - (\nabla V_\ell)|_{T_-}(x) \Big|.
\end{aligned}$$

Combining the estimate above with the trace inequality for polynomials, we obtain

$$|T_+|^{1/d} \| [(\boldsymbol{A}(\cdot, \nabla W_\ell) - \boldsymbol{A}(\cdot, \nabla V_\ell)) \cdot n] \|_{L^2(T_+ \cap T_-)}^2 \lesssim \| \nabla (W_\ell - V_\ell) \|_{L^2(T_+ \cup T_-)}^2. \tag{6.17}$$

This hidden constant depends only on the polynomial degree $p \in \mathbb{N}$ as well as the Lipschitz continuity of $\boldsymbol{A}(\cdot, \cdot)$ and the γ-shape regularity of $\mathcal{T}_\ell$. It remains to prove a similar estimate for the volume residual in (6.9), i.e.,

$$|T|^{2/d} \| \mathcal{L}|_T W_\ell - \mathcal{L}|_T V_\ell \|_{L^2(T)}^2 \lesssim \| \nabla (W_\ell - V_\ell) \|_{L^2(T)}^2 \quad \text{for all } T \in \mathcal{T}_\ell. \tag{6.18}$$

In the case of (i), this follows immediately from the combination of (6.16) and (6.10b) together with a standard inverse estimate. In the case of (ii), we observe that ∇U_ℓ is piecewise constant. Therefore, $\boldsymbol{A}(\nabla U_\ell)$ is also piecewise constant and hence $\mathcal{A}(\nabla U) = \operatorname{div} \boldsymbol{A}(\nabla U(\cdot)) = 0$. Thus, $\mathcal{L}|_T V_\ell = (\mathcal{K} V_\ell)|_T$, and it suffices to apply (6.10b) to prove (6.18). With the estimates (6.17)–(6.18), the proof of Lemma 3.1 still holds true with the obvious modifications. This concludes the proof. ∎

6.5.2. Auxiliary results. This section provides some technical lemmata, which are used to transfer the results from the linear case to the present nonlinear case.

LEMMA 6.7. *The residual error estimator satisfies reliability* (2.7) *as well as discrete reliability* (5.6). *Moreover, there holds convergence*

$$\| \nabla (u - U_\ell) \|_{L^2(\Omega)} \to 0 \quad \textit{as } \ell \to \infty. \tag{6.19}$$

Proof. The residual error estimator η_ℓ is well-defined by assumption in section 6.5.1. With the equivalence (6.12), the standard arguments apply to prove reliability (2.7), and also the proof of discrete reliability (5.6) follows analogously to [11]. The estimator reduction holds by assumption in section 6.5.1 and therefore Proposition 3.3 holds true and proves (6.19). ∎

LEMMA 6.8. *The operator* $(D\mathcal{L})|_{\mathcal{S}_0^p(\mathcal{T}_\infty)} u : \mathcal{S}_0^p(\mathcal{T}_\infty) \to \mathcal{S}_0^p(\mathcal{T}_\infty)^\star$ *is injective.*

Proof. With (6.11) and the definition of the Fréchet derivative, there holds for all $v \in \mathcal{S}_0^p(\mathcal{T}_\infty)$ with $\|\nabla v\|_{L^2(\Omega)} = 1$

$$\begin{aligned}
\langle ((D\mathcal{L})|_{\mathcal{S}_0^p(\mathcal{T}_\infty)} u)(v) \,,\, v \rangle &= \lim_{\delta \to 0} \delta^{-2} \langle \mathcal{L}(u + \delta v) - \mathcal{L} u \,,\, u + \delta v - u \rangle \\
&\gtrsim \lim_{\delta \to 0} \delta^{-2} \| \nabla (u + \delta v - u) \|_{L^2(\Omega)}^2 = 1.
\end{aligned}$$

Hence, we have $((D\mathcal{L})|_{\mathcal{S}_0^p(\mathcal{T}_\infty)} u)(v) \neq 0$ in $\mathcal{S}_0^p(\mathcal{T}_\infty)^\star$ for all $v \in \mathcal{S}_0^p(\mathcal{T}_\infty) \setminus \{0\}$. This concludes the proof. ∎

LEMMA 6.9 (Taylor). *For all* $v, w \in H_0^1(\Omega)$ *with* $\|\nabla(u - v)\|_{L^2(\Omega)} + \|\nabla(u - w)\|_{L^2(\Omega)} \leq \varepsilon_{\ell oc}$, *there holds*

$$\| \mathcal{L} w - \mathcal{L} v - D\mathcal{L}(w)(w - v) \|_{H^{-1}(\Omega)} \leq C_{\ell oc} \| \nabla (w - v) \|_{L^2(\Omega)}^2, \tag{6.20a}$$

$$\| \mathcal{A} w - \mathcal{A} v - D\mathcal{A}(w)(w - v) \|_{H^{-1}(\Omega)} \leq C_{\ell oc} \| \nabla (w - v) \|_{L^2(\Omega)}^2. \tag{6.20b}$$

Proof. The local boundedness (6.14) together with [13, Theorem 6.5] applied to the operators $\mathcal{L}$ and $\mathcal{A}$ proves the statement. ∎

6.5.3. Quasi-orthogonality. Following the steps of section 3.2, we derive a similar result for the present, nonlinear case.

LEMMA 6.10. *The sequence* $(e_\ell)_{\ell\in\mathbb{N}}$ *defined by*

$$e_\ell := \begin{cases} \frac{u-U_\ell}{\|\nabla(u-U_\ell)\|_{L^2(\Omega)}} & \text{for } u \neq U_\ell, \\ 0 & \text{else} \end{cases}$$

converges to zero, weakly in $H^1_0(\Omega)$ *in the sense of* (3.6).

Proof. With Galerkin orthogonality and the convention $\infty \cdot 0 = 0$, we obtain

$$\lim_{\ell\to\infty} \frac{\langle \mathcal{L}u - \mathcal{L}U_\ell\,,\, V_k\rangle}{\|\nabla(u-U_\ell)\|_{L^2(\Omega)}} = 0 \quad \text{for all } V_k \in \mathcal{S}^p_0(\mathcal{T}_k) \text{ and } k\in\mathbb{N}.$$

By continuity of the duality brackets, this results in convergence for all $v \in \mathcal{S}^p_0(\mathcal{T}_\infty)$,

$$\frac{\langle \mathcal{L}u - \mathcal{L}U_\ell\,,\, v\rangle}{\|\nabla(u-U_\ell)\|_{L^2(\Omega)}} \to 0 \quad \text{as } \ell\to\infty.$$

By use of (6.20a), we observe for all $v \in \mathcal{S}^p_0(\mathcal{T}_\infty)$

$$\frac{|\langle \mathcal{L}u - \mathcal{L}U_\ell\,,\, v\rangle|}{\|\nabla(u-U_\ell)\|_{L^2(\Omega)}} \geq \frac{|\langle (D\mathcal{L}u)(u-U_\ell)\,,\, v\rangle|}{\|\nabla(u-U_\ell)\|_{L^2(\Omega)}} - C_{\ell oc}\|\nabla(u-U_\ell)\|_{L^2(\Omega)}\|\nabla v\|_{L^2(\Omega)}.$$

With convergence $U_\ell \to u$ in $H^1_0(\Omega)$ from (6.19), this implies immediately

$$\frac{|\langle u - U_\ell\,,\, ((D\mathcal{L})|_{\mathcal{S}^p_0(\mathcal{T}_\infty)}u)^\star v\rangle|}{\|\nabla(u-U_\ell)\|_{L^2(\Omega)}} \to 0 \quad \text{as } \ell\to\infty \quad \text{for all } v\in\mathcal{S}^p_0(\mathcal{T}_\infty). \tag{6.21}$$

According to Lemma 6.8, $(D\mathcal{L})|_{\mathcal{S}^p_0(\mathcal{T}_\infty)}u$ is injective. Therefore, its adjoint operator $((D\mathcal{L})|_{\mathcal{S}^p_0(\mathcal{T}_\infty)}u)^\star$ is surjective onto $\mathcal{S}^p_0(\mathcal{T}_\infty)^\star$. Hence, (6.21) is equivalent to $e_\ell \rightharpoonup 0$ as $\ell\to\infty$. This concludes the proof. ∎

To abbreviate notation, we define the quasi metric

$$\mathrm{d\!l}(w,v)^2 := \langle \mathcal{L}w - \mathcal{L}v\,,\, w-v\rangle \quad \text{for all } w,v\in H^1_0(\Omega).$$

Note that due to (6.10)–(6.11), there holds

$$C_{\rm norm}^{-1}\|\nabla(w-v)\|_{L^2(\Omega)} \leq \mathrm{d\!l}(w,v) \leq C_{\rm norm}\|\nabla(w-v)\|_{L^2(\Omega)} \quad \text{for all } w,v\in H^1_0(\Omega) \tag{6.22}$$

with $C_{\rm norm} = \max\{2C_{\rm lip}, C_{\rm mon}^{-1}\} > 0$.

PROPOSITION 6.11. *For any* $\varepsilon > 0$*, there exists* $\ell_0\in\mathbb{N}$ *such that*

$$\mathrm{d\!l}(U_{\ell+1},U_\ell)^2 \leq \frac{1}{1-\varepsilon}\,\mathrm{d\!l}(u,U_\ell)^2 - \mathrm{d\!l}(u,U_{\ell+1})^2 \tag{6.23}$$

for all $\ell \geq \ell_0$.

Proof. Due to convergence $U_\ell\to u$ in $H^1_0(\Omega)$ (6.19), there exists $\ell_1\in\mathbb{N}$ such that for all $\ell\geq\ell_1$ we may apply (6.20b) to obtain

$$\begin{aligned}|\langle \mathcal{A}U_{\ell+1} - \mathcal{A}U_\ell\,,\, u-U_{\ell+1}\rangle| \leq{}& |\langle D\mathcal{A}(U_{\ell+1})(U_{\ell+1}-U_\ell)\,,\, u-U_{\ell+1}\rangle| \\ &+ C_{\ell oc}\|\nabla(U_{\ell+1}-U_\ell)\|^2_{L^2(\Omega)}\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}.\end{aligned}$$

Using the symmetry of $D\mathcal{A}(U_{\ell+1})$, we conclude

$$\begin{aligned}|\langle \mathcal{A}U_{\ell+1} - \mathcal{A}U_\ell\,,\, u - U_{\ell+1}\rangle| &\le |\langle D\mathcal{A}(U_{\ell+1})(u - U_{\ell+1})\,,\, U_{\ell+1} - U_\ell\rangle| \\ &\quad + C_{\ell oc}\|\nabla(U_{\ell+1} - U_\ell)\|^2_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)} \\ &\le |\langle \mathcal{A}u - \mathcal{A}U_{\ell+1}\,,\, U_{\ell+1} - U_\ell\rangle| \\ &\quad + C_{\ell oc}\|\nabla(U_{\ell+1} - U_\ell)\|^2_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)} \\ &\quad + C_{\ell oc}\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|^2_{L^2(\Omega)}.\end{aligned}$$

Analogously to the estimate above, we obtain

$$\begin{aligned}|\langle \mathcal{A}U_{\ell+1} - \mathcal{A}U_\ell\,,\, u - U_{\ell+1}\rangle| &\ge |\langle \mathcal{A}u - \mathcal{A}U_{\ell+1}\,,\, U_{\ell+1} - U_\ell\rangle| \\ &\quad - C_{\ell oc}\|\nabla(U_{\ell+1} - U_\ell)\|^2_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)} \\ &\quad - C_{\ell oc}\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|^2_{L^2(\Omega)}.\end{aligned}$$

For any $\delta > 0$, we may thus use convergence $U_\ell \to u$ as $\ell \to \infty$ to find an index $\ell_0 \in \mathbb{N}$ such that

$$\begin{aligned}&\big||\langle \mathcal{A}U_{\ell+1} - \mathcal{A}U_\ell\,,\, u - U_{\ell+1}\rangle| - |\langle \mathcal{A}u - \mathcal{A}U_{\ell+1}\,,\, U_{\ell+1} - U_\ell\rangle|\big| \\ &\quad \le \delta\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)}\end{aligned}$$

for all $\ell \ge \ell_0$. Since e_ℓ converges to zero weakly in $H^1_0(\Omega)$, we have strong convergence $e_\ell \to 0$ as $\ell \to \infty$ in $L^2(\Omega)$. This together with Lipschitz continuity (6.10b) allows us to estimate

$$|\langle \mathcal{K}U_{\ell+1} - \mathcal{K}U_\ell\,,\, u - U_{\ell+1}\rangle| \lesssim \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}\|e_{\ell+1}\|_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)}$$

and hence

$$|\langle \mathcal{K}U_{\ell+1} - \mathcal{K}U_\ell\,,\, u - U_{\ell+1}\rangle| \le \delta\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)}$$

for all $\ell \ge \ell_1$. The adjoint term follows analogously, since

$$|\langle \mathcal{K}u - \mathcal{K}U_{\ell+1}\,,\, U_{\ell+1} - U_\ell\rangle| \le |\langle \mathcal{K}u - \mathcal{K}U_{\ell+1}\,,\, U_{\ell+1} - u\rangle| + |\langle \mathcal{K}u - \mathcal{K}U_{\ell+1}\,,\, u - U_\ell\rangle|.$$

So far, we end up with

$$\begin{aligned}&|\langle \mathcal{K}U_{\ell+1} - \mathcal{K}U_\ell\,,\, u - U_{\ell+1}\rangle| + |\langle \mathcal{K}u - \mathcal{K}U_{\ell+1}\,,\, U_{\ell+1} - U_\ell\rangle| \\ &\quad \le \delta\big(\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)} \\ &\qquad + \|\nabla(u - U_{\ell+1})\|^2_{L^2(\Omega)} \\ &\qquad + \|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)}\|\nabla(u - U_\ell)\|_{L^2(\Omega)}\big) \\ &\quad \le \delta/2\|\nabla(U_{\ell+1} - U_\ell)\|^2_{L^2(\Omega)} + 2\delta\|\nabla(u - U_{\ell+1})\|^2_{L^2(\Omega)} \\ &\qquad + \delta/2\|\nabla(u - U_\ell)\|^2_{L^2(\Omega)}\end{aligned}$$

by use of Young's inequality. Putting everything together, we obtain

$$\begin{aligned}&|\langle(\mathcal{A} + \mathcal{K})U_{\ell+1} - (\mathcal{A} + \mathcal{K})U_\ell\,,\, u - U_{\ell+1}\rangle| \\ &\quad \le |\langle \mathcal{A}u - \mathcal{A}U_{\ell+1}\,,\, U_{\ell+1} - U_\ell\rangle| + \delta\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)} \\ &\qquad + |\langle \mathcal{K}U_{\ell+1} - \mathcal{K}U_\ell\,,\, u - U_{\ell+1}\rangle| \\ &\quad \le |\langle(\mathcal{A} + \mathcal{K})u - (\mathcal{A} + \mathcal{K})U_{\ell+1}\,,\, U_{\ell+1} - U_\ell\rangle| \\ &\qquad + \delta\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}\|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)} \\ &\qquad + |\langle \mathcal{K}U_{\ell+1} - \mathcal{K}U_\ell\,,\, u - U_{\ell+1}\rangle| + |\langle \mathcal{K}u - \mathcal{K}U_{\ell+1}\,,\, U_{\ell+1} - U_\ell\rangle| \\ &\quad \le 3\delta\big(\|\nabla(U_{\ell+1} - U_\ell)\|^2_{L^2(\Omega)} + \|\nabla(u - U_{\ell+1})\|^2_{L^2(\Omega)} + \|\nabla(u - U_\ell)\|^2_{L^2(\Omega)}\big),\end{aligned}$$

where we use Galerkin orthogonality $\langle (\mathcal{A}+\mathcal{K})u - (\mathcal{A}+\mathcal{K})U_{\ell+1}\,,\, U_{\ell+1}-U_\ell\rangle = 0$ to obtain the last estimate. With that at hand, we obtain similarly to (3.9)

$$\begin{aligned}\mathrm{d\!l}(U_{\ell+1},U_\ell)^2 &\le \mathrm{d\!l}(u,U_\ell)^2 - \mathrm{d\!l}(u,U_{\ell+1})^2 + |\langle (\mathcal{A}+\mathcal{K})U_{\ell+1} - (\mathcal{A}+\mathcal{K})U_\ell\,,\, u-U_{\ell+1}\rangle| \\ &\le \mathrm{d\!l}(u,U_\ell)^2 - \mathrm{d\!l}(u,U_{\ell+1})^2 + 3\delta\big(\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2 \\ &\qquad + \|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}^2 + \|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2\big).\end{aligned}$$

With the equivalence (6.22), we conclude

$$(1-3C_{\rm norm}\delta)\mathrm{d\!l}(U_{\ell+1},U_\ell)^2 \le (1+3C_{\rm norm}\delta)\mathrm{d\!l}(u,U_\ell)^2 - (1-3C_{\rm norm}\delta)\mathrm{d\!l}(u,U_{\ell+1})^2$$

for all $\ell \ge \ell_0$. Finally, we choose $\delta > 0$ sufficiently small such that $(1+3C_{\rm norm}\delta)/(1-3C_{\rm norm}\delta) \le 1/(1-\varepsilon)$ and conclude the proof. ∎

Together with the estimator reduction (3.2) which holds by assumption in section 6.5.1, the quasi-Galerkin orthogonality (6.23) of Proposition 6.11 allows us to prove the R-linear convergence of Theorem 4.1 if one exchanges $|\!|\!| u-U_{\ell+1} |\!|\!|$ and $|\!|\!| U_{\ell+1}-U_\ell |\!|\!|$ with $\mathrm{d\!l}(u,U_{\ell+1})$ and $\mathrm{d\!l}(U_{\ell+1},U_\ell)$, respectively. Therefore, all the results from section 5 hold (cf. the remarks after Theorem 4.1 and the proof of Theorem 5.3) and, in particular, we obtain the optimality result of Theorem 5.3.

REFERENCES

[1] M. Ainsworth and J. T. Oden, *A posteriori error estimation in finite element analysis*, Pure Appl. Math. (N.Y.), Wiley-Interscience, New York, 2000.

[2] M. Aurada, M. Feischl, J. Kemetmüller, M. Page, and D. Praetorius, *Each $H^{1/2}$-stable projection yields convergence and quasi-optimality of adaptive FEM with inhomogeneous Dirichlet data in $\mathbb{R}^d$*, M2AN Math. Model. Numer. Anal., 47 (2013), pp. 1207–1235.

[3] M. Aurada, S. Ferraz-Leite, and D. Praetorius, *Estimator reduction and convergence of adaptive BEM*, Appl. Numer. Math., 62 (2012), pp. 787–801.

[4] I. Babuska and M. Vogelius, *Feedback and adaptive finite element solution of one-dimensional boundary value problems*, Numer. Math., 44 (1984), pp. 75–102.

[5] S. Bartels, C. Carstensen, and G. Dolzmann, *Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis*, Numer. Math., 99 (2004), pp. 1–24.

[6] L. Belenki, L. Diening, and C. Kreuzer, *Optimality of an adaptive finite element method for the p-Laplacian equation*, IMA J. Numer. Anal., 32 (2012), pp. 484–510.

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

[8] P. Binev, W. Dahmen, R. DeVore, and P. Petrushev, *Approximation classes for adaptive methods*, Serdica Math. J., 28 (2002), pp. 391–416.

[9] A. Bonito and R. H. Nochetto, *Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method*, SIAM J. Numer. Anal., 48 (2010), pp. 734–771.

[10] D. Braess and J. Schöberl, *Equilibrated residual error estimator for edge elements*, Math. Comp., 77 (2008), pp. 651–672.

[11] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal., 46 (2008), pp. 2524–2550.

[12] J. M. Cascon and R. H. Nochetto, *Quasioptimal cardinality of AFEM driven by nonresidual estimators*, IMA J. Numeri. Anal., 32 (2012), pp. 1–29.

[13] R. F. Curtain and A. J. Pritchard, *Functional Analysis in Modern Applied Mathematics*, Mathematics in Science and Engineering 132, Academic Press, London, 1977.

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

[15] M. Feischl, M. Page, and D. Praetorius, *Convergence of adaptive FEM for elliptic obstacle problems*, Int. J. Numer. Anal. Model., 11 (2014), pp. 229–253.

[16] E. M. Garau, P. Morin, and C. Zuppa, *Quasi-optimal convergence rate of an AFEM for quasi-linear problems of monotone type*, Numer. Math. Theory Methods Appl., 5 (2012), pp. 131–156.

[17] F. Gaspoz and P. Morin, *Approximation classes for adaptive higher order finite element approximation*, DOI: http://dx.doi.org/10.1090/S0025-5718-2013-02777-9.
[18] M. Holst, G. Tsogtgerel, and Y. Zhu, *Local Convergence of Adaptive Methods for Nonlinear Partial Differential Equations*, arXiv:1001.1382, 2010.
[19] M. Karkulik, D. Pavlicek, and D. Praetorius, *On 2D newest vertex bisection: Optimality of mesh-closure and H^1-stability of L_2-projection*, Constr. Approx., 38 (2013), pp. 213–234.
[20] C. Kreuzer and K. G. Siebert, *Decay rates of adaptive finite elements with Dörfler marking*, Numer. Math., 117 (2011), pp. 679–716.
[21] K. Mekchay and R. H. Nochetto, *Convergence of adaptive finite element methods for general second order linear elliptic PDEs*, SIAM J. Numer. Anal., 43 (2005), pp. 1803–1827.
[22] M. S. Mommer and R. Stevenson, *A goal-oriented adaptive finite element method with convergence rates*, SIAM J. Numer. Anal., 47 (2009), pp. 861–886.
[23] P. Morin, R. H. Nochetto, and K. G. Siebert, *Data oscillation and convergence of adaptive FEM*, SIAM J. Numer. Anal., 38 (2000), pp. 466–488.
[24] P. Morin, K. G. Siebert, and A. Veeser, *A basic convergence result for conforming adaptive finite elements*, Math. Models Methods Appl. Sci., 18 (2008), pp. 707–737.
[25] W. Rudin, *Functional Analysis*, 2nd ed., Internat. Ser. in Pure and Appl. Math., McGraw-Hill, New York, 1991.
[26] R. Sacchi and A. Veeser, *Locally efficient and reliable a posteriori error estimators for Dirichlet problems*, Math. Models Methods Appl. Sci., 16 (2006), pp. 319–346.
[27] S. A. Sauter and C. Schwab, *Boundary Element Methods*, Springer Ser. in Comput. Math. 39, Springer-Verlag, Berlin, 2011.
[28] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[29] R. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., 7 (2007), pp. 245–269.
[30] R. Stevenson, *The completion of locally refined simplicial partitions created by bisection*, Math. Comp., 77 (2008), pp. 227–241.
[31] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, WileyTeubner, Chichester, 1996.
[32] E. Zeidler, *Nonlinear Functional Analysis and Its Applications. II/B*: Nonlinear Monotone Operators, Springer-Verlag, New York, 1990.