Robust energy a posteriori estimates for nonlinear elliptic problems

André Harnist, Koondanibha Mitra, Ari Rappaport, Martin Vohralík

Robust energy a posteriori estimates for nonlinear elliptic problems

André Harnist[†§1], Koondanibha Mitra[‡2], Ari Rappaport[†§3], and Martin Vohralík[†§4]

May 26, 2023

Abstract

In this paper, we design a posteriori estimates for finite element approximations of nonlinear elliptic problems satisfying strong-monotonicity and Lipschitz-continuity properties. These estimates include, and build on, any iterative linearization method that satisfies a few clearly identified assumptions; this encompasses the Picard, Newton, and Zarantonello linearizations. The estimates give a guaranteed upper bound on an augmented energy difference (reliability with constant one), as well as a lower bound (efficiency up to a generic constant). We prove that for the Zarantonello linearization, this generic constant only depends on the space dimension, the mesh shape regularity, and possibly the approximation polynomial degree in four or more space dimensions, making the estimates robust with respect to the strength of the nonlinearity. For the other linearizations, there is only a computable dependence on the local variation of the linearization operators. We also derive similar estimates for the energy difference. Numerical experiments illustrate and validate the theoretical results, for both smooth and singular solutions.

Key words: nonlinear elliptic problem, finite elements, iterative linearization, energy difference, a posteriori error estimate, robustness, equilibrated flux reconstruction

Contents

[1] andre.harnist@inria.fr, https://andreharnist.fr
[2] koondanibha.mitra@uhasselt.be, https://www.koondi.net
[3] ari.rappaport@inria.fr
[4] martin.vohralik@inria.fr, https://who.rocq.inria.fr/Martin.Vohralik
[†] Inria, 2 rue Simone Iff, 75589 Paris, France
[§] CERMICS, Ecole des Ponts, 77455 Marne-la-Vallée, France
[‡] Hasselt University, Belgium

1 Introduction

Nonlinear elliptic problems are of paramount importance in a broad range of domains such as physics, mechanics, economics, biology, and medicine, see, e.g. [2, 22, 32, 39, 40]. Numerical discretization methods then serve to deliver approximate solutions, upon employing iterative linearizations to resolve the arising discrete nonlinear systems, see, e.g., [7, 16, 18, 29, 34, 43] and the references therein.

Given a numerical approximation, there arises the important question of the error with respect to the exact solution. This is practically handled by the so-called a posteriori estimates. For nonlinear problems, these have been proposed, amongst others, in [3, 4, 10, 11, 17, 20, 26, 28, 29, 33, 35, 36, 38, 39, 43]. One step further, the estimates can be used to adaptively steer the numerical approximation, and recently, convergence and optimality results have been obtained in [5, 9, 24, 25, 27, 30, 31], see also the references therein.

A crucial property of an a posteriori estimate is its robustness, assessing whether its quality is independent of the parameters. In the present setting, we specifically use the term robustness if the chosen error measure and the associated estimate are uniformly equivalent, for any strength of the nonlinearity. This resumes to the same overestimation factor (effectivity index) for linear, mildly nonlinear, and highly nonlinear problems. Unfortunately, robustness is often not achieved; we illustrate this in Figure 1. There, we present the effectivity indices for three common error measures: the energy norm (L^2 norm of the difference of the weak gradients), the (square root of the) difference of the energies, and the dual norm of the residual (cf., e.g., [23] for their mutual comparisons). We employ guaranteed equilibrated flux estimates following [8, 12, 15], for the nonlinear problem of Example 5.2 below. We can observe that the estimate in the energy norm setting is not robust with respect to the strength of the nonlinearity (the effectivity index explodes as the ratio $a_{\mathrm{c}}/a_{\mathrm{m}}$ from (2.1) below grows). The dual norm of the residual leads to robustness, as proven in [17, 20]. Though the dual norm of the residual is indeed localizable, cf. [6] and the references therein, it may be criticized as it actually does not take into account the nonlinearity (an incorporation has recently been addressed in [37]). The energy difference is showing a numerical robustness, though known theoretical estimates depend on the ratio $a_{\mathrm{c}}/a_{\mathrm{m}}$.

In this paper, we focus on nonlinear elliptic problems of the form: find $u : \Omega \to \mathbb{R}$ such that

$$-\boldsymbol{\nabla}\cdot(a(\cdot,|\boldsymbol{\nabla} u|)\boldsymbol{\nabla} u) = f \quad \text{in } \Omega, \tag{1.1a}$$
$$u = 0 \quad \text{on } \partial\Omega, \tag{1.1b}$$

where $a : \Omega \times [0,\infty) \to (0,\infty)$ is a nonlinear function satisfying assumptions of Lipschitz continuity and strong monotonicity (cf. (2.1) below). We employ a finite element approximation of (1.1) and an iterative linearization, yielding the approximation u_ℓ^k on each mesh $\mathcal{T}_\ell$ and linearization step k. The iterative linearization method needs to satisfy a few clearly identified assumptions. We will show that this is satisfied for usual linearizations such as Picard, Newton, or Zarantonello.

We consider the energy difference $\mathcal{E}_{\mathrm{N},\ell}^k$ of the nonlinear problem (cf. (3.9a)) and its a posteriori estimator $\eta_{\mathrm{N},\ell}^k$ (cf. (3.9b)). In order to achieve our goals, we additionally consider a linearized energy difference $\mathcal{E}_{\mathrm{L},\ell}^k$ (cf. (3.11a)) and the associated estimator $\eta_{\mathrm{L},\ell}^k$ (cf. (3.11b)). We more precisely augment $\mathcal{E}_{\mathrm{N},\ell}^k$ by $\mathcal{E}_{\mathrm{L},\ell}^k$ to form $\mathcal{E}_\ell^k$ and similarly for the estimators. Our main result, Theorem 3.4, can be presented

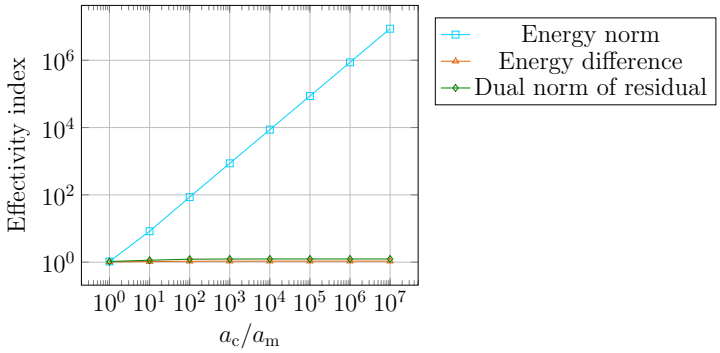


Figure 1: [Exponential nonlinearity (5.5), smooth solution (5.3), Newton solver, 25 DOFs] Comparison of the effectivity index (given as the ratio of the estimate over the error) of different error measures and associated a posteriori estimates.

as follows. If the iterative index k satisfies a computable stopping criterion and neglecting data oscillation and quadrature error terms, we have

$$\mathcal{E}_\ell^k \le \eta_\ell^k \lesssim C_\ell^k \mathcal{E}_\ell^k, \tag{1.2}$$

where the hidden constant depends only on the space dimension, the mesh shape-regularity, and possibly on the polynomial degree of the finite element approximation when the spatial dimension is greater than or equal to 4. Here, C_ℓ^k is a computable constant only depending on local variations of the linearization matrices. We show that $C_\ell^k = 1$ in the case of the Zarantonello iteration, making the estimate robust with respect to the strength of the nonlinearity in any setting. For the other linearizations, the estimates are robust if the computed constant C_ℓ^k is small (a posteriori verification of robustness for each given setting).

An additional a posteriori estimate is given in Theorem 4.1 for the energy difference $\mathcal{E}_{\mathrm{N},\ell}^k$ only, and can be summarized as follows. For every iterative index k, neglecting again data oscillation and quadrature error terms, we have

$$\mathcal{E}_{\mathrm{N},\ell}^k \le \eta_{\mathrm{N},\ell}^k \lesssim \widetilde{C}_\ell^k \mathcal{E}_{\mathrm{N},\ell}^k, \tag{1.3}$$

with the same dependence as in (1.2) for the hidden constant and where $\widetilde{C}_\ell^k$ only has a local (but here not computable) dependence on the nonlinearity and where $\widetilde{C}_\ell^k \le (a_c/a_m)^{1/2}$ in any case.

The rest of the paper is organized as follows. In Section 2, we detail the assumptions on the nonlinear function a. We next give the continuous weak formulation with its equivalent energy minimization. Then, we announce the discrete weak formulation with its associated discrete energy minimization, and finally the iterative linearization. In Section 3, we define the augmented error and estimator and state our main result, Theorem 3.4, giving details of (1.2). In Section 4, we study the original energy difference, which leads us to an a posteriori error estimate of Theorem 4.1, giving details of (1.3). In Section 5, we present a series of numerical experiments in order to illustrate our theoretical findings, for both smooth and singular solutions. In Section 6, we give a proof of Theorem 3.4, and then, in Section 7, a proof of Theorem 4.1. Finally, we summarize, in Appendix A, some useful properties of the nonlinear function and the assumptions required, and, in Appendix B, we show some technical results to determine the eigenvalues of the Newton linearization.

2 Weak formulation, energy minimization, finite element discretization, and iterative linearization

Let $\Omega \subset \mathbb{R}^d$, $d \ge 1$, be an open polytope with Lipschitz boundary $\partial\Omega$. We consider problem (1.1), where $f \in L^2(\Omega)$ represents a volumetric force term, while a is the diffusion coefficient which depends on the potential $u : \Omega \to \mathbb{R}$ only through the Euclidean norm of its gradient $|\boldsymbol{\nabla} u|$.

We consider the following assumption for the nonlinear function a (see, e.g., [43] for more details).

Assumption 2.1 (Nonlinear function a)**.** *We assume that the function $a : \Omega \times [0,\infty) \to (0,\infty)$ is measurable and that there exist constants $a_m \le a_c \in (0,\infty)$ such that, a.e. in Ω and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$,*

$$|a(\cdot,|\boldsymbol{x}|)\boldsymbol{x} - a(\cdot,|\boldsymbol{y}|)\boldsymbol{y}| \le a_c|\boldsymbol{x}-\boldsymbol{y}| \qquad \textit{(Lipschitz continuity)}, \tag{2.1a}$$

$$(a(\cdot,|\boldsymbol{x}|)\boldsymbol{x} - a(\cdot,|\boldsymbol{y}|)\boldsymbol{y}) \cdot (\boldsymbol{x}-\boldsymbol{y}) \ge a_m|\boldsymbol{x}-\boldsymbol{y}|^2 \qquad \textit{(strong monotonicity)}. \tag{2.1b}$$

2.1 Weak formulation and equivalent energy minimization

The weak formulation of problem (1.1) reads: find $u \in H^1_0(\Omega)$ such that

$$(a(\cdot, |\boldsymbol{\nabla} u|)\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) = (f, v) \qquad \forall v \in H^1_0(\Omega), \tag{2.2a}$$

where $(\cdot,\cdot)$ is the inner product of $L^2(\Omega)$. We denote by $\|\cdot\|$ its corresponding norm.

Referring to [43], the weak formulation (2.2a) is equivalent to the following minimization problem:

$$u = \arg \min_{v \in H^1_0(\Omega)} \mathcal{J}(v), \tag{2.2b}$$

with the energy functional $\mathcal{J} : H^1_0(\Omega) \to \mathbb{R}$ defined as,

$$\mathcal{J}(v) \coloneqq \int_\Omega \phi(\cdot, |\boldsymbol{\nabla} v|) - (f, v), \quad v \in H^1_0(\Omega), \tag{2.3}$$

where the function $\phi : \Omega \times [0,\infty) \to [0,\infty)$ is defined such that, a.e. in Ω and for all $r \in [0,\infty)$,

$$\phi(\cdot, r) \coloneqq \int_0^r a(\cdot, s)s \,\mathrm{d}s. \tag{2.4}$$

It is shown in [43] that, under Assumption 2.1, there exists a unique solution to problem (2.2). We refer to Appendix A for more details about equivalent assumptions on the nonlinear functions a and ϕ.

2.2 Finite element discretization

Let $\ell \geq 0$ be a mesh level index. We consider simplicial triangulations $\mathcal{T}_\ell$ of the domain Ω satisfying the following shape-regularity property: there exists a constant $\kappa_\mathcal{T} > 0$ such that for all $\ell \geq 0$ and all $K \in \mathcal{T}_\ell$, $h_K/\rho_K \leq \kappa_\mathcal{T}$, where h_K is the diameter of K and ρ_K is the diameter of the largest ball inscribed in K.

For a polynomial degree $p \geq 1$, denoting $\mathcal{P}_p(\mathcal{T}_\ell)$ the space of piecewise polynomials on the mesh $\mathcal{T}_\ell$ of total degree at most p, we define the discrete finite element space $V^p_\ell \coloneqq \mathcal{P}_p(\mathcal{T}_\ell) \cap H^1_0(\Omega)$. The finite element approximation of (2.2a) would be $u_\ell \in V^p_\ell$ such that

$$(a(\cdot, |\boldsymbol{\nabla} u_\ell|)\boldsymbol{\nabla} u_\ell, \boldsymbol{\nabla} v_\ell) = (f, v_\ell) \qquad \forall v_\ell \in V^p_\ell. \tag{2.5a}$$

As in (2.2b), $u_\ell \in V^p_\ell$ satisfies the minimization problem

$$u_\ell = \arg \min_{v_\ell \in V^p_\ell} \mathcal{J}(v_\ell). \tag{2.5b}$$

Below, we never work with u_ℓ.

2.3 Iterative linearization

We immediately consider an iterative linearization of (2.5a), which is anyhow necessary for a practical solution of (2.5a). Let $u^0_\ell \in V^p_\ell$ be a given initial guess. For an iterative linearization index $k \geq 1$, consider $\boldsymbol{A}^{k-1}_\ell : \Omega \to \mathbb{R}^{d\times d}$ and $\boldsymbol{b}^{k-1}_\ell : \Omega \to \mathbb{R}^d$, arising from a suitable linearization; details and examples are given below. We define the linearized finite element approximation: $u^k_\ell \in V^p_\ell$ to be such that

$$(\boldsymbol{A}^{k-1}_\ell \boldsymbol{\nabla} u^k_\ell, \boldsymbol{\nabla} v_\ell) = (f, v_\ell) + (\boldsymbol{b}^{k-1}_\ell, \boldsymbol{\nabla} v_\ell) \qquad \forall v_\ell \in V^p_\ell. \tag{2.6a}$$

As in (2.5b), this is equivalent to the discrete minimization problem

$$u^k_\ell = \arg \min_{v_\ell \in V^p_\ell} \mathcal{J}^{k-1}_\ell(v_\ell) \tag{2.6b}$$

with the linearized energy functional $\mathcal{J}^{k-1}_\ell : H^1_0(\Omega) \to \mathbb{R}$ defined for all $v \in H^1_0(\Omega)$ by

$$\mathcal{J}^{k-1}_\ell(v) \coloneqq \frac{1}{2} \left\| (\boldsymbol{A}^{k-1}_\ell)^{\frac{1}{2}} \boldsymbol{\nabla} v \right\|^2 - (f, v) - (\boldsymbol{b}^{k-1}_\ell, \boldsymbol{\nabla} v). \tag{2.7}$$

2.3.1 Assumptions on iterative linearization schemes

Let $k \geq 1$. We will suppose that $\boldsymbol{A}_\ell^{k-1} : \Omega \to \mathbb{R}^{d\times d}$ and $\boldsymbol{b}_\ell^{k-1} : \Omega \to \mathbb{R}^d$ from (2.6) satisfy the following assumption (for the sake of conciseness, we assume that they are well defined everywhere in Ω). We will use the following notation of the derivatives in the second argument of the functions a, ϕ, and others (cf. (3.6)–(3.7)): for all $r \in [0,\infty)$, $a'(\cdot, r) \coloneqq \frac{\partial}{\partial r} a(\boldsymbol{x}, r)$ and $\phi'(\cdot, r) \coloneqq \frac{\partial}{\partial r}\phi(\boldsymbol{x}, r)$.

Assumption 2.2 (Iterative linearization)**.** *For all points $\boldsymbol{x} \in \Omega$, we assume that $\boldsymbol{A}_\ell^{k-1}(\boldsymbol{x}) \in \mathbb{R}^{d\times d}$ is a bounded symmetric positive definite matrix. Specifically, denoting by $A_{\mathrm{m},\ell}^{k-1}(\boldsymbol{x})$ and $A_{\mathrm{c},\ell}^{k-1}(\boldsymbol{x})$ respectively its smallest and largest eigenvalues, we suppose, for all $\boldsymbol{\xi} \in \mathbb{R}^d$,*

$$|\boldsymbol{A}_\ell^{k-1}(\boldsymbol{x})\boldsymbol{\xi}| \leq A_{\mathrm{c},\ell}^{k-1}(\boldsymbol{x})|\boldsymbol{\xi}| \quad \textit{(boundedness)}, \tag{2.8a}$$

$$(\boldsymbol{A}_\ell^{k-1}(\boldsymbol{x})\boldsymbol{\xi})\cdot\boldsymbol{\xi} \geq A_{\mathrm{m},\ell}^{k-1}(\boldsymbol{x})|\boldsymbol{\xi}|^2 \quad \textit{(positive definiteness)}. \tag{2.8b}$$

Moreover, we suppose uniformity, i.e., that there exist $A_{\mathrm{m}} \leq A_{\mathrm{c}} \in (0,\infty)$ independent of k, ℓ, and $\boldsymbol{x}$ such that

$$A_{\mathrm{m}} \leq A_{\mathrm{m},\ell}^{k-1}(\boldsymbol{x}) \leq A_{\mathrm{c},\ell}^{k-1}(\boldsymbol{x}) \leq A_{\mathrm{c}}. \tag{2.8c}$$

Finally, we explicitly define $\boldsymbol{b}_\ell^{k-1}(\boldsymbol{x}) \in \mathbb{R}^d$ for all $\boldsymbol{x} \in \Omega$ by

$$\boldsymbol{b}_\ell^{k-1}(\boldsymbol{x}) \coloneqq \boldsymbol{A}_\ell^{k-1}(\boldsymbol{x})\boldsymbol{\nabla} u_\ell^{k-1}(\boldsymbol{x}) - a(\boldsymbol{x}, |\boldsymbol{\nabla} u_\ell^{k-1}(\boldsymbol{x})|)\boldsymbol{\nabla} u_\ell^{k-1}(\boldsymbol{x}). \tag{2.9}$$

In the following, we use the boldface font to denote the spaces of multi-dimensional functions, e.g., $\boldsymbol{L}^2(\Omega)$.

Remark 2.3 (Assumption 2.2)**.** *Equality (2.9) implies that (2.6a) can be equivalently written as a problem for the increment $u_\ell^k - u_\ell^{k-1}$ on the left-hand side and the residual of u_ℓ^{k-1} on the right-hand side:*

$$(\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}(u_\ell^k - u_\ell^{k-1}), \boldsymbol{\nabla} v_\ell) = (f, v_\ell) - (a(\cdot, |\boldsymbol{\nabla} u_\ell^{k-1}|)\boldsymbol{\nabla} u_\ell^{k-1}, \boldsymbol{\nabla} v_\ell) \quad \forall v_\ell \in V_\ell^p, \tag{2.10}$$

which is the form used in, e.g., [37]. Therefore, equality (2.9) ensures that the discrete problem (2.6a) is consistent with the discrete problem (2.5a), i.e., $\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla} u_\ell^k - \boldsymbol{b}_\ell^{k-1} \to a(\cdot, |\boldsymbol{\nabla} u_\ell|)\boldsymbol{\nabla} u_\ell$ in $\boldsymbol{L}^2(\Omega)$ when $u_\ell^k \to u_\ell$ in $H_0^1(\Omega)$. Indeed, (2.8a) implies that $\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}(u_\ell^k - u_\ell^{k-1}) \to \boldsymbol{0}$ in $\boldsymbol{L}^2(\Omega)$, and $a(\cdot, |\boldsymbol{\nabla} u_\ell^{k-1}|)\boldsymbol{\nabla} u_\ell^{k-1} \to a(\cdot, |\boldsymbol{\nabla} u_\ell|)\boldsymbol{\nabla} u_\ell$ in $\boldsymbol{L}^2(\Omega)$ thanks to (2.1a). Finally, we recall that the positive definiteness of $\boldsymbol{A}_\ell^{k-1}$ implies that $(\boldsymbol{A}_\ell^{k-1})^{-1}$ and $(\boldsymbol{A}_\ell^{k-1})^{\frac{1}{2}}$ exist, which is used below.

2.3.2 Examples of iterative linearization schemes

We now present some examples of linearization methods satisfying Assumption 2.2.

Example 2.4 (Picard)**.** *The Picard (fixed point) iteration, see, e.g. [16], is defined as*

$$\boldsymbol{A}_\ell^{k-1} = a(\cdot, |\boldsymbol{\nabla} u_\ell^{k-1}|)\boldsymbol{I}_d \quad \textit{with} \quad \boldsymbol{b}_\ell^{k-1} = \boldsymbol{0} \quad \textit{in } \Omega. \tag{2.11}$$

It satisfies Assumption 2.2 with $A_{\mathrm{m},\ell}^{k-1} = A_{\mathrm{c},\ell}^{k-1} = a(\cdot, |\boldsymbol{\nabla} u_\ell^{k-1}|)$, which leads to $A_{\mathrm{m}} = a_{\mathrm{m}}$ and $A_{\mathrm{c}} = a_{\mathrm{c}}$ thanks to (A.3).

Example 2.5 (Zarantonello)**.** *The Zarantonello iteration, introduced in [42], is defined as*

$$\boldsymbol{A}_\ell^{k-1} = \gamma\boldsymbol{I}_d \quad \textit{with} \quad \boldsymbol{b}_\ell^{k-1} = \big(\gamma - a(\cdot, |\boldsymbol{\nabla} u_\ell^{k-1}|)\big)\,\boldsymbol{\nabla} u_\ell^{k-1} \quad \textit{in } \Omega, \tag{2.12}$$

where $\gamma \in (0,\infty)$ is a constant parameter. To ensure contraction, one needs to assume that $\gamma \geq \frac{a_{\mathrm{c}}^2}{a_{\mathrm{m}}}$. The Zarantonello iteration converges linearly, but the convergence is slow as γ takes large values. It satisfies Assumption 2.2 with $A_{\mathrm{m},\ell}^{k-1} = A_{\mathrm{c},\ell}^{k-1} = \gamma$, which leads to $A_{\mathrm{m}} = A_{\mathrm{c}} = \gamma$.

Example 2.6 ((Damped) Newton)**.** *The (damped) Newton iteration, see, e.g. [16], is defined as*

$$\begin{aligned} \boldsymbol{A}_\ell^{k-1} &= a(\cdot, |\boldsymbol{\nabla} u_\ell^{k-1}|)\boldsymbol{I}_d + \theta\frac{a'(\cdot, |\boldsymbol{\nabla} u_\ell^{k-1}|)}{|\boldsymbol{\nabla} u_\ell^{k-1}|}\boldsymbol{\nabla} u_\ell^{k-1}\otimes\boldsymbol{\nabla} u_\ell^{k-1} \\ \textit{with} \quad \boldsymbol{b}_\ell^{k-1} &= \theta a'(\cdot, |\boldsymbol{\nabla} u_\ell^{k-1}|)|\boldsymbol{\nabla} u_\ell^{k-1}|\boldsymbol{\nabla} u_\ell^{k-1} \quad \textit{in } \Omega, \end{aligned} \tag{2.13}$$

where $\theta \in [0,1]$ is the damping parameter. Observe that $\theta = 1$ gives the Newton iteration, whereas $\theta = 0$ corresponds to the Picard iteration. If $\theta = 1$, the Newton method converges quadratically. However, it might not always converge. The (damped) Newton iteration satisfies Assumption 2.2 with, if the space dimension $d = 1$,

$$\begin{aligned} A_{\mathrm{m},\ell}^{k-1} = A_{\mathrm{c},\ell}^{k-1} &= a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|) + \theta a'(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|)|\boldsymbol{\nabla} u_\ell^{k-1}| \\ &\overset{(A.8)}{=} (1-\theta)a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|) + \theta \phi''(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|), \end{aligned} \tag{2.14}$$

and, if $d > 1$,

$$A_{\mathrm{m},\ell}^{k-1} = (1-\theta)a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|) + \theta \min(a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|), \phi''(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|)), \tag{2.15a}$$

$$A_{\mathrm{c},\ell}^{k-1} = (1-\theta)a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|) + \theta \max(a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|), \phi''(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|)). \tag{2.15b}$$

Indeed, denoting the spectrum of a matrix $\boldsymbol{A} \in \mathbb{R}^{d\times d}$ by $\mathrm{Spec}(\boldsymbol{A})$, we infer (2.15) by writing,

$$\begin{aligned} \mathrm{Spec}(\boldsymbol{A}_\ell^{k-1}) &\overset{(B.1)}{=} \{a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|), a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|) + \theta a'(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|)|\boldsymbol{\nabla} u_\ell^{k-1}|\} \\ &\overset{(A.8)}{=} \{(1-\theta)a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|) + \theta a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|), \\ &\qquad (1-\theta)a(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|) + \theta \phi''(\cdot,|\boldsymbol{\nabla} u_\ell^{k-1}|)\}. \end{aligned}$$

Finally, we can set $A_{\mathrm{m}} = a_{\mathrm{m}}$ and $A_{\mathrm{c}} = a_{\mathrm{c}}$ thanks to (A.3) and (A.5).

3 A posteriori estimate of an augmented energy difference

In this section, we define the augmented energy difference and a posteriori estimator of the finite element discretization u_ℓ^k given by (2.6) on mesh $\mathcal{T}_\ell$, $\ell \geq 0$, and linearization step $k \geq 1$.

3.1 Flux equilibration

Let $\mathcal{V}_\ell$ be the set of all mesh vertices and, for each vertex $\boldsymbol{a} \in \mathcal{V}_\ell$, define the patch $\mathcal{T}_\ell^{\boldsymbol{a}}$ as the collection of the simplicies of $\mathcal{T}_\ell$ sharing the vertex $\boldsymbol{a}$, as well as the patch subdomain $\omega_\ell^{\boldsymbol{a}}$ corresponding to $\mathcal{T}_\ell^{\boldsymbol{a}}$. For all $\boldsymbol{a} \in \mathcal{V}_\ell$, we define the space $\boldsymbol{V}_\ell^{\boldsymbol{a}} \coloneqq \boldsymbol{RTN}_p(\mathcal{T}_\ell^{\boldsymbol{a}}) \cap \boldsymbol{H}_0(\mathrm{div}, \omega_\ell^{\boldsymbol{a}})$. Here $\boldsymbol{RTN}_p(\mathcal{T}_\ell^{\boldsymbol{a}})$ denotes the broken space consisting of p-th order Raviart–Thomas–Nédélec space on each simplex, $\boldsymbol{RTN}_p(K) \coloneqq [\mathcal{P}(K)]^d + \boldsymbol{x}\mathcal{P}(K)$. Moreover, $\boldsymbol{H}_0(\mathrm{div}, \omega_\ell^{\boldsymbol{a}})$ is the subspace of $\boldsymbol{H}(\mathrm{div}, \omega_\ell^{\boldsymbol{a}})$ of functions with vanishing normal trace on $\partial\omega_\ell^{\boldsymbol{a}}$ if $\boldsymbol{a} \in \mathcal{V}_\ell$ is an interior vertex and of functions with vanishing normal trace on $\partial\omega_\ell^{\boldsymbol{a}} \backslash \{\psi_\ell^{\boldsymbol{a}} > 0\}$ if $\boldsymbol{a} \in \mathcal{V}_\ell$ is a boundary vertex.

Here, for all $\boldsymbol{a} \in \mathcal{V}_\ell$, the hat function $\psi_\ell^{\boldsymbol{a}}$ is the continuous, piecewise affine function equal to 1 in $\boldsymbol{a}$ and 0 in $\mathcal{V}_\ell \backslash \{\boldsymbol{a}\}$. We recall the partition of unity

$$\sum_{\boldsymbol{a} \in \mathcal{V}_\ell} \psi_\ell^{\boldsymbol{a}}(\boldsymbol{x}) = 1 \quad \forall \boldsymbol{x} \in \Omega. \tag{3.1}$$

We denote by $\Pi_{\ell,p}$ the L^2-orthogonal projection from $L^2(\Omega)$ to $\mathcal{P}_p(\mathcal{T}_\ell)$ and by $\boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}}$ the $\boldsymbol{L}^2$-orthogonal projection from $\boldsymbol{L}^2(\Omega)$ to $\boldsymbol{RTN}_{p-1}(\mathcal{T}_\ell)$; note that both are elementwise. Finally, we consider the equilibrated flux locally reconstructed from u_ℓ^k as

$$\boldsymbol{\sigma}_\ell^k \coloneqq \sum_{\boldsymbol{a} \in \mathcal{V}_\ell} \boldsymbol{\sigma}_\ell^{\boldsymbol{a},k}, \tag{3.2a}$$

where, for all vertices $\boldsymbol{a} \in \mathcal{V}_\ell$, following [8, 11, 15, 20, 21],

$$\begin{aligned} \boldsymbol{\sigma}_\ell^{\boldsymbol{a},k} &\coloneqq \arg \min_{\substack{\boldsymbol{w}_\ell \in \boldsymbol{V}_\ell^{\boldsymbol{a}} \\ \boldsymbol{\nabla}\cdot\boldsymbol{w}_\ell = \Pi_{\ell,p}\gamma_\ell^{\boldsymbol{a},k}}} \|(\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}}(\psi_\ell^{\boldsymbol{a}} \boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}} \boldsymbol{\xi}_\ell^k + \boldsymbol{w}_\ell)\|_{\omega_\ell^{\boldsymbol{a}}}, \\ &\text{with } \boldsymbol{\xi}_\ell^k \coloneqq \boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla} u_\ell^k - \boldsymbol{b}_\ell^{k-1} \text{ and } \gamma_\ell^{\boldsymbol{a},k} \coloneqq \psi_\ell^{\boldsymbol{a}} f - \boldsymbol{\nabla}\psi_\ell^{\boldsymbol{a}} \cdot \boldsymbol{\xi}_\ell^k. \end{aligned} \tag{3.2b}$$

Combining (3.2) and (3.1), we infer, as in, e.g., [19],

$$\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}_\ell^k &= \sum_{\boldsymbol{a} \in \mathcal{V}_\ell} \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}_\ell^{\boldsymbol{a},k} = \sum_{\boldsymbol{a} \in \mathcal{V}_\ell} \Pi_{\ell,p}(\psi_\ell^{\boldsymbol{a}} f) - \sum_{\boldsymbol{a} \in \mathcal{V}_\ell} \Pi_{\ell,p}(\boldsymbol{\nabla}\psi_\ell^{\boldsymbol{a}} \cdot \boldsymbol{\xi}_\ell^k) \\ &= \Pi_{\ell,p} \sum_{\boldsymbol{a} \in \mathcal{V}_\ell} (\psi_\ell^{\boldsymbol{a}} f) = \Pi_{\ell,p} f. \end{aligned} \tag{3.3}$$

In particular, $\boldsymbol{\sigma}_\ell^k \in \boldsymbol{RTN}_p(\mathcal{T}_\ell) \cap \boldsymbol{H}(\mathrm{div}, \Omega)$ and we have with the Green theorem, since $u_\ell^k \in V_\ell^p \subset \mathcal{P}_p(\mathcal{T}_\ell)$,

$$-(\boldsymbol{\sigma}_\ell^k, \boldsymbol{\nabla} u_\ell^k) = (f, u_\ell^k). \tag{3.4}$$

3.2 Convex conjugate function and dual energies

Recalling the primal energy $\mathcal{J}$ of (2.3), the corresponding dual energy $\mathcal{J}^* : \boldsymbol{H}(\mathrm{div}, \Omega) \to \mathbb{R}$, cf. [2, 32, 39, 43], is defined by

$$\mathcal{J}^*(\boldsymbol{w}) \coloneqq -\int_\Omega \phi^*(\cdot, |\boldsymbol{w}|), \quad \boldsymbol{w} \in \boldsymbol{H}(\mathrm{div}, \Omega), \tag{3.5}$$

where $\phi^* : \Omega \times [0, \infty) \to [0, \infty)$ is the convex conjugate of ϕ (also known as the Legendre dual or the Fenchel conjugate), which is defined such that, a.e. in Ω and for all $s \in [0, \infty)$,

$$\phi^*(\cdot, s) \coloneqq \sup_{r \in [0,\infty)} (sr - \phi(\cdot, r)), \tag{3.6a}$$

or equivalently, for all $s \in [0, \infty)$,

$$\phi^*(\cdot, s) = \int_0^s \phi'^{-1}(\cdot, r)\, \mathrm{d}r = s\phi'^{-1}(\cdot, s) - \phi(\cdot, \phi'^{-1}(\cdot, s)). \tag{3.6b}$$

We refer to [32] for more details and recall that the construction of ϕ^* yields

$$\phi^{*\prime} = \phi'^{-1} \quad \text{and} \quad \phi^{*\prime\prime} = \frac{1}{\phi'' \circ \phi'^{-1}}. \tag{3.7}$$

Consequently, under Assumption 2.1, ϕ^* is convex thanks to Remark A.3 below.

Finally, we define the linearized dual energy functional $\mathcal{J}_\ell^{*,k-1} : \boldsymbol{H}(\mathrm{div}, \Omega) \to \mathbb{R}$ such that

$$\mathcal{J}_\ell^{*,k-1}(\boldsymbol{w}) \coloneqq -\frac{1}{2} \left\| (\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}} (\boldsymbol{w} - \boldsymbol{b}_\ell^{k-1}) \right\|^2, \quad \boldsymbol{w} \in \boldsymbol{H}(\mathrm{div}, \Omega). \tag{3.8}$$

3.3 Energy difference and estimator of the nonlinear problem

We define the (square root of twice the) energy difference corresponding to the nonlinear problem (2.2) by

$$\mathcal{E}_{\mathrm{N},\ell}^k \coloneqq \left(2(\mathcal{J}(u_\ell^k) - \mathcal{J}(u))\right)^{\frac{1}{2}}. \tag{3.9a}$$

Note that $\mathcal{E}_{\mathrm{N},\ell}^k$ is well-defined from (2.2b) and the fact that $u_\ell^k \in H_0^1(\Omega)$. Actually, $\mathcal{E}_{\mathrm{N},\ell}^k \geq 0$ and 0 if and only if $u_\ell^k = u$ from the uniqueness of u in (2.2b). We then define the estimator $\eta_{\mathrm{N},\ell}^k$ corresponding to the nonlinear problem (2.2) as in, e.g., [3, 39, 43], by

$$\eta_{\mathrm{N},\ell}^k \coloneqq \left(2(\mathcal{J}(u_\ell^k) - \mathcal{J}^*(\boldsymbol{\sigma}_\ell^k))\right)^{\frac{1}{2}}. \tag{3.9b}$$

Note that $\eta_{\mathrm{N},\ell}^k$ is again well-defined, which follows from (5.7)–(5.8) below.

3.4 Energy difference and estimator of the linearized problem

In order to derive robust estimates, we further consider the linearized problem (2.6). We introduce the abstract linearization on the continuous level: $u_{\langle\ell\rangle}^k \in H_0^1(\Omega)$ such that

$$(\boldsymbol{A}_\ell^{k-1} \boldsymbol{\nabla} u_{\langle\ell\rangle}^k, \boldsymbol{\nabla} v) = (f, v) + (\boldsymbol{b}_\ell^{k-1}, \boldsymbol{\nabla} v) \qquad \forall v \in H_0^1(\Omega), \tag{3.10a}$$

which is a linear problem. Note that (3.10a) is equivalent to

$$u_{\langle\ell\rangle}^k \coloneqq \arg \min_{v \in H_0^1(\Omega)} \mathcal{J}_\ell^{k-1}(v), \tag{3.10b}$$

based on the linearized energy (2.7). Analogously to the nonlinear energy difference $\mathcal{E}_{\mathrm{N},\ell}^k$ (3.9a), we then define the energy difference $\mathcal{E}_{\mathrm{L},\ell}^k$ of the linearized problem (3.10)

$$\mathcal{E}_{\mathrm{L},\ell}^k \coloneqq \Big(2\big(\mathcal{J}_\ell^{k-1}(u_\ell^k) - \mathcal{J}_\ell^{k-1}(u_{\langle\ell\rangle}^k)\big)\Big)^{\frac{1}{2}} \overset{(6.5a)}{=} \|(\boldsymbol{A}_\ell^{k-1})^{\frac{1}{2}}\boldsymbol{\nabla}(u_\ell^k - u_{\langle\ell\rangle}^k)\|. \tag{3.11a}$$

Here, $\mathcal{E}_{\mathrm{L},\ell}^k \geq 0$ again and 0 if and only if $u_\ell^k = u_{\langle\ell\rangle}^k$ is obvious. Finally, we define the estimator $\eta_{\mathrm{L},\ell}^k$ of the linearized problem (3.10) by

$$\begin{aligned}\eta_{\mathrm{L},\ell}^k &\coloneqq \Big(2\big(\mathcal{J}_\ell^{k-1}(u_\ell^k) - \mathcal{J}_\ell^{*,k-1}(\boldsymbol{\sigma}_\ell^k)\big)\Big)^{\frac{1}{2}} \\ &\overset{(6.6a)}{=} \|(\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}}(\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}u_\ell^k - \boldsymbol{b}_\ell^{k-1} + \boldsymbol{\sigma}_\ell^k)\| \\ &\overset{(2.9)}{=} \|(\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}}(\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}(u_\ell^k - u_\ell^{k-1}) + a(\cdot,|\boldsymbol{\nabla}u_\ell^{k-1}|)\boldsymbol{\nabla}u_\ell^{k-1} + \boldsymbol{\sigma}_\ell^k)\|.\end{aligned} \tag{3.11b}$$

3.5 Augmented energy difference and estimator

We denote, a.e. in Ω,

$$a_\ell^k \coloneqq a(\cdot,|\boldsymbol{\nabla}u_\ell^k|) \quad \text{and} \quad a_u \coloneqq a(\cdot,|\boldsymbol{\nabla}u|). \tag{3.12}$$

We define the shifted estimator $\widehat{\eta}_{\mathrm{L},\ell}^k$ of the linearized problem (3.10) by

$$\widehat{\eta}_{\mathrm{L},\ell}^k \coloneqq \|(\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}}(a_\ell^k\boldsymbol{\nabla}u_\ell^k + \boldsymbol{\sigma}_\ell^k)\| \tag{3.13}$$

and the computable weight λ_ℓ^k defined from the estimators (3.9b) and (3.13) by

$$\lambda_\ell^k \coloneqq \frac{\eta_{\mathrm{N},\ell}^k}{\widehat{\eta}_{\mathrm{L},\ell}^k}. \tag{3.14}$$

Remark 3.1 (Weight λ_ℓ^k)**.** *The role of the weight λ_ℓ^k is to ensure a balance between $\mathcal{E}_{\mathrm{N},\ell}^k$ and $\lambda_\ell^k\mathcal{E}_{\mathrm{L},\ell}^k$, and $\eta_{\mathrm{N},\ell}^k$ and $\lambda_\ell^k\eta_{\mathrm{L},\ell}^k$. We employ $\widehat{\eta}_{\mathrm{L},\ell}^k$ (and not $\eta_{\mathrm{L},\ell}^k$ from (3.11b) which would give $\lambda_\ell^k\eta_{\mathrm{L},\ell}^k = \eta_{\mathrm{N},\ell}^k$) since as such λ_ℓ^k is uniformly bounded for every $k \geq 1$. Indeed, we have with (4.3a) and (7.4) below,*

$$\lambda_\ell^k \overset{(7.4)}{\leq} \frac{\|(a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k})^{-\frac{1}{2}}(a_\ell^k\boldsymbol{\nabla}u_\ell^k + \boldsymbol{\sigma}_\ell^k)\|}{\|(\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}}(a_\ell^k\boldsymbol{\nabla}u_\ell^k + \boldsymbol{\sigma}_\ell^k)\|} \overset{(2.8a)}{\leq} \operatorname*{ess\,sup}_\Omega \left(\frac{A_{\mathrm{c},\ell}^{k-1}}{a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k}}\right)^{\frac{1}{2}} \overset{(2.8c),(4.3a)}{\leq} \sqrt{\frac{A_\mathrm{c}}{a_\mathrm{m}}}. \tag{3.15}$$

Now, we define the augmented energy difference $\mathcal{E}_\ell^k$ by

$$\mathcal{E}_\ell^k \coloneqq \frac{1}{2}\left(\mathcal{E}_{\mathrm{N},\ell}^k + \lambda_\ell^k\mathcal{E}_{\mathrm{L},\ell}^k\right) \tag{3.16a}$$

and the corresponding augmented estimator η_ℓ^k by

$$\eta_\ell^k \coloneqq \frac{1}{2}\left(\eta_{\mathrm{N},\ell}^k + \lambda_\ell^k\eta_{\mathrm{L},\ell}^k\right). \tag{3.16b}$$

Remark 3.2 (Roles of $\mathcal{E}_\ell^k$ and η_ℓ^k)**.** *The objective of introducing $\mathcal{E}_\ell^k$ and η_ℓ^k is to augment the energy difference $\mathcal{E}_{\mathrm{N},\ell}^k$ and estimator $\eta_{\mathrm{N},\ell}^k$ of the nonlinear problem (2.2) (where $\mathcal{E}_{\mathrm{N},\ell}^k \leq \eta_{\mathrm{N},\ell}^k$ but not necessarily robustly) in order to obtain an a posteriori estimate that will be robust with respect to the nonlinear function a. The added linearized components $\mathcal{E}_{\mathrm{L},\ell}^k$ and $\eta_{\mathrm{L},\ell}^k$ are multiplied by the weight λ_ℓ^k which makes them comparable in size to $\mathcal{E}_{\mathrm{N},\ell}^k$ and $\eta_{\mathrm{N},\ell}^k$, respectively.*

Lemma 3.3 (Error consistency)**.** *Let the error $\mathcal{E}_\ell^k$ be given by (3.16a). We have $\mathcal{E}_\ell^k \to 0$ if and only if $u_\ell^k \to u$ in $H_0^1(\Omega)$.*

Proof. Assuming $\mathcal{E}_\ell^k \to 0$ implies $\mathcal{E}_{\mathrm{N},\ell}^k \to 0$ since $\mathcal{E}_{\mathrm{N},\ell}^k, \mathcal{E}_{\mathrm{L},\ell}^k, \lambda_\ell^k \geq 0$. Furthermore, inequality (7.2a) below together with (4.4) yields

$$a_\mathrm{m}\|\boldsymbol{\nabla}(u_\ell^k - u)\|^2 \leq (\mathcal{E}_{\mathrm{N},\ell}^k)^2 \leq a_\mathrm{c}\|\boldsymbol{\nabla}(u_\ell^k - u)\|^2, \tag{3.17}$$

hence $u_\ell^k \to u$ in $H_0^1(\Omega)$. Reciprocally, assume that $u_\ell^k \to u$ in $H_0^1(\Omega)$. We have again from (3.17) that $\mathcal{E}_{\mathrm{N},\ell}^k \to 0$. Furthermore, we have for all $k \geq 1$,

$$\begin{aligned}
A_{\mathrm{m}}\|\boldsymbol{\nabla}(u_{\langle\ell\rangle}^k - u_\ell^{k-1})\|^2 &\overset{(2.8)}{\leq} (\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}(u_{\langle\ell\rangle}^k - u_\ell^{k-1}), \boldsymbol{\nabla}(u_{\langle\ell\rangle}^k - u_\ell^{k-1})) \\
&\overset{(2.9)}{=} (\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}u_{\langle\ell\rangle}^k - \boldsymbol{b}_\ell^{k-1}, \boldsymbol{\nabla}(u_{\langle\ell\rangle}^k - u_\ell^{k-1})) - (a_\ell^{k-1}\boldsymbol{\nabla}u_\ell^{k-1}, \boldsymbol{\nabla}(u_{\langle\ell\rangle}^k - u_\ell^{k-1})) \\
&\overset{(3.10a)}{=} (f, u_{\langle\ell\rangle}^k - u_\ell^{k-1}) - (a_\ell^{k-1}\boldsymbol{\nabla}u_\ell^{k-1}, \boldsymbol{\nabla}(u_{\langle\ell\rangle}^k - u_\ell^{k-1})) \\
&\overset{(2.2a)}{=} (a_u\boldsymbol{\nabla}u - a_\ell^{k-1}\boldsymbol{\nabla}u_\ell^{k-1}, \boldsymbol{\nabla}(u_{\langle\ell\rangle}^k - u_\ell^{k-1})) \\
&\overset{(2.1a)}{\leq} a_{\mathrm{c}}\|\boldsymbol{\nabla}(u - u_\ell^{k-1})\|\|\boldsymbol{\nabla}(u_{\langle\ell\rangle}^k - u_\ell^{k-1})\|.
\end{aligned} \tag{3.18}$$

Hence, dividing (3.18) by $\|\boldsymbol{\nabla}(u_{\langle\ell\rangle}^k - u_\ell^{k-1})\|$ shows that $u_{\langle\ell\rangle}^k$ converges to the same limit of u_ℓ^k in $H_0^1(\Omega)$, i.e. u. Thus, the relation (6.5a) below together with (2.8) show that $\mathcal{E}_{\mathrm{L},\ell}^k \to 0$. Hence, since λ_ℓ^k is bounded (cf. (3.15)), we infer $\mathcal{E}_\ell^k \to 0$. □

3.6 Data oscillation and quadrature estimators

Following, e.g, [19], let

$$(\eta_{\mathrm{osc},\ell}^k)^2 \coloneqq \sum_{K\in\mathcal{T}_\ell}\left[\frac{h_K}{\pi \inf_K (A_{\mathrm{m},\ell}^{k-1})^{\frac{1}{2}}}\|(I-\Pi_{\ell,p})f\|_K\right]^2, \tag{3.19a}$$

$$(\widetilde{\eta}_{\mathrm{osc},\ell}^k)^2 \coloneqq \sum_{K\in\mathcal{T}_\ell}\left[\frac{h_K}{\pi a_{\mathrm{m}}^{\frac{1}{2}}}\|(I-\Pi_{\ell,p})f\|_K\right]^2, \tag{3.19b}$$

$$(\eta_{\mathrm{osc,q},\ell}^k)^2 \coloneqq \sum_{\boldsymbol{a}\in\mathcal{V}_\ell}(\eta_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k})^2 \quad\text{and}\quad (\eta_{\mathrm{q},\ell}^k)^2 \coloneqq \sum_{\boldsymbol{a}\in\mathcal{V}_\ell}(\eta_{\mathrm{q},\ell}^{\boldsymbol{a},k})^2, \tag{3.19c}$$

where, for all vertices $\boldsymbol{a} \in \mathcal{V}_\ell$, using the notations $\boldsymbol{\xi}_\ell^k$ and $\gamma_\ell^{\boldsymbol{a},k}$ of (3.2b),

$$(\eta_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k})^2 \coloneqq \frac{1}{\inf_{\omega_\ell^{\boldsymbol{a}}} A_{\mathrm{m},\ell}^{k-1}}\sum_{K\in\mathcal{T}_\ell^{\boldsymbol{a}}}\left[\frac{h_K}{\pi}\|(I-\Pi_{\ell,p})\gamma_\ell^{\boldsymbol{a},k}\|_K\right]^2, \tag{3.19d}$$

$$\eta_{\mathrm{q},\ell}^{\boldsymbol{a},k} \coloneqq \|(\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}}(\psi_\ell^{\boldsymbol{a}}(\boldsymbol{I}-\boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}})\boldsymbol{\xi}_\ell^k)\|_{\omega_\ell^{\boldsymbol{a}}}. \tag{3.19e}$$

These correspond to the oscillation in the data f and quadrature errors possibly arising from the nonlinear function a.

3.7 A posteriori estimate of the augmented energy difference

We now present our main result giving an a posteriori estimate based on the augmented energy difference and estimator defined in (3.16).

Theorem 3.4 (A posteriori estimate of the augmented energy difference)**.** *Suppose Assumption 2.1 and let u be the weak solution of (2.2). Let u_ℓ^k be its finite element approximation given by (2.6) on mesh $\mathcal{T}_\ell$, $\ell \geq 0$, and linearization step $k \geq 1$, for any iterative linearization satisfying Assumption 2.2. Let $u_{\langle\ell\rangle}^k$ be given by (3.10), the augmented error $\mathcal{E}_\ell^k$ by (3.16a), and the augmented estimator η_ℓ^k by (3.16b). We have*

$$\mathcal{E}_\ell^k \leq \eta_\ell^k + \widetilde{\eta}_{\mathrm{osc},\ell}^k + \frac{\lambda_\ell^k}{2}\eta_{\mathrm{osc},\ell}^k. \tag{3.20}$$

Moreover, assume that the discrete linearization iterate $(u_\ell^k)_{k\geq 1}$ converge to u_ℓ of (2.5) in $H_0^1(\Omega)$. If the linearization index k satisfies the stopping criterion

$$\widehat{\eta}_{\mathrm{L},\ell}^k \leq 2\eta_{\mathrm{L},\ell}^k \tag{3.21}$$

for the estimators of (3.11b) *and* (3.13)*, then*

$$\eta_\ell^k \lesssim C_\ell^k \mathcal{E}_\ell^k + \lambda_\ell^k (C_\ell^k \eta_{\mathrm{q},\ell}^k + \eta_{\mathrm{osc,q},\ell}^k), \tag{3.22}$$

where the hidden constant only depends on the space dimension d, the mesh shape-regularity $\kappa_{\mathcal{T}}$, and possibly, when $d \geq 4$, the polynomial degree p, with

$$C_\ell^k := \max_{\boldsymbol{a}\in\mathcal{V}_\ell} \left(\frac{\sup_{\omega_\ell^{\boldsymbol{a}}} A_{\mathrm{c},\ell}^{k-1}}{\inf_{\omega_\ell^{\boldsymbol{a}}} A_{\mathrm{m},\ell}^{k-1}} \right)^{\frac{1}{2}} \begin{cases} = 1 & \text{for the Zarantonello linearization,} \\ \leq \left(\dfrac{A_{\mathrm{c}}}{A_{\mathrm{m}}} \right)^{\frac{1}{2}} & \text{in general.} \end{cases} \tag{3.23}$$

Proof. See Section 6. □

Remark 3.5 (Robustness for the Zarantonello linearization)**.** *In the Zarantonello linearization case, since $C_\ell^k = 1$, we obtain an estimation of the augmented energy difference $\mathcal{E}_\ell^k$ by the estimator η_ℓ^k (under small oscillations and quadrature errors) whose quality is independent of the nonlinear function a.*

Remark 3.6 (Robustness in the general case)**.** *The constant C_ℓ^k can be easily calculated in practice, without knowing the continuous solution u of* (2.2)*. It in particular only relies on patchwise variations of the linearization matrix $\boldsymbol{A}_\ell^{k-1}$ (recall* (2.8)*). For the Picard and damped Newton iteration cases of Examples 2.4 and 2.6, in particular, the local ratio of the functions $A_{\mathrm{c},\ell}^{k-1}$ and $A_{\mathrm{m},\ell}^{k-1}$ is smaller than that of the global constants A_{c} and A_{m} given in* (2.8c)*, typically bringing C_ℓ^k close to 1 as in the Zarantonello case. This is indeed observed in the numerical experiments of Section 5. Importantly, C_ℓ^k allows us to quantify the quality of the estimates in any situation: whenever it is small, we can affirm robustness a posteriori.*

Remark 3.7 (Criterion (3.21))**.** *The computable criterion* (3.21) *allows us to determine when k is large enough to obtain* (3.22)*. Indeed, we have*

$$\frac{\widehat{\eta}_{\mathrm{L},\ell}^k}{\eta_{\mathrm{L},\ell}^k} \overset{(3.13),(3.11b)}{=} \frac{\|(\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}}(a_\ell^k \boldsymbol{\nabla} u_\ell^k + \boldsymbol{\sigma}_\ell^k)\|}{\|(\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}}(\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}(u_\ell^k - u_\ell^{k-1}) + a_\ell^{k-1}\boldsymbol{\nabla} u_\ell^{k-1} + \boldsymbol{\sigma}_\ell^k)\|}. \tag{3.24}$$

With the same reasoning as in Remark 2.3, $\widehat{\eta}_{\mathrm{L},\ell}^k$ and $\eta_{\mathrm{L},\ell}^k$ become equivalent for large k when $u_\ell^k \to u_\ell$ in $H_0^1(\Omega)$. Hence, if $\eta_{\mathrm{L},\ell}^k$ (and $\widehat{\eta}_{\mathrm{L},\ell}^k$) does not vanish, then (3.21) *holds for k large enough. Otherwise, $\eta_{\mathrm{N},\ell}^k$ vanishes too (cf.* (7.4) *below), so that η_ℓ^k vanishes. In this particular case, if $\widetilde{\eta}_{\mathrm{osc},\ell}^k > 0$, then $\eta_{\mathrm{osc,q},\ell}^k$ does not vanish, and* (3.21) *holds for k large enough. Otherwise, if $\widetilde{\eta}_{\mathrm{osc},\ell}^k = 0$, then* (3.20) *implies $\mathcal{E}_\ell^k \to 0$ which is equivalent to $u_\ell^k \to u$ thanks to Lemma 3.3, and* (3.22) *is then not necessary.*

4 A posteriori estimate of the energy difference

This section gives an a posteriori estimate for the energy difference $\mathcal{E}_{\mathrm{N},\ell}^k$ by the estimator $\eta_{\mathrm{N},\ell}^k$, cf. (3.9).

4.1 Equilibrated flux and data oscillation-quadrature estimators

We redefine here the local equilibrated flux given in (3.2b) by setting, for all $\boldsymbol{a} \in \mathcal{V}_\ell$,

$$\begin{aligned} \boldsymbol{\sigma}_\ell^{\boldsymbol{a},k} &:= \arg \min_{\substack{\boldsymbol{w}_\ell \in \boldsymbol{V}_\ell^{\boldsymbol{a}} \\ \boldsymbol{\nabla}\cdot\boldsymbol{w}_\ell = \Pi_{\ell,p}\gamma_\ell^{\boldsymbol{a},k}}} \|\psi_\ell^{\boldsymbol{a}} \boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}}(a_\ell^k \boldsymbol{\nabla} u_\ell^k) + \boldsymbol{w}_\ell\|_{\omega_\ell^{\boldsymbol{a}}}, \\ \text{with } \gamma_\ell^{\boldsymbol{a},k} &:= \psi_\ell^{\boldsymbol{a}} f - \boldsymbol{\nabla}\psi_\ell^{\boldsymbol{a}} \cdot (a_\ell^k \boldsymbol{\nabla} u_\ell^k). \end{aligned} \tag{4.1}$$

Observe that (3.3)–(3.4) still holds with this definition of $\boldsymbol{\sigma}_\ell^{\boldsymbol{a},k}$. We define the data oscillation–quadrature estimator

$$(\widetilde{\eta}_{\mathrm{osc,q},\ell}^k)^2 := \sum_{\boldsymbol{a}\in\mathcal{V}_\ell} \frac{1}{\operatorname{ess\,inf}_{\omega_\ell^{\boldsymbol{a}}} a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k}} (\eta_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k})^2, \tag{4.2a}$$

with for all $\boldsymbol{a} \in \mathcal{V}_\ell$,

$$\begin{aligned} (\widetilde{\eta}_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k})^2 := \sum_{K\in\mathcal{T}_\ell^{\boldsymbol{a}}} \Bigg[\frac{h_K}{\pi} \|(I - \Pi_{\ell,p})\gamma_\ell^{\boldsymbol{a},k}\|_K \Bigg]^2 \\ + \left\| \psi_\ell^{\boldsymbol{a}} (\boldsymbol{I} - \boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}})(a_\ell^k \boldsymbol{\nabla} u_\ell^k) \right\|_{\omega_\ell^{\boldsymbol{a}}}^2. \end{aligned} \tag{4.2b}$$

4.2 A posteriori estimate of the energy difference

We now present our second main result, giving an a posteriori estimate based on the energy difference and estimator defined in (3.9). We denote, a.e. in Ω,

$$a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k} \coloneqq \min\left(a_\ell^k, \operatorname*{ess\,inf}_{r\in(\phi'^{-1}(\cdot,|\boldsymbol{\sigma}_\ell^k|),|\boldsymbol{\nabla} u_\ell^k|)} \phi''(\cdot,r)\right) \overset{(A.3),(A.5)}{\in} [a_{\mathrm{m}}, a_{\mathrm{c}}]\,, \tag{4.3a}$$

$$a_{\mathrm{c},\ell}^{\boldsymbol{\sigma},k} \coloneqq \max\left(a_\ell^k, \operatorname*{ess\,sup}_{r\in(\phi'^{-1}(\cdot,|\boldsymbol{\sigma}_\ell^k|),|\boldsymbol{\nabla} u_\ell^k|)} \phi''(\cdot,r)\right) \overset{(A.3),(A.5)}{\in} [a_{\mathrm{m}}, a_{\mathrm{c}}] \tag{4.3b}$$

and

$$a_{\mathrm{m},\ell}^{\boldsymbol{\nabla} u,k} \coloneqq \min\left(a_\ell^k, \operatorname*{ess\,inf}_{r\in(|\boldsymbol{\nabla} u|,|\boldsymbol{\nabla} u_\ell^k|)} \phi''(\cdot,r)\right) \overset{(A.3),(A.5)}{\in} [a_{\mathrm{m}}, a_{\mathrm{c}}]\,, \tag{4.4a}$$

$$a_{\mathrm{c},\ell}^{\boldsymbol{\nabla} u,k} \coloneqq \max\left(a_\ell^k, \operatorname*{ess\,sup}_{r\in(|\boldsymbol{\nabla} u|,|\boldsymbol{\nabla} u_\ell^k|)} \phi''(\cdot,r)\right) \overset{(A.3),(A.5)}{\in} [a_{\mathrm{m}}, a_{\mathrm{c}}]\,. \tag{4.4b}$$

Theorem 4.1 (A posteriori estimate of the energy difference)**.** *Suppose Assumption 2.1 and let $u \in H^1_0(\Omega)$ be the weak solution of (2.2). Let u_ℓ^k be its finite element approximation given by (2.6) on mesh $\mathcal{T}_\ell$, $\ell \geq 0$, and linearization scheme step $k \geq 1$, for any iterative linearization satisfying Assumption 2.2. Then,*

$$\mathcal{E}_{\mathrm{N},\ell}^k \leq \eta_{\mathrm{N},\ell}^k + 2\widetilde{\eta}_{\mathrm{osc},\ell}^k, \tag{4.5a}$$

$$\eta_{\mathrm{N},\ell}^k \lesssim \widetilde{C}_\ell^k \mathcal{E}_{\mathrm{N},\ell}^k + \widetilde{\eta}_{\mathrm{osc,q},\ell}^k, \tag{4.5b}$$

where the hidden constant only depends on the space dimension d, the mesh shape-regularity $\kappa_{\mathcal{T}}$, and possibly, when $d \geq 4$, the polynomial degree p, with

$$\widetilde{C}_\ell^k \coloneqq \max_{\boldsymbol{a}\in\mathcal{V}_\ell} \left(\frac{\operatorname{ess\,sup}_{\omega_\ell^{\boldsymbol{a}}} a_{\mathrm{c},\ell}^{\boldsymbol{\nabla} u,k}}{\operatorname{ess\,inf}_{\omega_\ell^{\boldsymbol{a}}} a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k}}\right)^{\frac{1}{2}} \overset{(4.3a),(4.4b)}{\leq} \left(\frac{a_{\mathrm{c}}}{a_{\mathrm{m}}}\right)^{\frac{1}{2}}. \tag{4.6}$$

Proof. See Section 7. □

Remark 4.2 (Robustness)**.** *We can make a similar statement as in Remark 3.6, since the local ratio of the functions $a_{\mathrm{c},\ell}^{\boldsymbol{\nabla} u,k}$ and $a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k}$ on each patch $\omega_\ell^{\boldsymbol{a}}$ is sharper than the global constant ratio $a_{\mathrm{c}}/a_{\mathrm{m}}$. They can again compensate each other (according to the problem and the mesh), bringing $\widetilde{C}_\ell^k$ close to 1. However, in contrast to C_ℓ^k, the constant $\widetilde{C}_\ell^k$ cannot be fully computed since $\boldsymbol{\nabla} u$ appears, and is not purely given by the linearization method, cf. Remark 3.6.*

5 Numerical results

In this section, we present numerical experiments that serve to illustrate Theorems 3.4 and 4.1. Thus, we will primarily be interested in the effectivity indices

$$I_\ell^k \coloneqq \frac{\eta_\ell^k}{\mathcal{E}_\ell^k}, \quad I_{\mathrm{N},\ell}^k \coloneqq \frac{\eta_{\mathrm{N},\ell}^k}{\mathcal{E}_{\mathrm{N},\ell}^k}, \quad I_{\mathrm{L},\ell}^k \coloneqq \frac{\eta_{\mathrm{L},\ell}^k}{\mathcal{E}_{\mathrm{L},\ell}^k}. \tag{5.1}$$

In particular, our numerical experiments study the robustness of our estimates with respect to the ratio $a_{\mathrm{c}}/a_{\mathrm{m}}$ from (2.1) where we consider $a_{\mathrm{c}}/a_{\mathrm{m}} = 10^i$, $i \in \{0,\dots,7\}$. We present results for the three linearization methods of Examples 2.4, 2.5, and 2.6 after the convergence criterion

$$\|\boldsymbol{\nabla}(u_\ell^{\bar{k}-1} - u_\ell^{\bar{k}})\| < 10^{-6} \tag{5.2}$$

is satisfied. This in particular leads to the satisfaction of (3.21).

In addition to V_ℓ^p as defined in Section 2.2, we also consider a richer discrete space $V_{\tilde{\ell}}^{\tilde{p}}$ obtained by refining the mesh $\mathcal{T}_\ell$ and using higher order polynomials, $\ell < \tilde{\ell}$ and $p < \tilde{p}$. This space serves as an approximation to $H^1_0(\Omega)$, so that we can approximately compute $u_{\langle\ell\rangle}^k$ defined in (3.10b) at each iteration k of the linearization method. This only serves here for the evaluation of the error $\mathcal{E}_{\mathrm{L},\ell}^k$; it is not needed to evaluate our estimators.

For all examples, we use the method of manufactured solutions, i.e. we choose a solution u and construct f through (1.1). The boundary conditions are then enforced by the true solution. All experiments were conducted using the `Gridap.jl` finite element software package [1, 41].

5.1 Smooth solution

We consider a unit square domain $\Omega = (0,1)^2$. We set for all $(x,y) \in \Omega$,

$$u(x,y) \coloneqq 10\,x(x-1)y(y-1). \tag{5.3}$$

For the space V_ℓ^p, we use a polynomial degree $p = 1$ and a uniform triangular mesh consisting of 8192 elements for a total of 3969 DOFs. We consider three different nonlinear functions $a : [0,\infty) \to (0,\infty)$ (independent of the spatial coordinate $\boldsymbol{x} \in \Omega$) satisfying Assumption 2.1. We first consider the following example, in which the function a is monotone (decreasing).

Example 5.1 (Mean curvature nonlinearity)**.** *The mean curvature nonlinearity (cf. [14]) is defined such that for all $r \in [0,\infty)$,*

$$a(r) \coloneqq a_{\mathrm{m}} + \frac{a_{\mathrm{c}} - a_{\mathrm{m}}}{\sqrt{1+r^2}}, \tag{5.4}$$

where $a_{\mathrm{m}}, a_{\mathrm{c}} \in (0,\infty)$ with $a_{\mathrm{m}} \le a_{\mathrm{c}}$. Observe that Assumption 2.1 holds for the mean curvature nonlinearity. Indeed, we use Proposition A.2 observing that, for all $r \in [0,\infty)$,

$$\phi''(r) = a_{\mathrm{m}} + \frac{a_{\mathrm{c}} - a_{\mathrm{m}}}{(1+r^2)^{\frac{3}{2}}} \in [a_{\mathrm{m}}, a_{\mathrm{c}}].$$

The results for the effectivity indices (5.1) are presented in Figure 2, taking $a_{\mathrm{m}} = 1$. We first remark that they vary only very mildly with respect to the ratio $a_{\mathrm{m}}/a_{\mathrm{c}}$ and that all values are below 1.2. The robustness is anticipated for Zarantonello based on Remark 3.5. The fact that the ratios are so good for the Newton and Picard linearizations is the consequence of the fact that the constant $C_\ell^{\overline{k}}$ of (3.23) is quite close to 1.

We now consider the following nonmonotone nonlinear function a similar to the example given in [31, Section 5.3.2].

Example 5.2 (Exponential nonlinearity)**.** *The exponential nonlinearity is defined such that for all $r \in [0,\infty)$,*

$$a(r) \coloneqq a_{\mathrm{m}} + (a_{\mathrm{c}} - a_{\mathrm{m}})\frac{1 - e^{-\frac{3}{2}r^2}}{1 + 2e^{-\frac{3}{2}}}, \tag{5.5}$$

where again $a_{\mathrm{m}}, a_{\mathrm{c}} \in (0,\infty)$ with $a_{\mathrm{m}} \le a_{\mathrm{c}}$. Observe that Assumption 2.1 holds for the exponential nonlinearity. Indeed, we again use Proposition A.2 observing that, for all $r \in [0,\infty)$,

$$\phi''(r) = a_{\mathrm{m}} + (a_{\mathrm{c}} - a_{\mathrm{m}})\frac{1 + (3r^2-1)e^{-\frac{3}{2}r^2}}{1 + 2e^{-\frac{3}{2}}} \in [a_{\mathrm{m}}, a_{\mathrm{c}}].$$

The results, again with $a_{\mathrm{m}} = 1$, are presented in Figure 3. The Picard linearization is not included because the solver does not converge for large values of the ratio $a_{\mathrm{c}}/a_{\mathrm{m}}$. We observe that the results for the Zarantonello iteration are similar to those of Figure 2. However, the effectivity indices of the Newton iteration seem to start to deteriorate for large values of the ratio $a_{\mathrm{c}}/a_{\mathrm{m}}$. We can see that the reason is that the constant $C_\ell^{\overline{k}}$ is becoming very large. This constant is thus here a good indicator that the robustness may not be obtained when it is large. For this example, we also present, in Figure 4, the component errors $\lambda_\ell^{\overline{k}}\mathcal{E}_{\mathrm{L},\ell}^{\overline{k}}$ and $\mathcal{E}_{\mathrm{N},\ell}^{\overline{k}}$, as well as the factor $\lambda_\ell^{\overline{k}}$. We observe that the quantities $\lambda_\ell^{\overline{k}}\mathcal{E}_{\mathrm{L},\ell}^{\overline{k}}$ and $\mathcal{E}_{\mathrm{N},\ell}^{\overline{k}}$ stay very close, independently of the ratio $a_{\mathrm{c}}/a_{\mathrm{m}}$, which was our design in (3.14). Remark, though, that $\lambda_\ell^{\overline{k}} \simeq 1$ for Newton, whereas $\lambda_\ell^{\overline{k}} \simeq (a_{\mathrm{c}}/a_{\mathrm{m}})^{\frac{1}{2}}$ for Zarantonello.

5.2 Singular solution

We consider the L-shaped domain $\Omega = (-1,1)^2 \setminus ([0,1) \times (-1,0])$ and the singular solution u in polar coordinates $(\rho,\theta) \in [0,\infty) \times [0,2\pi]$

$$u(\rho,\theta) = \rho^\alpha \sin(\alpha\theta) \tag{5.6}$$

with $\alpha \coloneqq \frac{2}{3}$, so that $u \in H^{1+\frac{2}{3}-\varepsilon}(\Omega)$ for all $\varepsilon > 0$. We consider the exponential nonlinearity of Example 5.2 again with $a_{\mathrm{m}} = 1$; this choice of solution ensures that the right-hand side f belongs to $L^2(\Omega)$ despite the singularity in the norm of the gradient for the L-shaped solution (5.6).

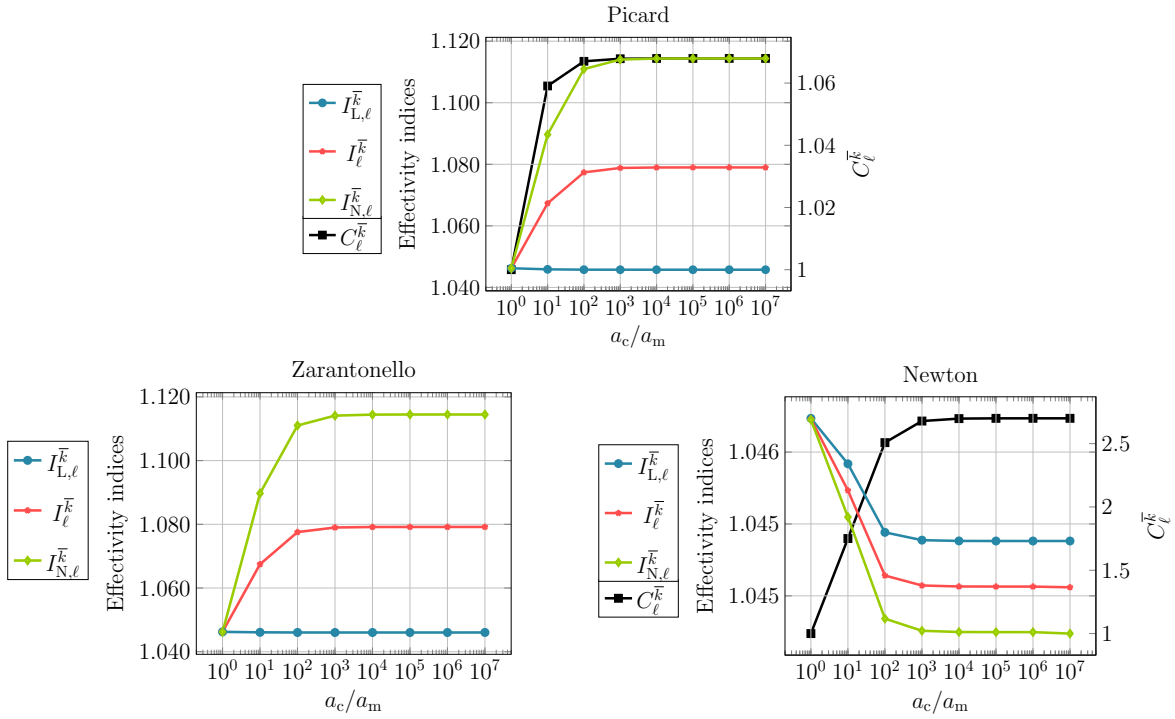


Figure 2: [Mean curvature nonlinearity (5.4), smooth solution (5.3), unit square domain, 3969 DOFs] Effectivity indices for the components and total quantities and the computable constant $C^{\overline{k}}_{\ell}$ from (3.23)

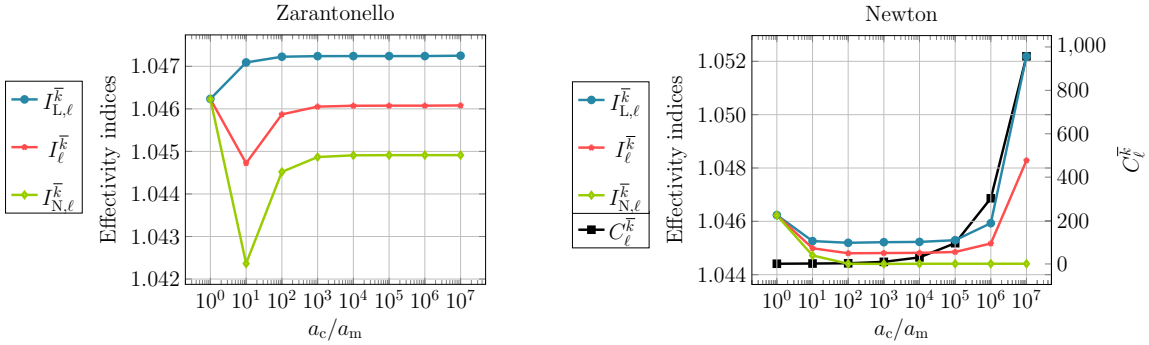


Figure 3: [Exponential nonlinearity (5.5), smooth solution (5.3), unit square domain, 3969 DOFs] Effectivity indices for the components and total quantities and the computable constant $C^{\overline{k}}_{\ell}$ from (3.23)

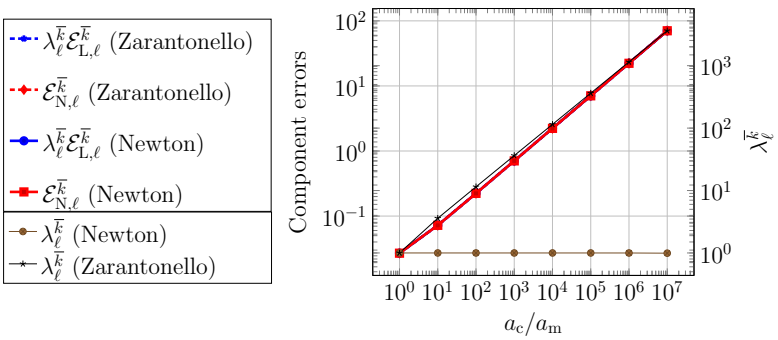
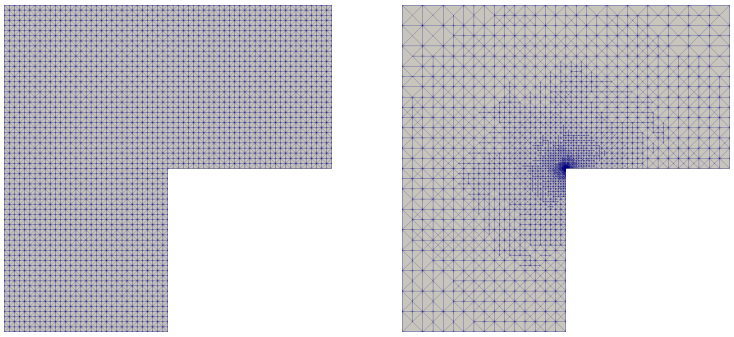


Figure 4: [Exponential nonlinearity (5.5), smooth solution (5.3), unit square domain, 3969 DOFs] Components $\lambda^{\overline{k}}_{\ell}\mathcal{E}^{\overline{k}}_{L,\ell}$ and $\mathcal{E}^{\overline{k}}_{N,\ell}$ from (3.16a) together with the weight $\lambda^{\overline{k}}_{\ell}$ from (3.14).



(a) 2945 DOFs (b) 2417 DOFs

Figure 5: Uniformly (left) and adaptively (right) refined meshes for the L-shaped domain with the singular solution (5.6). The adaptive mesh corresponds to the 28th iteration of Algorithm 5.3.

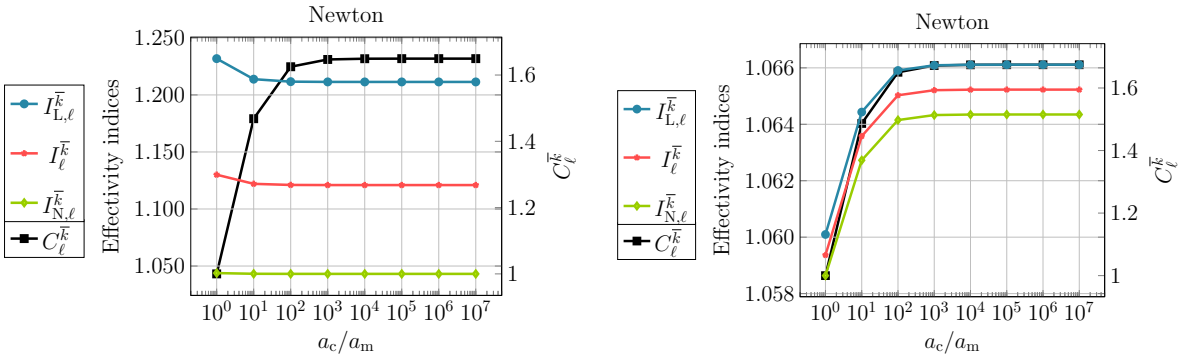


Figure 6: [Exponential nonlinearity (5.5), singular solution (5.6), L-shaped domain] Effectivity indices for the components and total quantities and the computable constant $C^{\bar{k}}_{\ell}$ from (3.23), for the uniform mesh (left) and the adaptive mesh (right) shown in Figure 5.

We consider two different meshes to analyze the results, see Figure 5. One mesh is obtained by taking an initial uniform triangulation of Ω, while the other one is adaptive following Algorithm 5.3.

The results are present in Figure 6. The Newton iteration presents, for both meshes, effectivity indices close to 1, which stabilize for large enough values of the ratio $a_{\mathrm{m}}/a_{\mathrm{c}}$. Moreover, the effectivity indices corresponding to the adaptive meshes are closer to 1 than those corresponding to the uniform meshes. Since $C^{\bar{k}}_{\ell}$ takes small values below 2, we can claim robustness a posteriori, see Remark 3.6.

5.3 Convergence on adaptively refined meshes

Due to the singularity in the solution of the previous section, it is of interest to consider a local adaptive mesh refinement strategy. We need to compute the contribution of the estimator $\eta^k_{\mathrm{N},\ell}$ to each element which is nonnegative, as discussed in, e.g., [4, Proposition 4.9]. In order to do that, we rewrite the estimator $\eta^k_{\mathrm{N},\ell}$ of (3.9b) as

$$\eta^k_{\mathrm{N},\ell} \overset{(7.4)}{=} \left(2\int_\Omega \left(\phi^*(\cdot,|\boldsymbol{\sigma}^k_\ell|)+\phi(\cdot,|\boldsymbol{\nabla} u^k_\ell|)+\boldsymbol{\sigma}^k_\ell\cdot\boldsymbol{\nabla} u^k_\ell\right)\right)^{\frac{1}{2}} = \left(\sum_{K\in\mathcal{T}_\ell}(\eta^k_{\mathrm{N},K})^2\right)^{\frac{1}{2}}, \tag{5.7a}$$

$$(\eta^k_{\mathrm{N},K})^2 := \int_K 2\left(\phi^*(\cdot,|\boldsymbol{\sigma}^k_\ell|)+\phi(\cdot,|\boldsymbol{\nabla} u^k_\ell|)+\boldsymbol{\sigma}^k_\ell\cdot\boldsymbol{\nabla} u^k_\ell\right), \quad K\in\mathcal{T}_\ell. \tag{5.7b}$$

Here, recalling the generalized Young inequality, cf. [32],

$$\phi(\cdot,|\boldsymbol{x}|)+\phi^*(\cdot,|\boldsymbol{y}|)+\boldsymbol{x}\cdot\boldsymbol{y}\geq 0 \qquad \forall \boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d, \tag{5.8}$$

it follows that for all $K\in\mathcal{T}_\ell$, $\eta^k_{\mathrm{N},K}\geq 0$. We then use the standard newest vertex bisection algorithm, see [9] and the references therein, as follows.

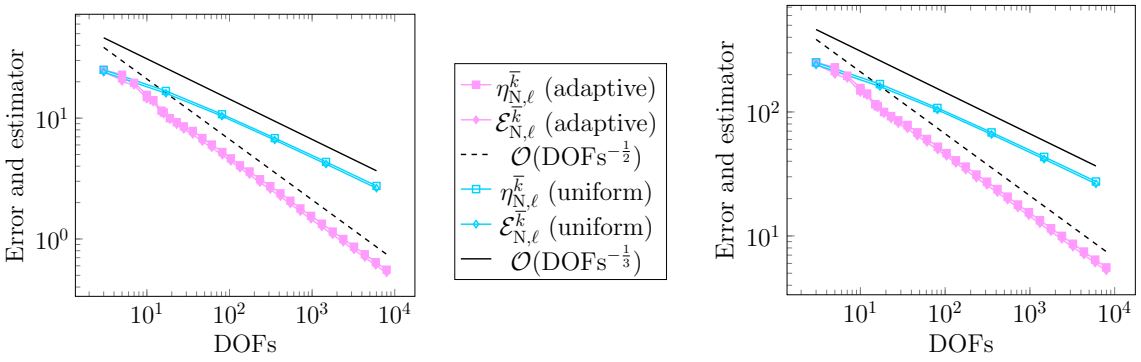


Figure 7: [Exponential nonlinearity (5.5), singular solution (5.6), L-shaped domain, uniform vs adaptive mesh refinement] Convergence rates of $\mathcal{E}^{\bar{k}}_{\mathrm{N},\ell}$ and $\eta^{\bar{k}}_{\mathrm{N},\ell}$ for uniform and adaptive mesh refinement for $a_{\mathrm{c}}/a_{\mathrm{m}}$ equal to 10^3 (left) and 10^6 (right).

Algorithm 5.3 (Adaptive refinement)**.** *Let $\varepsilon_{\mathrm{STOP}}$ and $\theta \in (0,1)$ be parameters, and let $\mathcal{T}_0$ be a conforming initial triangulation of Ω. Let $u_0^0 \in V_0^1$ be an initial linearization guess. For $\ell \geq 0$:*

1. ***Solve**: Starting from u_ℓ^0, solve the linearized problems (2.6) until the convergence criterion (5.2) is satisfied.*
2. ***Estimate**: Compute the elementwise estimators $(\eta^{\bar{k}}_{\mathrm{N},K})_{K\in\mathcal{T}_\ell}$ of (5.7b).*
   If $\eta^{\bar{k}}_{\mathrm{N},\ell} < \varepsilon_{\mathrm{STOP}}$, then stop.
3. ***Mark**: Choose a set $\mathcal{M}_\ell \subset \mathcal{T}_\ell$ with minimal cardinality such that*
$$\sum_{K\in\mathcal{M}_\ell} (\eta^{\bar{k}}_{\mathrm{N},K})^2 \geq \theta^2 \sum_{K\in\mathcal{T}_\ell} (\eta^{\bar{k}}_{\mathrm{N},K})^2. \tag{5.9}$$
4. ***Refine**: Perform the newest vertex bisection. Set $\ell := \ell+1$, $u_\ell^0 := u^{\bar{k}}_{\ell-1}$, and go to 1.*

The results of the refinement study are displayed in Figure 7, for the exponential nonlinearity (5.5) and the singular solution (5.6). We consider two values of the parameter $a_{\mathrm{c}}/a_{\mathrm{m}}$, namely 10^3 and 10^6. We note that for both values of the ratio, the asymptotic rates for the estimator and error agree with the theoretical optimal rate of $(\mathrm{DOFs})^{1/2}$ for the adaptive algorithm; we observe no distinguishable difference in this graphic representation between $a_{\mathrm{c}}/a_{\mathrm{m}} = 10^3$ and 10^6.

We also made an analogous study on the augmented error $\mathcal{E}^{\bar{k}}_\ell$ and estimator $\eta^{\bar{k}}_\ell$, with the same strategy of refinement, using the local version $\eta^{\bar{k}}_K := \eta^{\bar{k}}_{\mathrm{N},K} + \lambda^{\bar{k}}_\ell \eta^{\bar{k}}_{\mathrm{L},K}$, for all $K \in \mathcal{T}_\ell$, of the estimator $\eta^{\bar{k}}_\ell$. The results are displayed in Figure 8. We observe a similar behavior of the asymptotic rates as in Figure 7.

In conclusion, the adaptive mesh refinement is more efficient than the uniform mesh refinement since it requires a smaller number of DOFs for the same precision. The behavior seems to be independent of the strength of the nonlinearity $a_{\mathrm{c}}/a_{\mathrm{m}}$.

6 Proof of Theorem 3.4

For all vertices $\boldsymbol{a} \in \mathcal{V}_\ell$, we introduce the space $H^1_*(\omega_\ell^{\boldsymbol{a}})$ such that

$$H^1_*(\omega_\ell^{\boldsymbol{a}}) := \begin{cases} \{\varphi \in H^1(\omega_\ell^{\boldsymbol{a}}) : (\varphi,1)_{\omega_\ell^{\boldsymbol{a}}} = 0\} & \text{if } \boldsymbol{a} \in \mathcal{V}_\ell^{\mathrm{int}}, \\ \{\varphi \in H^1(\omega_\ell^{\boldsymbol{a}}) : \varphi_{|\partial\omega_\ell^{\boldsymbol{a}} \cap \{\psi_\ell^{\boldsymbol{a}} > 0\}} = 0\} & \text{if } \boldsymbol{a} \notin \mathcal{V}_\ell^{\mathrm{int}}, \end{cases} \tag{6.1}$$

where $\mathcal{V}_\ell^{\mathrm{int}}$ is the set of the vertices of $\mathcal{T}_\ell$ lying inside Ω. We will need:

Lemma 6.1 (Patchwise flux equilibration stability)**.** *Let $\boldsymbol{a} \in \mathcal{V}_\ell$ and let $\boldsymbol{\tau}_\ell^{\boldsymbol{a}} \in \boldsymbol{RTN}_p(\mathcal{T}_\ell^{\boldsymbol{a}})$ and $g_\ell^{\boldsymbol{a}} \in \mathcal{P}_p(\mathcal{T}_\ell^{\boldsymbol{a}})$ be such that $(g_\ell^{\boldsymbol{a}},1)_{\omega_\ell^{\boldsymbol{a}}} = 0$ if $\boldsymbol{a} \in \mathcal{V}_\ell^{\mathrm{int}}$. Then,*

$$\min_{\substack{\boldsymbol{v}_\ell^{\boldsymbol{a}} \in \boldsymbol{V}_\ell^{\boldsymbol{a}} \\ \boldsymbol{\nabla}\cdot\boldsymbol{v}_\ell^{\boldsymbol{a}} = g_\ell^{\boldsymbol{a}}}} \|\boldsymbol{v}_\ell^{\boldsymbol{a}} + \boldsymbol{\tau}_\ell^{\boldsymbol{a}}\|_{\omega_\ell^{\boldsymbol{a}}} \lesssim \sup_{\substack{\varphi \in H^1_*(\omega_\ell^{\boldsymbol{a}}) \\ \|\boldsymbol{\nabla}\varphi\|_{\omega_\ell^{\boldsymbol{a}}} = 1}} \left[(g_\ell^{\boldsymbol{a}},\varphi)_{\omega_\ell^{\boldsymbol{a}}} - (\boldsymbol{\tau}_\ell^{\boldsymbol{a}}, \boldsymbol{\nabla}\varphi)_{\omega_\ell^{\boldsymbol{a}}}\right], \tag{6.2}$$

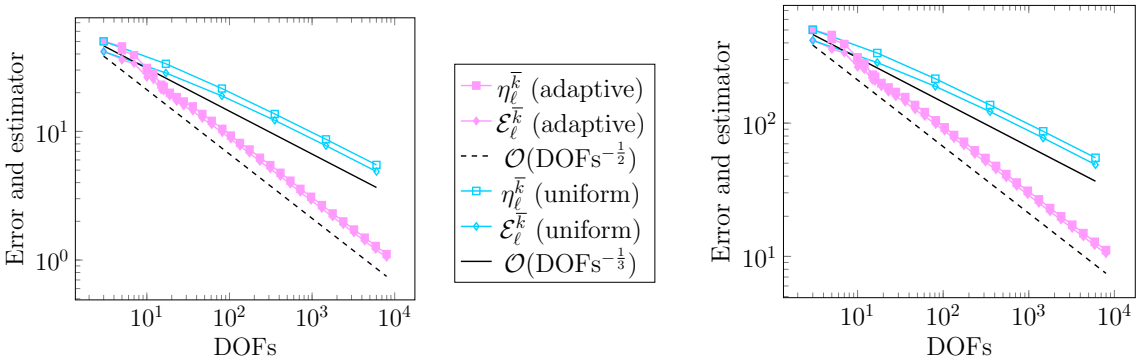


Figure 8: [Exponential nonlinearity (5.5), singular solution (5.6), L-shaped domain, uniform vs adaptive mesh refinement] Convergence rates of $\mathcal{E}_\ell^k$ and η_ℓ^k for uniform and adaptive mesh refinement for $a_{\mathrm{c}}/a_{\mathrm{m}}$ equal to 10^3 (left) and 10^6 (right).

where the hidden constant only depends on the space dimension d, the mesh shape-regularity $\kappa_{\mathcal{T}}$, and possibly, when $d \geq 4$, the polynomial degree p.

Proof. See [21, Corollary 3.3] or [19, Lemma 3.2]. □

We will use next the fact that, for all $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega)$ and $\boldsymbol{W} \in [L^2(\Omega)]^{d\times d}$,

$$\sum_{a\in\mathcal{V}_\ell} \|\boldsymbol{v}\|_{\omega_\ell^{\boldsymbol{a}}}^2 = \sum_{K\in\mathcal{T}_\ell} \sum_{a\in\mathcal{V}_K} \|\boldsymbol{v}\|_K^2 = (d+1) \sum_{K\in\mathcal{T}_\ell} \|\boldsymbol{v}\|_K^2 = (d+1)\|\boldsymbol{v}\|^2, \tag{6.3}$$

since every simplex $K \in \mathcal{T}_\ell$ has $(d+1)$ vertices, collected in the set $\mathcal{V}_K$, and

$$\begin{aligned} \|\boldsymbol{v} + \boldsymbol{W}\boldsymbol{\sigma}_\ell^k\|^2 &= \sum_{K\in\mathcal{T}_\ell} \|\boldsymbol{v} + \boldsymbol{W}\boldsymbol{\sigma}_\ell^k\|_K^2 \overset{(3.1),(3.2a)}{=} \sum_{K\in\mathcal{T}_\ell} \Big\| \sum_{a\in\mathcal{V}_K} (\psi_\ell^{\boldsymbol{a}}\boldsymbol{v} + \boldsymbol{W}\boldsymbol{\sigma}_\ell^{\boldsymbol{a},k}) \Big\|_K^2 \\ &\leq (d+1) \sum_{K\in\mathcal{T}_\ell} \sum_{a\in\mathcal{V}_K} \|\psi_\ell^{\boldsymbol{a}}\boldsymbol{v} + \boldsymbol{W}\boldsymbol{\sigma}_\ell^{\boldsymbol{a},k}\|_K^2 = (d+1) \sum_{a\in\mathcal{V}_\ell} \|\psi_\ell^{\boldsymbol{a}}\boldsymbol{v} + \boldsymbol{W}\boldsymbol{\sigma}_\ell^{\boldsymbol{a},k}\|_{\omega_\ell^{\boldsymbol{a}}}^2. \end{aligned} \tag{6.4}$$

Recalling the definition (2.7) of $\mathcal{J}_\ell^{k-1}$, we can rewrite $\mathcal{E}_{\mathrm{L},\ell}^k$ given by (3.11a) as

$$\begin{aligned} (\mathcal{E}_{\mathrm{L},\ell}^k)^2 &\overset{(2.7)}{=} \Big\|(\boldsymbol{A}_\ell^{k-1})^{\frac12}\boldsymbol{\nabla}u_\ell^k\Big\|^2 - \Big\|(\boldsymbol{A}_\ell^{k-1})^{\frac12}\boldsymbol{\nabla}u_{\langle\ell\rangle}^k\Big\|^2 \\ &\qquad - 2(\boldsymbol{b}_\ell^{k-1}, \boldsymbol{\nabla}(u_\ell^k - u_{\langle\ell\rangle}^k)) - 2(f, u_\ell^k - u_{\langle\ell\rangle}^k) \\ &\overset{(3.10a)}{=} \Big\|(\boldsymbol{A}_\ell^{k-1})^{\frac12}\boldsymbol{\nabla}u_\ell^k\Big\|^2 - \Big\|(\boldsymbol{A}_\ell^{k-1})^{\frac12}\boldsymbol{\nabla}u_{\langle\ell\rangle}^k\Big\|^2 - 2(\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}u_{\langle\ell\rangle}^k, \boldsymbol{\nabla}(u_\ell^k - u_{\langle\ell\rangle}^k)) \\ &= \|(\boldsymbol{A}_\ell^{k-1})^{\frac12}\boldsymbol{\nabla}(u_\ell^k - u_{\langle\ell\rangle}^k)\|^2 \overset{(6.3)}{=} \frac{1}{d+1} \sum_{\boldsymbol{a}\in\mathcal{V}_\ell} (\mathcal{E}_{\mathrm{L},\ell}^{\boldsymbol{a},k})^2, \end{aligned} \tag{6.5a}$$

where, for all $\boldsymbol{a} \in \mathcal{V}_\ell$,

$$\mathcal{E}_{\mathrm{L},\ell}^{\boldsymbol{a},k} \coloneqq \|(\boldsymbol{A}_\ell^{k-1})^{\frac12}\boldsymbol{\nabla}(u_\ell^k - u_{\langle\ell\rangle}^k)\|_{\omega_\ell^{\boldsymbol{a}}}. \tag{6.5b}$$

Similarly, using the definition (3.8) of $\mathcal{J}_\ell^{*,k-1}$, we can upper-bound $\eta_{\mathrm{L},\ell}^k$ of (3.11b),

$$\begin{aligned} (\eta_{\mathrm{L},\ell}^k)^2 &\overset{(3.8)}{=} \Big\|(\boldsymbol{A}_\ell^{k-1})^{\frac12}\boldsymbol{\nabla}u_\ell^k\Big\|^2 + \Big\|(\boldsymbol{A}_\ell^{k-1})^{-\frac12}(\boldsymbol{\sigma}_\ell^k - \boldsymbol{b}_\ell^{k-1})\Big\|^2 - 2(\boldsymbol{b}_\ell^{k-1}, \boldsymbol{\nabla}u_\ell^k) - 2(f, u_\ell^k) \\ &= \|(\boldsymbol{A}_\ell^{k-1})^{-\frac12}(\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}u_\ell^k - \boldsymbol{b}_\ell^{k-1} + \boldsymbol{\sigma}_\ell^k)\|^2 - 2(\boldsymbol{\sigma}_\ell^k, \boldsymbol{\nabla}u_\ell^k) - 2(f, u_\ell^k) \\ &\overset{(3.4)}{=} \|(\boldsymbol{A}_\ell^{k-1})^{-\frac12}(\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}u_\ell^k - \boldsymbol{b}_\ell^{k-1} + \boldsymbol{\sigma}_\ell^k)\|^2 \overset{(6.4)}{\leq} (d+1) \sum_{\boldsymbol{a}\in\mathcal{V}_\ell} (\eta_{\mathrm{L},\ell}^{\boldsymbol{a},k})^2, \end{aligned} \tag{6.6a}$$

where, for all $\boldsymbol{a} \in \mathcal{V}_\ell$,

$$\eta_{\mathrm{L},\ell}^{\boldsymbol{a},k} \coloneqq \|(\boldsymbol{A}_\ell^{k-1})^{-\frac12}(\psi_\ell^{\boldsymbol{a}}(\boldsymbol{A}_\ell^{k-1}\boldsymbol{\nabla}u_\ell^k - \boldsymbol{b}_\ell^{k-1}) + \boldsymbol{\sigma}_\ell^{\boldsymbol{a},k})\|_{\omega_\ell^{\boldsymbol{a}}}. \tag{6.6b}$$

In the following, for any $x, y \in [0,\infty)$, we use the notation $x \lesssim y$ when $x \leq Cy$ with $C \geq 0$ only depending on the space dimension d, the mesh shape-regularity $\kappa_{\mathcal{T}}$, and possibly, when $d \geq 4$, the polynomial degree p.

Proof of Theorem 3.4. Let $k \geq 1$. Proceeding as in [19, 20], we have $\mathcal{E}_{\mathrm{L},\ell}^k \leq \eta_{\mathrm{L},\ell}^k + \eta_{\mathrm{osc},\ell}^k$. Furthermore, as in [2, 39, 43] but including data oscillations, we have

$$\begin{aligned}(\mathcal{E}_{\mathrm{N},\ell}^k)^2 &\leq (\eta_{\mathrm{N},\ell}^k)^2 + 2(f - \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}_\ell^k, u - u_\ell^k)\\ &\overset{(3.3)}{\leq} (\eta_{\mathrm{N},\ell}^k)^2 + 2\left(\sum_{K\in\mathcal{T}_\ell}\left[\frac{h_K}{\pi}\|f - \Pi_{\ell,p}f\|_K\right]^2\right)^{\frac12}\|\boldsymbol{\nabla}(u-u_\ell^k)\|\\ &\overset{(3.17),(3.19a)}{\leq} (\eta_{\mathrm{N},\ell}^k)^2 + 2\widetilde{\eta}_{\mathrm{osc},\ell}^k\mathcal{E}_{\mathrm{N},\ell}^k,\end{aligned} \tag{6.7}$$

which gives by the quadratic formula,

$$\mathcal{E}_{\mathrm{N},\ell}^k \leq \frac12\left(2\widetilde{\eta}_{\mathrm{osc},\ell}^k + (4(\widetilde{\eta}_{\mathrm{osc},\ell}^k)^2 + 4(\eta_{\mathrm{N},\ell}^k)^2)^{\frac12}\right) \leq \eta_{\mathrm{N},\ell}^k + 2\widetilde{\eta}_{\mathrm{osc},\ell}^k. \tag{6.8}$$

Hence, we obtain (3.20) by writing

$$\begin{aligned}\mathcal{E}_\ell^k &\overset{(3.16a)}{=} \frac12(\mathcal{E}_{\mathrm{N},\ell}^k + \lambda_\ell^k\mathcal{E}_{\mathrm{L},\ell}^k)\\ &\overset{(6.8)}{\leq} \frac12(\eta_{\mathrm{N},\ell}^k + 2\widetilde{\eta}_{\mathrm{osc},\ell}^k + \lambda_\ell^k(\eta_{\mathrm{L},\ell}^k + \eta_{\mathrm{osc},\ell}^k)) \overset{(3.16b)}{=} \eta_\ell^k + \widetilde{\eta}_{\mathrm{osc},\ell}^k + \frac{\lambda_\ell^k}{2}\eta_{\mathrm{osc},\ell}^k.\end{aligned} \tag{6.9}$$

It remains to prove (3.22). Recalling (3.2b), we define for all vertices $\boldsymbol{a} \in \mathcal{V}_\ell$,

$$\widetilde{\boldsymbol{\sigma}}_\ell^{\boldsymbol{a},k} \coloneqq \arg\min_{\substack{\boldsymbol{w}_\ell\in\boldsymbol{V}_\ell^{\boldsymbol{a}}\\ \boldsymbol{\nabla}\cdot\boldsymbol{w}_\ell = \Pi_{\ell,p}\gamma_\ell^{\boldsymbol{a},k}}}\|\psi_\ell^{\boldsymbol{a}}\boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}}\boldsymbol{\xi}_\ell^k + \boldsymbol{w}_\ell\|_{\omega_\ell^{\boldsymbol{a}}}. \tag{6.10}$$

We next use Lemma 6.1 (applied with $\boldsymbol{\tau}_\ell^{\boldsymbol{a}} = \psi_\ell^{\boldsymbol{a}}\boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}}\boldsymbol{\xi}_\ell^k \in \boldsymbol{RTN}_p(\mathcal{T}_\ell^{\boldsymbol{a}})$ and $g_\ell^{\boldsymbol{a}} = \Pi_{\ell,p}\gamma_\ell^{\boldsymbol{a},k} \in \mathcal{P}_p(\mathcal{T}_\ell^{\boldsymbol{a}})$ with $(g_\ell^{\boldsymbol{a}},1)_{\omega_\ell^{\boldsymbol{a}}} = 0$ if $\boldsymbol{a} \in \mathcal{V}_\ell^{\mathrm{int}}$) to infer

$$\begin{aligned}\eta_{\mathrm{L},\ell}^{\boldsymbol{a},k} &\overset{(3.19e)}{\leq} \|(\boldsymbol{A}_\ell^{k-1})^{-\frac12}(\psi_\ell^{\boldsymbol{a}}\boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}}\boldsymbol{\xi}_\ell^k + \boldsymbol{\sigma}_\ell^{\boldsymbol{a},k})\|_{\omega_\ell^{\boldsymbol{a}}} + \eta_{\mathrm{q},\ell}^{\boldsymbol{a},k}\\ &\overset{(3.2b),(2.8b)}{\leq} \frac{1}{(\inf_{\omega_\ell^{\boldsymbol{a}}}A_{\mathrm{m},\ell}^{k-1})^{\frac12}}\|\psi_\ell^{\boldsymbol{a}}\boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}}\boldsymbol{\xi}_\ell^k + \widetilde{\boldsymbol{\sigma}}_\ell^{\boldsymbol{a},k}\|_{\omega_\ell^{\boldsymbol{a}}} + \eta_{\mathrm{q},\ell}^{\boldsymbol{a},k}\\ &\overset{(6.2),(3.23)}{\lesssim} \frac{1}{(\inf_{\omega_\ell^{\boldsymbol{a}}}A_{\mathrm{m},\ell}^{k-1})^{\frac12}}\sup_{\substack{\varphi\in H_*^1(\omega_\ell^{\boldsymbol{a}})\\ \|\boldsymbol{\nabla}\varphi\|_{\omega_\ell^{\boldsymbol{a}}}=1}}\left[(f,\psi_\ell^{\boldsymbol{a}}\varphi)_{\omega_\ell^{\boldsymbol{a}}} - (\boldsymbol{\xi}_\ell^k, \boldsymbol{\nabla}(\psi_\ell^{\boldsymbol{a}}\varphi))_{\omega_\ell^{\boldsymbol{a}}}\right]\\ &\qquad + (1+C_\ell^k)\eta_{\mathrm{q},\ell}^{\boldsymbol{a},k} + \eta_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k}\\ &\overset{(3.10a),(6.5b)}{\lesssim} C_\ell^k(\mathcal{E}_{\mathrm{L},\ell}^{\boldsymbol{a},k} + \eta_{\mathrm{q},\ell}^{\boldsymbol{a},k}) + \eta_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k}.\end{aligned} \tag{6.11}$$

Here, we have used in the third line the fact that $\psi_\ell^{\boldsymbol{a}}\varphi \in H_0^1(\omega_\ell^{\boldsymbol{a}})$ together with $\boldsymbol{\nabla}(\psi_\ell^{\boldsymbol{a}}\varphi) = \psi_\ell^{\boldsymbol{a}}\boldsymbol{\nabla}\varphi + \varphi\boldsymbol{\nabla}\psi_\ell^{\boldsymbol{a}}$ for $\varphi \in H_*^1(\omega_\ell^{\boldsymbol{a}})$, and concluded with $\|\boldsymbol{\nabla}(\psi_\ell^{\boldsymbol{a}}\varphi)\|_{\omega_\ell^{\boldsymbol{a}}} \lesssim \|\boldsymbol{\nabla}\varphi\|_{\omega_\ell^{\boldsymbol{a}}}$ as in [8, 19]. Therefore, we obtain (3.22) by writing

$$\begin{aligned}\eta_\ell^k &\overset{(3.14),(3.16b)}{=} \frac12\left(\frac{\widehat{\eta}_{\mathrm{L},\ell}^k}{\eta_{\mathrm{L},\ell}^k} + 1\right)\lambda_\ell^k\eta_{\mathrm{L},\ell}^k \overset{(3.21)}{\leq} \frac32\lambda_\ell^k\eta_{\mathrm{L},\ell}^k \overset{(6.6)}{\lesssim} \lambda_\ell^k\left(\sum_{\boldsymbol{a}\in\mathcal{V}_\ell}(\eta_{\mathrm{L},\ell}^{\boldsymbol{a},k})^2\right)^{\frac12}\\ &\overset{(6.11)}{\lesssim} \lambda_\ell^k\left[C_\ell^k\left(\sum_{\boldsymbol{a}\in\mathcal{V}_\ell}(\mathcal{E}_{\mathrm{L},\ell}^{\boldsymbol{a},k})^2\right)^{\frac12} + C_\ell^k\left(\sum_{\boldsymbol{a}\in\mathcal{V}_\ell}(\eta_{\mathrm{q},\ell}^{\boldsymbol{a},k})^2\right)^{\frac12} + \left(\sum_{\boldsymbol{a}\in\mathcal{V}_\ell}(\eta_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k})^2\right)^{\frac12}\right]\\ &\overset{(6.5),(3.19)}{\lesssim} \lambda_\ell^k\left[\frac12C_\ell^k\mathcal{E}_{\mathrm{L},\ell}^k + C_\ell^k\eta_{\mathrm{q},\ell}^k + \eta_{\mathrm{osc,q},\ell}^k\right] \overset{(3.16a)}{\leq} C_\ell^k\mathcal{E}_\ell^k + \lambda_\ell^k(C_\ell^k\eta_{\mathrm{q},\ell}^k + \eta_{\mathrm{osc,q},\ell}^k),\end{aligned} \tag{6.12}$$

where we used the triangle inequality on $\ell_2(\mathbb{R}^{|\mathcal{V}_\ell|})$. Finally, we get $C_\ell^k = 1$ for the Zarantonello iteration since $A_{\mathrm{m},\ell}^{k-1} = A_{\mathrm{c},\ell}^{k-1} = \gamma$ in Ω in this case. □

7 Proof of Theorem 4.1

In addition to (4.3)–(4.4), let

$$a_{\mathrm{m},\ell}^{u,k} \coloneqq \min\left(a_u, \operatorname*{ess\,inf}_{r\in(|\nabla u|,|\nabla u_\ell^k|)} \phi''(\cdot,r)\right) \overset{(A.3),(A.5)}{\in} [a_{\mathrm{m}}, a_{\mathrm{c}}]\,, \tag{7.1a}$$

$$a_{\mathrm{c},\ell}^{u,k} \coloneqq \max\left(a_u, \operatorname*{ess\,sup}_{r\in(|\nabla u|,|\nabla u_\ell^k|)} \phi''(\cdot,r)\right) \overset{(A.3),(A.5)}{\in} [a_{\mathrm{m}}, a_{\mathrm{c}}]\,. \tag{7.1b}$$

We first state the following result.

Lemma 7.1 (Local bounds for the energy difference and estimator)**.** *We have*

$$\|(a_{\mathrm{m},\ell}^{u,k})^{\frac12}(\nabla u_\ell^k - \nabla u)\|^2 \le (\mathcal{E}_{\mathrm{N},\ell}^k)^2 \le \|(a_{\mathrm{c},\ell}^{u,k})^{\frac12}(\nabla u_\ell^k - \nabla u)\|^2, \tag{7.2a}$$

$$\|(a_{\mathrm{c},\ell}^{\boldsymbol{\sigma},k})^{-\frac12}(a_\ell^k\nabla u_\ell^k + \boldsymbol{\sigma}_\ell^k)\|^2 \le (\eta_{\mathrm{N},\ell}^k)^2 \le \|(a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k})^{-\frac12}(a_\ell^k\nabla u_\ell^k + \boldsymbol{\sigma}_\ell^k)\|^2, \tag{7.2b}$$

$$\|(a_{\mathrm{c},\ell}^{\nabla u,k})^{-\frac12}(a_\ell^k\nabla u_\ell^k - a_u\nabla u)\|^2 \le (\mathcal{E}_{\mathrm{N},\ell}^k)^2 \le 2\|(a_{\mathrm{m},\ell}^{\nabla u,k})^{-\frac12}(a_\ell^k\nabla u_\ell^k - a_u\nabla u)\|^2. \tag{7.2c}$$

Proof. Observing that by integration by parts (IBP), $\frac{\partial}{\partial r}((|\boldsymbol{\sigma}_\ell^k|-r)\phi^{*\prime}(\cdot,r)) = (|\boldsymbol{\sigma}_\ell^k|-r)\phi^{*\prime\prime}(\cdot,r) - \phi^{*\prime}(\cdot,r)$, we obtain, a.e. in Ω,

$$\begin{aligned}
\phi^*(\cdot,|\boldsymbol{\sigma}_\ell^k|) - \phi^*(\cdot,a_\ell^k|\nabla u_\ell^k|) &= \int_{a_\ell^k|\nabla u_\ell^k|}^{|\boldsymbol{\sigma}_\ell^k|} \phi^{*\prime}(\cdot,r)\,\mathrm{d}r \\
&\overset{\text{(IBP)}}{=} \int_{a_\ell^k|\nabla u_\ell^k|}^{|\boldsymbol{\sigma}_\ell^k|} (|\boldsymbol{\sigma}_\ell^k|-r)\phi^{*\prime\prime}(\cdot,r)\,\mathrm{d}r - \Big[(|\boldsymbol{\sigma}_\ell^k|-r)\phi^{*\prime}(\cdot,r)\Big]_{a_\ell^k|\nabla u_\ell^k|}^{|\boldsymbol{\sigma}_\ell^k|} \\
&\overset{(3.7)}{=} \int_{a_\ell^k|\nabla u_\ell^k|}^{|\boldsymbol{\sigma}_\ell^k|} \frac{|\boldsymbol{\sigma}_\ell^k|-r}{\phi''(\cdot,\phi'^{-1}(\cdot,r))}\,\mathrm{d}r + (|\boldsymbol{\sigma}_\ell^k| - a_\ell^k|\nabla u_\ell^k|)|\nabla u_\ell^k|,
\end{aligned} \tag{7.3}$$

where we used the fact that $\phi'^{-1}(\cdot,a_\ell^k|\nabla u_\ell^k|) = \phi'^{-1}(\cdot,\phi'(\cdot,|\nabla u_\ell^k|)) = |\nabla u_\ell^k|$ thanks to (A.2). Therefore, recalling the definitions (2.3) of $\mathcal{J}$ and (3.5) of $\mathcal{J}^*$, we obtain the right-hand side of (7.2b) by writing

$$\begin{aligned}
(\eta_{\mathrm{N},\ell}^k)^2 &\overset{(3.9b)}{=} 2\int_\Omega \big(\phi^*(\cdot,|\boldsymbol{\sigma}_\ell^k|) + \phi(\cdot,|\nabla u_\ell^k|)\big) - (f,u_\ell^k) \\
&\overset{(3.4)}{=} 2\int_\Omega \big(\phi^*(\cdot,|\boldsymbol{\sigma}_\ell^k|) + \phi(\cdot,|\nabla u_\ell^k|) + \boldsymbol{\sigma}_\ell^k\cdot\nabla u_\ell^k\big) \\
&\overset{(3.6b)}{=} 2\int_\Omega \big(\phi^*(\cdot,|\boldsymbol{\sigma}_\ell^k|) - \phi^*(\cdot,a_\ell^k|\nabla u_\ell^k|) + a_\ell^k|\nabla u_\ell^k|^2 + \boldsymbol{\sigma}_\ell^k\cdot\nabla u_\ell^k\big) \\
&\overset{(7.3)}{=} 2\int_\Omega \left[\int_{a_\ell^k|\nabla u_\ell^k|}^{|\boldsymbol{\sigma}_\ell^k|} \frac{|\boldsymbol{\sigma}_\ell^k|-r}{\phi''(\cdot,\phi'^{-1}(\cdot,r))}\,\mathrm{d}r + \frac{1}{a_\ell^k}(|\boldsymbol{\sigma}_\ell^k||a_\ell^k\nabla u_\ell^k| + \boldsymbol{\sigma}_\ell^k\cdot(a_\ell^k\nabla u_\ell^k))\right] \\
&\overset{(4.3a)}{\le} \int_\Omega \frac{2}{a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k}}\left(\left[\frac12(|\boldsymbol{\sigma}_\ell^k|-r)^2\right]_{a_\ell^k|\nabla u_\ell^k|}^{|\boldsymbol{\sigma}_\ell^k|} + |\boldsymbol{\sigma}_\ell^k||a_\ell^k\nabla u_\ell^k| + \boldsymbol{\sigma}_\ell^k\cdot(a_\ell^k\nabla u_\ell^k)\right) \\
&= \int_\Omega \frac{1}{a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k}}\left[(|\boldsymbol{\sigma}_\ell^k| - a_\ell^k|\nabla u_\ell^k|)^2 + 2(|\boldsymbol{\sigma}_\ell^k||a_\ell^k\nabla u_\ell^k| + \boldsymbol{\sigma}_\ell^k\cdot(a_\ell^k\nabla u_\ell^k))\right] \\
&\overset{(A.4a)}{=} \int_\Omega \frac{1}{a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k}}|a_\ell^k\nabla u_\ell^k + \boldsymbol{\sigma}_\ell^k|^2,
\end{aligned} \tag{7.4}$$

where, in using (3.6b), we have set $s = a_\ell^k|\nabla u_\ell^k|$, and where we used (4.3a) noticing that $|\boldsymbol{\sigma}_\ell^k||a_\ell^k\nabla u_\ell^k| + \boldsymbol{\sigma}_\ell^k\cdot(a_\ell^k\nabla u_\ell^k) \ge 0$ thanks to the Cauchy–Schwarz inequality, together with the fact that $|\boldsymbol{\sigma}_\ell^k| - r$ is of the same sign as $|\boldsymbol{\sigma}_\ell^k| - a_\ell^k|\nabla u_\ell^k|$ for all $r \in (a_\ell^k|\nabla u_\ell^k|, |\boldsymbol{\sigma}_\ell^k|)$. The left-hand side of (7.2b) is obtained with the same reasoning.

Now, observing that $\frac{\partial}{\partial r}((|\nabla u_\ell^k| - r)\phi'(\cdot,r)) = (|\nabla u_\ell^k| - r)\phi''(\cdot,r) - \phi'(\cdot,r)$ by integration by parts

(IBP), we get

$$
\begin{aligned}
(\mathcal{E}_{\mathrm{N},\ell}^k)^2 &\overset{(3.9a),(2.3)}{=} 2\int_\Omega (\phi(\cdot,|\boldsymbol{\nabla} u_\ell^k|)-\phi(\cdot,|\boldsymbol{\nabla} u|)) - 2(f,u_\ell^k-u)\\
&\overset{(2.2a)}{=} 2\int_\Omega \int_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|} \phi'(\cdot,r)\,\mathrm{d}r - 2(a_u\boldsymbol{\nabla} u, \boldsymbol{\nabla}(u_\ell^k-u))\\
&\overset{(\mathrm{IBP})}{=} 2\int_\Omega \left(\int_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|} (|\boldsymbol{\nabla} u_\ell^k|-r)\phi''(\cdot,r)\,\mathrm{d}r - \Big[(|\boldsymbol{\nabla} u_\ell^k|-r)\phi'(\cdot,r)\Big]_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|}\right)\\
&\qquad - 2(a_u\boldsymbol{\nabla} u, \boldsymbol{\nabla}(u_\ell^k-u))\\
&\overset{(A.2)}{=} 2\int_\Omega \left(\int_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|} (|\boldsymbol{\nabla} u_\ell^k|-r)\phi''(\cdot,r)\,\mathrm{d}r + a_u(|\boldsymbol{\nabla} u_\ell^k||\boldsymbol{\nabla} u| - \boldsymbol{\nabla} u_\ell^k\cdot\boldsymbol{\nabla} u)\right).
\end{aligned}
\tag{7.5}
$$

From there, with the same reasoning as for (7.2b), we obtain the right-hand side of (7.2a), by writing

$$
\begin{aligned}
(\mathcal{E}_{\mathrm{N},\ell}^k)^2 &\overset{(7.5),(7.1)}{\leq} \int_\Omega a_{\mathrm{c},\ell}^{u,k}\left((|\boldsymbol{\nabla} u_\ell^k|-|\boldsymbol{\nabla} u|)^2 + 2(|\boldsymbol{\nabla} u_\ell^k||\boldsymbol{\nabla} u| - \boldsymbol{\nabla} u_\ell^k\cdot\boldsymbol{\nabla} u)\right)\\
&\overset{(A.4a)}{=} \int_\Omega a_{\mathrm{c},\ell}^{u,k}|\boldsymbol{\nabla} u_\ell^k - \boldsymbol{\nabla} u|^2.
\end{aligned}
\tag{7.6}
$$

The left-hand side of (7.2a) is obtained with the same reasoning. Moreover, since ϕ' is nondecreasing and $|\boldsymbol{\nabla} u_\ell^k|-r$ is of the same sign as $|\boldsymbol{\nabla} u_\ell^k|-|\boldsymbol{\nabla} u|$ for all $r\in(|\boldsymbol{\nabla} u|,|\boldsymbol{\nabla} u_\ell^k|)$, using the mean value inequality, we have a.e. in Ω,

$$
\begin{aligned}
\int_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|} (|\boldsymbol{\nabla} u_\ell^k|-r)\phi''(\cdot,r)\,\mathrm{d}r &\geq \int_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|} \frac{\phi'(\cdot,|\boldsymbol{\nabla} u_\ell^k|)-\phi'(\cdot,r)}{\operatorname*{ess\,sup}_{s\in(r,|\boldsymbol{\nabla} u_\ell^k|)}\phi''(\cdot,s)}\phi''(\cdot,r)\,\mathrm{d}r\\
&\overset{(4.4b)}{\geq} \frac{1}{a_{\mathrm{c},\ell}^{\boldsymbol{\nabla} u,k}}\int_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|}(\phi'(\cdot,|\boldsymbol{\nabla} u_\ell^k|)-\phi'(\cdot,r))\phi''(\cdot,r)\,\mathrm{d}r\\
&= \frac{1}{a_{\mathrm{c},\ell}^{\boldsymbol{\nabla} u,k}}\left[\phi'(\cdot,|\boldsymbol{\nabla} u_\ell^k|)\phi'(\cdot,r)-\frac{1}{2}\phi'(\cdot,r)^2\right]_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|}\\
&\overset{(A.2)}{=} \frac{1}{a_{\mathrm{c},\ell}^{\boldsymbol{\nabla} u,k}}\left(\frac{1}{2}\left(|a_\ell^k\boldsymbol{\nabla} u_\ell^k|^2+|a_u\boldsymbol{\nabla} u|^2\right) - a_\ell^k a_u|\boldsymbol{\nabla} u_\ell^k||\boldsymbol{\nabla} u|\right).
\end{aligned}
\tag{7.7}
$$

Thus, observing that $|\boldsymbol{\nabla} u_\ell^k||\boldsymbol{\nabla} u| - \boldsymbol{\nabla} u_\ell^k\cdot\boldsymbol{\nabla} u \geq 0$ thanks to the Cauchy–Schwarz inequality, we obtain the left-hand side of (7.2c) by writing

$$
\begin{aligned}
(\mathcal{E}_{\mathrm{N},\ell}^k)^2 &\overset{(7.5),(7.7),(4.4b)}{\geq} \int_\Omega \frac{2}{a_{\mathrm{c},\ell}^{\boldsymbol{\nabla} u,k}}\left(\frac{1}{2}\left(|a_\ell^k\boldsymbol{\nabla} u_\ell^k|^2+|a_u\boldsymbol{\nabla} u|^2\right) - (a_\ell^k\boldsymbol{\nabla} u_\ell^k)\cdot(a_u\boldsymbol{\nabla} u)\right)\\
&= \int_\Omega \frac{1}{a_{\mathrm{c},\ell}^{\boldsymbol{\nabla} u,k}}|a_\ell^k\boldsymbol{\nabla} u_\ell^k - a_u\boldsymbol{\nabla} u|^2.
\end{aligned}
\tag{7.8}
$$

Finally, since $|\boldsymbol{\nabla} u_\ell^k|-r$ is of the same sign as $|\boldsymbol{\nabla} u_\ell^k|-|\boldsymbol{\nabla} u|$ for all $r\in(|\boldsymbol{\nabla} u|,|\boldsymbol{\nabla} u_\ell^k|)$, we have

$$
\begin{aligned}
\int_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|} (|\boldsymbol{\nabla} u_\ell^k|-r)\phi''(\cdot,r)\,\mathrm{d}r &\leq (|\boldsymbol{\nabla} u_\ell^k|-|\boldsymbol{\nabla} u|)\int_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|}\phi''(\cdot,r)\,\mathrm{d}r\\
&\overset{(4.4a)}{\leq} \frac{1}{a_{\mathrm{m},\ell}^{\boldsymbol{\nabla} u,k}}\left(\int_{|\boldsymbol{\nabla} u|}^{|\boldsymbol{\nabla} u_\ell^k|}\phi''(\cdot,r)\,\mathrm{d}r\right)^2 \overset{(A.2)}{=} \frac{1}{a_{\mathrm{m},\ell}^{\boldsymbol{\nabla} u,k}}\left(|a_\ell^k\boldsymbol{\nabla} u_\ell^k|-|a_u\boldsymbol{\nabla} u|\right)^2.
\end{aligned}
\tag{7.9}
$$

Hence, observing again that $|\boldsymbol{\nabla} u_\ell^k||\boldsymbol{\nabla} u| - \boldsymbol{\nabla} u_\ell^k\cdot\boldsymbol{\nabla} u \geq 0$ thanks to the Cauchy–Schwarz inequality, we obtain the right-hand side of (7.2c) by writing

$$
\begin{aligned}
(\mathcal{E}_{\mathrm{N},\ell}^k)^2 &\overset{(7.5),(7.9),(4.4a)}{\leq} \int_\Omega \frac{2}{a_{\mathrm{m},\ell}^{\boldsymbol{\nabla} u,k}}\Big(\left(|a_\ell^k\boldsymbol{\nabla} u_\ell^k|-|a_u\boldsymbol{\nabla} u|\right)^2 + 2(|a_\ell^k\boldsymbol{\nabla} u_\ell^k||a_u\boldsymbol{\nabla} u|\\
&\qquad - (a_\ell^k\boldsymbol{\nabla} u_\ell^k)\cdot(a_u\boldsymbol{\nabla} u))\Big) \overset{(A.4a)}{=} \int_\Omega \frac{2}{a_{\mathrm{m},\ell}^{\boldsymbol{\nabla} u,k}}|a_\ell^k\boldsymbol{\nabla} u_\ell^k - a_u\boldsymbol{\nabla} u|^2.
\end{aligned}
\tag{7.10}
$$

□

Proof of Theorem 4.1. Inequality (4.5a) is already contained in the proof of Theorem 3.4. It remains to prove (4.5b). For all vertices $\boldsymbol{a} \in \mathcal{V}_\ell$, using Lemma 6.1 (applied with $\boldsymbol{\tau}_\ell^{\boldsymbol{a}} = \psi_\ell^{\boldsymbol{a}} \boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}} (a_\ell^k \boldsymbol{\nabla} u_\ell^k) \in \boldsymbol{RTN}_p(\mathcal{T}_\ell^{\boldsymbol{a}})$ and $g_\ell^{\boldsymbol{a}} = \Pi_{\ell,p} \gamma_\ell^{\boldsymbol{a},k} \in \mathcal{P}_p(\mathcal{T}_\ell^{\boldsymbol{a}})$ with $(g_\ell^{\boldsymbol{a}}, 1)_{\omega_\ell^{\boldsymbol{a}}} = 0$ if $\boldsymbol{a} \in \mathcal{V}_\ell^{\mathrm{int}}$) together with the fact that $\psi_\ell^{\boldsymbol{a}} \varphi \in H_0^1(\omega_\ell^{\boldsymbol{a}})$ and $\boldsymbol{\nabla}(\psi_\ell^{\boldsymbol{a}} \varphi) = \psi_\ell^{\boldsymbol{a}} \boldsymbol{\nabla} \varphi + \varphi \boldsymbol{\nabla} \psi_\ell^{\boldsymbol{a}}$, we get

$$
\begin{aligned}
\|\psi_\ell^{\boldsymbol{a}} a_\ell^k \boldsymbol{\nabla} u_\ell^k + \boldsymbol{\sigma}_\ell^{\boldsymbol{a},k}\|_{\omega_\ell^{\boldsymbol{a}}} &\le \|\psi_\ell^{\boldsymbol{a}} \boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}} (a_\ell^k \boldsymbol{\nabla} u_\ell^k) + \boldsymbol{\sigma}_\ell^{\boldsymbol{a},k}\|_{\omega_\ell^{\boldsymbol{a}}} \\
&\quad + \|\psi_\ell^{\boldsymbol{a}} (a_\ell^k \boldsymbol{\nabla} u_\ell^k - \boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}} (a_\ell^k \boldsymbol{\nabla} u_\ell^k))\|_{\omega_\ell^{\boldsymbol{a}}} \\
&\overset{(6.2),(4.2b)}{\lesssim} \sup_{\substack{\varphi \in H_*^1(\omega_\ell^{\boldsymbol{a}}) \\ \|\boldsymbol{\nabla}\varphi\|_{\omega_\ell^{\boldsymbol{a}}} = 1}} \big[(f, \psi_\ell^{\boldsymbol{a}} \varphi)_{\omega_\ell^{\boldsymbol{a}}} - (a_\ell^k \boldsymbol{\nabla} u_\ell^k, \boldsymbol{\nabla}(\psi_\ell^{\boldsymbol{a}} \varphi))_{\omega_\ell^{\boldsymbol{a}}} \big] + \widetilde{\eta}_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k} \\
&\overset{(2.2a)}{=} \sup_{\substack{\varphi \in H_*^1(\omega_\ell^{\boldsymbol{a}}) \\ \|\boldsymbol{\nabla}\varphi\|_{\omega_\ell^{\boldsymbol{a}}} = 1}} (a_u \boldsymbol{\nabla} u - a_\ell^k \boldsymbol{\nabla} u_\ell^k, \boldsymbol{\nabla}(\psi_\ell^{\boldsymbol{a}} \varphi))_{\omega_\ell^{\boldsymbol{a}}} + \widetilde{\eta}_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k} \\
&\lesssim \|a_\ell^k \boldsymbol{\nabla} u_\ell^k - a_u \boldsymbol{\nabla} u\|_{\omega_\ell^{\boldsymbol{a}}} + \widetilde{\eta}_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k},
\end{aligned}
\tag{7.11}
$$

where we have used $\|\boldsymbol{\nabla}(\psi_\ell^{\boldsymbol{a}} \varphi)\|_{\omega_\ell^{\boldsymbol{a}}} \lesssim \|\boldsymbol{\nabla}\varphi\|_{\omega_\ell^{\boldsymbol{a}}}$ as in [8, 19]. In conclusion, we obtain

$$
\begin{aligned}
\eta_{\mathrm{N},\ell}^k &\overset{(7.2b),(6.4)}{\lesssim} \left(\sum_{\boldsymbol{a} \in \mathcal{V}_\ell} \frac{1}{\operatorname{ess\,inf}_{\omega_\ell^{\boldsymbol{a}}} a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k}} \|\psi_\ell^{\boldsymbol{a}} a_\ell^k \boldsymbol{\nabla} u_\ell^k + \boldsymbol{\sigma}_\ell^{\boldsymbol{a},k}\|_{\omega_\ell^{\boldsymbol{a}}}^2 \right)^{\frac{1}{2}} \\
&\overset{(7.11)}{\lesssim} \left(\sum_{\boldsymbol{a} \in \mathcal{V}_\ell} \frac{1}{\operatorname{ess\,inf}_{\omega_\ell^{\boldsymbol{a}}} a_{\mathrm{m},\ell}^{\boldsymbol{\sigma},k}} \left(\|a_\ell^k \boldsymbol{\nabla} u_\ell^k - a_u \boldsymbol{\nabla} u\|_{\omega_\ell^{\boldsymbol{a}}} + \widetilde{\eta}_{\mathrm{osc,q},\ell}^{\boldsymbol{a},k} \right)^2 \right)^{\frac{1}{2}} \\
&\overset{(4.6),(4.2a)}{\le} \widetilde{C}_\ell^k \left(\sum_{\boldsymbol{a} \in \mathcal{V}_\ell} \frac{1}{\operatorname{ess\,sup}_{\omega_\ell^{\boldsymbol{a}}} a_{\mathrm{c},\ell}^{\boldsymbol{\nabla} u,k}} \|a_\ell^k \boldsymbol{\nabla} u_\ell^k - a_u \boldsymbol{\nabla} u\|_{\omega_\ell^{\boldsymbol{a}}}^2 \right)^{\frac{1}{2}} + \widetilde{\eta}_{\mathrm{osc,q},\ell}^k \\
&\overset{(6.3),(7.2c)}{\lesssim} \widetilde{C}_\ell^k \mathcal{E}_{\mathrm{N},\ell}^k + \widetilde{\eta}_{\mathrm{osc,q},\ell}^k,
\end{aligned}
\tag{7.12}
$$

where we have used the triangle inequality on $\ell_2(\mathbb{R}^{|\mathcal{V}_\ell|})$. □

A Equivalent assumptions on the nonlinear function

In this section, we show some useful properties of the nonlinear function a. In particular, we show that inequalities (2.1) admit equivalent versions with the function ϕ defined in (2.4), preserving the same constants a_{c} and a_{m}.

Proposition A.1 (Equivalent assumption on ϕ')**.** *Inequalities (2.1) are equivalent to the following ones: a.e. in Ω, for all $r, s \in [0, \infty)$,*

$$
|\phi'(\cdot, r) - \phi'(\cdot, s)| \le a_{\mathrm{c}} |r - s|, \tag{A.1a}
$$

$$
(\phi'(\cdot, r) - \phi'(\cdot, s))(r - s) \ge a_{\mathrm{m}} (r - s)^2. \tag{A.1b}
$$

Proof. Differentiating (2.4) gives, a.e. in Ω and for all $r \in [0, \infty)$,

$$
\phi'(\cdot, r) = a(\cdot, r) r. \tag{A.2}
$$

Thus, assuming and using (2.1) (with $\boldsymbol{x} = (r, 0, \dots, 0)^{\mathrm{t}}$ and $\boldsymbol{y} = (s, 0, \dots, 0)^{\mathrm{t}}$) together with (A.2) yields (A.1).

Reciprocally, assuming (A.1), we have in particular (with $s = 0$) together with (A.2) that, a.e. in Ω and for all $r \in [0, \infty)$,

$$
a_{\mathrm{m}} \le a(\cdot, r) \le a_{\mathrm{c}}. \tag{A.3}
$$

We conclude by using the fact that for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$, and all $\alpha, \beta \in [0, \infty)$, we have

$$|\boldsymbol{x} - \boldsymbol{y}|^2 = (|\boldsymbol{x}| - |\boldsymbol{y}|)^2 + 2(|\boldsymbol{x}||\boldsymbol{y}| - \boldsymbol{x} \cdot \boldsymbol{y}), \tag{A.4a}$$

$$(\alpha\boldsymbol{x} - \beta\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}) = (\alpha|\boldsymbol{x}| - \beta|\boldsymbol{y}|)(|\boldsymbol{x}| - |\boldsymbol{y}|) + (\alpha + \beta)(|\boldsymbol{x}||\boldsymbol{y}| - \boldsymbol{x} \cdot \boldsymbol{y}), \tag{A.4b}$$

to obtain, a.e. in Ω and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$,

$$\begin{aligned}
&|a(\cdot, |\boldsymbol{x}|)\boldsymbol{x} - a(\cdot, |\boldsymbol{y}|)\boldsymbol{y}|^2 \\
&\quad \overset{(A.4a)}{=} (a(\cdot, |\boldsymbol{x}|)|\boldsymbol{x}| - a(\cdot, |\boldsymbol{y}|)|\boldsymbol{y}|)^2 + 2a(\cdot, |\boldsymbol{x}|)a(\cdot, |\boldsymbol{y}|)(|\boldsymbol{x}||\boldsymbol{y}| - \boldsymbol{x} \cdot \boldsymbol{y}) \\
&\quad \overset{(A.2)}{=} (\phi'(\cdot, |\boldsymbol{x}|) - \phi'(\cdot, |\boldsymbol{y}|))^2 + 2a(\cdot, |\boldsymbol{x}|)a(\cdot, |\boldsymbol{y}|)(|\boldsymbol{x}||\boldsymbol{y}| - \boldsymbol{x} \cdot \boldsymbol{y}) \\
&\quad \overset{(A.1),(A.3)}{\leq} a_{\mathrm{c}}^2[(|\boldsymbol{x}| - |\boldsymbol{y}|)^2 + 2(|\boldsymbol{x}||\boldsymbol{y}| - \boldsymbol{x} \cdot \boldsymbol{y})] = a_{\mathrm{c}}^2|\boldsymbol{x} - \boldsymbol{y}|^2
\end{aligned}$$

and

$$\begin{aligned}
&(a(\cdot,|\boldsymbol{x}|)\boldsymbol{x} - a(\cdot, |\boldsymbol{y}|)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}) \\
&\quad \overset{(A.4b)}{=} (a(\cdot, |\boldsymbol{x}|)|\boldsymbol{x}| - a(\cdot, |\boldsymbol{y}|)|\boldsymbol{y}|)(|\boldsymbol{x}| - |\boldsymbol{y}|) + (a(\cdot, |\boldsymbol{x}|) + a(\cdot, |\boldsymbol{y}|))(|\boldsymbol{x}||\boldsymbol{y}| - \boldsymbol{x} \cdot \boldsymbol{y}) \\
&\quad \overset{(A.2)}{=} (\phi'(\cdot, |\boldsymbol{x}|) - \phi'(\cdot, |\boldsymbol{y}|))(|\boldsymbol{x}| - |\boldsymbol{y}|) + (a(\cdot, |\boldsymbol{x}|) + a(\cdot, |\boldsymbol{y}|))(|\boldsymbol{x}||\boldsymbol{y}| - \boldsymbol{x} \cdot \boldsymbol{y}) \\
&\quad \overset{(A.1),(A.3)}{\geq} a_{\mathrm{m}}[(|\boldsymbol{x}| - |\boldsymbol{y}|)^2 + 2(|\boldsymbol{x}||\boldsymbol{y}| - \boldsymbol{x} \cdot \boldsymbol{y})] = a_{\mathrm{m}}|\boldsymbol{x} - \boldsymbol{y}|^2,
\end{aligned}$$

hence (2.1). □

Proposition A.2 (Equivalent assumption on ϕ'')**.** *Inequalities* (A.1) *are equivalent to the facts that* ϕ' *is weakly differentiable and, a.e. in* Ω *and for a.e.* $r \in [0, \infty)$*,*

$$a_{\mathrm{m}} \leq \phi''(\cdot, r) \leq a_{\mathrm{c}}. \tag{A.5}$$

Proof. Assuming (A.1), since ϕ' is Lipschitz continuous thanks to (A.1a), ϕ' is weakly differentiable. Furthermore, defining the difference quotient, a.e. in Ω and for all $r, s \in [0, \infty)$, $r \neq s$, by

$$\tau(\cdot, r, s) \coloneqq \frac{\phi'(\cdot, r) - \phi'(\cdot, s)}{r - s},$$

inequalities (A.1) are equivalent to the fact that, a.e. in Ω and for all $r, s \in [0, \infty)$, $r \neq s$,

$$a_{\mathrm{m}} \leq \tau(\cdot, r, s) \leq a_{\mathrm{c}}. \tag{A.6}$$

Hence, letting s tend to r in (A.6) gives (A.5).

On the other hand, assuming and integrating (A.5) gives, a.e. in Ω and for a.e. $r, s \in [0, \infty)$ with $r > s$,

$$a_{\mathrm{m}}(r - s) \leq \phi'(\cdot, r) - \phi'(\cdot, s) \leq a_{\mathrm{c}}(r - s), \tag{A.7}$$

and dividing (A.7) by $r - s$ gives (A.6), which becomes also true for $r < s$ by symmetry. □

Remark A.3 (Convexity)**.** *Under Assumption 2.1, inequality* (A.1b) *implies that* ϕ' *is nondecreasing, i.e.,* ϕ *is convex. Moreover, from* (A.2)*, we have, a.e. in* Ω *and for a.e.* $r \in [0, \infty)$*,*

$$\phi''(\cdot, r) = a(\cdot, r) + a'(\cdot, r)r. \tag{A.8}$$

B Spectral properties of the tensor product

Lemma B.1 (Spectral properties of the tensor product)**.** *The following holds:*

$$\mathrm{Spec}(\alpha\boldsymbol{I}_d + \beta\boldsymbol{A}) = \{\alpha\} + \beta\,\mathrm{Spec}(\boldsymbol{A}) \qquad \forall \alpha, \beta \in \mathbb{R},\ \forall \boldsymbol{A} \in \mathbb{R}^{d\times d}, \tag{B.1a}$$

$$\mathrm{Spec}(\boldsymbol{\xi} \otimes \boldsymbol{\xi}) = \{0, |\boldsymbol{\xi}|^2\} \qquad \forall \boldsymbol{\xi} \in \mathbb{R}^d,\ d > 1. \tag{B.1b}$$

Proof. We refer to [13] for details about the tools used in the following. Denoting $P_{\boldsymbol{A}}$ the characteristic polynomial of $\boldsymbol{A}$, we obtain (B.1a) by writing for all $\lambda \in \mathbb{R}$,

$$P_{\alpha\boldsymbol{I}_d+\beta\boldsymbol{A}}(\alpha + \beta\lambda) = \det((\alpha + \beta\lambda)\boldsymbol{I}_d - (\alpha\boldsymbol{I}_d + \beta\boldsymbol{A})) = \beta^d \det(\lambda\boldsymbol{I}_d - \boldsymbol{A}) = \beta^d P_{\boldsymbol{A}}(\lambda).$$

Moving to (B.1b), since $(\boldsymbol{\xi} \otimes \boldsymbol{\xi})\boldsymbol{\tau} = (\boldsymbol{\xi} \cdot \boldsymbol{\tau})\boldsymbol{\xi}$ for all $\boldsymbol{\tau} \in \mathbb{R}^d$, $\dim(\mathrm{Ker}(\boldsymbol{\xi} \otimes \boldsymbol{\xi})) = \dim(\mathrm{Ker}(\langle \boldsymbol{\xi}, \cdot \rangle)) \geq d - 1$, i.e. $0 \in \mathrm{Spec}(\boldsymbol{\xi} \otimes \boldsymbol{\xi})$ with a multiplicity of at least $d - 1$. Thus, the sum of the eigenvalues of $\boldsymbol{\xi} \otimes \boldsymbol{\xi}$, being $\mathrm{tr}(\boldsymbol{\xi} \otimes \boldsymbol{\xi}) = |\boldsymbol{\xi}|^2$, is in $\mathrm{Spec}(\boldsymbol{\xi} \otimes \boldsymbol{\xi})$. Hence, 0 and $|\boldsymbol{\xi}|^2$ are the only elements of $\mathrm{Spec}(\boldsymbol{\xi} \otimes \boldsymbol{\xi})$, and we infer (B.1b). □

References

[1] S. Badia and F. Verdugo, *Gridap: an extensible finite element toolbox in Julia*, Journal of Open Source Software, 5 (2020), p. 2520, https://doi.org/10.21105/joss.02520.

[2] S. Bartels, *Numerical methods for nonlinear partial differential equations*, vol. 47 of Springer Series in Computational Mathematics, Springer, Cham, 2015, https://doi.org/10.1007/978-3-319-13797-1.

[3] S. Bartels and A. Kaltenbach, *Explicit and efficient error estimation for convex minimization problems*, 2022, https://doi.org/10.48550/ARXIV.2204.10745. ArXiv preprint 2204.10745.

[4] S. Bartels and M. Milicevic, *Primal-dual gap estimators for a posteriori error analysis of nonsmooth minimization problems*, ESAIM Math. Model. Numer. Anal., 54 (2020), pp. 1635–1660, https://doi.org/10.1051/m2an/2019074.

[5] 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, https://doi.org/10.1093/imanum/drr016.

[6] J. Blechta, J. Málek, and M. Vohralík, *Localization of the $W^{-1,q}$ norm for local a posteriori efficiency*, IMA J. Numer. Anal., 40 (2020), pp. 914–950, https://doi.org/10.1093/imanum/drz002.

[7] D. Boffi, F. Brezzi, and M. Fortin, *Mixed finite element methods and applications*, vol. 44 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2013, https://doi.org/10.1007/978-3-642-36519-5.

[8] D. Braess and J. Schöberl, *Equilibrated residual error estimator for edge elements*, Math. Comp., 77 (2008), pp. 651–672, https://doi.org/10.1090/S0025-5718-07-02080-7.

[9] C. Carstensen, M. Feischl, M. Page, and D. Praetorius, *Axioms of adaptivity*, Comput. Math. Appl., 67 (2014), pp. 1195–1253, https://doi.org/10.1016/j.camwa.2013.12.003.

[10] A. L. Chaillou and M. Suri, *Computable error estimators for the approximation of nonlinear problems by linearized models*, Comput. Methods Appl. Mech. Engrg., 196 (2006), pp. 210–224, https://doi.org/10.1016/j.cma.2006.03.008.

[11] R. Cottereau, L. Chamoin, and P. Díez, *Strict error bounds for linear and nonlinear solid mechanics problems using a patch-based flux-free method*, Mechanics & Industry, 11 (2010), pp. 249–254, https://doi.org/10.1051/meca/2010049.

[12] R. Cottereau, P. Díez, and A. Huerta, *Strict error bounds for linear solid mechanics problems using a subdomain-based flux-free method*, Comput. Mech., 44 (2009), pp. 533–547, https://doi.org/10.1007/s00466-009-0388-1.

[13] C. W. Curtis, *Linear algebra*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, fourth ed., 1993. An introductory approach.

[14] K. Deckelnick, G. Dziuk, and C. M. Elliott, *Computation of geometric partial differential equations and mean curvature flow*, Acta Numer., 14 (2005), pp. 139–232, https://doi.org/10.1017/S0962492904000224.

[15] P. Destuynder and B. Métivet, *Explicit error bounds in a conforming finite element method*, Math. Comp., 68 (1999), pp. 1379–1396, https://doi.org/10.1090/S0025-5718-99-01093-5.

[16] P. Deuflhard, *Newton methods for nonlinear problems*, vol. 35 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2011, https://doi.org/10.1007/978-3-642-23899-4. Affine invariance and adaptive algorithms, First softcover printing of the 2006 corrected printing.

[17] L. El Alaoui, A. Ern, and M. Vohralík, *Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 2782–2795, https://doi.org/10.1016/j.cma.2010.03.024.

[18] A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004, https://doi.org/10.1007/978-1-4757-4355-5.

[19] A. Ern, I. Smears, and M. Vohralík, *Discrete p-robust $\boldsymbol{H}$(div)-liftings and a posteriori estimates for elliptic problems with H^{-1} source terms*, Calcolo, 54 (2017), pp. 1009–1025, https://doi.org/10.1007/s10092-017-0217-4.

[20] A. Ern and M. Vohralík, *Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs*, SIAM J. Sci. Comput., 35 (2013), pp. A1761–A1791, https://doi.org/10.1137/120896918.

[21] A. Ern and M. Vohralík, *Stable broken H^1 and H(div) polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions*, Math. Comp., 89 (2020), pp. 551–594, https://doi.org/10.1090/mcom/3482.

[22] R. E. Ewing, *Finite element methods for nonlinear flows in porous media*, vol. 51, 1985, pp. 421–439, https://doi.org/10.1016/0045-7825(85)90041-6. FENOMECH '84, Part I, II (Stuttgart, 1984).

[23] F. Févotte, A. Rappaport, and M. Vohralík, *Adaptive regularization, discretization, and linearization for nonsmooth problems based on primal-dual gap estimators*. Preprint, May 2023, https://hal.science/hal-04105560.

[24] G. Gantner, A. Haberl, D. Praetorius, and B. Stiftner, *Rate optimal adaptive FEM with inexact solver for nonlinear operators*, IMA J. Numer. Anal., 38 (2018), pp. 1797–1831, https://doi.org/10.1093/imanum/drx050.

[25] E. M. Garau, P. Morin, and C. Zuppa, *Convergence of an adaptive Kačanov FEM for quasi-linear problems*, Appl. Numer. Math., 61 (2011), pp. 512–529, https://doi.org/10.1016/j.apnum.2010.12.001.

[26] M. Guo, W. Han, and H. Zhong, *Legendre-Fenchel duality and a generalized constitutive relation error*. ArXiv preprint 1611.05589, 2016, https://arxiv.org/abs/1611.05589.

[27] A. Haberl, D. Praetorius, S. Schimanko, and M. Vohralík, *Convergence and quasi-optimal cost of adaptive algorithms for nonlinear operators including iterative linearization and algebraic solver*, Numer. Math., 147 (2021), pp. 679–725, https://doi.org/10.1007/s00211-021-01176-w.

[28] W. Han, *A posteriori error analysis for linearization of nonlinear elliptic problems and their discretizations*, Math. Methods Appl. Sci., 17 (1994), pp. 487–508, https://doi.org/10.1002/mma.1670170702.

[29] W. Han, *A posteriori error analysis via duality theory*, vol. 8 of Advances in Mechanics and Mathematics, Springer-Verlag, New York, 2005. With applications in modeling and numerical approximations.

[30] P. Heid, D. Praetorius, and T. P. Wihler, *Energy contraction and optimal convergence of adaptive iterative linearized finite element methods*, Comput. Methods Appl. Math., 21 (2021), pp. 407–422, https://doi.org/10.1515/cmam-2021-0025.

[31] P. Heid and T. P. Wihler, *On the convergence of adaptive iterative linearized Galerkin methods*, Calcolo, 57 (2020), pp. Paper No. 24, 23, https://doi.org/10.1007/s10092-020-00368-4.

[32] J. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms I: Fundamentals*, Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg, 1996.

[33] P. Houston, E. Süli, and T. P. Wihler, *A posteriori error analysis of hp-version discontinuous Galerkin finite-element methods for second-order quasi-linear elliptic PDEs*, IMA J. Numer. Anal., 28 (2008), pp. 245–273.

[34] C. T. Kelley, *Iterative methods for linear and nonlinear equations*, vol. 16 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995, https://doi.org/10.1137/1.9781611970944. With separately available software.

[35] K.-Y. Kim, *A posteriori error estimators for locally conservative methods of nonlinear elliptic problems*, Appl. Numer. Math., 57 (2007), pp. 1065–1080, https://doi.org/10.1016/j.apnum.2006.09.010.

[36] P. Ladevèze and N. Moës, *A posteriori constitutive relation error estimators for nonlinear finite element analysis and adaptive control*, in Advances in adaptive computational methods in mechanics (Cachan, 1997), vol. 47 of Stud. Appl. Mech., Elsevier Sci. B. V., Amsterdam, 1998, pp. 231–256, https://doi.org/10.1016/S0922-5382(98)80013-5.

[37] K. Mitra and M. Vohralík, *Guaranteed, locally efficient, and robust a posteriori estimates for nonlinear elliptic problems in iteration-dependent norms. An orthogonal decomposition result based on iterative linearization.* To be submitted, 2023.

[38] P. Neittaanmaki and S. Repin, *A posteriori error identities for nonlinear variational problems*, Ann. Acad. Rom. Sci. Ser. Math. Appl., 7 (2015), pp. 157–172.

[39] S. Repin, *A posteriori estimates for partial differential equations*, vol. 4 of Radon Series on Computational and Applied Mathematics, Walter de Gruyter GmbH & Co. KG, Berlin, 2008, https://doi.org/10.1515/9783110203042.

[40] J. L. Vázquez, *An introduction to the mathematical theory of the porous medium equation*, in Shape optimization and free boundaries (Montreal, PQ, 1990), vol. 380 of NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci., Kluwer Acad. Publ., Dordrecht, 1992, pp. 347–389.

[41] F. Verdugo and S. Badia, *The software design of Gridap: a finite element package based on the Julia JIT compiler*, Computer Physics Communications, 276 (2022), p. 108341, https://doi.org/10.1016/j.cpc.2022.108341.

[42] E. H. Zarantonello, *Solving Functional Equations by Contractive Averaging*, Tech. Report 160, Mathematics Research Center, United States Army, University of Wisconsin, 1960.

[43] E. Zeidler, *Nonlinear Functional Analysis and Its Applications. II/B: Nonlinear Monotone Operators*, no. XV, 1992, https://doi.org/10.1007/978-1-4612-0981-2.