



# Convergence and quasi-optimal cost of adaptive algorithms for nonlinear operators including iterative linearization and algebraic solver

Alexander Haberl<sup>1</sup> · Dirk Praetorius<sup>1</sup> · Stefan Schimanko<sup>1</sup> · Martin Vohralík<sup>2,3</sup>

Received: 17 February 2020 / Revised: 18 October 2020 / Accepted: 24 December 2020 /  
Published online: 5 February 2021  
© The Author(s) 2021

## Abstract

We consider a second-order elliptic boundary value problem with strongly monotone and Lipschitz-continuous nonlinearity. We design and study its adaptive numerical approximation interconnecting a finite element discretization, the Banach–Picard linearization, and a contractive linear algebraic solver. In particular, we identify stopping criteria for the algebraic solver that on the one hand do not request an overly tight tolerance but on the other hand are sufficient for the inexact (perturbed) Banach–Picard linearization to remain contractive. Similarly, we identify suitable stopping criteria for the Banach–Picard iteration that leave an amount of linearization error that is not harmful for the residual *a posteriori* error estimate to steer reliably the adaptive mesh-refinement. For the resulting algorithm, we prove a contraction of the (doubly) inexact iterates after some amount of steps of mesh-refinement/linearization/algebraic solver, leading to its linear convergence. Moreover, for usual mesh-refinement rules, we also prove that the overall error decays at the optimal rate with respect to the number of elements (degrees of freedom) added with respect to the initial mesh. Finally, we prove that our fully adaptive algorithm drives the overall error down with the same optimal rate also with respect to the overall algorithmic cost expressed as the cumulated sum of the number of mesh elements over all mesh-refinement, linearization, and algebraic solver steps. Numerical experiments support these theoretical findings and illustrate the optimal overall algorithmic cost of the fully adaptive algorithm on several test cases.

**Mathematics Subject Classification** 65N12 · 65N15 · 65N30 · 65N50 · 68Q25

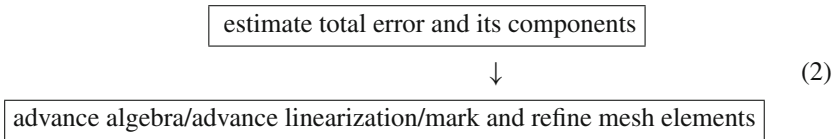
## 1 Introduction

Let  $\Omega \subset \mathbb{R}^d$  with  $d \geq 1$  be a bounded Lipschitz domain with polytopal boundary. Given  $f \in L^2(\Omega)$  and a nonlinear operator  $A: \mathbb{R}^d \rightarrow \mathbb{R}^d$ , we aim to numerically

approximate the weak solution  $u^* \in H_0^1(\Omega)$  of the nonlinear boundary value problem

$$\begin{aligned} -\operatorname{div} A(\nabla u^*) &= f \quad \text{in } \Omega, \\ u^* &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{1}$$

To this end, we propose an adaptive algorithm of the type



which monitors and adequately stops the iterative linearization and the linear algebraic solver as well as steers the local mesh-refinement. The goal of this contribution is to perform a first rigorous mathematical analysis of this algorithm in terms of convergence and quasi-optimal computational cost.

### 1.1 Finite element approximation and Banach–Picard iteration

Suppose that the nonlinearity  $A$  in (1) is Lipschitz-continuous (with constant  $L > 0$ ) and strongly monotone (with constant  $\alpha > 0$ ); see Sect. 2 for details. Then, the main theorem on monotone operators yields the existence and uniqueness of the weak solution  $u^* \in H_0^1(\Omega)$ ; see, e.g., [40, Theorem 25.B]. Given a triangulation  $\mathcal{T}_H$  of  $\Omega$ , the lowest-order finite element method (FEM) for problem (1) reads as follows: Find  $u_H^* \in \mathcal{X}_H := \{v_H \in C(\overline{\Omega}) : v_H|_T \text{ is affine for all } T \in \mathcal{T}_H \text{ and } v_H|_{\partial\Omega} = 0\} \subset H_0^1(\Omega)$  such that

$$(A(\nabla u_H^*), \nabla v_H)_\Omega = (f, v_H)_\Omega \quad \text{for all } v_H \in \mathcal{X}_H, \tag{3}$$

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

The discrete solution  $u_H^* \in \mathcal{X}_H$  again exists and is unique, but (3) corresponds to a *nonlinear discrete system* which can typically only be solved *inexactly*.

The most straightforward algorithm for *iterative linearization* of (3) stems from the proof of the main theorem on monotone operators which is constructive and relies on the Banach fixed point theorem: Define the (nonlinear) operator  $\Phi_H : \mathcal{X}_H \rightarrow \mathcal{X}_H$  by

$$(\nabla \Phi_H(w_H), \nabla v_H)_\Omega = (\nabla w_H, \nabla v_H)_\Omega - \frac{\alpha}{L^2} [(A(\nabla w_H), \nabla v_H)_\Omega - (f, v_H)_\Omega] \tag{4}$$

for all  $w_H, v_H \in \mathcal{X}_H$ . Note that (4) corresponds to a discrete Poisson problem and hence  $\Phi_H(w_H) \in \mathcal{X}_H$  is well-defined. Then, it holds that

$$\begin{aligned} \|\nabla(u_H^* - \Phi_H(w_H))\|_{L^2(\Omega)} &\leq q_{\text{Pic}} \|\nabla(u_H^* - w_H)\|_{L^2(\Omega)} \\ \text{with } q_{\text{Pic}} &:= (1 - \alpha^2/L^2)^{1/2} < 1; \end{aligned} \tag{5}$$

see, e.g., [40, Sect. 25.4]. Based on the contraction  $\Phi_H$ , the Banach–Picard iteration starts from an arbitrary discrete initial guess and applies  $\Phi_H$  inductively to generate a sequence of discrete functions which hence converge towards  $u_H^*$ . Note that the computation of  $\Phi_H(w_H)$  by means of the discrete variational formulation (4) corresponds to the solution of a (generically large) *linear discrete system* with symmetric and positive definite matrix that does not change during the iterations. In this work, we suppose that also (4) is solved *inexactly* by means of a *contractive iterative algebraic solver* (with contraction factor  $q_{\text{alg}} < 1$ ), e.g., PCG with optimal preconditioner; see, e.g., [34].

### 1.2 Fully adaptive algorithm

In our approach, we compute a *sequence* of discrete approximations  $u_\ell^{k,j}$  of  $u^*$  that have an index  $\ell$  for the *mesh-refinement*, an index  $k$  for the Banach–Picard *linearization* iteration, and an index  $j$  for the *algebraic solver* iteration.

First, we design a stopping criterion for the algebraic solver such that, at linearization step  $k - 1 \in \mathbb{N}_0$  on the mesh  $\mathcal{T}_\ell$ , we stop for some index  $\underline{j} \in \mathbb{N}$ . At the next linearization step  $k \in \mathbb{N}$ , the arising linear system reads as follows:

$$\begin{aligned} &\text{Find } u_\ell^{k,\star} \in \mathcal{X}_\ell \text{ such that, for all } v_\ell \in \mathcal{X}_\ell, \\ &(\nabla u_\ell^{k,\star}, \nabla v_\ell)_\Omega = (\nabla u_\ell^{k-1,\underline{j}}, \nabla v_\ell)_\Omega - \frac{\alpha}{L^2} [(A(\nabla u_\ell^{k-1,\underline{j}}), \nabla v_\ell)_\Omega - (f, v_\ell)_\Omega], \end{aligned} \tag{6}$$

with uniquely defined but not computed exact solution  $u_\ell^{k,\star} = \Phi_\ell(u_\ell^{k-1,\underline{j}})$  and computed iterates  $u_\ell^{k,j}$  that approximate  $u_\ell^{k,\star}$ . Note that (6) is a *perturbed* Banach–Picard iteration since it starts from the available  $u_\ell^{k-1,\underline{j}}$ , typically not equal to the unavailable  $u_\ell^{k-1,\star}$ .

Second, we design a stopping criterion for the perturbed Banach–Picard iteration at some index  $\underline{k}$ , producing a discrete approximation  $u_\ell^{\underline{k},\underline{j}}$ .

Finally, we locally refine the triangulation  $\mathcal{T}_\ell$  on the basis of the Dörfler marking criterion for the local contributions of the residual error estimator  $\eta_\ell(u_\ell^{\underline{k},\underline{j}})$ , and, to lower the computational effort, employ nested iteration in that the continuation on the new triangulation  $\mathcal{T}_{\ell+1}$  is started with the initial guess  $u_{\ell+1}^{0,0} := u_\ell^{\underline{k},\underline{j}}$ .

### 1.3 Previous contributions

#### 1.3.1 Inexact linearization

Performing an inexact solve of the linear system of form (6) gives rise to the “inexact Newton method”; see, e.g., [14,21] and the references therein. Under appropriate conditions, these can asymptotically preserve the convergence speed of the “exact” Newton method. Note, however, that these approaches only focus on the finite-

dimensional system of nonlinear algebraic equations of the form (3) but do not see/take into account the continuous problem (1), which is our central issue here.

### 1.3.2 Taking into account the discretization error

Solving the nonlinear algebraic systems (3) “exactly” (up to machine precision), only the *discretization* error is left. Then, convergence and optimal decay rates of the error  $\|\nabla(u^* - u_H^*)\|_{L^2(\Omega)}$  with respect to the degrees of freedom of FEM adapting the approximation space (mesh) were obtained in [3,16,26,39], following the seminal contributions [2,10,17,33,36] for linear problems. We also refer to [8] for a general framework of convergence of adaptive FEM with optimal convergence rates and an overview of the state of the art.

### 1.3.3 Taking into account the discretization and linearization errors

Solving only the linear algebraic systems (6) “exactly” but (3) inexactly leaves the *discretization* and *linearization* errors. Such a setting has been considered in, e.g., [12,19], where reliable (guaranteed) and efficient *a posteriori* error estimates were derived. Adaptive algorithms aiming at a balance of the linearization and discretization errors were proposed and their optimal performance was observed numerically; see, e.g., [1,4,13,28]. Later, theoretical proofs of *plain convergence* (without rates) were given in [25,30], where [30] builds on the unified framework of [29] encompassing also the Kačanov and (damped) Newton linearizations in addition to the Banach–Picard linearization (6).

The own works [23,24] considered that the linear systems (6) are solved exactly at *linear cost* (so that  $u_\ell^{k,j} = u_\ell^{k,*}$  with  $j(\ell, k) = \mathcal{O}(1)$  in the present notation), as in the seminal work [36] for the Poisson model problem and in [9] for an adaptive Laplace eigenvalue computation. Under this so-called realistic assumption on the algebraic solver, we have proved in [23] that the overall strategy leads to optimal convergence rates with respect to the *number of degrees of freedom* as well as to *almost optimal* convergence rates with respect to the *overall computational cost*. The latter means that, if the total error converges with rate  $s > 0$  with respect to the degrees of freedom, then, for all  $\varepsilon > 0$ , it also converges with rate  $s - \varepsilon > 0$  with respect to the overall computational cost. The proof of [23] was based on proving first that the estimator  $\eta_\ell(u_\ell^{k,*})$  for the final Picard iterates decays with optimal rate  $s$  and second that the number of Picard iterates satisfies  $\underline{k}(\ell) \lesssim 1 + \log[1 + \eta_\ell(u_{\ell+1}^{k,*})/\eta_\ell(u_\ell^{k,*})]$ . This logarithmic bound then led to the bound  $s - \varepsilon$  for the convergence rate with respect to the overall computational cost.

Recently in [24], we have improved the latter result and proved optimal computational cost (i.e.,  $\varepsilon = 0$ ), still relying on the assumption that the discrete Poisson problem (6) is solved exactly at linear cost. The core idea of the new proof follows ideas from adaptive Uzawa FEM for the Stokes model problem [15,32]. However, besides the nonlinearity, the structural difference is that the adaptive Uzawa FEM employs an outer iteration on the continuous level (i.e., we first linearize and then

discretize), while the approach of [13,23,24,29,30] is first to discretize and then to linearize.

### 1.3.4 Taking into account the discretization, linearization, and algebraic errors

As in the present setting, the “adaptive inexact Newton method” in [20] takes into account all *discretization*, *linearization*, and *algebraic* error components; see also [11, 18,35] for regularizations on coarse meshes ensuring well-posedness of the discrete systems in Newton-like linearizations. The goal of the present work is to perform a first rigorous mathematical analysis of such algorithms in terms of convergence and optimal decay rate of the error with respect to the computational cost.

We stress that such results have already been derived for adaptive wavelet discretizations [7,38] which provide inherent control of the residual error in terms of the wavelet coefficients, while the present analysis for standard finite element discretizations has to rely on the local information of appropriate *a posteriori* error estimators. Also, while the present analysis is closely related to that of [24], we stress that both works [23,24] focused only on linearization and discretization, while here, we also include the innermost algebraic loop into the adaptive algorithm. In particular, the technical challenges in the present analysis are much more involved than in the preceding work [24] due to the coupling of the two nested inexact solvers.

### 1.4 Main results: linear convergence, optimal decay rate, and quasi-optimal cost

The present contribution appears to be the first work that provides a thorough convergence analysis of fully adaptive strategies for nonlinear equations. To describe more precisely our results, we first note that the sequential nature of the fully adaptive algorithm of Sect. 1.2 gives rise to an index set

$$\mathcal{Q} := \{(\ell, k, j) \in \mathbb{N}_0^3 : \text{discrete approximation } u_\ell^{k,j} \text{ is computed by the algorithm}\}$$

together with an ordering

$$(\ell, k, j) < (\ell', k', j') \stackrel{\text{def}}{\iff} u_\ell^{k,j} \text{ is computed earlier than } u_{\ell'}^{k',j'}.$$

Our first main result, formulated in Theorem 3 below, proves that the proposed adaptive strategy is *contractive* after some amount of steps and *linearly convergent* in the sense of

$$\Delta_{\ell'}^{k',j'} \leq C_{\text{lin}} q_{\text{lin}}^{|\ell',k',j'| - |\ell,k,j|} \Delta_\ell^{k,j} \quad \text{for all } (\ell, k, j) \leq (\ell', k', j'), \tag{7}$$

where  $C_{\text{lin}} \geq 1$  and  $0 < q_{\text{lin}} < 1$  are generic constants and  $\Delta_\ell^{k,j}$  is an appropriate quasi-error quantity involving the error  $\|\nabla(u^* - u_\ell^{k,j})\|_{L^2(\Omega)}$  as well as the error estimator  $\eta_\ell(u_\ell^{k,j})$ . Second, we prove the *optimal error decay rate* with respect to the

number of *degrees of freedom* added with respect to the initial mesh in the sense that

$$\sup_{(\ell,k,j) \in \mathcal{Q}} (\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^s \Delta_\ell^{k,j} < \infty \quad (8)$$

whenever  $u^*$  is approximable at algebraic rate  $s > 0$ ; see Theorem 4 below for the details. Finally, estimate (7) appears to be also the key argument to prove our most eminent result, namely the *optimal error decay rate* with respect to the *overall computational cost* of the fully adaptive algorithm which steers the mesh-refinement, the perturbed Banach–Picard linearization, and the algebraic solver. In short, this reads

$$\sup_{(\ell',k',j') \in \mathcal{Q}} \left( \sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) \leq (\ell',k',j')}} \#\mathcal{T}_\ell \right)^s \Delta_{\ell'}^{k',j'} < \infty \quad (9)$$

whenever  $u^*$  is approximable at algebraic rate  $s > 0$ ; see Theorem 5 below for the details. We stress that under realistic assumptions the sum in (9) is indeed proportional to the overall computational cost invested into the fully adaptive numerical approximation of (1), if the cost of all procedures like matrix and right-hand-side assembly, one algebraic solver step, evaluation of the involved a posteriori error estimates, marking, and local adaptive mesh refinement is proportional to the number of mesh elements in  $\mathcal{T}_\ell$  (i.e., the number of degrees of freedom).

## 1.5 Outline

The remainder of the paper is organised as follows. In Sect. 2, we introduce an abstract setting in which all our results will be formulated, define the exact weak and finite elements solutions (none of which is available in our setting), and introduce our requirements on mesh-refinement and error estimator. We also give here precise requirements on the algebraic solver, state our adaptive algorithm and stopping criteria in all details, and present our main results, including some discussions. The proofs of some auxiliary results and of Proposition 2 (reliability in Algorithm 1), Theorem 3 (linear convergence), Theorem 4 (optimal decay rate with respect to the degrees of freedom), and Theorem 5 (optimal decay rate with respect to the overall computational cost) are respectively given in Sects. 3, 4, 5, and 6. Finally, numerical experiments in Sect. 7 underline the theoretical findings.

Throughout our work, we apply the following convention: In statements of theorems, lemmas, etc., we explicitly state all constants together with their dependencies. In proofs, however, we abbreviate  $A \leq cB$  with a generic constant  $c > 0$  by writing  $A \lesssim B$ . Moreover,  $A \simeq B$  abbreviates  $A \lesssim B \lesssim A$ .

## 2 Adaptive algorithm and main results

In this section, we introduce an abstract setting, in which all our results will be formulated, define the exact weak and finite elements solutions, introduce our requirements

on mesh-refinement, error estimator, and algebraic solver, state our adaptive algorithm, and present our main results, including some discussions.

### 2.1 Abstract setting

Let  $\mathcal{X}$  be a Hilbert space over  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$  with scalar product  $(\cdot, \cdot)$ , corresponding norm  $\|\cdot\|$ , and dual space  $\mathcal{X}'$  (with canonical operator norm  $\|\cdot\|'$ ). Let  $P : \mathcal{X} \rightarrow \mathbb{K}$  be Gâteaux-differentiable with derivative  $\mathcal{A} := dP : \mathcal{X} \rightarrow \mathcal{X}'$ , i.e.,

$$\langle \mathcal{A}w, v \rangle_{\mathcal{X}' \times \mathcal{X}} = \lim_{\substack{t \rightarrow 0 \\ t \in \mathbb{R}}} \frac{P(w + tv) - P(w)}{t} \quad \text{for all } v, w \in \mathcal{X}.$$

We suppose that the operator  $\mathcal{A}$  is *strongly monotone* and *Lipschitz-continuous*, i.e.,

$$\alpha \|w - v\|^2 \leq \operatorname{Re} \langle \mathcal{A}w - \mathcal{A}v, w - v \rangle_{\mathcal{X}' \times \mathcal{X}} \quad \text{and} \quad \|\mathcal{A}w - \mathcal{A}v\|' \leq L \|w - v\| \tag{10}$$

for all  $v, w \in \mathcal{X}$ , where  $0 < \alpha \leq L$  are generic real constants.

Given a linear and continuous functional  $F \in \mathcal{X}'$ , the main theorem on monotone operators [40, Sect. 25.4] yields existence and uniqueness of the solution  $u^* \in \mathcal{X}$  of

$$\langle \mathcal{A}u^*, v \rangle_{\mathcal{X}' \times \mathcal{X}} = \langle F, v \rangle_{\mathcal{X}' \times \mathcal{X}} \quad \text{for all } v \in \mathcal{X}. \tag{11}$$

The result actually holds true for any closed subspace  $\mathcal{X}_H \subseteq \mathcal{X}$ , which also gives rise to a unique  $u_H^* \in \mathcal{X}_H$  such that

$$\langle \mathcal{A}u_H^*, v_H \rangle_{\mathcal{X}' \times \mathcal{X}} = \langle F, v_H \rangle_{\mathcal{X}' \times \mathcal{X}} \quad \text{for all } v_H \in \mathcal{X}_H. \tag{12}$$

Finally, with the *energy functional*  $\mathcal{E} := \operatorname{Re}(P - F)$ , it holds that

$$\frac{\alpha}{2} \|v_H - u_H^*\|^2 \leq \mathcal{E}(v_H) - \mathcal{E}(u_H^*) \leq \frac{L}{2} \|v_H - u_H^*\|^2 \quad \text{for all } v_H \in \mathcal{X}_H; \tag{13}$$

see, e.g., [23, Lemma 5.1]. In particular,  $u^* \in \mathcal{X}$  (resp.  $u_H^* \in \mathcal{X}_H^*$ ) is the unique minimizer of the minimization problem

$$\mathcal{E}(u^*) = \min_{v \in \mathcal{X}} \mathcal{E}(v) \quad (\text{resp. } \mathcal{E}(u_H^*) = \min_{v_H \in \mathcal{X}_H} \mathcal{E}(v_H)). \tag{14}$$

As for linear elliptic problems, it follows from (10)–(12) that the present setting guarantees the C ea lemma (see, e.g., [40, Sect. 25.4])

$$\|u^* - u_H^*\| \leq C_{\text{C ea}} \|u^* - v_H\| \quad \text{for all } v_H \in \mathcal{X}_H \quad \text{with} \quad C_{\text{C ea}} := L/\alpha. \tag{15}$$

## 2.2 Mesh-refinement

Let  $\mathcal{T}_H$  be a conforming simplicial mesh of  $\Omega$ , i.e., a partition of  $\overline{\Omega}$  into compact simplices  $T$  such that  $\bigcup_{T \in \mathcal{T}_H} T = \overline{\Omega}$  and such that the intersection of two different simplices is either empty or their common vertex or their common  $d'$ -dimensional face (for some  $1 \leq d' \leq d - 1$ ). We assume that  $\text{refine}(\cdot)$  is a fixed mesh-refinement strategy, e.g., newest vertex bisection [37]. We write  $\mathcal{T}_h = \text{refine}(\mathcal{T}_H, \mathcal{M}_H)$  for the coarsest one-level refinement of  $\mathcal{T}_H$ , where all marked elements  $\mathcal{M}_H \subseteq \mathcal{T}_H$  have been refined, i.e.,  $\mathcal{M}_H \subseteq \mathcal{T}_H \setminus \mathcal{T}_h$ . We write  $\mathcal{T}_h \in \text{refine}(\mathcal{T}_H)$ , if  $\mathcal{T}_h$  can be obtained by finitely many steps of one-level refinement (with appropriate, yet arbitrary marked elements in each step). We define  $\mathbb{T} := \text{refine}(\mathcal{T}_0)$  as the set of all meshes which can be generated from the initial simplicial mesh  $\mathcal{T}_0$  of  $\Omega$  by use of  $\text{refine}(\cdot)$ . Finally, we associate to each  $\mathcal{T}_H \in \mathbb{T}$  a corresponding finite-dimensional subspace  $\mathcal{X}_H \subsetneq \mathcal{X}$ , where we suppose that  $\mathcal{X}_H \subseteq \mathcal{X}_h$  whenever  $\mathcal{T}_H, \mathcal{T}_h \in \mathbb{T}$  with  $\mathcal{T}_h \in \text{refine}(\mathcal{T}_H)$ .

For our analysis, we only employ that the shape-regularity of all meshes  $\mathcal{T}_H \in \mathbb{T}$  is uniformly bounded by that of  $\mathcal{T}_0$  together with the following structural properties (R1)–(R3), where  $C_{\text{son}} \geq 2$  and  $C_{\text{mesh}} > 0$  are generic constants:

- (R1) *splitting property*: Each refined element is split into finitely many sons, i.e., for all  $\mathcal{T}_H \in \mathbb{T}$  and all  $\mathcal{M}_H \subseteq \mathcal{T}_H$ , the mesh  $\mathcal{T}_h = \text{refine}(\mathcal{T}_H, \mathcal{M}_H)$  satisfies that

$$\#(\mathcal{T}_H \setminus \mathcal{T}_h) + \#\mathcal{T}_H \leq \#\mathcal{T}_h \leq C_{\text{son}} \#(\mathcal{T}_H \setminus \mathcal{T}_h) + \#(\mathcal{T}_H \cap \mathcal{T}_h);$$

- (R2) *overlay estimate*: For all meshes  $\mathcal{T} \in \mathbb{T}$  and  $\mathcal{T}_h, \mathcal{T}_{h'} \in \text{refine}(\mathcal{T})$ , there exists a common refinement  $\mathcal{T}_h \oplus \mathcal{T}_{h'} \in \text{refine}(\mathcal{T}_h) \cap \text{refine}(\mathcal{T}_{h'}) \subseteq \text{refine}(\mathcal{T})$  such that

$$\#(\mathcal{T}_h \oplus \mathcal{T}_{h'}) \leq \#\mathcal{T}_h + \#\mathcal{T}_{h'} - \#\mathcal{T};$$

- (R3) *mesh-closure estimate*: For each sequence  $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$  of successively refined meshes, i.e.,  $\mathcal{T}_{\ell+1} := \text{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$  with  $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$  for all  $\ell \in \mathbb{N}_0$ , it holds that

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

For newest vertex bisection, we refer to [2,10,27,31,36,37] for the validity of (R1)–(R3). For red-refinement with first-order hanging nodes, details are found in [6].

## 2.3 Error estimator

For each mesh  $\mathcal{T}_H \in \mathbb{T}$ , suppose that we can compute refinement indicators

$$\eta_H(T, v_H) \geq 0 \quad \text{for all } T \in \mathcal{T}_H \text{ and all } v_H \in \mathcal{X}_H. \quad (16)$$



We denote

$$\eta_H(\mathcal{V}_H, v_H) := \left( \sum_{T \in \mathcal{V}_H} \eta_H(T, v_H)^2 \right)^{1/2} \quad \text{for all } \mathcal{V}_H \subseteq \mathcal{T}_H \tag{17}$$

and abbreviate  $\eta_H(v_H) := \eta_H(\mathcal{T}_H, v_H)$ . As far as the estimator is concerned, we assume the following *axioms of adaptivity* from [8] for all  $\mathcal{T}_H \in \mathbb{T}$  and all  $\mathcal{T}_h \in \text{refine}(\mathcal{T}_H)$ , where  $C_{\text{stab}}, C_{\text{rel}} > 0$  and  $0 < q_{\text{red}} < 1$  are generic constants:

- (A1) stability:  $|\eta_h(\mathcal{V}_h, v_h) - \eta_H(\mathcal{V}_H, v_H)| \leq C_{\text{stab}} \|v_h - v_H\|$  for all  $v_h \in \mathcal{X}_h, v_H \in \mathcal{X}_H$  and all  $\mathcal{V}_h \subseteq \mathcal{T}_h \cap \mathcal{T}_H$ ;
- (A2) reduction:  $\mathcal{X}_H \subseteq \mathcal{X}_h$  and  $\eta_h(\mathcal{T}_h \setminus \mathcal{T}_H, v_H) \leq q_{\text{red}} \eta_H(\mathcal{T}_H \setminus \mathcal{T}_h, v_H)$  for all  $v_H \in \mathcal{X}_H$ ;
- (A3) reliability:  $\|u^* - u_H^*\| \leq C_{\text{rel}} \eta_H(u_H^*)$ ;
- (A4) discrete reliability:  $\|u_h^* - u_H^*\| \leq C_{\text{rel}} \eta_H(\mathcal{T}_H \setminus \mathcal{T}_h, u_H^*)$ .

We stress that the exact discrete solutions  $u_H^*$  (resp.  $u_h^*$ ) in (A3–A4) will never be computed but are only auxiliary quantities for the analysis.

We refer to Sect. 7.1 below for precise assumptions on the nonlinearity  $A(\cdot)$  of problem (1) such that the standard residual error estimator satisfies (A1–A4) for lowest-order Courant finite elements; see also Sects. 7.2–7.3.

### 2.4 Algebraic solver

For given linear and continuous functionals  $G \in \mathcal{X}'$ , we consider linear systems of algebraic equations of the type

$$(u_H^b, v_H) = G(v_H) \quad \text{for all } v_H \in \mathcal{X}_H \tag{18}$$

with unique (but not computed) exact solution  $u_H^b \in \mathcal{X}_H$ . We suppose here that we have at hand a *contractive* iterative algebraic solver for problems of the form (18). More precisely, let  $u_H^0 \in \mathcal{X}_H$  be an initial guess and let the solver produce a sequence  $u_H^j \in \mathcal{X}_H, j \geq 1$ . Then, we suppose that there exists a generic constant  $0 < q_{\text{alg}} < 1$  such that

$$\|u_H^b - u_H^j\| \leq q_{\text{alg}} \|u_H^b - u_H^{j-1}\| \quad \text{for all } j \geq 1. \tag{19}$$

Examples for such solvers are suitably preconditioned conjugate gradients or multigrid; see, e.g., Olshanskii and Tyrtshnikov [34] and the references therein.

### 2.5 Adaptive algorithm

For the numerical approximation of problem (11), the present work considers an adaptive algorithm which steers mesh-refinement with index  $\ell$ , a (perturbed) contractive Banach–Picard iteration with index  $k$ , and a contractive algebraic solver with index

**Table 1** Counters and discrete solutions in Algorithm 1

	Counter		Discrete solution		Unavailable exact
	Running	Stopping	Available		
			Running	Stopping	
Mesh	$\ell$	$\underline{\ell}$	$u_\ell^{k,j}$	$u_\ell^{k,j}$	$u_\ell^*$ from (20)
Linearization	$k$	$\underline{k}$	$u_\ell^{k,j}$	$u_\ell^{k,j}$	$u_\ell^{k,*}$ from (21)
Algebraic solver	$j$	$\underline{j}$	$u_\ell^{k,j}$	$u_\ell^{k,j}$	

$j$ . On each step  $(\ell, k, j)$ , it yields an approximation  $u_\ell^{k,j} \in \mathcal{X}_\ell$  to the unique but unavailable  $u_\ell^* \in \mathcal{X}_\ell$  on the mesh  $\mathcal{T}_\ell$  defined by

$$\langle \mathcal{A}u_\ell^*, v_\ell \rangle_{\mathcal{X}' \times \mathcal{X}} = \langle F, v_H \rangle_{\mathcal{X}' \times \mathcal{X}} \quad \text{for all } v_\ell \in \mathcal{X}_\ell. \tag{20}$$

Reporting for the summary of notation to Table 1, the algorithm reads as follows:

**Algorithm 1** *Input:* Initial mesh  $\mathcal{T}_0$  and initial guess  $u_0^{0,0} = u_0^{\underline{j}} \in \mathcal{X}_0$ , parameters  $0 < \theta \leq 1, 0 < \lambda_{\text{alg}} < 1, 0 < \lambda_{\text{Pic}}, 1 \leq C_{\text{mark}}$ , counters  $\ell = k = j = 0$ , tolerance  $\tau \geq 0$ .

**Repeat** the following steps (i–vi) (adaptive mesh refinement loop):

(i) **Repeat** the following steps (a–c) (linearization loop):

- (a) Define  $u_\ell^{k+1,0} := u_\ell^{k,j}$  and update counters  $k := k + 1$  as well as  $j := 0$ .
- (b) **Repeat** the following steps (I–IV) (algebraic solver loop):
  - (I) Update counter  $j := j + 1$ .
  - (II) Consider the problem of finding

$$u_\ell^{k,*} \in \mathcal{X}_\ell \text{ such that, for all } v_\ell \in \mathcal{X}_\ell, \tag{21}$$

$$(u_\ell^{k,*}, v_\ell) = (u_\ell^{k-1,\underline{j}}, v_\ell) - \frac{\alpha}{L^2} \langle \mathcal{A}u_\ell^{k-1,\underline{j}} - F, v_\ell \rangle_{\mathcal{X}' \times \mathcal{X}}$$

and do one step of the algebraic solver applied to (21) starting from  $u_\ell^{k,j-1}$ , which yields  $u_\ell^{k,j}$  (an approximation to  $u_\ell^{k,*}$ ).

- (III) Compute the local indicators  $\eta_\ell(T, u_\ell^{k,j})$  for all  $T \in \mathcal{T}_\ell$ .
- (IV) If  $\eta_\ell(u_\ell^{k,j}) + \|u_\ell^{k,j} - u_\ell^{k-1,\underline{j}}\| + \|u_\ell^{k,j} - u_\ell^{k,j-1}\| \leq \tau$ , then set  $\underline{\ell} := \ell, \underline{k}(\underline{\ell}) := k$ , and  $\underline{j}(\underline{\ell}, \underline{k}) := j$  and terminate Algorithm 1.

$$\textbf{Until} \quad \|u_\ell^{k,j} - u_\ell^{k,j-1}\| \leq \lambda_{\text{alg}} [\eta_\ell(u_\ell^{k,j}) + \|u_\ell^{k,j} - u_\ell^{k-1,\underline{j}}\|]. \tag{22}$$

(c) Define  $\underline{j} := \underline{j}(\underline{\ell}, \underline{k}) := j$ .

$$\text{Until } \|u_\ell^{k,j} - u_\ell^{k-1,j}\| \leq \lambda_{\text{Pic}} \eta_\ell(u_\ell^{k,j}). \tag{23}$$

- (ii) Define  $\underline{k} := \underline{k}(\ell) := k$ .
- (iii) If  $\eta_\ell(u_\ell^{\underline{k},j}) = 0$ , then set  $\underline{\ell} := \ell$  and terminate Algorithm 1.
- (iv) Determine a set  $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$  with up to the multiplicative constant  $C_{\text{mark}}$  minimal cardinality such that

$$\theta \eta_\ell(u_\ell^{\underline{k},j}) \leq \eta_\ell(\mathcal{M}_\ell, u_\ell^{\underline{k},j}). \tag{24}$$

- (v) Generate  $\mathcal{T}_{\ell+1} := \text{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$  and define  $u_{\ell+1}^{0,0} := u_{\ell+1}^{0,j} := u_\ell^{\underline{k},j}$ .
- (vi) Update counters  $\ell := \ell + 1, k := 0$ , and  $j := 0$  and continue with (i).

**Output:** Sequence of discrete solutions  $u_\ell^{k,j}$  and corresponding error estimators  $\eta_\ell(u_\ell^{k,j})$ .

Some remarks are in order to explain the nature of Algorithm 1. The innermost loop (Algorithm 1(ib)) steers the algebraic solver. Note here that the exact solution  $u_\ell^{k,\star}$  of (21) is not computed but only approximated by the computed iterates  $u_\ell^{k,j}$ . For the linear system (21), the contraction assumption (19) reads as

$$\|u_\ell^{k,\star} - u_\ell^{k,j}\| \leq q_{\text{alg}} \|u_\ell^{k,\star} - u_\ell^{k,j-1}\| \text{ for all } j \geq 1. \tag{25}$$

Then, the triangle inequality implies that

$$\frac{1 - q_{\text{alg}}}{q_{\text{alg}}} \|u_\ell^{k,\star} - u_\ell^{k,j}\| \leq \|u_\ell^{k,j} - u_\ell^{k,j-1}\| \leq (1 + q_{\text{alg}}) \|u_\ell^{k,\star} - u_\ell^{k,j-1}\|. \tag{26}$$

Hence, the term  $\|u_\ell^{k,j} - u_\ell^{k,j-1}\|$  provides a means to estimate the algebraic error  $\|u_\ell^{k,\star} - u_\ell^{k,j}\|$ . In particular, the approximation  $u_\ell^{k,j}$  is accepted and the algebraic solver is stopped if the algebraic error estimate  $\|u_\ell^{k,j} - u_\ell^{k,j-1}\|$  is, up to the threshold  $\lambda_{\text{alg}}$ , below the estimate  $\eta_\ell(u_\ell^{k,j}) + \|u_\ell^{k,j} - u_\ell^{k-1,j}\|$  of the discretization and linearization error; see (22). Since  $\|u_\ell^{k,1} - u_\ell^{k,0}\| = \|u_\ell^{k,1} - u_\ell^{k-1,j}\|$ , the stopping criterion (22) would always terminate the algebraic solver at the first step  $j = 1$  if  $\lambda_{\text{alg}}$  was chosen greater or equal to 1; this motivates the restriction  $\lambda_{\text{alg}} < 1$ .

The middle loop (Algorithm 1(i)) steers the linearization by means of the (perturbed) Banach–Picard iteration. Lemma 6 below shows that the term  $\|u_\ell^{k,j} - u_\ell^{k-1,j}\|$  estimates the linearization error  $\|u_\ell^\star - u_\ell^{k,j}\|$ . Note here that, a priori, only the non-perturbed Banach–Picard iteration corresponding to the (unavailable) exact solve of (21) yielding  $u_\ell^{k,\star}$  would lead to the contraction

$$\|u_\ell^\star - u_\ell^{k,\star}\| \leq q_{\text{Pic}} \|u_\ell^\star - u_\ell^{k-1,j}\| \text{ for all } (\ell, k, 0) \in \mathcal{Q} \text{ with } k \geq 1, \tag{27}$$

where

$$0 < q_{\text{Pic}} := (1 - \alpha^2/L^2)^{1/2} < 1. \tag{28}$$

The approximation  $u_\ell^{k,j}$  is accepted and the linearization is stopped if the linearization error estimate  $\|u_\ell^{k,j} - u_\ell^{k-1,j}\|$  is, up to the threshold  $\lambda_{pic}$ , below the discretization error estimate  $\eta_\ell(u_\ell^{k,j})$ ; see (23) (here  $\lambda_{pic} < 1$  is not necessary).

Finally, the outermost loop steers the local adaptive mesh-refinement. To this end, the Dörfler marking criterion (24) from [17] is employed to mark elements  $T \in \mathcal{M}_\ell$  for refinement, unless  $\eta_\ell(u_\ell^{k,j}) = 0$ , in which case Proposition 2 below ensures that the approximation  $u_\ell^{k,j}$  coincides with the exact solution  $u^*$  of (11). In practice, the computation is stopped as soon as the computed iterate  $u_\ell^{k,j}$  is sufficiently accurate with respect to a nonzero tolerance  $\tau$ . Based on the *a posteriori* error estimate from Proposition 2 below, this motivates the termination in Algorithm 1(IV).

### 2.6 Index set $\mathcal{Q}$ for the triple loop

To analyze the asymptotic convergence behavior of Algorithm 1 for tolerance  $\tau = 0$ , we define the index set

$$\mathcal{Q} := \{(\ell, k, j) \in \mathbb{N}_0^3 : \text{index triple } (\ell, k, j) \text{ is used in Algorithm 1}\}. \tag{29}$$

Since Algorithm 1 is sequential, the index set  $\mathcal{Q}$  is naturally ordered. For indices  $(\ell, k, j), (\ell', k', j') \in \mathcal{Q}$ , we write

$$(\ell, k, j) < (\ell', k', j') \stackrel{\text{def}}{\iff} (\ell, k, j) \text{ appears earlier in Algorithm 1 than } (\ell', k', j'). \tag{30}$$

With this order, we can define

$$|(\ell, k, j)| := \#\{(\ell', k', j') \in \mathcal{Q} : (\ell', k', j') < (\ell, k, j)\},$$

which is the *total step number* of Algorithm 1. We make the following definitions, which are consistent with that of Algorithm 1, and additionally define  $\underline{j}(\ell, 0) := 0$ :

$$\begin{aligned} \underline{\ell} &:= \sup \{ \ell \in \mathbb{N}_0 : (\ell, 0, 0) \in \mathcal{Q} \} \in \mathbb{N}_0 \cup \{\infty\}, \\ \underline{k}(\ell) &:= \sup \{ k \in \mathbb{N}_0 : (\ell, k, 0) \in \mathcal{Q} \} \in \mathbb{N}_0 \cup \{\infty\} \text{ if } (\ell, 0, 0) \in \mathcal{Q}, \\ \underline{j}(\ell, k) &:= \sup \{ j \in \mathbb{N}_0 : (\ell, k, j) \in \mathcal{Q} \} \in \mathbb{N}_0 \cup \{\infty\} \text{ if } (\ell, k, 0) \in \mathcal{Q}. \end{aligned}$$

Generically, it holds that  $\underline{\ell} = \infty$ , i.e., infinitely many steps of mesh-refinement take place when  $\tau = 0$ . However, our analysis also covers the cases that either the  $k$ -loop (linearization) or the  $j$ -loop (algebraic solver) does not terminate, i.e.,

$$\underline{k}(\ell) = \infty \text{ if } \underline{\ell} < \infty \text{ resp. } \underline{j}(\ell, k) = \infty \text{ if } \underline{\ell} < \infty \text{ and } \underline{k}(\ell) < \infty,$$

or that the exact solution  $u^*$  is hit at step (iii) of Algorithm 1 (note that  $\eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,\underline{j}}) = 0$  implies  $u^* = u_{\underline{\ell}}^{k,\underline{j}}$  by virtue of Proposition 2 below).

To abbreviate notation, we make the following convention: If the mesh index  $\ell \in \mathbb{N}_0$  is clear from the context, we simply write  $\underline{k} := \underline{k}(\ell)$ , e.g.,  $u_{\underline{\ell}}^{k,j} := u_{\underline{\ell}}^{k(\ell),j}$ . Similarly, we simply write  $\underline{j} := \underline{j}(\ell, k)$ , e.g.,  $u_{\underline{\ell}}^{k,\underline{j}} := u_{\underline{\ell}}^{k,j(\ell,k)}$ .

Note that there in particular holds  $u_{\underline{\ell}-1}^{k,\underline{j}} = u_{\underline{\ell}}^{0,0} = u_{\underline{\ell}}^{1,0}$  for all  $(\ell, 0, 0) \in \mathcal{Q}$  with  $\ell \geq 1$ . Hence, these approximate solutions are indexed three times. This is our notational choice that will not be harmful for what follows; alternatively, one could only index the approximate solutions that appear on step (i.b.II) of Algorithm 1.

### 2.7 Main results

Our first proposition provides computable upper bounds for the energy error  $\|u^* - u_{\underline{\ell}}^{k,j}\|$  of the iterates  $u_{\underline{\ell}}^{k,j}$  of Algorithm 1 at any step  $(\ell, k, j) \in \mathcal{Q}$ . In particular, we note that the stopping criteria (22) and (23) ensure reliability of  $\eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,\underline{j}})$  for the final perturbed Banach–Picard iterates  $u_{\underline{\ell}}^{k,\underline{j}}$ . The proof is postponed to Sect. 3.3.

**Proposition 2** (Reliability at various stages of Algorithm 1) *Suppose (A1) and (A3). Then, for all  $(\ell, k, j) \in \mathcal{Q}$ , it holds that*

$$\|u^* - u_{\underline{\ell}}^{k,j}\| \leq C'_{\text{rel}} \begin{cases} \eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,j}) + \|u_{\underline{\ell}}^{k,j} - u_{\underline{\ell}}^{k-1,\underline{j}}\| + \|u_{\underline{\ell}}^{k,j} - u_{\underline{\ell}}^{k,j-1}\| & \text{if } 0 < k \leq \underline{k}(\ell) \text{ and } 0 < j \leq \underline{j}(\ell, k), \\ \eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,\underline{j}}) + \|u_{\underline{\ell}}^{k,\underline{j}} - u_{\underline{\ell}}^{k-1,\underline{j}}\| & \text{if } 0 < k \leq \underline{k}(\ell) \text{ and } j = \underline{j}(\ell, k), \\ \eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,\underline{j}}) & \text{if } k = \underline{k}(\ell) \text{ and } \underline{j} = \underline{j}(\ell, \underline{k}), \\ \eta_{\underline{\ell}-1}(u_{\underline{\ell}-1}^{k,\underline{j}}) & \text{if } k = 0 \text{ and } \ell > 0. \end{cases} \tag{31}$$

The constant  $C'_{\text{rel}} > 0$  depends only on  $C_{\text{rel}}, C_{\text{stab}}, q_{\text{alg}}, \lambda_{\text{alg}}, q_{\text{Pic}},$  and  $\lambda_{\text{Pic}}$ .

The first main theorem states *linear convergence in each step* of the adaptive algorithm, i.e., algebraic solver or linearization or mesh-refinement. The proof is given in Sect. 4.

**Theorem 3** (linear convergence) *Suppose (A1)–(A3). Then, there exist  $\lambda_{\text{alg}}^*, \lambda_{\text{Pic}}^* > 0$  such that for arbitrary  $0 < \theta \leq 1$  as well as for all  $0 < \lambda_{\text{alg}} < 1$  and  $0 < \lambda_{\text{Pic}}$  with  $0 < \lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} < \lambda_{\text{alg}}^*$  and  $0 < \lambda_{\text{Pic}}/\theta < \lambda_{\text{Pic}}^*$ , there exist constants  $1 \leq C_{\text{lin}}$  and  $0 < q_{\text{lin}} < 1$  such that the quasi-error*

$$\Delta_{\underline{\ell}}^{k,j} := \|u^* - u_{\underline{\ell}}^{k,j}\| + \|u_{\underline{\ell}}^{k,*} - u_{\underline{\ell}}^{k,j}\| + \eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,j}), \tag{32}$$

composed of the overall error, the algebraic error, and the error estimator, is linearly convergent in the sense of

$$\Delta_{\ell'}^{k',j'} \leq C_{\text{lin}} q_{\text{lin}}^{|\ell',k',j'| - |(\ell,k,j)|} \Delta_{\ell}^{k,j} \tag{33}$$

for all  $(\ell, k, j), (\ell', k', j') \in \mathcal{Q}$  with  $(\ell', k', j') \geq (\ell, k, j)$ . The constants  $C_{\text{lin}}$  and  $q_{\text{lin}}$  depend only on  $C_{\text{rel}}, C_{\text{stab}}, q_{\text{red}}, \theta, q_{\text{alg}}, \lambda_{\text{alg}}, q_{\text{Pic}}, \lambda_{\text{Pic}}, \alpha$ , and  $L$ .

Note that  $\Delta_{\ell'}^{k',j'} = \Delta_{\ell}^{k,j}$  when  $(\ell', k', j') = (\ell, k, j)$ , and then (33) holds with equality for  $C_{\text{lin}} = 1$ . There are other cases where  $u_{\ell'}^{k',j'} = u_{\ell}^{k,j}$  and where  $u_{\ell'}^{k',j'} = u_{\ell}^{k,j}$  together with  $\mathcal{T}_{\ell'} = \mathcal{T}_{\ell}$ , and consequently  $\eta_{\ell'}(u_{\ell'}^{k',j'}) = \eta_{\ell}(u_{\ell}^{k,j})$ , related to our notational choice for  $\mathcal{Q}$  in (29) that also indexes nested iterates. The case with  $\ell' = \ell$  arises for instance when  $j = \underline{j}, j' = 0$ , and  $k' = k + 1$ ; see step (ia) of Algorithm 1. Note, however, that in such a situation, typically  $u_{\ell'}^{k',\star} \neq u_{\ell}^{k,\star}$ , and consequently  $\Delta_{\ell'}^{k',j'} \neq \Delta_{\ell}^{k,j}$ . A situation where  $\Delta_{\ell'}^{k',j'} = \Delta_{\ell}^{k,j}$  for  $(\ell', k', j') \neq (\ell, k, j)$  can nevertheless also appear and is covered in (33). For instance, in the above example, when  $j = \underline{j}, j' = 0, k' = k + 1$ , and  $\ell' = \ell$ , and where moreover  $u_{\ell}^{k,j} = u_{\ell}^{k,\star} = u_{\ell}^{\star}$  (so that  $u_{\ell}^{k,j} = u_{\ell}^{k,\star} = u_{\ell'}^{k',\star} = u_{\ell'}^{k',j'} = u_{\ell}^{\star}$ ), Algorithm 1 performs only one step of the algebraic solver on the linearization step  $k'$ , so that  $C_{\text{lin}} = 1/q_{\text{lin}}$  leads to equality in (33) where now  $|\ell', k', j'| - |(\ell, k, j)| = 1$ .

The second main result states optimal decay rate of the quasi-error  $\Delta_{\ell}^{k,j}$  of (32) (and consequently of the total error  $\|u^{\star} - u_{\ell}^{k,j}\|$ ) in terms of the number of degrees of freedom added in the space  $\mathcal{X}_{\ell}$  with respect to  $\mathcal{X}_0$ . More precisely, the result states that if the unknown weak solution  $u$  of (11) can be approximated at algebraic decay rate  $s$  with respect to the number of mesh elements added in the refinement of  $\mathcal{T}_0$  (plus one) for a best-possible mesh, then Algorithm 1 achieves the same decay rate  $s$  with respect to the number of elements actually added in Algorithm 1,  $(\#\mathcal{T}_{\ell} - \#\mathcal{T}_0 + 1)$ , up to a generic multiplicative constant. The proof of the following Theorem 4 is given in Sect. 5.

**Theorem 4** (optimal decay rate wrt. degrees of freedom) *Suppose (A1–A4) and (R1–R3). Recall  $\lambda_{\text{alg}}^{\star}, \lambda_{\text{Pic}}^{\star} > 0$  from Theorem 3. Let  $C_{\text{Pic}} := q_{\text{Pic}}/(1 - q_{\text{Pic}}) > 0, C_{\text{alg}} := q_{\text{alg}}/(1 - q_{\text{alg}}) > 0$ , and  $\theta_{\text{opt}} := (1 + C_{\text{stab}}^2 C_{\text{rel}}^2)^{-1}$ . Then, there exists  $\theta^{\star}$  such that for all  $0 < \lambda_{\text{alg}}, \lambda_{\text{Pic}}, \theta$  with  $0 < \theta < \min\{1, \theta^{\star}\}$  as well as  $\lambda_{\text{alg}} < 1, 0 < \lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} < \lambda_{\text{alg}}^{\star}$ , and  $0 < \lambda_{\text{Pic}}/\theta < \lambda_{\text{Pic}}^{\star}$ , it holds that*

$$0 < \theta' := \frac{\theta + C_{\text{stab}} \left( (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}} + [C_{\text{Pic}} + (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}}] \lambda_{\text{Pic}} \right)}{1 - \lambda_{\text{Pic}} / \lambda_{\text{Pic}}^{\star}} < \theta_{\text{opt}}, \tag{34}$$

where the constant  $\theta^{\star} > 0$  depends only on  $C_{\text{stab}}, q_{\text{Pic}}$ , and  $q_{\text{alg}}$ . Let  $s > 0$  and define

$$\|u^*\|_{\mathbb{A}_s} := \sup_{N \in \mathbb{N}_0} \left( (N + 1)^s \inf_{\mathcal{T}_{\text{opt}} \in \mathbb{T}(N)} \eta_{\text{opt}}(u^*_{\text{opt}}) \right) \in \mathbb{R}_{\geq 0} \cup \{\infty\}, \tag{35}$$

where  $\eta_{\text{opt}}(u^*_{\text{opt}})$  is the error estimator corresponding to the exact solution of (12) with respect to the mesh  $\mathcal{T}_{\text{opt}}$  and

$$\mathbb{T}(N) := \{ \mathcal{T} \in \mathbb{T} : \#\mathcal{T} - \#\mathcal{T}_0 \leq N \}.$$

Then, for tolerance  $\tau = 0$ , there exist  $c_{\text{opt}}, C_{\text{opt}} > 0$  such that

$$c_{\text{opt}}^{-1} \|u^*\|_{\mathbb{A}_s} \leq \sup_{(\ell, k, j) \in \mathcal{Q}} (\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^s \Delta_\ell^{k, j} \leq C_{\text{opt}} \max\{ \|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0} \}. \tag{36}$$

The constant  $c_{\text{opt}} > 0$  depends only on  $C_{\text{Céa}} = L/\alpha, C_{\text{stab}}, C_{\text{rel}}, C_{\text{son}}, \#\mathcal{T}_0, s$ , and, if  $\underline{\ell} < \infty$ , additionally on  $\underline{\ell}$ . The constant  $C_{\text{opt}} > 0$  depends only on  $C_{\text{stab}}, C_{\text{rel}}, C_{\text{mark}}, 1 - \lambda_{\text{Pic}}/\lambda_{\text{Pic}}^*, C_{\text{Céa}} = L/\alpha, C'_{\text{rel}}, C_{\text{mesh}}, C_{\text{lin}}, q_{\text{lin}}, \#\mathcal{T}_0$ , and  $s$ . The maximum in the right inequality is only needed if  $\ell = 0$ . If  $\ell \geq 1$ , the maximum  $\max\{ \|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0} \}$  can be replaced by  $\|u^*\|_{\mathbb{A}_s}$ .

Note that  $\Delta_0^{0,0}$  can be arbitrarily bad due to a bad initial guess  $u_0^{0,0}$ . However,  $\|u^*\|_{\mathbb{A}_s}$  as well as the constant  $C_{\text{opt}}$  are independent of the initial guess, so that the upper bound in (36) cannot avoid  $\max\{ \|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0} \}$  for the case  $\ell = 0$ . Such a phenomenon does not appear at later stages, since the stopping criteria (22) and (23) ensure that, though  $u_\ell^{k, j}$  does not in general coincide with  $u_\ell^*$ , it is sufficiently accurate. If one restricts the indices to  $(\ell, k, j) \in \mathcal{Q}$  with  $\ell \geq 1$ , then the upper bound in (36) may omit  $\Delta_0^{0,0}$ .

Our last main result states that Algorithm 1 drives the quasi-error down at each possible rate  $s$  not only with respect to the number of degrees of freedom added in the space  $\mathcal{X}_\ell$  in comparison with  $\mathcal{X}_0$ , but actually also with respect to the overall computational cost expressed as a cumulated sum of the number of degrees of freedom. This is an important improvement of Theorem 4. More precisely, under the same conditions as above, i.e., if the unknown weak solution  $u$  of (11) can be approximated at algebraic decay rate  $s$  with respect to the number of mesh elements added in the refinement of  $\mathcal{T}_0$  (plus one), Algorithm 1 generates a sequence of triple- $(\ell, k, j)$ -indexed approximations (mesh, linearization, algebraic solver) such that the quasi-error decays at rate  $s$  with respect to the overall algorithmic cost expressed as the sum of the number of simplices  $\#\mathcal{T}_\ell$  over all steps  $(\ell, k, j) \in \mathcal{Q}$  effectuated by Algorithm 1. The proof of the following Theorem 5 is given in Sect. 6.

**Theorem 5** (optimal decay rate wrt. overall computational cost) *Let the assumptions of Theorem 4 be verified. Then*

$$c_{\text{opt}}^{-1} \|u^*\|_{\mathbb{A}_s} \leq \sup_{(\ell', k', j') \in \mathcal{Q}} \left( \sum_{\substack{(\ell, k, j) \in \mathcal{Q} \\ (\ell, k, j) \leq (\ell', k', j')}} \#\mathcal{T}_\ell \right)^s \Delta_{\ell'}^{k', j'} \leq C'_{\text{opt}} \max\{ \|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0} \}. \tag{37}$$

The maximum in the right inequality is only needed if  $\ell = 0$ . If  $\ell \geq 1$ , the maximum  $\max\{\|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0}\}$  can be replaced by  $\|u^*\|_{\mathbb{A}_s}$ . While  $c_{\text{opt}} > 0$  is the constant of Theorem 4, the constant  $C'_{\text{opt}} > 0$  reads  $C'_{\text{opt}} := (\#\mathcal{T}_0)^s C_{\text{opt}} C_{\text{lin}} (1 - q_{\text{lin}}^{1/s})^{-s}$ .

Analogously to the comments after Theorem 4, the upper estimate in (37) cannot avoid  $\max\{\|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0}\}$  for the case  $\ell' = \ell = 0$ . As above, if one restricts the indices to  $(\ell', k', j'), (\ell, k, j) \in \mathcal{Q}$  with  $\ell', \ell \geq 1$ , then the upper bound in (37) may omit  $\Delta_0^{0,0}$ .

Note that for any reasonable algebraic solver on mesh  $\mathcal{T}_\ell$ , the cost of its one step is proportional to  $\#\mathcal{T}_\ell$ . This also holds true for matrix and right-hand-side assembly in (21), evaluation of the residual estimators  $\eta_\ell(u_\ell^{k,j})$ , Dörfler marking, and local adaptive mesh refinement by, e.g., newest vertex bisection, while the cost of evaluation of the stopping criteria (22) and (23) is of  $\mathcal{O}(1)$ . Thus, the sum in (37) is indeed proportional to the overall computational cost invested into the numerical approximation of (1) by Algorithm 1.

### 3 Auxiliary results

#### 3.1 Some observations on Algorithm 1

This section collects some elementary observations on Algorithm 1 in what concerns nested iteration and stopping criteria. The given initial value of Algorithm 1 reads

$$u_0^{0,0} = u_0^{0,j} = u_0^{0,\star} \in \mathcal{X}_0. \tag{38}$$

If  $(\ell, 0, 0) \in \mathcal{Q}$  with  $\ell \geq 1$ , then

$$u_\ell^{0,\star} := u_\ell^{0,0} := u_\ell^{0,j} := u_{\ell-1}^{k,j} \in \mathcal{X}_{\ell-1} \subseteq \mathcal{X}_\ell. \tag{39}$$

If  $(\ell, k, 0) \in \mathcal{Q}$ , then the initial guess for the algebraic solver reads

$$u_\ell^{k,0} = \begin{cases} u_0^{0,0} & \text{for } \ell = 0, \\ u_{\ell-1}^{k,j} & \text{if } k = 0 \text{ and } \ell \geq 1, \\ u_\ell^{k-1,j} & \text{if } k > 0, \end{cases} \tag{40}$$

i.e., the algebraic solver employs *nested iteration*. The stopping criterion (22) of Algorithm 1 guarantees that  $\underline{j}(\ell, k) \geq 1$  if  $k > 0$  and, for all  $(\ell, k, j) \in \mathcal{Q}$ , it holds that

$$\|u_\ell^{k,j} - u_\ell^{k,j-1}\| \leq \lambda_{\text{alg}} [\eta_\ell(u_\ell^{k,j}) + \|u_\ell^{k,j} - u_\ell^{k-1,j}\|] \quad \text{for } j = \underline{j}(\ell, k), \tag{41}$$

$$\|u_\ell^{k,j} - u_\ell^{k,j-1}\| > \lambda_{\text{alg}} [\eta_\ell(u_\ell^{k,j}) + \|u_\ell^{k,j} - u_\ell^{k-1,j}\|] \quad \text{for } j < \underline{j}(\ell, k), \tag{42}$$

i.e., the algebraic error estimate  $\|u_\ell^{k,j} - u_\ell^{k,j-1}\|$  only drops below the discretization plus linearization error estimate at the stopping iteration  $j = \underline{j}(\ell, k)$ .



The final iterates  $u_\ell^{k,j}$  of the algebraic solver are used to obtain the perturbed Banach–Picard iterates  $u_\ell^{k+1,j}$  for  $k > 0$ ; see (21). The stopping criterion (23) of Algorithm 1 guarantees that  $\underline{k}(\ell) \geq 1$  and, for all  $(\ell, k, j) \in \mathcal{Q}$ , it holds that

$$\|u_\ell^{k,j} - u_\ell^{k-1,j}\| \leq \lambda_{\text{Pic}} \eta_\ell(u_\ell^{k,j}) \quad \text{for } k = \underline{k}(\ell), \tag{43}$$

$$\|u_\ell^{k,j} - u_\ell^{k-1,j}\| > \lambda_{\text{Pic}} \eta_\ell(u_\ell^{k,j}) \quad \text{for } k < \underline{k}(\ell), \tag{44}$$

i.e., the linearization error estimate  $\|u_\ell^{k,j} - u_\ell^{k-1,j}\|$  only drops below the discretization error estimate at the stopping iteration  $\underline{k} = \underline{k}(\ell)$ .

### 3.2 Contraction of the perturbed Banach–Picard iteration

Assumption (19) immediately implies the algebraic solver contraction (25) and reliability (26) of the algebraic error estimate  $\|u_\ell^{k,j} - u_\ell^{k,j-1}\|$ . Similarly, one step of the non-perturbed Banach–Picard iteration (21) (i.e., with an exact algebraic solve of problem (21) with the datum  $u_\ell^{k-1,j}$ ) leads to contraction (27) and consequently to the reliability

$$\frac{1 - q_{\text{Pic}}}{q_{\text{Pic}}} \|u_\ell^* - u_\ell^{k,*}\| \leq \|u_\ell^{k,*} - u_\ell^{k-1,j}\| \leq (1 + q_{\text{Pic}}) \|u_\ell^* - u_\ell^{k-1,j}\| \tag{45}$$

of the unavailable linearization error estimate  $\|u_\ell^{k,*} - u_\ell^{k-1,j}\|$ . As our first result, we now show that, for sufficiently small stopping parameters  $0 < \lambda_{\text{alg}}$  in (22), we also get that the *perturbed* Banach–Picard iteration is a *contraction*. Recall that  $u_\ell^* \in \mathcal{X}_\ell$  is the (unavailable) exact discrete solution given by (20), that  $u_\ell^{k,*} \in \mathcal{X}_\ell$  is the (unavailable) exact linearization solution given by (21), and that  $u_\ell^{k,j} \in \mathcal{X}_\ell$  is the computed solution for which the algebraic solver is stopped; see (22) (resp. (41) and (42)) for the stopping criterion.

**Lemma 6** *There exists  $\lambda_{\text{alg}}^* > 0$  depending only on  $q_{\text{alg}}$  and  $q_{\text{Pic}}$  such that*

$$0 < q'_{\text{Pic}} := \frac{q_{\text{Pic}} + \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}}^*}{1 - \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}}^*} < 1. \tag{46}$$

*Moreover, for all stopping parameters  $0 < \lambda_{\text{alg}} < 1$  and  $0 < \lambda_{\text{Pic}}$  from (22) and (23) such that  $0 < \lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} < \lambda_{\text{alg}}^*$ , it holds that*

$$\|u_\ell^* - u_\ell^{k,j}\| \leq q'_{\text{Pic}} \|u_\ell^* - u_\ell^{k-1,j}\| \quad \text{for all } 1 \leq k < \underline{k}(\ell). \tag{47}$$

This also implies that

$$\frac{1 - q'_{\text{Pic}}}{q'_{\text{Pic}}} \|u_\ell^\star - u_\ell^{k,j}\| \leq \|u_\ell^{k,j} - u_\ell^{k-1,j}\| \leq (1 + q'_{\text{Pic}}) \|u_\ell^\star - u_\ell^{k-1,j}\|. \tag{48}$$

**Proof** Clearly, (48) follows from (47) by the triangle inequality as in (26) and (45). Moreover, (46) is obvious for sufficiently small  $\lambda_{\text{alg}}^\star$ , since  $q_{\text{Pic}} = (1 - \alpha^2/L^2)^{1/2} < 1$  from (28) and  $0 < q_{\text{alg}} < 1$  is fixed from (19). To see (47), first note that

$$\|u_\ell^\star - u_\ell^{k,j}\| \leq \|u_\ell^\star - u_\ell^{k,\star}\| + \|u_\ell^{k,\star} - u_\ell^{k,j}\| \stackrel{(27)}{\leq} q_{\text{Pic}} \|u_\ell^\star - u_\ell^{k-1,j}\| + \|u_\ell^{k,\star} - u_\ell^{k,j}\|,$$

where the first term corresponds to the unperturbed Banach–Picard iteration (21) and the second to the algebraic error. Second, note that, since  $1 \leq k < \underline{k}(\ell)$ ,

$$\begin{aligned} \|u_\ell^{k,\star} - u_\ell^{k,j}\| &\stackrel{(26)}{\leq} \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \|u_\ell^{k,j} - u_\ell^{k,j-1}\| \\ &\stackrel{(41)}{\leq} \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}} [\eta_\ell(u_\ell^{k,j}) + \|u_\ell^{k,j} - u_\ell^{k-1,j}\|] \\ &\stackrel{(44)}{<} \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} (\lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}}) \|u_\ell^{k,j} - u_\ell^{k-1,j}\| \\ &\leq \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} (\lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}}) [\|u_\ell^\star - u_\ell^{k,j}\| + \|u_\ell^\star - u_\ell^{k-1,j}\|]. \end{aligned}$$

Combining the latter estimates with the assumption  $\lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} < \lambda_{\text{alg}}^\star$ , we see that

$$\|u_\ell^\star - u_\ell^{k,j}\| \leq (q_{\text{Pic}} + \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}}^\star) \|u_\ell^\star - u_\ell^{k-1,j}\| + \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}}^\star \|u_\ell^\star - u_\ell^{k,j}\|.$$

If  $0 < \lambda_{\text{alg}}^\star$  is sufficiently small, this shows (46) and (47) and concludes the proof.  $\square$

### 3.3 Proof of Proposition 2 (reliable error control in Algorithm 1)

We are now ready to prove the estimates (31).

**Proof of Proposition 2** First, let  $(\ell, k, j) \in \mathcal{Q}$  with  $0 < k \leq \underline{k}(\ell)$  and  $0 < j \leq \underline{j}(\ell, k)$ . Due to stability (A1), reliability (A3), and the contraction properties (26) resp. (45), it holds that

$$\begin{aligned}
 \|u^* - u_\ell^{k,j}\| &\leq \|u^* - u_\ell^*\| + \|u_\ell^* - u_\ell^{k,j}\| \stackrel{(A3)}{\lesssim} \eta_\ell(u_\ell^*) + \|u_\ell^* - u_\ell^{k,j}\| \\
 &\stackrel{(A1)}{\lesssim} \eta_\ell(u_\ell^{k,j}) + \|u_\ell^* - u_\ell^{k,j}\| \leq \eta_\ell(u_\ell^{k,j}) + \|u_\ell^* - u_\ell^{k,*}\| + \|u_\ell^{k,*} - u_\ell^{k,j}\| \\
 &\stackrel{(45)}{\lesssim} \eta_\ell(u_\ell^{k,j}) + \|u_\ell^{k,*} - u_\ell^{k-1,j}\| + \|u_\ell^{k,*} - u_\ell^{k,j}\| \\
 &\leq \eta_\ell(u_\ell^{k,j}) + \|u_\ell^{k,j} - u_\ell^{k-1,j}\| + 2\|u_\ell^{k,*} - u_\ell^{k,j}\| \\
 &\stackrel{(26)}{\lesssim} \eta_\ell(u_\ell^{k,j}) + \|u_\ell^{k,j} - u_\ell^{k-1,j}\| + \|u_\ell^{k,j} - u_\ell^{k,j-1}\|. \tag{49}
 \end{aligned}$$

This proves (31) for the case  $0 < k \leq \underline{k}(\ell)$  and  $0 < j \leq \underline{j}(\ell, k)$ .

If  $j = \underline{j}(\ell, k)$ , we can improve estimate (49) using the stopping criterion (41). Similarly, if  $k = \underline{k}(\ell)$  and  $j = \underline{j}(\ell, \underline{k})$ , we can improve estimate (49) using the stopping criteria (41) and (43). Finally, for  $k = 0, \ell > 0$  and hence  $j = \underline{j} = 0$ , it directly follows from nested iteration (39) and the previous case  $k = \underline{k}(\ell - 1)$  resp.  $j = \underline{j}(\ell - 1, k)$  that

$$\|u^* - u_\ell^{0,0}\| \stackrel{(40)}{=} \|u^* - u_{\ell-1}^{\underline{k}, \underline{j}}\| \lesssim \eta_{\ell-1}(u_{\ell-1}^{\underline{k}, \underline{j}}). \tag{50}$$

This concludes the proof. □

### 3.4 An auxiliary adaptive algorithm

Due to Lemma 6, the iterates  $u_\ell^{k,j}$  are contractive in the index  $k$ . Consequently, Algorithm 1 fits into the framework of [23] upon defining  $u_\ell$  from [23] as  $u_\ell := u_\ell^{\underline{k}, \underline{j}}$  for the case where  $\underline{k}(\ell) < \infty$  and  $\underline{j}(\ell, k) < \infty$ , i.e., both the algebraic and the linearization solvers are stopped by (22) and (23) on the mesh  $\mathcal{T}_\ell$ . Note that the assumption  $(\ell + n + 1, 0, 0) \in \mathcal{Q}$  below ensures this for all meshes  $\mathcal{T}_{\ell'}$  with  $0 \leq \ell' \leq \ell + n$ . Then, we can rewrite [23, Lemma 4.9, eq. (4.10)] and [23, Theorem 5.3, eq. (5.5)] in the current setting to conclude two important properties: First, the estimators  $\eta_\ell(u_\ell^{\underline{k}, \underline{j}})$  available at step (iv) of Algorithm 1 are, up to a constant, equivalent to the estimators  $\eta_\ell(u_\ell^*)$  corresponding to the unavailable exact linearization  $u_\ell^*$  of (20). And second, the estimators  $\eta_\ell(u_\ell^{\underline{k}, \underline{j}})$  are linearly convergent.

**Lemma 7** [23, Lemma 4.9, Theorem 5.3] *Recall  $\lambda_{\text{alg}}^* > 0$  and  $0 < q'_{\text{Pic}} < 1$  from Lemma 6. Define  $\lambda_{\text{Pic}}^* := \frac{1 - q'_{\text{Pic}}}{q_{\text{Pic}} C_{\text{stab}}} > 0$  and note that it depends only on  $q_{\text{Pic}}, q_{\text{alg}}$ , and  $C_{\text{stab}}$ . Then, for all  $0 < \theta \leq 1$ , all  $0 < \lambda_{\text{alg}} < 1$  and  $0 < \lambda_{\text{Pic}}$  with  $0 < \lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} < \lambda_{\text{alg}}^*$  and  $0 < \lambda_{\text{Pic}}/\theta < \lambda_{\text{Pic}}^*$ , and all  $(\ell, \underline{k}, \underline{j}) \in \mathcal{Q}$  with  $\underline{k} < \infty$  and  $\underline{j} < \infty$ , it holds that*

$$(1 - \lambda_{\text{Pic}}/\lambda_{\text{Pic}}^*) \eta_\ell(u_\ell^{\underline{k}, \underline{j}}) \leq \eta_\ell(u_\ell^*) \leq (1 + \lambda_{\text{Pic}}/\lambda_{\text{Pic}}^*) \eta_\ell(u_\ell^{\underline{k}, \underline{j}}). \tag{51}$$

Moreover, there exist  $C_{\text{GHPS}} > 0$  and  $0 < q_{\text{GHPS}} < 1$  such that

$$\eta_{\ell+n}(u_{\ell+n}^{k,j}) \leq C_{\text{GHPS}} q_{\text{GHPS}}^n \eta_{\ell}(u_{\ell}^{k,j}) \text{ for all } (\ell + n + 1, 0, 0) \in \mathcal{Q}. \tag{52}$$

The constants  $C_{\text{GHPS}}$  and  $q_{\text{GHPS}}$  depend only on  $L, \alpha, C_{\text{rel}}, C_{\text{stab}}, q_{\text{red}}, q_{\text{alg}},$  and  $q_{\text{Pic}}$ , as well as on the adaptivity parameters  $\theta, \lambda_{\text{alg}},$  and  $\lambda_{\text{Pic}}$ .  $\square$

As a result of Lemma 7 and Proposition 2, we get the following lemma for the quasi-error of (32) on stopping indices  $\underline{k}(\ell), \underline{j}(\ell, k)$ . Please note that when  $\underline{\ell} < \infty$ , the summation below only goes to  $\underline{\ell} - 1$ , as the arguments rely on (52) which needs finite stopping indices  $\underline{k}(\ell)$  and  $\underline{j}(\ell, k)$  on each mesh  $\mathcal{T}_{\ell}$ .

**Lemma 8** *Suppose that  $0 < \lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} < \lambda_{\text{alg}}^*$  (from Lemma 6) as well as  $0 < \theta \leq 1$  and  $0 < \lambda_{\text{Pic}}/\theta < \lambda_{\text{Pic}}^*$  (from Lemma 7). With the convention  $\underline{\ell} - 1 = \infty$  if  $\underline{\ell} = \infty$ , there holds summability*

$$\sum_{\ell=\ell'+1}^{\underline{\ell}-1} \Delta_{\ell}^{k,j} \leq C \Delta_{\ell'}^{k,j} \text{ for all } (\ell', \underline{k}, \underline{j}) \in \mathcal{Q}, \tag{53}$$

where  $C > 0$  depends only on  $L, \alpha, C_{\text{rel}}, C_{\text{stab}}, q_{\text{red}}, \theta, q_{\text{alg}}, q_{\text{Pic}}, \lambda_{\text{alg}},$  and  $\lambda_{\text{Pic}}$ .

**Proof** Define  $\tilde{\Delta}_{\ell}^k := \|u^* - u_{\ell}^{k,j}\| + \eta_{\ell}(u_{\ell}^{k,j})$  as the sum of overall error plus error estimator. In comparison with (32),  $\tilde{\Delta}_{\ell}^k$  omits the algebraic error term but is only defined for the algebraic stopping indices  $\underline{j}(\ell, k)$ . With Proposition 2 and the linear convergence (52), we get that

$$\sum_{\ell=\ell'+1}^{\underline{\ell}-1} \tilde{\Delta}_{\ell}^k \stackrel{(31)}{\lesssim} \sum_{\ell=\ell'+1}^{\underline{\ell}-1} \eta_{\ell}(u_{\ell}^{k,j}) \stackrel{(52)}{\lesssim} \eta_{\ell'}(u_{\ell'}^{k,j}) \sum_{\ell=\ell'+1}^{\underline{\ell}-1} q_{\text{GHPS}}^{\ell-\ell'} \lesssim \tilde{\Delta}_{\ell'}^k.$$

Let  $(\ell', \underline{k}, \underline{j}) \in \mathcal{Q}$ . By definition (32), it holds that

$$\Delta_{\ell'}^{k,j} = \|u^* - u_{\ell'}^{k,j}\| + \|u_{\ell'}^{k,\star} - u_{\ell'}^{k,j}\| + \eta_{\ell'}(u_{\ell'}^{k,j}) = \tilde{\Delta}_{\ell'}^k + \|u_{\ell'}^{k,\star} - u_{\ell'}^{k,j}\|.$$

Moreover, note that

$$\begin{aligned} \|u_{\ell'}^{k,\star} - u_{\ell'}^{k,j}\| &\stackrel{(26)}{\lesssim} \|u_{\ell'}^{k,j} - u_{\ell'}^{k,j-1}\| \stackrel{(41)}{\lesssim} \eta_{\ell'}(u_{\ell'}^{k,j}) \\ &+ \|u_{\ell'}^{k,j} - u_{\ell'}^{k-1,j}\| \stackrel{(43)}{\lesssim} \eta_{\ell'}(u_{\ell'}^{k,j}) \leq \tilde{\Delta}_{\ell'}^k. \end{aligned}$$

This proves the equivalence  $\Delta_{\ell'}^{k,j} \simeq \tilde{\Delta}_{\ell'}^k$  for all  $(\ell', \underline{k}, \underline{j}) \in \mathcal{Q}$  and concludes the proof.  $\square$

### 4 Proof of Theorem 3 (linear convergence)

This section is dedicated to the proof of Theorem 3. The core is the following lemma that extends Lemma 8 to our setting with the triple indices.

**Lemma 9** *Suppose that  $0 < \lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} < \lambda_{\text{alg}}^*$  (from Lemma 6) as well as  $0 < \theta \leq 1$  and  $0 < \lambda_{\text{Pic}}/\theta < \lambda_{\text{Pic}}^*$  (from Lemma 7). Then, there exists  $C_{\text{sum}} > 0$  such that*

$$\sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} \Delta_\ell^{k,j} \leq C_{\text{sum}} \Delta_{\ell'}^{k',j'} \quad \text{for all } (\ell', k', j') \in \mathcal{Q}. \tag{54}$$

The constant  $C_{\text{sum}}$  depends only on  $C_{\text{rel}}, C_{\text{stab}}, q_{\text{red}}, \theta, q_{\text{alg}}, \lambda_{\text{alg}}, q_{\text{Pic}}, \lambda_{\text{Pic}}, \alpha$ , and  $L$ .

**Proof Step 1.** We prove that

$$\boxed{A_\ell^{k,j} := \|u_\ell^* - u_\ell^{k,j}\| + \|u_\ell^{k,*} - u_\ell^{k,j}\| + \eta_\ell(u_\ell^{k,j}) \simeq \Delta_\ell^{k,j} \quad \text{for all } (\ell, k, j) \in \mathcal{Q}.} \tag{55}$$

Note that  $A_\ell^{k,j}$  and  $\Delta_\ell^{k,j}$  only differ in the first term, where the overall error is replaced by the (inexact) linearization error. According to the Céa lemma (15), it holds that

$$\|u_\ell^* - u_\ell^{k,j}\| \leq \|u^* - u_\ell^{k,j}\| + \|u^* - u_\ell^*\| \stackrel{(15)}{\lesssim} \|u^* - u_\ell^{k,j}\| \leq \Delta_\ell^{k,j}.$$

This implies that  $A_\ell^{k,j} \lesssim \Delta_\ell^{k,j}$ . To see the converse inequality, note that

$$\begin{aligned} \|u^* - u_\ell^{k,j}\| &\leq \|u^* - u_\ell^*\| + \|u_\ell^* - u_\ell^{k,j}\| \stackrel{(A3)}{\lesssim} \eta_\ell(u_\ell^*) + \|u_\ell^* - u_\ell^{k,j}\| \\ &\stackrel{(A1)}{\lesssim} \eta_\ell(u_\ell^{k,j}) + \|u_\ell^* - u_\ell^{k,j}\| \leq A_\ell^{k,j}. \end{aligned}$$

This proves  $\Delta_\ell^{k,j} \lesssim A_\ell^{k,j}$  and concludes this step.

*Step 2* We prove some auxiliary estimates. First, we prove that the algebraic error  $\|u_\ell^{k,*} - u_\ell^{k,j-1}\|$  dominates the modified total error  $A_\ell^{k,j}$ , before the algebraic stopping criterion (22) is reached, i.e.,

$$\boxed{A_\ell^{k,j} \lesssim \|u_\ell^{k,*} - u_\ell^{k,j-1}\| \quad \text{for all } (\ell, k, j) \in \mathcal{Q} \text{ with } k \geq 1 \text{ and } 1 \leq j < \underline{j}(\ell, k).} \tag{56}$$

To this end, note that

$$\|u_\ell^* - u_\ell^{k,j}\| \leq \|u_\ell^* - u_\ell^{k,*}\| + \|u_\ell^{k,*} - u_\ell^{k,j}\| \stackrel{(45)}{\lesssim} \|u_\ell^{k,*} - u_\ell^{k-1,j}\| + \|u_\ell^{k,*} - u_\ell^{k,j}\|$$

$$\begin{aligned} &\leq 2 \|u_\ell^{k,\star} - u_\ell^{k,j}\| + \|u_\ell^{k,j} - u_\ell^{k-1,\underline{j}}\| \\ &\stackrel{(26)}{\lesssim} \|u_\ell^{k,j} - u_\ell^{k,j-1}\| + \|u_\ell^{k,j} - u_\ell^{k-1,\underline{j}}\|. \end{aligned}$$

Since  $1 \leq j < \underline{j}(\ell, k)$ , we conclude (56) from

$$\begin{aligned} A_\ell^{k,j} &= \|u_\ell^\star - u_\ell^{k,j}\| + \|u_\ell^{k,\star} - u_\ell^{k,j}\| + \eta_\ell(u_\ell^{k,j}) \\ &\lesssim \|u_\ell^{k,j} - u_\ell^{k,j-1}\| + \|u_\ell^{k,j} - u_\ell^{k-1,\underline{j}}\| + \eta_\ell(u_\ell^{k,j}) \\ &\stackrel{(42)}{\lesssim} \|u_\ell^{k,j} - u_\ell^{k,j-1}\| \stackrel{(26)}{\lesssim} \|u_\ell^{k,\star} - u_\ell^{k,j-1}\|. \end{aligned}$$

Second, we consider the use of nested iteration when passing to the next perturbed Banach–Picard step. We prove that

$$\boxed{\|u_\ell^{k,\star} - u_\ell^{k,0}\| \lesssim A_\ell^{k-1,\underline{j}} \text{ for all } (\ell, k, 0) \in \mathcal{Q} \text{ with } k \geq 1.} \tag{57}$$

To this end, simply note that

$$\|u_\ell^{k,\star} - u_\ell^{k,0}\| \stackrel{(40)}{=} \|u_\ell^{k,\star} - u_\ell^{k-1,\underline{j}}\| \stackrel{(45)}{\lesssim} \|u_\ell^\star - u_\ell^{k-1,\underline{j}}\| \leq A_\ell^{k-1,\underline{j}}.$$

Third, we prove that

$$\boxed{A_\ell^{k,\underline{j}} \lesssim A_\ell^{k,j} \text{ for all } (\ell, k, j) \in \mathcal{Q},} \tag{58}$$

related to the algebraic error contraction. Note that  $k = 0$  implies  $\underline{j} = 0$ , so that (58) trivially holds for  $k = 0$  with equality. Let now  $k \geq 1$ . We first consider the last but one algebraic iteration step  $j = \underline{j}(\ell, k) - 1 \geq 0$ . There holds that

$$\begin{aligned} A_\ell^{k,\underline{j}} &= \|u_\ell^\star - u_\ell^{k,\underline{j}}\| + \|u_\ell^{k,\star} - u_\ell^{k,\underline{j}}\| + \eta_\ell(u_\ell^{k,\underline{j}}) \\ &\leq \|u_\ell^\star - u_\ell^{k,\underline{j}-1}\| + \|u_\ell^{k,\star} - u_\ell^{k,\underline{j}-1}\| + \eta_\ell(u_\ell^{k,\underline{j}}) + 2 \|u_\ell^{k,\underline{j}} - u_\ell^{k,\underline{j}-1}\| \\ &\stackrel{(A1)}{\lesssim} A_\ell^{k,\underline{j}-1} + \|u_\ell^{k,\underline{j}} - u_\ell^{k,\underline{j}-1}\| \stackrel{(26)}{\lesssim} A_\ell^{k,\underline{j}-1} + \|u_\ell^{k,\star} - u_\ell^{k,\underline{j}-1}\| \simeq A_\ell^{k,\underline{j}-1}. \end{aligned}$$

This proves (58) for  $j = \underline{j}(\ell, k) - 1 \geq 0$ . Note that this argument also applies when  $\underline{j} = 1$ . If  $0 \leq j \leq \underline{j}(\ell, k) - 2$ , we employ the last estimate and (56) to conclude (58) from

$$A_\ell^{k,\underline{j}} \lesssim A_\ell^{k,\underline{j}-1} \stackrel{(56)}{\lesssim} \|u_\ell^{k,\star} - u_\ell^{k,\underline{j}-2}\| \stackrel{(25)}{\leq} \|u_\ell^{k,\star} - u_\ell^{k,j}\| \leq A_\ell^{k,j}.$$

Fourth, we prove that the linearization error  $\|u_\ell^* - u_\ell^{k-1,j}\|$  dominates the modified total error  $A_\ell^{k,j}$ , before the linearization stopping criterion (23) is reached, i.e.,

$$A_\ell^{k,j} \lesssim \|u_\ell^* - u_\ell^{k-1,j}\| \quad \text{for all } (\ell, k, \underline{j}) \in \mathcal{Q} \text{ with } 1 \leq k < \underline{k}(\ell). \tag{59}$$

To see this, note that  $1 \leq k < \underline{k}(\ell)$  yields that

$$\begin{aligned} A_\ell^{k,j} &= \|u_\ell^* - u_\ell^{k,j}\| + \|u_\ell^{k,*} - u_\ell^{k,j}\| + \eta_\ell(u_\ell^{k,j}) \\ &\stackrel{(26)}{\lesssim} \|u_\ell^* - u_\ell^{k,j}\| + \|u_\ell^{k,j} - u_\ell^{k,j-1}\| + \eta_\ell(u_\ell^{k,j}) \\ &\stackrel{(41)}{\lesssim} \|u_\ell^* - u_\ell^{k,j}\| + \|u_\ell^{k,j} - u_\ell^{k-1,j}\| + \eta_\ell(u_\ell^{k,j}) \stackrel{(48)}{\lesssim} \|u_\ell^{k,j} - u_\ell^{k-1,j}\| + \eta_\ell(u_\ell^{k,j}) \\ &\stackrel{(44)}{\lesssim} \|u_\ell^{k,j} - u_\ell^{k-1,j}\| \stackrel{(48)}{\lesssim} \|u_\ell^* - u_\ell^{k-1,j}\|, \end{aligned}$$

where we employ Lemma 6 and hence require  $0 < \lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}}$  to be sufficiently small.

Fifth, we consider the use of nested iteration when refining the mesh. We prove that

$$A_\ell^{0,j} \lesssim \eta_{\ell-1}(u_{\ell-1}^{k,j}) \leq A_{\ell-1}^{k,j} \quad \text{for all } (\ell, k, \underline{j}) \in \mathcal{Q}. \tag{60}$$

To this end, note that

$$\|u_\ell^* - u_{\ell-1}^{k,j}\| \leq \|u^* - u_\ell^*\| + \|u^* - u_{\ell-1}^{k,j}\| \stackrel{(15)}{\lesssim} \|u^* - u_{\ell-1}^{k,j}\| \stackrel{(31)}{\lesssim} \eta_{\ell-1}(u_{\ell-1}^{k,j}). \tag{61}$$

Next, recall from (39) that  $u_\ell^{0,*} = u_\ell^{0,j} = u_{\ell-1}^{k,j}$ . From (A1) used on non-refined mesh elements and (A2) used on refined mesh elements, we hence conclude that

$$A_\ell^{0,j} \stackrel{(39)}{=} \|u_\ell^* - u_{\ell-1}^{k,j}\| + \eta_\ell(u_{\ell-1}^{k,j}) \stackrel{(61)}{\lesssim} \eta_{\ell-1}(u_{\ell-1}^{k,j}) + \eta_\ell(u_{\ell-1}^{k,j}) \leq 2 \eta_{\ell-1}(u_{\ell-1}^{k,j}).$$

Sixth, we prove that

$$A_\ell^{k,j} \lesssim A_\ell^{k,j} \quad \text{for all } (\ell, k, \underline{j}) \in \mathcal{Q}. \tag{62}$$

We first consider  $k = \underline{k}(\ell) - 1 \geq 0$ . Note that

$$\|u_\ell^{k,*} - u_\ell^{k-1,j}\| \leq \|u_\ell^* - u_\ell^{k,*}\| + \|u_\ell^* - u_\ell^{k-1,j}\| \stackrel{(27)}{\lesssim} \|u_\ell^* - u_\ell^{k-1,j}\| \leq A_\ell^{k-1,j}. \tag{63}$$

Hence, the triangle inequality leads to

$$\begin{aligned}
 A_\ell^{k,j} &= \|u_\ell^\star - u_\ell^{k,j}\| + \|u_\ell^{k,\star} - u_\ell^{k,j}\| + \eta_\ell(u_\ell^{k,j}) \\
 &\stackrel{(A1)}{\lesssim} \|u_\ell^\star - u_\ell^{k-1,j}\| + \|u_\ell^{k,\star} - u_\ell^{k-1,j}\| + \|u_\ell^{k,j} - u_\ell^{k-1,j}\| + \eta_\ell(u_\ell^{k-1,j}) \\
 &\lesssim A_\ell^{k-1,j} + \|u_\ell^{k,j} - u_\ell^{k-1,j}\| \stackrel{(48)}{\lesssim} A_\ell^{k-1,j} + \|u_\ell^\star - u_\ell^{k-1,j}\| \leq 2A_\ell^{k-1,j}.
 \end{aligned}$$

This proves (62) for  $k = \underline{k}(\ell) - 1$ . Note that the same argument also applies when  $\underline{k} = 1$ . If  $0 \leq k \leq \underline{k}(\ell) - 2$ , then

$$A_\ell^{k,j} \lesssim A_\ell^{k-1,j} \stackrel{(59)}{\lesssim} \|u_\ell^\star - u_\ell^{k-2,j}\| \stackrel{(47)}{\leq} \|u_\ell^\star - u_\ell^{k,j}\| \leq A_\ell^{k,j},$$

also using that  $q'_{Pic} \leq 1$ . This concludes the proof of (62).

Seventh, we consider the use of nested iteration when passing to the next perturbed Banach–Picard step. We prove that

$$\boxed{A_\ell^{k,0} \lesssim A_\ell^{k-1,j} \text{ for all } (\ell, k, 0) \in \mathcal{Q} \text{ with } k \geq 1.} \tag{64}$$

Using (57) and recalling the definition  $u_\ell^{k,0} = u_\ell^{k-1,j}$ , it holds that

$$A_\ell^{k,0} = \|u_\ell^\star - u_\ell^{k-1,j}\| + \|u_\ell^{k,\star} - u_\ell^{k,0}\| + \eta_\ell(u_\ell^{k-1,j}) \lesssim A_\ell^{k-1,j},$$

which is the claim (64).

*Step 3* This step collects auxiliary estimates following from the geometric series and the contraction properties of the linearization and the algebraic solver. First, with the convention  $\underline{j}(\ell, k) - 1 = \infty$  when  $\underline{j}(\ell, k) = \infty$ , it holds that

$$\boxed{\sum_{j=i+1}^{\underline{j}(\ell,k)-1} A_\ell^{k,j} \lesssim \|u_\ell^{k,\star} - u_\ell^{k,i}\| \leq A_\ell^{k,i} \text{ for all } (\ell, k, i) \in \mathcal{Q} \text{ with } k \geq 1.} \tag{65}$$

This follows immediately from

$$\sum_{j=i+1}^{\underline{j}(\ell,k)-1} A_\ell^{k,j} \stackrel{(56)}{\lesssim} \sum_{j=i+1}^{\underline{j}(\ell,k)-1} \|u_\ell^{k,\star} - u_\ell^{k,j-1}\| \stackrel{(25)}{\leq} \|u_\ell^{k,\star} - u_\ell^{k,i}\| \sum_{j=i}^{\infty} q_{alg}^{j-i} \lesssim \|u_\ell^{k,\star} - u_\ell^{k,i}\|.$$



Analogously, with the convention that  $\underline{k}(\ell) - 1 = \infty$  when  $\underline{k}(\ell) = \infty$ , the contraction (47) of the perturbed Banach–Picard iteration leads to

$$\boxed{\sum_{k=i+1}^{\underline{k}(\ell)-1} A_\ell^{k,j} \lesssim \|u_\ell^* - u_\ell^{i,j}\| \leq A_\ell^{i,j} \text{ for all } (\ell, i, j) \in \mathcal{Q}.} \tag{66}$$

This follows immediately from

$$\sum_{k=i+1}^{\underline{k}(\ell)-1} A_\ell^{k,j} \stackrel{(59)}{\lesssim} \sum_{k=i+1}^{\underline{k}(\ell)-1} \|u_\ell^* - u_\ell^{k-1,j}\| \stackrel{(47)}{\lesssim} \|u_\ell^* - u_\ell^{i,j}\| \sum_{k=i}^{\infty} (q'_{\text{Pic}})^{k-i} \lesssim \|u_\ell^* - u_\ell^{i,j}\|.$$

With the analogous convention  $\underline{\ell} - 1 = \infty$  when  $\underline{\ell} = \infty$ , we finally prove that

$$\boxed{\sum_{\ell=i+1}^{\underline{\ell}-1} A_\ell^{k,j} \lesssim A_i^{k,j} \text{ for all } (i, k, j) \in \mathcal{Q}.} \tag{67}$$

This follows from Step 1 and

$$\sum_{\ell=i+1}^{\underline{\ell}-1} A_\ell^{k,j} \stackrel{(55)}{\simeq} \sum_{\ell=i+1}^{\underline{\ell}-1} \Delta_\ell^{k,j} \stackrel{(53)}{\lesssim} \Delta_i^{k,j} \stackrel{(55)}{\simeq} A_i^{k,j}.$$

*Step 4* From now on, let  $(\ell', k', j') \in \mathcal{Q}$  be arbitrary. Suppose first that  $\underline{\ell} = \infty$ , i.e., both algebraic and linearization solvers terminate at some finite values  $\underline{k}(\ell)$  for all  $\ell \geq 0$  and  $\underline{j}(\ell, k)$  for all  $\ell \geq 0$  and all  $k \leq \underline{k}(\ell)$ , whereas infinitely many steps of mesh-refinement take place. By the definition of our index set  $\mathcal{Q}$  in (29) (which in particular features nested iterates), it holds that

$$\begin{aligned} \sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} A_\ell^{k,j} &= \sum_{\ell=\ell'+1}^{\infty} \left( A_\ell^{0,0} + \sum_{k=1}^{\underline{k}(\ell)} \left( A_\ell^{k,0} + \sum_{j=1}^{\underline{j}(\ell,k)} A_\ell^{k,j} \right) \right) \\ &\quad + \sum_{k=k'+1}^{\underline{k}(\ell')} \left( A_{\ell'}^{k,0} + \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} \right) + \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j} \\ &\lesssim \sum_{\ell=\ell'+1}^{\infty} \sum_{k=1}^{\underline{k}(\ell)} \sum_{j=1}^{\underline{j}(\ell,k)} A_\ell^{k,j} + \sum_{k=k'+1}^{\underline{k}(\ell')} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} + \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j}, \end{aligned} \tag{68}$$

where we have employed estimates (60) and (64) in order to start all the summations from  $k = 1$  and  $j = 1$ .

We consider the three summands in (68) separately. For the first sum, we infer that

$$\begin{aligned}
 \sum_{\ell=\ell'+1}^{\infty} \sum_{k=1}^{\underline{k}(\ell)} \sum_{j=1}^{\underline{j}(\ell,k)} A_{\ell}^{k,j} &\stackrel{(65)}{\lesssim} \sum_{\ell=\ell'+1}^{\infty} \sum_{k=1}^{\underline{k}(\ell)} (A_{\ell}^{k,j} + \|u_{\ell}^{k,\star} - u_{\ell}^{k,0}\|) \\
 &\stackrel{(57)}{\lesssim} \sum_{\ell=\ell'+1}^{\infty} \sum_{k=1}^{\underline{k}(\ell)} (A_{\ell}^{k,j} + A_{\ell}^{k-1,j}) \lesssim \sum_{\ell=\ell'+1}^{\infty} \left( A_{\ell}^{0,j} + \sum_{k=1}^{\underline{k}(\ell)} A_{\ell}^{k,j} \right) \\
 &\stackrel{(66)}{\lesssim} \sum_{\ell=\ell'+1}^{\infty} (A_{\ell}^{0,j} + A_{\ell}^{k,j}) \stackrel{(60)}{\lesssim} \sum_{\ell=\ell'+1}^{\infty} (A_{\ell-1}^{k,j} + A_{\ell}^{k,j}) \\
 &\lesssim A_{\ell'}^{k,j} + \sum_{\ell=\ell'+1}^{\infty} A_{\ell}^{k,j} \stackrel{(67)}{\lesssim} A_{\ell'}^{k,j} \stackrel{(62)}{\lesssim} A_{\ell'}^{k',j} \stackrel{(58)}{\lesssim} A_{\ell'}^{k',j'}. \tag{69}
 \end{aligned}$$

If  $k' = \underline{k}(\ell')$ , the second sum in the bound (68) disappears. If  $k' < \underline{k}(\ell')$ , we infer that

$$\begin{aligned}
 \sum_{k=k'+1}^{\underline{k}(\ell')} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} &\stackrel{(65)}{\lesssim} \sum_{k=k'+1}^{\underline{k}(\ell')} (A_{\ell'}^{k,j} + \|u_{\ell'}^{k,\star} - u_{\ell'}^{k,0}\|) \stackrel{(57)}{\lesssim} \sum_{k=k'+1}^{\underline{k}(\ell')} (A_{\ell'}^{k,j} + A_{\ell'}^{k-1,j}) \\
 &\lesssim A_{\ell'}^{k',j} + \sum_{k=k'+1}^{\underline{k}(\ell')} A_{\ell'}^{k,j} \stackrel{(66)}{\lesssim} A_{\ell'}^{k',j} + A_{\ell'}^{k,j} \stackrel{(62)}{\leq} A_{\ell'}^{k',j} \stackrel{(58)}{\lesssim} A_{\ell'}^{k',j'}. \tag{70}
 \end{aligned}$$

If  $j' = \underline{j}(\ell', k')$ , the third sum in the bound (68) disappears. If  $j' < \underline{j}(\ell', k')$ , we infer that

$$\sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j} \stackrel{(65)}{\leq} A_{\ell'}^{k',j} + A_{\ell'}^{k',j'} \stackrel{(58)}{\lesssim} A_{\ell'}^{k',j'}. \tag{71}$$

Summing up (68)–(71), we see that, provided that  $\underline{\ell} = \infty$ ,

$$\boxed{\sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} A_{\ell}^{k,j} \lesssim A_{\ell'}^{k',j'}}. \tag{72}$$

*Step 5* Suppose that  $\underline{\ell} < \infty$  and  $\underline{k}(\ell) = \infty$ , i.e., for the mesh  $\mathcal{T}_{\underline{\ell}}$ , the linearization loop does not terminate. Moreover, let  $\ell' < \underline{\ell}$ . Then, it holds as in (68) that

$$\sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} A_{\ell}^{k,j} \lesssim \sum_{k=1}^{\infty} \sum_{j=1}^{\underline{j}(\ell,k)} A_{\underline{\ell}}^{k,j} + \sum_{\ell=\ell'+1}^{\underline{\ell}-1} \sum_{k=1}^{\underline{k}(\ell)} \sum_{j=1}^{\underline{j}(\ell,k)} A_{\ell}^{k,j}$$

$$+ \sum_{k=k'+1}^{\underline{k}(\ell')} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} + \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j}. \tag{73}$$

We argue as before to see that

$$\begin{aligned} \sum_{\ell=\ell'+1}^{\underline{\ell}-1} \sum_{k=1}^{\underline{k}(\ell)} \sum_{j=1}^{\underline{j}(\ell,k)} A_{\ell}^{k,j} &\lesssim A_{\ell'}^{k',j'}, & \sum_{k=k'+1}^{\underline{k}(\ell')} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} &\lesssim A_{\ell'}^{k',j'}, \quad \text{and} \\ \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j} &\lesssim A_{\ell'}^{k',j'}. \end{aligned} \tag{74}$$

It only remains to estimate

$$\begin{aligned} \sum_{k=1}^{\infty} \sum_{j=1}^{\underline{j}(\underline{\ell},k)} A_{\underline{\ell}}^{k,j} &\stackrel{(65)}{\lesssim} \sum_{k=1}^{\infty} (A_{\underline{\ell}}^{k,j} + \|u_{\underline{\ell}}^{k,\star} - u_{\underline{\ell}}^{k,0}\|) \stackrel{(57)}{\lesssim} A_{\underline{\ell}}^{0,j} + \sum_{k=1}^{\infty} A_{\underline{\ell}}^{k,j} \stackrel{(66)}{\lesssim} A_{\underline{\ell}}^{0,j} \\ &\stackrel{(60)}{\lesssim} A_{\underline{\ell}-1}^{k,j} \leq A_{\ell'}^{k,j} + \sum_{\ell=\ell'+1}^{\underline{\ell}-1} A_{\ell}^{k,j} \stackrel{(67)}{\lesssim} A_{\ell'}^{k,j} \stackrel{(62)}{\lesssim} A_{\ell'}^{k',j} \stackrel{(58)}{\lesssim} A_{\ell'}^{k',j'}. \end{aligned} \tag{75}$$

Altogether, provided that  $\ell' < \underline{\ell} < \infty$  and  $\underline{k}(\underline{\ell}) = \infty$ , we again obtain (72).

*Step 6* Suppose that  $\underline{\ell} < \infty$  and  $\underline{k}(\underline{\ell}) = \infty$ , i.e., for the mesh  $\mathcal{T}_{\underline{\ell}}$ , the linearization loop does not terminate, and moreover,  $\ell' = \underline{\ell}$ . Arguing as in (75) and (71), it holds that

$$\sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} A_{\ell}^{k,j} \lesssim \sum_{k=k'+1}^{\infty} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} + \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j} \lesssim A_{\ell'}^{k',j'}. \tag{76}$$

*Step 7* Suppose that  $\underline{\ell} < \infty$ , where  $\underline{k}(\underline{\ell}) < \infty$  and hence  $\underline{j}(\underline{\ell},k) = \infty$ , i.e., the linear solver does not terminate for the linearization step  $\underline{k}(\underline{\ell})$ . Suppose moreover  $\ell' < \underline{\ell}$ . Then, it holds that

$$\begin{aligned} \sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} A_{\ell}^{k,j} &\lesssim \sum_{j=1}^{\infty} A_{\underline{\ell}}^{k,j} + \sum_{k=1}^{\underline{k}(\underline{\ell})-1} \sum_{j=1}^{\underline{j}(\underline{\ell},k)} A_{\underline{\ell}}^{k,j} + \sum_{\ell=\ell'+1}^{\underline{\ell}-1} \sum_{k=1}^{\underline{k}(\ell)} \sum_{j=1}^{\underline{j}(\ell,k)} A_{\ell}^{k,j} \\ &+ \sum_{k=k'+1}^{\underline{k}(\ell')} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} + \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j}. \end{aligned} \tag{77}$$

We argue as before to see that

$$\sum_{\ell=\ell'+1}^{\underline{\ell}-1} \sum_{k=1}^{\underline{k}(\ell)} \sum_{j=1}^{\underline{j}(\ell,k)} A_{\ell}^{k,j} \stackrel{(69)}{\lesssim} A_{\ell'}^{k',j'}, \quad \sum_{k=k'+1}^{\underline{k}(\ell')} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} \stackrel{(70)}{\lesssim} A_{\ell'}^{k',j'}, \quad \text{and}$$

$$\sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j} \stackrel{(71)}{\lesssim} A_{\ell'}^{k',j'}.$$

For the first sum in (77), we get that

$$\sum_{j=1}^{\infty} A_{\underline{\ell}}^{k,j} \stackrel{(65)}{\lesssim} \|u_{\underline{\ell}}^{k,\star} - u_{\underline{\ell}}^{k,0}\| \stackrel{(57)}{\lesssim} A_{\underline{\ell}}^{k-1,j} \stackrel{(69)}{\lesssim} A_{\ell'}^{k',j'}. \tag{78}$$

Finally, the second sum in (77) can be bounded analogously to (75) in Step 5 by  $A_{\ell'}^{k',j'}$ . Altogether, we obtain (72) provided that  $\ell' < \underline{\ell} < \infty$ ,  $\underline{k}(\ell) < \infty$ , and  $\underline{j}(\underline{\ell}, \underline{k}) = \infty$ .

**Step 8.** Suppose that  $\underline{\ell} < \infty$ , where  $\underline{k}(\ell) < \infty$  and hence  $\underline{j}(\underline{\ell}, \underline{k}) = \infty$ , i.e., the linear solver does not terminate for the linearization step  $\underline{k}(\ell)$ . Suppose moreover  $\ell' = \underline{\ell}$  but  $k' < \underline{k}(\ell')$ . Then, it holds that

$$\sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} A_{\ell}^{k,j} \lesssim \sum_{j=1}^{\infty} A_{\ell'}^{k,j} + \sum_{k=k'+1}^{\underline{k}(\ell')-1} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} + \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j}. \tag{79}$$

We argue as before to see that

$$\sum_{j=1}^{\infty} A_{\ell'}^{k,j} \stackrel{(78)}{\lesssim} A_{\ell'}^{k',j'}, \quad \sum_{k=k'+1}^{\underline{k}(\ell')-1} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} \stackrel{(70)}{\lesssim} A_{\ell'}^{k',j'}, \quad \text{and} \quad \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j} \stackrel{(71)}{\lesssim} A_{\ell'}^{k',j'}.$$

Hence, we obtain (72) provided that  $\ell' = \underline{\ell} < \infty$ ,  $k' < \underline{k}(\ell') < \infty$ , and  $\underline{j}(\ell', \underline{k}) = \infty$ .

*Step 9* Suppose that  $\underline{\ell} < \infty$ , where  $\underline{k}(\ell) < \infty$  and hence  $\underline{j}(\underline{\ell}, \underline{k}) = \infty$ , i.e., the linear solver does not terminate for the linearization step  $\underline{k}(\ell)$ . Suppose  $\ell' = \underline{\ell}$  and  $k' = \underline{k}(\ell')$ . Then, the sum in (72) reduces to  $\sum_{j=j'+1}^{\infty} A_{\ell'}^{k',j}$  and (65) yields the inequality (72).

*Step 10* Suppose that  $\underline{\ell}, \underline{k}(\ell), \underline{j}(\underline{\ell}, \underline{k}(\ell)) < \infty$  and that Algorithm 1 finished on step (iii) when  $\eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,j}) = 0$  for tolerance  $\tau = 0$ . From (31), we see that  $\eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,j}) = 0$  implies  $u^{\star} = u_{\underline{\ell}}^{k,j}$ , i.e., the exact solution was found. Moreover, through the stopping criteria (23) and (22), we see that  $u_{\underline{\ell}}^{k-1,j} = u_{\underline{\ell}}^{k,j-1} = u_{\underline{\ell}}^{k,j}$ , so that (48) gives  $u_{\underline{\ell}}^{\star} = u_{\underline{\ell}}^{k,j}$ , and finally (26) gives  $u_{\underline{\ell}}^{k,\star} = u_{\underline{\ell}}^{k,j}$ . Thus  $A_{\underline{\ell}}^{k,j} = 0$ .

Let  $\ell' < \underline{\ell}$ . Then, as in (73),

$$\begin{aligned} \sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} A_{\ell}^{k,j} &\lesssim \sum_{k=1}^{\underline{k}(\underline{\ell})} \sum_{j=1}^{\underline{j}(\underline{\ell},k)} A_{\underline{\ell}}^{k,j} + \sum_{\ell=\ell'+1}^{\underline{\ell}-1} \sum_{k=1}^{\underline{k}(\underline{\ell})} \sum_{j=1}^{\underline{j}(\underline{\ell},k)} A_{\underline{\ell}}^{k,j} \\ &+ \sum_{k=k'+1}^{\underline{k}(\ell')} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} + \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j}. \end{aligned}$$

Here, the last three terms are estimated as in (74), whereas for the first one, we can proceed as in (75), crucially noting that the last summand  $A_{\underline{\ell}}^{k,j}$  is zero.

If  $\ell' = \underline{\ell}$ , three cases are possible. The first case is  $k' < \underline{k}$ . Then

$$\sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} A_{\ell}^{k,j} \lesssim \sum_{k=k'+1}^{\underline{k}(\ell')} \sum_{j=1}^{\underline{j}(\ell',k)} A_{\ell'}^{k,j} + \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j},$$

which is controlled as in (74). The second case is  $k' = \underline{k}$  but  $j' < \underline{j}$ , where directly

$$\sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} A_{\ell}^{k,j} = \sum_{j=j'+1}^{\underline{j}(\ell',k')} A_{\ell'}^{k',j} \stackrel{(65)}{\lesssim} A_{\ell'}^{k',j'},$$

since  $A_{\ell'}^{k',j} = 0$ . In the third case,  $k' = \underline{k}$  and  $j' = \underline{j}$ , the sum is void, and (72) follows.

*Step 11* Finally, if  $\underline{\ell}, \underline{k}(\underline{\ell}), \underline{j}(\underline{\ell}, \underline{k}(\underline{\ell})) < \infty$  and Algorithm 1 finished on step (IV) for tolerance  $\tau > 0$ , (72) follows immediately simplifying Step 4.

*Step 12* Combining Steps 4–11 that cover all possible runs of Algorithm 1 with Step 1, we finally see that

$$\sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} \Delta_{\ell}^{k,j} \stackrel{(55)}{\simeq} \sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) > (\ell',k',j')}} A_{\ell}^{k,j} \lesssim A_{\ell'}^{k',j'} \stackrel{(55)}{\simeq} \Delta_{\ell'}^{k',j'} \text{ for all } (\ell', k', j') \in \mathcal{Q}.$$

This concludes the proof of (54). □

**Proof of Theorem 3** The proof is split into two steps.

*Step 1* For the convenience of the reader, we recall an argument from the proof of [8, Lemma 4.9]: For  $M \in \mathbb{N} \cup \{\infty\}$ , let  $C > 0$  and  $\alpha_n \geq 0$  satisfy that

$$\sum_{n=N+1}^M \alpha_n \leq C \alpha_N \text{ for all } N \in \mathbb{N}_0 \text{ with } N < \min\{M, \infty\}.$$

Then,

$$(1 + C^{-1}) \sum_{n=N+1}^M \alpha_n \leq \sum_{n=N+1}^M \alpha_n + \alpha_N = \sum_{n=N}^M \alpha_n \quad \text{for all } N \in \mathbb{N}_0.$$

Inductively, it follows for all  $N, m \in \mathbb{N}_0$  with  $N + m < \min\{M + 1, \infty\}$  that

$$(1 + C^{-1})^m \sum_{n=N+m}^M \alpha_n \leq \sum_{n=N+1}^M \alpha_n + \alpha_N = \sum_{n=N}^M \alpha_n.$$

We thus conclude for all  $N, m \in \mathbb{N}_0$  with  $N + m < \min\{M + 1, \infty\}$  that

$$\alpha_{N+m} \leq \sum_{n=N+m}^M \alpha_n \leq (1 + C^{-1})^{-m} \sum_{n=N}^M \alpha_n \leq (1 + C) (1 + C^{-1})^{-m} \alpha_N.$$

*Step 2* Since the index set  $\mathcal{Q}$  is linearly ordered with respect to the total step counter  $|(\cdot, \cdot, \cdot)|$ , Lemma 9 and Step 1 imply that

$$\begin{aligned} \Delta_{\ell'}^{k',j'} &\leq C_{\text{lin}} q_{\text{lin}}^{|(\ell',k',j')| - |(\ell,k,j)|} \Delta_{\ell}^{k,j} \quad \text{for all } (\ell, k, j), (\ell', k', j') \\ &\in \mathcal{Q} \text{ with } (\ell', k', j') \geq (\ell, k, j), \end{aligned}$$

where  $C_{\text{lin}} = 1 + C_{\text{sum}}$  and  $q_{\text{lin}} = C_{\text{sum}} / (C_{\text{sum}} + 1)$ . This concludes the proof.  $\square$

### 5 Proof of Theorem 4 (optimal decay rate wrt. degrees of freedom)

The first result of this section proves the left inequality in (36):

**Lemma 10** *Suppose (R1) as well as (A1), (A2), and (A4). Let  $s > 0$  and assume  $\|u^*\|_{\mathbb{A}_s} > 0$ . For tolerance  $\tau = 0$ , it then holds that*

$$\|u^*\|_{\mathbb{A}_s} \leq c_{\text{opt}} \sup_{(\ell',k',j') \in \mathcal{Q}} (\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1)^s \Delta_{\ell'}^{k',j'}, \tag{80}$$

where the constant  $c_{\text{opt}} > 0$  depends only on  $C_{C\acute{e}a} = L/\alpha$ ,  $C_{\text{stab}}$ ,  $C_{\text{rel}}$ ,  $C_{\text{son}}$ ,  $\#\mathcal{T}_0$ ,  $s$ , and, if  $\underline{\ell} < \infty$ , additionally on  $\underline{\ell}$ .

**Proof** The proof is split into three steps. First, we recall from [5, Lemma 22] that

$$\#\mathcal{T}_h / \#\mathcal{T}_H \leq \#\mathcal{T}_h - \#\mathcal{T}_H + 1 \leq \#\mathcal{T}_h \quad \text{for all } \mathcal{T}_H \in \mathbb{T} \text{ and all } \mathcal{T}_h \in \text{refine}(\mathcal{T}_H). \tag{81}$$

*Step 1* We consider the three non-generic cases with  $\underline{\ell} < \infty$ . First, let  $\underline{k}(\underline{\ell}) < \infty$ , and  $\underline{j}(\underline{\ell}, k) < \infty$ . Then, Algorithm 1 was terminated in Step (iii) with  $\eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,j}) = 0$ . Due to the C ea lemma (15) and Proposition 2, it follows that

$$\|u^* - u_{\underline{\ell}}^*\| \stackrel{(15)}{\lesssim} \|u^* - u_{\underline{\ell}}^{k,j}\| \stackrel{(31)}{\lesssim} \eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,j}) = 0$$

and hence  $u^* = u_{\underline{\ell}}^* = u_{\underline{\ell}}^{k,\star} = u_{\underline{\ell}}^{k,j}$  and  $\eta_{\underline{\ell}}(u_{\underline{\ell}}^*) = 0$ .

Second, let  $\underline{k}(\underline{\ell}) < \infty$  but  $\underline{j}(\underline{\ell}, k) = \infty$ , i.e., the algebraic solver does not stop. According to Theorem 3, it holds that

$$\Delta_{\underline{\ell}}^{k,j} = \|u^* - u_{\underline{\ell}}^{k,j}\| + \|u_{\underline{\ell}}^{k,\star} - u_{\underline{\ell}}^{k,j}\| + \eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,j}) \rightarrow 0 \text{ as } j \rightarrow \infty.$$

Hence, we obtain that  $u^* = u_{\underline{\ell}}^* = u_{\underline{\ell}}^{k,\star}$ . From stability (A1), it follows that

$$0 \leq \eta_{\underline{\ell}}(u_{\underline{\ell}}^*) \stackrel{(A1)}{\lesssim} \eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,j}) + \|u_{\underline{\ell}}^* - u_{\underline{\ell}}^{k,j}\| \rightarrow 0 \text{ as } j \rightarrow \infty.$$

Hence, we see that  $\eta_{\underline{\ell}}(u_{\underline{\ell}}^*) = 0$ .

Finally, let  $\underline{k}(\underline{\ell}) = \infty$ , i.e., the linearization solver does not stop. Analogously to the previous case, we obtain that

$$\Delta_{\underline{\ell}}^{k,j} = \|u^* - u_{\underline{\ell}}^{k,j}\| + \|u_{\underline{\ell}}^{k,\star} - u_{\underline{\ell}}^{k,j}\| + \eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,j}) \rightarrow 0 \text{ as } k \rightarrow \infty.$$

Hence, we get that  $u^* = u_{\underline{\ell}}^*$ . Again, stability (A1) yields that  $\eta_{\underline{\ell}}(u_{\underline{\ell}}^*) = 0$ .

In any case,  $\underline{\ell} < \infty$  implies that  $\eta_{\underline{\ell}}(u_{\underline{\ell}}^*) = 0$  and hence

$$\|u^*\|_{\mathbb{A}_s} = \sup_{0 \leq N < \#\mathcal{T}_{\underline{\ell}} - \#\mathcal{T}_0} \left( (N + 1)^s \inf_{\mathcal{T}_{\text{opt}} \in \mathbb{T}(N)} \eta_{\text{opt}}(u_{\text{opt}}^*) \right).$$

The term  $N + 1$  within the supremum can be estimated by

$$N + 1 \leq \#\mathcal{T}_{\underline{\ell}} - \#\mathcal{T}_0 \stackrel{(R1)}{\leq} (C_{\text{son}}^{\underline{\ell}} - 1) \#\mathcal{T}_0.$$

Moreover, (A1), (A2), and (A4) yield quasi-monotonicity  $\eta_{\text{opt}}(u_{\text{opt}}^*) \lesssim \eta_0(u_0^*)$  (see, e.g., [8, Lemma 3.5]). Altogether, we thus arrive at

$$\|u^*\|_{\mathbb{A}_s} \lesssim \eta_0(u_0^*) \leq \sup_{\ell' \in \mathbb{N}_0} (\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1)^s \eta_{\ell'}(u_{\ell'}^*). \tag{82}$$

*Step 2* We consider the generic case that  $\underline{\ell} = \infty$  and  $\eta_{\underline{\ell}}(u_{\underline{\ell}}^{k,j}) > 0$  for all  $\ell \in \mathbb{N}_0$ . Algorithm 1 then guarantees that  $\#\mathcal{T}_{\underline{\ell}} \rightarrow \infty$  as  $\ell \rightarrow \infty$ . Thus, we can argue analogously to the proof of [8, Theorem 4.1]: Let  $N \in \mathbb{N}$ . Choose the maximal

$\ell' \in \mathbb{N}_0$  such that  $\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1 \leq N$ . Then,  $\mathcal{T}_{\ell'} \in \mathbb{T}(N)$ . The choice of  $N$  guarantees that

$$N + 1 \leq \#\mathcal{T}_{\ell'+1} - \#\mathcal{T}_0 + 1 \stackrel{(81)}{\leq} \#\mathcal{T}_{\ell'+1} \leq C_{\text{son}}\#\mathcal{T}_{\ell'} \stackrel{(81)}{\leq} C_{\text{son}}\#\mathcal{T}_0 (\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1). \tag{83}$$

This leads to

$$(N + 1)^s \inf_{\mathcal{T}_{\text{opt}} \in \mathbb{T}(N)} \eta_{\text{opt}}(u_{\text{opt}}^*) \lesssim (\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1)^s \eta_{\ell'}(u_{\ell'}^*),$$

and we immediately see that this also holds for  $N = 0$  with  $\ell' = 0$ . Taking the supremum over all  $N \in \mathbb{N}_0$ , we conclude that

$$\|u^*\|_{\mathbb{A}_s} \lesssim \sup_{\ell' \in \mathbb{N}_0} (\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1)^s \eta_{\ell'}(u_{\ell'}^*). \tag{84}$$

*Step 3* With stability (A1) and the Céa lemma (15), we see for all  $(\ell', 0, 0) \in \mathcal{Q}$  that

$$\eta_{\ell'}(u_{\ell'}^*) \stackrel{(A1)}{\lesssim} \|u^* - u_{\ell'}^*\| + \|u^* - u_{\ell'}^{0,0}\| + \eta_{\ell'}(u_{\ell'}^{0,0}) \stackrel{(15)}{\lesssim} \|u^* - u_{\ell'}^{0,0}\| + \eta_{\ell'}(u_{\ell'}^{0,0}) \leq \Delta_{\ell'}^{0,0}.$$

With (82) and (84), we thus obtain that

$$\|u^*\|_{\mathbb{A}_s} \lesssim \sup_{(\ell', 0, 0) \in \mathcal{Q}} (\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1)^s \eta_{\ell'}(u_{\ell'}^*) \leq \sup_{(\ell', k', j') \in \mathcal{Q}} (\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1)^s \Delta_{\ell'}^{k', j'}.$$

This concludes the proof. □

To prove the upper estimate in (36), we need the comparison lemma from [8, Lemma 4.14] for the error estimator of the exact discrete solution  $u_\ell^* \in \mathcal{X}_\ell$ .

**Lemma 11** *Suppose (R1–R2) as well as (A1), (A2), and (A4). Let  $0 < \theta' < \theta_{\text{opt}} := (1 + C_{\text{stab}}^2 C_{\text{rel}}^2)^{-1}$ . Then, there exist constants  $C_1, C_2 > 0$  such that for all  $s > 0$  with  $0 < \|u^*\|_{\mathbb{A}_s} < \infty$  and all  $\mathcal{T}_H \in \mathbb{T}$ , there exists  $\mathcal{R}_H \subseteq \mathcal{T}_H$  which satisfies*

$$\#\mathcal{R}_H \leq C_1 C_2^{-1/s} \|u^*\|_{\mathbb{A}_s}^{1/s} \eta_H(u_H^*)^{-1/s}, \tag{85}$$

as well as the Dörfler marking criterion

$$\theta' \eta_H(u_H^*) \leq \eta_H(\mathcal{R}_H, u_H^*). \tag{86}$$

The constants  $C_1, C_2$  depend only on  $C_{\text{stab}}$  and  $C_{\text{rel}}$ . □



**Proof of Theorem 4** The proof is split into four steps. Without loss of generality, we may assume that  $\|u^*\|_{\mathbb{A}_s} < \infty$ .

*Step 1* Due to the assumptions  $\lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} \leq \lambda_{\text{alg}}^*$  (from Lemma 6) and  $\lambda_{\text{Pic}}/\theta < \lambda_{\text{Pic}}^*$  (from Lemma 7), we get that  $\lambda_{\text{alg}} \leq \lambda_{\text{alg}}^* \lambda_{\text{Pic}} \leq \lambda_{\text{alg}}^* \lambda_{\text{Pic}}^* \theta$ . Hence, it holds that

$$\begin{aligned} \theta' &= \frac{\theta + C_{\text{stab}} \left( (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}} + [C_{\text{Pic}} + (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}}] \lambda_{\text{Pic}} \right)}{1 - \lambda_{\text{Pic}} / \lambda_{\text{Pic}}^*} \\ &\leq \frac{\theta + C_{\text{stab}} \left( (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}}^* \lambda_{\text{Pic}}^* \theta + [C_{\text{Pic}} + (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}}^* \lambda_{\text{Pic}}^* \theta] \lambda_{\text{Pic}}^* \theta \right)}{1 - \theta} \end{aligned}$$

which converges to 0 as  $\theta \rightarrow 0$ . As a consequence, (34) holds for sufficiently small  $\theta$ .

Clearly, the parameters  $\lambda_{\text{alg}}, \lambda_{\text{Pic}}, \theta > 0$  can be chosen such that all assumptions are fulfilled. First, choose  $\theta > 0$  such that  $0 < \theta < \min\{1, \theta^*\}$ . Then, choose  $\lambda_{\text{Pic}} > 0$  such that  $0 < \lambda_{\text{Pic}}/\theta < \lambda_{\text{Pic}}^*$ . Finally, choose  $0 < \lambda_{\text{alg}} < 1$  such that  $\lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} < \lambda_{\text{alg}}^*$ .

*Step 2* Recall that  $C_{\text{Pic}} = q_{\text{Pic}}/(1 - q_{\text{Pic}})$  and  $C_{\text{alg}} = q_{\text{alg}}/(1 - q_{\text{alg}})$ . Provided that  $(\ell + 1, 0, 0) \in \mathcal{Q}$ , it follows from the contraction properties (26) resp. (45), and the stopping criteria (41) resp. (43) that

$$\begin{aligned} \|u_\ell^* - u_\ell^{k,j}\| &\leq \|u_\ell^* - u_\ell^{k,*}\| + \|u_\ell^{k,*} - u_\ell^{k,j}\| \\ &\stackrel{(45)}{\leq} C_{\text{Pic}} \|u_\ell^{k,*} - u_\ell^{k-1,j}\| + \|u_\ell^{k,*} - u_\ell^{k,j}\| \\ &\leq (1 + C_{\text{Pic}}) \|u_\ell^{k,*} - u_\ell^{k,j}\| + C_{\text{Pic}} \|u_\ell^{k,j} - u_\ell^{k-1,j}\| \\ &\stackrel{(26)}{\leq} (1 + C_{\text{Pic}}) C_{\text{alg}} \|u_\ell^{k,j} - u_\ell^{k,j-1}\| + C_{\text{Pic}} \|u_\ell^{k,j} - u_\ell^{k-1,j}\| \\ &\stackrel{(41)}{\leq} (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}} \eta_\ell(u_\ell^{k,j}) + [C_{\text{Pic}} + (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}}] \|u_\ell^{k,j} - u_\ell^{k-1,j}\| \\ &\stackrel{(43)}{\leq} \left( (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}} + [C_{\text{Pic}} + (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}}] \lambda_{\text{Pic}} \right) \eta_\ell(u_\ell^{k,j}) \\ &\stackrel{(34)}{=} C_{\text{stab}}^{-1} \left( \theta' (1 - \lambda_{\text{Pic}}/\lambda_{\text{Pic}}^*) - \theta \right) \eta_\ell(u_\ell^{k,j}). \end{aligned}$$

*Step 3* Let  $\mathcal{R}_\ell \subseteq \mathcal{T}_\ell$  be the subset from Lemma 11 with  $\theta'$  from (34). From Step 2, we obtain that

$$\begin{aligned} \eta_\ell(\mathcal{R}_\ell, u_\ell^*) &\stackrel{(A1)}{\leq} \eta_\ell(\mathcal{R}_\ell, u_\ell^{k,j}) + C_{\text{stab}} \|u_\ell^* - u_\ell^{k,j}\| \\ &\leq \eta_\ell(\mathcal{R}_\ell, u_\ell^{k,j}) + \left( \theta' (1 - \lambda_{\text{Pic}}/\lambda_{\text{Pic}}^*) - \theta \right) \eta_\ell(u_\ell^{k,j}). \end{aligned} \tag{87}$$

With the equivalence (51), Lemma 11, and estimate (87), we see that

$$\theta' (1 - \lambda_{\text{Pic}}/\lambda_{\text{Pic}}^*) \eta_\ell(u_\ell^{k,j}) \stackrel{(51)}{\leq} \theta' \eta_\ell(u_\ell^*) \stackrel{(86)}{\leq} \eta_\ell(\mathcal{R}_\ell, u_\ell^*)$$

$$\stackrel{(87)}{\leq} \eta_\ell(\mathcal{R}_\ell, u_\ell^{k,j}) + (\theta'(1 - \lambda_{\text{Pic}}/\lambda_{\text{Pic}}^*) - \theta)\eta_\ell(u_\ell^{k,j}).$$

Thus, we are led to

$$\theta \eta_\ell(u_\ell^{k,j}) \leq \eta_\ell(\mathcal{R}_\ell, u_\ell^{k,j}).$$

Hence,  $\mathcal{R}_\ell$  satisfies the Dörfler marking criterion (24) used in Algorithm 1. By the (quasi-) minimality of  $\mathcal{M}_\ell$  in (24), we infer that

$$\#\mathcal{M}_\ell \lesssim \#\mathcal{R}_\ell \stackrel{(85)}{\lesssim} \|u^*\|_{\mathbb{A}_s}^{1/s} \eta_\ell(u_\ell^*)^{-1/s} \stackrel{(51)}{\simeq} \|u^*\|_{\mathbb{A}_s}^{1/s} \eta_\ell(u_\ell^{k,j})^{-1/s}.$$

Recall from (40) that  $u_{\ell+1}^{0,j} = u_\ell^{k,j}$ . Thus, (60) and the equivalence (55) lead to

$$\eta_\ell(u_\ell^{k,j})^{-1/s} \stackrel{(60)}{\lesssim} (A_{\ell+1}^{0,j})^{-1/s} \stackrel{(55)}{\simeq} (\Delta_{\ell+1}^{0,j})^{-1/s}.$$

Overall, we end up with

$$\#\mathcal{M}_\ell \lesssim \|u^*\|_{\mathbb{A}_s}^{1/s} (\Delta_{\ell+1}^{0,j})^{-1/s} \quad \text{for all } (\ell + 1, 0, 0) \in \mathcal{Q}. \tag{88}$$

The hidden constant depends only on  $C_{\text{stab}}, C_{\text{rel}}, C_{\text{mark}}, 1 - \lambda_{\text{Pic}}/\lambda_{\text{Pic}}^*, C_{\text{Céa}} = L/\alpha, C'_{\text{rel}}$ , and  $s$ .

*Step 4* With linear convergence (33) and the geometric series, we see that

$$\sum_{\substack{(\tilde{\ell}, \tilde{k}, \tilde{j}) \in \mathcal{Q} \\ (\tilde{\ell}, \tilde{k}, \tilde{j}) \leq (\ell, k, j)}} (\Delta_{\tilde{\ell}}^{\tilde{k}, \tilde{j}})^{-1/s} \stackrel{(33)}{\lesssim} (\Delta_\ell^{k,j})^{-1/s} \sum_{\substack{(\tilde{\ell}, \tilde{k}, \tilde{j}) \in \mathcal{Q} \\ (\tilde{\ell}, \tilde{k}, \tilde{j}) \leq (\ell, k, j)}} (q_{\text{lin}}^{1/s})^{|\ell, k, j| - |(\tilde{\ell}, \tilde{k}, \tilde{j})|} \lesssim (\Delta_\ell^{k,j})^{-1/s} \tag{89}$$

with hidden constants depending only on  $C_{\text{lin}}, q_{\text{lin}}$ , and  $s$ . For  $(\ell, k, j) \in \mathcal{Q}$  such that  $(\ell + 1, 0, 0) \in \mathcal{Q}$  and such that  $\mathcal{T}_\ell \neq \mathcal{T}_0$ , Step 3 and the closure estimate (R3) lead to

$$\begin{aligned} \#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1 &\simeq \#\mathcal{T}_\ell - \#\mathcal{T}_0 \stackrel{(R3)}{\lesssim} \sum_{\tilde{\ell}=0}^{\ell-1} \#\mathcal{M}_{\tilde{\ell}} \stackrel{(88)}{\lesssim} \|u^*\|_{\mathbb{A}_s}^{1/s} \sum_{\tilde{\ell}=0}^{\ell} (\Delta_{\tilde{\ell}}^{0,j})^{-1/s} \\ &\leq \|u^*\|_{\mathbb{A}_s}^{1/s} \sum_{\substack{(\tilde{\ell}, \tilde{k}, \tilde{j}) \in \mathcal{Q} \\ (\tilde{\ell}, \tilde{k}, \tilde{j}) \leq (\ell, k, j)}} (\Delta_{\tilde{\ell}}^{\tilde{k}, \tilde{j}})^{-1/s} \stackrel{(89)}{\lesssim} \|u^*\|_{\mathbb{A}_s}^{1/s} (\Delta_\ell^{k,j})^{-1/s}. \end{aligned}$$

Replacing  $\|u^*\|_{\mathbb{A}_s}$  with  $\max\{\|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0}\}$ , the overall estimate trivially holds for  $\mathcal{T}_\ell = \mathcal{T}_0$ . This proves that

$$(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^s \Delta_\ell^{k,j} \lesssim \begin{cases} \max\{\|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0}\}, & \text{if } (\ell + 1, 0, 0) \in \mathcal{Q} \text{ and } \ell \geq 0, \\ \|u^*\|_{\mathbb{A}_s}, & \text{if } (\ell + 1, 0, 0) \in \mathcal{Q} \text{ and } \ell \geq 1. \end{cases} \tag{90}$$

It remains to consider the cases where  $(\ell, k, j) \in \mathcal{Q}$  but  $(\ell + 1, 0, 0) \notin \mathcal{Q}$ , as well as the case  $\mathcal{T}_\ell = \mathcal{T}_0$ . In the first case, it holds that  $1 \leq \ell = \underline{\ell} < \infty$ , and one of the cases discussed in detail in Step 1 of Lemma 10 arises.

First, let  $2 \leq \ell = \underline{\ell} < \infty$ . Since  $\ell - 1 \geq 1$  and  $(\ell, 0, 0) \in \mathcal{Q}$ , (90) shows that

$$(\#\mathcal{T}_{\ell-1} - \#\mathcal{T}_0 + 1)^s \Delta_{\ell-1}^{k,j} \lesssim \|u^*\|_{\mathbb{A}_s}.$$

Moreover, Lemma 9 leads to  $\Delta_\ell^{k,j} \lesssim \Delta_{\ell-1}^{k,j}$ . Therefore, we obtain from (83) that

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1 \leq C_{\text{son}} \#\mathcal{T}_0 (\#\mathcal{T}_{\ell-1} - \#\mathcal{T}_0 + 1). \tag{91}$$

Altogether, (90) holds for this case as well.

Second, let  $\ell = \underline{\ell} = 1$ . Then, we can rely on the inequality

$$\begin{aligned} (\#\mathcal{T}_1 - \#\mathcal{T}_0 + 1)^s \Delta_1^{k,j} &\stackrel{(91)}{\leq} C_{\text{son}} (\#\mathcal{T}_0) \Delta_1^{k,j} \stackrel{(54)}{\lesssim} \Delta_0^{k,j} \\ &\stackrel{(32)}{=} \|u^* - u_0^{k,j}\| + \|u_0^{k,*} - u_0^{k,j}\| + \eta_0(u_0^{k,j}) \\ &\stackrel{(26)}{\lesssim} \|u^* - u_0^*\| + \|u_0^* - u_0^{k,j}\| + \|u_0^{k,j} - u_0^{k,j-1}\| + \eta_0(u_0^{k,j}) \\ &\stackrel{(43)}{\lesssim} \|u^* - u_0^*\| + \|u_0^* - u_0^{k,j}\| + \|u_0^{k,j} - u_0^{k-1,j}\| + \eta_0(u_0^{k,j}) \\ &\stackrel{(48)}{\lesssim} \|u^* - u_0^*\| + \|u_0^{k,j} - u_0^{k-1,j}\| + \eta_0(u_0^{k,j}) \\ &\stackrel{(43)}{\lesssim} \|u^* - u_0^*\| + \eta_0(u_0^{k,j}) \stackrel{(51)}{\lesssim} \|u^* - u_0^*\| + \eta_0(u_0^*) \stackrel{(A3)}{\lesssim} \eta_0(u_0^*) \leq \|u^*\|_{\mathbb{A}_s}. \end{aligned} \tag{92}$$

Thus, (90) holds for this case as well.

Finally, let  $\ell = \underline{\ell} = 0$ . Then, linear convergence (33) proves that

$$\Delta_0^{k,j} \stackrel{(33)}{\lesssim} \Delta_0^{0,0}. \tag{93}$$

Hence, (90) also holds for this case, and we conclude the proof of (36). □

### 6 Proof of Theorem 5 (optimal decay rate wrt. computational cost)

**Proof of Theorem 5** Note that  $\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1 = 1 \leq \#\mathcal{T}_0$  for  $\ell' = 0$  and  $\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1 \leq \#\mathcal{T}_{\ell'}$  for  $\ell' > 0$ , so that the left inequality in (37) immediately follows from the left inequality in (36). In order to prove the right inequality in (37), let  $(\ell', k', j') \in \mathcal{Q}$ . Employing the right inequality in (36) (cf. (90)), the geometric series proves that

$$\begin{aligned} \sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) \leq (\ell',k',j')}} \#\mathcal{T}_{\ell} &\stackrel{(81)}{\leq} \#\mathcal{T}_0 \sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) \leq (\ell',k',j')}} (\#\mathcal{T}_{\ell} - \#\mathcal{T}_0 + 1) \\ &\stackrel{(36)}{\leq} \#\mathcal{T}_0 C_{\text{opt}}^{1/s} \max\{\|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0}\}^{1/s} \sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) \leq (\ell',k',j')}} (\Delta_{\ell}^{k,j})^{-1/s} \\ &\stackrel{(33)}{\leq} \#\mathcal{T}_0 C_{\text{opt}}^{1/s} C_{\text{lin}}^{1/s} \frac{1}{1 - q_{\text{lin}}^{1/s}} \max\{\|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0}\}^{1/s} (\Delta_{\ell'}^{k',j'})^{-1/s}. \end{aligned}$$

Rearranging this estimate, we end up with

$$\sup_{(\ell',k',j') \in \mathcal{Q}} \left( \sum_{\substack{(\ell,k,j) \in \mathcal{Q} \\ (\ell,k,j) \leq (\ell',k',j')}} \#\mathcal{T}_{\ell} \right)^s \Delta_{\ell'}^{k',j'} \lesssim \max\{\|u^*\|_{\mathbb{A}_s}, \Delta_0^{0,0}\},$$

where the hidden constant depends only on  $C_{\text{stab}}, C_{\text{rel}}, C_{\text{mark}}, 1 - \lambda_{\text{Pic}}/\lambda_{\text{Pic}}^*, C_{\text{Céa}} = L/\alpha, C'_{\text{rel}}, C_{\text{mesh}}, C_{\text{lin}}, q_{\text{lin}}, \#\mathcal{T}_0$ , and  $s$ . This proves the right inequality in (37).  $\square$

### 7 Numerical experiments

In this section, we present numerical experiments to underpin our theoretical findings. We compare the performance of Algorithm 1 for

- different values of  $\lambda_{\text{alg}} \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$ ,
- different values of  $\lambda_{\text{Pic}} \in \{1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$ ,
- different values of  $\theta \in \{0.1, 0.3, 0.5, 0.7, 0.9, 1\}$ .

As model problems serve nonlinear boundary value problems which arise, e.g., from nonlinear material laws in magnetostatic computations, where the mesh-refinement is steered by newest vertex bisection.

As an algebraic solver for the linear problems arising from the Banach–Picard iteration, we use PCG with multilevel additive Schwarz preconditioner from [22, Sect. 7.4.1] which is an optimal preconditioner, i.e., the condition number of the preconditioned system is uniformly bounded; cf. also [24, Sect. 2.9].

### 7.1 Model problem

Let  $\Omega \subset \mathbb{R}^d, d \geq 2$ , be a bounded Lipschitz domain with polytopal boundary  $\Gamma = \partial\Omega$  split into relatively open and disjoint Dirichlet and Neumann boundaries  $\Gamma_D, \Gamma_N$  with  $|\Gamma_D| > 0$ , i.e.,  $\Gamma = \overline{\Gamma_D} \cup \overline{\Gamma_N}$ . While the numerical experiments in Sects. 7.4–7.5 only consider  $d = 2$ , we stress that the following model problem is covered by the abstract theory for any  $d \geq 2$ . For  $f \in L^2(\Omega)$  and  $g \in L^2(\Gamma)$ , find  $u^*$  such that:

$$\begin{aligned} -\operatorname{div}(\mu(x, |\nabla u^*(x)|^2) \nabla u^*(x)) &= f(x) && \text{in } \Omega, \\ u^*(x) &= 0 && \text{on } \Gamma_D, \\ \mu(x, |\nabla u^*(x)|^2) \partial_n u^*(x) &= g(x) && \text{on } \Gamma_N, \end{aligned} \tag{94}$$

where the scalar nonlinearity  $\mu : \Omega \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  satisfies the following properties (M1–M4), similarly considered in [23,26]:

(M1) There exist constants  $0 < \gamma_1 < \gamma_2 < \infty$  such that

$$\gamma_1 \leq \mu(x, t) \leq \gamma_2 \quad \text{for all } x \in \Omega \text{ and all } t \geq 0. \tag{95}$$

(M2) There holds  $\mu(x, \cdot) \in C^1(\mathbb{R}_{\geq 0}, \mathbb{R})$  for all  $x \in \Omega$ , and there exist constants  $0 < \tilde{\gamma}_1 < \tilde{\gamma}_2 < \infty$  such that

$$\tilde{\gamma}_1 \leq \mu(x, t) + 2t \frac{d}{dt} \mu(x, t) \leq \tilde{\gamma}_2 \quad \text{for all } x \in \Omega \text{ and all } t \geq 0. \tag{96}$$

(M3) Lipschitz continuity of  $\mu(x, t)$  in  $x$ , i.e., there exists a constant  $L_\mu > 0$  such that

$$|\mu(x, t) - \mu(y, t)| \leq L_\mu |x - y| \quad \text{for all } x, y \in \Omega \text{ and all } t \geq 0. \tag{97}$$

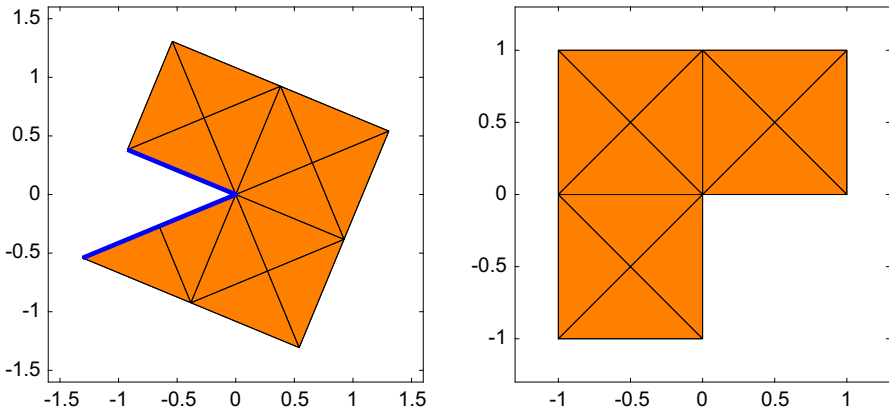
(M4) Lipschitz continuity of  $t \frac{d}{dt} \mu(x, t)$  in  $x$ , i.e., there exists a constant  $\tilde{L}_\mu > 0$  such that

$$\left| t \frac{d}{dt} \mu(x, t) - t \frac{d}{dt} \mu(y, t) \right| \leq \tilde{L}_\mu |x - y| \quad \text{for all } x, y \in \Omega \text{ and all } t \geq 0. \tag{98}$$

### 7.2 Weak formulation

The weak formulation of (94) reads as follows: Find  $u \in H_D^1(\Omega) := \{w \in H^1(\Omega) : w = 0 \text{ on } \Gamma_D\}$  such that

$$\int_{\Omega} \mu(x, |\nabla u^*(x)|^2) \nabla u^* \cdot \nabla v \, dx = \int_{\Omega} f v \, dx + \int_{\Gamma_N} g v \, ds \quad \text{for all } v \in H_D^1(\Omega). \tag{99}$$



**Fig. 1** Z-shaped domain  $\Omega \subset \mathbb{R}^2$  with initial mesh  $\mathcal{T}_0$  and  $\Gamma_D$  marked by a thick blue line (left) and L-shaped domain  $\Omega \subset \mathbb{R}^2$  with initial mesh  $\mathcal{T}_0$  (right)

With respect to the abstract framework of Sect. 2.1, we take  $\mathcal{X} = H_D^1(\Omega)$ ,  $\mathbb{K} = \mathbb{R}$ , and  $(\cdot, \cdot) = (\nabla \cdot, \nabla \cdot)$  with  $\|v\| = \|\nabla v\|_{L^2(\Omega)}$ . We obtain (11) with operators

$$\langle \mathcal{A}w, v \rangle_{\mathcal{X}' \times \mathcal{X}} = \int_{\Omega} \mu(x, |\nabla w(x)|^2) \nabla w(x) \cdot \nabla v(x) \, dx, \tag{100a}$$

$$\langle F, v \rangle_{\mathcal{X}' \times \mathcal{X}} = \int_{\Omega} f v \, dx + \int_{\Gamma_N} g v \, ds \tag{100b}$$

for all  $v, w \in \mathcal{X}$ . We recall from [23, Proposition 8.2] that (M1–M2) implies that  $\mathcal{A}$  is strongly monotone (with  $\alpha := \tilde{\gamma}_1$ ) and Lipschitz continuous (with  $L := \tilde{\gamma}_2$ ), so that (94) fits into the setting of Sect. 2.1. Moreover, (M3–M4) are required to prove the well-posedness and the properties (A1)–(A4) of the residual *a posteriori* error estimator.

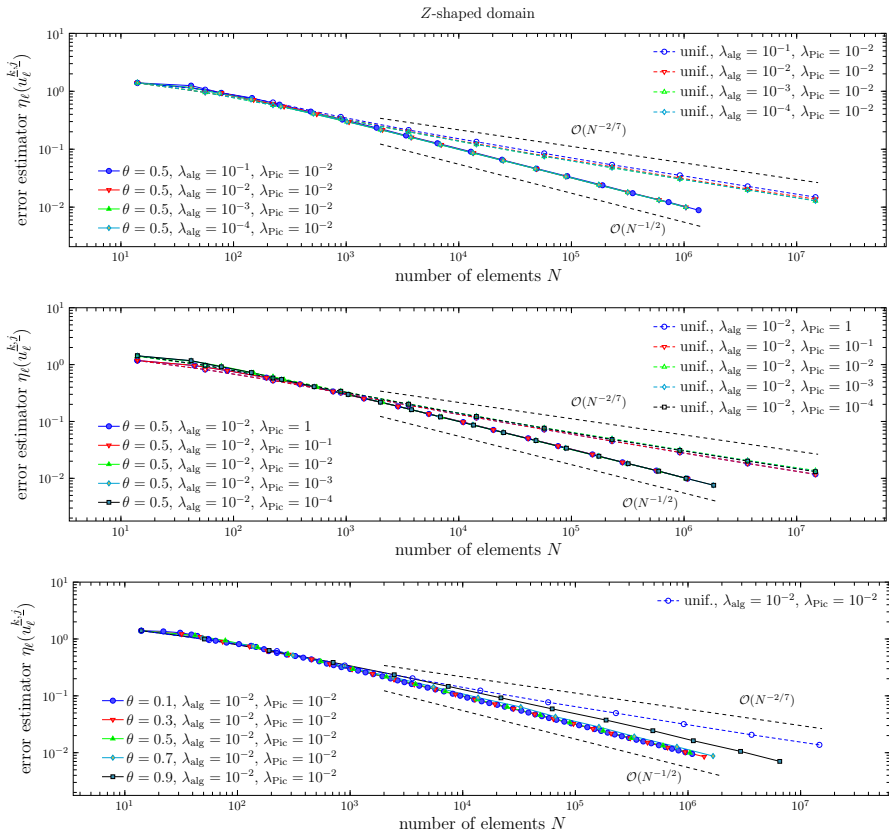
**7.3 Discretization and a posteriori error estimator**

Let  $\mathcal{T}_0$  be a conforming initial triangulation of  $\Omega$  into simplices  $T \in \mathcal{T}_0$ . For each  $\mathcal{T}_H \in \mathbb{T}$ , consider the lowest-order FEM space

$$\mathcal{X}_H := \{v \in C(\overline{\Omega}) : v|_{\Gamma} = 0 \text{ and } v|_T \in \mathcal{P}^1(T) \text{ for all } T \in \mathcal{T}_H\}. \tag{101}$$

As in [26, Sect. 3.2], we define for all  $T \in \mathcal{T}_H$  and all  $v_H \in \mathcal{X}_H$ , the corresponding weighted residual error indicators

$$\begin{aligned} \eta_H(T, v_H)^2 := & |T|^{2/d} \|f + \operatorname{div}(\mu(\cdot, |\nabla v_H|^2) \nabla v_H)\|_{L^2(T)}^2 \\ & + |T|^{1/d} \|[(\mu(\cdot, |\nabla v_H|^2) \nabla v_H) \cdot \mathbf{n}]\|_{L^2(\partial T \cap \Omega)}^2, \end{aligned} \tag{102}$$

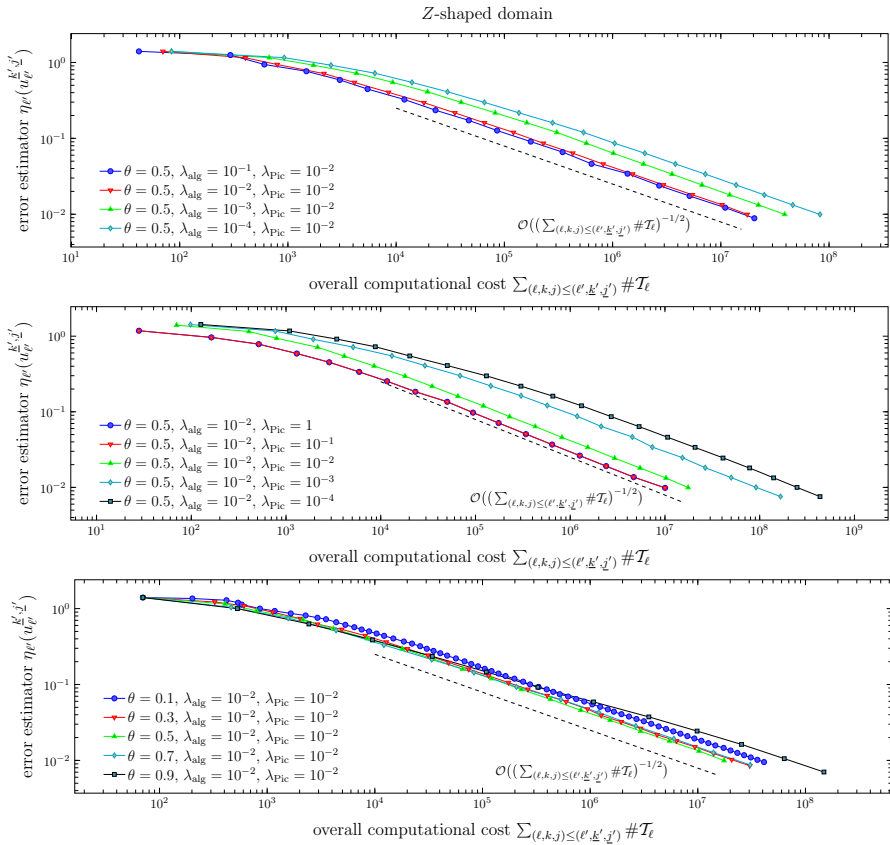


**Fig. 2** Optimal decay rate wrt. degrees of freedom, example from Sect. 7.4. Error estimator  $\eta_\ell(u_\ell^{k,j})$  on mesh  $\mathcal{T}_\ell$ , perturbed Banach–Picard iteration  $k$ , and PCG step  $j$  of Algorithm 1 with respect to the number of elements  $N$  of the mesh  $\mathcal{T}_\ell$  for various parameters  $\theta$ ,  $\lambda_{\text{pic}}$ , and  $\lambda_{\text{alg}}$

where  $[\cdot]$  denotes the usual jump of discrete functions across element interfaces, and  $\mathbf{n}$  is the outer normal vector of the considered element.

Due to (M3), the error estimator is well-posed, since the nonlinearity  $\mu(x, t)$  is Lipschitz continuous in  $x$ . Then, reliability (A3) and discrete reliability (A4) are proved as in the linear case; see, e.g., [10] for the linear case or [26, Theorem 3.3] and [26, Theorem 3.4], respectively, for strongly monotone nonlinearities.

The verification of stability (A1) and reduction (A2) requires the validity of a certain inverse estimate. For scalar nonlinearities and under the assumptions (M1–M4), the latter is proved in [26, Lemma 3.7]. Using this inverse estimate, the proof of (A1) and (A2) follows as for the linear case; see, e.g., [10] for the linear case or [26, Sect. 3.3] for scalar nonlinearities. We note that the necessary inverse estimate is, in particular, open for non-scalar nonlinearities. In any case, the arising constants in (A1–A4) depend also on the uniform shape regularity of the triangulations generated by newest vertex bisection.



**Fig. 3** Optimal decay rate wrt. overall computational cost, example from Sect. 7.4. Error estimator  $\eta_{\ell'}(u_{\ell'}^{k',j'})$  on mesh  $\mathcal{T}_{\ell'}$ , perturbed Banach–Picard iteration  $k'$ , and PCG step  $j'$  of Algorithm 1 with respect to the overall cost expressed as the cumulative sum  $\sum_{(\ell,k,j) \leq (\ell',k',j')} \#\mathcal{T}_{\ell}$  for various parameters  $\theta$ ,  $\lambda_{\text{pic}}$ , and  $\lambda_{\text{alg}}$ .

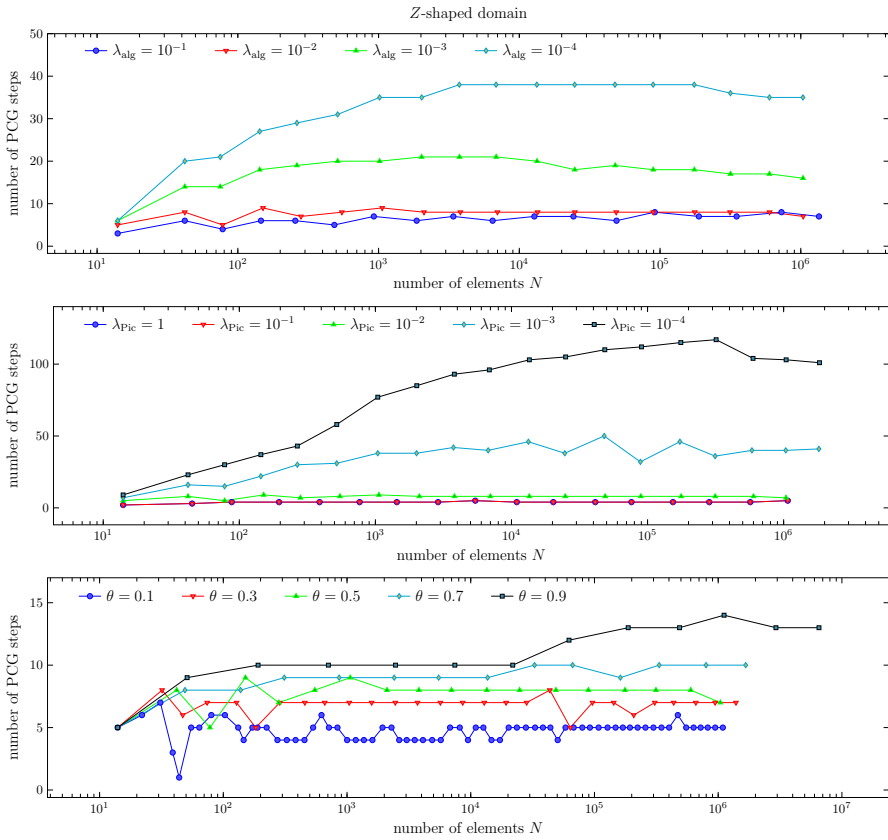
**7.4 Experiment with known solution**

We consider the Z-shaped domain  $\Omega \subset \mathbb{R}^2$  from Fig. 1 (left) with mixed boundary conditions and the nonlinear problem (94) with  $\mu(x, |\nabla u^*(x)|^2) := 2 + \frac{1}{\sqrt{1+|\nabla u^*(x)|^2}}$ . This leads to the bounds  $\alpha = 2$  and  $L = 3$  in (10). We prescribe the solution  $u^*$  in polar coordinates  $(x, y) = r(\cos \phi, \sin \phi)$  with  $\phi \in (-\pi, \pi)$

$$u^*(x, y) := r^\beta \cos(\beta \phi) \quad \text{with} \quad \beta = 4/7 \tag{103}$$

and compute  $f$  and  $g$  in (94) accordingly. We note that  $u^*$  has a generic singularity at the re-entrant corner  $(x, y) = (0, 0)$ .



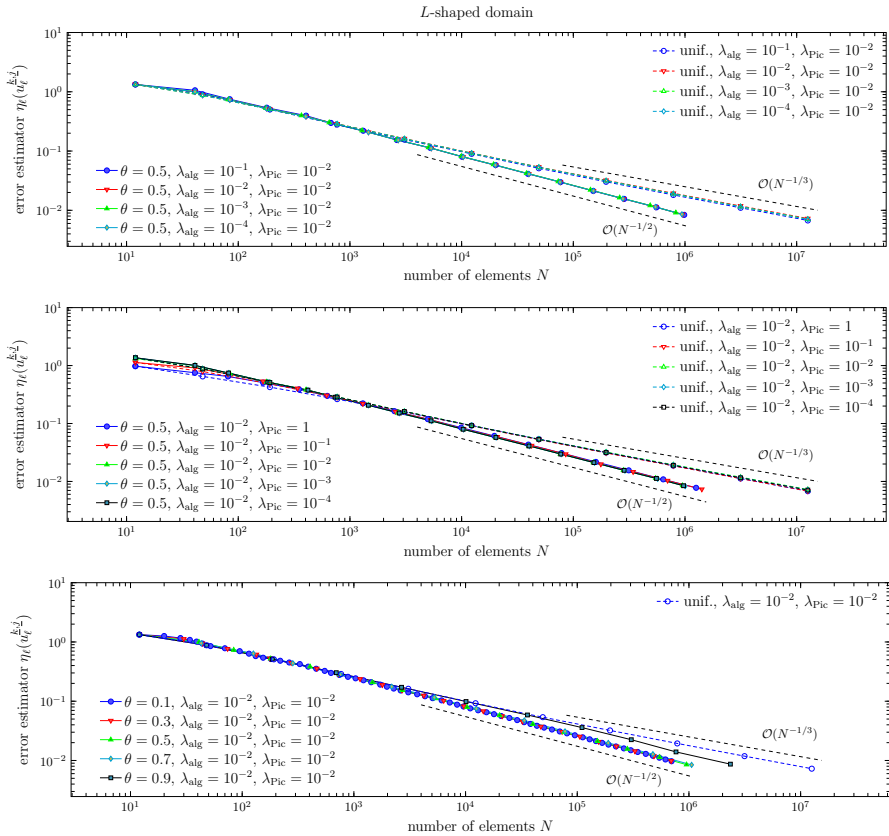


**Fig. 4** Number of PCG iterations wrt. the number of elements  $N := \#\mathcal{T}_\ell$ , example from Sect. 7.4, for  $\theta = 0.5$ ,  $\lambda_{\text{Pic}} = 10^{-2}$  (top), for  $\theta = 0.5$ ,  $\lambda_{\text{alg}} = 10^{-2}$  (middle), and for  $\lambda_{\text{alg}} = \lambda_{\text{Pic}} = 10^{-2}$  (bottom)

In Fig. 2, we compare uniform mesh-refinement ( $\theta = 1$ ) to adaptive mesh-refinement ( $0 < \theta < 1$ ) for different values of  $\lambda_{\text{alg}}$  and  $\lambda_{\text{Pic}}$ . We plot the error estimator  $\eta_\ell(u_\ell^{k,j})$  over the number of elements  $N := \#\mathcal{T}_\ell$ . First (top), we fix  $\theta = 0.5$ ,  $\lambda_{\text{Pic}} = 10^{-2}$ , and choose  $\lambda_{\text{alg}} \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$ . We see that uniform mesh-refinement leads to the suboptimal rate of convergence  $\mathcal{O}(N^{-2/7})$ , whereas Algorithm 1 with adaptive mesh-refinement regains the optimal rate of convergence  $\mathcal{O}(N^{-1/2})$ , independently of the actual choice of  $\lambda_{\text{alg}}$ . We observe the very same if we fix  $\theta = 0.5$ ,  $\lambda_{\text{alg}} = 10^{-2}$ , and choose  $\lambda_{\text{Pic}} \in \{1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$  (middle), or if we fix  $\lambda_{\text{alg}} = \lambda_{\text{Pic}} = 10^{-2}$  and vary  $\theta \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$  (bottom). Since we know from Proposition 2 and the estimate

$$\|u_\ell^{k,\star} - u_\ell^{k,j}\| \stackrel{(26)}{\lesssim} \|u_\ell^{k,j} - u_\ell^{k,j-1}\| \stackrel{(41)}{\lesssim} \eta_\ell(u_\ell^{k,j}) + \|u_\ell^{k,j} - u_\ell^{k-1,j}\| \stackrel{(43)}{\lesssim} \eta_\ell(u_\ell^{k,j})$$

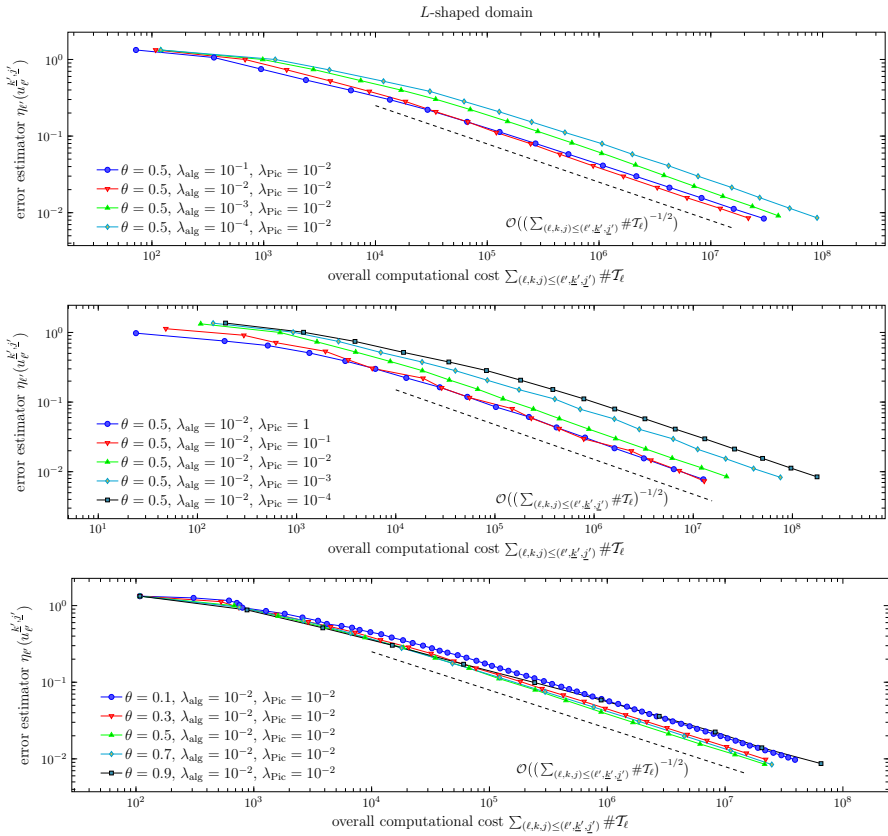
that  $\eta_\ell(u_\ell^{k,j}) \simeq \Delta_\ell^{k,j}$ , this empirically underpins Theorem 4.



**Fig. 5** Optimal decay rate wrt. degrees of freedom, example from Sect. 7.5. Error estimator  $\eta_\ell(u_\ell^{k,j})$  on mesh  $\mathcal{T}_\ell$ , perturbed Banach–Picard iteration  $k$ , and PCG step  $j$  of Algorithm 1 with respect to the number of elements  $N$  of the mesh  $\mathcal{T}_\ell$  for various parameters  $\theta$ ,  $\lambda_{\text{Pic}}$ , and  $\lambda_{\text{alg}}$

In Fig. 3, we analogously choose different combinations of  $\theta$ ,  $\lambda_{\text{alg}}$ , and  $\lambda_{\text{Pic}}$ . We plot the error estimator  $\eta_{\ell'}(u_{\ell'}^{k',j'})$  over the cumulative sum  $\sum_{(\ell,k,j) \leq (\ell',k',j')} \#\mathcal{T}_\ell$ . Independently of  $\theta$ ,  $\lambda_{\text{alg}}$ , and  $\lambda_{\text{Pic}}$ , we observe optimal order of convergence  $\mathcal{O}((\sum_{(\ell,k,j) \leq (\ell',k',j')} \#\mathcal{T}_\ell)^{-1/2})$  with respect to the overall computational cost, in accordance with Theorem 5.

In Fig. 4, we also consider the total number of PCG iterations cumulated over all Picard steps on the given mesh for different combinations of  $\theta$ ,  $\lambda_{\text{alg}}$ , and  $\lambda_{\text{Pic}}$ . We observe that independently of the choice of these parameters, the total number of PCG iterations stays uniformly bounded. Additionally, we see that for larger values of  $\lambda_{\text{alg}}$  and  $\lambda_{\text{Pic}}$ , as well as for smaller values of  $\theta$ , the total number of PCG iterations is smaller.

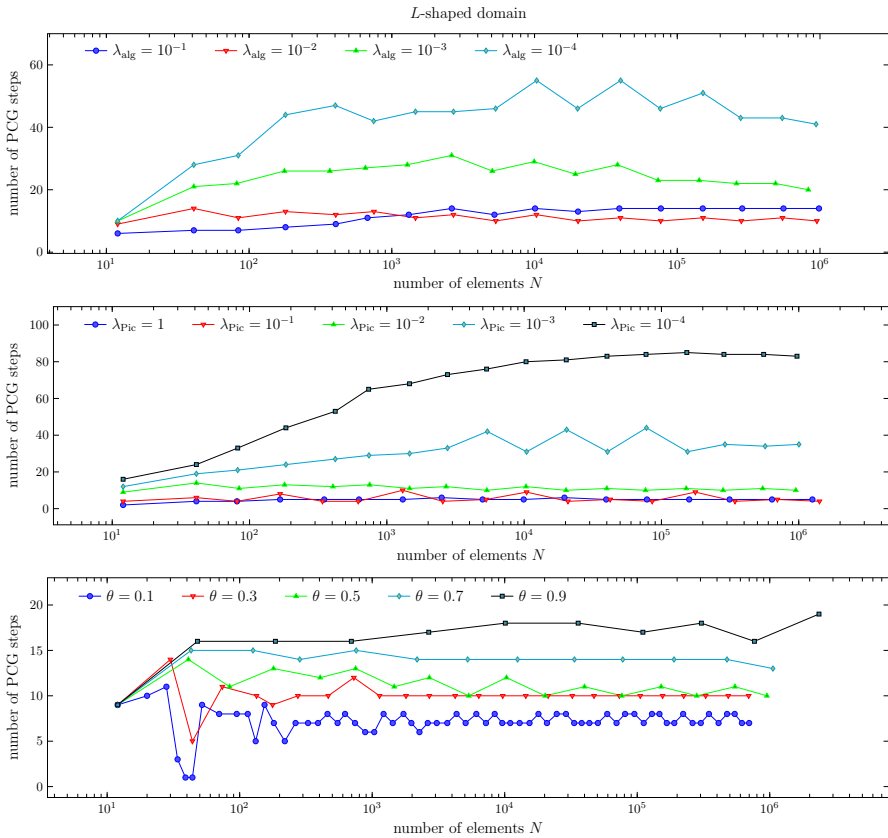


**Fig. 6** Optimal decay rate wrt. overall computational cost, example from Sect. 7.5. Error estimator  $\eta_{\ell}(u_{\ell}^{k',j'})$  on mesh  $\mathcal{T}_{\ell}$ , perturbed Banach–Picard iteration  $k'$ , and PCG step  $j'$  of Algorithm 1 with respect to the overall cost expressed as the cumulative sum  $\sum_{(\ell,k,j) \leq (\ell',k',j')} \#\mathcal{T}_{\ell}$  for various parameters  $\theta$ ,  $\lambda_{\text{Pic}}$ , and  $\lambda_{\text{alg}}$

### 7.5 Experiment with unknown solution

We consider the  $L$ -shaped domain  $\Omega \subset \mathbb{R}^2$  from Fig. 1 (right) and the nonlinear problem (94) with  $f(x) = 1$  and  $\mu(x, |\nabla u^*(x)|^2) := 1 + \frac{\ln(1+|\nabla u^*(x)|^2)}{1+|\nabla u^*(x)|^2}$ . Then, (M1–M4) hold with  $\alpha \approx 0.9582898$  and  $L \approx 1.5423438$ .

In Fig. 5, we again test Algorithm 1 with varying  $\theta$ ,  $\lambda_{\text{alg}}$ , and  $\lambda_{\text{Pic}}$ . We plot the error estimator  $\eta_{\ell}(u_{\ell}^{k,j})$  over the number of elements  $N := \#\mathcal{T}_{\ell}$ . Uniform mesh-refinement leads to the suboptimal rate of convergence  $\mathcal{O}(N^{-1/3})$ , whereas Algorithm 1 regains the optimal rate of convergence  $\mathcal{O}(N^{-1/2})$ . Again, this empirically confirms Theorem 4. The latter rate of convergence even appears to be robust with respect to  $\theta$ ,  $\lambda_{\text{alg}}$ , and  $\lambda_{\text{Pic}}$ .



**Fig. 7** Number of PCG iterations wrt. number of elements  $N := \#\mathcal{T}_\ell$ , example from Sect. 7.5, for  $\theta = 0.5$ ,  $\lambda_{\text{pic}} = 10^{-2}$  (top), for  $\theta = 0.5$ ,  $\lambda_{\text{alg}} = 10^{-2}$  (middle), and for  $\lambda_{\text{alg}} = \lambda_{\text{pic}} = 10^{-2}$  (bottom)

In Fig. 6, we plot the estimator  $\eta_{\ell'}(u_{\ell'}^{k',j'})$  over the cumulative sum  $\sum_{(\ell,k,j) \leq (\ell',k',j')} \#\mathcal{T}_\ell$ . Independently of the choice of the parameters  $\theta$ ,  $\lambda_{\text{alg}}$ , and  $\lambda_{\text{pic}}$ , we observe the optimal order of convergence  $\mathcal{O}((\sum_{(\ell,k,j) \leq (\ell',k',j')} \#\mathcal{T}_\ell)^{-1/2})$  with respect to the overall computational cost, which empirically underpins Theorem 5.

In Fig. 7, we finally consider the total number of PCG iterations cumulated over all Picard steps on the given mesh. We observe that independently of the choice of  $\theta$ ,  $\lambda_{\text{alg}}$ , and  $\lambda_{\text{pic}}$ , the total number of PCG iterations stays uniformly bounded. Additionally, we see that for larger values of  $\lambda_{\text{alg}}$  and  $\lambda_{\text{pic}}$ , as well as for smaller values of  $\theta$ , the total number of PCG iterations is smaller.

**Acknowledgements** The authors thankfully acknowledge support by the Austrian Science Fund (FWF) through the research projects *Optimal isogeometric boundary element method* (grant P29096), *Computational nonlinear PDEs* (grant P33216), and *Taming complexity in partial differential systems* (grant SFB F65). In addition, this project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No 647134 GATIPOR)

**Funding** Open Access funding provided by TU Wien (TUW).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Becker, Roland, Capatina, Daniela, Luce, Robert: Stopping criteria based on locally reconstructed fluxes. In: Abdulle, A., Deparis, S., Kressner, D., Nobile, F., Picasso, M. (eds.) *Numerical Mathematics and Advanced Applications—ENUMATH 2013*, pp. 243–251. Springer, Cham (2015)
2. Binev, Peter, Dahmen, Wolfgang, DeVore, Ron: Adaptive finite element methods with convergence rates. *Numer. Math.* **97**(2), 219–268 (2004)
3. Belenki, Liudmila, Diening, Lars, Kreuzer, Christian: Optimality of an adaptive finite element method for the  $p$ -Laplacian equation. *IMA J. Numer. Anal.* **32**(2), 484–510 (2012)
4. Bernardi, Christine, Dakroub, Jad, Mansour, Gihane, Sayah, Toni: A posteriori analysis of iterative algorithms for a nonlinear problem. *J. Sci. Comput.* **65**(2), 672–697 (2015)
5. Bespalov, Alex, Haberl, Alexander, Praetorius, Dirk: Adaptive FEM with coarse initial mesh guarantees optimal convergence rates for compactly perturbed elliptic problems. *Comput. Methods Appl. Mech. Engrg.* **317**, 318–340 (2017)
6. Bonito, Andrea, Nochetto, Ricardo H.: Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method. *SIAM J. Numer. Anal.* **48**(2), 734–771 (2010)
7. Cohen, Albert, Dahmen, Wolfgang, DeVore, Ronald: Adaptive wavelet schemes for nonlinear variational problems. *SIAM J. Numer. Anal.* **41**(5), 1785–1823 (2003)
8. Carstensen, Carsten, Feischl, Michael, Page, Marcus, Praetorius, Dirk: Axioms of adaptivity. *Comput. Math. Appl.* **67**(6), 1195–1253 (2014)
9. Carstensen, Carsten, Gedicke, Joscha: An adaptive finite element eigenvalue solver of asymptotic quasi-optimal computational complexity. *SIAM J. Numer. Anal.* **50**(3), 1029–1057 (2012)
10. Manuel Cascon, J., Kreuzer, Christian, Nochetto, Ricardo H., Siebert, Kunibert G.: Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.* **46**(5), 2524–2550 (2008)
11. Cancès, Clément, Pop, Iuliu Sorin, Vohralík, Martin: An a posteriori error estimate for vertex-centered finite volume discretizations of immiscible incompressible two-phase flow. *Math. Comp.* **83**(285), 153–188 (2014)
12. Chaillou, Alexandra L., Suri, Manil: A posteriori estimation of the linearization error for strongly monotone nonlinear operators. *J. Comput. Appl. Math.* **205**(1), 72–87 (2007)
13. Congreve, Scott, Wihler, Thomas P.: Iterative Galerkin discretizations for strongly monotone problems. *J. Comput. Appl. Math.* **311**, 457–472 (2017)
14. Deuffhard, Peter: Global inexact Newton methods for very large scale nonlinear problems. *Impact Comput. Sci. Engrg.* **3**(4), 366–393 (1991)
15. Di Fratta, Giovanni, Führer, Thomas, Gantner, Gregor, Praetorius, Dirk: Adaptive Uzawa algorithm for the Stokes equation. *ESAIM Math. Model. Numer. Anal.* **53**(6), 1841–1870 (2019)
16. Diening, Lars, Kreuzer, Christian: Linear convergence of an adaptive finite element method for the  $p$ -Laplacian equation. *SIAM J. Numer. Anal.* **46**(2), 614–638 (2008)
17. Dörfler, Willy: A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.* **33**(3), 1106–1124 (1996)
18. Di Pietro, Daniele A., Vohralík, Martin, Yousef, Soleiman: Adaptive regularization, linearization, and discretization and a posteriori error control for the two-phase Stefan problem. *Math. Comp.* **84**(291), 153–186 (2015)

19. El Alaoui, Linda, Ern, Alexandre, Vohralík, Martin: Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems. *Comput. Methods Appl. Mech. Engrg.* **200**(37–40), 2782–2795 (2011)
20. Ern, Alexandre, Vohralík, Martin: Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs. *SIAM J. Sci. Comput.* **35**(4), A1761–A1791 (2013)
21. Eisenstat, Stanley C., Walker, Homer F.: Globally convergent inexact Newton methods. *SIAM J. Optim.* **4**(2), 393–422 (1994)
22. Führer, Thomas: Zur Kopplung von finiten Elementen und Randelementen. PhD thesis, TU Wien, (2014)
23. Gantner, Gregor, Haberl, Alexander, Praetorius, Dirk, Stifner, Bernhard: Rate optimal adaptive FEM with inexact solver for nonlinear operators. *IMA J. Numer. Anal.* **38**(4), 1797–1831 (2018)
24. Gantner, Gregor, Haberl, Alexander, Praetorius, Dirk, Schimanko, Stefan: Rate optimality of adaptive finite element methods with respect to overall cost. Submitted, (2019)
25. Garau, Eduardo M., Morin, Pedro, Zuppa, Carlos: Convergence of an adaptive Kačanov FEM for quasi-linear problems. *Appl. Numer. Math.* **61**(4), 512–529 (2011)
26. Garau, Eduardo M., Morin, Pedro, Zuppa, Carlos: Quasi-optimal convergence rate of an AFEM for quasi-linear problems of monotone type. *Numer. Math. Theory Methods Appl.* **5**(2), 131–156 (2012)
27. Gallistl, Dietmar, Schedensack, Mira, Stevenson, Rob P.: A remark on newest vertex bisection in any space dimension. *Comput. Methods Appl. Math.* **14**(3), 317–320 (2014)
28. Houston, Paul, Wihler, Thomas P.: An  $hp$ -adaptive Newton-discontinuous-Galerkin finite element approach for semilinear elliptic boundary value problems. *Math. Comp.* **87**(314), 2641–2674 (2018)
29. Heid, Pascal, Wihler, Thomas P.: Adaptive iterative linearization Galerkin methods for nonlinear problems. *Math. Comp.* **89**, 2707–2734 (2020)
30. Heid, Pascal, Wihler, Thomas P.: On the convergence of adaptive iterative linearized Galerkin methods. *Calcolo* **57**, 24 (2020)
31. Karkulik, Michael, Pavlicek, David, Praetorius, Dirk: On 2D newest vertex bisection: optimality of mesh-closure and  $H^1$ -stability of  $L_2$ -projection. *Constr. Approx.* **38**(2), 213–234 (2013)
32. Kondratyuk, Yaroslav, Stevenson, Rob: An optimal adaptive finite element method for the Stokes problem. *SIAM J. Numer. Anal.* **46**(2), 747–775 (2008)
33. Morin, Pedro, Nochetto, Ricardo H., Siebert, Kunibert G.: Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.* **38**(2), 466–488 (2000)
34. Olshanskii, Maxim A., Tyrtshnikov, Eugene E.: Iterative methods for linear systems. Society for Industrial and Applied Mathematics, Philadelphia, PA (2014)
35. Pollock, Sara: Stabilized and inexact adaptive methods for capturing internal layers in quasilinear PDE. *J. Comput. Appl. Math.* **308**, 243–262 (2016)
36. Stevenson, Rob: Optimality of a standard adaptive finite element method. *Found. Comput. Math.* **7**(2), 245–269 (2007)
37. Stevenson, Rob: The completion of locally refined simplicial partitions created by bisection. *Math. Comp.* **77**(261), 227–241 (2008)
38. Stevenson, Rob: Adaptive wavelet methods for linear and nonlinear least-squares problems. *Found. Comput. Math.* **14**(2), 237–283 (2014)
39. Veese, Andreas: Convergent adaptive finite elements for the nonlinear Laplacian. *Numer. Math.* **92**(4), 743–770 (2002)
40. Zeidler, Eberhard: Nonlinear functional analysis and its applications. II/B. Springer, New York (1990)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Alexander Haberl<sup>1</sup> · Dirk Praetorius<sup>1</sup> · Stefan Schimanko<sup>1</sup> · Martin Vohralík<sup>2,3</sup>

✉ Dirk Praetorius  
dirk.praetorius@asc.tuwien.ac.at

Alexander Haberl  
alexander.haberl@asc.tuwien.ac.at

Stefan Schimanko  
stefan.schimanko@asc.tuwien.ac.at

Martin Vohralík  
martin.vohralik@inria.fr

- 1 TU Wien, Institute of Analysis and Scientific Computing, Wiedner Hauptstr. 8-10/E101/4, 1040 Vienna, Austria
- 2 Inria, 2 rue Simone Iff, 75589 Paris, France
- 3 CERMICS, Ecole des Ponts, 77455 Marne-la-Vallée, France