Convergence and quasi-optimal cost of adaptive algorithms for nonlinear operators including iterative linearization and algebraic solver
Alexander Haberl, Dirk Praetorius, Stefan Schimanko, Martin Vohralík

To cite this version:
Alexander Haberl, Dirk Praetorius, Stefan Schimanko, Martin Vohralík. Convergence and quasi-optimal cost of adaptive algorithms for nonlinear operators including iterative linearization and algebraic solver. Numerische Mathematik, Springer Verlag, 2021, 147 (3), pp.679-725. 10.1007/s00211-021-01176-w. hal-02557718v2

HAL Id: hal-02557718
https://hal.inria.fr/hal-02557718v2
Submitted on 22 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
CONVERGENCE AND QUASI-OPTIMAL COST OF ADAPTIVE ALGORITHMS FOR NONLINEAR OPERATORS INCLUDING ITERATIVE LINEARIZATION AND ALGEBRAIC SOLVER

ALEXANDER HABERL, DIRK PRAETORIUS, STEFAN SCHIMANKO, AND MARTIN VOHRALÍK

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.

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 \( A: \mathbb{R}^d \to \mathbb{R}^d \), we aim to numerically approximate the weak solution \( u^* \in H^1_0(\Omega) \) of the nonlinear boundary value problem

\[
- \text{div} A(\nabla u^*) = f \quad \text{in } \Omega, \\
\quad u^* = 0 \quad \text{on } \partial \Omega.
\]
To this end, we propose an adaptive algorithm of the type

\[
\text{estimate total error and its components} \downarrow \text{advance algebra/advance linearization/mark and refine mesh elements}
\]

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 Section 2 for details. Then, the main theorem on monotone operators yields the existence and uniqueness of the weak solution \( u^\star \in H_0^1(\Omega) \); see, e.g., [Zei90, Theorem 25.B]. Given a triangulation \( T_H \) of \( \Omega \), the lowest-order finite element method (FEM) for problem (1) reads as follows: Find \( u_H^\star \in X_H := \{ v_H \in C(\Omega) : v_H|_T \text{ is affine for all } T \in T_H \text{ and } v_H|_{\partial \Omega} = 0 \} \subset H_0^1(\Omega) \) such that

\[
(A(\nabla u_H^\star), \nabla v_H)_{\Omega} = (f, v_H)_{\Omega} \quad \text{for all } v_H \in X_H. \tag{3}
\]

The discrete solution \( u_H^\star \in X_H \) 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 \( \Phi_H : X_H \to X_H \) by

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

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

\[
\| \nabla (u_H^\star - \Phi_H(w_H)) \|_{L^2(\Omega)} \leq q_{\text{Pic}} \| \nabla (u_H^\star - w_H) \|_{L^2(\Omega)} \text{ with } q_{\text{Pic}} := (1 - \alpha^2/L^2)^{1/2} < 1; \tag{5}
\]

see, e.g., [Zei90, Section 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 \( u_H^\star \). 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., [OT14].

1.2. Fully adaptive algorithm. In our approach, we compute a sequence of discrete approximations \( u_{k,j}^\ell \) of \( u^\star \) 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 \( T_\ell \), we stop for some index \( j \in \mathbb{N} \). At the next linearization
step \( k \in \mathbb{N} \), the arising linear system reads as follows:

Find \( u_{k,j}^{k,*} \in X \) such that, for all \( v_{j} \in X \),

\[
(\nabla u_{k,j}^{k,*}, \nabla v_{j})_{\Omega} = \alpha \frac{1}{L^2} \left[ (A(\nabla u_{k-1,j}^{k-1}), \nabla v_{j})_{\Omega} - (f, v_{j})_{\Omega} \right],
\]

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

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

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

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., [Deu91, EW94] 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 \( \|\nabla (u^* - u_{T}^*)\|_{L^2(\Omega)} \) with respect to the degrees of freedom of FEM adapting the approximation space (mesh) were obtained in [Vee02, DK08, BDK12, GMZ12], following the seminal contributions [Dör96, MNS00, BDD04, Ste07, CKNS08] for linear problems. We also refer to [CFPP14] 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., [CS07, EAEV11], 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., [BDMS15, BCL15, CW17, HW18]. Later, theoretical proofs of plain convergence (without rates) were given in [GMZ11, HW20a], where [HW20a] builds on the unified framework of [HW20a] encompassing also the Kačanov and (damped) Newton linearizations in addition to the Banach–Picard linearization (6).

The own works [GHPS18, GHPS19] considered that the linear systems (6) are solved exactly at linear cost (so that \( u_{k,j}^{k,j} = u_{k,j}^{k,*} \) with \( j(\ell,k) = \mathcal{O}(1) \) in the present notation), as in the seminal work [Ste07] for the Poisson model problem and in [CG12] for an
adaptive Laplace eigenvalue computation. Under this so-called realistic assumption on the algebraic solver, we have proved in [GHPS18] 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 [GHPS18] was based on proving first that the estimator \( \eta_\ell(u^k_\ell) \) for the final Picard iterates decays with optimal rate \( s \) and second that the number of Picard iterates satisfies \( k(\ell) \lesssim 1 + \log[1 + \eta_\ell(u^k_{\ell+1})/\eta_\ell(u^k_\ell)] \). This logarithmic bound then led to the bound \( s - \varepsilon \) for the convergence rate with respect to the overall computational cost.

Recently in [GHPS19], 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 [0] is solved exactly at linear cost. The core idea of the new proof follows ideas from adaptive Uzawa FEM for the Stokes model problem [KS08, DFFGP19]. 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 [CW17, GHPS18, HW20a, HW20b, GHPS19] 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 [EV13] takes into account all discretization, linearization, and algebraic error components; see also [CPV14, DPVY15] and [Pol16] 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 computational cost.

We stress that such results have already been derived for adaptive wavelet discretizations [CDD03, Ste14] 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 [GHPS19], we stress that both works [GHPS18, GHPS19] 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 [GHPS19] 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 Section 1.2 gives rise to an index set

\[ Q := \{(\ell, k, j) \in \mathbb{N}_0^3 : \text{discrete approximation } u^k_j \text{ is computed by the algorithm}\} \]
together with an ordering

\[ |(\ell, k, j)| < |(\ell', k', j')| \quad \overset{\text{def}}{\iff} \quad u_k^j \text{ is computed earlier than } u_{k'}^{j'}.\]

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

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

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 \( \|\nabla (u^* - u_k^j)\|_{L^2(\Omega)} \) as well as the error estimator \( \eta_\ell (u_k^j) \).

Second, we prove the \textit{optimal error decay rate} with respect to the number of \textit{degrees of freedom} added with respect to the initial mesh in the sense that

\[ \sup_{(\ell, k, j) \in Q} (\# T_\ell - \# T_0 + 1)^s \Delta_{\ell}^{k,j} < \infty \quad (8) \]

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 \textit{optimal error decay rate} with respect to the \textit{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

\[ \sup_{(\ell', k', j') \in Q} \left( \sum_{(\ell, k, j) \in Q} (\# T_\ell) \right)^s \Delta_{\ell'}^{k',j'} < \infty \quad (9) \]

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 \( T_\ell \) (i.e., the number of degrees of freedom).

1.5. Outline. The remainder of the paper is organised as follows. In Section 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 Sections 3, 4, 5, and 6. Finally numerical experiments in Section 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 $A := dP : \mathcal{X} \rightarrow \mathcal{X}'$, i.e.,

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

We suppose that the operator $A$ is strongly monotone and Lipschitz-continuous, i.e.,

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

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 [Zei90, Section 25.4] yields existence and uniqueness of the solution $u^* \in \mathcal{X}$ of

$$\langle Au^*, v \rangle_{\mathcal{X}' \times \mathcal{X}} = \langle F, v \rangle_{\mathcal{X}' \times \mathcal{X}} \quad \text{for all } v \in \mathcal{X}.$$ (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

$$\langle Au^*_H, v \rangle_{\mathcal{X}' \times \mathcal{X}} = \langle F, v \rangle_{\mathcal{X}' \times \mathcal{X}} \quad \text{for all } v \in \mathcal{X}_H.$$ (12)

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

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

see, e.g., [GHPS18, Lemma 5.1]. In particular, $u^* \in \mathcal{X}$ (resp. $u^*_H \in \mathcal{X}_H$) is the unique minimizer of the minimization problem

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

As for linear elliptic problems, it follows from (10)–(12) that the present setting guarantees the Céa lemma (see, e.g., [Zei90, Section 25.4])

$$\| u^* - u^*_H \| \leq C_{\text{Céa}} \| u^* - v_H \| \quad \text{for all } v_H \in \mathcal{X}_H \quad \text{with } C_{\text{Céa}} := L/\alpha.$$ (15)

2.2. Mesh-refinement. Let $T_H$ be a conforming simplicial mesh of $\Omega$, i.e., a partition of $\overline{\Omega}$ into compact simplices $T$ such that $\bigcup_{T \in T_H} T = \overline{\Omega}$ and such that the intersection of two different simplices is either empty or their common vertex, edge, or face. We assume that refine($\cdot$) is a fixed mesh-refinement strategy, e.g., newest vertex bisection [Ste08]. We write $T_h = \text{refine}(T_H, \mathcal{M}_H)$ for the coarsest one-level refinement of $T_H$, where all marked elements $\mathcal{M}_H \subseteq T_H$ have been refined, i.e., $\mathcal{M}_H \subseteq T_H \setminus T_h$. We write $T_h \in \text{refine}(T_H)$, if $T_h$ can be obtained by finitely many steps of one-level refinement (with appropriate, yet arbitrary marked elements in each step). We define $T := \text{refine}(T_0)$ as the set of all meshes which can be generated from the initial simplicial mesh $T_0$ of...
\[ \Omega \text{ by use of } \text{refine}(\cdot). \] Finally, we associate to each \( \mathcal{T}_H \in \mathbb{T} \) a corresponding finite-dimensional subspace \( \mathcal{X}_H \subseteq \mathcal{X} \), 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 (\( \text{R1} \)–\( \text{R3} \)), where \( C_{\text{son}} \geq 2 \) and \( C_{\text{mesh}} > 0 \) are generic constants:

**\( \text{R1} \)** 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
\[
\#(\mathcal{T}_H \setminus \mathcal{T}_h) + \# \mathcal{T}_H \leq \# \mathcal{T}_h \leq C_{\text{son}} \#(\mathcal{T}_H \setminus \mathcal{T}_h) + \#(\mathcal{T}_H \cap \mathcal{T}_h);
\]

**\( \text{R2} \)** 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
\[
\#(\mathcal{T}_h \oplus \mathcal{T}'_h) \leq \# \mathcal{T}_h + \# \mathcal{T}'_h - \# \mathcal{T};
\]

**\( \text{R3} \)** 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
\[
\# \mathcal{T}_\ell - \# \mathcal{T}_0 \leq C_{\text{mesh}} \sum_{j=0}^{\ell-1} \# \mathcal{M}_j.
\]

For newest vertex bisection, we refer to \([\text{BDD04}, \text{Ste07}, \text{Ste08}, \text{CKNS08}, \text{KPP13}, \text{GSS14}]\) for the validity of (\( \text{R1} \)–\( \text{R3} \)). For red-refinement with first-order hanging nodes, details are found in \([\text{BN10}]\).

### 2.3. Error estimator

For each mesh \( \mathcal{T}_H \in \mathbb{T} \), suppose that we can compute refinement indicators
\[
\eta_H(\mathcal{T}, v_H) \geq 0 \quad \text{for all } \mathcal{T} \in \mathcal{T}_H \text{ and all } v_H \in \mathcal{X}_H.
\]

We denote
\[
\eta_H(\mathcal{V}_H, v_H) := \left( \sum_{T \in \mathcal{V}_H} \eta_H(\mathcal{T}, v_H)^\omega \right)^{1/2} \quad \text{for all } \mathcal{V}_H \subseteq \mathcal{T}_H
\]

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 \([\text{CFPP14}]\) 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:

- **\( \text{A1} \)** 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 \( \mathcal{V}_H \subseteq \mathcal{T}_H \cap \mathcal{T}_h \);
- **\( \text{A2} \)** 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 \( v_H \in \mathcal{X}_H \);
- **\( \text{A3} \)** reliability: \( \| u^* - u_H^* \| \leq C_{\text{rel}} \eta_H(\mathcal{U}_H) \);
- **\( \text{A4} \)** 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 (\( \text{A3} \)–\( \text{A4} \)) will never be computed but are only auxiliary quantities for the analysis.

We refer to Section 7.1 below for precise assumptions on the nonlinearity \( A(\cdot) \) of problem (1) such that the standard residual error estimator satisfies (\( \text{A1} \)–\( \text{A4} \)) for lowest-order Courant finite elements; see also Section 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

\[
(u_H^j, v_H) = G(v_H) \quad \text{for all } v_H \in \mathcal{X}_H
\]  

with unique (but not computed) exact solution \( u_H^\star \in \mathcal{X}_H \). 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

\[
\|u_H - u_H^j\| \leq q_{\text{alg}} \|u_H - u_H^{j-1}\| \quad \text{for all } j \geq 1.
\]  

Examples for such solvers are suitably preconditioned conjugate gradients or multigrid; see, e.g., Olshanskii and Tyrtyshnikov \([OT14]\) 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 \( j \). On each step \((\ell, k, j)\), it yields an approximation \( u_{\ell,j}^{k,j} \in \mathcal{X}_\ell \) to the unique but unavailable \( u_\ell^\star \in \mathcal{X}_\ell \) on the mesh \( T_\ell \) defined by

\[
\langle A u_\ell^\star - F, v_\ell \rangle_{\mathcal{X}' \times \mathcal{X}} = 0 \quad \text{for all } v_\ell \in \mathcal{X}_\ell.
\]  

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

**Algorithm 1.** Input: Initial mesh \( T_0 \) and initial guess \( u_0^0 = u_0^0 \in \mathcal{X}_0 \), parameters \( 0 \leq \theta \leq 1 \), \( 0 < \lambda_{\text{alg}} < 1 \), \( 0 < \lambda_{\text{Pic}} \), \( 1 \leq C_{\text{mark}} \), 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 \( u_{\ell}^{k,j} := u_{\ell}^{k,j} \) 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,j} \in \mathcal{X}_\ell \text{ such that, for all } v_\ell \in \mathcal{X}_\ell,

\]

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

\[
\text{from (20)}
\]  

| counter | discrete solution |
|---------|-------------------|
| running | stopping | running | stopping | exact |
| mesh    | \( \ell \) | \( \ell \) | \( u_{\ell}^{k,j} \) | \( u_{\ell}^{k,j} \) | \( u_\ell^\star \) from (20) |
| linearization | \( k \) | \( k \) | \( u_{\ell}^{k,j} \) | \( u_{\ell}^{k,j} \) | \( u_\ell^\star \) from (21) |
| algebraic solver | \( j \) | \( j \) | \( u_{\ell}^{k,j} \) | \( u_{\ell}^{k,j} \) |

Table 1. Counters and discrete solutions in Algorithm 1.
and do one step of the algebraic solver applied to (21) starting from $u^{k,j}_{\ell-1}$, which yields $u^{k,j}_{\ell}$ (an approximation to $u^{k,j*}_{\ell}$).

(III) Compute the local indicators $\eta(\ell, u^{k,j}_{\ell})$ for all $T \in T_{\ell}$.

(IV) If $\eta(\ell, u^{k,j}_{\ell}) + \|u^{k,j}_{\ell} - u^{k-1,j}_{\ell}\| + \|u^{k,j}_{\ell} - u^{k,j-1}_{\ell}\| \leq \tau$, then set $\ell := \ell$, $k(\ell) := k$, and $j(\ell, k) := j$ and terminate Algorithm 1.

Until $\|u^{k,j}_{\ell} - u^{k,j-1}_{\ell}\| \leq \lambda_{\text{alg}} \left[ \eta(\ell, u^{k,j}_{\ell}) + \|u^{k,j}_{\ell} - u^{k-1,j}_{\ell}\| \right]$. (22)

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

Until $\|u^{k,j}_{\ell} - u^{k-1,j}_{\ell}\| \leq \lambda_{\text{Pic}} \eta(\ell, u^{k,j}_{\ell})$. (23)

(ii) Define $k := k(\ell) := k$.

(iii) If $\eta(\ell, u^{k,j}_{\ell}) = 0$, then set $\ell := \ell$ and terminate Algorithm 1.

(iv) Determine a set $\mathcal{M}_{\ell} \subseteq T_{\ell}$ with up to the multiplicative constant $C_{\text{mark}}$ minimal cardinality such that

$$\theta \eta(\ell, u^{k,j}_{\ell}) \leq \eta(\ell, u^{k,j}_{\ell}).$$ (24)

(v) Generate $T_{\ell+1} := \text{refine}(T_{\ell}, \mathcal{M}_{\ell})$ and define $u^{0,j}_{\ell+1} := u^{0,j}_{\ell+1} := u^{k,j}_{\ell}$.

(vi) Update counters $\ell := \ell + 1$, $k := 0$, and $j := 0$ and continue with (i).

Output: Sequence of discrete solutions $u^{k,j}_{\ell}$ and corresponding error estimators $\eta(\ell, u^{k,j}_{\ell})$.

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

$$\|u^{k,j}_{\ell} - u^{k,j-1}_{\ell}\| \leq \lambda_{\text{alg}} \|u^{k,j*}_{\ell} - u^{k,j-1}_{\ell}\|$$ for all $j \geq 1$. (25)

Then, the triangle inequality implies that

$$\frac{1 - \lambda_{\text{alg}}}{\lambda_{\text{alg}}} \|u^{k,j*}_{\ell} - u^{k,j}_{\ell}\| \leq \|u^{k,j}_{\ell} - u^{k,j-1}_{\ell}\| \leq (1 + \lambda_{\text{alg}}) \|u^{k,j*}_{\ell} - u^{k,j-1}_{\ell}\|.$$ (26)

Hence, the term $\|u^{k,j}_{\ell} - u^{k,j-1}_{\ell}\|$ provides a means to estimate the algebraic error $\|u^{k,j*}_{\ell} - u^{k,j}_{\ell}\|$. In particular, the approximation $u^{k,j}_{\ell}$ is accepted and the algebraic solver is stopped if the algebraic error estimate $\|u^{k,j}_{\ell} - u^{k,j-1}_{\ell}\|$ is, up to the threshold $\lambda_{\text{alg}}$, below the estimate $\eta(\ell, u^{k,j}_{\ell}) + \|u^{k,j}_{\ell} - u^{k-1,j}_{\ell}\|$ of the discretization and linearization error; see (22).

Since $\|u^{k,j}_{\ell} - u^{k,j}_{\ell}\| = \|u^{k,j}_{\ell} - u^{k,j}_{\ell}\|$, 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 $\|u^{k,j}_{\ell} - u^{k,j-1,j}_{\ell}\|$ estimates the linearization error $\|u^{k,j*}_{\ell} - u^{k,j}_{\ell}\|$. Note here that, a priori, only the non-perturbed
Banach–Picard iteration corresponding to the (unavailable) exact solve of (21) yielding \( u_{k,\ell}^* \) would lead to the contraction

\[
\| u_{l}^{*} - u_{k,\ell}^{k+1} \| \leq q_{\text{Pic}} \| u_{l}^{*} - u_{k,\ell}^{k} \| \quad \text{for all } (\ell, k, 0) \in \mathcal{Q} \text{ with } k \geq 1,
\]

where

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

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

Finally, the outermost loop steers the local adaptive mesh-refinement. To this end, the Dörfler marking criterion (24) from [Dör96] is employed to mark elements \( T \in \mathcal{M}_{\ell} \) for refinement, unless \( \eta_{\ell}(u_{k,\ell}^{k}) = 0 \), in which case Proposition 2 below ensures that the approximation \( u_{k,\ell}^{k+1} \) coincides with the exact solution \( u^{*} \) of (11). In practice, the computation is stopped as soon as the computed iterate \( u_{k,\ell}^{k} \) 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

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

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

\[
(\ell, k, j) < (\ell', k', j') \iff (\ell, k, j) \text{ appears earlier in Algorithm 1 than } (\ell', k', j').
\]

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 \( j(\ell, 0) := 0 \):

\[
\ell := \sup \{ \ell \in \mathbb{N}_0 : (\ell, 0, 0) \in \mathcal{Q} \} \in \mathbb{N}_0 \cup \{ \infty \},
\]

\[
k(\ell) := \sup \{ k \in \mathbb{N}_0 : (\ell, k, 0) \in \mathcal{Q} \} \in \mathbb{N}_0 \cup \{ \infty \} \quad \text{if } (\ell, 0, 0) \in \mathcal{Q},
\]

\[
j(\ell, k) := \sup \{ j \in \mathbb{N}_0 : (\ell, k, j) \in \mathcal{Q} \} \in \mathbb{N}_0 \cup \{ \infty \} \quad \text{if } (\ell, k, 0) \in \mathcal{Q}.
\]

Generically, it holds that \( \ell = \infty \), 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.,

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

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

March 4, 2021 10
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 \( k := k(\ell) \), e.g., \( u^{k,j}_\ell := u^{k(\ell),j}_\ell \). Similarly, we simply write \( j := j(\ell, k) \), e.g., \( u^{k,j}_\ell := u^{k(\ell),j(\ell,k)}_\ell \).

Note that there in particular holds \( u^{k,j}_{\ell-1} = u^0_{\ell} = u^1_{\ell} \) 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 (1b.12) of Algorithm 1.

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

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

\[
\|u^* - u^{k,j}_\ell\| \leq C_{\text{rel}}^\star \begin{cases} 
\eta_{\ell}(u^{k,j}_\ell) + \|u^{k,j}_\ell - u^{k-1,j}_{\ell-1}\| & \text{if } 0 < k \leq k(\ell) \text{ and } 0 < j \leq j(\ell, k), \\
\eta_{\ell}(u^{k,j}_\ell) + \|u^{k,j}_\ell - u^{k-1,j}_{\ell-1}\| & \text{if } 0 < k \leq k(\ell) \text{ and } j = j(\ell, k), \\
\eta_{k}(u^{k,j}_\ell) & \text{if } k = k(\ell) \text{ and } j = j(\ell, k), \\
\eta_{k-1}(u^{k,j}_{\ell-1}) & \text{if } k = 0 \text{ and } \ell > 0.
\end{cases}
\]

The constant \( C_{\text{rel}}^\star > 0 \) depends only on \( C_{\text{rel}}, C_{\text{stab}}, q_{\text{alg}}, \lambda_{\text{alg}}, \lambda_{\text{Pic}}, q_{\text{Pic}}, \) 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 Section 4.

**Theorem 3 (linear convergence).** Suppose \( (\text{A1}) - (\text{A3}) \). 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 \( 0 < \lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} < \lambda_{\text{alg}}^\star \) and \( 0 < \lambda_{\text{Pic}}/\theta < \lambda_{\text{Pic}}^\star \), there exist constants \( 1 \leq C_{\text{lin}} \) and \( 0 < q_{\text{lin}} < 1 \) such that the quasi-error

\[
\Delta_{\ell}^{k,j} := \|u^* - u^{k,j}_\ell\| + \|u^{k,j}_\ell - u^{k,j}_{\ell}\| + \eta_{\ell}(u^{k,j}_\ell),
\]

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

\[
\Delta_{\ell}^{k,j'} \leq C_{\text{lin}} q_{\text{lin}}^{[(\ell', k', j')]-[(\ell, k, j)]} \Delta_{\ell}^{k,j}
\]

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 \( u^{k,j'}_{\ell} = u^{k,j}_\ell \) and where \( u^{k,j'}_{\ell} = u^{k,j}_\ell \) together with \( T_{\ell'} = T_{\ell} \), and consequently \( \eta_{\ell'}(u^{k,j'}_{\ell'}) = \eta_{\ell}(u^{k,j}_\ell) \), related to our notational choice for \( \mathcal{Q} \) in (29) that also indexes nested iterates. The case with \( \ell' = \ell \) arises for instance when \( j = j', j' = 0, \) and \( k' = k + 1; \) see step (ia) of Algorithm 1. Note, however, that in

March 4, 2021

11
such a situation, typically $u_{k',*}^{k,*} \neq u_{k}^{k,*}$, and consequently $\Delta_{k'}^{k',j} \neq \Delta_{k}^{k,j}$. A situation where $\Delta_{k'}^{k',j} = \Delta_{k}^{k,j}$ for $(\ell',k',j') \neq (\ell,k,j)$ can nevertheless also appear and is covered in (33). For instance, in the above example, when $j = j'$, $j' = 0$, $k' = k + 1$, and $\ell' = \ell$, and where moreover $u_{\ell}^{k,j} = u_{\ell}^{k,*,*} = u_{\ell}^{*}$ (so that $u_{\ell}^{k,j} = u_{\ell}^{k,*,*} = u_{\ell'}^{k',*,*} = u_{\ell'}^{*}$), 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 $\Delta_{k}^{k,j}$ of (32) (and consequently of the total error $\|u^{*} - u_{k}^{k,*}\|$) in terms of the number of degrees of freedom added in the space $X_{k}$ with respect to the mesh $T$. 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 $T_{0}$ (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. Theorem 4 (optimal decay rate wrt. degrees of freedom).

Suppose $\{A1\}$ - $\{A4\}$ and $\{R1\}$ - $\{R3\}$. Recall $\lambda_{\text{alg}}$, $\lambda_{\text{Pic}} > 0$ 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 $\theta^{*}$ such that for all $0 < \lambda_{\text{alg}}$, $\lambda_{\text{Pic}}$, $\theta$ with $0 < \theta < \min\{1, \theta^{*}\}$ as well as $\lambda_{\text{alg}} < 1$, $0 < C_{\text{alg}} + C_{\text{Pic}} C_{\text{alg}}$, and $0 < \lambda_{\text{Pic}}^*/\lambda_{\text{Pic}}$, it holds that

$$0 < \theta^{*} = \frac{\theta + C_{\text{stab}} (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}} + [C_{\text{Pic}} + (1 + C_{\text{Pic}}) C_{\text{alg}} \lambda_{\text{alg}}] \lambda_{\text{Pic}}}{1 - \lambda_{\text{Pic}}^*/\lambda_{\text{Pic}}} < \theta_{\text{opt}}, \quad (34)$$

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

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

where $\eta_{\text{opt}}(u_{s}^{*})$ is the error estimator corresponding to the exact solution of (12) with respect to the mesh $T_{\text{opt}}$ and

$$\mathcal{T}(N) := \{T \in \mathcal{T} : \#T - \#T_{0} \leq N\}.$$

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

$$c_{\text{opt}} \|u^{*}\|_{\lambda_{s}} \leq \sup_{(\ell,k,j) \in \Omega} (\#T_{\ell} - \#T_{0} + 1)^{s} \Delta_{k,j}^{k,j} \leq C_{\text{opt}} \max\{\|u^{*}\|_{\lambda_{s}}, \Delta_{0,0}^{0,0}\}. \quad (36)$$

The constant $c_{\text{opt}} > 0$ depends only on $C_{C_{\text{sta}}} = L/\alpha$, $C_{\text{stab}}, C_{\text{ref}}, C_{\text{son}}, \#T_{0}$, and $s$, and, if $\ell < \infty$, additionally on $\ell$. The constant $C_{\text{opt}} > 0$ depends only on $C_{\text{stab}}, C_{\text{ref}}, C_{\text{mark}}, 1 - \lambda_{\text{Pic}}^*/\lambda_{\text{Pic}}$, $C_{C_{\text{sta}}} = L/\alpha$, $C_{\text{ref}}, C_{\text{mesh}}, C_{\text{lin}}, q_{\text{lin}}, \#T_{0}$, and $s$. The maximum in the right inequality is only needed if $\ell = 0$. If $\ell \geq 1$, the maximum max$\{\|u^{*}\|_{\lambda_{s}}, \Delta_{0,0}^{0,0}\}$ can be replaced by $\|u^{*}\|_{\lambda_{s}}$.

Note that $\Delta_{0,0}^{0,0}$ can be arbitrarily bad due to a bad initial guess $u_{0,0}^{0,0}$. However, $\|u^{*}\|_{\lambda_{s}}$ as well as the constant $C_{\text{opt}}$ are independent of the initial guess, so that the upper bound in (36) cannot avoid max$\{\|u^{*}\|_{\lambda_{s}}, \Delta_{0,0}^{0,0}\}$ 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 \((\ell, k, j) \in Q\) 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 \( \sigma \) not only with respect to the number of degrees of freedom added in the space \( X_\ell \) in comparison with \( X_0 \), 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 \( \sigma \) with respect to the number of mesh elements added in the refinement of \( T_0 \) (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 \( \sigma \) with respect to the overall algorithmic cost expressed as the sum of the number of simplices \( \#T_\ell \) over all steps \((\ell, k, j) \in Q\) effectuated by Algorithm 1. The proof of the following Theorem 5 is given in Section 6.

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

\[
c_{\text{opt}}^{-1} \|u^*\|_{A_\sigma} \leq \sup_{(\ell', k', j') \in Q} \left( \sum_{(\ell, k, j) \in Q, (\ell, k, j) \leq (\ell', k', j')} \#T_\ell \right)^{\sigma} \Delta_\ell^{k', j'} \leq C_{\text{opt}}' \max\{\|u^*\|_{A_\sigma}, \Delta_0^{0,0}\}. \tag{37}
\]

The maximum in the right inequality is only needed if \( \ell = 0 \). If \( \ell \geq 1 \), the maximum \( \max\{\|u^*\|_{A_\sigma}, \Delta_0^{0,0}\} \) can be replaced by \( \|u^*\|_{A_\sigma} \). While \( c_{\text{opt}} > 0 \) is the constant of Theorem 4, the constant \( C_{\text{opt}}' > 0 \) reads

\[
C_{\text{opt}}' = (\#T_0)^s C_{\text{opt}} C_{\text{lin}} (1 - q_1^{1/\sigma})^{-s}.
\]

Analogously to the comments after Theorem 4, the upper estimate in (37) cannot avoid \( \max\{\|u^*\|_{A_\sigma}, \Delta_0^{0,0}\} \) for the case \( \ell' = \ell = 0 \). As above, if one restricts the indices to \((\ell', k', j'), (\ell, k, j) \in Q\) 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 \( T_\ell \), the cost of its one step is proportional to \( \#T_\ell \). This also holds true for matrix and right-hand-side assembly in (21), evaluation of the residual estimators \( \eta(\ell,j) \), 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 \( O(1) \). Thus, the sum in (37) is indeed proportional to the overall computational cost invested into the numerical approximation of \( u^* \) 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

\[
u_0^{0,0} = u_0 = u_0^0 \in X_0.
\]

If \((\ell, 0, 0) \in Q\) with \(\ell \geq 1\), then

\[
u_\ell^{0,*} := u_\ell^{0,0} := u_\ell^0 := u_{\ell-1}^{0,0} \in X_{\ell-1} \subseteq X_\ell.
\]
If \((\ell, k, 0) \in \mathcal{Q}\), then the initial guess for the algebraic solver reads

\[
 u^k_\ell = \begin{cases} 
 u_{0,0}^{\ell} & \text{for } \ell = 0, \\
 u_{k-j}^{\ell-1} & \text{if } k = 0 \text{ and } \ell \geq 1, \\
 u_{k-j}^{\ell} & \text{if } k > 0,
\end{cases}
\]

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

\[
\|u^{k,j}_\ell - u^{k-1,j}_\ell\| \leq \lambda_{\text{alg}} \left[ \eta_\ell(u^{k,j}_\ell) + \|u^{k,j}_\ell - u^{k-1,j}_\ell\| \right] \quad \text{for } j = j(\ell, k),
\]

\[
\|u^{k,j}_\ell - u^{k-1,j}_\ell\| > \lambda_{\text{alg}} \left[ \eta_\ell(u^{k,j}_\ell) + \|u^{k,j}_\ell - u^{k-1,j}_\ell\| \right] \quad \text{for } j < j(\ell, k),
\]
i.e., the algebraic error estimate \(\|u^{k,j}_\ell - u^{k-1,j}_\ell\|\) only drops below the discretization plus linearization error estimate at the stopping iteration \(j = j(\ell, k)\).

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

\[
\|u^{k,j}_\ell - u^{k-1,j}_\ell\| \leq \lambda_{\text{pic}} \eta_\ell(u^{k,j}_\ell) \quad \text{for } k = k(\ell),
\]

\[
\|u^{k,j}_\ell - u^{k-1,j}_\ell\| > \lambda_{\text{pic}} \eta_\ell(u^{k,j}_\ell) \quad \text{for } k < k(\ell),
\]
i.e., the linearization error estimate \(\|u^{k,j}_\ell - u^{k-1,j}_\ell\|\) only drops below the discretization error estimate at the stopping iteration \(k = k(\ell)\).

### 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 \(\|u^{k,j}_\ell - u^{k-1,j}_\ell\|\). 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

\[
\frac{1 - q_{\text{pic}}}{q_{\text{pic}}} \|u^*_{\ell} - u^k_{\ell}\| \leq \|u^{k,*}_{\ell} - u^{k-1,j}_\ell\| \leq (1 + q_{\text{pic}}) \|u^*_{\ell} - u^{k-1,j}_\ell\|
\]

of the unavailable linearization error estimate \(\|u^{k,*}_{\ell} - u^{k-1,j}_\ell\|\). 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 \(u^*_{\ell} \in X_\ell\) is the (unavailable) exact discrete solution given by \((20)\), that \(u^{k,*}_{\ell} \in X_\ell\) is the (unavailable) exact linearization solution given by \((21)\), and that \(u^{k,j}_\ell \in X_\ell\) is the computed solution for which the algebraic solver is stopped; see \((22)\) (resp. \((41)-(42)\)) for the stopping criterion.

**Lemma 6.** There exists \(\lambda^*_{\text{alg}} > 0\) depending only on \(q_{\text{alg}}\) and \(q_{\text{pic}}\) such that

\[
0 < q'_{\text{pic}} := \frac{q_{\text{pic}} + \frac{1}{1 - q_{\text{alg}}} \lambda^*_{\text{alg}}}{\lambda^*_{\text{alg}}} < 1.
\]
Moreover, for all stopping parameters \(0 < \lambda_{\text{alg}} < 1\) and \(0 < \lambda_{\text{Pic}}\) from (22)–(23) such that

\[
0 < \lambda_{\text{alg}} + \lambda_{\text{Pic}} < \lambda_{\text{alg}}^* \text{, it holds that}
\]

\[
\| u^* - u^k \| \leq q_{\text{Pic}}^k \| u^* - u^{k-1} \| \quad \text{for all } 1 \leq k < k(\ell).
\]

This also implies that

\[
\frac{1 - q_{\text{Pic}}^k}{q_{\text{Pic}}^k} \| u^* - u^k \| \leq \| u^k - u^{k-1} \| \leq (1 + q_{\text{Pic}}^k) \| u^* - u^{k-1} \|.
\]

**Proof.** Clearly, (47) follows from (47) by the triangle inequality as in (26) and (45). Moreover, (46) is obvious for sufficiently small \(\lambda_{\text{alg}}^*\), 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

\[
\| u^* - u^k \| \leq \| u^* - u^k \| + \| u^k - u^k \| \leq q_{\text{Pic}}\| u^* - u^k \| + \| u^k - u^k \|,
\]

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

\[
\| u^k - u^k \| \leq \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \| u^k - u^k \| \leq \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}} \| u^* - u^k - k, j \|
\]

\[
\leq \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \left( \lambda_{\text{alg}} + \lambda_{\text{alg}}/\lambda_{\text{Pic}} \right) \| u^k - u^k \|.
\]

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

\[
\| u^* - u^k \| \leq (q_{\text{Pic}} + \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}}) \| u^* - u^k - k, j \| + \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}} \| u^* - u^k - k, j \|.
\]

If \(0 < \lambda_{\text{alg}}^*\) is sufficiently small, this shows (46)–(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 \(0 < k \leq k(\ell)\) and \(0 < j \leq j(\ell, k)\). Due to stability (A1), reliability (A3), and the contraction properties (26) resp. (45), it holds that

\[
\| u^* - u^k \| \leq \| u^* - u^k \| + \| u^k - u^k \| \leq \eta(\ell) \| u^* - u^k \| + \| u^k - u^k \|
\]

\[
\leq \eta(\ell) \| u^k \| + \| u^k - u^k \| \leq \eta(\ell) \| u^k \| + \| u^k - u^k \|
\]

\[
\leq \eta(\ell) \| u^k \| + \| u^k - u^k \| + \| u^k - u^k \| + 2\| u^k - u^k \|
\]

\[
\leq \eta(\ell) \| u^k \| + \| u^k - u^k \| + \| u^k - u^k \|.
\]

This proves (31) for the case \(0 < k \leq k(\ell)\) and \(0 < j \leq j(\ell, k)\).
If \( j = j(\ell, k) \), we can improve estimate \( (49) \) using the stopping criterion \( (41) \). Similarly, if \( k = k(\ell) \) and \( j = j(\ell, k) \), we can improve estimate \( (49) \) using the stopping criteria \( (41) \) and \( (43) \). Finally, for \( k = 0, \ell > 0 \) and hence \( j = j = 0 \), it directly follows from nested iteration \( (39) \) and the previous case \( k = k(\ell - 1) \) resp. \( j = j(\ell - 1, k) \) that

\[
\|u^* - u^{0,0}_\ell\| \overset{(40)}{=} \|u^* - u^{k,j}_{\ell-1}\| \overset{(41)}{\lesssim} \eta_{\ell-1}(u^{k,j}_{\ell-1}).
\]

This concludes the proof. \( \square \)

### 3.4. An auxiliary adaptive algorithm.

Due to Lemma \ref{lem:4.9}, the iterates \( u^{k,j}_\ell \) are contractive in the index \( k \). Consequently, Algorithm \ref{alg:1} fits into the framework of \cite{GHPS18} upon defining \( u_\ell \) from \cite{GHPS18} as \( u_\ell := u^{k,j}_\ell \) for the case where \( k(\ell) < \infty \) and \( j(\ell, k) < \infty \), i.e., both the algebraic and the linearization solvers are stopped by \( \ell' \) and \( \ell'' \) 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 \cite{GHPS18}, Lemma 4.9, eq. \( (4.10) \) in the current setting to conclude two important properties: First, the estimators \( \eta_\ell(u^{k,j}_\ell) \) available at step (iv) of Algorithm \ref{alg: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 \( \eta_\ell(u^{k,j}_\ell) \) are linearly convergent.

**Lemma 7** ([\cite{GHPS18}, Lemma 4.9, Theorem 5.3]). Recall \( \lambda^*_\text{alg} > 0 \) and \( 0 < q_\text{Pic} < 1 \) from Lemma \ref{lem:6}. Define \( \lambda^*_\text{Pic} := \frac{1 - q_\text{Pic} }{q_\text{Pic} \lambda^*_\text{stab}} > 0 \) 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 \( (\ell, k, j) \in \mathcal{Q} \) with \( k < \infty \) and \( j < \infty \), it holds that

\[
(1 - \lambda_\text{Pic} / \lambda^*_\text{Pic}) \eta_\ell(u^{k,j}_\ell) \leq \eta_\ell(u^*_\ell) \leq (1 + \lambda_\text{Pic} / \lambda^*_\text{Pic}) \eta_\ell(u^{k,j}_\ell).
\]  

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

\[
\eta_{\ell+n}(u^{k,j}_{\ell+n}) \leq C_{\text{GHPS}} q_{\text{GHPS}}^n \eta_\ell(u^{k,j}_\ell) \quad \text{for all } (\ell + n + 1, 0, 0) \in \mathcal{Q}.
\]

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 \ref{lem:7} and Proposition \ref{prop:2}, we get the following lemma for the quasi-error of \( (32) \) on stopping indices \( k(\ell), j(\ell, k) \). Please note that when \( \ell < \infty \), the summation below only goes to \( \ell - 1 \), as the arguments rely on \( (52) \) which needs finite stopping indices \( k(\ell) \) and \( j(\ell, k) \) 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 \ref{lem:6}) as well as \( 0 < \theta < 1 \) and \( 0 < \lambda_\text{Pic} / \theta < \lambda^*_\text{Pic} \) (from Lemma \ref{lem:7}). With the convention \( \ell - 1 = \infty \) if \( \ell = \infty \), there holds summability

\[
\sum_{\ell = \ell + 1}^{\ell - 1} \Delta_{\ell}^{k,j} \leq C \Delta_{\ell'}^{k,j} \quad \text{for all } (\ell', k, j) \in \mathcal{Q},
\]

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 \( \lambda_\text{Pic} \).

March 4, 2021

16
Proof. Define \( \tilde{\Delta}_k := \|u^* - u_{k,j}^\ell\| + \eta \ell(u_{k,j}^\ell) \) as the sum of overall error plus error estimator. In comparison with (32), \( \Delta_k^\ell \) omits the algebraic error term but is only defined for the algebraic stopping indices \( j(\ell, k) \). With Proposition 2 and the linear convergence (52), we get that

\[
\sum_{\ell = \ell' + 1}^{\ell-1} \tilde{\Delta}_k^\ell \leq \sum_{\ell = \ell' + 1}^{\ell-1} \eta \ell(u_{k,j}^\ell) + \eta \ell(u_{k,j}^\ell) \leq \sum_{\ell = \ell' + 1}^{\ell-1} \eta \ell(u_{k,j}^\ell) + \|u_{k,j}^\ell - u_{k,j}^{\ell'}\|.
\]

Let \((\ell', k, j) \in Q\). By definition (32), it holds that

\[
\Delta_{k,j}^{\ell,j} = \|u^* - u_{k,j}^{\ell,j}\| + \|u_{k,j}^{\ell'} - u_{k,j}^{\ell,j}\| + \eta \ell(u_{k,j}^{\ell,j}) = \Delta_k^\ell + \|u_{k,j}^{\ell'} - u_{k,j}^{\ell,j}\|.
\]

Moreover, note that

\[
\|u_{k,j}^{\ell'} - u_{k,j}^{\ell,j}\| \leq \sum_{\ell = \ell' + 1}^{\ell-1} \eta \ell(u_{k,j}^\ell) + \|u_{k,j}^{\ell'} - u_{k,j}^{\ell-1}\| \leq \eta \ell(u_{k,j}^{\ell'}) \leq \Delta_k^\ell.
\]

This proves the equivalence \( \Delta_{k,j}^{\ell,j} \simeq \tilde{\Delta}_k^\ell \) for all \((\ell', k, j) \in Q\) and concludes the proof. □

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 4) 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

\[
\sum_{(\ell, k, j) \in Q} \Delta_{k,j}^{\ell,j} \leq C_{\text{sum}} \Delta_{k,j}^{\ell,j} \quad \text{for all } (\ell', k', j') \in Q.
\]

The constant \( C_{\text{sum}} \) depends only on \( C_{\text{rel}}, C_{\text{stab}}, q_{\text{rel}}, \theta, q_{\text{alg}}, \lambda_{\text{alg}}, q_{\text{pic}}, \lambda_{\text{pic}}, \alpha, \) and \( L \).

**Proof.** **Step 1.** We prove that

\[
\Delta_{k,j}^{\ell,j} := \|u^* - u_{k,j}^\ell\| + \|u_{k,j}^{\ell'} - u_{k,j}^\ell\| + \eta \ell(u_{k,j}^\ell) \simeq \Delta_k^\ell \quad \text{for all } (\ell, k, j) \in Q.
\]

Note that \( \Delta_{k,j}^{\ell,j} \) and \( \Delta_k^\ell \) 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

\[
\|u^* - u_{k,j}^\ell\| \leq \|u^* - u_{k,j}^{\ell'}\| + \|u_{k,j}^{\ell'} - u_{k,j}^\ell\| \overset{15}{\leq} \|u^* - u_{k,j}^{\ell'}\| \leq \Delta_k^\ell.
\]

This implies that \( \Delta_{k,j}^{\ell,j} \simeq \Delta_k^\ell \). To see the converse inequality, note that

\[
\|u^* - u_{k,j}^{\ell,j}\| \leq \|u^* - u_{k,j}^\ell\| + \|u_{k,j}^{\ell'} - u_{k,j}^{\ell,j}\| \overset{A_3}{\leq} \eta \ell(u_{k,j}^\ell) + \|u_{k,j}^{\ell'} - u_{k,j}^{\ell,j}\| \overset{A_3}{\leq} \eta \ell(u_{k,j}^{\ell'}) + \|u_{k,j}^{\ell'} - u_{k,j}^{\ell,j}\| \leq \Delta_k^\ell.
\]

This proves \( \Delta_{k,j}^{\ell,j} \simeq \Delta_k^\ell \) and concludes this step.
Step 2. We prove some auxiliary estimates. First, we prove that the algebraic error \( \| u_\ell^{k,j} - u_\ell^{k,j-1} \| \) dominates the modified total error \( A_\ell^{k,j} \), before the algebraic stopping criterion (22) is reached, i.e.,

\[
A_\ell^{k,j} \lesssim \| u_\ell^{k,*} - u_\ell^{k,j-1} \| \quad \text{for all } (\ell, k, j) \in \mathbb{Q} \text{ with } k \geq 1 \text{ and } 1 \leq j < j(\ell, k).
\] (56)

To this end, note that

\[
\| u_\ell^* - u_\ell^{k,j} \| \leq \| u_\ell^* - u_\ell^{k,j-1} \| + \| u_\ell^{k,j-1} - u_\ell^{k,j} \| \lesssim \| u_\ell^* - u_\ell^{k,j-1} \|
\]

Second, we consider the use of nested iteration when passing to the next perturbed algebraic iteration step before the linearization stopping criterion (23) is reached, i.e.,

\[
\| u_\ell^{k,j} - u_\ell^{k,j-1} \| \lesssim \| u_\ell^{k,j} - u_\ell^{k-1,j} \| \lesssim \| u_\ell^{k,j} - u_\ell^{k-1,j} \| \lesssim \| u_\ell^{k,j} - u_\ell^{k-1,j} \|.
\]

Third, we prove that

\[
A_\ell^{k,j} \lesssim A_\ell^{k,j} \quad \text{for all } (\ell, k, j) \in \mathbb{Q},
\] (58)

related to the algebraic error contraction. Note that \( k = 0 \) implies \( j = 0 \), 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 \( j = j(\ell, k) - 1 \geq 0 \). There holds that

\[
A_\ell^{k,j} = \| u_\ell^* - u_\ell^{k,j} \| + \| u_\ell^{k,j} - u_\ell^{k,j-1} \| + \eta_\ell(u_\ell^{k,j}) \lesssim \| u_\ell^{k,j} - u_\ell^{k,j-1} \| \lesssim \| u_\ell^{k,j} - u_\ell^{k-1,j} \| \lesssim \| u_\ell^{k,j} - u_\ell^{k-1,j} \|.
\]

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.,

\[
A_\ell^{k,j} \lesssim \| u_\ell^* - u_\ell^{k-1,j} \| \quad \text{for all } (\ell, k, j) \in \mathbb{Q} \text{ with } 1 \leq k < k(\ell).
\] (59)
To see this, note that $1 \leq k < k(\ell)$ yields that
\[
A_{k,j}^\ell \leq \|u_*^\ell - u_{k,j}^\ell\| + \|u_*^\ell - u_{k,j}^\ell\| + \eta_l(u_{k,j}^\ell) \leq \|u_*^\ell - u_{k,j}^\ell\| + \|u_*^\ell - u_{k,j}^\ell\| + \eta_l(u_{k,j}^\ell)
\]
which we employ Lemma 6 and hence require $0 < \lambda_{alg} + \lambda_{alg}/\lambda_{Pic}$ to be sufficiently small.

Fifth, we consider the use of nested iteration when refining the mesh. We prove that
\[
A_{k,j}^\ell \leq \eta_{l-1}(u_{k,j}^{\ell-1}) \leq A_{k,j}^{\ell-1} \quad \text{for all } (\ell, k, j) \in Q. \tag{60}
\]
To this end, note that
\[
\|u_*^\ell - u_{k,j}^{\ell-1}\| \leq \|u_*^\ell - u_{k,j}^{\ell}\| + \|u_*^\ell - u_{k,j}^{\ell-1}\| \leq \|u_*^\ell - u_{k,j}^{\ell}\| \leq \eta_{l-1}(u_{k,j}^{\ell-1}). \tag{61}
\]
Next, recall from (39) that $u_{k,j}^{0,\ell} = u_{k,j}^{0,\ell-1} = u_{k,j}^{\ell-1}$. From (A1) used on non-refined mesh elements and (A2) used on refined mesh elements, we hence conclude that
\[
A_{k,j}^{0,\ell} \leq \|u_*^\ell - u_{k,j}^{\ell-1}\| + \eta_l(u_{k,j}^{\ell-1}) \leq \eta_{l-1}(u_{k,j}^{\ell-1}) \leq 2 \eta_{l-1}(u_{k,j}^{\ell-1}). \tag{62}
\]
Sixth, we prove that
\[
A_{k,j}^{\ell} \leq A_{k,j}^{\ell-1} \quad \text{for all } (\ell, k, j) \in Q. \tag{63}
\]
We first consider $k = k(\ell) - 1 \geq 0$. Note that
\[
\|u_*^{k,j} - u_*^{k-1,j}\| \leq \|u_*^{k,j} - u_*^{k,j}\| + \|u_*^{k,j} - u_*^{k-1,j}\| \leq \|u_*^{k,j} - u_*^{k-1,j}\| \leq A_{k,j}^{k-1,j}. \tag{64}
\]
Hence, the triangle inequality leads to
\[
A_{k,j}^{k,j} = \|u_*^{k,j} - u_*^{k,j}\| + \|u_*^{k,j} - u_*^{k,j}\| + \eta_l(u_{k,j}^{k,j}) \leq \|u_*^{k,j} - u_*^{k,j}\| + \|u_*^{k,j} - u_*^{k,j}\| + \eta_l(u_{k,j}^{k,j}) \leq A_{k,j}^{k-1,j} + \|u_*^{k,j} - u_*^{k-1,j}\| \leq 2 A_{k,j}^{k-1,j}.
\]
This proves (62) for $k = k(\ell) - 1$. Note that the same argument also applies when $k = 1$. If $0 \leq k \leq k(\ell) - 2$, then
\[
A_{k,j}^{k,j} \leq A_{k,j}^{k-1,j} \leq \|u_*^{k,j} - u_*^{k-2,j}\| \leq \|u_*^{k,j} - u_*^{k,j}\| \leq A_{k,j}^{k,j},
\]
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
\[
A_{k,j}^{k,0} \leq A_{k,j}^{k-1,j} \quad \text{for all } (\ell, k, 0) \in Q \text{ with } k \geq 1. \tag{64}
\]
Using (57) and recalling the definition \( u_0^{k,0} = u_0^{k-1,j} \), it holds that
\[
A_0^{k,0} = \| u_0^k - u_0^{k-1,j} \| + \| u_0^{k,j} - u_0^{k,0} \| + \eta_0 (u_0^{k-1,j}) \lesssim A_0^{k-1,j},
\]
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 \( j(\ell,k) - 1 = \infty \) when \( j(\ell,k) = \infty \), it holds that
\[
\sum_{j=i+1}^{j(\ell,k)-1} A_{\ell}^{k,j} \lesssim \| u_0^{k,*} - u_0^{k,j} \| \leq A_{\ell}^{k,i} \quad \text{for all } (\ell,k) \in \mathcal{Q} \text{ with } k \geq 1.
\]
This follows immediately from
\[
\sum_{j=i+1}^{j(\ell,k)-1} A_{\ell}^{k,j} \lesssim \sum_{j=i+1}^{j(\ell,k)-1} \| u_0^{k,*} - u_0^{k,j-1} \| \lesssim \| u_0^{k,*} - u_0^{k,j} \| \sum_{j=i}^{\infty} q_{alg}^{j-i} \lesssim \| u_0^{k,*} - u_0^{k,j} \|.
\]
Analogously, with the convention that \( k(\ell) - 1 = \infty \) when \( k(\ell) = \infty \), the contraction (47) of the perturbed Banach–Picard iteration leads to
\[
\sum_{k=i+1}^{k(\ell)-1} A_{\ell}^{k,j} \lesssim \| u_0^{\ell,*} - u_0^{i,j} \| \leq A_{\ell}^{i,j} \quad \text{for all } (\ell,i,j) \in \mathcal{Q}.
\]
This follows immediately from
\[
\sum_{k=i+1}^{k(\ell)-1} A_{\ell}^{k,j} \lesssim \sum_{k=i+1}^{k(\ell)-1} \| u_0^{k,*} - u_0^{k-1,j} \| \lesssim \| u_0^{k,*} - u_0^{i,j} \| \sum_{k=i}^{\infty} (q_{Pic}^k)^{-k-i} \lesssim \| u_0^{k,*} - u_0^{i,j} \|.
\]
With the analogous convention \( \ell - 1 = \infty \) when \( \ell = \infty \), we finally prove that
\[
\sum_{\ell=1+1}^{\ell-1} A_{\ell}^{k,j} \lesssim A_{\ell}^{k,i} \quad \text{for all } (i,k,j) \in \mathcal{Q}.
\]
This follows from Step 1 and
\[
\sum_{\ell=1+1}^{\ell-1} A_{\ell}^{k,j} = \sum_{\ell=1+1}^{\ell-1} A_{\ell}^{k,j} \lesssim \sum_{\ell=1+1}^{\ell-1} \Delta_{\ell}^{k,j} \lesssim \Delta_{\ell}^{k,j} \lesssim A_{\ell}^{k,j}.
\]

**Step 4.** From now on, let \((\ell', k', j') \in \mathcal{Q}\) be arbitrary. Suppose first that \( \ell = \infty \), i.e., both algebraic and linearization solvers terminate at some finite values \( k(\ell) \) for all \( \ell \geq 0 \) and \( j(\ell,k) \) for all \( \ell \geq 0 \) and all \( k \leq k(\ell) \), 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

$$
\sum_{(\ell, k, j) \in \mathcal{Q}} A^k_j \sum_{\ell' + 1} \left( A^{k,0}_\ell + \sum_{k=1}^{k(\ell)} A^{j(k),0}_\ell + \sum_{j=1}^{j(\ell, k)} A^{j,k}_\ell \right)
$$

$$
+ \sum_{k=1}^{k(\ell)} \left( A^{k,0}_{\ell'} + \sum_{j=1}^{j(\ell, k)} A^{j,k}_{\ell'} \right) + \sum_{j=1}^{j(\ell, k')} A^{j,k}_{\ell'}
$$

$$
\lesssim \sum_{\ell' + 1} \sum_{k=1}^{k(\ell)} \sum_{j=1}^{j(\ell, k)} A^{k,j}_{\ell'} \lesssim \sum_{\ell' + 1} \sum_{k=1}^{k(\ell)} \left( A^{j,k}_{\ell'} + A^{j,k-1}_{\ell'} \right)
$$

$$
\lesssim \sum_{\ell' + 1} \left( A^{0,j}_{\ell'} + \sum_{k=1}^{k(\ell)} A^{k,j}_{\ell'} \right) \lesssim \sum_{\ell' + 1} \left( A^{0,j}_{\ell' - 1} + A^{0,j}_{\ell'} \right)
$$

$$
\lesssim A^{0,j}_{\ell' - 1} + \sum_{\ell' + 1} A^{0,j}_{\ell'} \lesssim A^{0,j}_{\ell'} \lesssim A^{0,j}_{\ell'} \lesssim A^{0,j}_{\ell'} \lesssim A^{0,j}_{\ell'}.
$$

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

$$
\sum_{k=k' + 1}^{k(\ell')} \sum_{j=1}^{j(\ell', k')} A^{k,j}_{\ell'} \lesssim \sum_{k=k' + 1}^{k(\ell')} \left( A^{k,j}_{\ell'} + \| u^{k,j}_{\ell'} - u^{k,0}_{\ell'} \| \right) \lesssim \sum_{k=k' + 1}^{k(\ell')} \left( A^{k,j}_{\ell'} + A^{k-1,j}_{\ell'} \right)
$$

$$
\lesssim A^{k,j}_{\ell'} + \sum_{k=k' + 1}^{k(\ell')} A^{k,j}_{\ell'} \lesssim A^{k,j}_{\ell'} + A^{k,j}_{\ell'} \lesssim A^{k,j}_{\ell'} \lesssim A^{k,j}_{\ell'}.
$$

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

$$
\sum_{j=j' + 1}^{j(\ell', k')} A^{k,j}_{\ell'} \lesssim A^{k,j}_{\ell'} + A^{k,j}_{\ell'} \lesssim A^{k,j}_{\ell'}.
$$

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

$$
\sum_{(\ell, k, j) \in \mathcal{Q}} A^{k,j}_{\ell} \lesssim A^{k,j}_{\ell'},
$$
Step 5. Suppose that $\ell < \infty$ and $k(\ell) = \infty$, i.e., for the mesh $T_{\ell}$, the linearization loop does not terminate. Moreover, let $\ell' < \ell$. Then, it holds as in (68) that
\[
\sum_{(\ell, k, j) \in Q} A^{k, j}_{\ell} \lesssim \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} A^{k, j}_{\ell} + \sum_{\ell = \ell' + 1}^{\ell-1} \sum_{k'=1}^{k(\ell')} \sum_{j=1}^{j(\ell,k,k')} A^{k, j}_{\ell'} + \sum_{k=k'+1}^{\ell-1} \sum_{j=1}^{j(\ell,k,k')} A^{k, j}_{\ell'} + \sum_{j=j'+1}^{j(\ell', k', k')} A^{k, j}_{\ell'}.
\] (73)

We argue as before to see that
\[
\sum_{\ell = \ell' + 1}^{\ell-1} \sum_{k'=1}^{k(\ell')} \sum_{j=1}^{j(\ell,k,k')} A^{k, j}_{\ell'} \lesssim A^{k', j'}_{\ell'}, \quad \sum_{k=k'+1}^{\ell-1} \sum_{j=1}^{j(\ell,k,k')} A^{k, j}_{\ell'} \lesssim A^{k', j'}_{\ell'}, \quad \text{and} \quad \sum_{j=j'+1}^{j(\ell', k', k')} A^{k, j}_{\ell'} \lesssim A^{k', j'}_{\ell'}.
\] (74)

It only remains to estimate
\[
\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} A^{k, j}_{\ell} \lesssim \sum_{k=1}^{\infty} (A^{k, j}_{\ell} + \| u_{k}^{x*} - u_{k}^{0} \|) \lesssim A^{0, j}_{\ell} + \sum_{k=1}^{\infty} A^{k, j}_{\ell} \lesssim A^{j}_{\ell}.
\] (75)

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

Step 6. Suppose that $\ell < \infty$ and $k(\ell) = \infty$, i.e., for the mesh $T_{\ell}$, the linearization loop does not terminate, and moreover, $\ell' = \ell$. Arguing as in (75) and (71), it holds that
\[
\sum_{(\ell, k, j) \in Q} A^{k, j}_{\ell} \lesssim \sum_{k=k'+1}^{\ell-1} \sum_{j=1}^{j(\ell,k,k')} A^{k, j}_{\ell'} + \sum_{j=j'+1}^{j(\ell', k', k')} A^{k, j}_{\ell'}.
\] (76)

Step 7. Suppose that $\ell < \infty$, where $k(\ell) < \infty$ and hence $j(\ell, k) = \infty$, i.e., the linear solver does not terminate for the linearization step $k(\ell)$. Suppose moreover $\ell' < \ell$. Then, it holds that
\[
\sum_{(\ell, k, j) \in Q} A^{k, j}_{\ell} \lesssim \sum_{k=1}^{\infty} A^{k, j}_{\ell} + \sum_{k=k'+1}^{\ell-1} \sum_{j=1}^{j(\ell,k,k')} A^{k, j}_{\ell'} + \sum_{k=k'+1}^{\ell-1} \sum_{j=1}^{j(\ell,k,k')} A^{k, j}_{\ell'} + \sum_{k=k'+1}^{\ell-1} \sum_{j=1}^{j(\ell', k', k')} A^{k, j}_{\ell'}.
\] (77)

We argue as before to see that
\[
\sum_{\ell = \ell' + 1}^{\ell-1} \sum_{k'=1}^{k(\ell')} \sum_{j=1}^{j(\ell,k,k')} A^{k, j}_{\ell'} \lesssim A^{k', j'}_{\ell'}, \quad \sum_{k=k'+1}^{\ell-1} \sum_{j=1}^{j(\ell,k,k')} A^{k, j}_{\ell'} \lesssim A^{k', j'}_{\ell'}, \quad \text{and} \quad \sum_{j=j'+1}^{j(\ell', k', k')} A^{k, j}_{\ell'} \lesssim A^{k', j'}_{\ell'}.
\] (78)

For the first sum in (77), we get that
\[
\sum_{j=1}^{\infty} A^{k, j}_{\ell} \lesssim \| u_{k}^{x*} - u_{k}^{0} \| \lesssim A^{k-1, j}_{\ell} \lesssim A^{j}_{\ell}.
\] (78)
Finally, the second sum in (71) can be bounded analogously to (75) in Step 5 by $A_{\ell'}^{k,j}$. Altogether, we obtain (72) provided that $\ell' < \ell < \infty$, $k(\ell) < \infty$, and $j(\ell,k) = \infty$.

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

$$
\sum_{(\ell,k,j) \in Q_{\ell'}} A_{\ell'}^{k,j} \lesssim \sum_{j=1}^{\infty} A_{\ell'}^{k,j} + \sum_{k=k'+1}^{k(\ell')} \sum_{j=1}^{j(\ell',k')} A_{\ell'}^{k,j'}, \quad \sum_{(\ell,k,j) \in Q_{\ell'}} A_{\ell'}^{k,j} \lesssim A_{\ell'}^{k,j}.
$$

(79)

We argue as before to see that

$$
\sum_{j=1}^{\infty} A_{\ell'}^{k,j} \lesssim A_{\ell'}^{k,j}, \quad \sum_{k=k'+1}^{k(\ell')} \sum_{j=1}^{j(\ell',k')} A_{\ell'}^{k,j} \lesssim A_{\ell'}^{k,j'}, \quad \sum_{j=1}^{j(\ell',k')} A_{\ell'}^{k,j'} \lesssim A_{\ell'}^{k,j'}.
$$

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

**Step 9.** Suppose that $\ell < \infty$, where $k(\ell) < \infty$ and hence $j(\ell,k) = \infty$, i.e., the linear solver does not terminate for the linearization step $k(\ell)$. Suppose $\ell' = \ell$ and $k' = k(\ell)$.

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

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

$$
\sum_{(\ell,k,j) \in Q_{\ell'}} A_{\ell'}^{k,j} \lesssim \sum_{k=1}^{k(\ell')} \sum_{j=1}^{j(\ell,k)} A_{\ell'}^{k,j} + \sum_{\ell' = \ell' + 1}^{k(\ell')} \sum_{k=1}^{k(\ell')} \sum_{j=1}^{j(\ell',k)} A_{\ell'}^{k,j} + \sum_{j=1}^{j(\ell',k)} A_{\ell'}^{k,j}.
$$

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 $A_{\ell'}^{k,j}$ is zero.

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

$$
\sum_{(\ell,k,j) \in Q_{\ell'}} A_{\ell'}^{k,j} \lesssim \sum_{k=k'+1}^{k(\ell')} \sum_{j=1}^{j(\ell',k')} A_{\ell'}^{k,j} + \sum_{j=j'+1}^{j(\ell',k')} A_{\ell'}^{k,j'},
$$

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

$$
\sum_{(\ell,k,j) \in Q_{\ell'}} A_{\ell'}^{k,j} \lesssim A_{\ell'}^{k,j'}.
$$

(65)

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

March 4, 2021 23
Lemma 10. Suppose \( R_1 \) as well as \([A1],[A2], \) and \([A4].\) Let \( s > 0 \) and assume \( \| u^\star \|_{\Lambda_s} > 0.\) For tolerance \( \tau = 0, \) it then holds that
\[
\| u^\star \|_{\Lambda_s} \leq c_{opt} \sup_{(\ell',k',j') \in Q} (\# T_{\ell'} - \# T_0 + 1)^s \Delta_{\ell',k',j'},
\]
where the constant \( c_{opt} > 0 \) depends only on \( C_{C\text{ea}} = L/\alpha, C_{\text{stab}}, C_{\text{rel}}, C_{\text{son}}, \# T_0, s, \) and, if \( \ell < \infty, \) additionally on \( \ell.\)
Proof. The proof is split into three steps. First, we recall from [BHP17, Lemma 22] that
\[ \#T_h/\#T_H \leq \#T_h - \#T_H + 1 \leq \#T_h \quad \text{for all } T_h \in \mathbb{T} \text{ and all } T_h \in \text{refine}(T_H). \] (81)

Step 1. We consider the three non-generic cases with \( \ell < \infty \). First, let \( k(\ell) < \infty \), and \( j(\ell, k) < \infty \). Then, Algorithm 1 was terminated in Step (iii) with \( \eta_{k,j}(u^k_{\ell}) = 0 \). Due to the Cea lemma [15] and Proposition 2, it follows that
\[ \|u^* - u^k_{\ell}\| \leq \|u^* - u^k_{\ell,j}\| + \eta_{k,j}(u^k_{\ell}) = 0 \]
and hence \( u^* = u^k_{\ell} = u^k_{\ell,j} \) and \( \eta_{k,j}(u^k_{\ell}) = 0 \).

Second, let \( k(\ell) < \infty \) but \( j(\ell, k) = \infty \), i.e., the algebraic solver does not stop. According to Theorem 3, it holds that
\[ \Delta_{k,j}^\ell = \|u^* - u^k_{\ell,j}\| + \|u^k_{\ell,j} - u^k_{\ell,j}\| + \eta_{k,j}(u^k_{\ell,j}) \to 0 \quad \text{as } j \to \infty. \]
Hence, we obtain that \( u^* = u^k_{\ell,j} = u^k_{\ell,j} \). From stability (A1), it follows that
\[ 0 \leq \eta_{k,j}(u^k_{\ell,j}) \leq \eta_{k,j}(u^k_{\ell,j}) + \|u^k_{\ell,j} - u^k_{\ell,j}\| \to 0 \quad \text{as } j \to \infty. \]
Hence, we see that \( \eta_{k,j}(u^k_{\ell}) = 0 \).

Finally, let \( k(\ell) = \infty \), i.e., the linearization solver does not stop. Analogously to the previous case, we obtain that
\[ \Delta_{k,j}^\ell = \|u^* - u^k_{\ell,j}\| + \|u^k_{\ell,j} - u^k_{\ell,j}\| + \eta_{k,j}(u^k_{\ell,j}) \to 0 \quad \text{as } k \to \infty. \]
Hence, we get that \( u^* = u^\star_{\ell} \). Again, stability (A1) yields that \( \eta_{k,j}(u^\star_{\ell}) = 0 \).

In any case, \( \ell < \infty \) implies that \( \eta_{k,j}(u^\star_{\ell}) = 0 \) and hence
\[ \|u^*\|_{\mathbb{A}} = \sup_{0 \leq N < \#T_L - \#T_0} (N + 1)^s \inf_{T_{opt} \in \mathbb{T}(N)} \eta_{opt}(u^*_{opt}). \]
The term \( N + 1 \) within the supremum can be estimated by
\[ N + 1 \leq \#T_L - \#T_0 \leq (C^\ell_{son} - 1) \#T_0. \]
Moreover, (A1), (A2), and (A4) yield quasi-monotonicity \( \eta_{opt}(u^*_{opt}) \lesssim \eta_0(u^\star_0) \) (see, e.g., [CFPP14, Lemma 3.5]). Altogether, we thus arrive at
\[ \|u^*\|_{A_s} \lesssim \eta_0(u^\star_0) \leq \sup_{\ell' \in \mathbb{N}_0} (\#T_{\ell'} - \#T_0 + 1)^s \eta_{\ell'}(u^\star_\ell). \] (82)

Step 2. We consider the generic case that \( \ell = \infty \) and \( \eta_{k,j}(u^k_{\ell,j}) > 0 \) for all \( \ell \in \mathbb{N}_0 \). Algorithm 1 then guarantees that \( \#T_{\ell} \to \infty \) as \( \ell \to \infty \). Thus, we can argue analogously to the proof of [CFPP14, Theorem 4.1]: Let \( N \in \mathbb{N} \). Choose the maximal \( \ell' \in \mathbb{N}_0 \) such that \( \#T_{\ell'} - \#T_0 + 1 \leq N \). Then, \( T_{\ell'} \in \mathbb{T}(N) \). The choice of \( N \) guarantees that
\[ N + 1 \leq \#T_{\ell' + 1} - \#T_0 + 1 \leq \#T_{\ell' + 1} \leq C_{son} \#T_{\ell'} \leq C_{son} \#T_0 \leq \#T_{\ell'} - \#T_0 + 1. \] (83)
This leads to
\[ (N + 1)^s \inf_{T_{opt} \in \mathbb{T}(N)} \eta_{opt}(u^*_{opt}) \lesssim (\#T_{\ell'} - \#T_0 + 1)^s \eta_{\ell'}(u^\star_{\ell'}). \]

March 4, 2021
25
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

$$\|u^*\|_{A_\ast} \lesssim \sup_{\ell' \in \mathbb{N}_0} (\#T_{\ell'} - \#T_0 + 1)^s \eta_{\ell'}(u_{\ell'}^*). \quad (84)$$

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

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

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

$$\|u^*\|_{A_\ast} \lesssim \sup_{(\ell',k') \in \mathcal{Q}} (\#T_{\ell'} - \#T_0 + 1)^s \eta_{\ell'}(u_{\ell'}^*) \leq \sup_{(\ell',k') \in \mathcal{Q}} (\#T_{\ell'} - \#T_0 + 1)^s \Delta_{\ell'}^{k',k'}. $$

This concludes the proof. $\square$

To prove the upper estimate in (36), we need the comparison lemma from CFPP14 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 $0 < \theta' < \theta_{opt} := (1 + C_{stab}^2 C_{rel}^2)^{-1}$. Then, there exists constants $C_1, C_2 > 0$ such that for all $s > 0$ with $0 < \|u^*\|_{A_\ast} < \infty$ and all $T_H \in \mathcal{T}$, there exists $\mathcal{R}_H \subseteq T_H$ which satisfies

$$\#\mathcal{R}_H \leq C_1 C_2^{-1/s} \|u^*\|_{A_\ast}^{1/s} \eta_H(u_{\mathcal{R}_H}^*)^{-1/s}, \quad (85)$$

as well as the Dörfler marking criterion

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

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

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

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

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

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

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

**Step 2.** Recall that $C_{Pic} = q_{Pic}/(1 - q_{Pic})$ and $C_{alg} = q_{alg}/(1 - q_{alg})$. Provided that $(\ell + 1, 0, 0) \in \mathcal{Q}$, it follows from the contraction properties (26) resp. (45), and the stopping conditions...
The hidden constant depends only on $C$. 

Recall from (40) that $u_{\ell}^{k_j \pm \frac{1}{2}} \leq \|u_{\ell}^* - u_{\ell}^{k_j \pm \frac{1}{2}}\| + \|u_{\ell}^{k_j \pm \frac{1}{2}} - u_{\ell}^{k_j \mp \frac{1}{2}}\|

\leq (1 + C_{\text{Pic}})\|u_{\ell}^* - u_{\ell}^{k_j \pm \frac{1}{2}}\| + C_{\text{Pic}}\|u_{\ell}^{k_j \pm \frac{1}{2}} - u_{\ell}^{k_j \mp \frac{1}{2}}\|

\leq (1 + C_{\text{Pic}})\|u_{\ell}^* - u_{\ell}^{k_j \pm \frac{1}{2}}\| + C_{\text{Pic}}\|u_{\ell}^{k_j \pm \frac{1}{2}} - u_{\ell}^{k_j \mp \frac{1}{2}}\|

\leq (1 + C_{\text{Pic}})C_{\text{stab}}\eta(u_{\ell}^{k_j \pm \frac{1}{2}}) + \|u_{\ell}^* - u_{\ell}^{k_j \pm \frac{1}{2}}\|

\leq \left((1 + C_{\text{Pic}})C_{\text{stab}}\eta(u_{\ell}^{k_j \pm \frac{1}{2}}) + \|u_{\ell}^* - u_{\ell}^{k_j \pm \frac{1}{2}}\|\right)\theta'(1 - \lambda_{\text{Pic}}'\lambda_{\text{Pic}}^* - \theta)\eta(u_{\ell}^{k_j \pm \frac{1}{2}}).

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

$$\eta(\mathcal{R}_\ell, u_{\ell}^*) \leq \eta(\mathcal{R}_\ell, u_{\ell}^{k_j \pm \frac{1}{2}}) + C_{\text{stab}}\|u_{\ell}^* - u_{\ell}^{k_j \pm \frac{1}{2}}\|$$

$$\leq \eta(\mathcal{R}_\ell, u_{\ell}^{k_j \pm \frac{1}{2}}) + \left(\theta'(1 - \lambda_{\text{Pic}}'\lambda_{\text{Pic}}^* - \theta)\eta(u_{\ell}^{k_j \pm \frac{1}{2}})\right).$$

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

$$\theta'(1 - \lambda_{\text{Pic}}'\lambda_{\text{Pic}}^*\eta(u_{\ell}^{k_j \pm \frac{1}{2}}) \leq \eta(\mathcal{R}_\ell, u_{\ell}^*) \leq \eta(\mathcal{R}_\ell, u_{\ell}^{k_j \pm \frac{1}{2}}) + \left(\theta'(1 - \lambda_{\text{Pic}}'\lambda_{\text{Pic}}^* - \theta)\eta(u_{\ell}^{k_j \pm \frac{1}{2}})\right).$$

Thus, we are led to

$$\theta\eta(u_{\ell}^{k_j \pm \frac{1}{2}}) \leq \eta(\mathcal{R}_\ell, u_{\ell}^{k_j \pm \frac{1}{2}}).$$

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

$$\#\mathcal{M}_\ell \leq \#\mathcal{R}_\ell \leq \|u^*\|_{A_{s+1}}^{1/s} \eta(u_{\ell}^*)^{-1/s} \leq \|u^*\|_{A_{s+1}}^{1/s} \eta(u_{\ell}^{k_j \pm \frac{1}{2}})^{-1/s}.$$

Recall from (40) that $u_{\ell+1}^{0, j} = u_{\ell}^{k_j \pm \frac{1}{2}}$. Thus, (63) and the equivalence (55) lead to

$$\eta(u_{\ell}^{k_j \pm \frac{1}{2}})^{-1/s} \leq (A_{s+1})^{-1/s} \leq (\Delta_{\ell+1})^{-1/s}.$$

Overall, we end up with

$$\#\mathcal{M}_\ell \leq \|u^*\|_{A_{s+1}}^{1/s} (\Delta_{\ell+1})^{-1/s} \quad \text{for all } (\ell + 1, 0, 0) \in \mathcal{Q}. \quad (88)$$

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 $s$. 

March 4, 2021 27
Step 4. With linear convergence \textsuperscript{[33]} and the geometric series, we see that
\[
\sum_{((\ell, k, j) \in \mathcal{Q}, (\ell, k, j) \leq (\ell, k, j))} (\Delta_{\ell}^{k,j})^{-1/s} \lesssim (\Delta_{\ell}^{k,j})^{-1/s} \sum_{((\ell, k, j) \in \mathcal{Q}, (\ell, k, j) \leq (\ell, k, j))} (q_{\text{lin}}^{1/s})^{((\ell, k, j))-(((\ell, k, j)))} \lesssim (\Delta_{\ell}^{k,j})^{-1/s}
\]  \text{(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 \textsuperscript{[R3]} lead to
\[
\# \mathcal{T}_{\ell} - \# \mathcal{T}_{0} + 1 \simeq \# \mathcal{T}_{\ell} - \# \mathcal{T}_{0} \simeq \sum_{\ell=0}^{\ell-1} \# \mathcal{M}_{\ell} \simeq \|u^*\|_{H_s} \sum_{\ell=0}^{\ell} (\Delta_{\ell}^{k,j})^{-1/s} \lesssim \|u^*\|_{H_s} (\Delta_{\ell}^{k,j})^{-1/s}.
\]
Replacing \(\|u^*\|_{H_s}\) with \(\max\{\|u^*\|_{H_s}, \Delta_{0}^{0,0}\}\), the overall estimate trivially holds for \(\mathcal{T}_{\ell} = \mathcal{T}_{0}\). This proves that
\[
(\# \mathcal{T}_{\ell} - \# \mathcal{T}_{0} + 1)^s \Delta_{\ell}^{k,j} \lesssim \left\{ \begin{array}{ll}
\max\{\|u^*\|_{H_s}, \Delta_{0}^{0,0}\}, & \text{if } (\ell + 1, 0, 0) \in \mathcal{Q} \text{ and } \ell \geq 0, \\
\|u^*\|_{H_s}, & \text{if } (\ell + 1, 0, 0) \in \mathcal{Q} \text{ and } \ell \geq 1.
\end{array} \right.
\]  \text{(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, in holds that \(1 \leq \ell = \ell < \infty\), and one of the cases discussed in detail in Step 1 of Lemma \textsuperscript{[10]} arises.

First, let \(2 \leq \ell = \ell < \infty\). Since \(\ell - 1 \geq 1\) and \((\ell, 0, 0) \in \mathcal{Q}\), \textsuperscript{[90]} shows that
\[
(\# \mathcal{T}_{\ell-1} - \# \mathcal{T}_{0} + 1)^s \Delta_{\ell}^{k,j} \lesssim \|u^*\|_{H_s}.
\]
Moreover, Lemma \textsuperscript{[9]} leads to \(\Delta_{\ell}^{k,j} \lesssim \Delta_{\ell-1}^{k,j}\). Therefore, we obtain from \textsuperscript{[83]} that
\[
\# \mathcal{T}_{\ell} - \# \mathcal{T}_{0} + 1 \leq C_{\text{son}} \# \mathcal{T}_{0}(\# \mathcal{T}_{\ell-1} - \# \mathcal{T}_{0} + 1).
\]  \text{(91)}
Altogether, \textsuperscript{[90]} holds for this case as well.

Second, let \(\ell = \ell = 1\). Then, we can rely on the inequality
\[
(\# \mathcal{T}_{1} - \# \mathcal{T}_{0} + 1)^s \Delta_{1}^{k,j} \lesssim C_{\text{son}}(\# \mathcal{T}_{0}) \Delta_{1}^{k,j} \lesssim \Delta_{0}^{k,j}
\]
\[
\lesssim \|u^* - u_{0}^{k,j}\| + \|u_{0}^{k,j} - u_{0}^{k,j} - u_{0}^{k,j}\| + \eta_{0}(u_{0}^{k,j})
\]
\[
\lesssim \|u^* - u_{0}^{k,j}\| + \|u_{0}^{k,j} - u_{0}^{k,j} - u_{0}^{k,j}\| + \|u_{0}^{k,j} - u_{0}^{k,j} - u_{0}^{k,j}\| + \eta_{0}(u_{0}^{k,j})
\]
\[
\lesssim \|u^* - u_{0}^{k,j}\| + \|u_{0}^{k,j} - u_{0}^{k,j} - u_{0}^{k,j}\| + \|u_{0}^{k,j} - u_{0}^{k,j} - u_{0}^{k,j}\| + \eta_{0}(u_{0}^{k,j})
\]
\[
\lesssim \|u^* - u_{0}^{k,j}\| + \eta_{0}(u_{0}^{k,j}) \lesssim \|u^* - u_{0}^{k,j}\| + \eta_{0}(u_{0}^{k,j}) \lesssim \eta_{0}(u_{0}^{k,j}) \lesssim \|u^*\|_{H_s}.
\]  \text{(92)}
Thus, \textsuperscript{[90]} holds for this case as well.

March 4, 2021 28
Finally, let $\ell = \ell' = 0$. Then, linear convergence (33) proves that
\[
\Delta_k^{\ell} \lesssim \Delta_0^{0,0} \enspace (93)
\]
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 $\#T_\ell - \#T_0 + 1 = 1 \leq \#T_\ell' - \#T_0 + 1$ 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 $(\ell', k', j') \in Q$. Employing the right inequality in (36) (cf. (90)), the geometric series proves that
\[
\sum_{(\ell,k,j) \in Q} \#T_\ell \lesssim \sum_{(\ell,k,j) \leq (\ell',k',j')} (\#T_\ell - \#T_0 + 1) \leq \#T_0 C_{\text{opt}}^{1/s} \max\{\|u^*\|_{A_{\text{fs}}}, \Delta_0^{0,0}\}^{1/s} \sum_{(\ell,k,j) \leq (\ell',k',j')} (\Delta_k^{\ell,j})^{-1/s}
\]
\[
\leq \#T_0 C_{\text{opt}}^{1/s} C_{\text{lin}}^{1/s} \frac{1}{1 - q_{\text{lin}}^{1/s}} \max\{\|u^*\|_{A_{\text{fs}}}, \Delta_0^{0,0}\}^{1/s} (\Delta_k^{\ell,j})^{-1/s}
\]
Rearranging this estimate, we end up with
\[
\sup_{(\ell',k',j') \in Q} \left( \sum_{(\ell,k,j) \in Q} \#T_\ell \right)^{s} \Delta_k^{\ell,j'} \lesssim \max\{\|u^*\|_{A_{\text{fs}}}, \Delta_0^{0,0}\},
\]
where the hidden constant depends only on $C_{\text{stab}}, C_{\text{rel}}, C_{\text{mark}}, 1 - \lambda_{\text{Pic}} / \lambda_{\text{Pic}}^{\alpha}, C_{\text{Cea}} = L / \alpha, C_{\text{rel}}^{\alpha}, C_{\text{mesh}}, C_{\text{lin}}, q_{\text{lin}}, \#T_0$, and $s$. This proves the right inequality in (37). □

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 [Füh14 Section 7.4.1] which is an optimal preconditioner, i.e., the condition number of the preconditioned system is uniformly bounded; cf. also [GHPS19 Section 2.9].

March 4, 2021 29
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., $\Gamma = \Gamma_D \cup \Gamma_N$. While the numerical experiments in Section 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:

$$
- \text{div} (\mu(x, |\nabla u^*(x)|^2) \nabla u^*(x)) = f(x) \quad \text{in } \Omega,
$$

$$
u^*(x) = 0 \quad \text{on } \Gamma_D,
$$

$$
\mu(x, |\nabla u^*(x)|^2) \partial_n u^*(x) = g(x) \quad \text{on } \Gamma_N,
$$

where the scalar nonlinearity $\mu : \Omega \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ satisfies the following properties ($\text{M1}$–$\text{M4}$), similarly considered in [GMZ12, GHP18]:

- ($\text{M1}$) There exist constants $0 < \gamma_1 < \gamma_2 < \infty$ such that

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

- ($\text{M2}$) 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

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

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

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

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

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

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

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

(99)

With respect to the abstract framework of Section 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

$$
\langle Av, w \rangle_{\mathcal{X}^\prime \times \mathcal{X}} = \int_\Omega \mu(x, |\nabla w(x)|^2) \nabla w(x) \cdot \nabla v(x) \, dx,
$$

(100a)

$$
\langle F, v \rangle_{\mathcal{X}^\prime \times \mathcal{X}} = \int_\Omega f v \, dx + \int_{\Gamma_N} g v \, ds
$$

(100b)

for all $v, w \in \mathcal{X}$. We recall from [GHP18, Proposition 8.2] that ($\text{M1}$–$\text{M2}$) implies that $A$ is strongly monotone (with $\alpha := \tilde{\gamma}_1$) and Lipschitz continuous (with $L := \tilde{\gamma}_2$), so that (94) fits into the setting of Section 2.1. Moreover, ($\text{M3}$–$\text{M4}$) are required to prove the well-posedness and the properties (A1)–(A4) of the residual $a \text{ posteriori}$ 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 $T_H \in \mathcal{T}$, consider the lowest-order FEM space $X_H := \{ v \in C(\Omega) : v|_\Gamma = 0 \text{ and } v|_T \in P^1(T) \text{ for all } T \in \mathcal{T}_H \}$. (101)

As in [GMZ12, Section 3.2], we define for all $T \in \mathcal{T}_H$ and all $v_H \in X_H$, the corresponding weighted residual error indicators

$$\eta_H(T, v_H)^2 := |T|^{2/d} \| f + \text{div} (\mu(\cdot, |\nabla v_H|^2)\nabla v_H) \|_{L^2(T)}^2 + |T|^{1/d} \| [\mu(\cdot, |\nabla v_H|^2)\nabla v_H \cdot n] \|_{L^2(\partial T \cap \Omega)^*},$$

where $[\cdot]$ denotes the usual jump of discrete functions across element interfaces, and $n$ 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., [CKNS08] for the linear case or [GMZ12, Theorem 3.3] and [GMZ12, 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 [GMZ12, Lemma 3.7]. Using this inverse estimate, the proof of (A1) and (A2) follows as for the linear case; see, e.g., [CKNS08] for the linear case or [GMZ12, Section 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.

7.4. Experiment with known solution. We consider the Z-shaped domain $\Omega \subset \mathbb{R}^2$ from Figure 1 (left) with mixed boundary conditions and the nonlinear problem (94) with
μ(x, |∇u⋆(x)|^2) := 2 + \frac{1}{\sqrt{1 + |∇u⋆(x)|^2}}. This leads to the bounds α = 2 and L = 3 in (10). We prescribe the solution u⋆ in polar coordinates (x, y) = r(\cos φ, \sin φ) with φ ∈ (−π, π)

\[ u⋆(x, y) := r^β \cos(β \phi) \quad \text{with} \quad β = 4/7 \] (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).

In Figure 2, we compare uniform mesh-refinement (θ = 1) to adaptive mesh-refinement (0 < θ < 1) for different values of λalg and λPic. We plot the error estimator \( η(\ell u_κ) \) over the number of elements \( N := \#T_κ \). First (top), we fix \( θ = 0.5 \), λPic = 10^{-2}, and choose \( λalg \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\} \). We see that uniform mesh-refinement leads to the suboptimal rate of convergence \( O(N^{-2/7}) \), whereas Algorithm 1 with adaptive mesh-refinement regains the optimal rate of convergence \( O(N^{-1/2}) \), independently of the actual choice of λalg. We observe the very same if we fix \( θ = 0.5 \), λalg = 10^{-2}, and choose \( λPic \in \{1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\} \) (middle), or if we fix λalg = λPic = 10^{-2} and vary \( θ \in \{0.1, 0.3, 0.5, 0.7, 0.9\} \) (bottom). Since we know from Proposition 2 and the estimate

\[ \|u_k^{u⋆} - u_k^{u}\| ≲ \|u_k^{u} - u_k^{u-1}\| \equiv η(k^{u}\ell) \|u_k^{u} - u_k^{-1}\| \equiv η\ell(k^{u}\ell) \]

that \( η\ell(k^{u}\ell) ≈ Δu^{u}\ell \), this empirically underpins Theorem 4.

In Figure 3, we analogously choose different combinations of \( θ, λalg, \) and λPic. We plot the error estimator \( η\ell(k^{u} \ell \ell) \) over the cumulative sum \( \sum_{(\ell,k,j) ≤ (\ell',k',j')} \#T_\ell \). Independently of \( θ, λalg, \) and λPic, we observe optimal order of convergence \( O\left(\sum_{(\ell,k,j) ≤ (\ell',k',j')} \#T_\ell \right)^{-1/2} \) with respect to the overall computational cost, in accordance with Theorem 5.

In Figure 4, we also consider the total number of PCG iterations cumulated over all Picard steps on the given mesh for different combinations of \( θ, λalg, \) and λPic. 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 λalg and λPic, as well as for smaller values of θ, the total number of PCG iterations is smaller.

7.5. Experiment with unknown solution. We consider the L-shaped domain Ω ⊂ \( \mathbb{R}^2 \) from Figure 1 (right) and the nonlinear problem (94) with f(x) = 1 and μ(x, |∇u⋆(x)|^2) := 1 + \frac{\ln(1 + |∇u⋆(x)|^2)}{1 + |∇u⋆(x)|^2}. Then, (M1)–(M4) hold with α ≈ 0.9582989 and L ≈ 1.5423438.

In Figure 5, we again test Algorithm 1 with varying \( θ, λalg, \) and λPic. We plot the error estimator \( η\ell(u_{\ell}^{u\ell}) \) over the number of elements \( N := \#T_\ell \). Uniform mesh-refinement leads to the suboptimal rate of convergence \( O(N^{-1/3}) \), whereas Algorithm 1 regains the optimal rate of convergence \( O(N^{-1/2}) \). Again, this empirically confirms Theorem 4. The latter rate of convergence even appears to be robust with respect to θ, λalg, and λPic.

In Figure 6, we plot the estimator \( η\ell(u_{\ell}^{u\ell}) \) over the cumulative sum \( \sum_{(\ell,k,j) ≤ (\ell',k',j')} \#T_\ell \). Independently of the choice of the parameters \( θ, λalg, \) and λPic, we observe the optimal order of convergence \( O\left(\sum_{(\ell,k,j) ≤ (\ell',k',j')} \#T_\ell \right)^{-1/2} \) with respect to the overall computational cost, which empirically underpins Theorem 5.

In Figure 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 θ, λalg,
number of elements $N$ of the mesh $\mathcal{T}_\ell$ for various parameters $\theta$, $\lambda_{\text{Pic}}$, and $\lambda_{\text{alg}}$.

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.

**References**

[BCL15] Roland Becker, Daniela Capatina, and Robert Luce. Stopping criteria based on locally reconstructed fluxes. In *Numerical mathematics and advanced applications—ENUMATH 2013*, volume 103 of *Lect. Notes Comput. Sci. Eng.*, pages 243–251. Springer, Cham, 2015.

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

March 4, 2021
Figure 3. Optimal decay rate wrt. overall computational cost, example from Section 7.4. Error estimator \( \eta_{\ell}(u^{(\ell)'}) \) on mesh \( T_{\ell}' \), perturbed Banach–Picard iteration \( k' \), and PCG step \( j' \) of Algorithm 1 with respect to the overall cost expressed as the cumulative sum \( \sum_{(\ell,k,j) \leq (\ell',k',j')} #T_{\ell} \) for various parameters \( \theta \), \( \lambda_{\text{Pic}} \), and \( \lambda_{\text{alg}} \).
Figure 4. Number of PCG iterations wrt. the number of elements \( N := \#T_\ell \), example from Section 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).

[CDD03] Albert Cohen, Wolfgang Dahmen, and Ronald Devore. Adaptive wavelet schemes for non-linear variational problems. *SIAM J. Numer. Anal.*, 41(5):1785–1823, 2003.

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

[CG12] Carsten Carstensen and Joscha Gedicke. An adaptive finite element eigenvalue solver of asymptotic quasi-optimal computational complexity. *SIAM J. Numer. Anal.*, 50(3):1029–1057, 2012.

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

[CPV14] Clément Cancès, Iuliu Sorin Pop, and Martin Vohralík. An a posteriori error estimate for vertex-centered finite volume discretizations of immiscible incompressible two-phase flow. *Math. Comp.*, 83(285):153–188, 2014.
Figure 5. Optimal decay rate w.r.t. degrees of freedom, example from Section 7.5. Error estimator $\eta(u_j^{k_{\ell}})$ on mesh $T_{\ell}$, perturbed Banach–Picard iteration $k_{\ell}$, and PCG step $j$ of Algorithm 1 with respect to the number of elements $N$ of the mesh $T_{\ell}$ for various parameters $\theta$, $\lambda_{\text{Pic}}$, and $\lambda_{\text{alg}}$.

[CS07] Alexandra L. Chailou and Manil Suri. A posteriori estimation of the linearization error for strongly monotone nonlinear operators. J. Comput. Appl. Math., 205(1):72–87, 2007.
[CW17] Scott Congreve and Thomas P. Wihler. Iterative Galerkin discretizations for strongly monotone nonlinear operators. J. Comput. Appl. Math., 311:457–472, 2017.
[Deu91] Peter Deuflhard. Global inexact Newton methods for very large scale nonlinear problems. Impact Comput. Sci. Engrg., 3(4):366–393, 1991.
[DFFGP19] Giovanni Di Fratta, Thomas Führer, Gregor Gantner, and Dirk Praetorius. Adaptive Uzawa algorithm for the Stokes equation. ESAIM Math. Model. Numer. Anal., 53(6):1841–1870, 2019.
[DK08] Lars Diening and Christian Kreuzer. Linear convergence of an adaptive finite element method for the $p$-Laplacian equation. SIAM J. Numer. Anal., 46(2):614–638, 2008.
Figure 6. Optimal decay rate wrt. overall computational cost, example from Section 7.5 Error estimator $\eta^e(y^{k,f}, f^e)$ on mesh $T^e$, perturbed Banach–Picard iteration $k^f$, and PCG step $j$ of Algorithm 1 with respect to the overall cost expressed as the cumulative sum $\sum_{(\ell,k,j) \leq (\ell',k',j')} T_\ell$ for various parameters $\theta$, $\lambda_{\text{Pic}}$, and $\lambda_{\text{alg}}$.

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

[DPVY15] Daniele A. Di Pietro, Martin Vohralík, and Soleiman Yousef. Adaptive regularization, linearization, and discretization and a posteriori error control for the two-phase Stefan problem. Math. Comp., 84(291):153–186, 2015.

[EAEV11] Linda El Alaoui, Alexandre Ern, and Martin Vohralík. 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.
Figure 7. Number of PCG iterations wrt. number of elements $N := \#T_L$, example from Section 7.5 for $\theta = 0.5$, $\lambda_{PC} = 10^{-2}$ (top), for $\theta = 0.5$, $\lambda_{\text{alg}} = 10^{-2}$ (middle), and for $\lambda_{\text{alg}} = \lambda_{PC} = 10^{-2}$ (bottom).

[EV13] Alexandre Ern and Martin Vohralík. Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs. *SIAM J. Sci. Comput.*, 35(4):A1761–A1791, 2013.

[EW94] Stanley C. Eisenstat and Homer F. Walker. Globally convergent inexact Newton methods. *SIAM J. Optim.*, 4(2):393–422, 1994.

[Füh14] Thomas Führer. *Zur Kopplung von finiten Elementen und Randelementen*. PhD thesis, TU Wien, 2014.

[GHPS18] Gregor Gantner, Alexander Haberl, Dirk Praetorius, and Bernhard Stiftner. Rate optimal adaptive FEM with inexact solver for nonlinear operators. *IMA J. Numer. Anal.*, 38(4):1797–1831, 2018.

[GHPS19] Gregor Gantner, Alexander Haberl, Dirk Praetorius, and Stefan Schimanko. Rate optimality of adaptive finite element methods with respect to overall cost. Submitted, 2019.
[GMZ11] Eduardo M. Garau, Pedro Morin, and Carlos Zuppa. Convergence of an adaptive Kačanov FEM for quasi-linear problems. *Appl. Numer. Math.*, 61(4):512–529, 2011.

[GMZ12] Eduardo M. Garau, Pedro Morin, and Carlos Zuppa. Quasi-optimal convergence rate of an AFEM for quasi-linear problems of monotone type. *Numer. Math. Theory Methods Appl.*, 5(2):131–156, 2012.

[GSS14] Dietmar Gallistl, Mira Schedensack, and Rob P. Stevenson. A remark on newest vertex bisection in any space dimension. *Comput. Methods Appl. Math.*, 14(3):317–320, 2014.

[HW18] Paul Houston and Thomas P. Wihler. An $hp$-adaptive Newton-discontinuous-Galerkin finite element approach for semilinear elliptic boundary value problems. *Math. Comp.*, 87(314):2641–2674, 2018.

[HW20a] Pascal Heid and Thomas P. Wihler. Adaptive iterative linearization Galerkin methods for nonlinear problems. *Math. Comp.*, 89:2707–2734, 2020.

[HW20b] Pascal Heid and Thomas P. Wihler. On the convergence of adaptive iterative linearized Galerkin methods. *Calcolo*, 57:24, 2020.

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

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

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

[OT14] Maxim A. Olshanskii and Eugene E. Tyrtyshnikov. *Iterative methods for linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2014.

[Pol16] Sara Pollock. Stabilized and inexact adaptive methods for capturing internal layers in quasilinear PDE. *J. Comput. Appl. Math.*, 308:243–262, 2016.

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

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

[Ste14] Rob Stevenson. Adaptive wavelet methods for linear and nonlinear least-squares problems. *Found. Comput. Math.*, 14(2):237–283, 2014.

[Vee02] Andreas Veeser. Convergent adaptive finite elements for the nonlinear Laplacian. *Numer. Math.*, 92(4):743–770, 2002.

[Zei90] Eberhard Zeidler. *Nonlinear functional analysis and its applications. II/B*. Springer-Verlag, New York, 1990.