



Philipp Bringmann, Maximilian Brunner, Dirk Praetorius and Julian Streitberger\*

# Optimal complexity of goal-oriented adaptive FEM for nonsymmetric linear elliptic PDEs

<https://doi.org/10.1515/jnma-2023-0150>

Received December 1, 2023; accepted July 6, 2024; published online November 4, 2024

**Abstract:** We analyze a goal-oriented adaptive algorithm that aims to efficiently compute the quantity of interest  $G(u^*)$  with a linear goal functional  $G$  and the solution  $u^*$  to a general second-order nonsymmetric linear elliptic partial differential equation. The current state of the analysis of iterative algebraic solvers for nonsymmetric systems lacks the contraction property in the norms that are prescribed by the functional analytic setting. This seemingly prevents their application in the optimality analysis of goal-oriented adaptivity. As a remedy, this paper proposes a goal-oriented adaptive iteratively symmetrized finite element method (GOAIS-FEM). It employs a nested loop with a contractive symmetrization procedure, e.g., the Zangantello iteration, and a contractive algebraic solver, e.g., an optimal multigrid solver. The various iterative procedures require well-designed stopping criteria such that the adaptive algorithm can effectively steer the local mesh refinement and the computation of the inexact discrete approximations. The main results consist of full linear convergence of the proposed adaptive algorithm and the proof of optimal convergence rates with respect to both degrees of freedom and total computational cost (i.e., optimal complexity). Numerical experiments confirm the theoretical results and investigate the selection of the parameters.

**Keywords:** goal-oriented adaptive finite element method; linear quantity of interest; iterative solver; nonsymmetric partial differential equations; optimal convergence rates; optimal complexity

**MSC 2010 Classification:** 41A25; 65N15; 65N30; 65N50; 65Y20

## 1 Introduction

Adaptive finite element methods (AFEMs) are a cornerstone in the numerical solution of partial differential equations (PDEs). The abundant literature emphasizes significant progress and manifests a matured understanding of the topic; see, e.g. [1]–[9], for linear elliptic PDEs.

The variational formulation of a *nonsymmetric* second-order linear elliptic PDE with bilinear form  $b(\cdot, \cdot)$  and right-hand side functional  $F$  on the Sobolev space  $\mathcal{X} := H_0^1(\Omega)$  seeks a weak solution  $u^*$  to

$$b(u^*, v) = F(v) \quad \forall v \in \mathcal{X}. \quad (1)$$

While standard AFEM aims at an efficient approximation of the solution  $u^* \in \mathcal{X}$ , goal-oriented AFEM (GOAFEM) strives only to approximate a *quantity of interest*  $G(u^*)$ ; see [10]–[13] for early prominent contributions. However, to accurately approximate  $G(u^*)$  for a continuous linear *goal functional*  $G: \mathcal{X} \rightarrow \mathbb{R}$ , following the generic approach  $G(u_H) \approx G(u^*)$  leads to convergence rates determined by the error of the approximation  $u_H \approx u^*$  to

\*Corresponding author: Julian Streitberger, TU Wien, Institute of Analysis and Scientific Computing, Wien, Austria,

E-mail: julian.streitberger@asc.tuwien.ac.at. <https://orcid.org/0000-0003-1189-0611>

Philipp Bringmann, Maximilian Brunner and Dirk Praetorius, TU Wien, Institute of Analysis and Scientific Computing, Wien, Austria,

E-mail: philipp.bringmann@asc.tuwien.ac.at (P. Bringmann), maximilian.brunner@asc.tuwien.ac.at (M. Brunner),

dirk.praetorius@asc.tuwien.ac.at (D. Praetorius). <https://orcid.org/0000-0002-4546-5165> (P. Bringmann). <https://orcid.org/0000-0003-0636-1491> (M. Brunner). <https://orcid.org/0000-0002-1977-9830> (D. Praetorius)

Open Access. © 2024 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

the *primal problem* (1). Instead, GOAFEM adopts a duality technique by additionally approximating  $z_H \approx z^* \in \mathcal{X}$  solving the *dual problem*

$$b(v, z^*) = G(v) \quad \forall v \in \mathcal{X}. \quad (2)$$

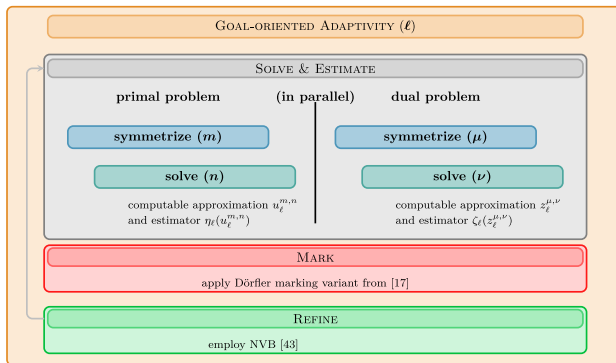
Following [13], a discrete approximation  $G_H(u_H, z_H) \approx G(u^*)$  enables the control of the error for any  $u_H, z_H \in \mathcal{X}$  by

$$|G(u^*) - G_H(u_H, z_H)| \leq |b(u^* - u_H, z^* - z_H)| \leq L \|u^* - u_H\| \|z^* - z_H\|, \quad (3)$$

where  $L > 0$  is the continuity constant of  $b(\cdot, \cdot)$  with respect to the energy norm  $\|\cdot\|$ ; see Section 2 for details. As seen in (3), this approach allows to add the convergence rates of the primal and dual problem. Moreover, it is not necessary – and may even lead to unnecessary computational expense – to compute approximations  $u_H \approx u^*$  and  $z_H \approx z^*$  across the entire domain with the same accuracy. Instead, a careful marking of elements for refinement enables a considerable reduction of the computational costs and makes GOAFEM highly relevant in both practical applications and mathematical research.

First rigorous convergence results of GOAFEM are found in [14]–[18], recent contributions in this context include [19], [20] and for a dual weighted-residual approach see, e.g., [21]–[23]. The works [14], [16], [17], [19], [20] focus on optimal convergence rates with respect to the degrees of freedom. However, the cumulative nature of adaptivity calls for optimal convergence rates with respect to the total computational effort, i.e., the overall computational time. Coined as *optimal complexity* initially for wavelet-based discretizations [24], [25], this notion was later adopted for AFEM with contributions including, e.g., [4], [26]–[28]. In the setting of GOAFEM, optimal complexity was established first in [14] for the Poisson problem and sufficiently small adaptivity parameters, and extended to a general second-order *symmetric* linear elliptic PDE with uniformly contractive algebraic solver in [29]. Since uniform contraction with respect to the PDE-related energy norm for nonsymmetric algebraic solvers such as GMRES is still open, as a remedy, the proof of the Lax–Milgram lemma motivates the application of an iterative symmetrization [28]. This results in a sequence of symmetric algebraic systems that allow the application of optimal algebraic solvers, e.g., [30]–[32]. Figure 1 illustrates the nested structure of the resulting goal-oriented adaptive iteratively symmetrized finite element method (GOAISFEM). The detailed Algorithm 1 is presented in Section 3 below. Table 1 displays the notation of the associated indices and quasi-error quantities, which are equivalent to the total error.

The first challenge in the analysis of the GOAISFEM algorithm consists of the nonlinear product structure attained by the combined quasi-error product as displayed in Table 1. The resulting nonlinear remainder term significantly complicates the proof compared to treating only the primal problem as in [28] and requires the application of a novel proof strategy from [33] that only utilizes summability of the remainder, denoted as *tail-summability* throughout. The second challenge arises from the combination of the primal and dual marking leading to a merged marked set. Thereby, either only the primal or only the dual estimator is guaranteed to satisfy the estimator reduction property. Since the estimator belongs to the quasi-error, this also leads to a failure of contraction for one of the two involved quasi-errors. While [29] solves this issue in the symmetric case, the additional symmetrization loop results in a more involved situation at hand. Adapting the novel approach of



**Figure 1:** Schematic overview of the GOAISFEM algorithm with nested symmetrization and inexact solver.

**Table 1:** Iteration counters and quasi-errors for the GOAISFEM algorithm. We note that for the combination of the index sets, the quasi-errors are extended to the full index set by the last available quasi-error. We refer to Section 3 for details on the iteration counters and index sets and to the beginning of Section 5 for a detailed description of the quasi-errors and their extension to the full index set  $Q$ .

Iteration	Mesh refinement		Symmetrization		Algebraic solver		Index set	Quasi-error
	Running	Final	Running	Final	Running	Final		
Primal	$\ell$	$\underline{\ell}$	$m$	$\underline{m}$	$n$	$\underline{n}$	$Q^u$ in (24a)	$H_\ell^{m,n}$ in (44a)
Dual	$\ell$	$\underline{\ell}$	$\mu$	$\underline{\mu}$	$\nu$	$\underline{\nu}$	$Q^z$ in (24b)	$Z_\ell^{\mu,\nu}$ in (44b)
Combined	$\ell$	$\underline{\ell}$	$k$	$\underline{k} = \max\{\underline{m}, \underline{\mu}\}$	$j$	$\underline{j} = \max\{\underline{n}, \underline{\nu}\}$	$Q = Q^u \cup Q^z$	$H_\ell^{k,j} Z_\ell^{k,j}$ in (45)

the tail-summability criterion from [33] enables the proof of full linear convergence and optimal complexity for the nonlinear quasi-error product in this paper. The analysis employs the generalized quasi-orthogonality from [34] to remedy the lack of a Pythagorean identity for nonsymmetric problems.

Our main result asserts full linear convergence of the quasi-error product  $H_\ell^{k,j} Z_\ell^{k,j}$  with respect to the total step counter  $|\cdot, \cdot, \cdot|$  (measuring the total solver steps in the index set). Therein, we allow for an arbitrary symmetrization stopping parameter  $\lambda_{\text{sym}}$  and only require a small algebraic solver parameter  $\lambda_{\text{alg}}$  such that the product  $\lambda_{\text{sym}}\lambda_{\text{alg}}$  is sufficiently small. More precisely, Theorem 1 states that there exist constants  $C_{\text{lin}} > 0$  and  $0 < q_{\text{lin}} < 1$  such that, for all  $(\ell, k, j), (\ell', k', j') \in Q$  with  $|\ell', k', j'| \leq |\ell, k, j|$ ,

$$H_\ell^{k,j} Z_\ell^{k,j} \leq C_{\text{lin}} q_{\text{lin}}^{|\ell,k,j| - |\ell',k',j'|} H_{\ell'}^{k',j'} Z_{\ell'}^{k',j'}.$$

Note that, unlike [28], where full linear convergence is guaranteed only for sufficiently large  $\ell \geq \ell_0$ , the current result is stronger in the sense that the result holds for  $\ell_0 = 0$  owing to a generalized quasi-orthogonality from [34]. An immediate consequence of full linear convergence and the geometric series in Corollary 1 states that the rates with respect to the degrees of freedom coincide with the rates with respect to the cumulative computational work (i.e., computational time), i.e., for all  $r > 0$ , there holds

$$\sup_{(\ell,k,j) \in Q} (\#\mathcal{T}_\ell)^r H_\ell^{k,j} Z_\ell^{k,j} \leq \sup_{(\ell,k,j) \in Q} \left( \sum_{\substack{(\ell',k',j') \in Q \\ |\ell',k',j'| \leq |\ell,k,j|}} \#\mathcal{T}_{\ell'} \right)^r H_\ell^{k,j} Z_\ell^{k,j} \leq C_{\text{cost}} \sup_{(\ell,k,j) \in Q} (\#\mathcal{T}_\ell)^r H_\ell^{k,j} Z_\ell^{k,j}$$

along the sequence of meshes  $\mathcal{T}_\ell$  generated by the GOAISFEM algorithm. The second main result of Theorem 2 proves that, for sufficiently small adaptivity parameters and any achievable rates  $s, t > 0$  of the primal resp. dual problem (stated in terms of nonlinear approximation classes), the algorithm guarantees optimal complexity, i.e.,

$$\sup_{(\ell,k,j) \in Q} \left( \sum_{\substack{(\ell',k',j') \in Q \\ |\ell',k',j'| \leq |\ell,k,j|}} \#\mathcal{T}_{\ell'} \right)^{s+t} H_\ell^{k,j} Z_\ell^{k,j} \leq C_{\text{opt}} \max \left\{ \|u^*\|_{\mathbb{A}_s}, \|z^*\|_{\mathbb{A}_t}, H_0^{0,0} Z_0^{0,0} \right\}.$$

This means the convergence of the algorithm attains the optimal rate  $s + t$  with respect to the overall computational work, where  $\|u^*\|_{\mathbb{A}_s} < \infty$  means that  $u^*$  can be approximated at rate  $s$  (along a sequence of unavailable optimal meshes) and likewise for  $z^*$ .

The remaining parts of the paper are organized as follows. The preliminary Section 2 introduces the model problem, the assumptions on the solvers, and the axioms of adaptivity from [9], including the general quasi-orthogonality from [34]. Following the algorithm in Section 3 and its contraction properties in Sections 4 and 5 presents full linear convergence as the first main result of this paper. This allows to prove optimal complexity in Section 6 as the second main result, which is underlined by the numerical experiments in Section 7 including a thorough investigation of the adaptivity parameters. The paper concludes with a summary in Section 8.

## 2 Setting

In this section, we introduce the problem and explain the key components needed to design the adaptive algorithm in Section 3.

### 2.1 Continuous model problem

Let  $\Omega \subset \mathbb{R}^d$  with  $d \geq 1$  be a polygonal Lipschitz domain. Given right-hand sides  $f \in L^2(\Omega)$  and  $\mathbf{f} \in [L^2(\Omega)]^d$ , we consider a general second-order linear elliptic PDE

$$-\operatorname{div}(A\nabla u^*) + \mathbf{b} \cdot \nabla u^* + c u^* = f - \operatorname{div}(\mathbf{f}) \quad \text{in } \Omega \quad \text{subject to } u^* = 0 \quad \text{on } \partial\Omega, \quad (4)$$

with a pointwise symmetric and positive definite diffusion matrix  $A \in [L^\infty(\Omega)]_{\text{sym}}^{d \times d}$ , a convection coefficient  $\mathbf{b} \in [L^\infty(\Omega)]^d$ , and a reaction coefficient  $c \in L^\infty(\Omega)$ . For well-definedness of the *a posteriori* error estimator in Section 2.6 below, we additionally require that  $A|_T \in [W^{1,\infty}(T)]_{\text{sym}}^{d \times d}$  and  $\mathbf{f}|_T \in [H^1(T)]^d$  for all  $T \in \mathcal{T}_0$ , where  $\mathcal{T}_0$  is an initial triangulation that subdivides  $\Omega$  into compact simplices. Let  $\langle \cdot, \cdot \rangle$  denote the  $L^2(\Omega)$ -scalar product. With the principal part  $a(u, v) := \langle A\nabla u, \nabla v \rangle$ , the variational formulation of (4) seeks a solution  $u^* \in \mathcal{X} := H_0^1(\Omega)$  to the so-called *primal problem*

$$b(u^*, v) := a(u^*, v) + \langle \mathbf{b} \cdot \nabla u^* + c u^*, v \rangle = \langle f, v \rangle + \langle \mathbf{f}, \nabla v \rangle =: F(v) \quad \forall v \in \mathcal{X}. \quad (5)$$

We suppose that the bilinear form  $b(\cdot, \cdot)$  from (5) is continuous and elliptic with respect to the norm  $\|\cdot\|_{\mathcal{X}}$  on  $\mathcal{X}$ , i.e., there exist constants  $L', \alpha' > 0$  such that

$$b(u, v) \leq L' \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}, \quad b(v, v) \geq \alpha' \|v\|_{\mathcal{X}}^2 \quad \forall u, v \in \mathcal{X}. \quad (6)$$

Then, the Lax–Milgram lemma proves existence and uniqueness of the solution  $u^*$  to (5). An elementary compactness argument shows that (6) implies ellipticity of the principal part  $a(\cdot, \cdot)$  and thus  $a(\cdot, \cdot)$  is a scalar product on  $\mathcal{X}$  with induced energy norm  $a(\cdot, \cdot)^{1/2} =: \|\cdot\| \simeq \|\cdot\|_{\mathcal{X}}$ , cf. [35, Remark 3]. Therefore,  $b(\cdot, \cdot)$  is also continuous and elliptic with respect to  $\|\cdot\|$ , i.e., there exist constants  $L, \alpha > 0$  such that

$$b(u, v) \leq L \|\cdot\| \|\cdot\|, \quad b(v, v) \geq \alpha \|\cdot\|^2 \quad \forall u, v \in \mathcal{X}. \quad (7)$$

In the present paper, we suppose that the *quantity of interest*  $G$  is linear and reads for given data  $g \in L^2(\Omega)$  and  $\mathbf{g} \in [L^2(\Omega)]^d$ ,

$$G(v) := \int_{\Omega} (g v + \mathbf{g} \cdot \nabla v) dx.$$

In order to guarantee well-definedness of the error estimator in Section 2.6 below, we suppose  $\mathbf{g}|_T \in [H^1(T)]^d$  for all initial simplices  $T \in \mathcal{T}_0$ . In view of the continuity and coercivity of  $b(\cdot, \cdot)$ , the Lax–Milgram lemma yields existence and uniqueness of the solution  $z^* \in \mathcal{X}$  of the so-called *dual problem*: Find  $z^* \in \mathcal{X}$  such that

$$b(v, z^*) = G(v) \quad \forall v \in \mathcal{X}. \quad (8)$$

### 2.2 Finite element discretization and discrete goal

For a polynomial degree  $p \in \mathbb{N}$  and a conforming simplicial triangulation  $\mathcal{T}_H$  of  $\Omega$ , the discrete ansatz space reads

$$\mathcal{X}_H := \{v_H \in \mathcal{X} : \forall T \in \mathcal{T}_H, v_H|_T \text{ is a polynomial of total degree } \leq p\}. \quad (9)$$

Since  $\mathcal{X}_H \subset \mathcal{X}$  is conforming, the Lax–Milgram lemma ensures the existence and uniqueness of primal and dual discrete solutions  $u_H^*, z_H^* \in \mathcal{X}_H$  satisfying

$$b(u_H^*, v_H) = F(v_H), \quad b(v_H, z_H^*) = G(v_H) \quad \forall v_H \in \mathcal{X}_H. \quad (10)$$

It is well-known that conforming FEMs are quasi-optimal, i.e., there hold Céa-type estimates with constant  $C_{\text{Céa}} = L/\alpha$

$$\|u^* - u_H^*\| \leq C_{\text{Céa}} \min_{v_H \in \mathcal{X}_H} \|u^* - v_H\|, \quad \|z^* - z_H^*\| \leq C_{\text{Céa}} \min_{v_H \in \mathcal{X}_H} \|z^* - v_H\|. \quad (11)$$

For arbitrary approximations  $u_H, z_H \in \mathcal{X}_H$  the linearity of the quantity of interest  $G$  as well as the primal and the dual problem (1) and (2) show that

$$\begin{aligned} G(u^*) - G(u_H) &= G(u^* - u_H) \stackrel{(2)}{=} b(u^* - u_H, z^*) \\ &\stackrel{(1)}{=} b(u^* - u_H, z^* - z_H) + [F(z_H) - b(u_H, z_H)]. \end{aligned}$$

The definition of the discrete goal quantity by  $G_H(u_H, z_H) := G(u_H) + [F(z_H) - b(u_H, z_H)]$  allows to control the goal error by continuity of  $b(\cdot, \cdot)$

$$|G(u^*) - G_H(u_H, z_H)| \leq |b(u^* - u_H, z^* - z_H)| \leq L \|u^* - u_H\| \|z^* - z_H\|. \quad (12)$$

We emphasize that (12) holds for any  $u_H, z_H$  and, in particular, for those stemming from an iterative solution step. Moreover, if  $u_H = u_H^*$ , then  $G(u_H, z_H) = G(u_H^*)$  as expected.

### 2.3 Zarantonello iteration

The discrete formulations (10) lead to positive definite, but *nonsymmetric* linear systems of equations. To reduce the formulation to symmetric and positive definite (SPD) problems, we follow previous own work [28] for the primal problem and employ the Zarantonello iteration [36]. Typically, the latter is used in the up-to-date proof of the Lax–Milgram lemma and also defines a linearization scheme for the treatment of a certain class of nonlinear elliptic PDEs (see, e.g., [33], [37]–[39]). In its core, it is a fixed-point method, thus also applicable in the nonsymmetric setting at hand. For a damping parameter  $\delta > 0$  and given  $u_H, z_H \in \mathcal{X}_H$ , the Zarantonello iterations  $\Phi_H^u, \Phi_H^z: (0, \infty) \times \mathcal{X}_H \rightarrow \mathcal{X}_H$  compute the unique solutions  $\Phi_H^u(\delta; u_H), \Phi_H^z(\delta; z_H) \in \mathcal{X}_H$  to the symmetric variational formulations

$$a(\Phi_H^u(\delta; u_H), v_H) = a(u_H, v_H) + \delta [F(v_H) - b(u_H, v_H)] \quad \forall v_H \in \mathcal{X}_H, \quad (13a)$$

$$a(v_H, \Phi_H^z(\delta; z_H)) = a(v_H, z_H) + \delta [G(v_H) - b(v_H, z_H)] \quad \forall v_H \in \mathcal{X}_H. \quad (13b)$$

The Riesz–Fischer theorem (and also the Lax–Milgram lemma) guarantees existence and uniqueness of  $\Phi_H^u(\delta; u_H), \Phi_H^z(\delta; z_H) \in \mathcal{X}_H$ , i.e., the Zarantonello operators  $\Phi_H^u(\delta; \cdot)$  and  $\Phi_H^z(\delta; \cdot)$  are well-defined. In particular, the exact discrete solutions  $u_H^* = \Phi_H^u(\delta; u_H^*)$  and  $z_H^* = \Phi_H^z(\delta; z_H^*)$  are the unique fixed points for all  $\delta > 0$ . Moreover, for a sufficiently small damping parameter  $\delta$ , i.e.,  $0 < \delta < \delta^* := 2\alpha/L^2$ , the Banach fixed-point theorem [40, Section 25.4] guarantees that  $\Phi_H^u(\delta; \cdot)$  and  $\Phi_H^z(\delta; \cdot)$  are contractive with constant  $0 < q_{\text{sym}}^* := [1 - \delta(2\alpha - \delta L^2)]^{1/2} < 1$ , i.e., for all functions  $v_H, w_H \in \mathcal{X}_H$ , it holds that

$$\max\{\|\Phi_H^u(\delta; v_H) - \Phi_H^u(\delta; w_H)\|, \|\Phi_H^z(\delta; v_H) - \Phi_H^z(\delta; w_H)\|\} \leq q_{\text{sym}}^* \|v_H - w_H\|. \quad (14)$$

The optimal value  $\delta_{\text{opt}} = \alpha/L^2$  yields the minimal contraction value  $q_{\text{sym}}^* = 1 - \alpha^2/L^2$ .

### 2.4 Algebraic solver

A canonical candidate for solving (10) directly is a generalized minimal residual method [41], [42] with optimal preconditioner for the symmetric part. While this guarantees uniform contraction of the algebraic residuals in a *discrete vector norm*, the link between the algebraic residuals and the functional setting is still open [28]. Instead, after a symmetrization with the Zarantonello iteration, it remains to solve the SPD systems (13). Since large SPD problems are still computationally expensive and the exact solution cannot be computed in linear computational complexity, we employ an iterative algebraic solver whose iteration is expressed by the operator

$\Psi_H: \mathcal{X}' \times \mathcal{X}_H \rightarrow \mathcal{X}_H$ . More precisely, given a bounded linear functional  $\psi \in \mathcal{X}'$  and an approximation  $w_H \in \mathcal{X}_H$  of the exact solution  $w_H^* \in \mathcal{X}_H$  to  $a(w_H^*, v_H) = \psi(v_H)$  for all  $v_H \in \mathcal{X}_H$ , the algebraic solver returns an improved approximation  $\Psi_H(\psi; w_H) \in \mathcal{X}_H$  in the sense that there exists  $0 < q_{\text{alg}} < 1$  independent of  $\psi$  and  $\mathcal{X}_H$  such that

$$\|w_H^* - \Psi_H(\psi; w_H)\| \leq q_{\text{alg}} \|w_H^* - w_H\| \quad \forall w_H \in \mathcal{X}_H. \quad (15)$$

To simplify notation, we shall identify  $\psi$  with its Riesz representative  $w_H^* \in \mathcal{X}_H$  and write  $\Psi_H(w_H^*; \cdot)$  instead of  $\Psi_H(\psi; \cdot)$ , even though  $w_H^*$  is unknown in practice and will only be approximated by an optimal algebraic solver, e.g., [30]–[32]. In the following, we use the *hp*-robust multigrid method from [32] with localized lowest-order smoothing on intermediate levels and patchwise higher-order smoothing on the finest mesh as an innermost algebraic solver loop.

## 2.5 Mesh refinement

The mesh refinement employs newest-vertex bisection (NVB). We refer to [43] for NVB with admissible initial triangulation  $\mathcal{T}_0$  and  $d \geq 2$ , to [44], [45] for NVB with general  $\mathcal{T}_0$  for  $d \in \{1, 2\}$ , and to the recent work [46] for NVB with general  $\mathcal{T}_0$  in any dimension  $d \geq 2$ . For each triangulation  $\mathcal{T}_H$  and marked elements  $\mathcal{M}_H \subseteq \mathcal{T}_H$ , let  $\mathcal{T}_h := \text{refine}(\mathcal{T}_H, \mathcal{M}_H)$  be the coarsest conforming refinement of  $\mathcal{T}_H$  such that at least all  $T \in \mathcal{M}_H$  have been refined, i.e.,  $\mathcal{M}_H \subseteq \mathcal{T}_h \setminus \mathcal{T}_H$ . We write  $\mathcal{T}_h \in \mathbb{T}(\mathcal{T}_H)$  if  $\mathcal{T}_h$  can be obtained from  $\mathcal{T}_H$  by finitely many steps of NVB, and  $\mathcal{T}_h \in \mathbb{T}_N(\mathcal{T}_H)$  if  $\mathcal{T}_h \in \mathbb{T}(\mathcal{T}_H)$  with  $\#\mathcal{T}_h - \#\mathcal{T}_H \leq N$  with the number of additional elements  $N \in \mathbb{N}_0$ . To simplify notation, we write  $\mathbb{T} := \mathbb{T}(\mathcal{T}_0)$  and  $\mathbb{T}_N := \mathbb{T}_N(\mathcal{T}_0)$ . We note that the nestedness of meshes  $\mathcal{T}_h \in \mathbb{T}(\mathcal{T}_H)$  implies nestedness of the corresponding finite element spaces  $\mathcal{X}_h \subseteq \mathcal{X}_H \subseteq \mathcal{X}$  from (9).

## 2.6 A posteriori error estimation

For a triangle  $T \in \mathcal{T}_H \in \mathbb{T}$  and  $v_H \in \mathcal{X}_H$ , let  $\mathbf{n}$  denote the outer unit normal vector and  $[[\cdot]]$  the jump along inner edges of  $\mathcal{T}_H$ . We define the refinement indicators  $\eta_H(T; v_H) \geq 0$  and  $\zeta_H(T; v_H) \geq 0$  for the primal and dual problem from (10), respectively, by

$$\begin{aligned} \eta_H(T; v_H)^2 &:= |T|^{2/d} \| -\text{div}(\mathbf{A}\nabla v_H - \mathbf{f}) + \mathbf{b} \cdot \nabla v_H + c v_H - f \|_{L^2(T)}^2 \\ &\quad + |T|^{1/d} \| [[(\mathbf{A}\nabla v_H - \mathbf{f}) \cdot \mathbf{n}]] \|_{L^2(\partial T \cap \Omega)}^2, \\ \zeta_H(T; v_H)^2 &:= |T|^{2/d} \| -\text{div}(\mathbf{A}\nabla v_H - \mathbf{g}) - \mathbf{b} \cdot \nabla v_H + (c - \text{div}(\mathbf{b})) v_H - g \|_{L^2(T)}^2 \\ &\quad + |T|^{1/d} \| [[(\mathbf{A}\nabla v_H - \mathbf{g}) \cdot \mathbf{n}]] \|_{L^2(\partial T \cap \Omega)}^2. \end{aligned} \quad (16a)$$

For any subset  $\mathcal{U}_H \subseteq \mathcal{T}_H$ , we abbreviate

$$\eta_H(\mathcal{U}_H; v_H)^2 := \sum_{T \in \mathcal{U}_H} \eta_H(T; v_H)^2, \quad \zeta_H(\mathcal{U}_H; v_H)^2 := \sum_{T \in \mathcal{U}_H} \zeta_H(T; v_H)^2 \quad (16b)$$

as well as  $\eta_H(v_H) := \eta_H(\mathcal{T}_H; v_H)$  and  $\zeta_H(v_H) := \zeta_H(\mathcal{T}_H; v_H)$  for all  $v_H \in \mathcal{X}_H$ .

For details on residual-based error estimators, we refer to [47], [48]. Throughout the paper, the index of the estimators refer to the underlying mesh, e.g.,  $\eta_h$  and  $\zeta_h$  on the refinement  $\mathcal{T}_h \in \mathbb{T}(\mathcal{T}_H)$  or  $\eta_\ell$  and  $\zeta_\ell$  on a sequence of meshes  $\mathcal{T}_\ell$  with  $\ell \in \mathbb{N}_0$ . It is well-known that  $\eta_H, \zeta_H$  satisfy the following *axioms of adaptivity*.

**Lemma 1** (see [9, Section 6.1]). *The error estimators  $\eta_H, \zeta_H$  from (16) satisfy the following properties with constants  $C_{\text{stab}}, C_{\text{rel}}, C_{\text{drel}}, C_{\text{mon}} > 0$  and  $0 < q_{\text{red}} < 1$  for any triangulation  $\mathcal{T}_H \in \mathbb{T}$  and any conforming refinement*

$\mathcal{T}_h \in \mathbb{T}(\mathcal{T}_H)$  with the corresponding Galerkin solutions  $u_H^*, z_H^* \in \mathcal{X}_H, u_h^*, z_h^* \in \mathcal{X}_h$  to (10), any subset  $\mathcal{U}_H \subseteq \mathcal{T}_H \cap \mathcal{T}_h$ , and arbitrary  $v_H \in \mathcal{X}_H, v_h \in \mathcal{X}_h$ :

**(A1) stability:**  $|\eta_h(\mathcal{U}_H; v_h) - \eta_H(\mathcal{U}_H; v_H)| + |\zeta_h(\mathcal{U}_H; v_h) - \zeta_H(\mathcal{U}_H; v_H)| \leq C_{\text{stab}} \|v_h - v_H\|$ ;

**(A2) reduction:**  $\eta_h(\mathcal{T}_h \setminus \mathcal{T}_H; v_H) \leq q_{\text{red}} \eta_H(\mathcal{T}_H \setminus \mathcal{T}_h; v_H)$  and  $\zeta_h(\mathcal{T}_h \setminus \mathcal{T}_H; v_H) \leq q_{\text{red}} \zeta_H(\mathcal{T}_H \setminus \mathcal{T}_h; v_H)$ ;

**(A3) reliability:**  $\|u^* - u_H^*\| \leq C_{\text{rel}} \eta_H(u_H^*)$  and  $\|z^* - z_H^*\| \leq C_{\text{rel}} \zeta_H(z_H^*)$ ;

**(A3+) discrete reliability:**  $\|u_h^* - u_H^*\| \leq C_{\text{drel}} \eta_H(\mathcal{T}_H \setminus \mathcal{T}_h, u_H^*)$  and  $\|z_h^* - z_H^*\| \leq C_{\text{drel}} \zeta_H(\mathcal{T}_H \setminus \mathcal{T}_h, z_H^*)$ ;

**(QM) quasi-monotonicity:**  $\eta_h(u_h^*) \leq C_{\text{mon}} \eta_H(u_H^*)$  and  $\zeta_h(z_h^*) \leq C_{\text{mon}} \zeta_H(z_H^*)$ .

The constant  $C_{\text{rel}}$  depends only on the uniform  $\gamma$ -shape regularity of all  $\mathcal{T}_H \in \mathbb{T}$  and on the space dimension  $d$ , while  $C_{\text{stab}}$  and  $C_{\text{drel}}$  additionally depend on the polynomial degree  $p$ . For NVB, reduction (A2) holds with  $q_{\text{red}} := 2^{-1/(2d)}$ . Moreover, the constant in quasi-monotonicity (QM) satisfies  $C_{\text{mon}} \leq \min\{1 + C_{\text{stab}}(1 + C_{\text{Céa}})C_{\text{rel}}, 1 + C_{\text{stab}}C_{\text{drel}}\}$ .

Reliability (A3) and stability (A1) verify

$$\begin{aligned} \|u^* - u_H\| &\leq \max\{C_{\text{rel}}, 1 + C_{\text{stab}} C_{\text{rel}}\} [\eta_H(u_H) + \|u_H^* - u_H\|], \\ \|z^* - z_H\| &\leq \max\{C_{\text{rel}}, 1 + C_{\text{stab}} C_{\text{rel}}\} [\zeta_H(z_H) + \|z_H^* - z_H\|]. \end{aligned}$$

In combination with the estimate (12), we finally conclude for  $C_{\text{goal}} := L \max\{C_{\text{rel}}, 1 + C_{\text{stab}} C_{\text{rel}}\}^2$  the reliable goal-error estimate

$$|G(u^*) - G_H(u_H, z_H)| \leq C_{\text{goal}} [\eta_H(u_H) + \|u_H^* - u_H\|] [\zeta_H(z_H) + \|z_H^* - z_H\|], \quad (17)$$

which provides the core estimate of the proposed adaptive algorithm in Section 3 below.

The ellipticity of  $b(\cdot, \cdot)$  from (7) ensures inf-sup stability of the elliptic problem at hand. Recall from [34] that inf-sup stability implies the generalized quasi-orthogonality, which will be an important tool in the subsequent analysis.

**Proposition 1** (validity of quasi-orthogonality [34, Eq. (8)]). *For any sequence  $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1} \subset \mathcal{X}$  of nested discrete subspaces with  $\ell \geq 0$ , there holds*

**(A4) quasi-orthogonality:** *There exist constants  $C_{\text{orth}} > 0$  and  $0 < \delta < 1$  such that the corresponding Galerkin solutions  $u_\ell^*, z_\ell^* \in \mathcal{X}_\ell$  to (10) satisfy, for all  $\ell, M \in \mathbb{N}_0$ ,*

$$\sum_{\ell'=\ell}^{\ell+M} \|u_{\ell'+1}^* - u_{\ell'}^*\|^2 \leq C_{\text{orth}} (M+1)^{1-\delta} \|u^* - u_\ell^*\|^2, \quad (18a)$$

$$\sum_{\ell'=\ell}^{\ell+M} \|z_{\ell'+1}^* - z_{\ell'}^*\|^2 \leq C_{\text{orth}} (M+1)^{1-\delta} \|z^* - z_\ell^*\|^2. \quad (18b)$$

The constants  $C_{\text{orth}}$  and  $\delta$  depend only on the dimension  $d$ , the elliptic bilinear form  $b(\cdot, \cdot)$ , and the chosen norm  $\|\cdot\|$ , but are independent of the spaces  $\mathcal{X}_\ell$ .

### 3 Adaptive algorithm

In this section, we introduce our goal-oriented adaptive iteratively symmetrized algorithm. It utilizes specific stopping indices denoted by an underline, e.g.,  $\underline{\ell}, \underline{m}[\ell], \underline{n}[\ell, k] \in \mathbb{N}_0$ . For an overview, see Table 1 above. However, we may omit the dependence whenever it is apparent from the context, such as in the abbreviation  $\underline{n} := \underline{n}[\ell, m]$  for  $u_\ell^{\underline{m}, \underline{n}}$ .

**Algorithm 1** (GOAISFEM).

**Input:** Initial mesh  $\mathcal{T}_0$ , polynomial degree  $p \in \mathbb{N}$ , marking parameters  $0 < \theta \leq 1$ ,  $C_{\text{mark}} \geq 1$ , solver parameters  $\lambda_{\text{sym}} > 0$ ,  $\lambda_{\text{alg}} > 0$ , Zarantonello damping parameter  $\delta > 0$ , and initial guesses  $u_0^{0,0} = u_0^{0,n}$ ,  $z_0^{0,0} = z_0^{0,v} \in \mathcal{X}_0$ .

**Adaptive loop:** For all  $\ell = 0, 1, 2, \dots$ , repeat the following steps (I)–(IV):

(I) SOLVE & ESTIMATE (PRIMAL). For all  $m = 1, 2, 3, \dots$ , repeat (a)–(c):

- (a) Set  $u_\ell^{m,0} := u_\ell^{m-1,n}$  and define for theoretical reasons  $u_\ell^{m,\star} := \Phi_\ell^u(\delta; u_\ell^{m-1,n})$ .  
 (b) For all  $n = 1, 2, 3, \dots$ , repeat the following steps (i)–(ii):

- (i) Compute  $u_\ell^{m,n} := \Psi_\ell(u_\ell^{m,\star}; u_\ell^{m,n-1})$  and corresponding refinement indicators  $\eta_\ell(T; u_\ell^{m,n})$  for all  $T \in \mathcal{T}_\ell$ .  
 (ii) Terminate  $n$ -loop and define  $\underline{n}[\ell, m] := n$  if

$$\|u_\ell^{m,n} - u_\ell^{m,n-1}\| \leq \lambda_{\text{alg}} \left[ \lambda_{\text{sym}} \eta_\ell(u_\ell^{m,n}) + \|u_\ell^{m,n} - u_\ell^{m,0}\| \right]. \quad (19)$$

(c) Terminate  $m$ -loop and define  $\underline{m}[\ell] := m$  if

$$\|u_\ell^{m,n} - u_\ell^{m,0}\| \leq \lambda_{\text{sym}} \eta_\ell(u_\ell^{m,n}). \quad (20)$$

(II) SOLVE & ESTIMATE (DUAL). For all  $\mu = 1, 2, 3, \dots$ , repeat (a)–(c):

- (a) Set  $z_\ell^{\mu,0} := z_\ell^{\mu-1,v}$  and define for theoretical reasons  $z_\ell^{\mu,\star} := \Phi_\ell^z(\delta; z_\ell^{\mu-1,v})$ .  
 (b) For all  $\nu = 1, 2, 3, \dots$ , repeat the following steps (i)–(ii):

- (i) Compute  $z_\ell^{\mu,\nu} := \Psi_\ell(z_\ell^{\mu,\star}; z_\ell^{\mu,\nu-1})$  and corresponding refinement indicators  $\zeta_\ell(T; z_\ell^{\mu,\nu})$  for all  $T \in \mathcal{T}_\ell$ .  
 (ii) Terminate  $\nu$ -loop and define  $\underline{\nu}[\ell, \mu] := \nu$  if

$$\|z_\ell^{\mu,\nu} - z_\ell^{\mu,\nu-1}\| \leq \lambda_{\text{alg}} \left[ \lambda_{\text{sym}} \zeta_\ell(z_\ell^{\mu,\nu}) + \|z_\ell^{\mu,\nu} - z_\ell^{\mu,0}\| \right]. \quad (21)$$

(c) Terminate  $\mu$ -loop and define  $\underline{\mu}[\ell] := \mu$  if

$$\|z_\ell^{\mu,\nu} - z_\ell^{\mu,0}\| \leq \lambda_{\text{sym}} \zeta_\ell(z_\ell^{\mu,\nu}). \quad (22)$$

(III) MARK. Determine sets

$$\begin{aligned} \overline{\mathcal{M}}_\ell^u &\in \mathbb{M}_\ell^u[\theta, u_\ell^{m,n}] := \left\{ \mathcal{U}_\ell \subseteq \mathcal{T}_\ell : \theta \eta_\ell(u_\ell^{m,n})^2 \leq \eta_\ell(\mathcal{U}_\ell, u_\ell^{m,n})^2 \right\}, \\ \overline{\mathcal{M}}_\ell^z &\in \mathbb{M}_\ell^z[\theta, z_\ell^{\mu,\nu}] := \left\{ \mathcal{U}_\ell \subseteq \mathcal{T}_\ell : \theta \zeta_\ell(z_\ell^{\mu,\nu})^2 \leq \zeta_\ell(\mathcal{U}_\ell, z_\ell^{\mu,\nu})^2 \right\} \end{aligned}$$

satisfying the following Dörfler criterion [1] with quasi-minimal cardinality

$$\#\overline{\mathcal{M}}_\ell^u \leq C_{\text{mark}} \min_{\mathcal{U}_\ell^* \in \mathbb{M}_\ell^u[\theta, u_\ell^{m,n}]} \#\mathcal{U}_\ell^*, \quad \#\overline{\mathcal{M}}_\ell^z \leq C_{\text{mark}} \min_{\mathcal{U}_\ell^* \in \mathbb{M}_\ell^z[\theta, z_\ell^{\mu,\nu}]} \#\mathcal{U}_\ell^*. \quad (23)$$

As in [17], define the set of marked elements  $\mathcal{M}_\ell := \mathcal{M}_\ell^u \cup \mathcal{M}_\ell^z$ , where  $\mathcal{M}_\ell^u \subseteq \overline{\mathcal{M}}_\ell^u$  and  $\mathcal{M}_\ell^z \subseteq \overline{\mathcal{M}}_\ell^z$  satisfy  $\#\mathcal{M}_\ell^u = \#\mathcal{M}_\ell^z = \min\{\#\overline{\mathcal{M}}_\ell^u, \#\overline{\mathcal{M}}_\ell^z\}$ .

(IV) REFINE. Generate the new mesh  $\mathcal{T}_{\ell+1} := \text{refine}(\mathcal{M}_\ell, \mathcal{T}_\ell)$  by NVB and define  $u_{\ell+1}^{0,0} := u_{\ell+1}^{0,n} := u_{\ell+1}^{0,\star} := u_{\ell+1}^{m,n}$  and  $z_{\ell+1}^{0,0} := z_{\ell+1}^{0,v} := z_{\ell+1}^{0,\star} := z_{\ell+1}^{\mu,\nu}$  (nested iteration).

**Output:** Sequences of successively refined triangulations  $\mathcal{T}_\ell$ , successive discrete approximations  $u_\ell^{m,n}$ ,  $z_\ell^{\mu,\nu}$ , and corresponding error estimators  $\eta_\ell(u_\ell^{m,n})$ ,  $\zeta_\ell(z_\ell^{\mu,\nu})$ .



**Remark 1.** (i) Although the primal loop (I) and dual loop (II) in Algorithm 1 are displayed sequentially, they are independent of each other. Therefore, a practical implementation will realize these iterations simultaneously since the system matrix is the same (thanks to the symmetrization step).

(ii) In order to investigate the asymptotic behavior, it is reasonable to analyze Algorithm 1 in the present formulation with infinitely many steps. We note that a practical implementation will terminate with  $\underline{\ell} := \ell$  provided that the estimator product is smaller than a user-specified tolerance.

For the analysis of Algorithm 1, we define the index set  $Q := Q^u \cup Q^z$  with

$$Q^u := \{(\ell, m, n) \in \mathbb{N}_0^3: u_\ell^{m,n} \text{ is used in Algorithm 1}\}, \quad (24a)$$

$$Q^z := \{(\ell, \mu, \nu) \in \mathbb{N}_0^3: z_\ell^{\mu,\nu} \text{ is used in Algorithm 1}\}. \quad (24b)$$

Furthermore, we require the following final indices and notice that these are consistent with those defined in Algorithm 1:

$$\underline{\ell} := \sup\{\ell \in \mathbb{N}_0: (\ell, 0, 0) \in Q^u \text{ or } (\ell, 0, 0) \in Q^z\} \in \mathbb{N}_0 \cup \{\infty\}, \quad (25a)$$

$$\underline{m}[\ell] := \sup\{m \in \mathbb{N}: (\ell, m, 0) \in Q^u\}, \quad \underline{\mu}[\ell] := \sup\{\mu \in \mathbb{N}: (\ell, \mu, 0) \in Q^z\}, \quad (25b)$$

$$\underline{n}[\ell, m] := \sup\{n \in \mathbb{N}: (\ell, m, n) \in Q^u\}, \quad \underline{\nu}[\ell, \mu] := \sup\{\nu \in \mathbb{N}: (\ell, \mu, \nu) \in Q^z\}. \quad (25c)$$

In addition, we set  $\underline{k}[\ell] := \max\{\underline{m}[\ell], \underline{\mu}[\ell]\}$  as well as  $\underline{j}[\ell, k] := \max\{\underline{n}[\ell, k], \underline{\nu}[\ell, k]\}$ .

Finally, we introduce the total step counter  $|\cdot, \cdot, \cdot|$  defined for all  $(\ell, k, j) \in Q$  by

$$|\ell, k, j| = \sum_{\ell'=0}^{\ell-1} \sum_{k'=0}^{-k[\ell']-j[\ell',k']} 1 + \sum_{k'=0}^{k-1} \sum_{j'=0}^{-j[\ell,k']} 1 + \sum_{j'=0}^{j-1} 1.$$

This definition indeed provides a lexicographic ordering on  $Q$ , if the solver steps Algorithm 1(I) for  $u_\ell^{m,n}$  and Algorithm 1(II) for  $z_\ell^{\mu,\nu}$  are done in *parallel*. We note that one solver step of an optimal geometric multigrid method on graded meshes can be performed in  $\mathcal{O}(\#\mathcal{T}_\ell)$  operations; see, e.g., [30], [32]. For given  $u_\ell^{m,n}, z_\ell^{\mu,\nu} \in \mathcal{X}_\ell$ , the simultaneous computation of the refinement indicators  $\eta_\ell(T, u_\ell^{m,n})$  and  $\zeta_\ell(T, z_\ell^{\mu,\nu})$  requires  $\mathcal{O}(\#\mathcal{T}_\ell)$  operations, hence the steps Algorithm 1(I)–(II) require  $\mathcal{O}(\#\mathcal{T}_\ell)$  operations as well. Furthermore, Dörfler marking can be performed in  $\mathcal{O}(\#\mathcal{T}_\ell)$  operations; see, e.g., [4], [49]. Therefore, the total work to compute  $u_\ell^{m,n}$  and  $z_\ell^{\mu,\nu}$  is (up to a constant) given by

$$\text{cost}(\ell, k, j) := \sum_{\substack{(\ell', m', n') \in Q^u \\ |\ell', m', n'| \leq |\ell, k, j|}} \#\mathcal{T}_{\ell'} + \sum_{\substack{(\ell', \mu', \nu') \in Q^z \\ |\ell', \mu', \nu'| \leq |\ell, k, j|}} \#\mathcal{T}_{\ell'} \simeq \sum_{\substack{(\ell', k', j') \in Q \\ |\ell', k', j'| \leq |\ell, k, j|}} \#\mathcal{T}_{\ell'}. \quad (26)$$

Since  $\#Q = \infty$ , we have either  $\underline{\ell} = \infty$ , or  $\underline{k}[\underline{\ell}] = \infty$ , or  $\underline{j}[\underline{\ell}, \underline{k}] = \infty$ . A further observation about Algorithm 1 is that the nested algebraic solver loop within the Zarantonello loop is guaranteed to terminate, and the latter case  $\underline{j}[\underline{\ell}, \underline{k}] = \infty$  is therefore excluded.

**Lemma 2** (finite termination of algebraic solver [28, Lemma 3.2]). *Independently of the algorithmic parameters  $\delta, \theta, \lambda_{\text{sym}}$ , and  $\lambda_{\text{alg}}$ , the innermost  $n$ - and  $\nu$ -loops of Algorithm 1 always terminate. In particular,  $\underline{j}[\underline{\ell}, \underline{k}] < \infty$  for all  $(\ell, k, 0) \in Q$ .*

## 4 A posteriori error analysis

Algorithm 1 does not provide the exact algebraic solutions  $u_\ell^{m,\star}$  and  $z_\ell^{\mu,\star}$  to (13) but instead uses an inexact algebraic solver. However, the following result from [28] applies to the primal and the dual problem alike and shows that these inexact Zarantonello iterations remain contractions except for the final iterate on each mesh (see also [50] for an extended version).

**Lemma 3** (contraction of inexact Zarantonello iteration [28, Lemma 5.1]). *Choose any damping parameter  $0 < \delta < \delta^* = 2\alpha/L^2$  to ensure the contraction (14) of the Zarantonello iteration and*

$$0 < \lambda_{\text{alg}}^* < \frac{(1 - q_{\text{sym}}^*)(1 - q_{\text{alg}})}{4q_{\text{alg}}} \quad \text{such that} \quad 0 < q_{\text{sym}} := \frac{q_{\text{sym}}^* + 2 \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}}^*}{1 - 2 \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}}^*} < 1. \quad (27)$$

*Then, for arbitrary  $\lambda_{\text{sym}} > 0$  and any  $0 < \lambda_{\text{alg}} \leq \lambda_{\text{alg}}^*$ , we have for all  $(\ell, m, \underline{n}) \in \mathcal{Q}^u$  with  $1 \leq m < \underline{m}[\ell]$  and all  $(\ell, \mu, \underline{\nu}) \in \mathcal{Q}^z$  with  $1 \leq \mu < \underline{\mu}[\ell]$  that*

$$\|u_\ell^* - u_\ell^{m, \underline{n}}\| \leq q_{\text{sym}} \|u_\ell^* - u_\ell^{m-1, \underline{n}}\|, \quad \|z_\ell^* - z_\ell^{\mu, \underline{\nu}}\| \leq q_{\text{sym}} \|z_\ell^* - z_\ell^{\mu-1, \underline{\nu}}\|. \quad (28)$$

*Moreover, for  $m = \underline{m}[\ell]$  resp.  $\mu = \underline{\mu}[\ell]$ , it holds that*

$$\begin{aligned} \|u_\ell^* - u_\ell^{m, \underline{n}}\| &\leq q_{\text{sym}}^* \|u_\ell^* - u_\ell^{m-1, \underline{n}}\| + \frac{2 q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}} \lambda_{\text{sym}} \eta_\ell \left( u_\ell^{m, \underline{n}} \right), \\ \|z_\ell^* - z_\ell^{\mu, \underline{\nu}}\| &\leq q_{\text{sym}}^* \|z_\ell^* - z_\ell^{\mu-1, \underline{\nu}}\| + \frac{2 q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}} \lambda_{\text{sym}} \zeta_\ell \left( z_\ell^{\mu, \underline{\nu}} \right). \end{aligned} \quad (29)$$

The subsequent lemma gathers *a posteriori* error estimates following directly from the corresponding contraction of the symmetrization, algebraic solver, and the inexact Zarantonello iteration. Further details of the elementary proof are omitted.

**Lemma 4** (stability and a posteriori error control). *For all  $(\ell, m, 0) \in \mathcal{Q}^u$  with  $m \geq 1$ , contraction (14) shows*

$$\frac{1 - q_{\text{sym}}^*}{q_{\text{sym}}^*} \|u_\ell^* - u_\ell^{m, \star}\| \leq \|u_\ell^{m, \star} - u_\ell^{m-1, \underline{n}}\| \leq (1 + q_{\text{sym}}^*) \|u_\ell^* - u_\ell^{m-1, \underline{n}}\|. \quad (30)$$

*Analogously, for all  $(\ell, m, n) \in \mathcal{Q}^u$  with  $n \geq 1$ , the contraction (15) ensures*

$$\frac{1 - q_{\text{alg}}}{q_{\text{alg}}} \|u_\ell^{m, \star} - u_\ell^{m, n}\| \leq \|u_\ell^{m, n} - u_\ell^{m, n-1}\| \leq (1 + q_{\text{alg}}) \|u_\ell^{m, \star} - u_\ell^{m, n-1}\|. \quad (31)$$

*For all  $(\ell, m, \underline{n}) \in \mathcal{Q}^u$  with  $1 \leq m < \underline{m}[\ell]$ , the contraction (28) leads to*

$$\frac{1 - q_{\text{sym}}}{q_{\text{sym}}} \|u_\ell^* - u_\ell^{m, \underline{n}}\| \leq \|u_\ell^{m, \underline{n}} - u_\ell^{m-1, \underline{n}}\| \leq (1 + q_{\text{sym}}) \|u_\ell^* - u_\ell^{m-1, \underline{n}}\|. \quad (32)$$

*The analogous estimates are also valid for the dual variable.*

Finally, the following lemma shows that in the case of finitely many mesh-refinement steps, the Zarantonello iteration does not terminate and one of the two exact continuous solutions is already the discrete solution to (10).

**Lemma 5** (case of finite mesh-refinement steps). *Suppose that the inexact Zarantonello iteration satisfies contraction (28) and that  $\eta$  and  $\zeta$  satisfy (A1)–(A3). If  $\underline{\ell} < \infty$ , then  $\underline{k}[\underline{\ell}] = \infty$  and  $\eta_{\underline{\ell}} \left( u_{\underline{\ell}}^* \right) = 0$  (so that  $u^* = u_{\underline{\ell}}^*$ ) or  $\zeta_{\underline{\ell}} \left( z_{\underline{\ell}}^* \right) = 0$  (so that  $z^* = z_{\underline{\ell}}^*$ ).*

*Proof.* By Lemma 2, we have  $\underline{j}[\underline{\ell}, k] < \infty$ . If  $\underline{\ell} < \infty$ , then  $\underline{k}[\underline{\ell}] = \infty$  and, hence,

$$\eta_{\underline{\ell}} \left( u_{\underline{\ell}}^{m, \underline{n}} \right) < \lambda_{\text{sym}}^{-1} \|u_{\underline{\ell}}^{m, \underline{n}} - u_{\underline{\ell}}^{m-1, \underline{n}}\| \quad \forall m \in \mathbb{N} \quad (33)$$

or

$$\zeta_{\underline{\ell}} \left( z_{\underline{\ell}}^{\mu, \underline{\nu}} \right) < \lambda_{\text{sym}}^{-1} \|z_{\underline{\ell}}^{\mu, \underline{\nu}} - z_{\underline{\ell}}^{\mu-1, \underline{\nu}}\| \quad \forall \mu \in \mathbb{N}. \quad (34)$$

If (33) holds, then the inexact Zarantonello iterates  $u_{\ell}^{m,n}$  are convergent with limit  $u_{\ell}^*$  and we obtain by stability (A1) that

$$\eta_{\ell}(u_{\ell}^*) \stackrel{(A1)}{\leq} \eta_{\ell}(u_{\ell}^{m,n}) + C_{\text{stab}} \|u_{\ell}^* - u_{\ell}^{m,n}\| \stackrel{(33)}{\lesssim} \|u_{\ell}^{m,n} - u_{\ell}^{m-1,n}\| \xrightarrow{m \rightarrow \infty} 0.$$

This proves that  $\eta_{\ell}(u_{\ell}^*) = 0$ , and we infer from reliability (A3) that  $u_{\ell}^* = u^*$ . The same arguments apply to  $z_{\ell}^*$  in the case of (34).  $\square$

Due to the contraction of the inexact Zarantonello iteration (28), we have the following a posteriori error estimates for the final iterates.

**Lemma 6** (stability of final iterates). *Suppose that the inexact Zarantonello iteration satisfies (28). Then, for all  $(\ell + 1, \underline{m}, \underline{n}) \in Q^u$  and  $(\ell + 1, \underline{\mu}, \underline{\nu}) \in Q^z$ , there holds*

$$\|u_{\ell+1}^* - u_{\ell+1}^{m-1,n}\| \leq \|u_{\ell+1}^* - u_{\ell+1}^{m,n}\|, \quad \|z_{\ell+1}^* - z_{\ell+1}^{\mu-1,\nu}\| \leq \|z_{\ell+1}^* - z_{\ell+1}^{\mu,\nu}\|, \quad (35)$$

$$\|u_{\ell+1}^{m,n} - u_{\ell}^{m,n}\| \leq 4 \|u_{\ell+1}^* - u_{\ell}^{m,n}\|, \quad \|z_{\ell+1}^{\mu,\nu} - z_{\ell}^{\mu,\nu}\| \leq 4 \|z_{\ell+1}^* - z_{\ell}^{\mu,\nu}\|, \quad (36)$$

$$\|u_{\ell}^{m,n} - u_{\ell}^{m-1,n}\| \leq 4 \|u_{\ell}^* - u_{\ell}^{m-1,n}\|, \quad \|z_{\ell}^{\mu,\nu} - z_{\ell}^{\mu-1,\nu}\| \leq 4 \|z_{\ell}^* - z_{\ell}^{\mu-1,\nu}\|. \quad (37)$$

*Proof.* For  $(\ell + 1, \underline{m}, \underline{n}) \in Q^u$ , nested iteration  $u_{\ell+1}^{0,n} = u_{\ell}^{m,n}$  together with the contraction of the inexact Zarantonello iteration (28) and  $\underline{m}[\ell + 1] \geq 1$  prove (35) by

$$\|u_{\ell+1}^* - u_{\ell+1}^{m-1,n}\| \stackrel{(28)}{\leq} q_{\text{sym}}^{m[\ell+1]-1} \|u_{\ell+1}^* - u_{\ell+1}^{0,n}\| \leq \|u_{\ell+1}^* - u_{\ell}^{m,n}\|.$$

Let  $(\ell, \underline{m}, \underline{n}) \in Q^u$ . Contraction of the algebraic solver (15), the fact  $\underline{n}[\ell, \underline{m}] \geq 1$ , and nested iteration  $u_{\ell}^{m,0} = u_{\ell}^{m-1,n}$  show that

$$\|u_{\ell}^{m,*} - u_{\ell}^{m,n}\| \stackrel{(15)}{\leq} q_{\text{alg}}^{\underline{n}[\ell, \underline{m}]} \|u_{\ell}^{m,*} - u_{\ell}^{m,0}\| \leq q_{\text{alg}} \|u_{\ell}^{m,*} - u_{\ell}^{m-1,n}\|. \quad (38)$$

This and with the contraction of the exact Zarantonello iteration (14) result in

$$\begin{aligned} \|u_{\ell}^* - u_{\ell}^{m,n}\| &\leq \|u_{\ell}^* - u_{\ell}^{m,*}\| + \|u_{\ell}^{m,*} - u_{\ell}^{m,n}\| \\ &\stackrel{(38)}{\leq} (1 + q_{\text{alg}}) \|u_{\ell}^* - u_{\ell}^{m,*}\| + q_{\text{alg}} \|u_{\ell}^* - u_{\ell}^{m-1,n}\| \\ &\stackrel{(14)}{\leq} \left[ (1 + q_{\text{alg}}) q_{\text{sym}}^* + q_{\text{alg}} \right] \|u_{\ell}^* - u_{\ell}^{m-1,n}\| \leq 3 \|u_{\ell}^* - u_{\ell}^{m-1,n}\|. \end{aligned} \quad (39)$$

Consequently, the combination of (39) and (35) validates (36) via

$$\begin{aligned} \|u_{\ell+1}^{m,n} - u_{\ell}^{m,n}\| &\leq \|u_{\ell+1}^* - u_{\ell+1}^{m,n}\| + \|u_{\ell+1}^* - u_{\ell}^{m,n}\| \\ &\stackrel{(39)}{\leq} 3 \|u_{\ell+1}^* - u_{\ell+1}^{m-1,n}\| + \|u_{\ell+1}^* - u_{\ell}^{m,n}\| \stackrel{(35)}{\leq} 4 \|u_{\ell+1}^* - u_{\ell}^{m,n}\|. \end{aligned}$$

The estimate (39) also implies (37), because

$$\|u_{\ell}^{m,n} - u_{\ell}^{m-1,n}\| \leq \|u_{\ell}^* - u_{\ell}^{m,n}\| + \|u_{\ell}^* - u_{\ell}^{m-1,n}\| \stackrel{(39)}{\leq} 4 \|u_{\ell}^* - u_{\ell}^{m-1,n}\|.$$

The same arguments prove the estimates for the dual variable and conclude the proof.  $\square$

The subsequent lemma states the estimator reduction for only one of the two error estimators. This poses a significant challenge in the proof of full linear convergence due to the required contraction of the nonlinear quasi-error product in Lemma 8 below.

**Lemma 7** (estimator reduction and stability). *Define the constant  $0 < q(\theta) := [1 - (1 - q_{\text{red}}^2) \theta]^{1/2} < 1$  and suppose that the estimators  $\eta$  and  $\zeta$  satisfy (A1)–(A2). If the primal error estimator satisfies the Dörfler criterion, i.e.,  $\mathcal{M}_\ell^u = \overline{\mathcal{M}}_\ell^u \subseteq \mathcal{M}_\ell$  in Algorithm 1(III), then*

$$\begin{aligned} \eta_{\ell+1}\left(u_{\ell+1}^{\underline{m}, \underline{n}}\right) &\leq q(\theta) \eta_\ell\left(u_\ell^{\underline{m}, \underline{n}}\right) + 4 C_{\text{stab}} \left\| u_{\ell+1}^* - u_\ell^{\underline{m}, \underline{n}} \right\| & \forall (\ell + 1, \underline{m}, \underline{n}) \in \mathcal{Q}^u, \\ \zeta_{\ell+1}\left(z_{\ell+1}^{\underline{\mu}, \underline{\nu}}\right) &\leq \zeta_\ell\left(z_\ell^{\underline{\mu}, \underline{\nu}}\right) + 4 C_{\text{stab}} \left\| z_{\ell+1}^* - z_\ell^{\underline{\mu}, \underline{\nu}} \right\| & \forall (\ell + 1, \underline{\mu}, \underline{\nu}) \in \mathcal{Q}^z. \end{aligned} \quad (40)$$

*If the dual error estimator satisfies the Dörfler criterion, i.e.,  $\mathcal{M}_\ell^z = \overline{\mathcal{M}}_\ell^z \subseteq \mathcal{M}_\ell$  in Algorithm 1(III), then*

$$\begin{aligned} \eta_{\ell+1}\left(u_{\ell+1}^{\underline{m}, \underline{n}}\right) &\leq \eta_\ell\left(u_\ell^{\underline{m}, \underline{n}}\right) + 4 C_{\text{stab}} \left\| u_{\ell+1}^* - u_\ell^{\underline{m}, \underline{n}} \right\| & \forall (\ell + 1, \underline{m}, \underline{n}) \in \mathcal{Q}^u, \\ \zeta_{\ell+1}\left(z_{\ell+1}^{\underline{\mu}, \underline{\nu}}\right) &\leq q(\theta) \zeta_\ell\left(z_\ell^{\underline{\mu}, \underline{\nu}}\right) + 4 C_{\text{stab}} \left\| z_{\ell+1}^* - z_\ell^{\underline{\mu}, \underline{\nu}} \right\| & \forall (\ell + 1, \underline{\mu}, \underline{\nu}) \in \mathcal{Q}^z. \end{aligned} \quad (41)$$

*Proof.* For  $(\ell + 1, 0, 0) \in \mathcal{Q}^u$ , stability (A1) and reduction (A2) yield that

$$\begin{aligned} \eta_{\ell+1}\left(u_\ell^{\underline{m}, \underline{n}}\right)^2 &= \eta_{\ell+1}\left(\mathcal{T}_{\ell+1} \cap \mathcal{T}_\ell; u_\ell^{\underline{m}, \underline{n}}\right)^2 + \eta_{\ell+1}\left(\mathcal{T}_{\ell+1} \setminus \mathcal{T}_\ell; u_\ell^{\underline{m}, \underline{n}}\right)^2 \\ &\leq \eta_\ell\left(\mathcal{T}_{\ell+1} \cap \mathcal{T}_\ell; u_\ell^{\underline{m}, \underline{n}}\right)^2 + q_{\text{red}}^2 \eta_\ell\left(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}; u_\ell^{\underline{m}, \underline{n}}\right)^2 \\ &= \eta_\ell\left(u_\ell^{\underline{m}, \underline{n}}\right)^2 - (1 - q_{\text{red}}^2) \eta_\ell\left(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}; u_\ell^{\underline{m}, \underline{n}}\right)^2. \end{aligned} \quad (42)$$

The Dörfler marking in Algorithm 1(III) for the primal error estimator  $\eta$  and  $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$  prove the contraction in (40)

$$\eta_{\ell+1}\left(u_\ell^{\underline{m}, \underline{n}}\right)^2 \leq \eta_\ell\left(u_\ell^{\underline{m}, \underline{n}}\right)^2 - (1 - q_{\text{red}}^2) \eta_\ell\left(\mathcal{M}_\ell; u_\ell^{\underline{m}, \underline{n}}\right)^2 \leq q(\theta)^2 \eta_\ell\left(u_\ell^{\underline{m}, \underline{n}}\right)^2. \quad (43)$$

For  $(\ell + 1, \underline{m}, \underline{n}) \in \mathcal{Q}^u$ , this and (36) lead to

$$\begin{aligned} \eta_{\ell+1}\left(u_{\ell+1}^{\underline{m}, \underline{n}}\right) &\stackrel{(A1)}{\leq} \eta_{\ell+1}\left(u_\ell^{\underline{m}, \underline{n}}\right) + C_{\text{stab}} \left\| u_{\ell+1}^{\underline{m}, \underline{n}} - u_\ell^{\underline{m}, \underline{n}} \right\| \\ &\stackrel{(43)}{\leq} q(\theta) \eta_\ell\left(u_\ell^{\underline{m}, \underline{n}}\right) + C_{\text{stab}} \left\| u_{\ell+1}^{\underline{m}, \underline{n}} - u_\ell^{\underline{m}, \underline{n}} \right\| \\ &\stackrel{(36)}{\leq} q(\theta) \eta_\ell\left(u_\ell^{\underline{m}, \underline{n}}\right) + 4 C_{\text{stab}} \left\| u_{\ell+1}^* - u_\ell^{\underline{m}, \underline{n}} \right\|. \end{aligned}$$

For  $(\ell + 1, \underline{\mu}, \underline{\nu}) \in \mathcal{Q}^z$ , we argue analogously to (42) in order to obtain that  $\zeta_{\ell+1}\left(z_\ell^{\underline{\mu}, \underline{\nu}}\right) \leq \zeta_\ell\left(z_\ell^{\underline{\mu}, \underline{\nu}}\right)$ . Together with (36), it follows that

$$\zeta_{\ell+1}\left(z_{\ell+1}^{\underline{\mu}, \underline{\nu}}\right) \stackrel{(A1)}{\leq} \zeta_{\ell+1}\left(z_\ell^{\underline{\mu}, \underline{\nu}}\right) + C_{\text{stab}} \left\| z_{\ell+1}^{\underline{\mu}, \underline{\nu}} - z_\ell^{\underline{\mu}, \underline{\nu}} \right\| \stackrel{(36)}{\leq} \zeta_\ell\left(z_\ell^{\underline{\mu}, \underline{\nu}}\right) + 4 C_{\text{stab}} \left\| z_{\ell+1}^* - z_\ell^{\underline{\mu}, \underline{\nu}} \right\|.$$

The proof holds verbatim in the case of Dörfler marking for the dual error estimator, albeit with reversed roles. This concludes the proof.  $\square$

## 5 Full linear convergence

This section presents full linear convergence of Algorithm 1 as the first main result of this work. Recall the goal-error estimate from (17) motivating the product structure of the respective primal and dual error components. Thus, we define the quasi-errors

$$\mathbb{H}_\ell^{m, n} := \left\| u_\ell^* - u_\ell^{m, n} \right\| + \left\| u_\ell^{m, \star} - u_\ell^{m, n} \right\| + \eta_\ell\left(u_\ell^{m, n}\right) \quad \forall (\ell, m, n) \in \mathcal{Q}^u, \quad (44a)$$

$$Z_\ell^{\mu,\nu} := \|z_\ell^\star - z_\ell^{\mu,\nu}\| + \|z_\ell^{\mu,\star} - z_\ell^{\mu,\nu}\| + \zeta_\ell(z_\ell^{\mu,\nu}) \quad \forall (\ell, \mu, \nu) \in \mathcal{Q}^z. \quad (44b)$$

The quasi-errors naturally extend to the full index set  $(\ell, k, j) \in \mathcal{Q}$  by

$$H_\ell^{k,j} := \begin{cases} H_\ell^{k,n} & \text{if } (\ell, k, 0) \in \mathcal{Q}^u \text{ but } (\ell, k, j) \notin \mathcal{Q}^u, \\ H_\ell^{m,n} & \text{if } (\ell, k, 0) \notin \mathcal{Q}^u, \end{cases} \quad (45)$$

$$Z_\ell^{k,j} := \begin{cases} Z_\ell^{k,\nu} & \text{if } (\ell, k, 0) \in \mathcal{Q}^z \text{ but } (\ell, k, j) \notin \mathcal{Q}^z, \\ Z_\ell^{\mu,\nu} & \text{if } (\ell, k, 0) \notin \mathcal{Q}^z. \end{cases}$$

The following theorem asserts full linear convergence of the quasi-error product.

**Theorem 1** (full linear convergence). *Suppose that the estimators  $\eta$  and  $\zeta$  satisfy (A1)–(A3) and (QM) and suppose (A4). Recall  $\lambda_{\text{alg}}^\star$  and  $q_{\text{sym}}$  from Lemma 3. With the constant  $q(\theta)$  from Lemma 7 and  $\bar{q} := \max\{q(\theta)^{1/2}, (1 + q_{\text{sym}}^\star)/2\} < 1$ , let*

$$0 < \lambda^\star := \frac{(1 - q_{\text{alg}})(\bar{q} - q_{\text{sym}}^\star)(1 - \bar{q})}{10 q_{\text{alg}} C_{\text{stab}}}. \quad (46)$$

Then, for arbitrary marking parameter  $0 < \theta \leq 1$  and any solver parameters  $\lambda_{\text{sym}} > 0$  and  $0 < \lambda_{\text{alg}} \leq \lambda_{\text{alg}}^\star$  with  $\lambda_{\text{sym}} \lambda_{\text{alg}} \leq \lambda^\star$ , Algorithm 1 guarantees full linear convergence: There exist constants  $C_{\text{lin}} \geq 1$  and  $0 < q_{\text{lin}} < 1$  such that the quasi-error product satisfies, for all  $(\ell, k, j), (\ell', k', j') \in \mathcal{Q}$  with  $|\ell', k', j'| \leq |\ell, k, j|$

$$H_\ell^{k,j} Z_\ell^{k,j} \leq C_{\text{lin}} q_{\text{lin}}^{|\ell,k,j| - |\ell',k',j'|} H_{\ell'}^{k',j'} Z_{\ell'}^{k',j'}. \quad (47)$$

The constants  $C_{\text{lin}}$  and  $q_{\text{lin}}$  depend only on  $C_{\text{stab}}, C_{\text{rel}}, C_{\text{mon}}, C_{\text{orth}}, C_{\text{Céa}}, \theta, q_{\text{red}}, q_{\text{sym}}, q_{\text{sym}}^\star, q_{\text{alg}}, \lambda_{\text{sym}}$  and  $\lambda_{\text{alg}}$ .

Three lemmas are required to prove Theorem 1. The characterization of  $R$ -linear convergence from [33, Lemma 5 and 10] is the primary tool for the proof of Theorem 1; see (70) below. The proof of Theorem 1 departs with the contraction of the quasi-error for the final iterates of the inexact Zarantonello loop up to a remainder on the mesh level  $\ell$ . To this end, we define the simplified weighted quasi-error

$$H_\ell := \left[ \|u_\ell^\star - u_\ell^{m,n}\| + \gamma \eta_\ell(u_\ell^{m,n}) \right], \quad Z_\ell := \left[ \|z_\ell^\star - z_\ell^{\mu,\nu}\| + \gamma \zeta_\ell(z_\ell^{\mu,\nu}) \right] \quad \forall (\ell, \underline{k}, \underline{j}) \in \mathcal{Q}, \quad (48)$$

where  $\gamma > 0$  is a free parameter chosen in (51) below. This quasi-error quantity satisfies contraction up to a tail-summable remainder due to estimator reduction (40), (41).

**Lemma 8** (contraction in mesh level up to tail-summable remainder). *Under the assumptions of Theorem 1, there exists  $0 < q < 1$  such that the quasi-error product  $H_\ell Z_\ell$  from (48) satisfies contraction up to a remainder  $R_\ell \geq 0$ ,*

$$H_{\ell+1} Z_{\ell+1} \leq q H_\ell Z_\ell + q R_\ell \quad \forall (\ell + 1, \underline{k}, \underline{j}) \in \mathcal{Q}. \quad (49)$$

The remainder  $R_\ell$  satisfies

$$R_{\ell+M} \lesssim H_\ell Z_\ell \quad \text{and} \quad \sum_{\ell'=\ell}^{\ell+M} R_{\ell'}^2 \lesssim (M+1)^{1-\delta} H_\ell^2 Z_\ell^2 \quad \forall \ell, M \in \mathbb{N}_0 \text{ with } \ell + M < \underline{\ell}. \quad (50)$$

*Proof.* The proof consists of four steps.

**Step 1 (choice of constants).** Recall the constants  $0 < q(\theta) < 1$  from Lemma 7 and  $\lambda^\star > 0$  and  $0 < \bar{q} < 1$  defined in the statement of Theorem 1 and define the constants

$$C(\gamma, \lambda) := 1 + \frac{2 q_{\text{alg}}}{1 - q_{\text{alg}}} \frac{\lambda}{\gamma} > 1, \quad 0 < q_{\text{ctr}} := \max\left\{ q_{\text{sym}}^\star + 4C_{\text{stab}} C(\gamma, \lambda) \gamma, q(\theta)C(\gamma, \lambda) \right\}.$$

Elementary calculations show that the choice of

$$\gamma := \frac{\bar{q}(\bar{q} - q_{\text{sym}}^*)}{4 C_{\text{stab}}} < 1 \quad (51)$$

ensures  $q_{\text{sym}}^* C(\gamma, \lambda) + 4 C_{\text{stab}} \gamma C(\gamma, \lambda)^2 < 1$  as well as, for all  $0 < \lambda < \lambda^*$ ,

$$C(\gamma, \lambda) = 1 + \frac{2 q_{\text{alg}}}{1 - q_{\text{alg}}} \frac{\lambda}{\gamma} < 1 + \frac{1 - \bar{q}}{\bar{q}} = \frac{1}{q(\theta)^{1/2}}. \quad (52)$$

Consequently, we have  $q(\theta)C(\gamma, \lambda)^2 < 1$  and thus  $0 < q'_{\text{ctr}} := C(\gamma, \lambda) q_{\text{ctr}} < 1$  and  $q_{\text{ctr}} < 1$ .

**Step 2 (contraction of  $H_\ell$  and  $Z_\ell$ ).** Abbreviate  $\lambda := \lambda_{\text{alg}} \lambda_{\text{sym}}$ . Recall that marking in Algorithm 1(III) ensures that the estimate (40) or (41) hold. If (40) is satisfied, the quasi-contraction of the inexact Zarantonello iteration (29) for the final iterate, the stability estimate (35), and the estimator reduction (40) lead, for all  $(\ell + 1, \underline{k}, \underline{j}) \in \mathcal{Q}^u$ , to

$$\begin{aligned} H_{\ell+1} &\stackrel{(29)}{\leq} q_{\text{sym}}^* \left\| u_{\ell+1}^* - u_{\ell+1}^{m-1, n} \right\| + C(\gamma, \lambda) \gamma \eta_{\ell+1} \left( u_{\ell+1}^{m, n} \right) \\ &\stackrel{(35)}{\leq} q_{\text{sym}}^* \left\| u_{\ell+1}^* - u_{\ell+1}^{m, n} \right\| + C(\gamma, \lambda) \gamma \eta_{\ell+1} \left( u_{\ell+1}^{m, n} \right) \\ &\stackrel{(40)}{\leq} \left( q_{\text{sym}}^* + 4 C_{\text{stab}} C(\gamma, \lambda) \gamma \right) \left\| u_{\ell+1}^* - u_{\ell+1}^{m, n} \right\| + q(\theta) C(\gamma, \lambda) \gamma \eta_{\ell} \left( u_{\ell}^{m, n} \right) \\ &\leq q_{\text{ctr}} \left[ \left\| u_{\ell+1}^* - u_{\ell+1}^{m, n} \right\| + \gamma \eta_{\ell} \left( u_{\ell}^{m, n} \right) \right]. \end{aligned} \quad (53)$$

The same arguments yield, for all  $(\ell + 1, \underline{\mu}, \underline{\nu}) \in \mathcal{Q}^z$ ,

$$\begin{aligned} Z_{\ell+1} &\stackrel{(29)}{\leq} q_{\text{sym}}^* \left\| z_{\ell+1}^* - z_{\ell+1}^{\mu-1, \nu} \right\| + C(\gamma, \lambda) \gamma \zeta_{\ell+1} \left( z_{\ell+1}^{\mu, \nu} \right) \\ &\stackrel{(35)}{\leq} q_{\text{sym}}^* \left\| z_{\ell+1}^* - z_{\ell+1}^{\mu, \nu} \right\| + C(\gamma, \lambda) \gamma \zeta_{\ell+1} \left( z_{\ell+1}^{\mu, \nu} \right) \\ &\stackrel{(40)}{\leq} C(\gamma, \lambda) \left[ q_{\text{ctr}} \left\| z_{\ell+1}^* - z_{\ell+1}^{\mu, \nu} \right\| + \gamma \zeta_{\ell} \left( z_{\ell}^{\mu, \nu} \right) \right]. \end{aligned} \quad (54)$$

For  $0 < q_{\text{ctr}} < q'_{\text{ctr}} = C(\gamma, \lambda) q_{\text{ctr}} < 1$ , the product of (53) and (54) reads

$$\begin{aligned} H_{\ell+1} Z_{\ell+1} &\leq C(\gamma, \lambda) q_{\text{ctr}} \left[ \left\| u_{\ell+1}^* - u_{\ell+1}^{m, n} \right\| + \gamma \eta_{\ell} \left( u_{\ell}^{m, n} \right) \right] \left[ \left\| z_{\ell+1}^* - z_{\ell+1}^{\mu, \nu} \right\| + \gamma \zeta_{\ell} \left( z_{\ell}^{\mu, \nu} \right) \right] \\ &= q'_{\text{ctr}} \left[ \left\| u_{\ell+1}^* - u_{\ell+1}^{m, n} \right\| + \gamma \eta_{\ell} \left( u_{\ell}^{m, n} \right) \right] \left[ \left\| z_{\ell+1}^* - z_{\ell+1}^{\mu, \nu} \right\| + \gamma \zeta_{\ell} \left( z_{\ell}^{\mu, \nu} \right) \right]. \end{aligned} \quad (55)$$

If (41) is satisfied, we obtain the same estimate with reversed roles in the derivation.

**Step 3 (quasi-monotonicity of  $H_\ell$  and  $Z_\ell$ ).** The C ea estimate (11), nestedness of the discrete spaces, reliability (A3), quasi-monotonicity (QM), stability (A1), and the definition (48) prove, for all  $\ell \leq \ell' \leq \ell'' \leq \underline{\ell}$  with  $(\ell, \underline{m}, \underline{n}) \in \mathcal{Q}^u$  and  $(\ell, \underline{\mu}, \underline{\nu}) \in \mathcal{Q}^z$ , that

$$\left\| u_{\ell'}^* - u_{\ell'}^* \right\| \stackrel{(11)}{\lesssim} \left\| u_{\ell'}^* - u_{\ell'}^* \right\| \stackrel{(A3)}{\lesssim} \eta_{\ell'} \left( u_{\ell'}^* \right) \stackrel{(QM)}{\lesssim} \eta_{\ell'} \left( u_{\ell'}^* \right) \stackrel{(A1)}{\lesssim} \eta_{\ell'} \left( u_{\ell'}^{m, n} \right) + \left\| u_{\ell'}^* - u_{\ell'}^{m, n} \right\| \stackrel{(48)}{\simeq} H_{\ell'}, \quad (56a)$$

$$\left\| z_{\ell'}^* - z_{\ell'}^* \right\| \stackrel{(11)}{\lesssim} \left\| z_{\ell'}^* - z_{\ell'}^* \right\| \stackrel{(A3)}{\lesssim} \zeta_{\ell'} \left( z_{\ell'}^* \right) \stackrel{(QM)}{\lesssim} \zeta_{\ell'} \left( z_{\ell'}^* \right) \stackrel{(A1)}{\lesssim} \zeta_{\ell'} \left( z_{\ell'}^{m, \nu} \right) + \left\| z_{\ell'}^* - z_{\ell'}^{m, \nu} \right\| \stackrel{(48)}{\simeq} Z_{\ell'}, \quad (56b)$$

where the hidden constants depend only on  $\gamma^{-1}$ ,  $C_{\text{C ea}}$ ,  $C_{\text{stab}}$ ,  $C_{\text{rel}}$ , and  $C_{\text{mon}}$ .

Similarly to (53), the inexact Zaronello contraction (29), stability (A1), and the stability estimate (36) yield for  $\ell < \ell' < \underline{\ell}$  and  $\lambda = \lambda_{\text{sym}} \lambda_{\text{alg}}$ ,

$$\begin{aligned}
\|u_\ell^* - u_\ell^{m,n}\| &\stackrel{(29)}{\leq} q_{\text{sym}}^* \|u_\ell^* - u_\ell^{m-1,n}\| + 2 \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda \eta_\ell(u_\ell^{m,n}) \\
&\stackrel{(28)}{\leq} q_{\text{sym}}^* q_{\text{sym}}^{m[\ell']-1} \|u_\ell^* - u_{\ell'-1}^{m,n}\| + 2 \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda \eta_\ell(u_\ell^{m,n}) \\
&\stackrel{(A1), (42)}{\leq} q_{\text{sym}}^* \|u_\ell^* - u_{\ell'-1}^{m,n}\| + 2 \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda \eta_{\ell'-1}(u_{\ell'-1}^{m,n}) + 2C_{\text{stab}} \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda \|u_{\ell'}^{m,n} - u_{\ell'-1}^{m,n}\| \\
&\stackrel{(36)}{\leq} \left( q_{\text{sym}}^* + 8C_{\text{stab}} \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda \right) \|u_\ell^* - u_{\ell'-1}^{m,n}\| + 2 \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda \eta_{\ell'-1}(u_{\ell'-1}^{m,n}) \\
&\stackrel{(A1)}{\leq} \left( q_{\text{sym}}^* + 10C_{\text{stab}} \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda \right) \|u_{\ell'-1}^* - u_{\ell'-1}^{m,n}\| \\
&\quad + \left( q_{\text{sym}}^* + 8C_{\text{stab}} \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda \right) \|u_\ell^* - u_{\ell'-1}^*\| + 2 \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda \eta_{\ell'-1}(u_{\ell'-1}^*) \\
&\stackrel{(56a)}{\leq} \left( q_{\text{sym}}^* + 10C_{\text{stab}} \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda \right) \|u_{\ell'-1}^* - u_{\ell'-1}^{m,n}\| \\
&\quad + \left( q_{\text{sym}}^* + 2 \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda C_{\text{mon}} [4 C_{\text{stab}} (1 + C_{\text{Céa}}) C_{\text{rel}} + 1] \right) \eta_\ell(u_\ell^*).
\end{aligned} \tag{57}$$

The choice of  $\lambda \leq \lambda^*$  with  $\lambda^*$  from (46) ensures

$$0 < q := q_{\text{sym}}^* + 10 C_{\text{stab}} \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda < 1. \tag{58}$$

With  $C := q_{\text{sym}}^* + 2 \frac{q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda C_{\text{mon}} [4 C_{\text{stab}} (1 + C_{\text{Céa}}) C_{\text{rel}} + 1]$ , a successive application of (57) and the geometric series shows

$$\begin{aligned}
\|u_{\ell+M}^* - u_{\ell+M}^{m,n}\| &\stackrel{(57)}{\leq} q \|u_{\ell+M-1}^* - u_{\ell+M-1}^{m,n}\| + C \eta_\ell(u_\ell^*) \\
&\stackrel{(57)}{\leq} q^M \|u_\ell^* - u_\ell^{m,n}\| + \left( C \sum_{j=0}^{M-1} q^j \right) \eta_\ell(u_\ell^*) \\
&\stackrel{(56a)}{\lesssim} \|u_\ell^* - u_\ell^{m,n}\| + \eta_\ell(u_\ell^*) \lesssim H_\ell \quad \forall \ell, M \in \mathbb{N}_0 \text{ with } \ell + M < \underline{\ell}.
\end{aligned} \tag{59}$$

Hence, we have quasi-monotonicity of the quasi-error

$$\begin{aligned}
H_{\ell+M} &\simeq \|u_{\ell+M}^* - u_{\ell+M}^{m,n}\| + \eta_{\ell+M}(u_{\ell+M}^{m,n}) \stackrel{(A1)}{\simeq} \|u_{\ell+M}^* - u_{\ell+M}^{m,n}\| + \eta_{\ell+M}(u_{\ell+M}^*) \\
&\stackrel{(56a)}{\lesssim} \|u_{\ell+M}^* - u_{\ell+M}^{m,n}\| + H_\ell \stackrel{59}{\lesssim} H_\ell \quad \forall \ell, M \in \mathbb{N}_0 \text{ with } \ell + M < \underline{\ell}.
\end{aligned} \tag{60a}$$

The same argument proves

$$Z_{\ell+M} \lesssim Z_\ell \quad \forall \ell, M \in \mathbb{N}_0 \text{ with } \ell + M < \underline{\ell}. \tag{60b}$$

**Step 4 (contraction of  $H_\ell Z_\ell$  up to tail-summable remainder).** Define

$$\begin{aligned}
R_\ell &:= \|u_{\ell+1}^* - u_\ell^*\| \left[ \|z_\ell^* - z_\ell^{\mu,\nu}\| + \|z_{\ell+1}^* - z_\ell^*\| + \gamma \zeta_\ell(z_\ell^{\mu,\nu}) \right] \\
&\quad + \|z_{\ell+1}^* - z_\ell^*\| \left[ \|u_\ell^* - u_\ell^{m,n}\| + \|u_{\ell+1}^* - u_\ell^*\| + \gamma \eta_\ell(u_\ell^{m,n}) \right].
\end{aligned}$$

The contraction (55) proves the quasi-contraction (49) via

$$\begin{aligned} \mathbf{H}_{\ell+1} \mathbf{Z}_{\ell+1} &\stackrel{(55)}{\leq} q'_{\text{ctr}} \left[ \left\| \mathbf{u}_{\ell+1}^* - \mathbf{u}_{\ell}^{m,n} \right\| + \gamma \eta_{\ell} \left( \mathbf{u}_{\ell}^{m,n} \right) \right] \left[ \left\| \mathbf{z}_{\ell+1}^* - \mathbf{z}_{\ell}^{H,v} \right\| + \gamma \zeta_{\ell} \left( \mathbf{z}_{\ell}^{H,v} \right) \right] \\ &\leq q'_{\text{ctr}} \left[ \left\| \mathbf{u}_{\ell}^* - \mathbf{u}_{\ell}^{m,n} \right\| + \left\| \mathbf{u}_{\ell+1}^* - \mathbf{u}_{\ell}^* \right\| + \gamma \eta_{\ell} \left( \mathbf{u}_{\ell}^{m,n} \right) \right] \\ &\quad \times \left[ \left\| \mathbf{z}_{\ell}^* - \mathbf{z}_{\ell}^{H,v} \right\| + \left\| \mathbf{z}_{\ell+1}^* - \mathbf{z}_{\ell}^* \right\| + \gamma \zeta_{\ell} \left( \mathbf{z}_{\ell}^{H,v} \right) \right] \\ &\leq q'_{\text{ctr}} \mathbf{H}_{\ell} \mathbf{Z}_{\ell} + q'_{\text{ctr}} R_{\ell}. \end{aligned}$$

The remainder term  $R_{\ell}$  can be estimated via (56) and the Young inequality by

$$R_{\ell}^2 \stackrel{(56)}{\lesssim} \left( \left\| \mathbf{u}_{\ell+1}^* - \mathbf{u}_{\ell}^* \right\| \mathbf{Z}_{\ell} + \left\| \mathbf{z}_{\ell+1}^* - \mathbf{z}_{\ell}^* \right\| \mathbf{H}_{\ell} \right)^2 \lesssim \left\| \mathbf{u}_{\ell+1}^* - \mathbf{u}_{\ell}^* \right\|^2 \mathbf{Z}_{\ell}^2 + \left\| \mathbf{z}_{\ell+1}^* - \mathbf{z}_{\ell}^* \right\|^2 \mathbf{H}_{\ell}^2. \quad (61)$$

Thus, the quasi-monotonicity (60) verifies

$$R_{\ell+M} \lesssim \mathbf{H}_{\ell+M} \mathbf{Z}_{\ell+M} \stackrel{(60)}{\lesssim} \mathbf{H}_{\ell} \mathbf{Z}_{\ell} \quad \forall \ell, M \in \mathbb{N} \text{ with } \ell + M < \underline{\ell}.$$

Quasi-orthogonality (A4), reliability (A3), and the estimates (56) imply, for all  $\ell, M \in \mathbb{N}_0$  with  $\ell + M < \underline{\ell}$ ,

$$\begin{aligned} \sum_{\ell'=\ell}^{\ell+M} \left\| \mathbf{u}_{\ell'+1}^* - \mathbf{u}_{\ell'}^* \right\|^2 &\stackrel{(A4)}{\lesssim} (M+1)^{1-\delta} \left\| \mathbf{u}^* - \mathbf{u}_{\ell}^* \right\|^2 \stackrel{(A3)}{\lesssim} (M+1)^{1-\delta} \eta_{\ell} \left( \mathbf{u}_{\ell}^* \right)^2 \stackrel{(56a)}{\lesssim} (M+1)^{1-\delta} \mathbf{H}_{\ell}^2, \\ \sum_{\ell'=\ell}^{\ell+M} \left\| \mathbf{z}_{\ell'+1}^* - \mathbf{z}_{\ell'}^* \right\|^2 &\stackrel{(A4)}{\lesssim} (M+1)^{1-\delta} \left\| \mathbf{z}^* - \mathbf{z}_{\ell}^* \right\|^2 \stackrel{(A3)}{\lesssim} (M+1)^{1-\delta} \zeta_{\ell} \left( \mathbf{z}_{\ell}^* \right)^2 \stackrel{(56b)}{\lesssim} (M+1)^{1-\delta} \mathbf{Z}_{\ell}^2. \end{aligned} \quad (62)$$

Using (61), the quasi-monotonicity (60), and (62), we conclude the proof of (50), for all  $\ell, M \in \mathbb{N}_0$  with  $\ell + M < \underline{\ell}$ ,

$$\begin{aligned} \sum_{\ell'=\ell}^{\ell+M} R_{\ell'}^2 &\stackrel{(61)}{\lesssim} \sum_{\ell'=\ell}^{\ell+M} \left\| \mathbf{u}_{\ell'+1}^* - \mathbf{u}_{\ell'}^* \right\|^2 \mathbf{Z}_{\ell'}^2 + \sum_{\ell'=\ell}^{\ell+M} \left\| \mathbf{z}_{\ell'+1}^* - \mathbf{z}_{\ell'}^* \right\|^2 \mathbf{H}_{\ell'}^2 \\ &\stackrel{(60)}{\lesssim} \mathbf{Z}_{\ell}^2 \sum_{\ell'=\ell}^{\ell+M} \left\| \mathbf{u}_{\ell'+1}^* - \mathbf{u}_{\ell'}^* \right\|^2 + \mathbf{H}_{\ell}^2 \sum_{\ell'=\ell}^{\ell+M} \left\| \mathbf{z}_{\ell'+1}^* - \mathbf{z}_{\ell'}^* \right\|^2 \stackrel{(62)}{\lesssim} (M+1)^{1-\delta} \mathbf{H}_{\ell}^2 \mathbf{Z}_{\ell}^2. \end{aligned}$$

□

The tail-summability in  $\ell$  provides the basis for the proof of tail-summability on the mesh level  $\ell$  together with the Zarantonello symmetrization index  $k$  for the final iterates of the algebraic solver. The main ingredients in the proof of tail-summability in  $(\ell, k)$  are Lemma 8 and the following quasi-contraction in the symmetrization index  $k$ .

**Lemma 9** (quasi-contraction of inexact Zarantonello symmetrization). *There holds*

$$\mathbf{H}_{\ell}^{k,j} \mathbf{Z}_{\ell}^{k,j} \lesssim q_{\text{sym}}^{k-k} \mathbf{H}_{\ell}^{k,j} \mathbf{Z}_{\ell}^{k,j} \quad \forall (\ell, k', j) \in \mathcal{Q} \text{ with } 0 \leq k \leq k' \leq \underline{k}[\ell], \quad (63)$$

$$\mathbf{H}_{\ell}^{0,j} \mathbf{Z}_{\ell}^{0,j} \lesssim \mathbf{H}_{\ell-1} \mathbf{Z}_{\ell-1} \quad \forall (\ell, 0, 0) \in \mathcal{Q} \text{ with } \ell \geq 1. \quad (64)$$

*Proof.* First, we note that the a posteriori error control (31) and the stopping criteria of the algebraic solver (19) and of the symmetrization (20) lead, for  $(\ell, \underline{m}, \underline{n}) \in \mathcal{Q}^u$ , to

$$\left\| \mathbf{u}_{\ell}^{m,*} - \mathbf{u}_{\ell}^{m,n} \right\| \stackrel{(31)}{\lesssim} \left\| \mathbf{u}_{\ell}^{m,n} - \mathbf{u}_{\ell}^{m,n-1} \right\| \stackrel{(19)}{\lesssim} \eta_{\ell} \left( \mathbf{u}_{\ell}^{m,n} \right) + \left\| \mathbf{u}_{\ell}^{m,n} - \mathbf{u}_{\ell}^{m,0} \right\| \stackrel{(20)}{\lesssim} \eta_{\ell} \left( \mathbf{u}_{\ell}^{m,n} \right) \lesssim \mathbf{H}_{\ell}.$$



Since the two notions of quasi-errors  $H_\ell$  and  $H_\ell^{k,j}$  only differ by the middle term  $\|u_\ell^{m,\star} - u_\ell^{m,n}\|$  and the fixed constant factor  $0 < \gamma < 1$ , this and the analogous estimate for the dual variable show

$$H_\ell \leq H_\ell^{k,j} \lesssim H_\ell, \quad Z_\ell \leq Z_\ell^{k,j} \lesssim Z_\ell \quad \forall (\ell, \underline{k}, \underline{j}) \in \mathcal{Q}. \quad (65)$$

For  $0 \leq k < k' < \underline{m}[\ell] < \underline{k}[\ell]$  (i.e., the primal iteration stops earlier than the dual iteration), the validity of the stopping criterion (19) for the algebraic solver and the failure of criterion (20) for the inexact Zarantonello symmetrization prove that

$$\begin{aligned} H_\ell^{k',n} &\stackrel{(31)}{\lesssim} \|u_\ell^\star - u_\ell^{k',n}\| + \|u_\ell^{k',n} - u_\ell^{k',n-1}\| + \eta_\ell(u_\ell^{k',n}) \\ &\stackrel{(19)}{\lesssim} \|u_\ell^\star - u_\ell^{k'-1,n}\| + \|u_\ell^{k',n} - u_\ell^{k'-1,n}\| + \eta_\ell(u_\ell^{k',n}) \\ &\stackrel{(20)}{\lesssim} \|u_\ell^\star - u_\ell^{k',n}\| + \|u_\ell^{k',n} - u_\ell^{k'-1,n}\| \\ &\stackrel{(32)}{\leq} \|u_\ell^\star - u_\ell^{k'-1,n}\| \stackrel{(28)}{\lesssim} q_{\text{sym}}^{k'-k} \|u_\ell^\star - u_\ell^{k,n}\| \lesssim q_{\text{sym}}^{k'-k} H_\ell^{k,n}. \end{aligned} \quad (66)$$

Moreover, for  $0 \leq k < k' = \underline{m}[\ell]$ , stability (A1) and the estimate (37) verify

$$\begin{aligned} H_\ell^{m,n} &\stackrel{(65)}{\simeq} \|u_\ell^\star - u_\ell^{m,n}\| + \eta_\ell(u_\ell^{m,n}) \stackrel{(A1)}{\lesssim} \|u_\ell^\star - u_\ell^{m,n}\| + \|u_\ell^{m,n} - u_\ell^{m-1,n}\| + \eta_\ell(u_\ell^{m-1,n}) \\ &\lesssim H_\ell^{m-1,n} + \|u_\ell^{m,n} - u_\ell^{m-1,n}\| \stackrel{(37)}{\lesssim} H_\ell^{m-1,n} \stackrel{(66)}{\lesssim} q_{\text{sym}}^{m[\ell]-1-k} H_\ell^{k,n} \simeq q_{\text{sym}}^{m[\ell]-k} H_\ell^{k,n}. \end{aligned}$$

For  $0 \leq k \leq \underline{m}[\ell] < k' \leq \underline{k}[\ell]$ , it follows  $H_\ell^{k',n} = H_\ell^{m,n} \lesssim q_{\text{sym}}^{m[\ell]-k} H_\ell^{k,n}$ . Finally, for  $\underline{m}[\ell] \leq k < k' \leq \underline{k}[\ell]$ , we have  $H_\ell^{k',n} = H_\ell^{m[\ell],n} = H_\ell^{k,n}$ . Notice that the same argumentation holds for the dual quasi-error  $Z_\ell^{k,v}$  in the remaining cases with  $\mu[\ell] < \underline{k}[\ell]$  (i.e., the dual iteration stops earlier than the primal iteration).

Since  $\underline{k}[\ell] = \underline{m}[\ell]$  or  $\underline{k}[\ell] = \mu[\ell]$  by definition, we obtain, for all  $(\ell, k', j) \in \mathcal{Q}$  with  $0 \leq k \leq k' \leq \underline{k}[\ell]$ ,

$$H_\ell^{k',j} \lesssim q_{\text{sym}}^{k'-k} H_\ell^{k,j} \quad \text{if } \underline{k}[\ell] = \underline{m}[\ell] \quad \text{or} \quad Z_\ell^{k',j} \lesssim q_{\text{sym}}^{k'-k} Z_\ell^{k,j} \quad \text{if } \underline{k}[\ell] = \mu[\ell].$$

Furthermore, there holds  $H_\ell^{k',j} \lesssim H_\ell^{k,j}$  and  $Z_\ell^{k',j} \lesssim Z_\ell^{k,j}$  in any case. This yields (63) via

$$H_\ell^{k',j} Z_\ell^{k',j} \lesssim q_{\text{sym}}^{k'-k} H_\ell^{k,j} Z_\ell^{k,j} \quad \forall (\ell, k', j) \in \mathcal{Q} \text{ with } 0 \leq k \leq k' \leq \underline{k}[\ell],$$

where the hidden constant depends only on  $C_{\text{stab}}$ ,  $\lambda_{\text{sym}}$ , and  $q_{\text{sym}}$ .

Nested iteration  $u_{\ell-1}^{m,n} = u_{\ell-1}^{0,n}$  and  $z_{\ell-1}^{\mu,v} = z_{\ell-1}^{0,v}$  and the estimates (56) yield, for all  $(\ell, 0, 0) \in \mathcal{Q}$  with  $\ell > 0$ ,

$$\begin{aligned} H_\ell^{0,j} &\stackrel{(65)}{\simeq} \|u_\ell^\star - u_{\ell-1}^{m,n}\| + \eta_\ell(u_{\ell-1}^{m,n}) \leq \|u_\ell^\star - u_{\ell-1}^\star\| + H_{\ell-1}^{k,j} \stackrel{(56)}{\lesssim} H_{\ell-1} + H_{\ell-1}^{k,j} \stackrel{(65)}{\simeq} H_{\ell-1}, \\ Z_\ell^{0,j} &\stackrel{(65)}{\simeq} \|z_\ell^\star - z_{\ell-1}^{\mu,v}\| + \zeta_\ell(z_{\ell-1}^{\mu,v}) \leq \|z_\ell^\star - z_{\ell-1}^\star\| + Z_{\ell-1}^{k,j} \stackrel{(56)}{\lesssim} Z_{\ell-1} + Z_{\ell-1}^{k,j} \stackrel{(65)}{\simeq} Z_{\ell-1}. \end{aligned}$$

A multiplication of the two previous estimates proves (64).  $\square$

Finally, the quasi-contraction in  $(\ell, k)$  from Lemma 9 together with a quasi-contraction in the algebraic solver index  $j$  leads to tail-summability in  $(\ell, k, j)$ .

**Lemma 10** (quasi-contraction and stability by algebraic solver). *There holds*

$$H_\ell^{k,j'} Z_\ell^{k,j'} \lesssim q_{\text{alg}}^{j'-j} H_\ell^{k,j} Z_\ell^{k,j} \quad \forall (\ell, k, j') \in \mathcal{Q} \text{ with } 0 \leq j \leq j' \leq \underline{j}[\ell, k] \quad (67)$$

and, with the abbreviation  $(m-1)_+ := \max\{m-1, 0\}$ ,

$$H_\ell^{m,0} \leq 3 H_\ell^{(m-1)_+, \underline{n}} \quad \text{and} \quad Z_\ell^{\mu,0} \leq 3 Z_\ell^{(\mu-1)_+, \underline{\nu}} \quad \forall (\ell, m, 0) \in \mathcal{Q}^u, (\ell, \mu, 0) \in \mathcal{Q}^z. \quad (68)$$

*Proof.* We recall that  $u_\ell^{0,0} = u_\ell^{0,\underline{n}} = u_\ell^{0,\star}$  by definition and, hence,  $H_\ell^{0,0} = H_\ell^{0,\underline{n}} = H_\ell^{0,\underline{j}}$ . Nested iteration  $u_\ell^{m,0} = u_\ell^{m-1,\underline{n}}$  implies that

$$\|u_\ell^{m,\star} - u_\ell^{m,0}\| \stackrel{(30)}{\leq} (q_{\text{sym}}^\star + 1) \|u_\ell^\star - u_\ell^{m-1,\underline{n}}\| \leq 2 H_\ell^{m-1,\underline{j}} \quad \forall (\ell, m, 0) \in \mathcal{Q}^u.$$

Therewith, we derive (68).

The combination of a posteriori error control (30) for the exact Zarantonello iteration, for the algebraic solver (31), and the failure of the stopping criterion (19) in Algorithm 1(I.b.ii) for the algebraic solver proves, for  $0 \leq j < j' < \underline{n}[\ell, m] < \underline{j}[\ell, m]$ ,

$$\begin{aligned} H_\ell^{m,j'} &\leq \|u_\ell^\star - u_\ell^{m,\star}\| + 2 \|u_\ell^{m,\star} - u_\ell^{m,j}\| + \eta_\ell(u_\ell^{m,j}) \\ &\stackrel{(30)}{\leq} \frac{q_{\text{sym}}^\star}{1 - q_{\text{sym}}^\star} \|u_\ell^{m,j} - u_\ell^{m-1,\underline{j}}\| + \left(2 + \frac{q_{\text{sym}}^\star}{1 - q_{\text{sym}}^\star}\right) \|u_\ell^{m,\star} - u_\ell^{m,j}\| + \eta_\ell(u_\ell^{m,j}) \\ &\stackrel{(31)}{\lesssim} \|u_\ell^{m,j} - u_\ell^{m-1,\underline{j}}\| + \|u_\ell^{m,j} - u_\ell^{m,j-1}\| + \eta_\ell(u_\ell^{m,j}) \stackrel{(19)}{\lesssim} \|u_\ell^{m,j} - u_\ell^{m,j-1}\| \\ &\stackrel{(31)}{\lesssim} \|u_\ell^{m,\star} - u_\ell^{m,j-1}\| \stackrel{(15)}{\leq} q_{\text{alg}}^{(j-1)-j} \|u_\ell^{m,\star} - u_\ell^{m,j}\| \lesssim q_{\text{alg}}^{j-j} H_\ell^{m,j}. \end{aligned} \quad (69)$$

For  $0 \leq j < \underline{n}[\ell, m] \leq j' \leq \underline{j}[\ell, m]$ , stability (A1) and contraction of the algebraic solver (15) verify that

$$\begin{aligned} H_\ell^{m,j'} &= H_\ell^{m,\underline{n}} \stackrel{(15)}{\leq} \|u_\ell^\star - u_\ell^{m,\underline{n}-1}\| + \|u_\ell^{m,\underline{n}} - u_\ell^{m,\underline{n}-1}\| + q_{\text{alg}} \|u_\ell^{m,\star} - u_\ell^{m,\underline{n}-1}\| + \eta_\ell(u_\ell^{m,\underline{n}}) \\ &\stackrel{(A1)}{\leq} H_\ell^{m,\underline{n}-1} + (2 + C_{\text{stab}}) \|u_\ell^{m,\underline{n}} - u_\ell^{m,\underline{n}-1}\| \\ &\stackrel{(31)}{\lesssim} H_\ell^{m,\underline{n}-1} + \|u_\ell^{m,\star} - u_\ell^{m,\underline{n}-1}\| \lesssim H_\ell^{m,\underline{n}-1} \stackrel{(69)}{\lesssim} q_{\text{alg}}^{\underline{n}[\ell]-j} H_\ell^{m,j}. \end{aligned}$$

For  $\underline{n}[\ell, m] \leq j < j' \leq \underline{j}[\ell, m]$ , it holds that  $H_\ell^{m,j} = H_\ell^{m,\underline{n}} = H_\ell^{m,j'}$ . Since  $\underline{j}[\ell, k] = \underline{n}[\ell, k]$  or  $\underline{j}[\ell, k] = \underline{\nu}[\ell, k]$ , we have, for all  $(\ell, k, j') \in \mathcal{Q}$  with  $0 \leq j \leq j' \leq \underline{j}[\ell, k]$ ,

$$H_\ell^{k,j} \lesssim q_{\text{alg}}^{j-j'} H_\ell^{k,j'} \quad \text{if } \underline{j}[\ell, k] = \underline{n}[\ell, k] \quad \text{or} \quad Z_\ell^{k,j} \lesssim q_{\text{alg}}^{j-j'} Z_\ell^{k,j'} \quad \text{if } \underline{j}[\ell, k] = \underline{\nu}[\ell, k].$$

Furthermore, we have  $H_\ell^{k,j} \lesssim H_\ell^{k,j'}$  and  $Z_\ell^{k,j} \lesssim Z_\ell^{k,j'}$  in any case. Hence, we obtain

$$H_\ell^{k,j} Z_\ell^{k,j} \lesssim q_{\text{alg}}^{j-j'} H_\ell^{k,j'} Z_\ell^{k,j'} \quad \forall (\ell, k, j) \in \mathcal{Q} \text{ with } 0 \leq j' \leq j \leq \underline{j}[\ell, k],$$

where the hidden constant depends only on  $q_{\text{sym}}^\star$ ,  $\lambda_{\text{sym}}$ ,  $q_{\text{alg}}$ ,  $\lambda_{\text{alg}}$ , and  $C_{\text{stab}}$ .  $\square$

Ultimately, synthesizing the preceding lemmas yields tail-summability of the quasi-error product and thus leads to the following proof of Theorem 1.

*Proof of Theorem 1.* The proof consists of four steps.

**Step 1 (tail-summability in mesh level  $\ell$ ).** We apply the tail-summability criterion from [33, Lemma 5] to the sequences  $a_\ell := H_\ell Z_\ell$  and  $b_\ell := q_{\text{ctr}}' R_\ell$ . Therein, it is shown that  $R$ -linear convergence is equivalent to tail-summability and that, for tail-summability, it is sufficient to guarantee

$$a_{\ell+1} \leq qa_{\ell} + b_{\ell}, \quad b_{\ell+M} \leq C_1 a_{\ell}, \quad \sum_{\ell'=\ell}^{\ell+M} b_{\ell'}^2 \leq C_2 (M+1)^{1-\delta} a_{\ell}^2 \quad \forall \ell, M \in \mathbb{N}_0. \quad (70)$$

Indeed, contraction up to a remainder from (49), the estimate of the remainder from (50), and the quasi-monotonicity of  $H_{\ell}$  and  $Z_{\ell}$  from (60) validate the assumptions of the tail-summability criterion (70) and lead to tail-summability

$$\sum_{\ell'=\ell+1}^{\ell-1} H_{\ell'} Z_{\ell'} \lesssim H_{\ell} Z_{\ell} \quad \forall (\ell, \underline{k}, \underline{j}) \in Q. \quad (71)$$

**Step 2 (tail-summability in  $(\ell, \underline{k})$ ).** For  $(\ell, \underline{k}, \underline{j}) \in Q$ , the estimates (63), (64) and the geometric series prove tail-summability

$$\begin{aligned} \sum_{\substack{(\ell', \underline{k}', \underline{j}') \in Q \\ |\ell', \underline{k}', \underline{j}'| > |\ell, \underline{k}, \underline{j}|}} H_{\ell'}^{k', j'} Z_{\ell'}^{k', j'} &= \sum_{k'=k+1}^{\underline{k}[\ell]} H_{\ell'}^{k', j'} Z_{\ell'}^{k', j'} + \sum_{\ell'=\ell+1}^{\ell} \sum_{k'=0}^{\underline{k}[\ell']} H_{\ell'}^{k', j'} Z_{\ell'}^{k', j'} \\ &\stackrel{(63)}{\lesssim} H_{\ell}^{k, j} Z_{\ell}^{k, j} + \sum_{\ell'=\ell+1}^{\ell} H_{\ell'}^{0, j} Z_{\ell'}^{0, j} \stackrel{(64)}{\lesssim} H_{\ell}^{k, j} Z_{\ell}^{k, j} + \sum_{\ell'=\ell}^{\ell-1} H_{\ell'} Z_{\ell'} \\ &\stackrel{(71)}{\lesssim} H_{\ell}^{k, j} Z_{\ell}^{k, j} + H_{\ell} Z_{\ell} \stackrel{(65)}{\lesssim} H_{\ell}^{k, j} Z_{\ell}^{k, j}. \end{aligned} \quad (72)$$

**Step 3 (tail-summability in  $(\ell, \underline{k}, \underline{j})$ ).** Finally, for all  $(\ell, \underline{k}, \underline{j}) \in Q$ , we observe that

$$\begin{aligned} \sum_{\substack{(\ell', \underline{k}', \underline{j}') \in Q \\ |\ell', \underline{k}', \underline{j}'| > |\ell, \underline{k}, \underline{j}|}} H_{\ell'}^{k', j'} Z_{\ell'}^{k', j'} &= \sum_{j'=j+1}^{\underline{j}[\ell, \underline{k}]} H_{\ell'}^{k, j'} Z_{\ell'}^{k, j'} + \sum_{k'=k+1}^{\underline{k}[\ell]} \sum_{j'=0}^{\underline{j}[\ell, \underline{k}']} H_{\ell'}^{k', j'} Z_{\ell'}^{k', j'} + \sum_{\ell'=\ell+1}^{\ell} \sum_{k'=0}^{\underline{k}[\ell']} \sum_{j'=0}^{\underline{j}[\ell', \underline{k}']} H_{\ell'}^{k', j'} Z_{\ell'}^{k', j'} \\ &\stackrel{(67)}{\lesssim} H_{\ell}^{k, j} Z_{\ell}^{k, j} + \sum_{k'=k+1}^{\underline{k}[\ell]} H_{\ell}^{k', 0} Z_{\ell}^{k', 0} + \sum_{\ell'=\ell+1}^{\ell} \sum_{k'=0}^{\underline{k}[\ell']} H_{\ell'}^{k', 0} Z_{\ell'}^{k', 0} \\ &\stackrel{(68)}{\lesssim} H_{\ell}^{k, j} Z_{\ell}^{k, j} + \sum_{\substack{(\ell', \underline{k}', \underline{j}') \in Q \\ |\ell', \underline{k}', \underline{j}'| > |\ell, \underline{k}, \underline{j}|}} H_{\ell'}^{k', j'} Z_{\ell'}^{k', j'} \stackrel{(72)}{\lesssim} H_{\ell}^{k, j} Z_{\ell}^{k, j} + H_{\ell}^{k, j} Z_{\ell}^{k, j} \stackrel{(67)}{\lesssim} H_{\ell}^{k, j} Z_{\ell}^{k, j}. \end{aligned}$$

**Step 4.** Since the index set  $Q$  is linearly ordered with respect to the total step counter  $|\cdot, \cdot, \cdot|$ , tail-summability in Step 3 and the equivalence of tail-summability and  $R$ -linear convergence from [33, Lemma 10] conclude the proof of (47) in Theorem 1.  $\square$

## 6 Optimal complexity of Algorithm 1

Full linear convergence (47) has a simple but crucial consequence. Using a geometric series argument, one can prove that the cumulative computational cost up to a given level is bounded by the cost of the said level; see [33, Corollary 14], where only the primal quasi-error  $H_{\ell}^{k, j}$  has to be replaced by the quasi-error product  $H_{\ell}^{k, j} Z_{\ell}^{k, j}$ . As a consequence, the convergence rates with respect to the number of degrees of freedom (defined as  $M(r)$  in (73) below) and the rates with respect to the overall computational cost (cf. (26) and the discussion following the statement of Algorithm 1) coincide.

**Corollary 1** (rates = complexity [33, Corollary 14]). *Suppose the assumptions of Theorem 1. For all  $r > 0$ , the output  $(\mathcal{T}_{\ell})_{\ell \in \mathbb{N}_0}$  of Algorithm 1 satisfies*

$$M(r) := \sup_{(\ell,k,j) \in \mathcal{Q}} (\#\mathcal{T}_\ell)^r H_\ell^{k,j} Z_\ell^{k,j} \leq \sup_{(\ell,k,j) \in \mathcal{Q}} \left( \sum_{\substack{(\ell',k',j') \in \mathcal{Q} \\ |\ell',k',j'| \leq |\ell,k,j|}} \#\mathcal{T}_{\ell'} \right)^r H_\ell^{k,j} Z_\ell^{k,j} \leq C_{\text{cost}}(r)M(r) \quad (73)$$

with the constant  $C_{\text{cost}}(r) := C_{\text{lin}}/(1 - q_{\text{lin}}^{1/r})^r > 0$ .

While Theorem 1 only concerns  $R$ -linear convergence, a sufficiently small choice of the adaptivity parameters  $\theta$ ,  $\lambda_{\text{sym}}$ , and  $\lambda_{\text{alg}}$  even guarantees the optimal convergence rate  $r = s + t$  with respect to computational cost, i.e., the overall computational time. Here, we suppose that the primal solution  $u^*$  to (5) can be approximated at rate  $s$  and the dual solution  $z^*$  to (8) can be approximated at rate  $t$ . To formalize this idea, we introduce the notion of approximation classes [3]–[5], [9]. For  $s, t > 0$ , define

$$\|u^*\|_{\mathbb{A}_s} := \sup_{N \in \mathbb{N}_0} \left( (N+1)^s \min_{\mathcal{T}_{\text{opt}} \in \mathbb{T}_N} \eta_{\text{opt}}(u_{\text{opt}}^*) \right), \quad \|z^*\|_{\mathbb{A}_t} := \sup_{N \in \mathbb{N}_0} \left( (N+1)^t \min_{\mathcal{T}_{\text{opt}} \in \mathbb{T}_N} \zeta_{\text{opt}}(z_{\text{opt}}^*) \right),$$

where  $\eta_{\text{opt}}(\cdot)$  and  $\zeta_{\text{opt}}(\cdot)$  denote the estimator values for the exact discrete solutions  $u_{\text{opt}}^*$  and  $z_{\text{opt}}^*$  on the unavailable optimal triangulations  $\mathcal{T}_{\text{opt}} \in \mathbb{T}_N(\mathcal{T})$ . We stress that  $\|u^*\|_{\mathbb{A}_s}$  and  $\|z^*\|_{\mathbb{A}_t}$  can equivalently be defined by energy error plus data oscillations [8], [9].

**Theorem 2** (optimal complexity). *Suppose that the estimators  $\eta$  and  $\zeta$  satisfy (A1)–(A3<sup>+</sup>) and (QM) and suppose quasi-orthogonality (A4). Recall  $\lambda_{\text{alg}}^*$  from Lemma 3 and  $\lambda^*$  from (46) in Theorem 1. Define the constants*

$$\lambda_{\text{sym}}^* := \min \left\{ 1, C_{\text{stab}}^{-1} C_{\text{alg}}^{-1} \right\} \leq 1 \quad \text{with} \quad C_{\text{alg}} := \frac{1}{1 - q_{\text{sym}}^*} \left( \frac{2 q_{\text{alg}}}{1 - q_{\text{alg}}} \lambda_{\text{alg}}^* + q_{\text{sym}}^* \right), \quad (74)$$

$$\theta^* := (1 + C_{\text{stab}}^2 C_{\text{rel}}^2)^{-1} < 1.$$

Suppose that  $\theta$ ,  $\lambda_{\text{sym}}$ , and  $\lambda_{\text{alg}}$  are sufficiently small in the sense of

$$0 < \lambda_{\text{alg}} \leq \lambda_{\text{alg}}^*, \quad 0 < \lambda_{\text{sym}} < \lambda_{\text{sym}}^*, \quad \lambda_{\text{alg}} \lambda_{\text{sym}} < \lambda^*, \quad (75)$$

$$0 < \theta_{\text{mark}} := \frac{(\theta^{1/2} + \lambda_{\text{sym}}/\lambda_{\text{sym}}^*)^2}{(1 - \lambda_{\text{sym}}/\lambda_{\text{sym}}^*)^2} < \theta^* < 1.$$

Then, Algorithm 1 guarantees, for all  $s, t > 0$ , that

$$\sup_{(\ell,k,j) \in \mathcal{Q}} \left( \sum_{\substack{(\ell',k',j') \in \mathcal{Q} \\ |\ell',k',j'| \leq |\ell,k,j|}} \#\mathcal{T}_{\ell'} \right)^{s+t} H_\ell^{k,j} Z_\ell^{k,j} \leq C_{\text{opt}} \max \left\{ \|u^*\|_{\mathbb{A}_s}, \|z^*\|_{\mathbb{A}_t}, H_0^{0,0} Z_0^{0,0} \right\}. \quad (76)$$

The constant  $C_{\text{opt}}$  depends only on  $C_{\text{stab}}$ ,  $C_{\text{rel}}$ ,  $C_{\text{drel}}$ ,  $C_{\text{mark}}$ ,  $C_{\text{mesh}}$ ,  $C_{\text{lin}}$ ,  $q_{\text{lin}}$ ,  $\#\mathcal{T}_0$ , and  $s + t$ . In particular, there holds optimal complexity of Algorithm 1.

The proof of Theorem 2 employs the following result from [50] providing estimator equivalence between the (unavailable) estimators for the exact discrete solutions  $u_\ell^*$ ,  $z_\ell^*$  and the estimators at the computed approximations  $u_\ell^{m,n}$ ,  $z_\ell^{\mu,\nu}$ .

**Lemma 11** (estimator equivalence [50, Lemma 15]). *Recall the constants  $\lambda_{\text{sym}}^*$ ,  $C_{\text{alg}} > 0$  from (74) and  $\lambda_{\text{alg}}^* > 0$  from Lemma 3. Then, for all  $0 < \theta \leq 1$ ,  $0 < \lambda_{\text{alg}} \leq \lambda_{\text{alg}}^*$ ,  $0 < \lambda_{\text{sym}} < \lambda_{\text{sym}}^*$ , it holds that*

$$\begin{aligned} (1 - \lambda_{\text{sym}}/\lambda_{\text{sym}}^*) \eta_{\ell} \left( u_{\ell}^{\underline{m}, \underline{n}} \right) &\leq \eta_{\ell} \left( u_{\ell}^* \right) \leq (1 + \lambda_{\text{sym}}/\lambda_{\text{sym}}^*) \eta_{\ell} \left( u_{\ell}^{\underline{m}, \underline{n}} \right) \quad \forall (\ell, \underline{m}, \underline{n}) \in \mathcal{Q}^u, \\ (1 - \lambda_{\text{sym}}/\lambda_{\text{sym}}^*) \zeta_{\ell} \left( z_{\ell}^{\underline{\mu}, \underline{\nu}} \right) &\leq \zeta_{\ell} \left( z_{\ell}^* \right) \leq (1 + \lambda_{\text{sym}}/\lambda_{\text{sym}}^*) \zeta_{\ell} \left( z_{\ell}^{\underline{\mu}, \underline{\nu}} \right) \quad \forall (\ell, \underline{\mu}, \underline{\nu}) \in \mathcal{Q}^z. \end{aligned} \quad (77)$$

*Proof of Theorem 2.* By Corollary 1, it suffices to prove that, for any  $s, t > 0$ ,

$$\sup_{(\ell, k, j) \in \mathcal{Q}} (\#\mathcal{T}_{\ell})^{s+t} H_{\ell}^{k, j} Z_{\ell}^{k, j} \lesssim \max \left\{ \|u^*\|_{\mathbb{A}_s} \|z^*\|_{\mathbb{A}_t}, H_0^{0,0} Z_0^{0,0} \right\}. \quad (78)$$

Since the inequality becomes trivial if either  $\|u^*\|_{\mathbb{A}_s} = \infty$  or  $\|z^*\|_{\mathbb{A}_t} = \infty$ , we may assume  $\|u^*\|_{\mathbb{A}_s} \|z^*\|_{\mathbb{A}_t} < \infty$ . The proof consists of three steps.

**Step 1.** With  $0 < \theta_{\text{mark}} := (\theta^{1/2} + \lambda_{\text{sym}}/\lambda_{\text{sym}}^*)^2 (1 - \lambda_{\text{sym}}/\lambda_{\text{sym}}^*)^{-2} < \theta^*$ , the validity of (A3<sup>+</sup>) for both estimators and [16, Lemma 14] guarantee the existence of sets  $\mathcal{R}_{\ell'} \subseteq \mathcal{T}_{\ell'}$  with  $0 \leq \ell' < \underline{\ell}$  such that

$$\#\mathcal{R}_{\ell'} \lesssim (\|u^*\|_{\mathbb{A}_s} \|z^*\|_{\mathbb{A}_t})^{1/(s+t)} [\eta_{\ell'}(u_{\ell'}^*) \zeta_{\ell'}(z_{\ell'}^*)]^{-1/(s+t)}, \quad (79a)$$

$$\theta_{\text{mark}} \eta_{\ell'}(u_{\ell'}^*) \leq \eta_{\ell'}(\mathcal{R}_{\ell'}, u_{\ell'}^*) \quad \text{or} \quad \theta_{\text{mark}} \zeta_{\ell'}(z_{\ell'}^*) \leq \zeta_{\ell'}(\mathcal{R}_{\ell'}, z_{\ell'}^*). \quad (79b)$$

For  $0 \leq \ell' < \underline{\ell}$ , the estimator equivalence (77) in Lemma 11 leads to

$$(1 - \lambda_{\text{sym}}/\lambda_{\text{sym}}^*) \eta_{\ell'} \left( u_{\ell'}^{\underline{m}, \underline{n}} \right) \leq \eta_{\ell'} \left( u_{\ell'}^* \right), \quad (1 - \lambda_{\text{sym}}/\lambda_{\text{sym}}^*) \zeta_{\ell'} \left( z_{\ell'}^{\underline{\mu}, \underline{\nu}} \right) \leq \zeta_{\ell'} \left( z_{\ell'}^* \right)$$

and consequently with (79a) to

$$\#\mathcal{R}_{\ell'} \lesssim (\|u^*\|_{\mathbb{A}_s} \|z^*\|_{\mathbb{A}_t})^{1/(s+t)} \left[ \eta_{\ell'} \left( u_{\ell'}^{\underline{m}, \underline{n}} \right) \zeta_{\ell'} \left( z_{\ell'}^{\underline{\mu}, \underline{\nu}} \right) \right]^{-1/(s+t)}. \quad (80)$$

Note that the stopping criteria (20) and (22) lead to

$$H_{\ell'} \simeq \|u_{\ell'}^* - u_{\ell'}^{\underline{m}, \underline{n}}\| + \eta_{\ell'} \left( u_{\ell'}^{\underline{m}, \underline{n}} \right) \stackrel{(20)}{\lesssim} \eta_{\ell'} \left( u_{\ell'}^{\underline{m}, \underline{n}} \right), \quad Z_{\ell'} \simeq \|z_{\ell'}^* - z_{\ell'}^{\underline{\mu}, \underline{\nu}}\| + \zeta_{\ell'} \left( z_{\ell'}^{\underline{\mu}, \underline{\nu}} \right) \stackrel{(22)}{\lesssim} \zeta_{\ell'} \left( z_{\ell'}^{\underline{\mu}, \underline{\nu}} \right)$$

and with (64) to

$$H_{\ell'+1}^{0, j} Z_{\ell'+1}^{0, j} \stackrel{(64)}{\lesssim} H_{\ell'} Z_{\ell'} \lesssim \eta_{\ell'} \left( u_{\ell'}^{\underline{m}, \underline{n}} \right) \zeta_{\ell'} \left( z_{\ell'}^{\underline{\mu}, \underline{\nu}} \right). \quad (81)$$

Hence, the combination of (80) and (81) reads

$$\#\mathcal{R}_{\ell'} \lesssim (\|u^*\|_{\mathbb{A}_s} \|z^*\|_{\mathbb{A}_t})^{1/(s+t)} \left[ H_{\ell'+1}^{0, j} Z_{\ell'+1}^{0, j} \right]^{-1/(s+t)}. \quad (82)$$

**Step 2.** Recall from [29, Theorem 8] that the set  $\mathcal{R}_{\ell'}$  satisfies the Dörfler criterion from Algorithm 1 with the same parameter  $\theta$ . The quasi-minimality of  $\mathcal{M}_{\ell'}$  implies

$$\#\mathcal{M}_{\ell'} \leq C_{\text{mark}} \#\mathcal{R}_{\ell'} \quad \forall 0 \leq \ell' < \underline{\ell} \quad (83)$$

with the constant  $C_{\text{mark}} \geq 1$  from Algorithm 1.

**Step 3.** Let  $(\ell, k, j) \in \mathcal{Q}$ . Full linear convergence (47) from Theorem 1 yields that

$$\sum_{\substack{(\ell', k', j') \in \mathcal{Q} \\ |\ell', k', j'| \leq |\ell, k, j|}} \left( H_{\ell'}^{k', j'} Z_{\ell'}^{k', j'} \right)^{-1/(s+t)} \stackrel{(47)}{\lesssim} \left( H_{\ell}^{k, j} Z_{\ell}^{k, j} \right)^{-1/(s+t)} \sum_{\substack{(\ell', k', j') \in \mathcal{Q} \\ |\ell', k', j'| \leq |\ell, k, j|}} (q_{\text{lin}}^{1/s})^{|\ell', k, j| - |\ell', k', j'|} \lesssim \left( H_{\ell}^{k, j} Z_{\ell}^{k, j} \right)^{-1/(s+t)}. \quad (84)$$

NVB refinement satisfies the mesh-closure estimate [9, Eq. (2.9)] reading,

$$\#\mathcal{T}_{\ell} - \#\mathcal{T}_0 \leq C_{\text{mesh}} \sum_{\ell'=0}^{\ell-1} \#\mathcal{M}_{\ell'} \quad \forall \ell \geq 1, \quad (85)$$

where  $C_{\text{mesh}} > 1$  depends only on  $\mathcal{T}_0$ . Thus, for  $(\ell, k, j) \in \mathcal{Q}$ , we have by the mesh-closure estimate (85), quasi-optimality of Dörfler marking (83), and the result (84) that

$$\begin{aligned} \#\mathcal{T}_\ell - \#\mathcal{T}_0 &\stackrel{(85)}{\lesssim} \sum_{\ell'=0}^{\ell-1} \#\mathcal{M}_{\ell'} \stackrel{(83)}{\lesssim} \sum_{\ell'=0}^{\ell-1} \#\mathcal{R}_{\ell'} \stackrel{(82)}{\lesssim} (\|u^*\|_{\mathbb{A}_s} \|z^*\|_{\mathbb{A}_t})^{1/(s+t)} \sum_{\ell'=0}^{\ell-1} \left( H_{\ell'+1}^{0,j} Z_{\ell'+1}^{0,j} \right)^{-1/(s+t)} \\ &\leq (\|u^*\|_{\mathbb{A}_s} \|z^*\|_{\mathbb{A}_t})^{1/(s+t)} \sum_{\substack{(\ell', k', j') \in \mathcal{Q} \\ |\ell', k', j'| \leq |\ell, k, j|}} \left( H_{\ell'}^{k', j'} Z_{\ell'}^{k', j'} \right)^{-1/(s+t)} \\ &\stackrel{(84)}{\lesssim} (\|u^*\|_{\mathbb{A}_s} \|z^*\|_{\mathbb{A}_t})^{1/(s+t)} \left( H_\ell^{k, j} Z_\ell^{k, j} \right)^{-1/(s+t)}. \end{aligned}$$

Rearranging the terms and noting that  $1 \leq \#\mathcal{T}_\ell - \#\mathcal{T}_0$  implies  $\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1 \leq 2(\#\mathcal{T}_\ell - \#\mathcal{T}_0)$ , we obtain, for  $\ell > 0$ , that

$$(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^{s+t} H_\ell^{k, j} Z_\ell^{k, j} \lesssim \|u^*\|_{\mathbb{A}_s} \|z^*\|_{\mathbb{A}_t}. \quad (86a)$$

Moreover, full linear convergence (47) proves that

$$(\#\mathcal{T}_0 - \#\mathcal{T}_0 + 1)^{s+t} H_0^{k, j} Z_0^{k, j} = H_0^{k, j} Z_0^{k, j} \lesssim H_0^{0,0} Z_0^{0,0}. \quad (86b)$$

We recall from [35, Lemma 22] that, for all  $\mathcal{T}_\ell \in \mathbb{T}$ , it holds

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1 \leq \#\mathcal{T}_\ell \leq \#\mathcal{T}_0 (\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1). \quad (87)$$

This shows, for all  $(\ell, k, j) \in \mathcal{Q}$ ,

$$(\#\mathcal{T}_\ell)^{s+t} H_\ell^{k, j} Z_\ell^{k, j} \stackrel{(87)}{\lesssim} (\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^{s+t} H_\ell^{k, j} Z_\ell^{k, j} \stackrel{(86)}{\lesssim} \max \left\{ \|u^*\|_{\mathbb{A}_s} \|z^*\|_{\mathbb{A}_t}, H_0^{0,0} Z_0^{0,0} \right\}$$

and concludes the proof of (78).  $\square$

## 7 Numerical examples

In this section, we present numerical experiments using the open source software package MooAFEM [51].<sup>1</sup> In the following, Steps (I) and (II) of Algorithm 1 employ the optimal *hp*-robust local multigrid method from [32] as an algebraic solver. If not explicitly stated otherwise, we choose the parameters  $\theta = 0.5$ ,  $\delta = 0.5$ ,  $\lambda_{\text{sym}} = \lambda_{\text{alg}} = 0.7$  in Algorithm 1 throughout the numerical experiments.

### 7.1 Singularity in the goal functional

The first model problem is a nonsymmetric variant of the benchmark problem from [29, Section 4.1] with a singularity only in the goal functional. On the unit square  $\Omega = (0, 1)^2 \subset \mathbb{R}^2$ , we consider

$$-\Delta u^* + x \cdot \nabla u^* + u^* = f \quad \text{in } \Omega \quad \text{subject to } u^* = 0 \quad \text{on } \partial\Omega, \quad (88)$$

where the right-hand side is chosen such that the exact solution  $u^*$  reads

$$u^*(x) = x_1 x_2 (1 - x_1) (1 - x_2).$$

<sup>1</sup> All experiments presented in this paper are reproducible with the openly available software package under <https://www.tuwien.at/mg/asc/praetorius/software/mooafem>.

Consider  $g = 0$  and  $\mathbf{g} = \chi_K(1, 0)^\top$  in the quantity of interest

$$G(u^*) := \int_K \partial_{x_1} u^* dx = \frac{11}{960} \quad \text{with } K := \text{conv}\{(1/2, 1), (1, 1/2), (1, 1)\}.$$

Figure 2 (left) displays a mesh generated by Algorithm 1 and the support  $K$  of  $\mathbf{g}$ . The error estimator captures and resolves the two point singularities induced by  $G$ .

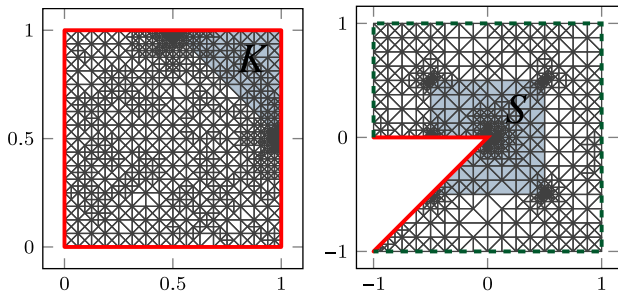
### 7.2 Geometric singularity and strong convection

The second benchmark problem investigates  $\Omega = (-1, 1)^2 \setminus \text{conv}\{(0, 0), (-1, 0), (-1, -1)\} \subset \mathbb{R}^2$  with the Dirichlet boundary  $\Gamma_D = \text{conv}\{(-1, 0), (0, 0)\} \cup \text{conv}\{(0, 0), (-1, -1)\}$  and Neumann boundary  $\Gamma_N = \partial\Omega \setminus \Gamma_D$ ; see Figure 2 (right) for a visualization of the geometry. We consider

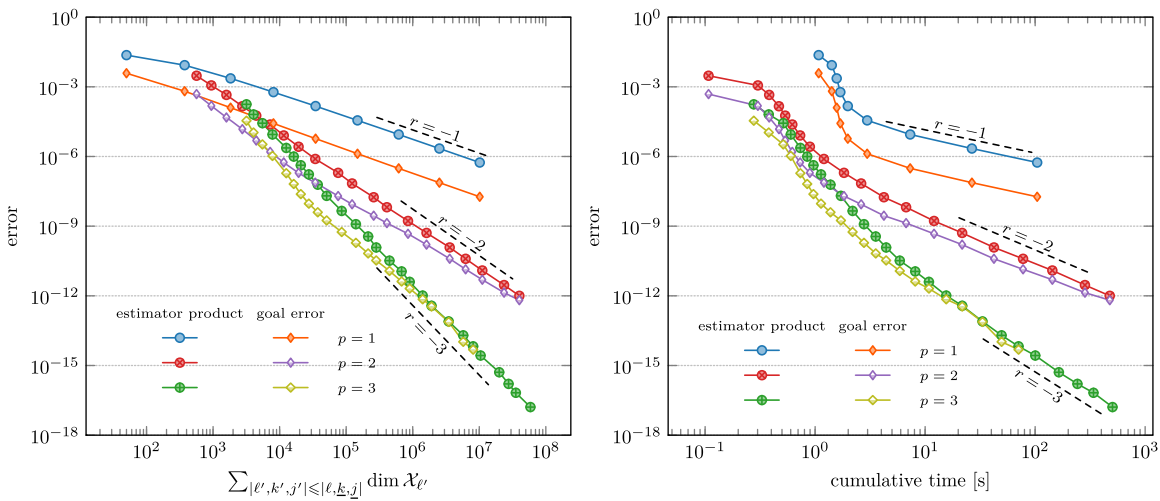
$$-\Delta u^* + (5, 5)^\top \cdot \nabla u^* = 1 \quad \text{in } \Omega \quad \text{subject to } u^* = 0 \quad \text{on } \Gamma_D \quad \text{and } \nabla u^* \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_N. \quad (89)$$

Consider  $g = 0$  and  $\mathbf{g} = \chi_S(1, 1)^\top$  in the quantity of interest

$$G(u^*) = \int_S \partial_{x_1} u^* + \partial_{x_2} u^* dx \quad \text{with } S := (-1/2, 1/2)^2 \cap \Omega.$$



**Figure 2:** Left: Mesh  $\mathcal{T}_{15}$  for the problem (88) generated by Algorithm 1 with  $\#\mathcal{T}_{15} = 2315$ . Right: Mesh  $\mathcal{T}_{18}$  for the problem (89) with  $\#\mathcal{T}_{18} = 2130$ , where the Dirichlet boundary part  $\Gamma_D$  is marked by red solid lines and the Neumann boundary part  $\Gamma_N$  by green dashed lines.

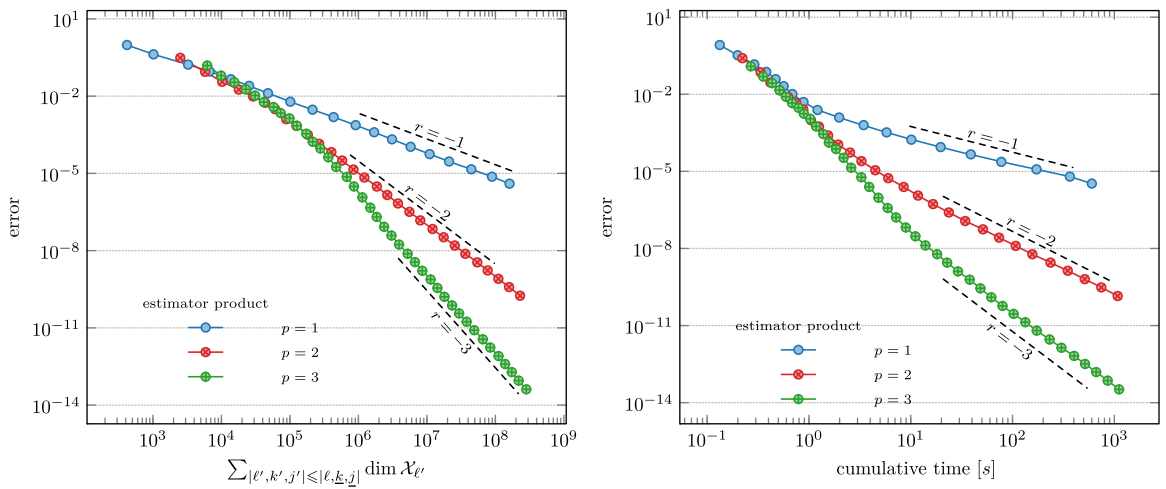


**Figure 3:** Convergence history plot of estimator product  $\eta_\ell(u_\ell^{m,n}) \zeta_\ell(z_\ell^{u,v})$  indicated by bullets and goal error from (17) indicated by diamonds with respect to the cumulative computational work (left) and with respect to the cumulative computational time (right) for the benchmark problem (88).

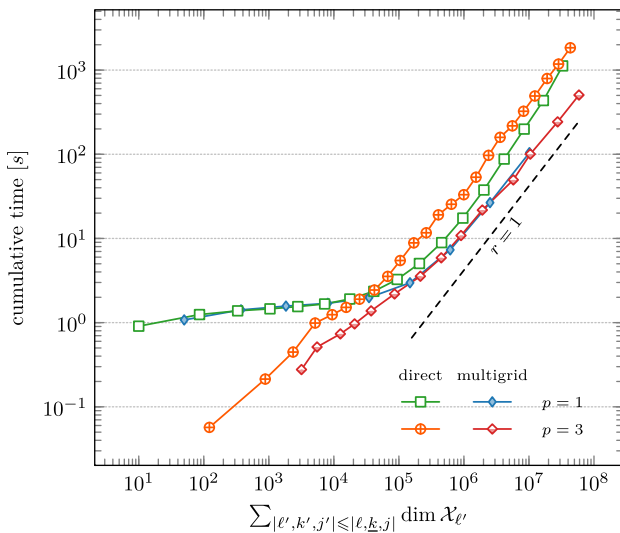
The exact solution  $u^*$  is not known analytically in this case so that we do not have access to the exact goal error  $|G(u^*) - G_\ell(u_\ell^{m,n}, z_\ell^{\mu,\nu})|$ . Figure 2 (right) shows a mesh generated by Algorithm 1 as well as the configuration, i.e., the support  $S$  of  $g$  in blue, the Dirichlet boundary in red solid lines, and the Neumann boundary in green dashed lines.

### 7.2.1 Optimality of Algorithm 1

Figure 3 displays the estimator product  $\eta_\ell(u_\ell^{m,n}) \zeta_\ell(z_\ell^{\mu,\nu})$  and the goal error  $|G(u^*) - G_\ell(u_\ell^{m,n}, z_\ell^{\mu,\nu})|$  from (17) for the problem (88), due to higher-order approximations, we only show results prior to machine precision. For all investigated polynomial degrees  $p$ , the goal error and the estimator product are indeed equivalent. Algorithm 1 achieves the optimal rate  $-p$  with respect to the cumulative computational work and with respect to the cumulative computational time in Figure 3 for problem (88) and Figure 4 for problem (89). Figure 5 shows that the proposed algorithm indeed achieves linear complexity and is substantially faster than the MATLAB



**Figure 4:** Convergence history plot of estimator product  $\eta_\ell(u_\ell^{m,n}) \zeta_\ell(z_\ell^{\mu,\nu})$  with respect to the cumulative computational cost (left) and the cumulative computational time (right) for the benchmark problem (89).



**Figure 5:** Comparison of cumulative time of the local multigrid solver with the MATLAB built-in direct solver `mldivide` with respect to the cumulative computational cost for the benchmark problem (89).



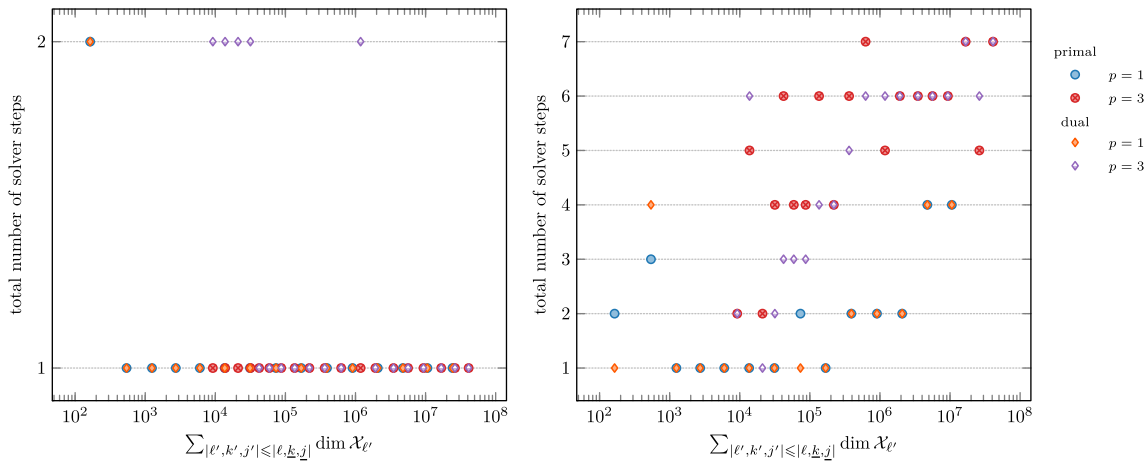
built-in direct solver as the slightly larger slope of the latter indicates super-linear complexity. Table 2 displays the weighted costs

$$\eta_\ell \left( u_\ell^{\underline{m}, \underline{n}} \right) \zeta_\ell \left( z_\ell^{\underline{\mu}, \underline{\nu}} \right) \left( \sum_{\substack{(\ell', k', j') \in Q \\ |\ell', k', j'| \leq |\ell, \underline{k}, \underline{j}|}} \text{time}(\ell', k', j') \right)^p \quad (90)$$

of Algorithm 1 for polynomial degree  $p = 2$  with  $\text{time}(\ell', k', j')$  in seconds and highlights the corresponding optimal choices of the parameters. This justifies the selection of  $\theta = 0.5$  together with larger symmetrization parameter  $\lambda_{\text{sym}} = 0.7$ , and algebraic solver parameter  $\lambda_{\text{alg}} = 0.7$ . The table for the second benchmark problem from (89) leads to similar results and is therefore omitted. While the choice of the damping parameter  $0 < \delta < 2\alpha/L^2$  in (13) is crucial in the case of large convection to guarantee the contraction property (14), the adaptivity parameters appear more robust with respect to other coefficients in (4).

**Table 2:** Optimal selection of parameters with respect to the cumulative computational costs (overall computation time in seconds) for the experiment (88) with fixed polynomial degree  $p = 2$  and  $\delta = 0.5$ . For comparison, the table displays the value of the weighted costs from (90) (in  $10^{-7}$ ) with overall stopping criterion  $\eta_\ell \left( u_\ell^{\underline{m}, \underline{n}} \right) \zeta_\ell \left( z_\ell^{\underline{\mu}, \underline{\nu}} \right) < 5 \cdot 10^{-10}$  for various choices of  $\lambda_{\text{sym}}$ ,  $\lambda_{\text{alg}}$ , and  $\theta$ . For each  $\theta$ -block, we mark the row-wise optimal values in blue, the column-wise optimal values in yellow, and in green if both optimal values coincide.

$\times 10^{-7}$	$\theta = 0.1$					$\theta = 0.3$					$\theta = 0.5$				
$\lambda_{\text{alg}}/\lambda_{\text{sym}}$	0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9
0.1	38.7	33.4	29.6	22.1	24.4	10.2	5.12	4.90	4.83	4.74	6.18	4.48	4.66	4.89	5.25
0.3	36.2	24.7	24.5	21.8	23.1	7.28	4.98	3.53	3.27	3.26	4.18	4.54	4.79	5.01	5.13
0.5	24.3	24.7	24.7	23.4	23.6	5.84	3.64	3.39	3.27	3.37	3.41	2.71	2.52	2.49	2.68
0.7	24.1	24.8	23.8	22.2	24.0	4.95	3.59	3.30	3.25	3.42	2.74	2.35	2.41	2.24	2.46
0.9	23.5	24.6	22.3	24.4	23.8	4.90	3.58	3.29	3.26	3.41	2.81	2.30	2.43	2.27	2.41
	$\theta = 0.7$					$\theta = 0.8$					$\theta = 0.9$				
0.1	5.82	5.18	5.43	5.40	5.93	8.53	6.10	7.31	6.67	7.77	11.6	8.86	9.12	9.87	9.97
0.3	4.65	4.86	5.35	5.98	6.67	6.27	5.92	7.20	7.46	7.57	8.62	8.40	9.27	10.6	11.5
0.5	3.69	2.89	2.88	2.95	3.13	5.09	3.61	3.66	3.63	3.66	7.27	5.32	4.84	4.93	5.12
0.7	2.99	2.56	2.64	2.62	2.89	3.75	3.12	3.23	3.03	3.11	4.58	3.95	4.04	4.43	4.79
0.9	2.89	2.49	2.65	2.66	2.89	3.79	3.11	3.19	3.13	3.27	4.67	4.06	4.16	4.35	4.61



**Figure 6:** Number of total solver steps  $|\ell, \underline{m}, \underline{n}| - |\ell, 0, 0|$  resp.  $|\ell, \underline{\mu}, \underline{\nu}| - |\ell, 0, 0|$  on each mesh level for the benchmark problems (88) (left) and (89) (right).

Finally, in Figure 6, we display the number of total solver steps  $|\ell, \underline{m}, \underline{n}| - |\ell, 0, 0|$  resp.  $|\ell, \underline{\mu}, \underline{\nu}| - |\ell, 0, 0|$  on each mesh level for both benchmark problems (88) and (89). The plots show that the two iterations often stop after the same number of steps.

## 8 Summary

In this work, we developed a cost-optimal goal-oriented adaptive finite element method for the efficient computation of the quantity of interest  $G(u^*)$  with solution  $u^*$  to the general second-order linear elliptic partial differential equation (4). Since the current analysis of iterative algebraic solvers for nonsymmetric systems with optimal preconditioner only leads to contraction of the residual in a vector norm, we proposed a nested iterative solver for the primal and dual problem in parallel. The strategy consists of the Zarantonello iteration (13) as an outer solver loop and an optimal multigrid solver for the arising SPD system as an innermost solver loop. In recent own work [33], we have shown that the link between convergence rates with respect to the degrees of freedom and the total computational cost is full linear convergence of the quasi-error  $H_\ell^{k,j} Z_\ell^{k,j}$ . To this end, Theorem 1 shows that the proposed adaptive algorithm contracts (up to a multiplicative constant) the quasi-error product  $H_\ell^{k,j} Z_\ell^{k,j}$  in every step, independently of the algorithmic decision to employ mesh refinement, symmetrization, or the algebraic solver. A particular problem in the analysis is that the nested iterative solver procedure only guarantees contraction as long as  $1 \leq k < \underline{k}[\ell]$ , whereas contraction for the final iterate is only guaranteed up to an estimator term (cf. (29)). Another difficulty arises from the nonsymmetric setting with a quasi-Pythagorean estimate (18) replacing the usual Pythagorean estimate. Therefore, the proof of Theorem 1 employs the equivalence of R-linear convergence and tail-summability of the quasi-error product  $H_\ell^{k,j} Z_\ell^{k,j}$  and leads to mild restriction on the product  $\lambda_{\text{sym}} \lambda_{\text{alg}}$  of the involved solver stopping parameters. The key ingredients to cost-optimality are an adaptive mesh-refinement algorithm with optimal convergence rate with respect to the number of degrees of freedom (under the assumption of exact solution) and an algebraic solver for the linear system of equations that is contractive with respect to the underlying Sobolev norm. In this regard, the analysis in this paper may guide the generalization to conforming discretizations of vector-valued elliptic problems. Finally, the numerical experiments in Section 7 suggest that the proposed strategy allows for large stopping parameter in practice and that a larger choice is beneficial in terms of total runtime. Admittedly, the development of an optimal solver for the nonsymmetric problem (10) would allow to prove full linear convergence with an arbitrary selection of the stopping parameter.

**Research ethics:** Not applicable.

**Author contributions:** All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

**Use of Large Language Models, AI and Machine Learning Tools:** None declared.

**Competing interests:** The authors state no conflict of interest.

**Research funding:** This research was funded in whole or in part by the Austrian Science Fund (FWF) through [10.55776/F65], [10.55776/I6802], and [10.55776/P33216]. For open access purposes, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. Additionally, Maximilian Brunner and Julian Streitberger are supported by the Vienna School of Mathematics.

**Data availability:** Not applicable.

## References

- [1] W. Dörfler, “A convergent adaptive algorithm for Poisson’s equation,” *SIAM J. Numer. Anal.*, vol. 33, no. 3, pp. 1106–1124, 1996.
- [2] P. Morin, R. H. Nochetto, and K. G. Siebert, “Data oscillation and convergence of adaptive FEM,” *SIAM J. Numer. Anal.*, vol. 38, no. 2, pp. 466–488, 2000.
- [3] P. Binev, W. Dahmen, and R. DeVore, “Adaptive finite element methods with convergence rates,” *Numer. Math.*, vol. 97, no. 2, pp. 219–268, 2004.

- [4] R. Stevenson, “Optimality of a standard adaptive finite element method,” *Found. Comput. Math.*, vol. 7, no. 2, pp. 245–269, 2007.
- [5] J. M. Cascón, C. Kreuzer, R. H. Nochetto, and K. G. Siebert, “Quasi-optimal convergence rate for an adaptive finite element method,” *SIAM J. Numer. Anal.*, vol. 46, no. 5, pp. 2524–2550, 2008.
- [6] C. Kreuzer and K. G. Siebert, “Decay rates of adaptive finite elements with Dörfler marking,” *Numer. Math.*, vol. 117, no. 4, pp. 679–716, 2011.
- [7] J. M. Cascón and R. H. Nochetto, “Quasioptimal cardinality of AFEM driven by nonresidual estimators,” *IMA J. Numer. Anal.*, vol. 32, no. 1, pp. 1–29, 2012.
- [8] M. Feischl, T. Führer, and D. Praetorius, “Adaptive FEM with optimal convergence rates for a certain class of nonsymmetric and possibly nonlinear problems,” *SIAM J. Numer. Anal.*, vol. 52, no. 2, pp. 601–625, 2014.
- [9] C. Carstensen, M. Feischl, M. Page, and D. Praetorius, “Axioms of adaptivity,” *Comput. Math. Appl.*, vol. 67, no. 6, pp. 1195–1253, 2014.
- [10] R. Becker and R. Rannacher, “An optimal control approach to a posteriori error estimation in finite element methods,” *Acta Numer.*, vol. 10, pp. 1–102, 2001.
- [11] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Basel, Springer Science & Business Media, 2003.
- [12] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, “Introduction to adaptive methods for differential equations,” *Acta Numer.*, vol. 4, pp. 105–158, 1995.
- [13] M. B. Giles and E. Süli, “Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality,” *Acta Numer.*, vol. 11, pp. 145–236, 2002.
- [14] M. S. Mommer and R. Stevenson, “A goal-oriented adaptive finite element method with convergence rates,” *SIAM J. Numer. Anal.*, vol. 47, no. 2, pp. 861–886, 2009.
- [15] R. Becker, E. Estecahandy, and D. Trujillo, “Weighted marking for goal-oriented adaptive finite element methods,” *SIAM J. Numer. Anal.*, vol. 49, no. 6, pp. 2451–2469, 2011.
- [16] M. Feischl, G. Gantner, A. Haberl, D. Praetorius, and T. Führer, “Adaptive boundary element methods for optimal convergence of point errors,” *Numer. Math.*, vol. 132, no. 3, pp. 541–567, 2016.
- [17] M. Feischl, D. Praetorius, and K. G. van der Zee, “An abstract analysis of optimal goal-oriented adaptivity,” *SIAM J. Numer. Anal.*, vol. 54, no. 3, pp. 1423–1448, 2016.
- [18] M. Holst and S. Pollock, “Convergence of goal-oriented adaptive finite element methods for nonsymmetric problems,” *Numer. Methods Partial Differ. Equ.*, vol. 32, no. 2, pp. 479–509, 2016.
- [19] R. Becker, M. Innerberger, and D. Praetorius, “Optimal convergence rates for goal-oriented FEM with quadratic goal functional,” *Comput. Methods Appl. Math.*, vol. 21, no. 2, pp. 267–288, 2021.
- [20] R. Becker, M. Brunner, M. Innerberger, J. M. Melenk, and D. Praetorius, “Rate-optimal goal-oriented adaptive FEM for semilinear elliptic PDEs,” *Comput. Math. Appl.*, vol. 118, pp. 18–35, 2022.
- [21] B. Endtmayer, U. Langer, and T. Wick, “Multigoal-oriented error estimates for non-linear problems,” *J. Numer. Math.*, vol. 27, no. 4, pp. 215–236, 2019.
- [22] B. Endtmayer, U. Langer, and T. Wick, “Two-side a posteriori error estimates for the dual-weighted residual method,” *SIAM J. Sci. Comput.*, vol. 42, no. 1, pp. A371–A394, 2020.
- [23] V. Dolejší, O. Bartoš, and F. Roskovec, “Goal-oriented mesh adaptation method for nonlinear problems including algebraic errors,” *Comput. Math. Appl.*, vol. 93, pp. 178–198, 2021.
- [24] A. Cohen, W. Dahmen, and R. DeVore, “Adaptive wavelet methods for elliptic operator equations: convergence rates,” *Math. Comput.*, vol. 70, no. 233, pp. 27–75, 2001.
- [25] A. Cohen, W. Dahmen, and R. DeVore, “Adaptive wavelet schemes for nonlinear variational problems,” *SIAM J. Numer. Anal.*, vol. 41, no. 5, pp. 1785–1823, 2003.
- [26] C. Carstensen and J. Gedicke, “An adaptive finite element eigenvalue solver of asymptotic quasi-optimal computational complexity,” *SIAM J. Numer. Anal.*, vol. 50, no. 3, pp. 1029–1057, 2012.
- [27] G. Gantner, A. Haberl, D. Praetorius, and S. Schimanko, “Rate optimality of adaptive finite element methods with respect to overall computational costs,” *Math. Comput.*, vol. 90, no. 331, pp. 2011–2040, 2021.
- [28] M. Brunner, M. Innerberger, A. Miraçi, D. Praetorius, J. Streitberger, and P. Heid, “Adaptive FEM with quasi-optimal overall cost for nonsymmetric linear elliptic PDEs,” *IMA J. Numer. Anal.*, vol. 44, no. 3, pp. 1560–1596, 2024.
- [29] R. Becker, G. Gantner, M. Innerberger, and D. Praetorius, “Goal-oriented adaptive finite element methods with optimal computational complexity,” *Numer. Math.*, vol. 153, no. 1, pp. 111–140, 2023.
- [30] J. Wu and H. Zheng, “Uniform convergence of multigrid methods for adaptive meshes,” *Appl. Numer. Math.*, vol. 113, pp. 109–123, 2017.
- [31] L. Chen, R. H. Nochetto, and J. Xu, “Optimal multilevel methods for graded bisection grids,” *Numer. Math.*, vol. 120, no. 1, pp. 1–34, 2012.
- [32] M. Innerberger, A. Miraçi, D. Praetorius, and J. Streitberger, “hp-robust multigrid solver on locally refined meshes for FEM discretizations of symmetric elliptic PDEs,” *ESAIM: Math. Modell. Numer. Anal.*, vol. 58, no. 1, pp. 247–272, 2024.

- [33] P. Bringmann, M. Feischl, A. Miraçi, D. Praetorius, and J. Streitberger, “On full linear convergence and optimal complexity of adaptive FEM with inexact solver,” *arXiv:2311.15738*, 2023.
- [34] M. Feischl, “Inf-sup stability implies quasi-orthogonality,” *Math. Comput.*, vol. 91, no. 337, pp. 2059–2094, 2022.
- [35] A. Bespalov, A. Haberl, and D. Praetorius, “Adaptive FEM with coarse initial mesh guarantees optimal convergence rates for compactly perturbed elliptic problems,” *Comput. Methods Appl. Mech. Eng.*, vol. 317, pp. 318–340, 2017.
- [36] E. Zangtanello, “Solving functional equations by contractive averaging, math,” *Res. Cent. Rep.*, vol. 160, 1960.
- [37] S. Congreve and T. P. Wihler, “Iterative Galerkin discretizations for strongly monotone problems,” *J. Comput. Appl. Math.*, vol. 311, pp. 457–472, 2017.
- [38] G. Gantner, A. Haberl, D. Praetorius, and B. Stifftner, “Rate optimal adaptive FEM with inexact solver for nonlinear operators,” *IMA J. Numer. Anal.*, vol. 38, no. 4, pp. 1797–1831, 2018.
- [39] A. Haberl, D. Praetorius, S. Schimanko, and M. Vohralík, “Convergence and quasi-optimal cost of adaptive algorithms for nonlinear operators including iterative linearization and algebraic solver,” *Numer. Math.*, vol. 147, no. 3, pp. 679–725, 2021.
- [40] E. Zeidler, *Nonlinear Functional Analysis and its Applications. Part II/A*, New York, Springer-Verlag, 1990.
- [41] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed. Philadelphia, PA, Society for Industrial and Applied Mathematics, 2003.
- [42] Y. Saad and M. H. Schultz, “GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems,” *SIAM J. Sci. Comput.*, vol. 7, no. 3, pp. 856–869, 1986.
- [43] R. Stevenson, “The completion of locally refined simplicial partitions created by bisection,” *Math. Comput.*, vol. 77, no. 261, pp. 227–241, 2008.
- [44] M. Aurada, M. Feischl, T. Führer, M. Karkulik, and D. Praetorius, “Energy norm based error estimators for adaptive BEM for hypersingular integral equations,” *Appl. Numer. Math.*, vol. 95, pp. 15–35, 2015.
- [45] M. Karkulik, D. Pavlicek, and D. Praetorius, “On 2D newest vertex bisection: optimality of mesh-closure and  $H^1$ -stability of  $L_2$ -projection,” *Constr. Approx.*, vol. 38, no. 2, pp. 213–234, 2013.
- [46] L. Diening, L. Gehring, and J. Storn, “Adaptive Mesh Refinement for Arbitrary Initial Triangulations,” *arXiv:2306.02674*, 2023.
- [47] M. Ainsworth and J. T. Oden, *A Posteriori Error Estimation in Finite Element Analysis, Ser. Pure and Applied Mathematics*, New York, Wiley-Interscience, 2000.
- [48] R. Verfürth, “A posteriori error estimation and adaptive mesh-refinement techniques,” in *Proceedings of the Fifth International Congress on Computational and Applied Mathematics (Leuven, 1992)*, vol. 50, pp. 67–83, 1994.
- [49] C.-M. Pfeiler and D. Praetorius, “Dörfler marking with minimal cardinality is a linear complexity problem,” *Math. Comput.*, vol. 89, no. 326, pp. 2735–2752, 2020.
- [50] M. Brunner, M. Innerberger, A. Miraçi, D. Praetorius, J. Streitberger, and P. Heid, “Corrigendum to: adaptive FEM with quasi-optimal overall cost for nonsymmetric linear elliptic PDEs,” *IMA J. Numer. Anal.*, vol. 44, no. 3, pp. 1903–1909, 2024.
- [51] M. Innerberger and D. Praetorius, “MooAFEM: an object oriented matlab code for higher-order adaptive FEM for (nonlinear) elliptic PDEs,” *Appl. Math. Comput.*, vol. 442, Art. no. 127731, 2023.