ASC Report No. 34/2018

Adaptive Uzawa algorithm for the Stokes equation

G. Di Fratta. T. Führer, G. Gantner, and D. Praetorius

Institute for Analysis and Scientific Computing
Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-00-1

Most recent ASC Reports

33/2018 *C.-M. Pfeiler, M. Ruggeri, B. Stiftner, L. Exl, M. Hochsteger, G. Hrkac, J. Schöberl, N.J. Mauser, and D. Praetorius*
Computational micromagnetics with Commics

32/2018 *A. Gerstenmayer and A. Jüngel*
Comparison of a finite-element and finite-volume scheme for a degenerate cross-diffusion system for ion transport

31/2018 *A. Bespalov, D. Praetorius, L. Rocchi, and M. Ruggeri*
Convergence of adaptive stochastic
Galerkin FEM

30/2018 *L. Chen, E.S. Daus, and A. Jüngel*
Rigorous mean-field limit and cross diffusion

29/2018 *X. Huo, A. Jüngel, and A.E. Tzavaras*
High-friction limits of Euler flows for multicomponent systems

28/2018 *A. Jüngel, U. Stefanelli, and L. Trussardi*
Two time discretizations for gradient flows exactly replicating energy dissipation

27/2018 *A. Jüngel and M. Ptashnyk*
Homogenization of degenerate cross-diffusion systems

26/2018 *F. Auer and E.B. Weinmüller*
Convergence of the collocation schemes for systems of nonlinear ODEs with a time singularity

25/2018 *M. Karkulik, J.M. Melenk*
H-matrix approximability of inverses of discretizations of the fractional Laplacian

24/2018 *M. Bernkopf and J.M. Melenk*
Analysis of the hp-version of a first order system least squares method for the Helmholtz equation

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: `admin@asc.tuwien.ac.at`
WWW: `http://www.asc.tuwien.ac.at`
FAX: `+43-1-58801-10196`

ISBN 978-3-902627-00-1

ADAPTIVE UZAWA ALGORITHM FOR THE STOKES EQUATION

GIOVANNI DI FRATTA, THOMAS FÜHRER, GREGOR GANTNER, AND DIRK PRAETORIUS

Abstract. Based on the Uzawa algorithm, we consider an adaptive finite element method for the Stokes system. We prove linear convergence with optimal algebraic rates, if the arising linear systems are solved iteratively, e.g., by PCG. Our analysis avoids the use of efficiency estimates for the residual error estimator. Unlike prior work, our adaptive Uzawa algorithm can thus avoid to discretize the given data and does not rely on an interior node property for the refinement.

1. Introduction

The mathematical analysis of adaptive finite element methods (AFEMs) has significantly increased over the last years. Nowadays, AFEMs are recognized as a powerful and rigorous tool to efficiently solve partial differential equations arising in physics and engineering.

1.1. Model problem. In this paper, we focus on an adaptive algorithm for the solution of the steady-state Stokes equations, which after a suitable normalization read

$$\begin{aligned} -\Delta \boldsymbol{u} + \nabla p &= \boldsymbol{f} \quad \text{in } \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 \quad \text{in } \Omega, \\ \boldsymbol{u} &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{1}$$

In the literature, the first equation is referred to as *momentum equation*, the second as *mass equation*, and the third as *no-slip boundary condition*. Here, $\Omega \subset \mathbb{R}^d$ with $d \in \{2, 3\}$ is a bounded polygonal resp. polyhedral Lipschitz domain. Given the body force $\boldsymbol{f}$, one seeks the velocity field $\boldsymbol{u}$ of an incompressible fluid and the associated pressure p. With

$$\mathbb{V} := H_0^1(\Omega)^d, \quad \mathbb{P} := \left\{ q \in L^2(\Omega) : \int_\Omega q \, dx = 0 \right\}, \tag{2}$$

it is well-known that the Stokes problem admits a unique solution $(\boldsymbol{u}, p) \in \mathbb{V} \times \mathbb{P}$, where p can be characterized as the unique null average solution of the elliptic Schur complement equation; see, e.g., [Bra03]. More precisely, the pressure solves the elliptic equation

$$Sp = \nabla \cdot \Delta^{-1} \boldsymbol{f} \quad \text{with the Schur complement operator} \quad S := \nabla \cdot \Delta^{-1} \nabla : \mathbb{P} \to \mathbb{P}. \tag{3}$$

Date: December 20, 2018.

2010 *Mathematics Subject Classification.* 65N30, 65N50, 65N15, 41A25.

Key words and phrases. adaptive finite element method; optimal convergence; Uzawa algorithm; Stokes equation.

Acknowledgement. The authors thankfully acknowledge the support by the Austrian Science Fund (FWF) through grant P27005 (DP), P29096 (GG, DP), as well as grant F65 (GDF, DP) and by CONICYT through FONDECYT project P11170050 (TF). Moreover, GG thanks Peter Binev for his explanations on [BD04, Bin15].

The latter equation can be reformulated as a fixpoint problem for the operator

$$N_\alpha : \mathbb{P} \to \mathbb{P}, \quad q \mapsto (I - \alpha S)q + \alpha \nabla \cdot \Delta^{-1} \boldsymbol{f}. \tag{4}$$

Note that S is self-adjoint. Since the norm of self-adjoint operators coincides with their spectral radius and S has positive spectrum, one has that $\|I - \alpha S\| < 1$ whenever $|1-\alpha\|S\|| < 1$. It follows that N_α is a contraction for $0 < \alpha < 2\,\|S\|^{-1}$; see Appendix A. Moreover, elementary calculation proves that $\|S\| \leq 1$. Hence, for all $0 < \alpha < 2$ and any initial guess $p^0 \in \mathbb{P}$, the generalized Richardson iteration

$$p^{j+1} := N_\alpha p^j = (I - \alpha S)p^j + \alpha \nabla \cdot \Delta^{-1} \boldsymbol{f} \tag{5}$$

converges to the exact pressure of the Stokes problem. It follows that $\boldsymbol{u} = \lim_{j\to\infty} \boldsymbol{u}[p^j]$ in $\mathbb{V}$ with $\boldsymbol{u}[p^j] := -\Delta^{-1}(\boldsymbol{f} - \nabla p^j)$, so that, at the continuous level, the full iterative process can be expressed in the form

$$\begin{aligned} \boldsymbol{u}[p^j] &= -\Delta^{-1}(\boldsymbol{f} - \nabla p^j), \\ p^{j+1} &= p^j - \alpha \nabla \cdot \boldsymbol{u}[p^j]. \end{aligned} \tag{6}$$

In the spirit of [KS08], the iterative scheme (6), usually referred to as *Uzawa algorithm* for the Stokes problem, is the starting point of our AFEM analysis.

1.2. State of the art. Although AFEMs for the analysis of mixed variational problems issuing from fluid dynamics have a long history in the engineering and physics literature, only in the last decade, [DDU02] introduced an adaptive wavelet method based on the Uzawa algorithm for solving the Stokes problem. In [BMN02], the adaptive wavelet method is replaced by an AFEM. Their numerical experiments suggested that the latter algorithm leads to optimal algebraic convergence rates. Indeed, by addition of a mesh-coarsening step to this method, [Kon06] proved optimal convergence rates. Later, in [KS08], the original algorithm of [BMN02] was modified by adding an additional loop, which separately controls the triangulations on which the pressure is discretized.

We also note that for a standard conforming AFEM with Taylor–Hood elements, the first proofs of convergence were presented in [MSV08, Sie10]. The work [Gan14] gives an optimality proof under the assumption that some *general quasi-orthogonality* is satisfied. This assumption has only recently been verified in [Fei17]. For adaptive nonconforming finite element methods, convergence and optimal rates have been investigated and proved in [BM11, HX13, CPR13].

1.3. Adaptive Uzawa FEM. In this work, we further investigate the algorithm of [KS08], which is described in the following: Given a possibly non-conforming partition $\mathcal{P}_i$ of Ω, we denote by $p_i \in \mathbb{P}_i$ the best approximation to p, with respect to the S-induced energy norm $\|\cdot\|_{\mathbb{P}}$, from the corresponding discrete space $\mathbb{P}_i \subset \mathbb{P}$ of piecewise polynomials of degree $m-1$ with vanishing integral mean. With the corresponding velocity $\boldsymbol{u}_i := \boldsymbol{u}[p_i]$ defined analogously to (6) and the L^2-orthogonal projection $\Pi_i : L^2(\Omega) \to \mathbb{P}_i$, one can show that $(\boldsymbol{u}_i, p_i)$ is the unique solution of the reduced problem

$$\begin{aligned} -\Delta \boldsymbol{u}_i + \nabla p_i &= \boldsymbol{f} \quad \text{in } \Omega, \\ \Pi_i \nabla \cdot \boldsymbol{u}_i &= 0 \quad \text{in } \Omega, \\ \boldsymbol{u}_i &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{7}$$

In general, the velocity $\boldsymbol{u}_i$ is not discrete, and hence this problem can still not be solved in practice. In an inner loop, the velocity $\boldsymbol{u}_i$ is approximated by some FEM approximation $\boldsymbol{U}_{ijk} \in \mathbb{V}_{ijk}$ via a standard adaptive algorithm of the form

$$\boxed{\text{SOLVE}} \longrightarrow \boxed{\text{ESTIMATE}} \longrightarrow \boxed{\text{MARK}} \longrightarrow \boxed{\text{REFINE}}$$

for the vector-valued Poisson problem steered by a weighted-residual error estimator η_{ijk}. Here, $\mathbb{V}_{ijk} \subset \mathbb{V}$ denotes the space of all continuous piecewise polynomials on some conforming triangulation $\mathcal{T}_{ijk}$, which is a refinement of the possibly non-conforming $\mathcal{P}_i$. In the next loop, we apply a discretized version of the Uzawa algorithm (6) to obtain an approximation $P_{ij} \in \mathbb{P}_i$ of p_i. Here, the update reads $P_{i(j+1)} = P_{ij} - \Pi_i \nabla \cdot \boldsymbol{U}_{ijk}$. The last loop employs an adaptive tree approximation algorithm from [BD04] to obtain a better approximation $p_{i+1} \in \mathbb{P}_{i+1}$ of p on a refinement $\mathcal{P}_{i+1}$ of the partition $\mathcal{P}_i$ such that $\vartheta \, \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega \le \|\Pi_{i+1} \nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega$ for some bulk parameter $0 < \vartheta < 1$. We will see in Section 3.1 that $\|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega$ is related to $\|p - p_i\|_{\mathbb{P}}$ and $\|\Pi_{i+1} \nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega$ to $\|p_{i+1} - p_i\|_{\mathbb{P}}$. In contrast to [KS08], in [BMN02] the latter loop was not present, since the same triangulation for the discretization of the pressure and the velocity, i.e., $\mathcal{P}_i = \mathcal{T}_{ijk}$ was used.

Under the assumption that the right-hand side $\boldsymbol{f}$ is a piecewise polynomial of degree $m-1$, [KS08] proved that the approximations $\boldsymbol{U}_{ijk}$ and P_{ij} converge with optimal algebraic rate to the exact solutions $\boldsymbol{u}$ and p. To generalize this result for arbitrary $\boldsymbol{f}$, as in the seminal work [Ste07], which proves optimal convergence of a standard AFEM for the Poisson problem, [KS08] applies an additional outer loop to resolve the data oscillations appropriately. However, [KS08] only outlines the proof of this generalization. Moreover, as in the seminal work [Ste07], the analysis of [KS08] hinges on the following interior node property: Given marked elements $\mathcal{M}_{ijk}$ of the current velocity triangulation $\mathcal{T}_{ijk}$, the next velocity triangulation $\mathcal{T}_{ij(k+1)}$ is the coarsest refinement via newest vertex bisection (NVB) such that all $T \in \mathcal{M}_{ijk}$ and all $T' \in \mathcal{T}_{ijk}$, which share a common $(n-1)$-dimensional hyperface, contain a vertex of $\mathcal{T}_{ij(k+1)}$ in their interior. In particular for $n = 3$, this property is highly demanding; see, e.g., the 3D refinement pattern in [EGP18].

1.4. Contributions of present work. In the spirit of [CKNS08], which generalizes [Ste07], we prove that the algorithm of [KS08] without the data approximation loop leads to convergence of the combined error estimator $\eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega$ (which is equivalent to the error plus data oscillations) at optimal algebraic rate with respect to the number of elements $\#\mathcal{T}_{ijk}$ if one uses standard newest vertex bisection (without interior node property) for the velocity triangulations. We also prove that the combined estimator sequence converges linearly in each step, i.e., it essentially contracts uniformly in each step. Moreover, our algorithm allows for the inexact solution of the arising linear systems for the discrete velocities by iterative solvers like PCG.

On a conceptual level, our proofs show that even for general saddle point problems and adaptive strategies based on Richardson-type iterations, the analysis of rate optimal adaptivity can be conducted without exploiting efficiency estimates of the corresponding *a posteriori* error estimators.

1.5. Outline. The paper is organized as follows: Section 2 rewrites the Stokes problem in its variational form, introduces newest vertex bisection, and fixes some notation for the discrete ansatz spaces. In Section 3, we consider the reduced Stokes problem and the corresponding Galerkin approximations, recall some well-known results on *a posteriori* error estimation, and introduce the tree approximation Algorithm 3.6 from [BD04] as well as our variant of the adaptive Uzawa Algorithm 3.6 from [KS08]. In Section 4, we state and prove linear convergence of the resulting combined error estimator in each step of the algorithm (Theorem 4.1). To this end, we show that each increase of either i, j, or k essentially leads to a uniform contraction of the combined error estimator. Finally, Section 5 is dedicated to the main Theorem 5.3 on optimal convergence rates for the combined error estimator and its proof. As an auxiliary result of general interest, Lemma 5.1 proves that the two different definitions of approximation classes from the literature, which are either based on the accuracy $\varepsilon > 0$ (see, e.g., [Ste08, KS08]) or the number of elements N (see, e.g., [CKNS08, CFPP14]), are exactly the same.

While all constants in statements of theorems, lemmas, etc. are explicitly given, we abbreviate the notation in proofs: For scalar terms A and B, we write $A \lesssim B$ to abbreviate $A \le C\,B$, where the generic constant $C > 0$ is clear from the context. Moreover, $A \simeq B$ abbreviates $A \lesssim B \lesssim A$.

2. Preliminaries

2.1. Continuous Stokes problem. The vector-valued velocity fields $\boldsymbol{v} \in \mathbb{V}$ are denoted in boldface, the scalar pressures $q \in \mathbb{P}$ in normal font. Let $\langle \cdot\,,\,\cdot \rangle_\Omega$ be the $L^2(\Omega)$ scalar product with the corresponding $L^2(\Omega)$ norm $\|\cdot\|_\Omega$. With the bilinear forms $a : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ and $b : \mathbb{V} \times \mathbb{P} \to \mathbb{R}$ defined by

$$a(\boldsymbol{w}, \boldsymbol{v}) := \langle \nabla \boldsymbol{w}\,,\,\nabla \boldsymbol{v} \rangle_\Omega \quad \text{and} \quad b(\boldsymbol{v}, q) := -\langle \nabla \cdot \boldsymbol{v}\,,\,q \rangle_\Omega,$$

the mixed variational formulation of the Stokes problem (1) reads as follows: Given $\boldsymbol{f} \in L^2(\Omega)^d$, let $(\boldsymbol{u}, p) \in \mathbb{V} \times \mathbb{P}$ be the unique solution to

$$\begin{array}{rclcll} a(\boldsymbol{u}, \boldsymbol{v}) & + & b(\boldsymbol{v}, p) & = & \langle \boldsymbol{f}\,,\,\boldsymbol{v} \rangle_\Omega & \text{for all } \boldsymbol{v} \in \mathbb{V}, \\ b(\boldsymbol{u}, q) & & & = & 0 & \text{for all } q \in \mathbb{P}. \end{array} \tag{8}$$

On the velocity space $\mathbb{V}$, we consider the $a(\cdot,\cdot)$-induced energy norm $\|\boldsymbol{v}\|_{\mathbb{V}} := a(\boldsymbol{v}, \boldsymbol{v})^{1/2} = \|\nabla \boldsymbol{v}\|_\Omega \simeq \|\boldsymbol{v}\|_{H^1(\Omega)}$. We note that $\nabla \cdot \boldsymbol{v} \in \mathbb{P}$ for all $\boldsymbol{v} \in \mathbb{V}$ and

$$\|\nabla \cdot \boldsymbol{v}\|_\Omega \le \|\nabla \boldsymbol{v}\|_\Omega = \|\boldsymbol{v}\|_{\mathbb{V}} \quad \text{for all } \boldsymbol{v} \in \mathbb{V}, \tag{9}$$

which follows from integration by parts; see Appendix B.

Define the operators $A : \mathbb{V} \to \mathbb{V}^*$, $B : \mathbb{V} \to \mathbb{P}^*$, and $B' : \mathbb{P} \to \mathbb{V}^*$ by

$$A\boldsymbol{v} := a(\boldsymbol{v}, \cdot), \quad B\boldsymbol{v} := b(\boldsymbol{v}, \cdot), \quad B'q := b(\cdot, q).$$

Then, the Schur complement operator $S := BA^{-1}B' : \mathbb{P} \to \mathbb{P}^* \sim \mathbb{P}$ is bounded, symmetric, and elliptic; see [KS08, Lemma 2.2]. Thus, it holds that $\|q\|_{\mathbb{P}} := \langle Sq\,,\,q \rangle_\Omega^{1/2} \simeq \|q\|_\Omega$ on $\mathbb{P}$. More precisely, there exists a constant $C_{\rm div} \ge 1$, which depends only on Ω, such that

$$C_{\rm div}^{-1}\,\|q\|_\Omega \le \|q\|_{\mathbb{P}} \le \|q\|_\Omega \quad \text{for all } q \in \mathbb{P}. \tag{10}$$

Here, the upper bound with constant 1 follows from $\|S\| \le 1$, which itself follows from (9).

2.2. Partitions, triangulations, and newest vertex bisection (NVB). Throughout, $\mathcal{P}$ is a finite (possibly non-conforming) partition of Ω into compact (non-degenerate) simplices, which is used to discretize $\mathbb{P}$, while $\mathcal{T}$ is a finite (conforming) triangulation of Ω into compact (non-degenerate) simplices, which is used to discretize $\mathbb{V}$. Throughout, we use NVB refinement; see, e.g., [Ste08, KPP13] for the precise mesh-refinement rules.

We write $\mathcal{P}' := \texttt{bisect}(\mathcal{P}, \mathcal{M})$ for the partition obtained by *one* bisection of all marked elements $\mathcal{M} \subseteq \mathcal{P}$, i.e., $\mathcal{M} = \mathcal{P} \backslash \mathcal{P}'$ and $\#\mathcal{M} = \#\mathcal{P}' - \#\mathcal{P}$. We write $\mathcal{P}' \in \mathbb{T}^{\rm nc}(\mathcal{P})$, if there exists $J \in \mathbb{N}_0$ and partitions $\mathcal{P}_j$ and $\mathcal{M}_j \subseteq \mathcal{P}_j$ for all $j = 0, \dots, J$, such that

$$\mathcal{P} = \mathcal{P}_0, \quad \mathcal{P}_j = \texttt{bisect}(\mathcal{P}_{j-1}, \mathcal{M}_{j-1}) \text{ for all } j = 1, \dots, J, \quad \text{and} \quad \mathcal{P}' = \mathcal{P}_J.$$

We write $\mathcal{T}' := \texttt{refine}(\mathcal{T}, \mathcal{M})$ for the coarsest triangulation such that (at least) all marked elements $\mathcal{M} \subseteq \mathcal{T}$ have been bisected, i.e., $\mathcal{M} \subseteq \mathcal{T} \backslash \mathcal{T}'$. We write $\mathcal{T}' \in \mathbb{T}^{\rm c}(\mathcal{T})$, if there exists $J \in \mathbb{N}_0$ and triangulations $\mathcal{T}_j$ and $\mathcal{M}_j \subseteq \mathcal{T}_j$ for all $j = 0, \dots, J$, such that

$$\mathcal{T} = \mathcal{T}_0, \quad \mathcal{T}_j = \texttt{refine}(\mathcal{T}_{j-1}, \mathcal{M}_{j-1}) \text{ for all } j = 1, \dots, J, \quad \text{and} \quad \mathcal{T}' = \mathcal{T}_J.$$

Let $\mathcal{T}_{\rm init}$ be a given initial (conforming) triangulation of Ω. We define the sets

$$\mathbb{T}^{\rm nc} := \mathbb{T}^{\rm nc}(\mathcal{T}_{\rm init}) \quad \text{and} \quad \mathbb{T}^{\rm c} := \mathbb{T}^{\rm c}(\mathcal{T}_{\rm init}) \tag{11}$$

of all non-conforming and conforming NVB refinements of $\mathcal{T}_{\rm init}$. Clearly, $\mathbb{T}^{\rm c} \subset \mathbb{T}^{\rm nc}$. We write $\mathcal{T} := \texttt{close}(\mathcal{P})$ if $\mathcal{P} \in \mathbb{T}^{\rm nc}$ is a partition and $\mathcal{T} \in \mathbb{T}^{\rm c}$ is the coarsest (conforming) refinement of $\mathcal{P}$. Existence and uniqueness of $\mathcal{T}$ follow from the fact that NVB is a binary refinement rule, and the order of the bisections does not matter. In particular, this also implies that $\texttt{refine}(\mathcal{T}, \mathcal{M}) = \texttt{close}(\texttt{bisect}(\mathcal{T}, \mathcal{M}))$ for all $\mathcal{T} \in \mathbb{T}^{\rm c}$ and $\mathcal{M} \subseteq \mathcal{T}$.

It follows from elementary geometric observations that NVB refinement leads only to finitely many shapes of simplices T; see, e.g., [Ste08]. Hence, all NVB refinements are uniformly γ-shape regular, i.e.,

$$\gamma := \sup_{\mathcal{P} \in \mathbb{T}^{\rm nc}} \max_{T \in \mathcal{P}} \frac{\operatorname{diam}(T)}{|T|^{1/d}} < \infty. \tag{12}$$

Finally, we recall the following properties of NVB, where $C_{\rm son}, C_{\rm cls} > 0$ are constants, which depend only on $\mathcal{T}_{\rm init}$ and the space dimension $d \geq 2$:

- **(M1) overlay estimate:** For all $\mathcal{P}, \mathcal{P}' \in \mathbb{T}^{\rm nc}$, there exists a (unique) coarsest common refinement $\mathcal{P} \oplus \mathcal{P}' \in \mathbb{T}^{\rm nc}(\mathcal{P}) \cap \mathbb{T}^{\rm nc}(\mathcal{P}')$. It holds that $\#(\mathcal{P} \oplus \mathcal{P}') \leq \#\mathcal{P} + \#\mathcal{P}' - \#\mathcal{T}_{\rm init}$. If $\mathcal{P}, \mathcal{P}' \in \mathbb{T}^{\rm c}$ are conforming, it also holds that $\mathcal{P} \oplus \mathcal{P}' \in \mathbb{T}^{\rm c}$.
- **(M2) finite number of sons:** For all $\mathcal{T} \in \mathbb{T}^{\rm c}$, $\mathcal{M} \subseteq \mathcal{T}$, and $\mathcal{T}' := \texttt{refine}(\mathcal{T}, \mathcal{M})$, it holds that $\bigcup\{T' \in \mathcal{T}' : T' \subseteq T\} = T$ and $\#\{T' \in \mathcal{T}' : T' \subseteq T\} \leq C_{\rm son}$ for all $T \in \mathcal{T}$.
- **(M3) mesh-closure estimate:** For all sequences $\mathcal{T}_j \in \mathbb{T}^{\rm c}$ such that $\mathcal{T}_0 = \mathcal{T}_{\rm init}$ and $\mathcal{T}_j = \texttt{refine}(\mathcal{T}_{j-1}, \mathcal{M}_{j-1})$ with $\mathcal{M}_{j-1} \subseteq \mathcal{T}_{j-1}$ for all $j \in \mathbb{N}$, it holds that

$$\#\mathcal{T}_J - \#\mathcal{T}_{\rm init} \leq C_{\rm cls} \sum_{j=0}^{J-1} \#\mathcal{M}_j \quad \text{for all } J \in \mathbb{N}_0. \tag{13}$$

- **(M4) conformity estimate:** For all partitions $\mathcal{P} \in \mathbb{T}^{\rm nc}$, it holds that

$$\#\texttt{close}(\mathcal{P}) - \#\mathcal{T}_{\rm init} \leq C_{\rm cls}(\#\mathcal{P} - \#\mathcal{T}_{\rm init}). \tag{14}$$

The overlay estimate (M1) is first proved in [Ste07] for $d = 2$, but the proof transfers to arbitrary dimension $d \geq 2$; see [CKNS08]. For $d = 2$, (M2) obviously holds with $C_{\mathrm{son}} = 4$, while it is proved in [GSS14] for general dimension $d \geq 2$. The closure estimate (M3) is first proved in [BDD04] for $d = 2$. For general $d \geq 2$, it is proved in [Ste08]. While the proofs of [BDD04, Ste08] require an admissibility condition on $\mathcal{T}_{\mathrm{init}}$, the work [KPP13] proves (M3) for $d = 2$, but arbitrary conforming triangulation $\mathcal{T}_{\mathrm{init}}$. We refer to Appendix D for the fact that (M3) implies (M4).

2.3. Discrete function spaces. Given a fixed polynomial degree $m \in \mathbb{N}$ as well as $\mathcal{P} \in \mathbb{T}^{\mathrm{nc}}$ and $\mathcal{T} \in \mathbb{T}^{\mathrm{c}}$, we consider the discrete spaces

$$\begin{aligned} \mathbb{P}(\mathcal{P}) &:= \{Q_{\mathcal{P}} \in \mathbb{P} \ : \forall T \in \mathcal{P} \quad Q_{\mathcal{P}}|_T \text{ is polynomial of degree } \leq m-1\}, \\ \mathbb{V}(\mathcal{T}) &:= \{\boldsymbol{V}_{\mathcal{T}} \in \mathbb{V} \ : \forall T \in \mathcal{T} \quad \boldsymbol{V}_{\mathcal{T}}|_T \text{ is polynomial of degree } \leq m\}, \end{aligned} \tag{15}$$

which consist of piecewise polynomials.

2.4. Auxiliary problems. Let $\mathcal{P} \in \mathbb{T}^{\mathrm{nc}}$. Then, $p_{\mathcal{P}} \in \mathbb{P}(\mathcal{P})$ denotes the best approximation of the exact pressure p with respect to $\|\cdot\|_{\mathbb{P}}$, i.e.,

$$\|p - p_{\mathcal{P}}\|_{\mathbb{P}} = \min_{Q_{\mathcal{P}} \in \mathbb{P}(\mathcal{P})} \|p - Q_{\mathcal{P}}\|_{\mathbb{P}}. \tag{16}$$

By definition of the operator S from (3), there exists a unique $\boldsymbol{u}_{\mathcal{P}} \in \mathbb{V}$ such that $(\boldsymbol{u}_{\mathcal{P}}, p_{\mathcal{P}}) \in \mathbb{V} \times \mathbb{P}(\mathcal{P})$ is the unique solution to the *reduced Stokes problem*

$$\begin{array}{rcll} a(\boldsymbol{u}_{\mathcal{P}}, \boldsymbol{v}) \ + \ b(\boldsymbol{v}, p_{\mathcal{P}}) & = & \langle \boldsymbol{f} \,,\, \boldsymbol{v}\rangle_\Omega & \text{for all } \boldsymbol{v} \in \mathbb{V}, \\ b(\boldsymbol{u}_{\mathcal{P}}, Q_{\mathcal{P}}) & = & 0 & \text{for all } Q_{\mathcal{P}} \in \mathbb{P}(\mathcal{P}); \end{array} \tag{17}$$

see [KS08, Section 4]. Note that the second condition can equivalently be stated as $\Pi_{\mathcal{P}} \nabla \cdot \boldsymbol{u}_{\mathcal{P}} = 0$ in Ω, where $\Pi_{\mathcal{P}} : L^2(\Omega) \to \mathbb{P}(\mathcal{P})$ is the orthogonal projection with respect to $\|\cdot\|_\Omega$. Thus, (17) is just the variational formulation of (7) (with $\mathcal{P}_i$ replaced by $\mathcal{P}$).

Even though, $p_{\mathcal{P}}$ is a discrete function, it can hardly be computed (since p is unknown). Given $q \in \mathbb{P}$, let $\boldsymbol{u}[q] \in \mathbb{V}$ be the unique solution to the (vector-valued) Poisson equation

$$a(\boldsymbol{u}[q], \boldsymbol{v}) = \langle \boldsymbol{f} \,,\, \boldsymbol{v}\rangle_\Omega - b(\boldsymbol{v}, q) \quad \text{for all } \boldsymbol{v} \in \mathbb{V}. \tag{18}$$

Note that $\boldsymbol{u}_{\mathcal{P}} = \boldsymbol{u}[p_{\mathcal{P}}]$.

Finally, let $\mathcal{T} \in \mathbb{T}^{\mathrm{nc}}(\mathcal{P}) \cap \mathbb{T}^{\mathrm{c}}$ be a conforming refinement of $\mathcal{P}$. Then, $\boldsymbol{U}_{\mathcal{T}}[q] \in \mathbb{V}(\mathcal{T})$ is the unique solution to the Galerkin discretization of (18)

$$a(\boldsymbol{U}_{\mathcal{T}}[q], \boldsymbol{V}_{\mathcal{T}}) = \langle \boldsymbol{f} \,,\, \boldsymbol{V}_{\mathcal{T}}\rangle_\Omega - b(\boldsymbol{V}_{\mathcal{T}}, q) \quad \text{for all } \boldsymbol{V}_{\mathcal{T}} \in \mathbb{V}(\mathcal{T}). \tag{19}$$

Note that $\boldsymbol{U}_{\mathcal{T}}[q]$ is the Galerkin approximation to $\boldsymbol{u}[q]$ in $\mathbb{V}(\mathcal{T})$. Since $\|\cdot\|_{\mathbb{V}}$ denotes the energy norm corresponding to $a(\cdot,\cdot)$, there holds the Céa lemma

$$\|\boldsymbol{u}[q] - \boldsymbol{U}_{\mathcal{T}}[q]\|_{\mathbb{V}} = \min_{\boldsymbol{V}_{\mathcal{T}} \in \mathbb{V}(\mathcal{T})} \|\boldsymbol{u}[q] - \boldsymbol{V}_{\mathcal{T}}\|_{\mathbb{V}}, \tag{20}$$

Recall the operators A, B, B' from Section 2.1. Note that $\boldsymbol{u}[q] - \boldsymbol{u}[r] = A^{-1}B'(r-q)$ for arbitrary $q, r \in \mathbb{P}$, which yields that $\|\boldsymbol{u}[q] - \boldsymbol{u}[r]\|_{\mathbb{V}}^2 = \langle B'(r-q) \,,\, A^{-1}B'(r-q)\rangle_{\mathbb{V}^* \times \mathbb{V}}$. By definition of the operator $S = BA^{-1}B'$ and the norm $\|\cdot\|_{\mathbb{P}}$, we thus see that

$$\|\boldsymbol{U}_{\mathcal{T}}[q] - \boldsymbol{U}_{\mathcal{T}}[r]\|_{\mathbb{V}} \leq \|\boldsymbol{u}[q] - \boldsymbol{u}[r]\|_{\mathbb{V}} = \|q - r\|_{\mathbb{P}}. \tag{21}$$

2.5. Notational conventions. Throughout this work, $(\boldsymbol{u}, p) \in \mathbb{V} \times \mathbb{P}$ denotes the exact solution of the continuous Stokes problem (8). All occurring functions $\boldsymbol{u}_{\mathcal{P}}$, $\boldsymbol{u}[q]$, and $\boldsymbol{U}_{\mathcal{T}}[q]$ are approximations of $\boldsymbol{u}$. All occurring functions $p_{\mathcal{P}}$ and $P_{\mathcal{P}}$ are approximations of p. We employ bold face symbols for velocity functions, e.g., $\boldsymbol{v} \in \mathbb{V}$ or $\boldsymbol{V}_{\mathcal{T}} \in \mathbb{V}(\mathcal{T})$, and normal font for pressure functions, e.g., $q \in \mathbb{P}$, $Q_{\mathcal{P}} \in \mathbb{P}(\mathcal{P})$. Finally, small letters indicate functions, which are continuous or not computable, e.g., $\boldsymbol{u}$, p, and $p_{\mathcal{P}}$, while *computable discrete* functions are written with capital letters, e.g., $\boldsymbol{U}_{\mathcal{T}}[Q_{\mathcal{P}}]$. The corresponding partitions $\mathcal{P} \in \mathbb{T}^{\mathrm{nc}}$ resp. triangulations $\mathcal{T} \in \mathbb{T}^{\mathrm{c}}$ are always indicated by indices.

2.6. Abbreviate notation for adaptive algorithm. The adaptive algorithm below generates nested partitions $\mathcal{P}_i \in \mathbb{T}^{\mathrm{nc}}$ and triangulations $\mathcal{T}_{ijk} \in \mathbb{T}^{\mathrm{c}}$ for certain indices $(i, j, k) \in \mathcal{Q} \subset \mathbb{N}_0^3$ such that $\mathcal{T}_{ijk} \in \mathbb{T}^{\mathrm{nc}}(\mathcal{P}_i) \cap \mathbb{T}^{\mathrm{c}}$. Furthermore, it provides approximations

$$p \approx P_{ij} \in \mathbb{P}_i := \mathbb{P}(\mathcal{P}_i) \quad \text{as well as} \quad \boldsymbol{u} \approx \boldsymbol{U}_{ijk} \in \mathbb{V}_{ijk} := \mathbb{V}(\mathcal{T}_{ijk}). \tag{22}$$

More precisely and with the notation from Section 2.4, it holds that[1]

$$P_{ij} \approx p_i := p_{\mathcal{P}_i} \quad \text{as well as} \quad \boldsymbol{U}_{ijk} \approx \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}] \approx \boldsymbol{u}[P_{ij}] =: \boldsymbol{u}_{ij}. \tag{23}$$

Besides this notation, let

$$\Pi_i := \Pi_{\mathcal{P}_i} : L^2(\Omega) \to \mathbb{P}(\mathcal{P}_i) \tag{24}$$

be the $L^2(\Omega)$-orthogonal projection (with respect to $\|\cdot\|_\Omega$) and let

$$\eta_{ijk} := \eta(\mathcal{T}_{ijk}; \boldsymbol{U}_{ijk}, P_{ij}) \approx \eta(\mathcal{T}_{ijk}; \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}], P_{ij}) \tag{25}$$

be the computable *a posteriori* error estimator from Section 3.1 below.

3. Adaptive Uzawa algorithm

3.1. *A posteriori* error estimation. Throughout this section, let $\mathcal{P} \in \mathbb{T}^{\mathrm{nc}}$ be a partition of $\Omega \subset \mathbb{R}^d$ and $\mathcal{T} \in \mathbb{T}^{\mathrm{nc}}(\mathcal{P}) \cap \mathbb{T}^{\mathrm{c}}$ be a conforming refinement. We recall the residual *a posteriori* error estimator: For $T \in \mathcal{T}$, $Q_{\mathcal{P}} \in \mathbb{P}(\mathcal{P})$, and $\boldsymbol{V}_{\mathcal{T}} \in \mathbb{V}(\mathcal{T})$, define

$$\eta_T(\boldsymbol{V}_{\mathcal{T}}, Q_{\mathcal{P}})^2 := |T|^{2/n} \, \|\boldsymbol{f} - \nabla Q_{\mathcal{P}} + \Delta \boldsymbol{V}_{\mathcal{T}}\|_T^2 + |T|^{1/n} \, \|[\![Q_{\mathcal{P}} \, \boldsymbol{n} - \nabla \boldsymbol{V}_{\mathcal{T}} \cdot \boldsymbol{n}]\!]\|_{\partial T \cap \Omega}^2, \tag{26}$$

where $[\![\cdot]\!]$ denotes the jump of its argument over ∂T. Then, the error estimator reads

$$\eta(\mathcal{M}; \boldsymbol{V}_{\mathcal{T}}, Q_{\mathcal{P}})^2 := \sum_{T \in \mathcal{M}} \eta_T(\boldsymbol{V}_{\mathcal{T}}, Q_{\mathcal{P}})^2 \quad \text{for all } \mathcal{M} \subset \mathcal{T}. \tag{27}$$

In the following, we recall some important properties of η from [CKNS08, KS08]. We start with the available reliability results.

Lemma 3.1 (*reliability* [KS08, Prop. 5.1, Prop. 5.5]). *There exists a constant $C_{\mathrm{rel}} > 0$ such that, for all $Q_{\mathcal{P}} \in \mathbb{P}(\mathcal{P})$, it holds that*

$$\|\boldsymbol{u}[Q_{\mathcal{P}}] - \boldsymbol{U}_{\mathcal{T}}[Q_{\mathcal{P}}]\|_{\mathbb{V}} \le C_{\mathrm{rel}} \, \eta(\mathcal{T}; \boldsymbol{U}_{\mathcal{T}}[Q_{\mathcal{P}}], Q_{\mathcal{P}}). \tag{28}$$

Moreover, it holds that

$$\|\boldsymbol{u}_{\mathcal{P}} - \boldsymbol{U}_{\mathcal{T}}[Q_{\mathcal{P}}]\|_{\mathbb{V}} + \|p_{\mathcal{P}} - Q_{\mathcal{P}}\|_{\mathbb{P}} \le C_{\mathrm{rel}} \left(\eta(\mathcal{T}; U_{\mathcal{T}}[Q_{\mathcal{P}}], Q_{\mathcal{P}}) + \|\Pi_{\mathcal{P}} \nabla \cdot \boldsymbol{U}_{\mathcal{T}}[Q_{\mathcal{P}}]\|_\Omega \right) \tag{29}$$

[1] Do not mistake the pressure p_i with the iterates p^j of the exact Uzawa algorithm (6).

as well as

(30) $$\|\boldsymbol{u} - \boldsymbol{U}_{\mathcal{T}}[Q_{\mathcal{P}}]\|_{\mathbb{V}} + \|p - Q_{\mathcal{P}}\|_{\mathbb{P}} \le C_{\mathrm{rel}} \left(\eta(\mathcal{T}; U_{\mathcal{T}}[Q_{\mathcal{P}}], Q_{\mathcal{P}}) + \|\nabla \cdot \boldsymbol{U}_{\mathcal{T}}[Q_{\mathcal{P}}]\|_{\Omega} \right).$$

The constant C_{rel} depends only on γ-shape regularity. ■

For some fixed discrete pressure $Q_{\mathcal{P}}$, we recall the localized upper bound in the current form of [CKNS08], which improves [KS08, Prop. 5.1].

Lemma 3.2 (*discrete reliability* [CKNS08, Lemma 3.6]). *Let $\widehat{\mathcal{T}} \in \mathbb{T}^{\mathrm{c}}(\mathcal{T})$. There exists a constant $C_{\mathrm{drel}} > 0$ such that, for all $Q_{\mathcal{P}} \in \mathbb{P}(\mathcal{P})$, it holds that*

(31) $$\|\boldsymbol{U}_{\widehat{\mathcal{T}}}[Q_{\mathcal{P}}] - \boldsymbol{U}_{\mathcal{T}}[Q_{\mathcal{P}}]\|_{\mathbb{V}} \le C_{\mathrm{drel}}\, \eta(\mathcal{T} \setminus \widehat{\mathcal{T}}; \boldsymbol{U}[Q_{\mathcal{P}}], Q_{\mathcal{P}}).$$

The constant C_{drel} depends only on γ-shape regularity. ■

Next, we note that the estimator depends Lipschitz continuously on the arguments. The result is slightly stronger than [KS08, Prop. 5.4], but the proof is standard [CKNS08].

Lemma 3.3 (*stability* [CKNS08, Prop. 3.3]). *Let $\widehat{\mathcal{T}} \in \mathbb{T}^{\mathrm{c}}(\mathcal{T})$. There exists a constant $C_{\mathrm{stab}} > 0$ such that, for all $\boldsymbol{V}_{\widehat{\mathcal{T}}} \in \mathbb{V}(\widehat{\mathcal{T}})$, $\boldsymbol{W}_{\mathcal{T}} \in \mathbb{V}(\mathcal{T})$, and $Q_{\mathcal{P}}, R_{\mathcal{P}} \in \mathbb{P}(\mathcal{P})$, it holds that*

(32) $$|\eta(\mathcal{T} \cap \widehat{\mathcal{T}}; \boldsymbol{V}_{\widehat{\mathcal{T}}}, Q_{\mathcal{P}}) - \eta(\mathcal{T} \cap \widehat{\mathcal{T}}; \boldsymbol{W}_{\mathcal{T}}, R_{\mathcal{P}})| \le C_{\mathrm{stab}} \left(\|\boldsymbol{V}_{\widehat{\mathcal{T}}} - \boldsymbol{W}_{\mathcal{T}}\|_{\mathbb{V}} + \|Q_{\mathcal{P}} - R_{\mathcal{P}}\|_{\mathbb{P}} \right).$$

The constant C_{stab} depends only on the polynomial degree m and γ-shape regularity. ■

The following reduction property follows from the reduction of the mesh-size on refined elements. The proof is standard [CKNS08].

Lemma 3.4 (*reduction* [CKNS08, Proof of Cor. 3.4]). *Let $\widehat{\mathcal{T}} \in \mathbb{T}^{\mathrm{c}}(\mathcal{T})$. Let $Q_{\mathcal{P}} \in \mathbb{P}(\mathcal{P})$. Then, with $q_{\mathrm{red}} = 2^{-1/(n+1)}$, there holds the reduction property*

(33) $$\eta(\widehat{\mathcal{T}} \setminus \mathcal{T}; \boldsymbol{U}_{\widehat{\mathcal{T}}}[Q_{\mathcal{P}}], Q_{\mathcal{P}}) \le q_{\mathrm{red}}\, \eta(\mathcal{T} \setminus \widehat{\mathcal{T}}; \boldsymbol{U}_{\mathcal{T}}[Q_{\mathcal{P}}], Q_{\mathcal{P}}) + C_{\mathrm{red}}\, \|\boldsymbol{U}_{\widehat{\mathcal{T}}}[Q_{\mathcal{P}}] - \boldsymbol{U}_{\mathcal{T}}[Q_{\mathcal{P}}]\|_{\mathbb{V}}.$$

The constant $C_{\mathrm{red}} > 0$ depends only on the polynomial degree m and γ-shape regularity. ■

Finally, for the divergence contribution to the Stokes error estimator, we recall the following equivalence. The result is slightly stronger than [KS08, Prop. 5.7].

Lemma 3.5. *Let $C_{\mathrm{div}} \ge 1$ be the norm equivalence constant from (10). Let $\Pi_{\mathcal{T}} : L^2(\Omega) \to \mathbb{P}(\mathcal{T})$ be the $L^2(\Omega)$-orthogonal projection. If $Q_{\mathcal{P}} \in \mathbb{P}(\mathcal{P})$, then it holds that*

(34) $$\|\Pi_{\mathcal{T}} \nabla \cdot \boldsymbol{u}[Q_{\mathcal{P}}]\|_{\Omega} \le \|\nabla \cdot (\boldsymbol{u}_{\mathcal{T}} - \boldsymbol{u}[Q_{\mathcal{P}}])\|_{\Omega} \le \|p_{\mathcal{T}} - Q_{\mathcal{P}}\|_{\mathbb{P}} \le C_{\mathrm{div}}\, \|\Pi_{\mathcal{T}} \nabla \cdot \boldsymbol{u}[Q_{\mathcal{P}}]\|_{\Omega}.$$

Moreover, it holds that

(35) $$\|\nabla \cdot \boldsymbol{u}[Q_{\mathcal{P}}]\|_{\Omega} \le \|p - Q_{\mathcal{P}}\|_{\mathbb{P}} \le C_{\mathrm{div}}\, \|\nabla \cdot \boldsymbol{u}[Q_{\mathcal{P}}]\|_{\Omega}.$$

Proof. From the definition of the Schur complement operator, we have that

(36) $$\nabla \cdot (\boldsymbol{u}_{\mathcal{T}} - \boldsymbol{u}[Q_{\mathcal{P}}]) = S(p_{\mathcal{T}} - Q_{\mathcal{P}}).$$

Taking into account (10), we obtain that

$$
\begin{aligned}
\|\nabla\cdot(\boldsymbol{u}_{\mathcal T}-\boldsymbol{u}[Q_{\mathcal P}])\|_\Omega^2 &\overset{(36)}{=} \langle S(p_{\mathcal T}-Q_{\mathcal P})\,,\,\nabla\cdot(\boldsymbol{u}_{\mathcal T}-\boldsymbol{u}[Q_{\mathcal P}])\rangle_\Omega\\
&= \langle p_{\mathcal T}-Q_{\mathcal P}\,,\,\nabla\cdot(\boldsymbol{u}_{\mathcal T}-\boldsymbol{u}[Q_{\mathcal P}])\rangle_{\mathbb P} \le \|p_{\mathcal T}-Q_{\mathcal P}\|_{\mathbb P}\,\|\nabla\cdot(\boldsymbol{u}_{\mathcal T}-\boldsymbol{u}[Q_{\mathcal P}])\|_{\mathbb P}\\
&\overset{(10)}{\le} \|p_{\mathcal T}-Q_{\mathcal P}\|_{\mathbb P}\,\|\nabla\cdot(\boldsymbol{u}_{\mathcal T}-\boldsymbol{u}[Q_{\mathcal P}])\|_\Omega.
\end{aligned}
$$

Together with $\Pi_{\mathcal T}\nabla\cdot\boldsymbol{u}_{\mathcal T}=0$, this proves that

$$
\|\Pi_{\mathcal T}\nabla\cdot\boldsymbol{u}[Q_{\mathcal P}]\|_\Omega \le \|\nabla\cdot(\boldsymbol{u}_{\mathcal T}-\boldsymbol{u}[Q_{\mathcal P}])\|_\Omega \le \|p_{\mathcal T}-Q_{\mathcal P}\|_{\mathbb P}.
$$

On the other hand, note that $\mathcal T\in\mathbb T^{\rm nc}(\mathcal P)$ implies that $\Pi_{\mathcal T}(p_{\mathcal T}-Q_{\mathcal P})=p_{\mathcal T}-Q_{\mathcal P}$. The norm equivalence (10) and the Cauchy-Schwarz inequality thus imply that

$$
\begin{aligned}
C_{\rm div}\|p_{\mathcal T}-Q_{\mathcal P}\|_{\mathbb P}\,\|\Pi_{\mathcal T}\nabla\cdot\boldsymbol{u}[Q_{\mathcal P}]\|_\Omega &\overset{(10)}{\ge} \|p_{\mathcal T}-Q_{\mathcal P}\|_\Omega\,\|\Pi_{\mathcal T}\nabla\cdot\boldsymbol{u}[Q_{\mathcal P}]\|_\Omega\\
&\ge -\langle p_{\mathcal T}-Q_{\mathcal P}\,,\,\Pi_{\mathcal T}\nabla\cdot\boldsymbol{u}[Q_{\mathcal P}]\rangle_\Omega = \langle p_{\mathcal T}-Q_{\mathcal P}\,,\,\Pi_{\mathcal T}\nabla\cdot(\boldsymbol{u}_{\mathcal T}-\boldsymbol{u}[Q_{\mathcal P}])\rangle_\Omega\\
&= \langle p_{\mathcal T}-Q_{\mathcal P}\,,\,\nabla\cdot(\boldsymbol{u}_{\mathcal T}-\boldsymbol{u}[Q_{\mathcal P}])\rangle_\Omega \overset{(36)}{=} \langle S(p_{\mathcal T}-Q_{\mathcal P})\,,\,p_{\mathcal T}-Q_{\mathcal P}\rangle_\Omega = \|p_{\mathcal T}-Q_{\mathcal P}\|_{\mathbb P}^2
\end{aligned}
$$

and therefore $\|p_{\mathcal T}-Q_{\mathcal P}\|_{\mathbb P}\le C_{\rm div}\|\Pi_{\mathcal T}\nabla\cdot\boldsymbol{u}[Q_{\mathcal P}]\|_\Omega$. This concludes the proof of (34). The proof of (35) follows along the same lines (with $p=p_{\mathcal T}$ and hence $0=\nabla\cdot\boldsymbol{u}=\nabla\cdot\boldsymbol{u}_{\mathcal T}$). ■

3.2. Adaptive refinement of pressure triangulation. To refine the partitions $\mathcal P_i$, we apply the following algorithm from [Bin15, Section 2] (which slightly differs from the well-known *thresholding second algorithm* of [BD04]):

Algorithm 3.6. INPUT: *Partition* $\mathcal P':=\mathcal P\in\mathbb T^{\rm nc}$, *triangulation* $\mathcal T\in\mathbb T^{\rm nc}(\mathcal P)\cap\mathbb T^{\rm c}$, *function* $\boldsymbol V_{\mathcal T}\in\mathbb V(\mathcal T)$, *adaptivity parameter* $0<\vartheta\le 1$.

LOOP: *Iterate the following steps* (i)–(iii) until $\vartheta\,\|\nabla\cdot\boldsymbol V_{\mathcal T}\|_\Omega\le\|\Pi_{\mathcal P'}\nabla\cdot\boldsymbol V_{\mathcal T}\|_\Omega$:

(i) Compute $e(T):=\inf\{\|\nabla\cdot\boldsymbol V_{\mathcal T}-Q\|_T^2\ :\ Q$ polynomial of degree $m-1\}$ for all $T\in\mathcal P'$, for which $e(T)$ has not been already computed.

(ii) For all $T\in\mathcal P'$ for which $\widetilde e(T)$ has not been already defined, define $\widetilde e(T):=e(T)$ if $T\in\mathcal P$ and $\widetilde e(T):=e(T)\widetilde e(\widetilde T)/(e(T)+\widetilde e(\widetilde T))$ otherwise, where $\widetilde T$ denotes the unique father element of T.

(iii) Choose one element $T\in\mathcal P'$ with $\widetilde e(T)=\max_{T'\in\mathcal P'}\widetilde e(T')$ and employ newest vertex bisection to generate $\mathcal P':=\texttt{bisect}(\mathcal P',\{T\})$.

OUTPUT: Triangulation $\mathcal P'=\texttt{bisect}(\mathcal P,\mathcal T,\boldsymbol V_{\mathcal T};\vartheta)\in\mathbb T^{\rm nc}(\mathcal P)$ with $\mathcal T\in\mathbb T^{\rm nc}(\mathcal P')\cap\mathbb T^{\rm c}$.

According to [Bin15, Theorem 2.1], the output $\mathcal P'$ is a quasi-optimal mesh in $\mathbb T^{\rm nc}(\mathcal P)$ with $\vartheta\,\|\nabla\cdot\boldsymbol V_{\mathcal T}\|_\Omega\le\|\Pi_{\mathcal P'}\nabla\cdot\boldsymbol V_{\mathcal T}\|_\Omega$: This means that for all $\vartheta<\vartheta'<1$ and all $\widetilde{\mathcal P}\in\mathbb T^{\rm nc}(\mathcal P)$ with $\vartheta'\,\|\nabla\cdot\boldsymbol V_{\mathcal T}\|_\Omega\le\|\Pi_{\widetilde{\mathcal P}}\nabla\cdot\boldsymbol V_{\mathcal T}\|_\Omega$, it holds that $\#\mathcal P'-\#\mathcal P\le C_{\rm bin}\,(\#\widetilde{\mathcal P}-\#\mathcal P)$ for some $C_{\rm bin}>1$, which depends only on the ratio $(1-\vartheta'^2)/(1-\vartheta^2)$. The same reference states that the effort of Algorithm 3.6 is $\mathcal O(\#\mathcal T\log(\#\mathcal T))$ if $0<\vartheta<1$.

To obtain optimal algebraic convergence rates of the error estimator, one has to choose ϑ sufficiently small and ϑ' sufficiently close to ϑ; see Theorem 5.3 below.

3.3. Adaptive Uzawa algorithm. We investigate the following adaptive Uzawa algorithm, which goes back to [KS08, Section 7].

Algorithm 3.7. INPUT: *Conforming initial triangulation* $\mathcal{P}_0 := \mathcal{T}_{000} := \mathcal{T}_{\rm init}$ *of* Ω, *initial approximation* $P_{00} = 0$, *counters* $i = j = k = 0$, *adaptivity parameters* $0 \le \kappa_1 < 1$, $0 < \kappa_2 < 1$, $0 < \kappa_3 < 1$, $0 < \vartheta \le 1$, $0 < \theta \le 1$, *and* $C_{\rm mark} \ge 1$.

LOOP: *Iterate the following steps* (i)–(iv)*:*

(i) *Compute* $\boldsymbol{U}_{ijk} \in \mathbb{V}_{ijk}$ *as well as (all local contributions of) the corresponding error estimator* $\eta_{ijk} = \eta(\mathcal{T}_{ijk}; \boldsymbol{U}_{ijk}, P_{ij})$ *such that the exact Galerkin approximation* $\boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}] \in \mathbb{V}_{ijk}$ *of* $\boldsymbol{u}_{ij}$ *satisfies that* $\|\boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}] - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} \le \kappa_1\, \eta_{ijk}$.

(ii) `while` $\quad \eta_{ijk} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega \le \kappa_2 \big(\eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega\big) \quad$ `do`

- *Determine* $\mathcal{P}_{i+1} := \texttt{bisect}(\mathcal{P}_i, \mathcal{T}_{ijk}, \boldsymbol{U}_{ijk}; \vartheta)$ *by Algorithm 3.6.*
- *Define* $\mathcal{M}_{ijk} := \emptyset$, $P_{(i+1)0} := P_{ij}$, *and* $\mathcal{T}_{(i+1)00} := \mathcal{T}_{ijk}$.
- *Update counters* $(i,j,k) \mapsto (i+1, 0, 0)$.

`end while`

(iii) `if` $\quad \eta_{ijk} \le \kappa_3\, \|\Pi_i \nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega \quad$ `then`

- *Define* $\mathcal{M}_{ijk} := \emptyset$, $P_{i(j+1)} := P_{ij} - \Pi_i \nabla \cdot \boldsymbol{U}_{ijk} \in \mathbb{P}_i$, *and* $\mathcal{T}_{i(j+1)0} := \mathcal{T}_{ijk}$.
- *Update counters* $(i,j,k) \mapsto (i, j+1, 0)$.

(iv) `else`

- *Determine a set* $\mathcal{M}_{ijk} \subseteq \mathcal{T}_{ijk}$ *of (up to the fixed factor* $C_{\rm mark}$*) minimal cardinality, which satisfies the Dörfler marking criterion*

$$\theta\, \eta_{ijk}^2 \le \eta(\mathcal{M}_{ijk}; P_{ij}, \boldsymbol{U}_{ijk})^2. \tag{37}$$

- *Generate* $\mathcal{T}_{ij(k+1)} := \texttt{refine}(\mathcal{T}_{ijk}, \mathcal{M}_{ijk})$.
- *Update counters* $(i,j,k) \mapsto (i, j, k+1)$.

`end if` ■

Remark 3.8. *The actual implementation of Algorithm 3.7 will replace the triple indices* (i,j,k) *by one single index* $n \in \mathbb{N}_0$, *which is increased in each step* (ii)–(iv). *However, the present statement of the algorithm makes the numerical analysis more accessible.* ■

Lemma 3.9. *Define the index set* $\mathcal{Q} := \{(i,j,k) \in \mathbb{N}_0^3 : \boldsymbol{U}_{ijk}$ *is defined by Algorithm 3.7*$\}$. *Then, for* $(i,j,k) \in \mathbb{N}_0^3$, *there hold the following assertions* (a)–(c)*:*

(a) *If* $(i,j,k+1) \in \mathcal{Q}$, *then* $(i,j,k) \in \mathcal{Q}$.

(b) *If* $(i,j{+}1,0) \in \mathcal{Q}$, *then* $(i,j,0) \in \mathcal{Q}$ *and* $\underline{k}(i,j) := \max\{k \in \mathbb{N}_0 \colon (i,j,k) \in \mathcal{Q}\} < \infty$.

(c) *If* $(i{+}1,0,0) \in \mathcal{Q}$, *then* $(i,0,0) \in \mathcal{Q}$ *and* $\underline{j}(i) := \max\{j \in \mathbb{N}_0 : (i,j,0) \in \mathcal{Q}\} < \infty$.

Throughout, we shall make the following conventions for the triple index: If we write $\eta_{ij\underline{k}}$ *etc. (see, e.g., Lemma 4.5), then (implicitly)* $\underline{k} = \underline{k}(i,j)$. *If we write* $\eta_{i\underline{j}\underline{k}}$ *etc. (see, e.g., Lemma 4.6), then (implicitly)* $\underline{j} = \underline{j}(i)$ *and* $\underline{k} = \underline{k}(i,\underline{j})$.

Proof. Each step (ii)–(iv) of the algorithm increases either i or j or k by one. ■

Remark 3.10. *Unlike the algorithm from* [KS08], *our formulation of the adaptive Uzawa algorithm avoids any special treatment of the data oscillations (i.e., to resolve* $\boldsymbol{f}$ *by a* *piecewise polynomial in an additional outer loop). This is achieved by the fact that our analysis avoids to exploit any efficiency of the* a posteriori *estimator η.* ■

Remark 3.11. *We note that the choice $\boldsymbol{U}_{ijk} := \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}]$ (i.e., $\kappa_1 = 0$) is admissible in step* (i) *of Algorithm 3.7. In the spirit of* [FHPS18], *one can also employ the PCG algorithm* [GVL13, Algorithm 11.5.1] *with optimal preconditioner. With κ_1' and an additional index $\ell \in \mathbb{N}_0$ for the PCG iteration and initially $\ell := 0$, repeat the following three steps, until $U_{ijk} := U_{ijk(\ell+1)}$ satisfies $\|\boldsymbol{U}_{ijk(\ell+1)} - \boldsymbol{U}_{ijk\ell}\|_{\mathbb{V}} \le \kappa_1' \, \eta_{ijk(\ell+1)}$:*

- *Do one PCG step to obtain $\boldsymbol{U}_{ijk(\ell+1)} \in \mathbb{V}_{ijk}$ from $\boldsymbol{U}_{ijk\ell} \in \mathbb{V}_{ijk}$.*
- *Compute (all local contributions of) the estimator $\eta_{ijk(\ell+1)} := \eta(\mathcal{T}_{ijk}; \boldsymbol{U}_{ijk(\ell+1)}, P_{ij})$.*
- *Update counters $(i,j,k,\ell) \mapsto (i,j,k,\ell+1)$.*

If the preconditioner is optimal, i.e., the preconditioned linear system has uniformly bounded condition number, then it follows that PCG is a uniform contraction [FHPS18, Section 2.6]*: There exists $0 < q_{\rm pcg} < 1$ such that*

$$\|U_{\mathcal{T}_{ijk}}[P_{ij}] - \boldsymbol{U}_{ijk(\ell+1)}\|_{\mathbb{V}} \le q_{\rm pcg} \, \|U_{\mathcal{T}_{ijk}}[P_{ij}] - \boldsymbol{U}_{ijk\ell}\|_{\mathbb{V}} \quad \text{for all } \ell \in \mathbb{N}_0.$$

Hence, the PCG loop terminates, and the triangle inequality proves that

$$\|U_{\mathcal{T}_{ijk}}[P_{ij}] - \boldsymbol{U}_{ijk(\ell+1)}\|_{\mathbb{V}} \le \frac{q_{\rm pcg}}{1-q_{\rm pcg}} \, \|\boldsymbol{U}_{ijk(\ell+1)} - \boldsymbol{U}_{ijk\ell}\|_{\mathbb{V}} \le \frac{q_{\rm pcg}}{1-q_{\rm pcg}} \, \kappa_1' \, \eta_{ijk(\ell+1)},$$

i.e., the criterion of step (i) *of Algorithm 3.7 is satisfied for $\kappa_1 := \kappa_1' q_{\rm pcg}/(1-q_{\rm pcg})$.* ■

---

4. Convergence

4.1. Main theorem on linear convergence. To state linear convergence, we need an ordering of the set $\mathcal{Q}$ from Lemma 3.9: For $(i,j,k), (i',j',k') \in \mathcal{Q}$, write $(i',j',k') < (i,j,k)$ if the index (i',j',k') appears earlier in Algorithm 3.7 than (i,j,k). Define

$$|(i,j,k)| := \#\{(i',j',k') \in \mathcal{Q} \, : \, (i',j',k') < (i,j,k)\} \in \mathbb{N}_0. \tag{38}$$

Note that $|(i,j,k)|$ coincides with the single index n from Remark 3.8. Then, we have the following theorem. The proof is given in Section 4.3.

Theorem 4.1. *Let $0 < \kappa_1 < \theta^{1/2}/C_{\rm stab}$. Suppose that $0 < \kappa_2, \kappa_3 < 1$ are sufficiently small as in Lemma 4.5 and Lemma 4.6 below. Let $0 < \vartheta \le 1$ and $0 < \theta \le 1$. Then, there exist constants $C_{\rm lin} > 0$ and $0 < q_{\rm lin} < 1$ such that*

$$\eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega \le C_{\rm lin} q_{\rm lin}^{|(i,j,k)| - |(i',j',k')|} \left(\eta_{i'j'k'} + \|\nabla \cdot \boldsymbol{U}_{i'j'k'}\|_\Omega\right) \tag{39}$$

for all $(i',j',k'), (i,j,k) \in \mathcal{Q}$ with $(i',j',k') < (i,j,k)$. The constants $C_{\rm lin}$ and $q_{\rm lin}$ depend only on the domain Ω, γ-shape regularity, the polynomial degree m, and the parameters κ_1, κ_2, κ_3, ϑ, and θ.

Remark 4.2. *The adaptive Uzawa algorithm from* [BMN02] *employs only one triangulation for both, the pressure and the velocity. Similarly, we can additionally update $\mathcal{P}_i := \mathcal{T}_{ij(k+1)}$ in step* (iv) *of Algorithm 3.7. Since $0 < \kappa_2 < 1$ and $\Pi_i \nabla \cdot U_{ijk} = \nabla U_{ijk}$,* *then the condition in* (ii) *will always fail. We note that the convergence analysis of Section 4.2 and in particular, linear convergence (Theorem 4.1) clearly remain valid for this modified algorithm, while our proof of optimal convergence rates (Theorem 5.3) fails.* ■

4.2. Auxiliary results. The first lemma provides links between the exact Galerkin solutions $\boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}]$ and its approximations $\boldsymbol{U}_{ijk}$.

Lemma 4.3. *Let $(i,j,k) \in \mathcal{Q}$. For all $\mathcal{S} \subseteq \mathcal{T}_{ijk}$, it holds that*

$$|\eta(\mathcal{S}; \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}], P_{ij}) - \eta(\mathcal{S}; \boldsymbol{U}_{ijk}, P_{ij})| \le \kappa_1 \, C_{\mathrm{stab}} \, \eta_{ijk}, \tag{40}$$

where $C_{\mathrm{stab}} > 0$ is the constant from Lemma 3.3. This particularly yields the equivalence

$$(1 - \kappa_1 \, C_{\mathrm{stab}}) \, \eta_{ijk} \le \eta(\mathcal{T}_{ijk}; \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}], P_{ij}) \le (1 + \kappa_1 \, C_{\mathrm{stab}}) \, \eta_{ijk}. \tag{41}$$

as well as the reliability estimates

$$\|\boldsymbol{u}_{ij} - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} \le C'_{\mathrm{rel}}(\kappa_1) \, \eta_{ijk}, \tag{42}$$

$$\|\boldsymbol{u}_i - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} + \|p_i - P_{ij}\|_{\mathbb{P}} \le C'_{\mathrm{rel}}(\kappa_1) \, \big(\eta_{ijk} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{ijk}\|_{\Omega}\big), \tag{43}$$

$$\|\boldsymbol{u} - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} + \|p - P_{ij}\|_{\mathbb{P}} \le C'_{\mathrm{rel}}(\kappa_1) \, \big(\eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_{\Omega}\big), \tag{44}$$

where $C'_{\mathrm{rel}}(\kappa_1) := ((1 + \kappa_1 C_{\mathrm{stab}}) C_{\mathrm{rel}} + \kappa_1 (C_{\mathrm{rel}} + 1)) \ge C_{\mathrm{rel}}$ with $C_{\mathrm{rel}} > 0$ from Lemma 3.1.

Proof. To shorten notation, we set $\eta^\star_{ijk} := \eta(\mathcal{T}_{ijk}; \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}], P_{ij})$. The stability (40) follows from Lemma 3.3 and $\|\boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}] - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} \le \kappa_1 \, \eta_{ijk}$, which is guaranteed by step (i) of Algorithm 3.7. Taking $\mathcal{S} = \mathcal{T}_{ijk}$, (41) is an immediate consequence. To see (42), we use reliability (28), step (i) of Algorithm 3.7, and (41) to see that

$$\|\boldsymbol{u}_{ij} - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} \overset{(28)}{\le} C_{\mathrm{rel}} \, \eta^\star_{ijk} + \|\boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}] - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} \overset{(41)}{\le} ((1 + \kappa_1 C_{\mathrm{stab}}) C_{\mathrm{rel}} + \kappa_1) \, \eta_{ijk}.$$

To prove (43), we apply (29)

$$\begin{aligned} \|\boldsymbol{u}_i - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} + \|p_i - P_{ij}\|_{\mathbb{P}} &\overset{(29)}{\le} C_{\mathrm{rel}} \big(\eta^\star_{ijk} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}]\|_{\Omega}\big) + \|\boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}] - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} \\ &\overset{(41)}{\le} ((1 + \kappa_1 C_{\mathrm{stab}}) C_{\mathrm{rel}} + \kappa_1) \, \eta_{ijk} + C_{\mathrm{rel}} \, \|\Pi_i \nabla \cdot \boldsymbol{U}_{ijk}\|_{\Omega}. \end{aligned}$$

Similarly, (44) follows from (30). ■

The following three lemmas prove that Algorithm 3.7 leads to contraction if either i, j, or k is increased. Throughout, let $0 < \vartheta \le 1$, $0 < \theta \le 1$, and, if not stated otherwise, $0 \le \kappa_1 < 1$, $0 < \kappa_2, \kappa_3 < 1$.

Lemma 4.4. *Let $(i,j,0) \in \mathcal{Q}$ and define $\underline{k} := \max\{k \in \mathbb{N}_0 : (i,j,k) \in \mathcal{Q}\} \in \mathbb{N}_0 \cup \{\infty\}$. If $0 \le \kappa_1 < \theta^{1/2}/C_{\mathrm{stab}}$, then, there exist constants $0 < q_1 < 1$ and $C_1 > 0$, which depend only on γ-shape regularity, the polynomial degree m, κ_1, and θ, such that*

$$\eta_{ij(k+n)} \le C_1 \, q_1^n \, \eta_{ijk} \quad \textit{for all } k, n \in \mathbb{N}_0 \textit{ with } k \le k + n \le \underline{k}. \tag{45}$$

Moreover, it holds that

$$\eta_{ijk} \le \eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_{\Omega} \le \frac{1}{\kappa_2} \Big(1 + \frac{1}{\kappa_3}\Big) \, \eta_{ijk} \quad \textit{for all } 0 \le k < \underline{k}. \tag{46}$$

If $\underline{k} = \infty$, this yields that $\|\boldsymbol{u} - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} + \|p - P_{ij}\|_{\mathbb{P}} \to 0$ as $k \to \infty$ with $p = p_i = P_{ij}$.

Proof. We split the proof into three steps.

Step 1. If $\boldsymbol{U}_{ijk} = \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}]$ for all $(i,j,k) \in \mathcal{Q}$, step (iv) of Algorithm 3.7 is the usual adaptive step in an adaptive algorithm for, e.g., the (vector-valued) Poisson model problem. Hence, (45) follows from reliability (28), stability (32) and reduction (33); see, e.g., [CFPP14, Theorem 4.1 (i)]. For general $\boldsymbol{U}_{ijk}$, (45) follows from [CFPP14, Theorem 7.2] under the constraint $0 \le \kappa_1 < \theta^{1/2}/C_{\rm stab}$.

Step 2. If $k < \underline{k}$, the structure of Algorithm 3.7 implies that the conditions in step (ii) and (iii) are false, i.e.,

$$\eta_{ijk} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega > \kappa_2 \big(\eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega\big) \quad \text{and} \quad \eta_{ijk} > \kappa_3 \, \|\Pi_i \nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega.$$

Hence,

$$\eta_{ijk} \le \eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega < \frac{1}{\kappa_2} \big(\eta_{ijk} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega\big) < \frac{1}{\kappa_2}\Big(1 + \frac{1}{\kappa_3}\Big)\, \eta_{ijk}$$

which proves (46).

Step 3. For $\underline{k} = \infty$, the estimates (45)–(46) imply that

$$\|\boldsymbol{u} - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} + \|p - P_{ij}\|_{\mathbb{P}} \overset{(44)}{\lesssim} \eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega \overset{(46)}{\simeq} \eta_{ijk} \xrightarrow{k \to \infty} 0.$$

Note that $\underline{k} = \infty$ also implies that neither i nor j are increased, i.e., P_{ij} remains constant as $k \to \infty$. Hence, $p = P_{ij} \in \mathbb{P}_i$ and therefore also $p = p_i$. ∎

Lemma 4.5. *Let $(i,0,0) \in \mathcal{Q}$ and define $\underline{j} := \max\{j \in \mathbb{N}_0 \,:\, (i,j,0) \in \mathcal{Q}\} \in \mathbb{N}_0 \cup \{\infty\}$. If $0 < \kappa_3 \ll 1$ is sufficiently small (see (55) in the proof below), then there exist constants $0 < q_2 < 1$ and $C_2 > 0$ such that*

$$(47) \qquad \|p_i - P_{i(j+n)}\|_{\mathbb{P}} \le q_2^n \, \|p_i - P_{ij}\|_{\mathbb{P}} \quad \textit{for all } j, n \in \mathbb{N}_0 \textit{ with } j \le j + n \le \underline{j}.$$

Moreover, it holds that

$$(48) \qquad C_2^{-1} \, \|p_i - P_{ij}\|_{\mathbb{P}} \le \eta_{ij\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{ij\underline{k}}\|_\Omega \le C_2 \, \|p_i - P_{ij}\|_{\mathbb{P}} \quad \textit{for all } 0 \le j < \underline{j}.$$

If $\underline{j} = \infty$, this yields convergence $\|\boldsymbol{u} - \boldsymbol{U}_{ij\underline{k}}\|_{\mathbb{V}} + \|p - P_{ij}\|_{\mathbb{P}} \to 0$ as $j \to \infty$. While q_2 depends only on the domain Ω, γ-shape regularity, κ_1, and κ_3, the constant C_2 depends additionally on κ_2.

Proof. We split the proof into three steps.

Step 1. If $j < \underline{j}(i)$ and $k = \underline{k}(i,j)$, then necessarily $\underline{k}(i,j) < \infty$. The structure of Algorithm 3.7 implies that the condition in step (ii) is false, while the condition in step (iii) is true, i.e.,

$$(49) \quad \eta_{ij\underline{k}} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{ij\underline{k}}\|_\Omega > \kappa_2 \, (\eta_{ij\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{ij\underline{k}}\|_\Omega) \quad \text{and} \quad \eta_{ij\underline{k}} \le \kappa_3 \, \|\Pi_i \nabla \cdot \boldsymbol{U}_{ij\underline{k}}\|_\Omega.$$

First, this proves that

$$(50) \quad \begin{aligned} \kappa_2 \, (\eta_{ij\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{ij\underline{k}}\|_\Omega) &< \eta_{ij\underline{k}} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{ij\underline{k}}\|_\Omega \le (1+\kappa_3) \, \|\Pi_i \nabla \cdot \boldsymbol{U}_{ij\underline{k}}\|_\Omega \\ &\le (1+\kappa_3) \, \|\nabla \cdot \boldsymbol{U}_{ij\underline{k}}\|_\Omega \le (1+\kappa_3) \, (\eta_{ij\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{ij\underline{k}}\|_\Omega)\,. \end{aligned}$$

Second, reliability (42) gives that

(51)
$$\|\Pi_i \nabla\cdot(\boldsymbol{u}_{ij}-\boldsymbol{U}_{ij\underline{k}})\|_\Omega \le \|\boldsymbol{u}_{ij}-\boldsymbol{U}_{ij\underline{k}}\|_{\mathbb{V}} \overset{(42)}{\le} C'_{\mathrm{rel}}(\kappa_1)\,\eta_{ijk} \overset{(49)}{\le} \kappa_3 C'_{\mathrm{rel}}(\kappa_1)\,\|\Pi_i\nabla\cdot\boldsymbol{U}_{ij\underline{k}}\|_\Omega.$$

The triangle inequality yields that

$$(1-\kappa_3 C'_{\mathrm{rel}}(\kappa_1))\,\|\Pi_i\nabla\cdot\boldsymbol{U}_{ij\underline{k}}\|_\Omega \overset{(51)}{\le} \|\Pi_i\nabla\cdot\boldsymbol{u}_{ij}\|_\Omega \overset{(51)}{\le} (1+\kappa_3 C'_{\mathrm{rel}}(\kappa_1))\|\Pi_i\nabla\cdot\boldsymbol{U}_{ij\underline{k}}\|_\Omega. \tag{52}$$

This leads us to

$$\begin{aligned} C_{\mathrm{div}}^{-1}\frac{1-\kappa_3 C'_{\mathrm{rel}}(\kappa_1)}{1+\kappa_3 C'_{\mathrm{rel}}(\kappa_1)}\|p_i-P_{ij}\|_{\mathbb{P}} &\overset{(34)}{\le} \frac{1-\kappa_3 C'_{\mathrm{rel}}(\kappa_1)}{1+\kappa_3 C'_{\mathrm{rel}}(\kappa_1)}\,\|\Pi_i\nabla\cdot\boldsymbol{u}_{ij}\|_\Omega \\ &\overset{(52)}{\le} (1-\kappa_3 C'_{\mathrm{rel}}(\kappa_1))\,\|\Pi_i\nabla\cdot\boldsymbol{U}_{ij\underline{k}}\|_\Omega \overset{(52)}{\le} \|\Pi_i\nabla\cdot\boldsymbol{u}_{ij}\|_\Omega \overset{(34)}{\le} \|p_i-P_{ij}\|_{\mathbb{P}}. \end{aligned} \tag{53}$$

If $\kappa_3 C'_{\mathrm{rel}}(\kappa_1)<1$, the combination of (53) and (50) proves (48).

Step 2. Starting from P_{ij}, one step of the *exact* Uzawa iteration for the *reduced* Stokes problem (leading to the auxiliary quantity $p_{i(j+1)}$) guarantees the existence of some $0<q_{\mathrm{Uzawa}}<1$ such that the following contraction holds (see [KS08, Eq. (4.3)]):

$$\|p_i-p_{i(j+1)}\|_{\mathbb{P}} \le q_{\mathrm{Uzawa}}\,\|p_i-P_{ij}\|_{\mathbb{P}} \quad \text{with} \quad p_{i(j+1)} = P_{ij}-\Pi_i\nabla\cdot\boldsymbol{u}_{ij}. \tag{54}$$

The contraction constant q_{Uzawa} is the norm of the operator from (4) with $\alpha=1$. Indeed, the proof of (54) works exactly as in Appendix A if $S:\mathbb{P}\to\mathbb{P}$ is replaced by the operator $\Pi_i S:\mathbb{P}_i\to\mathbb{P}_i$. In particular, q_{Uzawa} does neither depend on i nor on j. Since $P_{i(j+1)}=P_{ij}-\Pi_i\nabla\cdot\boldsymbol{U}_{ij\underline{k}}$, we are thus led to

$$\begin{aligned} \|p_i-P_{i(j+1)}\|_{\mathbb{P}} &\le \|p_i-p_{i(j+1)}\|_{\mathbb{P}}+\|p_{i(j+1)}-P_{i(j+1)}\|_{\mathbb{P}} \\ &\le q_{\mathrm{Uzawa}}\,\|p_i-P_{ij}\|_{\mathbb{P}} + \|\Pi_i\nabla\cdot(\boldsymbol{u}_{ij}-\boldsymbol{U}_{ij\underline{k}})\|_{\mathbb{P}} \\ &\overset{(51)}{\le} q_{\mathrm{Uzawa}}\,\|p_i-P_{ij}\|_{\mathbb{P}} + \kappa_3 C'_{\mathrm{rel}}(\kappa_1)\,\|\Pi_i\nabla\cdot\boldsymbol{U}_{ij\underline{k}}\|_\Omega \\ &\overset{(53)}{\le} \left(q_{\mathrm{Uzawa}}+\frac{\kappa_3 C'_{\mathrm{rel}}(\kappa_1)}{1-\kappa_3 C'_{\mathrm{rel}}(\kappa_1)}\right)\|p_i-P_{ij}\|_{\mathbb{P}} =: q_2\,\|p_i-P_{ij}\|_{\mathbb{P}}. \end{aligned}$$

Let $0<\kappa_3\ll 1$ be sufficiently small, i.e.,

$$0<\kappa_3 C'_{\mathrm{rel}}(\kappa_1)<1 \quad\text{and}\quad 0<q_2 := q_{\mathrm{Uzawa}}+\frac{\kappa_3 C'_{\mathrm{rel}}(\kappa_1)}{1-\kappa_3 C'_{\mathrm{rel}}(\kappa_1)}<1. \tag{55}$$

Then, induction proves that $\|p_i-P_{i(j+n)}\|_{\mathbb{P}} \le q_2^n\,\|p_i-P_{ij}\|_{\mathbb{P}}$ for every $j,n\in\mathbb{N}_0$ with $j\le j+n\le \underline{j}$. This proves (47).

Step 3. For $\underline{j}=\infty$, the estimates (47)–(48) imply that

$$\|\boldsymbol{u}-\boldsymbol{U}_{ij\underline{k}}\|_{\mathbb{V}}+\|p-P_{ij}\|_{\mathbb{P}} \overset{(44)}{\lesssim} \eta_{ij\underline{k}}+\|\nabla\cdot\boldsymbol{U}_{ij\underline{k}}\|_\Omega \overset{(48)}{\simeq} \|p_i-P_{ij}\|_{\mathbb{P}} \xrightarrow{j\to\infty} 0.$$

This concludes the proof. ■

Note that $\underline{i}:=\max\{i\in\mathbb{N}_0 : (i,0,0)\in\mathcal{Q}\}<\infty$ in Algorithm 3.7 implies that either $\underline{j}:=\underline{j}(\underline{i})=\infty$ or $\underline{k}(\underline{i},\underline{j})=\infty$. According to Lemma 4.4 (for $\underline{k}=\infty$) and Lemma 4.5 (for $\underline{j}=\infty$), it only remains to analyze the case $\underline{i}=\infty$.

Lemma 4.6. *Let $\underline{i} := \max\{i \in \mathbb{N}_0 : (i,0,0) \in \mathcal{Q}\} \in \mathbb{N}_0 \cup \{\infty\}$. If $0 < \kappa_2 \ll 1$ is sufficiently small (see (61) in the proof below), then there exist constants $0 < q_3 < 1$ and $C_3 > 0$ such that*

$$\|p - P_{(i+n)\underline{j}}\|_{\mathbb{P}} \le q_3^n \, \|p - P_{i\underline{j}}\|_{\mathbb{P}} \quad \text{for all } i, n \in \mathbb{N}_0 \text{ with } i \le i+n \le \underline{i}. \tag{56}$$

Moreover, it holds that

$$C_3^{-1} \, \|p - P_{i\underline{j}}\|_{\mathbb{P}} \le \eta_{i\underline{j}\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega \le C_3 \, \|p - P_{i\underline{j}}\|_{\mathbb{P}} \quad \text{for all } 0 \le i < \underline{i}. \tag{57}$$

While C_3 depends only on the domain Ω, γ-shape regularity, κ_1 and κ_2, the contraction constant q_3 depends additionally on $0 < \vartheta \le 1$. If $\underline{i} = \infty$, this yields convergence $\|\boldsymbol{u} - \boldsymbol{U}_{i\underline{j}\underline{k}}\|_{\mathbb{V}} + \|p - P_{i\underline{j}}\|_{\mathbb{P}} \to 0$ as $i \to \infty$.

Proof. We split the proof into five steps.

Step 1. According to Algorithm 3.7, it holds that

$$\eta_{i\underline{j}\underline{k}} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega \le \kappa_2 \big(\eta_{i\underline{j}\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega\big). \tag{58}$$

For $0 < \kappa_2 < 1$, this implies that

$$\eta_{i\underline{j}\underline{k}} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega \le \frac{\kappa_2}{1-\kappa_2} \, \|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega.$$

Recall that

$$\|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega \le \|\nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega + \|\nabla \cdot (\boldsymbol{u}_{i\underline{j}} - \boldsymbol{U}_{i\underline{j}\underline{k}})\|_\Omega \overset{(42)}{\le} \|\nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega + C'_{\mathrm{rel}}(\kappa_1) \, \eta_{i\underline{j}\underline{k}}$$

We abbreviate $C(\kappa_1, \kappa_2) := C'_{\mathrm{rel}}(\kappa_1)\, \kappa_2/(1-\kappa_2)$. For sufficiently small $0 < \kappa_2 \ll 1$ with $0 < C(\kappa_1,\kappa_2) < 1$, the combination of the last two estimates implies that $\|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega \le (1 - C(\kappa_1,\kappa_2))^{-1} \, \|\nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega$. With

$$C'(\kappa_1, \kappa_2) := \frac{C(\kappa_1,\kappa_2)}{1 - C(\kappa_1,\kappa_2)},$$

we are hence led to

$$\begin{aligned} \|\boldsymbol{u}_{i\underline{j}} - \boldsymbol{U}_{i\underline{j}\underline{k}}\|_{\mathbb{V}} &\overset{(42)}{\le} C'_{\mathrm{rel}}(\kappa_1) \big(\eta_{i\underline{j}\underline{k}} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega\big) \le C(\kappa_1,\kappa_2) \, \|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega \\ &\le C'(\kappa_1,\kappa_2) \, \|\nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega \overset{(35)}{\le} C'(\kappa_1,\kappa_2) \, \|p - P_{i\underline{j}}\|_{\mathbb{P}}. \end{aligned} \tag{59}$$

Conversely,

$$\begin{aligned} \|p - P_{i\underline{j}}\|_{\mathbb{P}} &\overset{(35)}{\le} C_{\mathrm{div}} \|\nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega \le C_{\mathrm{div}} \big(\|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega + \|\nabla \cdot (\boldsymbol{u}_{i\underline{j}} - \boldsymbol{U}_{i\underline{j}\underline{k}})\|_\Omega\big) \\ &\overset{(42)}{\le} \max\{1, C'_{\mathrm{rel}}(\kappa_1)\} \, C_{\mathrm{div}} \big(\|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega + \eta_{i\underline{j}\underline{k}}\big). \end{aligned}$$

In particular, this proves (57).

Step 2. Recall from Step 1 that

$$\begin{aligned} \|\nabla \cdot (\boldsymbol{u}_{i\underline{j}} - \boldsymbol{U}_{i\underline{j}\underline{k}})\|_\Omega + \|\Pi_i \nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega &\overset{(42)}{\le} \max\{1, C'_{\mathrm{rel}}(\kappa_1)\} \big(\eta_{i\underline{j}\underline{k}} + \|\Pi_i \nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega\big) \\ &\overset{(59)}{\le} \max\{1, C'_{\mathrm{rel}}(\kappa_1)\} \, C'(\kappa_1,\kappa_2) \, \|p - P_{i\underline{j}}\|_{\mathbb{P}}. \end{aligned} \tag{60}$$

We hence observe that

$$\|p_i - P_{i\underline{j}}\|_{\mathbb{P}} \overset{(34)}{\leq} C_{\text{div}} \|\Pi_i \nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega \leq C_{\text{div}} \big(\|\Pi_i \nabla \cdot (\boldsymbol{u}_{i\underline{j}} - \boldsymbol{U}_{i\underline{j}\underline{k}})\|_\Omega + \|\Pi_i \nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega \big)$$

$$\overset{(60)}{\leq} C_{\text{div}} \max\{1, C'_{\text{rel}}(\kappa_1)\} \, C'(\kappa_1, \kappa_2) \, \|p - P_{i\underline{j}}\|_{\mathbb{P}}.$$

Step 3. From Algorithm 3.6, we obtain that

$$\vartheta \, \|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega \leq \|\Pi_{i+1} \nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega.$$

According to (59), it holds that

$$\|\nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega \leq \|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega + \|\nabla \cdot (\boldsymbol{u}_{i\underline{j}} - \boldsymbol{U}_{i\underline{j}\underline{k}})\|_\Omega \overset{(59)}{\leq} (1 + C(\kappa_1, \kappa_2)) \, \|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega,$$

as well as

$$\|\Pi_{i+1} \nabla \cdot (\boldsymbol{u}_{i\underline{j}} - \boldsymbol{U}_{i\underline{j}\underline{k}})\|_\Omega \leq \|\boldsymbol{u}_{i\underline{j}} - \boldsymbol{U}_{i\underline{j}\underline{k}}\|_{\mathbb{V}} \overset{(59)}{\leq} C'(\kappa_1, \kappa_2) \, \|\nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega.$$

Combining the last three estimates, we see that

$$\|\Pi_{i+1} \nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega \geq \|\Pi_{i+1} \nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_\Omega - \|\Pi_{i+1} \nabla \cdot (\boldsymbol{u}_{i\underline{j}} - \boldsymbol{U}_{i\underline{j}\underline{k}})\|_\Omega$$

$$\geq \Big(\frac{\vartheta}{1 + C(\kappa_1, \kappa_2)} - C'(\kappa_1, \kappa_2) \Big) \, \|\nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega.$$

Recall the constant $C_{\text{div}} > 0$ from Lemma 3.5. If $0 < \kappa_2 \ll 1$ is sufficiently small, it holds that $C''(\kappa_2) := \big(\frac{\vartheta}{1 + C(\kappa_1, \kappa_2)} - C'(\kappa_1, \kappa_2) \big) / C_{\text{div}} > 0$. This implies that

$$\|p_{i+1} - P_{i\underline{j}}\|_{\mathbb{P}} \overset{(34)}{\geq} \|\Pi_{i+1} \nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega \geq \Big(\frac{\vartheta}{1 + C(\kappa_1, \kappa_2)} - C'(\kappa_1, \kappa_2) \Big) \, \|\nabla \cdot \boldsymbol{u}_{i\underline{j}}\|_\Omega$$

$$\overset{(35)}{\geq} C''(\kappa_2) \, \|p - P_{i\underline{j}}\|_{\mathbb{P}}.$$

Together with the Pythagoras theorem, we are hence led to

$$\|p - p_{i+1}\|_{\mathbb{P}}^2 = \|p - P_{i\underline{j}}\|_{\mathbb{P}}^2 - \|p_{i+1} - P_{i\underline{j}}\|_{\mathbb{P}}^2 \leq (1 - C''(\kappa_2)^2) \, \|p - P_{i\underline{j}}\|_{\mathbb{P}}^2.$$

Step 4. Combining Step 2 and Step 3, we obtain that

$$\|p - P_{(i+1)\underline{j}}\|_{\mathbb{P}}^2 = \|p - p_{i+1}\|_{\mathbb{P}}^2 + \|p_{i+1} - P_{(i+1)\underline{j}}\|_{\mathbb{P}}^2$$

$$\leq \big(1 - C''(\kappa_2)^2\big) \, \|p - P_{i\underline{j}}\|_{\mathbb{P}}^2 + C_{\text{div}}^2 \, \max\{1, C'_{\text{rel}}(\kappa_1)^2\} \, C'(\kappa_1, \kappa_2)^2 \, \|p - P_{(i+1)\underline{j}}\|_{\mathbb{P}}^2.$$

For sufficiently small $0 < \kappa_2 \ll 1$, i.e.,

$$\begin{aligned} &C(\kappa_1, \kappa_2) = \frac{C'_{\text{rel}}(\kappa_1)\kappa_2}{1 - \kappa_2} < 1, \; 0 < C''(\kappa_2) = \Big(\frac{\vartheta}{1 + C(\kappa_1, \kappa_2)} - \frac{C(\kappa_1, \kappa_2)}{1 - C(\kappa_1, \kappa_2)} \Big) C_{\text{div}}^{-1}, \\ &0 < q_3^2 := \frac{1 - C''(\kappa_2)^2}{1 - C_{\text{div}}^2 \, \max\{1, C'_{\text{rel}}(\kappa_1)^2\} \, C'(\kappa_1, \kappa_2)^2} < 1, \end{aligned} \tag{61}$$

we hence see that

$$\|p - P_{(i+1)\underline{j}}\|_{\mathbb{P}}^2 \leq q_3^2 \, \|p - P_{i\underline{j}}\|_{\mathbb{P}}^2.$$

By induction, we conclude (56).

Step 5. For $\underline{i} = \infty$, the estimates (56)–(57) imply that

$$\|\boldsymbol{u} - \boldsymbol{U}_{i\underline{j}\underline{k}}\|_{\mathbb{V}} + \|p - P_{i\underline{j}}\|_{\mathbb{P}} \overset{(44)}{\lesssim} \eta_{i\underline{j}\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{i\underline{j}\underline{k}}\|_{\Omega} \overset{(57)}{\lesssim} \|p - P_{i\underline{j}}\|_{\mathbb{P}} \xrightarrow{i\to\infty} 0.$$

This concludes the proof. ∎

4.3. Proof of Theorem 4.1. To prove Theorem 4.1, we need the following two lemmas. A slightly weaker version of the first lemma is already proved in [CFPP14, Lemma 4.9]. The proof, however, immediately extends to the following generalization and is therefore omitted.

Lemma 4.7. *Let $(a_\ell)_{\ell\in\mathbb{N}_0}$ be a sequence with $a_\ell \geq 0$ for all $\ell \in \mathbb{N}_0$. With the convention $0^{-1/s} := \infty$, the following three statements are pairwise equivalent:*

(a) *There exist a constant $C > 0$ such that $\sum_{n=\ell}^{\infty} a_n \leq C a_\ell$ for all $\ell \in \mathbb{N}_0$.*
(b) *For all $s > 0$, there exists $C > 0$ such that $\sum_{n=0}^{\ell} a_n^{-1/s} \leq C a_\ell^{-1/s}$ for all $\ell \in \mathbb{N}_0$.*
(c) *There exist $0 < q < 1$ and $C > 0$ such that $a_{\ell+n} \leq C q^n a_\ell$ for all $n, \ell \in \mathbb{N}_0$.*

Here, in each statement, the constants $C > 0$ may differ. ∎

Lemma 4.8. *Let $0 < \kappa_1 < \theta^{1/2}/C_{\mathrm{stab}}$. Suppose that κ_2, κ_3 are sufficiently small as in Lemma 4.5 and Lemma 4.6. Let $(i,j,0) \in \mathcal{Q}$. Then, there hold the assertions* (a)–(d)*:*

(a) *If $i \geq 1$, then $\eta_{i00} + \|\nabla \cdot \boldsymbol{U}_{i00}\|_{\Omega} \leq C_{\mathrm{mon}}\big(\eta_{(i-1)\underline{j}\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{(i-1)\underline{j}\underline{k}}\|_{\Omega}\big)$.*
(b) *If $j \geq 1$, then $\eta_{ij0} + \|\nabla \cdot \boldsymbol{U}_{ij0}\|_{\Omega} \leq C_{\mathrm{mon}}\big(\eta_{i(j-1)\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{i(j-1)\underline{k}}\|_{\Omega}\big)$.*
(c) *$\eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_{\Omega} \leq C_{\mathrm{mon}}\big(\eta_{ijk'} + \|\nabla \cdot \boldsymbol{U}_{ijk'}\|_{\Omega}\big)$ for all $0 \leq k' \leq k \leq \underline{k}(i,j)$.*
(d) *$\eta_{ij\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{ij\underline{k}}\|_{\Omega} \leq C_{\mathrm{mon}}\big(\eta_{ij'\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{ij'\underline{k}}\|_{\Omega}\big)$ for all $0 \leq j' \leq j < \underline{j}(i)$.*

The constant $C_{\mathrm{mon}} > 0$ depends only on Ω, C_{stab}, C_{rel}, C_1, and C_2.

Proof. To shorten notation, we set $\eta_{ijk}^\star := \eta(\mathcal{T}_{ijk}; \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}], P_{ij})$ and $\boldsymbol{U}_{ijk}^\star := \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}]$. To prove (a), recall from step (ii) of Algorithm 3.7 that $\mathcal{T}_{i00} = \mathcal{T}_{(i-1)\underline{j}\underline{k}}$ as well as $P_{i0} = P_{(i-1)\underline{j}}$. Hence, $\boldsymbol{U}_{i00}^\star = \boldsymbol{U}_{(i-1)\underline{j}\underline{k}}^\star$ and consequently $\eta_{i00}^\star = \eta_{(i-1)\underline{j}\underline{k}}^\star$ as well as $\|\nabla \cdot \boldsymbol{U}_{i00}^\star\|_{\Omega} = \|\nabla \cdot \boldsymbol{U}_{(i-1)\underline{j}\underline{k}}^\star\|_{\Omega}$. Since $\kappa_1 < \theta^{1/2} C_{\mathrm{stab}}^{-1} \leq C_{\mathrm{stab}}^{-1}$, we can apply the equivalence (41) in both directions. With step (i) of Algorithm 3.7, we see that

$$\begin{aligned}
\eta_{i00} + \|\nabla \cdot \boldsymbol{U}_{i00}\|_{\Omega} &\overset{(41)}{\lesssim} \eta_{i00}^\star + \|\nabla \cdot \boldsymbol{U}_{i00}^\star\|_{\Omega} + \|\boldsymbol{U}_{i00}^\star - \boldsymbol{U}_{i00}\|_{\mathbb{V}} \lesssim \eta_{i00}^\star + \|\nabla \cdot \boldsymbol{U}_{i00}^\star\|_{\Omega} + \eta_{i00} \\
&\overset{(41)}{\lesssim} \eta_{i00}^\star + \|\nabla \cdot \boldsymbol{U}_{i00}^\star\|_{\Omega} = \eta_{(i-1)\underline{j}\underline{k}}^\star + \|\nabla \cdot \boldsymbol{U}_{(i-1)\underline{j}\underline{k}}^\star\|_{\Omega} \overset{(41)}{\lesssim} \eta_{(i-1)\underline{j}\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{(i-1)\underline{j}\underline{k}}\|_{\Omega} \\
&\quad + \|\boldsymbol{U}_{(i-1)\underline{j}\underline{k}}^\star - \boldsymbol{U}_{(i-1)\underline{j}\underline{k}}\|_{\mathbb{V}} \lesssim \eta_{(i-1)\underline{j}\underline{k}} + \|\nabla \cdot \boldsymbol{U}_{(i-1)\underline{j}\underline{k}}\|_{\Omega}.
\end{aligned}$$

To prove (b), recall from step (iii) of Algorithm 3.7 that $\mathcal{T}_{ij0} = \mathcal{T}_{i(j-1)\underline{k}}$ and $P_{ij} = P_{i(j-1)} - \Pi_i \nabla \cdot \boldsymbol{U}_{i(j-1)\underline{k}}$. According to the discrete variational form (19), it holds that

$$a(\boldsymbol{U}_{ij0}^\star - \boldsymbol{U}_{i(j-1)\underline{k}}^\star, \boldsymbol{V}_{ij0}) = b(\boldsymbol{V}_{ij0}, \Pi_i \nabla \cdot \boldsymbol{U}_{i(j-1)\underline{k}}) \quad \text{for all } \boldsymbol{V}_{ij0} \in \mathbb{V}(\mathcal{T}_{ij0}) = \mathbb{V}(\mathcal{T}_{(i-1)\underline{j}\underline{k}}).$$

This proves that $\|\boldsymbol{U}^\star_{ij0} - \boldsymbol{U}^\star_{i(j-1)\underline{k}}\|_{\mathbb{V}} \lesssim \|\Pi_i \nabla\cdot \boldsymbol{U}_{i(j-1)\underline{k}}\|_\Omega \le \|\nabla\cdot \boldsymbol{U}_{i(j-1)\underline{k}}\|_\Omega$. First, it follows that

$$
\begin{aligned}
\|\nabla\cdot \boldsymbol{U}_{ij0}\|_\Omega &\le \|\nabla\cdot \boldsymbol{U}_{i(j-1)\underline{k}}\|_\Omega + \|\boldsymbol{U}_{ij0} - \boldsymbol{U}_{i(j-1)\underline{k}}\|_{\mathbb{V}} \le \|\nabla\cdot \boldsymbol{U}_{i(j-1)\underline{k}}\|_\Omega + \|\boldsymbol{U}^\star_{ij0} - \boldsymbol{U}^\star_{i(j-1)\underline{k}}\|_{\mathbb{V}} \\
&\quad + \|\boldsymbol{U}^\star_{ij0} - \boldsymbol{U}_{ij0}\|_{\mathbb{V}} + \|\boldsymbol{U}^\star_{i(j-1)\underline{k}} - \boldsymbol{U}_{i(j-1)\underline{k}}\|_{\mathbb{V}} \le \|\nabla\cdot \boldsymbol{U}_{i(j-1)\underline{k}}\|_\Omega + \kappa_1\,\eta_{ij0} + \kappa_1\,\eta_{i(j-1)\underline{k}}.
\end{aligned}
$$

Second, stability of the error estimator (Lemma 3.3), $\mathcal{T}_{ij0} = \mathcal{T}_{i(j-1)\underline{k}}$ and the previous estimate prove that

$$
\begin{aligned}
\eta_{ij0} &\overset{(32)}{\le} \eta_{i(j-1)\underline{k}} + C_{\mathrm{stab}}\big(\|\boldsymbol{U}_{ij0} - \boldsymbol{U}_{i(j-1)\underline{k}}\|_{\mathbb{V}} + \|\Pi_i\nabla\cdot \boldsymbol{U}_{i(j-1)\underline{k}}\|_\Omega\big) \\
&\le (1+\kappa_1 C_{\mathrm{stab}})\,\eta_{i(j-1)\underline{k}} + C_{\mathrm{stab}}\,\|\nabla\cdot \boldsymbol{U}_{i(j-1)\underline{k}}\|_\Omega + \kappa_1 C_{\mathrm{stab}}\,\eta_{ij0}.
\end{aligned}
$$

Recall that $\kappa_1 C_{\mathrm{stab}} < \theta^{1/2} \le 1$. Thus, combining the last two estimates, we conclude the proof of (b).

To prove (c), note that Lemma 4.4 implies that

$$
\eta_{ijk} \overset{(45)}{\le} C_1\,\eta_{ijk'} \quad \text{for all } 0 \le k' < k \le \underline{k} := \underline{k}(i,j). \tag{62}
$$

Moreover, the Pythagoras theorem, reliability (28), and the equivalence (41) prove that

$$
\begin{aligned}
\|\nabla\cdot \boldsymbol{U}_{ijk}\|_\Omega &\le \|\nabla\cdot \boldsymbol{U}_{ijk'}\|_\Omega + \|\boldsymbol{U}^\star_{ijk} - \boldsymbol{U}^\star_{ijk'}\|_{\mathbb{V}} + \|\boldsymbol{U}^\star_{ijk} - \boldsymbol{U}_{ijk}\|_{\mathbb{V}} + \|\boldsymbol{U}^\star_{ijk'} - \boldsymbol{U}_{ijk'}\|_{\mathbb{V}} \\
&\le \|\nabla\cdot \boldsymbol{U}_{ijk'}\|_\Omega + \|\boldsymbol{u}_{ij} - \boldsymbol{U}^\star_{ijk'}\|_{\mathbb{V}} + \kappa_1\,\eta_{ijk} + \kappa_1\,\eta_{ijk'} \\
&\overset{(28)+(62)}{\lesssim} \|\nabla\cdot \boldsymbol{U}_{ijk'}\|_\Omega + \eta^\star_{ijk'} + \eta_{ijk'} \\
&\overset{(41)}{\lesssim} \|\nabla\cdot \boldsymbol{U}_{ijk'}\|_\Omega + \eta_{ijk'}.
\end{aligned}
$$

To prove (d), note that Lemma 4.5 implies that

$$
\eta_{ij\underline{k}} + \|\nabla\cdot \boldsymbol{U}_{ij\underline{k}}\|_\Omega \overset{(48)}{\simeq} \|p_i - P_{ij}\|_{\mathbb{P}} \overset{(47)}{\le} \|p_i - P_{ij'}\|_{\mathbb{P}} \overset{(48)}{\simeq} \eta_{ij'\underline{k}} + \|\nabla\cdot \boldsymbol{U}_{ij'\underline{k}}\|_\Omega.
$$

This concludes the proof. ∎

Proof of Theorem 4.1. For all $0 \le i' \le i \le \underline{i}$, define $\underline{\underline{j}}(i) \in \mathbb{N}_0$ by

$$
\underline{\underline{j}}(i) := \begin{cases} 0 & \text{if } i' < i, \\ j' & \text{if } i' = i. \end{cases}
$$

For all $0 \le i' \le i \le \underline{i}$ and all $\underline{\underline{j}}(i) \le j \le \underline{j}(i)$, define $\underline{\underline{k}}(i,j) \in \mathbb{N}_0$ by

$$
\underline{\underline{k}}(i,j) := \begin{cases} 0 & \text{if } i' < i \text{ or } j' < j, \\ k' & \text{if } i' = i \text{ and } j' = j. \end{cases}
$$

As for $\underline{j}$ and $\underline{k}$, we write $\underline{\underline{j}} = \underline{\underline{j}}(i)$ and $\underline{\underline{k}} = \underline{\underline{k}}(i,j)$ if i and j are clear from the context. Further, we abbreviate

$$
\mu_{ijk} := \eta_{ijk} + \|\nabla\cdot \boldsymbol{U}_{ijk}\|_\Omega.
$$

With this notation and according to Lemma 4.7, (39) is equivalent to

$$\sum_{\substack{(i,j,k)\in\mathcal{Q}\\(i',j',k')\leq(i,j,k)}} \mu_{ijk} = \sum_{i=i'}^{\underline{i}} \sum_{j=\underline{\underline{j}}(i)}^{\underline{j}(i)} \sum_{k=\underline{\underline{k}}(i,j)}^{\underline{k}(i,j)} \mu_{ijk} \lesssim \mu_{i'j'k'} \quad \text{for all } (i',j',k')\in\mathcal{Q}. \tag{63}$$

We prove (63) in the following three steps.

Step 1. For $\underline{\underline{k}}(i,j) < \underline{k}(i,j) < \infty$, Lemma 4.8 (c) proves that $\mu_{ij\underline{k}} \lesssim \mu_{ij\underline{\underline{k}}}$ Hence, Lemma 4.4 in combination with the geometric series allows to estimate the sum over k

$$\begin{aligned}\sum_{i=i'}^{\underline{i}} \sum_{j=\underline{\underline{j}}(i)}^{\underline{j}(i)} \sum_{k=\underline{\underline{k}}(i,j)}^{\underline{k}(i,j)} \mu_{ijk} &\overset{\text{(c)}}{\lesssim} \sum_{i=i'}^{\underline{i}} \sum_{j=\underline{\underline{j}}(i)}^{\underline{j}(i)} \sum_{k=\underline{\underline{k}}(i,j)}^{\underline{k}(i,j)-1} \mu_{ijk} \overset{(46)}{\simeq} \sum_{i=i'}^{\underline{i}} \sum_{j=\underline{\underline{j}}(i)}^{\underline{j}(i)} \sum_{k=\underline{\underline{k}}(i,j)}^{\underline{k}(i,j)-1} \eta_{ijk} \overset{(45)}{\lesssim} \sum_{i=i'}^{\underline{i}} \sum_{j=\underline{\underline{j}}(i)}^{\underline{j}(i)} \eta_{ij\underline{\underline{k}}}\\ &\leq \sum_{i=i'}^{\underline{i}} \sum_{j=\underline{\underline{j}}(i)}^{\underline{j}(i)} \mu_{ij\underline{\underline{k}}} = \sum_{j=\underline{\underline{j}}(i')}^{\underline{j}(i')} \mu_{i'j\underline{\underline{k}}} + \sum_{i=i'+1}^{\underline{i}} \sum_{j=\underline{\underline{j}}(i)}^{\underline{j}(i)} \mu_{ij\underline{\underline{k}}} = \sum_{j=j'}^{\underline{j}(i')} \mu_{i'j\underline{\underline{k}}} + \sum_{i=i'+1}^{\underline{i}} \sum_{j=0}^{\underline{j}(i)} \mu_{ij0}.\end{aligned} \tag{64}$$

Step 2. In this step, we bound the first summand of (64) by $\mu_{i'j'k'}$. It holds that

$$\sum_{j=j'}^{\underline{j}(i')} \mu_{i'j\underline{\underline{k}}} = \mu_{i'j'\underline{\underline{k}}} + \sum_{j=j'+1}^{\underline{j}(i')} \mu_{i'j\underline{\underline{k}}} = \mu_{i'j'k'} + \sum_{j=j'+1}^{\underline{j}(i')} \mu_{i'j0}.$$

Lemma 4.8 (b) and Lemma 4.5 in combination with the geometric series show that

$$\sum_{j=j'+1}^{\underline{j}(i')} \mu_{i'j0} \overset{\text{(b)}}{\lesssim} \sum_{j=j'+1}^{\underline{j}(i')} \mu_{i'(j-1)\underline{k}} = \sum_{j=j'}^{\underline{j}(i')-1} \mu_{i'j\underline{k}} \overset{(48)}{\simeq} \sum_{j=j'}^{\underline{j}(i')-1} \|p_{i'} - P_{i'j'}\|_{\mathbb{P}} \overset{(47)}{\lesssim} \|p_{i'} - P_{i'j}\|_{\mathbb{P}} \overset{(29)}{\lesssim} \mu_{i'j'k'}.$$

Step 3. In this step, we bound the second summand of (64) by $\mu_{i'j'k'}$. First, we consider only the terms where $j > 0$. As in Step 2, Lemma 4.8 (b) and Lemma 4.5 in combination with the geometric series show that

$$\sum_{i=i'+1}^{\underline{i}} \sum_{j=1}^{\underline{j}(i)} \mu_{ij0} \overset{\text{(b)}}{\lesssim} \sum_{i=i'+1}^{\underline{i}} \sum_{j=1}^{\underline{j}(i)} \mu_{i(j-1)\underline{k}} = \sum_{i=i'+1}^{\underline{i}} \sum_{j=0}^{\underline{j}(i)-1} \mu_{ij\underline{k}} \overset{\text{Lem.4.5}}{\lesssim} \sum_{i=i'+1}^{\underline{i}} \mu_{i0\underline{k}} \overset{\text{(c)}}{\lesssim} \sum_{i=i'+1}^{\underline{i}} \mu_{i00}.$$

Hence, it holds that

$$\sum_{i=i'+1}^{\underline{i}} \sum_{j=0}^{\underline{j}(i)} \mu_{ij0} = \sum_{i=i'+1}^{\underline{i}} \mu_{i00} + \sum_{i=i'+1}^{\underline{i}} \sum_{j=1}^{\underline{j}(i)} \mu_{ij0} \lesssim \sum_{i=i'+1}^{\underline{i}} \mu_{i00}.$$

Lemma 4.8 (a) and Lemma 4.6 in combination with the geometric series show that

$$\sum_{i=i'+1}^{\underline{i}} \mu_{i00} \overset{\text{(a)}}{\lesssim} \sum_{i=i'+1}^{\underline{i}} \mu_{(i-1)\underline{j}\underline{k}} = \sum_{i=i'}^{\underline{i}-1} \mu_{i\underline{j}\underline{k}} \overset{(57)}{\simeq} \sum_{i=i'}^{\underline{i}-1} \|p - P_{i\underline{j}}\|_{\mathbb{P}} \overset{(56)}{\lesssim} \|p - P_{i'\underline{j}}\|_{\mathbb{P}} \overset{(30)}{\lesssim} \mu_{i'\underline{j}\underline{k}}.$$

If $j' = \underline{j}(i')$, then Lemma 4.8 (c) yields that $\mu_{i'\underline{j}\underline{k}} = \mu_{i'j'\underline{k}} \lesssim \mu_{i'j'k'}$. Otherwise, if $j' < \underline{j}(i')$, then Lemma 4.8 (b)–(d) yield that

$$\mu_{i'\underline{j}\underline{k}} \overset{\text{(c)}}{\lesssim} \mu_{i'\underline{j}0} \overset{\text{(b)}}{\lesssim} \mu_{i'(\underline{j}-1)\underline{k}} \overset{\text{(d)}}{\lesssim} \mu_{i'j'\underline{k}} \overset{\text{(c)}}{\lesssim} \mu_{i'j'k'}.$$

Altogether, we have derived (63), which concludes the proof. ∎

5. Convergence rates

5.1. Main theorem on optimal convergence rates. The first lemma relates two different characterizations of approximation classes from the literature, which are either based on the accuracy $\varepsilon > 0$ (see, e.g., [Ste08, KS08]) or the number of elements N (see, e.g., [CKNS08, CFPP14]).

Lemma 5.1. *Recall that $\mathbb{T}^{\rm c} = \mathbb{T}^{\rm c}(\mathcal{T}_{\rm init})$. Let $\varrho : \mathbb{T}^{\rm c} \to \mathbb{R}_{\geq 0}$ satisfy that $\inf_{\mathcal{T}\in\mathbb{T}^{\rm c}} \varrho(\mathcal{T}) = 0$. Let $s > 0$ and define*

$$\mathbb{A}^{\rm c}_s(\varrho) := \sup_{N\in\mathbb{N}_0} \Big((N+1)^s \min_{\mathcal{T}\in\mathbb{T}^{\rm c}_N} \varrho(\mathcal{T})\Big), \quad \textit{where } \mathbb{T}^{\rm c}_N := \{\mathcal{T} \in \mathbb{T}^{\rm c} \,:\, \#\mathcal{T} - \#\mathcal{T}_{\rm init} \leq N\}. \tag{65}$$

With $\mathbb{T}^{\rm c}_\varepsilon(\varrho) := \{\mathcal{T} \in \mathbb{T}^{\rm c} \,:\, \varrho(\mathcal{T}) \leq \varepsilon\} \neq \emptyset$ for $\varepsilon > 0$, there holds the equality

$$\mathbb{A}^{\rm c}_s(\varrho) = \sup_{\varepsilon>0} \Big(\varepsilon \min_{\mathcal{T}\in\mathbb{T}^{\rm c}_\varepsilon(\varrho)} (\#\mathcal{T} - \#\mathcal{T}_{\rm init})^s\Big). \tag{66}$$

The minimum in (65) exists, since all $\mathbb{T}^{\rm c}_N$ are finite sets. The minimum in (66) exists, since the cardinality is a mapping $\# : \mathbb{T}^{\rm nc} \to \mathbb{N}$. In either case, the minimizers might not be unique. If $\mathbb{T}^{\rm c} = \mathbb{T}^{\rm c}(\mathcal{T}_{\rm init})$ is replaced by $\mathbb{T}^{\rm nc} = \mathbb{T}^{\rm nc}(\mathcal{T}_{\rm init})$, one can define $\mathbb{A}^{\rm nc}_s$, $\mathbb{T}^{\rm nc}_N$, and $\mathbb{T}^{\rm nc}_\varepsilon(\varrho)$ similarly, and the assertion (66) holds accordingly.

Proof. We only consider the set $\mathbb{T}^{\rm c}$ of conforming triangulations, the proof for the set $\mathbb{T}^{\rm nc}$ of non-conforming triangulations follows along the same lines. For $N \in \mathbb{N}_0$, define $\varepsilon_N := \min_{\mathcal{T}\in\mathbb{T}^{\rm c}_N} \varrho(\mathcal{T}) \geq 0$.

Step 1. To prove "$\geq$" in (66), let $\varepsilon > 0$. If $0 < \varepsilon < \varepsilon_0$, there exists a minimal $N \in \mathbb{N}_0$ such that $\min_{\mathcal{T}\in\mathbb{T}^{\rm c}_N} \varrho(\mathcal{T}) \leq \varepsilon$. In particular, it follows that $N > 0$, $\mathbb{T}^{\rm c}_N \cap \mathbb{T}^{\rm c}_\varepsilon(\varrho) \neq \emptyset$, and $\varepsilon < \min_{\mathcal{T}\in\mathbb{T}^{\rm c}_{N-1}} \varrho(\mathcal{T})$. This yields that

$$\varepsilon \min_{\mathcal{T}\in\mathbb{T}^{\rm c}_\varepsilon(\varrho)} (\#\mathcal{T} - \#\mathcal{T}_{\rm init})^s \leq \min_{\mathcal{T}\in\mathbb{T}^{\rm c}_{N-1}} \varrho(\mathcal{T})\, N^s \leq \sup_{N\in\mathbb{N}_0} \Big((N+1)^s \min_{\mathcal{T}\in\mathbb{T}^{\rm c}_N} \varrho(\mathcal{T})\Big) = \mathbb{A}^{\rm c}_s(\varrho). \tag{67}$$

If $\varepsilon_0 \leq \varepsilon$, then $\mathcal{T}_{\rm init} \in \mathbb{T}^{\rm c}_{\varepsilon_0}(\varrho) \subseteq \mathbb{T}^{\rm c}_\varepsilon(\varrho)$ and hence the left-hand side of (67) is zero, and (67) thus remains true. Taking the supremum over all $\varepsilon > 0$, we prove "$\geq$" in (66).

Step 2. To prove "$\leq$" in (66), let $N \in \mathbb{N}_0$. If $\varepsilon_N > 0$, the definition of ε_N yields that $\#\mathcal{T} - \#\mathcal{T}_{\rm init} \geq N + 1$ for all $\mathcal{T} \in \mathbb{T}^{\rm c}_{\lambda\varepsilon_N}(\varrho)$ and all $0 < \lambda < 1$. This proves that

$$(N+1)^s \min_{\mathcal{T}\in\mathbb{T}^{\rm c}_N} \varrho(\mathcal{T}) \leq \min_{\mathcal{T}\in\mathbb{T}^{\rm c}_{\lambda\varepsilon_N}(\varrho)} (\#\mathcal{T} - \#\mathcal{T}_{\rm init})^s\, \varepsilon_N \leq \frac{1}{\lambda} \sup_{\varepsilon>0} \big(\varepsilon \min_{\mathcal{T}\in\mathbb{T}^{\rm c}_\varepsilon(\varrho)} (\#\mathcal{T} - \#\mathcal{T}_{\rm init})^s\big). \tag{68}$$

If $\varepsilon_N = 0$, then the left-hand side of (68) is zero, and the overall estimate thus remains true. Taking the supremum over all $N \in \mathbb{N}_0$, we prove "$\leq$" in (66) for the limit $\lambda \to 1$. ∎

The following lemma specifies $\varrho(\mathcal{T})$ and hence introduces the precise approximation class of the present work.

Lemma 5.2. *For $s > 0$, let*

$$\mathbb{A}^{\rm c}_s := \mathbb{A}^{\rm c}_s(\varrho), \quad \textit{where} \quad \varrho(\mathcal{T}) := \eta(\mathcal{T}; \boldsymbol{U}_\mathcal{T}[p_\mathcal{T}], p_\mathcal{T}) + \|\nabla\cdot\boldsymbol{U}_\mathcal{T}[p_\mathcal{T}]\|_\Omega \quad \textit{for } \mathcal{T} \in \mathbb{T}^{\rm c}. \tag{69}$$

Then, ϱ satisfies the assumptions of Lemma 5.1. Moreover, there exists a constant $C > 0$, which depends only on $C_{\rm stab}$ and $C_{\rm rel}$, such that

$$\varrho(\mathcal{T}) \le C \min_{Q_\mathcal{T} \in \mathbb{P}(\mathcal{T})} \big(\eta(\mathcal{T}; \boldsymbol{U}_\mathcal{T}[Q_\mathcal{T}], Q_\mathcal{T}) + \|\nabla \cdot \boldsymbol{U}_\mathcal{T}[Q_\mathcal{T}]\|_\Omega\big). \tag{70}$$

Proof. Let $Q_\mathcal{T} \in \mathbb{P}(\mathcal{T})$. According to (21), we have that $\|\boldsymbol{U}_\mathcal{T}[p_\mathcal{T}] - \boldsymbol{U}_\mathcal{T}[Q_\mathcal{T}]\|_\mathbb{V} \le \|p_\mathcal{T} - Q_\mathcal{T}\|_\mathbb{P}$. Since $p_\mathcal{T}$ is the best approximation of p in $\mathbb{P}(\mathcal{T})$, it holds that $\|p_\mathcal{T} - Q_\mathcal{T}\|_\mathbb{P} \le \|p - Q_\mathcal{T}\|_\mathbb{P}$. Hence, stability (32) and reliability (30) of the error estimator prove that

$$\begin{aligned}
\varrho(\mathcal{T}) &= \eta(\mathcal{T}; \boldsymbol{U}_\mathcal{T}[p_\mathcal{T}], p_\mathcal{T}) + \|\nabla \cdot \boldsymbol{U}_\mathcal{T}[p_\mathcal{T}]\|_\Omega \\
&\overset{(32)}{\lesssim} \eta(\mathcal{T}; \boldsymbol{U}_\mathcal{T}[Q_\mathcal{T}], Q_\mathcal{T}) + \|\boldsymbol{U}_\mathcal{T}[p_\mathcal{T}] - \boldsymbol{U}_\mathcal{T}[Q_\mathcal{T}]\|_\mathbb{V} + \|p_\mathcal{T} - Q_\mathcal{T}\|_\mathbb{P} + \|\nabla \cdot \boldsymbol{U}_\mathcal{T}[Q_\mathcal{T}]\|_\Omega \\
&\lesssim \eta(\mathcal{T}; \boldsymbol{U}_\mathcal{T}[Q_\mathcal{T}], Q_\mathcal{T}) + \|\nabla \cdot \boldsymbol{U}_\mathcal{T}[Q_\mathcal{T}]\|_\Omega + \|p - Q_\mathcal{T}\|_\mathbb{P}. \\
&\overset{(30)}{\lesssim} \eta(\mathcal{T}; \boldsymbol{U}_\mathcal{T}[Q_\mathcal{T}], Q_\mathcal{T}) + \|\nabla \cdot \boldsymbol{U}_\mathcal{T}[Q_\mathcal{T}]\|_\Omega.
\end{aligned}$$

This proves (70). With linear convergence (Theorem 4.1), this yields that

$$\inf_{\mathcal{T} \in \mathbb{T}^{\rm c}} \varrho(\mathcal{T}) \le \inf_{(i,j,k) \in \mathcal{Q}} \varrho(\mathcal{T}_{ijk}) \lesssim \inf_{(i,j,k) \in \mathcal{Q}} \big(\eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega\big) = 0.$$

This concludes the proof. ■

Together with Theorem 4.1, the following theorem is the main result of this work. It states optimal convergence of Algorithm 3.7. The proof is given in Section 5.2.

Theorem 5.3. *Let $0 < \vartheta < C_{\rm div}^{-1}$ and $0 < \theta < \theta_{\rm opt} := (1 + C_{\rm stab}^2 C_{\rm drel}^2)^{-1}$. Suppose that*

$$\kappa_1 < \theta^{1/2} C_{\rm stab} \quad \text{and} \quad \theta < \sup_{\delta > 0} \frac{(1 - \kappa_1 C_{\rm stab})^2 \theta_{\rm opt} - (1 + \delta^{-1}) \kappa_1^2 C_{\rm stab}^2}{1 + \delta}, \tag{71}$$

i.e., $0 \le \kappa_1 < 1$ is sufficiently small. Moreover, let $0 < \kappa_2, \kappa_3 < 1$ be sufficiently small in the sense of Lemma 4.5, Lemma 4.6, and Lemma 5.6 below. Then, for all $s > 0$, there exist constants $c_{\rm opt}, C_{\rm opt} > 0$ such that

$$c_{\rm opt}\, \mathbb{A}_s^{\rm c} \le \sup_{(i,j,k) \in \mathcal{Q}} \big(\eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega\big) \big(\#\mathcal{T}_{ijk} - \#\mathcal{T}_{\rm init} + 1\big)^s \le C_{\rm opt}\, (1 + \mathbb{A}_s^{\rm c}). \tag{72}$$

The constant $c_{\rm opt}$ depends only on the initial triangulation $\mathcal{T}_{\rm init}$ and the polynomial degree m, while $C_{\rm opt}$ depends additionally on the domain Ω, the parameters $\kappa_1, \kappa_2, \kappa_3, \vartheta, d, \theta$, $C_{\rm mark}$, s, and $\boldsymbol{f}$.

The following remark relates our definition of the approximation class from Lemma 5.2 to that of [KS08]. We refer to Appendix C for the proof.

Remark 5.4. (i) The seminal work [KS08] employs two approximation classes:

- $\mathbb{A}_s^{\rm nc}(p) := \mathbb{A}_s^{\rm nc}(\varrho_p)$ for $\varrho_p(\mathcal{P}) := \min_{Q_\mathcal{P} \in \mathbb{P}(\mathcal{P})} \|p - Q_\mathcal{P}\|_\mathbb{P} = \|p - p_\mathcal{P}\|_\mathbb{P}$,
- $\mathbb{A}_s^{\rm c}(\boldsymbol{u}) := \mathbb{A}_s^{\rm c}(\varrho_u)$ for $\varrho_{\boldsymbol{u}}(\mathcal{T}) := \min_{\boldsymbol{V}_\mathcal{T} \in \mathbb{V}(\mathcal{T})} \|\boldsymbol{u} - \boldsymbol{V}_\mathcal{T}\|_\mathbb{V}$.

Clearly, the definitions of ϱ_p and $\varrho_{\boldsymbol{u}}$ satisfy the assumptions of Lemma 5.1. Moreover,

$$\mathbb{A}_s^{\rm nc}(p) \simeq \mathbb{A}_s^{\rm c}(p) := \mathbb{A}_s^{\rm c}(\varrho_p). \tag{73}$$

(ii) If we additionally define

- $\mathbb{A}_s^{\rm c}(\boldsymbol{u},p) := \mathbb{A}_s^{\rm c}(\varrho_{u,p})$ for $\varrho_{u,p}(\mathcal{T}) := \min\limits_{Q_\mathcal{T}\in\mathbb{P}(\mathcal{T})} \|p - Q_\mathcal{T}\|_\mathbb{P} + \min\limits_{\boldsymbol{V}_\mathcal{T}\in\mathbb{V}(\mathcal{T})} \|\boldsymbol{u} - \boldsymbol{V}_\mathcal{T}\|_\mathbb{V}$,

then it holds that

$$\frac{1}{2}\left(\mathbb{A}_s^{\rm c}(p) + \mathbb{A}_s^{\rm c}(\boldsymbol{u})\right) \le \mathbb{A}_s^{\rm c}(\boldsymbol{u},p) \le 2^s\left(\mathbb{A}_s^{\rm c}(p) + \mathbb{A}_s^{\rm c}(\boldsymbol{u})\right) \quad \text{for all } s>0. \tag{74}$$

(iii) Finally, the reliability (30) implies that $\mathbb{A}_s^{\rm c}(\boldsymbol{u},p) \le C_{\rm rel}^s\,\mathbb{A}_s^{\rm c}$. Conversely, if the volume force $\boldsymbol{f}$ is a $\mathcal{T}_{\rm init}$-piecewise polynomial, it holds that

$$\mathbb{A}_s^{\rm c} \lesssim \mathbb{A}_s^{\rm c}(\boldsymbol{u},p) \le C_{\rm rel}\,\mathbb{A}_s^{\rm c}, \tag{75}$$

i.e., if the volume force $\boldsymbol{f}$ is a $\mathcal{T}_{\rm init}$-piecewise polynomial, then our approximation class coincides with that of [KS08]. ∎

5.2. Proof of Theorem 5.3. We start with an auxiliary lemma, which was originally proved in [KS08, Lemma 6.3].

Lemma 5.5. *Let $0<\vartheta<\vartheta'<C_{\rm div}^{-1}$. Let $0<\omega<1$ be sufficiently small such that*

$$0 < q := C_{\rm div}\,\frac{\omega+\vartheta'}{1-\omega} < 1, \tag{76}$$

Let $\mathcal{P}\in\mathbb{T}^{\rm nc}$ and $\mathcal{T}\in\mathbb{T}^{\rm c}(\mathcal{P})$. Let $Q_\mathcal{P}\in\mathbb{P}(\mathcal{P})$. Let $\boldsymbol{V}_\mathcal{T}\in\mathbb{V}(\mathcal{T})$ satisfy that

$$\|\nabla\cdot(\boldsymbol{u}[Q_\mathcal{P}] - \boldsymbol{V}_\mathcal{T})\|_\Omega \le \omega\,\|\nabla\cdot\boldsymbol{V}_\mathcal{T}\|_\Omega. \tag{77}$$

Then, `bisect`$(\mathcal{P},\mathcal{T},\boldsymbol{V}_\mathcal{T};\vartheta)$ *from Algorithm 3.6 returns $\mathcal{P}'\in\mathbb{T}^{\rm nc}(\mathcal{P})$ such that the following implication is satisfied for all $\overline{\mathcal{P}}\in\mathbb{T}^{\rm nc}(\mathcal{P})$*

$$\|p - p_{\overline{\mathcal{P}}}\|_\mathbb{P}^2 \le (1-q^2)\,\|p - Q_\mathcal{P}\|_\mathbb{P}^2 \quad\Longrightarrow\quad \#\mathcal{P}' - \#\mathcal{P} \le C_{\rm bin}\,(\#\overline{\mathcal{P}} - \#\mathcal{T}_{\rm init}). \tag{78}$$

Proof. To see (78), let $\overline{\mathcal{P}}\in\mathbb{T}^{\rm nc}(\mathcal{P})$ with $\|p-p_{\overline{\mathcal{P}}}\|_\mathbb{P}^2 \le (1-q^2)\,\|p-Q_\mathcal{P}\|_\mathbb{P}^2$. Note that

$$\|p - p_{\widetilde{\mathcal{P}}}\|_\mathbb{P}^2 \le \|p - p_{\overline{\mathcal{P}}}\|_\mathbb{P}^2 \le (1-q^2)\,\|p - Q_\mathcal{P}\|_\mathbb{P}^2, \quad \text{where} \quad \widetilde{\mathcal{P}} := \mathcal{P}\oplus\overline{\mathcal{P}} \in \mathbb{T}^{\rm nc}(\mathcal{P}). \tag{79}$$

The triangle inequality and assumption (77) show that

$$\|\nabla\cdot\boldsymbol{V}_\mathcal{T}\|_\Omega \le \|\nabla\cdot u[Q_\mathcal{P}]\|_\Omega + \|\nabla\cdot(u[Q_\mathcal{P}] - V_\mathcal{T})\|_\Omega \overset{(77)}{\le} \|\nabla\cdot\boldsymbol{u}[Q_\mathcal{P}]\|_\Omega + \omega\,\|\nabla\cdot\boldsymbol{V}_\mathcal{T}\|_\Omega.$$

Hence, Lemma 3.5 yields that

$$\begin{aligned} &q^2(1-\omega)^2\,\|\nabla\cdot\boldsymbol{V}_\mathcal{T}\|_\Omega^2 \le q^2\,\|\nabla\cdot\boldsymbol{u}[Q_\mathcal{P}]\|_\Omega^2 \\ &\overset{(35)}{\le} q^2\,\|p-Q_\mathcal{P}\|_\mathbb{P}^2 \overset{(79)}{\le} \|p-Q_\mathcal{P}\|_\mathbb{P}^2 - \|p-p_{\widetilde{\mathcal{P}}}\|_\mathbb{P}^2 = \|p_{\widetilde{\mathcal{P}}} - Q_\mathcal{P}\|_\mathbb{P}^2 \overset{(34)}{\le} C_{\rm div}^2\|\Pi_{\widetilde{\mathcal{P}}}\nabla\cdot\boldsymbol{u}[Q_\mathcal{P}]\|_\Omega^2. \end{aligned}$$

The triangle inequality together with (77) shows that

$$\|\Pi_{\widetilde{\mathcal{P}}}\nabla\cdot\boldsymbol{u}[Q_\mathcal{P}]\|_\Omega \le \|\Pi_{\widetilde{\mathcal{P}}}\nabla\cdot\boldsymbol{V}_\mathcal{T}\|_\Omega + \|\Pi_{\widetilde{\mathcal{P}}}\nabla\cdot(\boldsymbol{u}[Q_\mathcal{P}] - \boldsymbol{V}_\mathcal{T})\|_\Omega \overset{(77)}{\le} \|\Pi_{\widetilde{\mathcal{P}}}\nabla\cdot\boldsymbol{V}_\mathcal{T}\|_\Omega + \omega\,\|\nabla\cdot\boldsymbol{V}_\mathcal{T}\|_\Omega.$$

Altogether, we derive that

$$q(1-\omega)\,\|\nabla\cdot\boldsymbol{V}_\mathcal{T}\|_\Omega \le C_{\rm div}\|\Pi_{\widetilde{\mathcal{P}}}\nabla\cdot\boldsymbol{u}[Q_\mathcal{P}]\|_\Omega \le C_{\rm div}\big(\|\Pi_{\widetilde{\mathcal{P}}}\nabla\cdot\boldsymbol{V}_\mathcal{T}\|_\Omega + \omega\,\|\nabla\cdot\boldsymbol{V}_\mathcal{T}\|_\Omega\big).$$

By choice of q in (76), this is equivalent to

$$\vartheta' \, \|\nabla \cdot \boldsymbol{V}_{\mathcal{T}}\|_\Omega = \frac{q(1-\omega) - C_{\mathrm{div}}\,\omega}{C_{\mathrm{div}}} \, \|\nabla \cdot \boldsymbol{V}_{\mathcal{T}}\|_\Omega \le \|\Pi_{\widetilde{\mathcal{P}}} \nabla \cdot \boldsymbol{V}_{\mathcal{T}}\|_\Omega.$$

By definition, Algorithm 3.6 returns $\mathcal{P}' \in \mathbb{T}^{\mathrm{nc}}(\mathcal{P})$ such that

$$\#\mathcal{P}' - \#\mathcal{P} \le C_{\mathrm{bin}}\,(\#\widetilde{\mathcal{P}} - \#\mathcal{P}) \overset{\text{(M1)}}{\le} C_{\mathrm{bin}}\,(\#\overline{\mathcal{P}} - \#\mathcal{T}_{\mathrm{init}}).$$

This concludes the proof. ■

The heart of the proof of Theorem 4.1 is the following auxiliary lemma.

Lemma 5.6. *Let $(i,j,k) \in \mathcal{Q}$ with $k < \underline{k}(i,j)$ and $s > 0$. Let $0 < \vartheta < C_{\mathrm{div}}^{-1}$ and $0 < \theta < \theta_{\mathrm{opt}} = (1 + C_{\mathrm{stab}}^2 C_{\mathrm{drel}}^2)^{-1}$. Let $0 \le \kappa_1 < 1$ be sufficiently small such that (71) is satisfied. For sufficiently small $0 < \kappa_2 \ll 1$ (see (88) in the proof below), there exists C_{comp} such that*

$$\#\mathcal{M}_{ijk} \le C_{\mathrm{comp}}(1 + (\mathbb{A}_s^{\mathrm{c}})^{1/s})\big(\eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega\big)^{-1/s}. \tag{80}$$

The constant $C_{\mathrm{comp}} > 0$ depends only on the domain Ω, γ-shape regularity, the polynomial degree m, the parameters $\kappa_1, \kappa_2, \kappa_3, \vartheta, \theta$ C_{mark}, and s.

Proof. The proof is split into five steps.

Step 1. Choose

$$\varepsilon := \eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega. \tag{81}$$

Without loss of generality, we may assume that $\varepsilon > 0$. Since $\mathbb{A}_s^{\mathrm{c}} < \infty$, Lemma 5.1 and Lemma 5.2 guarantee the existence of $\overline{\mathcal{T}} \in \mathbb{T}^{\mathrm{c}}$ such that

$$\#\overline{\mathcal{T}} - \#\mathcal{T}_{\mathrm{init}} \le (\mathbb{A}_s^{\mathrm{c}}/\varepsilon)^{1/s} \quad \text{and} \quad \eta(\overline{\mathcal{T}}; \boldsymbol{U}_{\overline{\mathcal{T}}}[p_{\overline{\mathcal{T}}}], p_{\overline{\mathcal{T}}}) + \|\nabla \cdot \boldsymbol{U}_{\overline{\mathcal{T}}}[p_{\overline{\mathcal{T}}}]\|_\Omega \le \varepsilon. \tag{82}$$

Step 2. Define the uniformly refined triangulations

$$\widehat{\mathcal{T}}_0 := \mathtt{close}(\mathcal{P}_i) \oplus \overline{\mathcal{T}} \quad \text{and} \quad \widehat{\mathcal{T}}_{n+1} := \mathtt{refine}(\widehat{\mathcal{T}}_n, \widehat{\mathcal{T}}_n) \quad \text{for all } n \in \mathbb{N}_0.$$

Note that $P_{ij} \in \mathbb{P}(\mathcal{P}_i) \subseteq \mathbb{P}(\widehat{\mathcal{T}}_n)$. We recall some standard arguments for adaptive mesh-refinement for the (vector-valued) Poisson model problem. Reliability (28), stability (32), and reduction (33) guarantee the existence of $C_{\mathrm{ctr}} > 0$ and $0 < q_{\mathrm{ctr}} < 1$ such that

$$\eta(\widehat{\mathcal{T}}_n; \boldsymbol{U}_{\widehat{\mathcal{T}}_n}[P_{ij}], P_{ij}) \le C_{\mathrm{ctr}}\, q_{\mathrm{ctr}}^n\, \eta(\widehat{\mathcal{T}}_0; \boldsymbol{U}_{\widehat{\mathcal{T}}_0}[P_{ij}], P_{ij});$$

see, e.g., [CFPP14, Theorem 4.1 (i)]. According to, e.g., [CFPP14, Section 3.4], there exists $C_{\mathrm{mon}} > 0$ such that for all $\widehat{\mathcal{T}} \in \mathbb{T}^{\mathrm{c}}$, $\widehat{\mathcal{T}}' \in \mathbb{T}^{\mathrm{c}}(\widehat{\mathcal{T}})$, $P_{\widehat{\mathcal{T}}} \in \mathbb{P}(\widehat{\mathcal{T}})$

$$\eta(\widehat{\mathcal{T}}'; \boldsymbol{U}_{\widehat{\mathcal{T}}'}[P_{\widehat{\mathcal{T}}}], P_{\widehat{\mathcal{T}}}) \le C_{\mathrm{mon}} \eta(\widehat{\mathcal{T}}; \boldsymbol{U}_{\widehat{\mathcal{T}}}[P_{\widehat{\mathcal{T}}}], P_{\widehat{\mathcal{T}}}) \tag{83}$$

Note that C_{ctr}, q_{ctr}, and C_{mon} depend only on γ-shape regularity and the polynomial degree m. With stability (32) and quasi-monotonicity (83), it follows that

$$\begin{aligned}
&\eta(\widehat{\mathcal{T}}_n; \boldsymbol{U}_{\widehat{\mathcal{T}}_n}[P_{ij}], P_{ij}) \le C_{\mathrm{ctr}}\, q_{\mathrm{ctr}}^n\, \eta(\widehat{\mathcal{T}}_0; \boldsymbol{U}_{\widehat{\mathcal{T}}_0}[P_{ij}], P_{ij}) \\
&\overset{(32)}{\le} C_{\mathrm{ctr}}\, q_{\mathrm{ctr}}^n \left[\eta(\widehat{\mathcal{T}}_0; \boldsymbol{U}_{\widehat{\mathcal{T}}_0}[p_{\overline{\mathcal{T}}}], p_{\overline{\mathcal{T}}}) + C_{\mathrm{stab}}\big(\|\boldsymbol{U}_{\widehat{\mathcal{T}}_0}[P_{ij}] - \boldsymbol{U}_{\widehat{\mathcal{T}}_0}[p_{\overline{\mathcal{T}}}]\|_{\mathbb{V}} + \|P_{ij} - p_{\overline{\mathcal{T}}}\|_{\mathbb{P}}\big)\right] \\
&\overset{(83)}{\le} C_{\mathrm{ctr}}\, q_{\mathrm{ctr}}^n \left[C_{\mathrm{mon}}\, \eta(\overline{\mathcal{T}}; \boldsymbol{U}_{\overline{\mathcal{T}}}[p_{\overline{\mathcal{T}}}], p_{\overline{\mathcal{T}}}) + C_{\mathrm{stab}}\big(\|\boldsymbol{U}_{\widehat{\mathcal{T}}_0}[P_{ij}] - \boldsymbol{U}_{\widehat{\mathcal{T}}_0}[p_{\overline{\mathcal{T}}}]\|_{\mathbb{V}} + \|P_{ij} - p_{\overline{\mathcal{T}}}\|_{\mathbb{P}}\big)\right].
\end{aligned}$$

With (21), we hence obtain that

$$\eta(\widehat{\mathcal{T}}_n; \boldsymbol{U}_{\widehat{\mathcal{T}}_n}[P_{ij}], P_{ij}) \le C_{\rm ctr}\, q_{\rm ctr}^n \big[C_{\rm mon}\, \eta(\overline{\mathcal{T}}; \boldsymbol{U}_{\overline{\mathcal{T}}}[p_{\overline{\mathcal{T}}}], p_{\overline{\mathcal{T}}}) + 2C_{\rm stab}\, \|P_{ij} - p_{\overline{\mathcal{T}}}\|_{\mathbb{P}}\big].$$

According to the reliability estimates (30) and (44), it holds that

$$\begin{aligned}\|P_{ij} - p_{\overline{\mathcal{T}}}\|_{\mathbb{P}} &\le \|p - p_{\overline{\mathcal{T}}}\|_{\mathbb{P}} + \|p - P_{ij}\|_{\mathbb{P}}\\ &\le C'_{\rm rel}(\kappa_1)\, \big\{\big(\eta(\overline{\mathcal{T}}; \boldsymbol{U}_{\overline{\mathcal{T}}}[p_{\overline{\mathcal{T}}}], p_{\overline{\mathcal{T}}}) + \|\nabla\cdot \boldsymbol{U}_{\overline{\mathcal{T}}}[p_{\overline{\mathcal{T}}}]\|_\Omega\big) + \big(\eta_{ijk} + \|\nabla\cdot\boldsymbol{U}_{ijk}\|_\Omega\big)\big\}.\end{aligned}$$

By choice of $\overline{\mathcal{T}}$ in Step 1 and for $k < \underline{k}(i,j)$, we overall obtain that

$$\begin{aligned}\eta(\widehat{\mathcal{T}}_n; \boldsymbol{U}_{\widehat{\mathcal{T}}_n}[P_{ij}], P_{ij}) &\le q_{\rm ctr}^n\, C_{\rm ctr}\, \big[C_{\rm mon} + 4\, C_{\rm stab}\, C'_{\rm rel}(\kappa_1)\big]\, \big(\eta_{ijk} + \|\nabla\cdot\boldsymbol{U}_{ijk}\|_\Omega\big)\\ &\overset{(46)}{\le} q_{\rm ctr}^n\, C_{\rm ctr}\, \big[C_{\rm mon} + 4\, C_{\rm stab}\, C'_{\rm rel}(\kappa_1)\big]\, \frac{1}{\kappa_2}\Big(1 + \frac{1}{\kappa_3}\Big)\,\eta_{ijk}.\end{aligned} \tag{84}$$

Step 3. To shorten notation, we set $\eta^\star_{ijk} := \eta(\mathcal{T}_{ijk}; \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}], P_{ij})$ and $\boldsymbol{U}^\star_{ijk} := \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}]$. Note that discrete reliability (31) and stability (32) imply optimality of Dörfler marking (see, e.g., [CFPP14, Section 4.5]): For any $0 < \theta_\star < \theta_{\rm opt}$, there exists some $0 < \lambda = \lambda(\theta_\star) \ll 1$ such that, for all $\check{\mathcal{T}} \in \mathbb{T}^{\rm c}(\mathcal{T}_{ijk})$, it holds that

$$\eta(\check{\mathcal{T}}; \boldsymbol{U}_{\check{\mathcal{T}}}[P_{ij}], P_{ij}) \le \lambda\, \eta^\star_{ijk} \quad\Longrightarrow\quad \theta_\star\, (\eta^\star_{ijk})^2 \le \eta(\mathcal{T}_{ijk} \setminus \check{\mathcal{T}}; \boldsymbol{U}_{ijk}, P_{ij})^2. \tag{85}$$

The second inequality in (85), Lemma 4.3, and the Young inequality imply for $\delta > 0$ that

$$\begin{aligned}(1 - \kappa_1 C_{\rm stab})^2 \theta_\star\, \eta^2_{ijk} &\overset{(41)}{\le} \theta_\star (\eta^\star_{ijk})^2 \overset{(85)}{\le} \eta(\mathcal{T}_{ijk} \setminus \check{\mathcal{T}}; \boldsymbol{U}^\star_{ijk}, P_{ij})^2\\ &\overset{(40)}{\le} (1+\delta)\eta(\mathcal{T}_{ijk} \setminus \check{\mathcal{T}}; \boldsymbol{U}_{ijk}, P_{ij})^2 + (1+\delta^{-1})\kappa_1^2 C_{\rm stab}^2 \eta^2_{ijk}.\end{aligned}$$

Due to (71), we can choose $0 < \theta_\star < \theta_{\rm opt}$ sufficiently close to $\theta_{\rm opt}$ such that

$$\theta\, \eta^2_{ijk} \overset{(71)}{\le} \sup_{\delta > 0} \frac{(1-\kappa_1 C_{\rm stab})^2 \theta_\star - (1+\delta^{-1})\kappa_1^2 C^2_{\rm stab}}{1+\delta}\, \eta^2_{ijk} \le \eta(\mathcal{T}_{ijk} \setminus \check{\mathcal{T}}; \boldsymbol{U}_{ijk}, P_{ij})^2. \tag{86}$$

Let $\ell \in \mathbb{N}_0$ be the minimal integer such that

$$q^\ell_{\rm ctr}\, \frac{C_{\rm mon}}{1 - \kappa_1 C_{\rm stab}} C_{\rm ctr}\, \big[C_{\rm mon} + 4 C_{\rm stab} C'_{\rm rel}(\kappa_1)\big]\, \frac{1}{\kappa_2}\Big(1 + \frac{1}{\kappa_3}\Big) \le \lambda.$$

Recall $\widehat{\mathcal{T}}_\ell$ from Step 2. For $\check{\mathcal{T}} := \widehat{\mathcal{T}}_\ell \oplus \mathcal{T}_{ijk}$, it then holds that

$$\eta(\check{\mathcal{T}}; \boldsymbol{U}_{\check{\mathcal{T}}}[P_{ij}], P_{ij}) \overset{(83)}{\le} C_{\rm mon}\, \eta(\widehat{\mathcal{T}}_\ell; \boldsymbol{U}_{\widehat{\mathcal{T}}_\ell}[P_{ij}], P_{ij}) \overset{(84)}{\le} \lambda\, (1 - \kappa_1 C_{\rm stab})\, \eta_{ijk} \overset{(41)}{\le} \lambda \eta^\star_{ijk}.$$

Hence, (85)–(86) imply that $\theta\, \eta^2_{ijk} \le \eta(\mathcal{T}_{ijk} \setminus \check{\mathcal{T}}; \boldsymbol{U}_{ijk}, P_{ij})^2$.

Step 4. Since $\mathcal{M}_{ijk} \subseteq \mathcal{T}_{ijk}$ in Algorithm 3.7 (iv) has (up to some fixed factor $C_{\rm mark}$) minimal cardinality, the overlay estimate (M1)implies that

$$\begin{aligned}C^{-1}_{\rm mark} \#\mathcal{M}_{ijk} &\overset{(85)}{\le} \#(\mathcal{T}_{ijk} \setminus \check{\mathcal{T}}) \le \#\check{\mathcal{T}} - \#\mathcal{T}_{ijk} \overset{(M1)}{\le} \#\widehat{\mathcal{T}}_\ell - \#\mathcal{T}_{\rm init} \overset{(M2)}{\le} C^\ell_{\rm son} \#\widehat{\mathcal{T}}_0\\ &\overset{(M1)}{\le} C^\ell_{\rm son}\big(\#\texttt{close}(\mathcal{P}_i) + \#\overline{\mathcal{T}} - \#\mathcal{T}_{\rm init}\big) \overset{(82)}{\lesssim} (\mathbb{A}^{\rm c}_s)^{1/s}\, \big(\eta_{ijk} + \|\nabla\cdot\boldsymbol{U}_{ijk}\|_\Omega\big)^{-1/s} + \#\texttt{close}(\mathcal{P}_i).\end{aligned}$$

Elementary calculation (see, e.g., [BHP17, Lemma 22]) shows that

$$\#\mathcal{P} - \#\mathcal{T}_{\rm init} + 1 \le \#\mathcal{P} \le \#\mathcal{T}_{\rm init}\, \big(\#\mathcal{P} - \#\mathcal{T}_{\rm init} + 1\big) \quad \text{for all } \mathcal{P} \in \mathbb{T}^{\rm nc}.$$

With $\#\mathcal{T}_{\rm init} \simeq 1 \lesssim \big(\eta_{ijk} + \|\nabla\cdot \boldsymbol{U}_{ijk}\|_\Omega\big)^{-1/s}$, the conformity estimate (M4) yields that

$$\#\texttt{close}(\mathcal{P}_i) \lesssim \big(\eta_{ijk} + \|\nabla\cdot \boldsymbol{U}_{ijk}\|_\Omega\big)^{-1/s} + (\#\mathcal{P}_i - \#\mathcal{T}_{\rm init}).$$

Altogether, this step thus concludes that

$$\#\mathcal{M}_{ijk} \lesssim (1 + (\mathbb{A}_s^{\rm c})^{1/s})\big(\eta_{ijk} + \|\nabla\cdot \boldsymbol{U}_{ijk}\|_\Omega\big)^{-1/s} + (\#\mathcal{P}_i - \#\mathcal{T}_{\rm init}). \tag{87}$$

Step 5. Reliability (42) as well as Algorithm 3.7 (ii) show for all $0 \le i' < i$ that

$$\|\nabla\cdot(\boldsymbol{u}_{i'\underline{j}} - \boldsymbol{U}_{i'\underline{jk}})\|_\Omega \le \|\boldsymbol{u}_{i'\underline{j}} - \boldsymbol{U}_{i'\underline{jk}}\|_{\mathbb{V}} \le C'_{\rm rel}(\kappa_1)\,\eta_{i'\underline{jk}} \le C'_{\rm rel}(\kappa_1)\,\frac{\kappa_2}{1-\kappa_2}\,\|\nabla\cdot \boldsymbol{U}_{i'\underline{jk}}\|_\Omega.$$

Let $0 < \vartheta < \vartheta' < C_{\rm div}^{-1}$ and $\omega := C'_{\rm rel}(\kappa_1)\kappa_2/(1-\kappa_2)$. For $0 < \kappa_2 \ll 1$ with

$$0 < q := C_{\rm div}\,\frac{\omega + \vartheta'}{1-\omega} < 1, \tag{88}$$

Lemma 5.5 applies and proves for all $\overline{\mathcal{P}}_{i'} \in \mathbb{T}^{\rm nc}(\mathcal{P}_{i'})$ that

$$\|p - p_{\overline{\mathcal{P}}_{i'}}\|_{\mathbb{P}} \le (1-q^2)^{1/2}\,\|p - P_{i'\underline{j}}\|_{\mathbb{P}} \quad\Longrightarrow\quad \#\mathcal{P}_{i'+1} - \#\mathcal{P}_{i'} \lesssim \#\overline{\mathcal{P}}_{i'} - \#\mathcal{T}_{\rm init}.$$

We choose $\overline{\mathcal{P}}_{i'}$ from the definition (66) of the approximation norm $\mathbb{A}_s^{\rm c}$ such that

$$\begin{aligned}\#\overline{\mathcal{P}}_{i'} - \#\mathcal{T}_{\rm init} \le (\mathbb{A}_s^{\rm c}/\varepsilon_{i'})^{1/s} \quad \text{with} \quad & \eta(\overline{\mathcal{P}}_{i'}; \boldsymbol{U}_{\overline{\mathcal{P}}_{i'}}[p_{\overline{\mathcal{P}}_{i'}}], p_{\overline{\mathcal{P}}_{i'}}) + \|\nabla\cdot \boldsymbol{U}_{\overline{\mathcal{P}}_{i'}}[p_{\overline{\mathcal{P}}_{i'}}]\|_\Omega \\ & \le \varepsilon_{i'} := \frac{(1-q^2)^{1/2}}{C'_{\rm rel}(\kappa_1)}\,\|p - P_{i'\underline{j}}\|_{\mathbb{P}}.\end{aligned}$$

Reliability (30) shows that $\|p - p_{\overline{\mathcal{P}}_{i'}}\|_{\mathbb{P}} \le C_{\rm rel}\big(\eta(\overline{\mathcal{P}}_{i'}; \boldsymbol{U}_{\overline{\mathcal{P}}_{i'}}[p_{\overline{\mathcal{P}}_{i'}}], p_{\overline{\mathcal{P}}_{i'}}) + \|\nabla\cdot \boldsymbol{U}_{\overline{\mathcal{P}}_{i'}}[p_{\overline{\mathcal{P}}_{i'}}]\|_\Omega\big)$. With $C_{\rm rel} \le C'_{\rm rel}(\kappa_1)$, Lemma 4.6 and Lemma 4.7 (b) yield that

$$\#\mathcal{P}_i - \#\mathcal{T}_{\rm init} = \sum_{i'=0}^{i-1}(\#\mathcal{P}_{i'+1} - \#\mathcal{P}_{i'}) \lesssim (\mathbb{A}_s^{\rm c})^{1/s} \sum_{i'=0}^{i-1} \|p - P_{i'\underline{j}}\|_{\mathbb{P}}^{-1/s} \overset{\rm (b)}{\lesssim} (\mathbb{A}_s^{\rm c})^{1/s}\,\|p - P_{(i-1)\underline{j}}\|_{\mathbb{P}}^{-1/s}.$$

Next, we prove that $\|p - P_{(i-1)\underline{j}}\|_{\mathbb{P}}^{-1/s} \lesssim \big(\eta_{ijk} + \|\nabla\cdot \boldsymbol{U}_{ijk}\|_\Omega\big)^{-1/s}$. To this end, we apply Lemma 4.8 (a)–(d) and Lemma 4.6. For $i, j > 0$, it holds that

$$\begin{aligned}\eta_{ijk} + \|\nabla\cdot \boldsymbol{U}_{ijk}\|_\Omega &\overset{\rm (c)}{\lesssim} \eta_{ij0} + \|\nabla\cdot \boldsymbol{U}_{ij0}\|_\Omega \overset{\rm (b)}{\lesssim} \eta_{i(j-1)\underline{k}} + \|\nabla\cdot \boldsymbol{U}_{i(j-1)\underline{k}}\|_\Omega \overset{\rm (d)}{\lesssim} \eta_{i0\underline{k}} + \|\nabla\cdot \boldsymbol{U}_{i0\underline{k}}\|_\Omega \\ &\overset{\rm (c)}{\lesssim} \eta_{i00} + \|\nabla\cdot \boldsymbol{U}_{i00}\|_\Omega \overset{\rm (a)}{\lesssim} \eta_{(i-1)\underline{jk}} + \|\nabla\cdot \boldsymbol{U}_{(i-1)\underline{jk}}\|_\Omega \overset{(57)}{\simeq} \|p - P_{(i-1)\underline{j}}\|_{\mathbb{P}}.\end{aligned}$$

Note that the overall estimate is also true if $j = 0$. This proves that $\#\mathcal{P}_i - \#\mathcal{T}_{\rm init} \lesssim (\mathbb{A}_s^{\rm c})^{1/s}\big(\eta_{ijk} + \|\nabla\cdot \boldsymbol{U}_{ijk}\|_\Omega\big)^{-1/s}$. With (87), we obtain that

$$\#\mathcal{M}_{ijk} \lesssim (1 + (\mathbb{A}_s^{\rm c})^{1/s})\big(\eta_{ijk} + \|\nabla\cdot \boldsymbol{U}_{ijk}\|_\Omega\big)^{-1/s}.$$

This concludes the proof. ∎

Proof of Theorem 5.3. The proof is split into two steps.

Step 1. We show the lower bound in (72). Recall that $P_{ij} \in \mathbb{P}(\mathcal{P}_i) \subseteq \mathbb{P}(\mathcal{T}_{ijk})$ for all $(i,j,k) \in \mathcal{Q}$. Therefore, Lemma 5.2 gives that

$$\varrho(\mathcal{T}_{ijk}) \overset{(70)}{\lesssim} \eta(\mathcal{T}_{ijk}; \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}], P_{ij}) + \|\nabla\cdot \boldsymbol{U}_{\mathcal{T}_{ijk}}[P_{ij}]\|_\Omega \overset{(41)}{\simeq} \eta_{ijk} + \|\nabla\cdot \boldsymbol{U}_{ijk}\|. \tag{89}$$

If there exists some $(i,j,k) \in \mathcal{Q}$ such that $\mathcal{T}_{ijk} = \mathcal{T}_{i'j'k'}$ for all $(i',j',k') \in \mathcal{Q}$ with $(i,j,k) \le (i',j',k')$, then, $\varrho(\mathcal{T}_{i'j'k'}) = \varrho(\mathcal{T}_{ijk})$, (70), and convergence (39) yield that $\varrho(\mathcal{T}_{i'j'k'}) = 0$ and hence $\mathbb{A}_s^{\rm c} < \infty$. Otherwise, let $N \in \mathbb{N}_0$ and let $(i,j,k) \in \mathcal{Q}$ be the largest possible index (with respect to "$\le$") such that $\#\mathcal{T}_{ijk} - \#\mathcal{T}_{\rm init} \le N$, i.e., $\mathcal{T}_{ijk} \in \mathbb{T}_N^{\rm c}$. Clearly, it holds that $k < \underline{k}(i,j)$. Therefore, the son estimate (M2) yields that

$$N + 1 < \#\mathcal{T}_{ij(k+1)} - \#\mathcal{T}_{\rm init} + 1 \simeq \#\mathcal{T}_{ij(k+1)} \overset{\rm (M2)}{\simeq} \#\mathcal{T}_{ijk} \simeq \#\mathcal{T}_{ijk} - \#\mathcal{T}_{\rm init} + 1.$$

Together with (89), this leads to

$$\min_{\mathcal{T} \in \mathbb{T}_N^{\rm c}} (N+1)^s \varrho(\mathcal{T}) \lesssim (\#\mathcal{T}_{ijk} - \#\mathcal{T}_{\rm init} + 1)^s \varrho(\mathcal{T}_{ijk}).$$

Taking the supremum over all $(i,j,k) \in \mathcal{Q}$, and then over all $N \in \mathbb{N}_0$, we conclude the first step.

Step 2. We show the upper bound in (72). According to the closure estimate (M3) and Lemma 5.6, it holds for all $(i',j',k') \in \mathcal{Q}$ with $\mathcal{T}_{i'j'k'} \neq \mathcal{T}_{\rm init}$ that

$$\begin{aligned}\#\mathcal{T}_{i'j'k'} - \#\mathcal{T}_{\rm init} + 1 \simeq \#\mathcal{T}_{i'j'k'} - \#\mathcal{T}_{\rm init} &\overset{\rm (M3)}{\lesssim} \sum_{(i,j,k) \le (i',j',k')} \#\mathcal{M}_{ijk} \\ &\overset{(80)}{\lesssim} (1 + (\mathbb{A}_s^{\rm c})^{1/s}) \sum_{(i,j,k) \le (i',j',k')} \big(\eta_{ijk} + \|\nabla \cdot \boldsymbol{U}_{ijk}\|_\Omega\big)^{-1/s}.\end{aligned}$$

Hence, linear convergence (39) in combination with Lemma 4.7 (a) gives that

$$\#\mathcal{T}_{i'j'k'} - \#\mathcal{T}_{\rm init} + 1 \lesssim (1 + (\mathbb{A}_s^{\rm c})^{1/s})^s \big(\eta_{i'j'k'} + \|\nabla \cdot \boldsymbol{U}_{i'j'k'}\|_\Omega\big)^{-1/s}$$

for all for all $(i',j',k') \in \mathcal{Q}$ with $\mathcal{T}_{i'j'k'} \neq \mathcal{T}_{\rm init}$. For all other $(i',j',k') \in \mathcal{Q}$ with $\mathcal{T}_{i'j'k'} = \mathcal{T}_{\rm init}$, the latter estimate is clear. With $(1 + (\mathbb{A}_s^{\rm c})^{1/s})^s \lesssim 1 + \mathbb{A}_s^{\rm c}$, we conclude the proof. ■

Appendix A. Contraction property of N_α

The norm of a self-adjoint operator $T : H \to H$ on a Hilbert space H satisfies that

$$\|T\| = \max\{|\mu|, |M|\}, \quad \text{where} \quad \mu := \inf_{x \in H\setminus\{0\}} \frac{\langle Tx\,,\,x\rangle_H}{\|x\|_H^2} \text{ and } M := \sup_{x \in H\setminus\{0\}} \frac{\langle Tx\,,\,x\rangle_H}{\|x\|_H^2}.$$

If T is positive semi-definite (i.e., $\langle Tx\,,\,x\rangle_H \ge 0$ for all $x \in H$), then

$$\|T\| = \sup_{x \in H\setminus\{0\}} \frac{\langle Tx\,,\,x\rangle_H}{\|x\|_H^2}.$$

Consider $H = \mathbb{P}$. Let $0 < \alpha < 2\,\|S\|^{-1}$. Since the Schur complement operator $S = \nabla \cdot \Delta^{-1} \nabla : \mathbb{P} \to \mathbb{P}$ is self-adjoint, also the operator $T := I - \alpha S$ is self-adjoint. Moreover, S is positive definite. Hence,

$$\mu = \inf_{q \in \mathbb{P}\setminus\{0\}} \frac{\langle (I - \alpha S)q\,,\,q\rangle_\Omega}{\|q\|_\Omega^2} = 1 - \alpha \sup_{q \in \mathbb{P}\setminus\{0\}} \frac{\langle Sq\,,\,q\rangle_\Omega}{\|q\|_\Omega^2} = 1 - \alpha\,\|S\| > -1$$

as well as

$$M = \sup_{q \in \mathbb{P}\setminus\{0\}} \frac{\langle (I - \alpha S)q\,,\,q\rangle_\Omega}{\|q\|_\Omega^2} = 1 - \alpha \inf_{q \in \mathbb{P}\setminus\{0\}} \frac{\langle Sq\,,\,q\rangle_\Omega}{\|q\|_\Omega^2} < 1.$$

Altogether, $\|I-\alpha S\| = \max\{|\mu|,|M|\} < 1$ and thus $N_\alpha : \mathbb{P} \to \mathbb{P}$ from (4) is a contraction.

Appendix B. Proof of (9)

It suffices to prove the inequality for $\boldsymbol{v}$ in the dense subspace $C_c^\infty(\Omega)^n \subseteq H_0^1(\Omega) = \mathbb{V}$. Integration by parts and the fact that $\partial_k \partial_j \boldsymbol{v}_j = \partial_j \partial_k \boldsymbol{v}_j$ show that

$$\begin{aligned}
\|\nabla \cdot \boldsymbol{v}\|_\Omega^2 &= \sum_{j,k=1}^n \langle \partial_j \boldsymbol{v}_j \,,\, \partial_k \boldsymbol{v}_k \rangle_\Omega = -\sum_{j,k=1}^n \langle \partial_k \partial_j \boldsymbol{v}_j \,,\, \boldsymbol{v}_k \rangle_\Omega = -\sum_{j,k=1}^n \langle \partial_j \partial_k \boldsymbol{v}_j \,,\, \boldsymbol{v}_k \rangle_\Omega \\
&= \sum_{j,k=1}^n \langle \partial_k \boldsymbol{v}_j \,,\, \partial_j \boldsymbol{v}_k \rangle_\Omega \le \sum_{j,k=1}^n \|\partial_k \boldsymbol{v}_j\|_\Omega \|\partial_j \boldsymbol{v}_k\|_\Omega \le \frac{1}{2} \sum_{j,k=1}^n \big(\|\partial_k \boldsymbol{v}_j\|_\Omega^2 + \|\partial_j \boldsymbol{v}_k\|_\Omega^2 \big) = \|\nabla \boldsymbol{v}\|_\Omega^2.
\end{aligned}$$

Appendix C. Proof of Remark 5.4

Proof of (73). First, $\mathbb{A}_s^{\mathrm{nc}}(p) \le \mathbb{A}_s^{\mathrm{c}}(p)$ is trivially satisfied due to $\mathbb{T}^{\mathrm{c}} \subseteq \mathbb{T}^{\mathrm{nc}}$. To see the converse inequality, let $N \in \mathbb{N}_0$ be arbitrary and $\mathcal{P}' \in \mathbb{T}_N^{\mathrm{nc}}$ with $\varrho_p(\mathcal{P}') = \min_{\mathcal{P} \in \mathbb{T}_N^{\mathrm{nc}}} \varrho_p(\mathcal{P})$. According to (M4), we have that $\mathtt{close}(\mathcal{P}) \in \mathbb{T}^{\mathrm{c}}_{C_{\mathrm{cls}} N}$. Thus, monotonicity of ϱ_p gives that

$$\begin{aligned}
\min_{\mathcal{T} \in \mathbb{T}^{\mathrm{c}}_{\lfloor C_{\mathrm{cls}} N \rfloor}} (C_{\mathrm{cls}} N + 1)^s \varrho_p(\mathcal{T}) &\le (C_{\mathrm{cls}} N + 1)^s \varrho_p(\mathtt{close}(\mathcal{P}')) \le (C_{\mathrm{cls}} + 1)^s (N+1)^s \varrho_p(\mathcal{P}') \\
&= (C_{\mathrm{cls}} + 1)^s (N+1)^s \min_{\mathcal{P} \in \mathbb{T}_N^{\mathrm{nc}}} \varrho_p(\mathcal{P}) \le (C_{\mathrm{cls}} + 1)^s \mathbb{A}_s^{\mathrm{nc}}(p).
\end{aligned}$$

Finally, elementary estimation yields for arbitrary $M \in \mathbb{N}_0$ and $N := \lfloor M / C_{\mathrm{cls}} \rfloor$ that

$$\min_{\mathcal{T} \in \mathbb{T}_M^{\mathrm{c}}} (M+1)^s \varrho_p(\mathcal{T}) \lesssim \min_{\mathcal{T} \in \mathbb{T}^{\mathrm{c}}_{\lfloor C_{\mathrm{cls}} N \rfloor}} (C_{\mathrm{cls}} N + 1)^s \varrho_p(\mathcal{T}) \le 2^s \mathbb{A}_s^{\mathrm{nc}}(p).$$

Taking the supremum over all $M \in \mathbb{N}_0$, we conclude the proof. ∎

Proof of (74). Clearly, $\varrho_p(\mathcal{T}) + \varrho_{\boldsymbol{u}}(\mathcal{T}) = \varrho_{\boldsymbol{u},p}(\mathcal{T})$. Hence, $\mathbb{A}_s^{\mathrm{c}}(p) + \mathbb{A}_s^{\mathrm{c}}(\boldsymbol{u}) \le 2\, \mathbb{A}_s^{\mathrm{c}}(\boldsymbol{u}, p)$. Moreover, the overlay estimate (M1) also proves the converse estimate. To see this, let $N \in \mathbb{N}_0$. If N is even, choose $n' = N/2 = n'' \in \mathbb{N}_0$. If N is odd, choose $n' = (N-1)/2$, $n'' = (N+1)/2 \in \mathbb{N}_0$. Choose $\mathcal{T}' \in \mathbb{T}^{\mathrm{c}}_{n'}$ such that $\varrho_p(\mathcal{T}') = \min_{\mathcal{T} \in \mathbb{T}^{\mathrm{c}}_{n'}} \varrho_p(\mathcal{T})$. Choose $\mathcal{T}'' \in \mathbb{T}^{\mathrm{c}}_{n''}$ such that $\varrho_{\boldsymbol{u}}(\mathcal{T}'') = \min_{\mathcal{T} \in \mathbb{T}^{\mathrm{c}}_{n''}} \varrho_{\boldsymbol{u}}(\mathcal{T})$. Then, $n' + n'' = N$ and hence $\mathcal{T} := \mathcal{T}' \oplus \mathcal{T}'' \in \mathbb{T}_N^{\mathrm{c}}$. Moreover,

$$\begin{aligned}
(N+1)^s \, \varrho_{\boldsymbol{u},p}(\mathcal{T}) &\le \Big(\frac{N+1}{n'+1} \Big)^s (n'+1)^s \, \varrho_p(\mathcal{T}') + \Big(\frac{N+1}{n''+1} \Big)^s (n''+1)^s \, \varrho_{\boldsymbol{u}}(\mathcal{T}'') \\
&\le \Big(\frac{N+1}{n'+1} \Big)^s \big(\mathbb{A}_s^{\mathrm{c}}(p) + \mathbb{A}_s^{\mathrm{c}}(\boldsymbol{u}) \big).
\end{aligned}$$

Since $(N+1)/(n'+1) \le 2$, this concludes the proof. ∎

Proof of (75). Reliability (30) implies that $\mathbb{A}_s^{\mathrm{c}}(\boldsymbol{u}, p) \le C_{\mathrm{rel}}^s \, \mathbb{A}_s^{\mathrm{c}}$. If the volume force $\boldsymbol{f}$ is a $\mathcal{T}_{\mathrm{init}}$-piecewise polynomial, efficiency [KS08, Prop. 5.6] yields that, for all $P_\mathcal{T} \in \mathbb{P}(\mathcal{T})$,

$$\begin{aligned}
\eta(\mathcal{T}; \boldsymbol{U}_\mathcal{T}[P_\mathcal{T}], P_\mathcal{T}) + \|\nabla \cdot \boldsymbol{U}_\mathcal{T}[P_\mathcal{T}]\|_\Omega &\overset{\text{[KS08]}}{\lesssim} \|\boldsymbol{u}[p_\mathcal{T}] - \boldsymbol{U}_\mathcal{T}[P_\mathcal{T}]\|_\mathbb{V} + \|p_\mathcal{T} - P_\mathcal{T}\|_\mathbb{P} \\
&\le \|\boldsymbol{u} - \boldsymbol{U}_\mathcal{T}[P_\mathcal{T}]\|_\mathbb{V} + \|\boldsymbol{u} - \boldsymbol{u}[p_\mathcal{T}]\|_\mathbb{V} + \|p_\mathcal{T} - P_\mathcal{T}\|_\mathbb{P} \\
&\overset{(21)}{=} \|\boldsymbol{u} - \boldsymbol{U}_\mathcal{T}[P_\mathcal{T}]\|_\mathbb{V} + \|p - p_\mathcal{T}\|_\mathbb{P} + \|p_\mathcal{T} - P_\mathcal{T}\|_\mathbb{P} \lesssim \|\boldsymbol{u} - \boldsymbol{U}_\mathcal{T}[P_\mathcal{T}]\|_\mathbb{V} + \|p - P_\mathcal{T}\|_\mathbb{P}.
\end{aligned}$$

The hidden constant depends only on $\mathcal{T}_{\rm init}$ and the polynomial degree of $\boldsymbol{f}$. Moreover, it holds that $\boldsymbol{U}_\mathcal{T} := \operatorname{argmin}_{\boldsymbol{V}_\mathcal{T} \in \mathbb{V}(\mathcal{T})} \|\boldsymbol{u} - \boldsymbol{V}_\mathcal{T}\|_\mathbb{V} = \boldsymbol{U}_\mathcal{T}[p]$. Hence, (21) shows that

$$\|\boldsymbol{u} - \boldsymbol{U}_\mathcal{T}[p_\mathcal{T}]\|_\mathbb{V} \le \|\boldsymbol{u} - \boldsymbol{U}_\mathcal{T}\|_\mathbb{V} + \|\boldsymbol{U}_\mathcal{T}[p] - \boldsymbol{U}_\mathcal{T}[p_\mathcal{T}]\|_\mathbb{V} \overset{(21)}{\le} \|\boldsymbol{u} - \boldsymbol{U}_\mathcal{T}\|_\mathbb{V} + C_{\rm div}\, \|p - p_\mathcal{T}\|_\mathbb{P}.$$

Combining the latter two estimates, we prove for $\mathcal{T}_{\rm init}$-piecewise polynomial $\boldsymbol{f}$ that

$$\eta(\mathcal{T}; \boldsymbol{U}_\mathcal{T}[p_\mathcal{T}], p_\mathcal{T}) + \|\nabla \cdot \boldsymbol{U}_\mathcal{T}[p_\mathcal{T}]\|_\Omega \lesssim \min_{\boldsymbol{V}_\mathcal{T} \in \mathbb{V}(\mathcal{T})} \|\boldsymbol{u} - \boldsymbol{V}_\mathcal{T}\|_\mathbb{V} + \min_{Q_\mathcal{T} \in \mathbb{P}(\mathcal{T})} \|p - Q_\mathcal{T}\|_\mathbb{P}.$$

Overall, we thus get the converse estimate $\mathbb{A}_s^{\rm c} \lesssim \mathbb{A}_s^{\rm c}(\boldsymbol{u}, p)$ and hence obtain (75). ∎

Appendix D. Mesh closure estimate (M3) implies (M4)

For $\mathcal{P} \in \mathbb{T}^{\rm nc}$, there exists $J \in \mathbb{N}_0$ as well as $\mathcal{P}_0, \dots, \mathcal{P}_J \in \mathbb{T}^{\rm nc}$ and $\mathcal{M}_j \subseteq \mathcal{P}_j$ such that $\mathcal{P}_0 = \mathcal{T}_{\rm init}$, $\mathcal{P}_j = \texttt{bisect}(\mathcal{P}_{j-1}, \mathcal{M}_{j-1})$ for all $j = 1, \dots, J$, and $\mathcal{P}_J = \mathcal{P}$. Note that $\#\mathcal{P} - \#\mathcal{T}_{\rm init} = \sum_{j=0}^{J-1} \#\mathcal{M}_j$. We define $\mathcal{T}_0, \dots, \mathcal{T}_J \in \mathbb{T}^{\rm c}$ inductively by $\mathcal{T}_0 := \mathcal{T}_{\rm init}$ and $\mathcal{T}_j := \texttt{refine}(\mathcal{T}_{j-1}, \mathcal{M}_{j-1} \cap \mathcal{T}_{j-1})$ for all $j = 1, \dots, J$. By induction on j, we show that $\mathcal{T}_j$ is finer than $\mathcal{P}_j$, i.e., $\mathcal{T}_j \in \mathbb{T}^{\rm nc}(\mathcal{P}_j)$ for all $j = 0, \dots, J$. By definition, the claim is true for $j = 0$ with $\mathcal{P}_0 = \mathcal{T}_0 = \mathcal{T}_{\rm init}$. Hence, we may assume that $\mathcal{T}_{j-1} \in \mathbb{T}^{\rm nc}(\mathcal{P}_{j-1})$, and it remains to show that $\mathcal{T}_j \in \mathbb{T}^{\rm nc}(\mathcal{P}_j)$. Since NVB is a binary refinement rule, $\mathcal{T}_{j-1}$ is already strictly finer than $\mathcal{P}_{j-1}$ on each $T \in \mathcal{M}_{j-1} \backslash \mathcal{T}_{j-1}$, i.e., $\mathcal{T}_{j-1} \in \mathbb{T}^{\rm nc}(\mathcal{P}_j^-)$ with $\mathcal{P}_j^- := \texttt{bisect}(\mathcal{P}_{j-1}, \mathcal{M}_{j-1} \backslash \mathcal{T}_{j-1})$. Note that $\mathcal{P}_j = \texttt{bisect}(\mathcal{P}_j^-, \mathcal{M}_{j-1} \cap \mathcal{T}_{j-1})$. Therefore, $\mathcal{T}_j = \texttt{refine}(\mathcal{T}_{j-1}, \mathcal{M}_{j-1} \cap \mathcal{T}_{j-1}) \in \mathbb{T}^{\rm nc}(\mathcal{P}_j)$, which concludes the induction step. In particular, it holds that $\mathcal{T}_J \in \mathbb{T}^{\rm c} \cap \mathbb{T}^{\rm nc}(\mathcal{P})$ and hence also $\mathcal{T}_J \in \mathbb{T}^{\rm c}(\texttt{close}(\mathcal{P}))$. Thus, the application of (M3) yields that

$$\#\texttt{close}(\mathcal{P}) - \#\mathcal{T}_{\rm init} \le \#\mathcal{T}_J - \#\mathcal{T}_{\rm init} \le C_{\rm cls} \sum_{j=0}^{J-1} \#(\mathcal{M}_{j-1} \cap \mathcal{T}_{j-1}) \le C_{\rm cls} \sum_{j=0}^{J-1} \#\mathcal{M}_{j-1}.$$

Since the last sum equals $\#\mathcal{P} - \#\mathcal{T}_{\rm init}$, this concludes the proof.

References

[BD04] Peter Binev and Ronald DeVore. Fast computation in adaptive tree approximation. *Numer. Math.*, 97(2):193–217, 2004.

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

[BHP17] Alex Bespalov, Alexander Haberl, and Dirk Praetorius. Adaptive FEM with coarse initial mesh guarantees optimal convergence rates for compactly perturbed elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 317:318–340, 2017.

[Bin15] Peter Binev. Tree approximation for hp-adaptivity. *IMI Preprint Series*, 07, 2015.

[BM11] Roland Becker and Shipeng Mao. Quasi-optimality of adaptive nonconforming finite element methods for the Stokes equations. *SIAM J. Numer. Anal.*, 49(3):970–991, 2011.

[BMN02] Eberhard Bänsch, Pedro Morin, and Ricardo H. Nochetto. An adaptive Uzawa FEM for the Stokes problem: Convergence without the inf-sup condition. *SIAM J. Numer. Anal.*, 40(4):1207–1229, 2002.

[Bra03] James H. Bramble. A proof of the inf–sup condition for the Stokes equations on Lipschitz domains. *Math. Models Methods Appl. Sci.*, 13(03):361–371, 2003.

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

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

[CPR13] Carsten Carstensen, Daniel Peterseim, and Hella Rabus. Optimal adaptive nonconforming FEM for the Stokes problem. *Numer. Math.*, 123(2):291–308, 2013.

[DDU02] Stephan Dahlke, Wolfgang Dahmen, and Karsten Urban. Adaptive wavelet methods for saddle point problems—optimal convergence rates. *SIAM J. Numer. Anal.*, 40(4):1230–1262, 2002.

[EGP18] Christoph Erath, Gregor Gantner, and Dirk Praetorius. Optimal convergence behavior of adaptive FEM driven by simple $(h-h/2)$-type error estimators. *Preprint*, arXiv:1805.00715, 2018.

[Fei17] Michael Feischl. Optimal adaptivity for a standard finite element method for the Stokes problem. *Preprint*, arXiv:1710.08289, 2017.

[FHPS18] Thomas Führer, Alexander Haberl, Dirk Praetorius, and Stefan Schimanko. Adaptive BEM with inexact PCG solver yields almost optimal computational costs. *Numer. Math.*, published online first, 2018.

[Gan14] Tsogtgerel Gantumur. On the convergence theory of adaptive mixed finite element methods for the Stokes problem. *Preprint*, arXiv:1403.0895, 2014.

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

[GVL13] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.

[HX13] Jun Hu and Jinchao Xu. Convergence and optimality of the adaptive nonconforming linear element method for the Stokes problem. *J. Sci. Comput.*, 55(1):125–148, 2013.

[Kon06] Yaroslav Kondratyuk. Adaptive finite element algorithms for the Stokes problem: Convergence rates and optimal computational complexity. *Preprint, Department of Mathematics, Utrecht University, The Netherlands*, 1346, 2006.

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

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

[MSV08] Pedro Morin, Kunibert G Siebert, and Andreas Veeser. A basic convergence result for conforming adaptive finite elements. *Math. Models Methods Appl. Sci.*, 18(05):707–737, 2008.

[Sie10] Kunibert G. Siebert. A convergence proof for adaptive finite elements without lower bound. *IMA J. Numer. Anal.*, 31(3):947–970, 2010.

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

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

TU Wien, Institute for Analysis and Scientific Computing, Wiedner Hauptstr. 8–10/E101/4, 1040 Wien, Austria

E-mail address: `{ giovanni.difratta , dirk.praetorius } @asc.tuwien.ac.at`

E-mail address: `gregor.gantner@asc.tuwien.ac.at` (corresponding author)

Pontificia Universidad Católica de Chile, Facultad de Matemáticas, Vickuña Mackenna 4860, Santiago, Chile

E-mail address: `tofuhrer@mat.uc.cl`