

OPTIMALITY OF A STANDARD ADAPTIVE FINITE ELEMENT  
METHOD FOR THE STOKES PROBLEM\*

MICHAEL FEISCHL†

**Abstract.** We prove that the standard adaptive algorithm for the Taylor–Hood discretization of the stationary Stokes problem converges with optimal rate. This is done by developing an abstract framework for quite general problems, which allows us to prove *general quasi-orthogonality* proposed in [C. Carstensen et al., *Comput. Math. Appl.*, 67 (2014), pp. 1195–1253]. This property is the main obstacle towards the optimality proof and therefore is the main focus of this work. The key ingredient is a new connection between the mentioned quasi-orthogonality and *LU*-factorization of infinite matrices.

**Key words.** Stokes, adaptive algorithm, optimality, finite element method, quasi-orthogonality

**AMS subject classification.** 65N30

**DOI.** 10.1137/17M1153170

**1. Introduction.** We consider an adaptive mixed finite element method for the stationary Stokes problem

$$\begin{aligned} -\Delta u + \nabla p &= f, \\ \operatorname{div} u &= 0 \end{aligned}$$

with standard Dirichlet boundary conditions in two space dimensions. We discretize the problem with Taylor–Hood elements and employ a standard adaptive algorithm with Dörfler marking. The goal of this work is to provide a missing link in the theory of rate optimality: linear convergence for indefinite problems.

The theory of rate optimal adaptive algorithms for finite element methods originated in the seminal paper [41] by Stevenson and was further improved in [13] by Cascon et al. (Note that the notion of rate optimality differs from that of instance optimality for adaptive finite element methods introduced in the seminal papers [7, 19].) These papers prove essentially, that a standard adaptive algorithm of the form

Solve → Estimate → Mark → Refine

generates asymptotically optimal meshes for the approximation of the solution of a Poisson problem. The new ideas sparked a multitude of papers applying and extending the techniques to different problems; see, e.g., [14, 34] for conforming methods, [5, 6, 11, 36, 38] for nonconforming methods, [12, 15, 30] for mixed formulations, and [2, 21, 22, 24, 25] for boundary element methods (the list is not exhaustive; see also [10] and the references therein). All the mentioned results, however, focus on symmetric and definite problems in the sense that the underlying equation induces a symmetric and definite operator. The missing link required to extend the theory to nonsymmetric and

---

\*Received by the editors October 20, 2017; accepted for publication (in revised form) March 11, 2019; published electronically May 23, 2019.

<http://www.siam.org/journals/sinum/57-3/M115317.html>

**Funding:** The work of the author was supported by the Australian Research Council (ARC) under grant DE170100222 and by the Deutsche Forschungsgemeinschaft (DFG) under grant CRC 1173.

†Institute for Numerical Simulation, University of Bonn, 53115 Bonn, Germany (michael.feischl@uni-bonn.de).

indefinite problems is the quasi-orthogonality and consequently the linear convergence of the error. The first is usually an estimate of the form

$$(1.1) \quad \|u_{\ell+1} - u_\ell\|^2 \leq c\|u - u_\ell\|^2 - C\|u - u_{\ell+1}\|^2 + \text{terms of higher order}$$

with  $c \approx C \approx 1$ . However, for indefinite and nonsymmetric problems such an estimate does not seem to hold. The first proof of rate optimality for a nonsymmetric problem which does not rely on additional assumptions on the initial mesh is given in [23] for a general second order elliptic operator with nonvanishing diffusion coefficient of the form

$$-\operatorname{div}(A\nabla u) + b \cdot \nabla u + cu = f.$$

This approach, however, relies heavily on the fact that the nonsymmetric part of the operator ( $b \cdot \nabla u + cu$ ) is only a compact perturbation (one differentiation instead of two for the diffusion part). The first optimality proof of a strongly nonsymmetric problem was given in the recent work [20] for a finite element/boundary element discretization of a transmission problem. The proof relies on a novel form of quasi-orthogonality introduced in [10, 23] called *general quasi-orthogonality*. Heuristically, instead of requiring (1.1) to hold for all  $\ell \in \mathbb{N}$ , it must only be true in a cumulative sense. This notion of quasi-orthogonality is the last missing building block to apply the theory developed in [13, 41] and culminating in [10] to prove linear convergence and ultimately rate optimal convergence.

The present work aims to generalize the approach from [20] to include indefinite problems. While the work is concerned with the Stokes problem, the applied methods are quite general and may be useful for other indefinite problems.

Currently available convergence and optimality theory for the Stokes problem is building on the seminal works [4, 16]. For certain nonstandard (Uzawa-type) algorithms for the Stokes problem, the work [32] proves optimal convergence. For nonconforming finite element methods, rate optimality and convergence has been investigated and achieved in [5, 11, 28] by use of the fact that the Crouzeix–Raviart discretization of the Stokes problem essentially transforms it into a symmetric and elliptic problem.

For the standard Taylor–Hood element, the first proofs of adaptive convergence were presented in [37, 40], while the first a posteriori error estimator was presented in [43]. The work [26] gives an optimality proof under the assumption that general quasi-orthogonality is satisfied. This assumption is verified in the present work.

Since the level of technicality is already considerably high in the present two dimensional case, we refrain from presenting the general case. However, there doesn't seem to be any inherent barrier and the proof techniques are expected to transfer.

The remainder of the work is organized as follows: We present the setting and the main theorem in section 2, discuss the key ideas of the proof in section 3, construct a suitable Riesz basis in section 4, and prove the main result in section 5. We conclude the work with sections 6 and 7 containing technical results which are required for the proof.

## 2. General assumptions and the Stokes problem.

**2.1. Preliminaries and notation.** In the following,  $\Omega \subseteq \mathbb{R}^2$  is a polygonal domain with Lipschitz boundary  $\partial\Omega$ . The restriction to  $\mathbb{R}^2$  is not fundamental but simplifies some constructions later on in the technical sections. Given a Lipschitz domain  $\omega \subseteq \mathbb{R}^2$ , we denote by  $H^s(\omega)$  the usual Sobolev spaces for  $s \geq 0$ .

For noninteger values of  $s$ , we use real interpolation to define  $H^s(\omega)$ . Their dual spaces  $\tilde{H}^{-s}(\omega)$  are defined by extending the  $L^2$ -scalar product. We denote by  $H_0^1(\Omega)$  the  $H^1$ -functions with vanishing trace and by  $H^{-1}(\Omega)$  the dual space. The subscript  $\star$  denotes vanishing mean of the functions in the given space, i.e.,  $L_\star^2(\Omega) := \{v \in L^2(\Omega) : \langle v, 1 \rangle_\Omega = 0\}$  (the definition generalizes to  $\tilde{H}^{-s}(\Omega)$  in a straightforward fashion). Finally,  $\mathcal{P}^p(\omega)$  denotes the polynomials of total degree less than or equal to  $p$ .

By  $\|\cdot\|_{\ell_2}$ , we denote the Euclidean distance of vectors in  $\mathbb{R}^n$  as well as the  $\ell_2$ -sequence norm. The norm  $\|\cdot\|_{\ell_2}$  induces the spectral norm for (infinite) matrices denoted by  $\|\cdot\|_2$ . The symbol  $I$  denotes the identity matrix. The size of the matrix is always determined by the context.

**2.2. Mesh refinement.** Let  $\mathcal{T}_0$  be a triangulation of  $\Omega$ . Given two triangulations  $\mathcal{T}, \mathcal{T}'$ , we write  $\mathcal{T}' = \text{refine}(\mathcal{T}, \mathcal{M})$  for some  $\mathcal{M} \subseteq \mathcal{T}$  if  $\mathcal{T}'$  is generated from  $\mathcal{T}$  by refinement of all  $T \in \mathcal{M}$  (and subsequent mesh closure) via the newest vertex bisection. In contrast to the general case where a certain initial labeling of the edges is required to obtain an optimal refinement algorithm [7], the paper [31] shows that in two space dimensions those initial conditions are not necessary. We write  $\mathcal{T}' \in \text{refine}(\mathcal{T})$  if  $\mathcal{T}'$  is generated from  $\mathcal{T}$  by a finite number  $j \in \mathbb{N}$  of iterated newest-vertex-bisection refinements. To keep mesh conformity, one usually has to refine more elements than initially marked for refinement (mesh closure). The number of extra refinements is bounded cumulatively in the sense that

$$\#\mathcal{T}' - \#\mathcal{T}_0 \lesssim \sum_{i=0}^j \#\mathcal{M}_i,$$

where  $\mathcal{M}_j$  are elements marked for refinement in the individual steps of the refinement procedure; see [7, 31, 42]. We denote the set of all possible refinements by  $\mathbb{T} := \text{refine}(\mathcal{T}_0)$ . Given  $\omega \subseteq \Omega$ , we call  $\mathcal{T}'|_\omega$  a local refinement of  $\mathcal{T}$ , if there exists  $\mathcal{T}'' \in \text{refine}(\mathcal{T})$  such that  $\mathcal{T}'|_\omega = \mathcal{T}''|_\omega$ . Given  $T \in \mathcal{T}$  for some  $\mathcal{T} \in \mathbb{T}$ ,  $\text{level}(T)$  denotes the number of bisections necessary to generate  $T$  from a parent element in  $\mathcal{T}_0$ .

We define  $\mathcal{N}(\mathcal{T})$  as the set of nodes of  $\mathcal{T}$  and  $\mathcal{E}(\mathcal{T})$  as the set of edges of  $\mathcal{T}$ . For any triangulation, we define

$$\begin{aligned} \mathcal{P}^p(\mathcal{T}) &:= \{v \in L^2(\Omega) : v|_T \in \mathcal{P}^p(T), T \in \mathcal{T}\}, \\ \mathcal{S}^p(\mathcal{T}) &:= \mathcal{P}^p(\mathcal{T}) \cap H^1(\Omega), \\ \mathcal{S}_0^p(\mathcal{T}) &:= \mathcal{P}^p(\mathcal{T}) \cap H_0^1(\Omega), \\ \mathcal{S}_\star^p(\mathcal{T}) &:= \mathcal{P}^p(\mathcal{T}) \cap H_\star^1(\Omega). \end{aligned}$$

We define  $h_{\mathcal{T}} \in \mathcal{P}^0(\mathcal{T})$  as the mesh-size function by  $h_{\mathcal{T}}|_T := \text{diam}(T)$  for all  $T \in \mathcal{T}$ .

Given a subset  $\Omega' \subseteq \Omega$ , we define the patch

$$\omega(\Omega', \mathcal{T}) := \left\{ T_1 \in \mathcal{T} : \exists T_2 \in \mathcal{T}, T_1 \cap T_2 \neq \emptyset, \int_{T_2 \cap \Omega'} 1 \, dx > 0 \right\}.$$

The extended patches  $\omega^k(\Omega', \mathcal{T})$  are defined iteratively by

$$\omega^1(\Omega', \mathcal{T}) := \omega(\Omega', \mathcal{T}) \quad \text{and} \quad \omega^k(\Omega', \mathcal{T}) := \omega\left(\bigcup \omega^{k-1}(\Omega', \mathcal{T}), \mathcal{T}\right).$$

**2.3. The Stokes problem and the Taylor–Hood element.** Our main goal here is to prove general quasi-orthogonality and hence optimality of an adaptive al-

gorithm for the stationary Stokes problem, which reads

$$(2.1) \quad \begin{aligned} -\Delta \mathbf{u} + \nabla \mathbf{p} &= f \quad \text{in } \Omega, \\ \operatorname{div} \mathbf{u} &= 0 \quad \text{in } \Omega, \\ \mathbf{u} &= 0 \quad \text{on } \partial\Omega, \\ \int_{\Omega} \mathbf{p} dx &= 0 \end{aligned}$$

for given functions  $f \in L^2(\Omega)$  with weak solutions  $\mathbf{u} \in H_0^1(\Omega)^2$  and  $\mathbf{p} \in L^2(\Omega)$ . We define the space  $\mathcal{X} := H_0^1(\Omega)^2 \times L_*^2(\Omega)$ .

The weak formulation of (2.1) reads, find  $(\mathbf{u}, \mathbf{p}) \in \mathcal{X}$  such that all  $(v, q) \in \mathcal{X}$  satisfy

$$(2.2) \quad a((\mathbf{u}, \mathbf{p}), (v, q)) := \int_{\Omega} \nabla \mathbf{u} \cdot \nabla v dx - \int_{\Omega} \mathbf{p} \operatorname{div} v dx - \int_{\Omega} q \operatorname{div} \mathbf{u} dx = \int_{\Omega} f v dx.$$

For the purpose of discretization, we choose standard Taylor–Hood elements defined by

$$\mathcal{X}_{\mathcal{T}} := \mathcal{S}_0^2(\mathcal{T})^2 \times \mathcal{S}_*^1(\mathcal{T}).$$

Thus, the Galerkin formulation reads, find  $(\mathbf{u}_{\mathcal{T}}, \mathbf{p}_{\mathcal{T}}) \in \mathcal{X}_{\mathcal{T}}$  such that all  $(v, q) \in \mathcal{X}_{\mathcal{T}}$  satisfy

$$(2.3) \quad a((\mathbf{u}_{\mathcal{T}}, \mathbf{p}_{\mathcal{T}}), (v, q)) := \int_{\Omega} \nabla \mathbf{u}_{\mathcal{T}} \cdot \nabla v dx - \int_{\Omega} \mathbf{p}_{\mathcal{T}} \operatorname{div} v dx - \int_{\Omega} q \operatorname{div} \mathbf{u}_{\mathcal{T}} dx = \int_{\Omega} f v dx.$$

We use a locally equivalent variation proposed in [26] of the classical error estimator proposed by Verfürth [43], i.e., for all  $T \in \mathcal{T}$  define

$$\begin{aligned} \eta_T(\mathcal{T})^2 &:= \operatorname{diam}(T)^2 \|f + \Delta \mathbf{u}_{\mathcal{T}} - \nabla \mathbf{p}_{\mathcal{T}}\|_{L^2(T)}^2 + \operatorname{diam}(T) \|[\partial_n \mathbf{u}_{\mathcal{T}}]\|_{L^2(\partial T \cap \Omega)}^2 \\ &\quad + \operatorname{diam}(T) \|\operatorname{div}(\mathbf{u}_{\mathcal{T}})|_T\|_{L^2(\partial T)}^2, \end{aligned}$$

where  $[.]$  denotes the jump across an edge of  $\mathcal{T}$ . (Note that there are also other error estimators which could be used here, e.g., those in [37].) The overall estimator reads

$$\eta(\mathcal{T}) := \left( \sum_{T \in \mathcal{T}} \eta_T(\mathcal{T})^2 \right)^{1/2} \quad \text{for all } \mathcal{T} \in \mathbb{T}$$

and satisfies upper and lower error bounds, i.e.,

$$(2.4) \quad C_{\text{rel}}^{-1} \|\mathbf{u} - \mathbf{u}_{\mathcal{T}}, \mathbf{p} - \mathbf{p}_{\mathcal{T}}\|_{\mathcal{X}} \leq \eta(\mathcal{T})^2 \leq C_{\text{eff}} \left( \|\mathbf{u} - \mathbf{u}_{\mathcal{T}}, \mathbf{p} - \mathbf{p}_{\mathcal{T}}\|_{\mathcal{X}}^2 + \operatorname{osc}(\mathcal{T})^2 \right)^{1/2},$$

where the data oscillation term reads

$$\operatorname{osc}(\mathcal{T})^2 := \min_{g \in \mathcal{P}^0(\mathcal{T})} \sum_{T \in \mathcal{T}} \operatorname{diam}(T)^2 \|f - g\|_{L^2(T)}^2.$$

We define  $\mathcal{X}_{\ell} := \mathcal{S}^2(\mathcal{T}_{\ell})^2 \times \mathcal{S}_*^1(\mathcal{T}_{\ell})$  to denote the adaptively generated spaces from the algorithm in section 2.5 below.

**2.4. Inf-sup condition and some notation.** To simplify presentation, we collect velocity and pressure in one variable, i.e.,  $u = (\mathbf{u}, \mathbf{p}) \in \mathcal{X}$  and  $u_{\mathcal{T}} = (\mathbf{u}_{\mathcal{T}}, \mathbf{p}_{\mathcal{T}}) \in \mathcal{X}_{\mathcal{T}}$ . According to [9], the Stokes problem  $a(\cdot, \cdot) : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  satisfies

$$(2.5) \quad \inf_{u \in \mathcal{X}} \sup_{v \in \mathcal{X}} \frac{a(u, v)}{\|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}} > 0.$$

If  $\mathcal{T}_0$  contains at least three triangles, [8] even proves

$$(2.6) \quad \inf_{\mathcal{T} \in \mathbb{T}} \inf_{u \in \mathcal{X}_{\mathcal{T}}} \sup_{v \in \mathcal{X}_{\mathcal{T}}} \frac{a(u, v)}{\|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}} > 0.$$

This shows existence and uniqueness of the weak solutions defined above. In the new notation, we have  $u \in \mathcal{X}$  as well as the discrete solution  $u_{\mathcal{T}} \in \mathcal{X}_{\mathcal{T}}$  by

$$(2.7) \quad a(u, v) = \langle f, v \rangle \quad \text{for all } v \in \mathcal{X} \quad \text{and} \quad a(u_{\mathcal{T}}, v) = \langle f, v \rangle \quad \text{for all } v \in \mathcal{X}_{\mathcal{T}}.$$

**2.5. Adaptive algorithm.** Given a triangulation  $\mathcal{T} \in \mathbb{T}$ , we assume that we can compute the error estimator  $\eta(\mathcal{T}) = \sqrt{\sum_{T \in \mathcal{T}} \eta_T(\mathcal{T})^2}$  exactly. Due to technical reasons, we have to restrict ourselves to adaptive triangulations with mild grading in the sense that there exists  $D_{\text{grad}} \in \mathbb{N}$  such that

$$(2.8) \quad |\text{level}(T) - \text{level}(T')| \leq 1 \quad \text{for all } T' \in \omega^{D_{\text{grad}}}(T, \mathcal{T}) \text{ and all } T \in \mathcal{T}.$$

This condition seems to be required for the present proof and also appears in [18] to prove optimal convergence in the  $L^2$ -norm. By  $\mathbb{T}_{\text{grad}} \subseteq \mathbb{T}$ , we denote all triangulations which satisfy (2.8) for a given  $D_{\text{grad}} \in \mathbb{N}$ . The result [20, Lemma 2.3] shows that the restriction does not alter the optimal convergence rate. Numerical experiments suggest that the restriction is not even necessary for optimal convergence rate, and thus might just be an artifact of the proof. In the following, we assume that  $D_{\text{grad}}$  is sufficiently large to satisfy all the conditions in the proofs below.

We assume that the sequence  $(\mathcal{T}_{\ell})_{\ell \in \mathbb{N}_0} \subset \mathbb{T}_{\text{grad}}$  is generated by an adaptive algorithm of the following form.

---

**Algorithm 1.** Adaptive algorithm

---

Input:  $\ell = 0$ ,  $\mathcal{T}_0$ ,  $D_{\text{grad}} \in \mathbb{N}$ ,  $0 < \theta \leq 1$ ,  $f \in \mathcal{X}^*$ .

For  $\ell = 0, 1, \dots$  do:

1. Compute  $u_{\ell} \in \mathcal{X}_{\ell}$  as the unique solution of

$$(2.9) \quad a(u_{\ell}, v) = \langle f, v \rangle \quad \text{for all } v \in \mathcal{X}_{\ell}.$$

2. Compute error estimator  $\eta_T(\mathcal{T}_{\ell})$  for all  $T \in \mathcal{T}_{\ell}$ .
3. Mark set of minimal cardinality  $\mathcal{M}_{\ell} \subseteq \mathcal{T}_{\ell}$  such that

$$(2.10) \quad \sum_{T \in \mathcal{M}_{\ell}} \eta_T(\mathcal{T}_{\ell}) \geq \theta \sum_{T \in \mathcal{T}_{\ell}} \eta_T(\mathcal{T}_{\ell}).$$

4. Refine at least the elements  $\mathcal{M}_{\ell}$  of  $\mathcal{T}_{\ell}$  to obtain  $\mathcal{T}_{\ell+1}$ .
5. Refine additional elements to ensure that  $\mathcal{T}_{\ell+1}$  satisfies (2.8) (see, e.g., [18, section A.3] for a valid mesh-refinement algorithm).

Output: sequence of meshes  $\mathcal{T}_{\ell}$  and corresponding solutions  $u_{\ell}$ .

---

**2.6. Rate optimality.** We aim to analyze the best possible algebraic convergence rate which can be obtained by the adaptive algorithm. This is mathematically characterized as follows: For the exact solution  $u \in \mathcal{X}$ , we define an approximation class  $\mathbf{A}_s$  by

$$(2.11) \quad u \in \mathbf{A}_s \iff \|u\|_{\mathbf{A}_s} := \sup_{N \in \mathbb{N}} \min_{\substack{\tau \in \mathbb{T} \\ \#\tau - \#\tau_0 \leq N}} N^s \eta(\tau) < \infty.$$

By definition, a convergence rate  $\eta(\tau) = \mathcal{O}(N^{-s})$  is theoretically possible if the optimal meshes are chosen. In view of mildly graded triangulations, we define

$$u \in \mathbf{A}_s^{\text{grad}} \iff \|u\|_{\mathbf{A}_s^{\text{grad}}} := \sup_{N \in \mathbb{N}} \min_{\substack{\tau \in \mathbb{T}_{\text{grad}} \\ \#\tau - \#\tau_0 \leq N}} N^s \eta(\tau) < \infty.$$

In [20, Lemma 2.3], we show that in many situations (including the present setting)  $\mathbf{A}_s^{\text{grad}} = \mathbf{A}_s$  for algebraic rates  $s > 0$ . However, condition (2.8) influences  $\|u\|_{\mathbf{A}_s^{\text{grad}}}$  and leads to a slightly nonstandard adaptive finite element method. In the spirit of [10], rate optimality of the adaptive algorithm means that there exists a constant  $C_{\text{opt}} > 0$  such that

$$C_{\text{opt}}^{-1} \|u\|_{\mathbf{A}_s} \leq \sup_{\ell \in \mathbb{N}_0} \frac{\eta(\tau_\ell)}{(\#\tau_\ell - \#\tau_0 + 1)^{-s}} \leq C_{\text{opt}} \|u\|_{\mathbf{A}_s}$$

for all  $s > 0$  with  $\|u\|_{\mathbf{A}_s} < \infty$ .

**2.7. An abstract theory of rate optimality.** As proved in [10], we need to check the assumptions (A1)–(A4) to ensure rate optimality for a given adaptive algorithm: There exist constant  $C_{\text{red}}, C_{\text{stab}}, C_{\text{qo}}, C_{\text{dlr}}, C_{\text{ref}} > 0$ , and  $0 \leq q_{\text{red}} < 1$  such that the following assumptions hold.

- (A1) *Stability on nonrefined elements:* For all refinements  $\hat{\mathcal{T}} \in \mathbb{T}$  of a triangulation  $\mathcal{T} \in \mathbb{T}$ , for all subsets  $\mathcal{S} \subseteq \mathcal{T} \cap \hat{\mathcal{T}}$  of nonrefined elements, it holds that

$$\left| \left( \sum_{T \in \mathcal{S}} \eta_T(\hat{\mathcal{T}})^2 \right)^{1/2} - \left( \sum_{T \in \mathcal{S}} \eta_T(\mathcal{T})^2 \right)^{1/2} \right| \leq C_{\text{stab}} \|u_T - u_{\hat{\mathcal{T}}}\|_{\mathcal{X}}.$$

- (A2) *Reduction property on refined elements:* Any refinement  $\hat{\mathcal{T}} \in \mathbb{T}$  of a triangulation  $\mathcal{T} \in \mathbb{T}$  satisfies

$$\sum_{T \in \hat{\mathcal{T}} \setminus \mathcal{T}} \eta_T(\hat{\mathcal{T}})^2 \leq q_{\text{red}} \sum_{T \in \mathcal{T} \setminus \hat{\mathcal{T}}} \eta_T(\mathcal{T})^2 + C_{\text{red}} \|u_T - u_{\hat{\mathcal{T}}}\|_{\mathcal{X}}^2.$$

- (A3) *General quasi-orthogonality:* For one sufficiently small  $\varepsilon \geq 0$  (see [10, section 3] for details) the output of the algorithm in section 2.5 satisfies for all  $\ell, N \in \mathbb{N}_0$

$$(2.12) \quad \sum_{k=\ell}^{\ell+N} \left( \|u_{k+1} - u_k\|_{\mathcal{X}}^2 - C_{\text{qo}} \varepsilon \|u - u_k\|_{\mathcal{X}}^2 \right) \leq C_{\text{qo}} \|u - u_{\ell}\|_{\mathcal{X}}^2.$$

- (A4) *Discrete reliability:* For all refinements  $\hat{\mathcal{T}} \in \mathbb{T}$  of a triangulation  $\mathcal{T} \in \mathbb{T}$ , there exists a subset  $\mathcal{R}(\mathcal{T}, \hat{\mathcal{T}}) \subseteq \mathcal{T}$  with  $\mathcal{T} \setminus \hat{\mathcal{T}} \subseteq \mathcal{R}(\mathcal{T}, \hat{\mathcal{T}})$  and  $|\mathcal{R}(\mathcal{T}, \hat{\mathcal{T}})| \leq C_{\text{ref}} |\mathcal{T} \setminus \hat{\mathcal{T}}|$  such that

$$\|u_{\hat{\mathcal{T}}} - u_{\mathcal{T}}\|_{\mathcal{X}}^2 \leq C_{\text{dlr}}^2 \sum_{T \in \mathcal{R}(\mathcal{T}, \hat{\mathcal{T}})} \eta_T(\mathcal{T})^2.$$

While the assumptions (A1), (A2), and (A4) are already known in various forms in the literature, the general quasi-orthogonality (A3) seems to be the main obstacle to the optimality proof.

**2.8. The main result.** The following result shows rate optimality of the adaptive algorithm and is the main result of the paper.

**THEOREM 2.1** (optimality of the adaptive algorithm). *Given sufficiently small  $\theta > 0$  and sufficiently large  $D_{\text{grad}} \geq 1$ , the algorithm in section 2.5 applied to the stationary Stokes problem as described above guarantees rate-optimal convergence, i.e., there exists a constant  $C_{\text{opt}} > 0$  such that*

$$C_{\text{opt}}^{-1} \|u\|_{\mathbf{A}_s^{\text{grad}}} \leq \sup_{\ell \in \mathbb{N}_0} \frac{\eta(\mathcal{T}_\ell)}{(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^{-s}} \leq C_{\text{opt}} \|u\|_{\mathbf{A}_s^{\text{grad}}},$$

as well as

$$C_{\text{eff}}^{-1} C_{\text{opt}}^{-1} \|u\|_{\mathbf{A}_s^{\text{grad}}} \leq \sup_{\ell \in \mathbb{N}_0} \frac{\sqrt{\|u - u_\ell, p - p_\ell\|_{\mathcal{X}}^2 + \text{osc}(\mathcal{T}_\ell)^2}}{(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^{-s}} \leq (1 + C_{\text{rel}}) C_{\text{opt}} \|u\|_{\mathbf{A}_s^{\text{grad}}}$$

for all  $s > 0$  with  $\|u\|_{\mathbf{A}_s} < \infty$ .

*Proof.* Note that both statements are equivalent due to the upper and lower bounds in (2.4). To prove the first statement, we have to specify a mesh-refinement strategy which ensures (2.8) and fits into the framework of [10, section 2.4]. To that end, we use the strategy specified in [18, section A.3] (note that the condition in [18, section A.3] and (2.8) are equivalent up to shape regularity). Then, the result follows immediately from [10, Theorem 4.1] and [20, Lemma 2.3], after we prove the assumptions (A1)–(A4) in section 5. (Note that instead of this proof, we could also use Theorem 5.4 together with [26], in which general quasi-orthogonality is assumed in order to prove optimality.)  $\square$

*Remark.* The condition on  $\theta > 0$  being sufficiently small can be quantified in terms of the constants in (A1)–(A4) (see [10]). However, in general one does not know the constants and hence has to guess  $\theta$ . Practical examples show that a choice of  $\theta = 1/2$  is very robust in many cases. This is in stark contrast to the maximum marking strategy [7, 19], where one can prove that any marking parameter leads to optimality. The maximum strategy, however, is not proven to produce work optimal algorithms. We also would like to mention the work [33] on the Stokes problem, which proves instance optimality for a maximum marking strategy. This work, however, relies heavily on the fact that the Crouzeix–Raviart discretization essentially turns the Stokes problem into a symmetric and elliptic problem.

**3. The key ideas of the proof.** This section presents the two insights which enable the proof of general quasi-orthogonality and thus optimal convergence. The first one in section 3.1 is that general quasi-orthogonality is strongly related to the existence and boundedness of the *LU*-factorization of an infinite stiffness matrix. The second one in section 3.2 is that a certain exponential decay property of infinite matrices guarantees the boundedness of the *LU*-factors.

*Sketch of the proof.* The first idea to prove (A3) and hence the main theorem (Theorem 2.1) is as follows: Find an  $\mathcal{X}$ -orthogonal basis  $\mathbf{v}_1, \mathbf{v}_2, \dots$  of  $\mathcal{X}$  such that all the adaptive spaces generated by the algorithm in section 2.5 are spanned by it, i.e.,  $\mathcal{X}_\ell = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{N_\ell}\}$  for all  $\ell \in \mathbb{N}$  and numbers  $N_\ell \in \mathbb{N}$ . Then, consider the infinite stiffness matrix  $A \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$ ,  $A_{ij} := a(\mathbf{v}_j, \mathbf{v}_i)$  and rewrite the original problem as an infinite matrix problem. Section 3.1 shows that a (block)-*LU*-factorization  $A = LU$  with bounded factors implies general quasi-orthogonality (A3). A bound on

the factors  $L$  and  $U$  would hence conclude the proof as we already mentioned above that the other requirements (A1), (A2), and (A4) are well known.

The main problem here is that the  $LU$ -factorization results of section 3.2.1 are not strong enough to show that the matrix  $A$  has a bounded  $LU$ -factorization. Hence, we require more structure (bandedness, exponential decay) of the infinite stiffness matrix and therefore have to replace the orthogonal basis  $\mathbf{v}_1, \mathbf{v}_2, \dots$  by a hierarchical (wavelet-type) Riesz basis which is local in a certain sense. Unfortunately, this introduces some technical difficulties as well as the need for mildly graded meshes. It seems that the topic of the stability of the  $LU$ -factorization is very complex and not well understood in the literature. Advances in this direction could significantly reduce the length of this work.

One can say that sections 3.1 and 3.2 contain the most interesting ideas and are problem independent in the sense that they only require a bilinear form and a conforming method to be applicable. This means that the main ideas of this work are expected to carry over to other challenging problems in the domain of adaptive algorithms. The remaining sections (particularly section 4) solve the technical details and are tailored to the space  $\mathcal{X}$  in which the Stokes problem is formulated. The techniques are quite general and might also be useful for other problem spaces. However, the proofs are not likely to carry over directly.

**3.1. LU-factorization and general quasi-orthogonality.** Before we proceed, we need to clarify the notion of block structure of matrices.

**DEFINITION 3.1.** *Given a block structure in the sense that there exist numbers  $n_1, n_2, \dots \in \mathbb{N}$  with  $n_1 = 1$  and  $n_i < n_j$  for all  $i \leq j$ , we denote matrix blocks by*

$$M(i, j) := M|_{\{n_i, \dots, n_{i+1}-1\} \times \{n_j, \dots, n_{j+1}-1\}} \in \mathbb{R}^{(n_{i+1}-n_i) \times (n_{j+1}-n_j)}.$$

*By  $M[k] \in \mathbb{R}^{(n_{k+1}-1) \times (n_{k+1}-1)}$ , we denote the restriction of  $M$  to the first  $k \times k$  blocks.*

A matrix  $M \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  has an  $LU$ -factorization if  $M = LU$  for matrices  $L, U \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  such that

$$L_{ii} = 1 \text{ and } L_{ij} = U_{ji} = 0 \quad \text{for all } i, j \in \mathbb{N}, i < j.$$

Given a block structure in the sense of Definition 3.1, we say that  $M$  has a block- $LU$ -factorization if  $A = LU$  and  $L$  and  $U$  satisfy

$$L(i, i) = I \text{ and } L(i, j) = U(j, i) = 0 \quad \text{for all } i, j \in \mathbb{N}, i < n_r \leq j, \text{ for some } r \in \mathbb{N}.$$

A sufficient condition for the existence of a block- $LU$ -factorization is the matrix inf-sup condition

$$(3.1) \quad \sup_{k \in \mathbb{N}} \|(M[k])^{-1}\|_2 = C_{\inf-\sup} < \infty.$$

The following result connects the existence of bounded  $LU$ -factors with general quasi-orthogonality.

**THEOREM 3.2.** *Let there exist a Riesz bases  $(w_n)_{n \in \mathbb{N}}$  of  $\mathcal{X}$  and a constant  $C > 0$  such that all  $x = \sum_{n \in \mathbb{N}} \lambda_n w_n \in \mathcal{X}$  satisfy*

$$(3.2) \quad C^{-1} \|x\|_2^2 \leq \sum_{n \in \mathbb{N}} \lambda_n^2 \leq C \|x\|_2^2$$

*and there holds  $\mathcal{X}_\ell = \text{span}\{w_n : n \in \{1, \dots, N_\ell\}\}$  for some constants  $N_\ell \in \mathbb{N}$  with  $N_\ell < N_{\ell+1}$  (a block structure in the sense of Definition 3.1). Assume that*

for some  $\varepsilon > 0$ , there exists  $M^\varepsilon \in \mathbb{R}^{N \times N}$  such that  $M_{ij} := a(w_j, w_i)$ ,  $M \in \mathbb{R}^{N \times N}$  satisfies  $\|M - M^\varepsilon\|_2 \leq \varepsilon$ . If  $M$  and  $M^\varepsilon$  satisfy (3.1) and  $M^\varepsilon = LU$  has a block-LU-factorization such that  $U, U^{-1}: \ell_2 \rightarrow \ell_2$  are bounded operators, then there holds general quasi-orthogonality (2.12). The constant  $C_{\text{qo}} > 0$  depends only on the basis  $(w_n)$ ,  $a$ ,  $C_{\text{inf-sup}}$ ,  $C$ , and on the norms of  $U, U^{-1}$ .

We postpone the proof of Theorem 3.2 to the end of this section.

**LEMMA 3.3.** *Let  $(w_n)_{n \in \mathbb{N}}$  denote the Riesz basis from the statement of Theorem 3.2. If  $M_{ij} := a(w_j, w_i)$  and  $M$  has a block-LU-factorization with bounded operators  $U, U^{-1}: \ell_2 \rightarrow \ell_2$ , then there holds general quasi-orthogonality (2.12) even with  $\varepsilon = 0$ . The constant  $C_{\text{qo}}$  depends only on the basis  $(w_n)$ ,  $C_{\text{inf-sup}}$ , the norm of  $a(\cdot, \cdot)$ ,  $C$ , and  $\mathcal{X}$ .*

*Proof.* In the following proof, we identify vectors  $\lambda \in \mathbb{R}^N$  with vectors  $\lambda \in \mathbb{R}^N$  by adding zero entries on the right, i.e.,  $\lambda = (\lambda_1, \dots, \lambda_N, 0, \dots)$ . The same is done with matrices  $M \in \mathbb{R}^{N \times N}$  by embedding them into the upper-left corner of an infinite zero matrix. By (3.2) and with (2.5)–(2.6), the matrix  $M$  is bounded and satisfies (3.1). Let  $u = \sum_{n=1}^{\infty} \lambda_n w_n$  and  $u_\ell = \sum_{n=1}^{\infty} \lambda(\ell)_n w_n$ . With  $\lambda := (\lambda_1, \lambda_2, \dots)$  and  $\lambda(\ell) := (\lambda(\ell)_1, \lambda(\ell)_2, \dots, \lambda(\ell)_{N_\ell})$ ,  $F := (\langle f, w_1 \rangle, \langle f, w_2 \rangle, \dots) \in \mathbb{R}^N$ , there holds

$$(3.3) \quad M\lambda = F \quad \text{and} \quad M[\ell]\lambda(\ell) = F[\ell] \quad \text{for all } \ell \in \mathbb{N},$$

where  $F[\ell] := (F_1, \dots, F_{N_\ell})$ . Moreover, there holds  $M[\ell] = L[\ell]U[\ell]$  for any block-LU-factorization. Due to the lower-triangular structure of  $L$  (note that  $L(i, i) = I$ ), there holds for all  $1 \leq i \leq N_\ell$  that

$$(L[\ell]U[\ell]\lambda(\ell))_i = (M[\ell]\lambda(\ell))_i = F[\ell]_i = F_i = (M\lambda)_i = (LU\lambda)_i = (L[\ell]U\lambda)_i.$$

Since  $L$  and hence also  $L[\ell]$  is regular, this shows that  $(U\lambda)_i = (U[\ell]\lambda(\ell))_i$  for all  $1 \leq i \leq N_\ell$ . Moreover, there holds  $U[\ell]\lambda(\ell) = U\lambda(\ell)$  due to the block upper-triangular structure of  $U$ . Altogether, this proves

$$(U\lambda)_i = (U\lambda(\ell))_i \quad \text{for all } 1 \leq i \leq N_\ell \quad \text{and} \quad (U\lambda(\ell))_i = 0 \quad \text{for all } i > N_\ell.$$

Hence, we have, by use of the boundedness of  $U$  and  $U^{-1}$  and (3.2), that

$$\begin{aligned} \|u_{k+1} - u_k\|_{\mathcal{X}} &\simeq \|\lambda(k+1) - \lambda(k)\|_{\ell_2} \\ &\simeq \|U\lambda(k+1) - U\lambda(k)\|_{\ell_2} = \|(U\lambda)|_{\{N_k, \dots, N_{k+1}-1\}}\|_{\ell_2}. \end{aligned}$$

This shows

$$\begin{aligned} \sum_{k=\ell}^{\infty} \|u_{k+1} - u_k\|_{\mathcal{X}}^2 &\simeq \sum_{k=\ell}^{\infty} \|(U\lambda)|_{\{N_k, \dots, N_{k+1}-1\}}\|_{\ell_2}^2 = \|(U\lambda)|_{\{N_\ell, \dots\}}\|_{\ell_2}^2 \\ &= \|U\lambda - U\lambda(\ell)\|_{\ell_2}^2 \simeq \|\lambda - \lambda(\ell)\|_{\ell_2}^2 \simeq \|u - u_\ell\|_{\mathcal{X}}^2. \end{aligned}$$

Hence, we conclude the proof.  $\square$

*Proof of Theorem 3.2.* In the following, instead of function spaces, we consider an  $\ell_2$ -setting of infinite vectors and apply Lemma 3.3 to it, i.e.,  $\mathcal{X} = \ell_2$ . With (3.2), the boundedness of  $a(\cdot, \cdot)$  implies  $\|M\|_2 + \|M^\varepsilon\|_2 < \infty$ . With the notation from the proof of Lemma 3.3, we apply Lemma 3.3 to the bilinear form  $a^\varepsilon: \ell_2 \times \ell_2 \rightarrow \mathbb{R}$ ,  $a^\varepsilon(x, y) := \langle M^\varepsilon x, y \rangle_{\ell_2}$  and  $f^\varepsilon := M^\varepsilon \lambda \in \ell_2$  and the spaces

$$\mathcal{X}_\ell := \{x \in \ell_2 : x_i = 0 \text{ for } i > N_\ell\}.$$

Boundedness of  $M^\varepsilon$  together with (3.1) imply boundedness and the inf-sup condition (2.5) for  $a^\varepsilon(\cdot, \cdot)$ . We use the  $\ell_2$  unit vectors as the Riesz basis to obtain with Lemma 3.3

$$(3.4) \quad \sum_{k=\ell}^{\infty} \|\lambda^\varepsilon(k+1) - \lambda^\varepsilon(k)\|_{\ell_2}^2 \lesssim \|\lambda - \lambda^\varepsilon(\ell)\|_{\ell_2}^2.$$

Here, we used that  $\lambda^\varepsilon = \lambda$  by the definition of  $f^\varepsilon$ .

As in the previous proof, we identify vectors in  $\mathbb{R}^n$  with vectors in  $\mathbb{R}^{\mathbb{N}}$  by adding zeros to the right. Then, there holds with (2.5)

$$\begin{aligned} \|\lambda^\varepsilon(k) - \lambda(k)\|_{\ell_2} &\lesssim \sup_{\substack{\mu \in \mathcal{X}_k \\ \|\mu\|_{\ell_2}=1}} a^\varepsilon(\lambda^\varepsilon(k) - \lambda(k), \mu) \\ &= \sup_{\substack{\mu \in \mathcal{X}_k \\ \|\mu\|_{\ell_2}=1}} a^\varepsilon(\lambda - \lambda(k), \mu) \\ &= \sup_{\substack{\mu \in \mathcal{X}_k \\ \|\mu\|_{\ell_2}=1}} \langle (M^\varepsilon - M)(\lambda - \lambda(k)), \mu \rangle_{\ell_2} \\ &\leq \|M^\varepsilon - M\|_2 \|\lambda - \lambda(k)\|_{\ell_2}. \end{aligned}$$

Hence, we have

$$|\|\lambda^\varepsilon(k+1) - \lambda^\varepsilon(k)\|_{\ell_2} - \|\lambda(k+1) - \lambda(k)\|_{\ell_2}| \lesssim \varepsilon (\|\lambda - \lambda(k)\|_{\ell_2} + \|\lambda - \lambda(k+1)\|_{\ell_2}),$$

where the hidden constant is independent of  $\varepsilon > 0$ . With (3.4) and (3.2), this concludes

$$\sum_{k=\ell}^{\infty} \left( \|u_{k+1} - u_k\|_{\mathcal{X}}^2 - 2C^2\varepsilon \|u - u_k\|_{\mathcal{X}}^2 \right) \lesssim C^2 \|u - u_\ell\|_{\mathcal{X}}^2$$

and hence the assertion follows.  $\square$

**3.2. Jaffard class matrices and LU-factorization.** This section discusses the second main insight of this work, namely, that a particular mixture of exponential decay and block-bandedness ensures bounded LU-factorization. (This may not be viewed as new in the general matrix context; however, the available literature does not seem to be general enough to be applicable to this problem.)

Jaffard class matrices generalize the notion of matrices which decay exponentially away from the diagonal. The generalization allows one to replace the distance  $|i - j|$  between indices by a general metric  $d(i, j)$ . This class was introduced and analyzed in [29].

**DEFINITION 3.4** (Jaffard class). *We say that an infinite matrix  $M \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  is of Jaffard class,  $M \in \mathcal{J}(d, \gamma, C)$  for some metric  $d(\cdot, \cdot): \mathbb{N} \times \mathbb{N} \rightarrow [0, \infty)$  and some  $\gamma > 0$ , if for all  $0 < \gamma' < \gamma$  there exists  $C(\gamma') > 0$  such that*

$$(3.5) \quad |M_{ij}| \leq C(\gamma') \exp(-\gamma' d(i, j)) \quad \text{for all } i, j \in \mathbb{N}.$$

Moreover, the metric  $d(\cdot, \cdot)$  must satisfy for all  $\varepsilon > 0$

$$(3.6) \quad \sup_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} \exp(-\varepsilon d(i, j)) < \infty.$$

We also write  $M \in \mathcal{J}$  to state the existence of parameters  $d, \gamma, C$  such that  $M \in \mathcal{J}(d, \gamma, C)$ .

DEFINITION 3.5 (banded matrix). *We say that an infinite matrix  $M \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  is banded with respect to some metric  $d(\cdot, \cdot): \mathbb{N} \times \mathbb{N} \rightarrow [0, \infty)$  if there exists a bandwidth  $b \geq 1$  such that*

$$(3.7) \quad d(i, j) > b \implies M_{ij} = 0 \quad \text{for all } i, j \in \mathbb{N}.$$

*In this case, we write  $M \in \mathcal{B}(d, b)$ . Note that we do not require  $d(\cdot, \cdot)$  to satisfy (3.6). We also write  $M \in \mathcal{B}$  or  $M \in \mathcal{B}(d)$  to state that the missing parameters exist.*

The following technical lemmas state some straightforward facts about infinite matrices and can be found in [20].

LEMMA 3.6. *Let  $M^{i,j} \in \mathcal{B}(d, b_j)$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ , for some  $m, n \in \mathbb{N}$  with respect to some metric  $d(\cdot, \cdot)$  and respective bandwidths  $b_j \in \mathbb{N}$ . Then, there holds*

$$\sum_{i=1}^n \prod_{j=1}^m M^{i,j} \in \mathcal{B}\left(d, \sum_{j=1}^m b_j\right).$$

LEMMA 3.7. *Let  $M \in \mathcal{J}(d, \gamma, C)$ , then  $|M|: \ell_2 \rightarrow \ell_2$  is a bounded operator (the modulus  $|M|$  is understood entrywise).*

LEMMA 3.8. *Let  $M \in \mathcal{J}(d, \gamma, C)$  and assume a block structure  $n_1, n_2, \dots \in \mathbb{N}$  such that  $M$  satisfies (3.1). Then,*

$$\overline{M} \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}} \quad \text{with} \quad \overline{M}_{ij} := \sup_{\substack{k \in \mathbb{N} \\ n_k \geq \max\{i, j\}}} |(M[k])^{-1}|_{ij}$$

*is in  $\mathcal{J}(d, \tilde{\gamma}, \tilde{C})$  and thus a bounded operator  $\overline{M}: \ell_2 \rightarrow \ell_2$ . The constant  $\tilde{\gamma}$  depends only on  $C_{\inf-\sup} > 0$ ,  $d$ , and  $\gamma$ , whereas for all  $0 < \gamma' < \tilde{\gamma}$ ,  $\tilde{C}(\gamma')$  depends only on an upper bound for  $C(\gamma')$  and on  $C_{\inf-\sup} > 0$ .*

*Proof.* The result [29, Proposition 2] shows that  $(M[k])^{-1} \in \mathcal{J}(d, \tilde{\gamma}, \tilde{C})$ . Inspection of the proof reveals that  $\tilde{\gamma}$  depends only on  $\gamma$ ,  $d$ , and  $\tilde{C}(\gamma')$  depends only on an upper bound for  $C(\gamma')$  from Definition 3.4 and on  $C_{\inf-\sup} > 0$ . Therefore, we have for all  $0 < \gamma' < \tilde{\gamma}$

$$|\overline{M}_{ij}| \leq \tilde{C}(\gamma') \exp(-\gamma' d(i, j)) \quad \text{for all } i, j \in \mathbb{N}$$

and hence  $\overline{M} \in \mathcal{J}(d, \tilde{\gamma}, \tilde{C})$ . Lemma 3.7 concludes the proof.  $\square$

**3.2.1. LU-factorization.** The following results show existence of a bounded *LU*-factorization for particular Jaffard class matrices.

THEOREM 3.9. *Let  $M \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}} \in \mathcal{J}(d, \gamma, C)$  and additionally satisfy (3.1) for some given block structure in the sense of Definition 3.1. Then,  $M$  has a block-LU-factorization such that  $L, U, L^{-1}, U^{-1}: \ell_2 \rightarrow \ell_2$  are bounded operators with operator norms depending only on  $\mathcal{J}(d, \gamma, C)$ ,  $C_{\inf-\sup}$ , and  $\|M\|_2$ .*

*Proof.* We repeat the proof of [1, Theorem 2] to show that it also works for block-LU-factorization. The existence of block-lower/block-upper triangular matrices  $L, U \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  such that  $M = LU$  is well known. The identity (3.8) shows, for  $x \in \ell_2$  and  $\overline{M}$  from Lemma 3.8

$$\begin{aligned} \|U^{-1}x\|_{\ell_2}^2 &\leq \sum_{i=1}^{\infty} \left\| \sum_{k=i}^{\infty} |M[k]^{-1}(i, k)| |x|_{\{n_k, \dots, n_{k+1}-1\}} \right\|_{\ell_2}^2 \\ &\leq \sum_{i=1}^{\infty} \left\| \sum_{k=i}^{\infty} \overline{M}(i, k) |x|_{\{n_k, \dots, n_{k+1}-1\}} \right\|_{\ell_2}^2 \leq \|\overline{M}\|_2 \|x\|_{\ell_2} \lesssim \|x\|_{\ell_2}. \end{aligned}$$

This shows that  $\|U^{-1}\|_2 < \infty$  and we deduce immediately  $\|L\|_2 \leq \|M\|_2 \|U^{-1}\|_2 < \infty$ . Lemma 3.12 shows additionally that the block-diagonal matrix  $D(i,i) := U(i,i)$ ,  $i \in \mathbb{N}$ , is bounded  $\|D\|_2 < \infty$  in terms of  $C_{\inf-\sup}$  and  $\|M\|_2$ . The same argument proves for  $M^T = \tilde{L}\tilde{U}$  that  $\|\tilde{L}\|_2 + \|\tilde{U}^{-1}\|_2 < \infty$ . From (3.9), we see

$$\|U\|_2 = \|\tilde{L}\|_2 \|D^T\|_2 < \infty \quad \text{and hence} \quad \|L^{-1}\|_2 \leq \|U\|_2 \|M^{-1}\|_2 < \infty.$$

This concludes the proof.  $\square$

Below we present the main theorem of this section. It proves that the *LU*-factors of a banded matrix can be approximated by banded matrices. In some sense, the result is an extension of [35] to a particular case.

**THEOREM 3.10.** *Let  $M \in \mathcal{B}(d, b_0)$  for some  $b_0 \in \mathbb{N}$  and some metric  $d(\cdot, \cdot)$  such that  $M: \ell_2 \rightarrow \ell_2$  is bounded and satisfies (3.1) for some block structure in the sense of Definition 3.1. Let  $M$  be block-banded in the sense  $M(i,j) = 0$  for  $|i-j| > b_0$ . Then, for all  $\varepsilon > 0$ , there exists an approximate block-LDU-decomposition  $\|M - LDU\|_2 \leq \varepsilon$  for block-lower/block-upper triangular factors  $L, U$  such that  $L(i,i) = U(i,i) = I$  for all  $i \in \mathbb{N}$ , and a block diagonal factor  $D$ . The factors  $L, D, U: \ell_2 \rightarrow \ell_2$  are bounded with bounded inverses uniformly in  $\varepsilon$  and satisfy  $L^{-1}, D, U^{-1} \in \mathcal{B}(d, b)$  for some bandwidth  $b$ . Moreover,  $L^{-1}, U^{-1}$  are block-banded with bandwidth  $b$ , i.e.,  $L^{-1}(i,j) = U^{-1}(i,j) = 0$  for all  $|i-j| > b$ . Additionally,  $D$  satisfies (3.1). The constant  $b$  depends only on  $M$ ,  $b_0$ , and  $\varepsilon$ .*

*Remark.* The main idea behind the proof of Theorem 3.10 is to exploit the well-known property of the *LU*-factorization to inherit skyline structure and particularly (block-)bandedness of the parent matrix. The theorem goes one step further and proves that the combination of block-bandedness and bandedness in the sense of Definition 3.5 can be preserved approximately. We use this fact to show that the approximate *LU*-factors are inversely bounded (which seems to be inaccessible without any additional structure).

We postpone the proof of Theorem 3.10 to the end of this section.

**LEMMA 3.11.** *Let  $M \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  such that  $M: \ell_2 \rightarrow \ell_2$  is bounded and satisfies (3.1) for some block structure in the sense of Definition 3.1. Moreover, let  $M \in \mathcal{B}(d, b_0)$ . Then, given  $\varepsilon > 0$ , there exists a bandwidth  $b \in \mathbb{N}$  such that for all  $k \in \mathbb{N}$ , there exist  $R, R_k \in \mathcal{B}(d, b)$  such that*

$$\|M^{-1} - R\|_2 + \sup_{k \in \mathbb{N}} \|M[k]^{-1} - R_k\|_2 \leq \varepsilon.$$

*If  $M$  is additionally block-banded in the sense  $M(i,j) = 0$  for all  $|i-j| > b_0$ , then,  $R_k$  and  $R$  will additionally be block-banded with bandwidth  $b$ . If  $M$  is block diagonal,  $R$  and  $R_k$  will also be block diagonal. The bandwidth  $b$  depends only on  $b_0$ ,  $C_{\inf-\sup}$ ,  $\|M\|_2$ , and  $\varepsilon$ .*

*Proof.* Let  $A := M[k]^T M[k]$  or  $A := M^T M$ . Due to (3.1),  $A$  is elliptic with some constant  $C_{\inf-\sup}^{-2}$ . We obtain for  $\alpha := C_{\inf-\sup}^{-2} \|M\|_2^{-4}$  and  $x \in \ell_2$

$$\begin{aligned} \|x - \alpha A x\|_{\ell_2}^2 &= \|x\|_{\ell_2}^2 - 2\alpha \langle x, Ax \rangle_{\ell_2} + \alpha^2 \|Ax\|_{\ell_2}^2 \\ &\leq (1 - 2\alpha C_{\inf-\sup}^{-2} + \alpha^2 \|A\|_2^2) \|x\|_{\ell_2}^2 \leq (1 - C_{\inf-\sup}^{-4}/\|M\|_2^4) \|x\|_{\ell_2}^2. \end{aligned}$$

This shows  $\|I - \alpha A\|_2 \leq (1 - C_{\inf-\sup}^{-4}/\|M\|_2^4) := q < 1$ . We obtain

$$A^{-1} = \alpha(\alpha A)^{-1} = \alpha(I - (I - \alpha A))^{-1} = \alpha \sum_{j=0}^{\infty} (I - \alpha A)^j.$$

We define  $R := (\sum_{j=0}^N (I - \alpha A)^j) M^T$  and  $R_k := (\sum_{j=0}^N (I - \alpha A)^j) M[k]^T$  for some  $N \in \mathbb{N}$  such that  $\|\sum_{j=N+1}^{\infty} (I - \alpha A)^j\|_2 \leq \sum_{j=N+1}^{\infty} q^j \leq \varepsilon/\|M\|_2$ . Since  $(I - \alpha A) \in \mathcal{B}(d, b_0)$  and  $M \in \mathcal{B}(d, b_0)$ , Lemma 3.6 shows that  $R$  and  $R_k$  are banded as well. The bandwidth depends only on  $b_0$ ,  $q$ , and  $N$ . If  $M$  is additionally block-banded,  $A$  and  $(I - \alpha A)$  will also be block-banded with bandwidth  $2b_0$ . Hence  $(I - \alpha A)^k$  will be block-banded with bandwidth  $2kb_0$ . The same argumentation proves the statement for block diagonal  $M$ . This concludes the proof.  $\square$

The following results prove that block-banded matrices  $M$  hand down some structure to their  $LU$ -factors. This section is similar to [20, section 3.2], however, with the difference that we include indefinite matrices instead of elliptic ones.

**LEMMA 3.12.** *Let  $M \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  such that  $M: \ell_2 \rightarrow \ell_2$  is bounded and satisfies (3.1) for a block structure in the sense of Definition 3.1. Then, the block-LU-factorization  $M = LU$  for block-lower/block-upper triangular matrices  $L, U \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  such that  $L(i, i) = I$  for all  $i \in \mathbb{N}$  exists and satisfies*

$$\sup_{k \in \mathbb{N}} \|U(k, k)\|_2 \leq (1 + \|M\|_2)^2 C_{\inf-\sup}.$$

*Proof.* It is well known that the block-LU-factorization exists. We further note that  $M[k] = L[k]U[k]$  for all  $k \in \mathbb{N}$ , i.e., restriction to principal submatrices commutes with the block-LU-factorization. Therefore, to see that  $U(k, k)$  is uniformly bounded, we may restrict ourselves to  $M[k]$  for  $k \in \mathbb{N}$ . For matrices  $R_1, R_2 \in \mathbb{R}^{(n_k-1) \times (n_{k+1}-n_k)}$  and  $R_3 \in \mathbb{R}^{(n_{k+1}-n_k) \times (n_{k+1}-n_k)}$  the  $(2 \times 2)$ -block-LU-factorization reads

$$\begin{aligned} M[k] &= \begin{pmatrix} M[k-1] & R_1 \\ R_2^T & R_3 \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ R_2^T M[k-1]^{-1} & I \end{pmatrix} \begin{pmatrix} M[k-1] & R_1 \\ 0 & R_3 - R_2^T M[k-1]^{-1} R_1 \end{pmatrix}. \end{aligned}$$

Uniqueness of normalized block-LU-factorization (further factorization of  $M[k-1]$  will not alter the lower-right block of the  $U$ -factor) implies  $U(k, k) = R_3 - R_2^T M[k-1]^{-1} R_1$  and hence  $\|U(k, k)\|_2 \leq \|M[k]\|_2 + \|M[k]\|_2^2 C_{\inf-\sup} \leq (1 + \|M\|_2)^2 C_{\inf-\sup}$ , where we used that the norm of the submatrices  $R_1, R_2, R_3$  is bounded by the norm of the matrix  $M[k]$  as well as  $\|M[k-1]^{-1}\|_2 \leq C_{\inf-\sup}$ .  $\square$

**LEMMA 3.13.** *Let  $M \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  such that  $M: \ell_2 \rightarrow \ell_2$  is bounded and satisfies (3.1) for some block structure in the sense of Definition 3.5. Moreover, let  $M$  be block-banded in the sense  $M(i, j) = 0$  for  $|i - j| > b_0$  for some  $b_0 \in \mathbb{N}$ . Then, the block-LU-factorization  $M = LU$  for block-lower/block-upper triangular matrices  $L, U \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  such that  $L(i, i) = I$  for all  $i \in \mathbb{N}$  exists, is block-banded with bandwidth  $b_0$ , and satisfies*

$$\|L\|_2 + \|U\|_2 + \|L^{-1}\|_2 + \|U^{-1}\|_2 < \infty.$$

*Moreover, the block-diagonal matrix  $D \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$ ,  $D(i, i) := U(i, i)$  as well as its inverse are bounded and satisfy (3.1).*

*Proof.* Since  $M[k]$  is invertible for all  $k \in \mathbb{N}$ , it is a well-known fact that the block-LU-factorization exists. Since  $M[k]$  and  $L[k]$  are invertible by definition,  $U[k]$  is invertible for all  $k \in \mathbb{N}$ . The block-triangular structure guarantees that  $L[k]^{-1} = (L^{-1})[k]$  as well as  $U[k]^{-1} = (U^{-1})[k]$  and, hence, the existence of  $L^{-1}, U^{-1}$  as matrices in  $\mathbb{R}^{\mathbb{N} \times \mathbb{N}}$ . Moreover, it is well known that  $L$  and  $U$  are block-banded with bandwidth  $b_0$ . By definition, we have  $M[k]^{-1} = U[k]^{-1}L[k]^{-1}$ . Since  $L$  is lower triangular with normalized block diagonal (only identities in the diagonal blocks), the same is true for  $L[k]^{-1}$ . Therefore, we obtain

$$(3.8) \quad M[k]^{-1}(i, k) = \sum_{r=1}^k U[k]^{-1}(i, r)L[k]^{-1}(r, k) = U[k]^{-1}(i, k) = U^{-1}(i, k),$$

where the last identity follows from the fact that  $U^{-1}$  is upper-block triangular.

We see that  $\sup_{i,j} \|U^{-1}(i, j)\|_2 \leq \sup_{k \in \mathbb{N}} \|M[k]^{-1}\|_2 \leq C_{\inf-\sup} < \infty$ . Hence, there holds that  $\|D(k, k)^{-1}\|_2 = \|U^{-1}(k, k)\|_2 \leq C_{\inf-\sup}$ . Moreover, we obtain

$$\sup_{i,j \in \mathbb{N}} \|L(i, j)\|_2 \leq \sup_{i,j \in \mathbb{N}} \sum_{\substack{k \in \mathbb{N}_0 \\ |k-i| \leq b_0}} \|M(i, k)U^{-1}(k, j)\|_2 \leq 2b_0\|M\|_2 \sup_{i,j} \|U^{-1}(i, j)\|_2 < \infty.$$

Since  $L$  is block-banded with bandwidth  $b_0$ , the result [20, Lemma 8.4] shows  $\|L\|_2 < \infty$ . This implies  $\|U^{-1}\|_2 \leq \|M^{-1}\|_2\|L\|_2 < \infty$ . Lemma 3.12 shows that  $D(k, k) = U(k, k)$  is uniformly bounded. Thus, we proved  $\|D\|_2 + \|D^{-1}\|_2 < \infty$  depending only on  $C_{\inf-\sup}$  and  $\|M\|_2$ .

Let  $M^T = \tilde{L}\tilde{U}$  be the analogous block-LU-factorization for the transposed matrix (note that  $M^T$  still satisfies (3.1) and is bounded and banded). Since normalized LU-factorizations are unique, we see that

$$(3.9) \quad \tilde{L} = U^T D^{-T} \quad \text{and} \quad \tilde{U} = D^T L^T.$$

Repeating the above arguments shows  $\|\tilde{L}\|_2 + \|\tilde{U}^{-1}\|_2 < \infty$ . With boundedness of  $D$  and  $D^{-1}$ , (3.9) shows  $\|U\|_2 + \|L^{-1}\|_2 < \infty$  and hence concludes the proof.  $\square$

**LEMMA 3.14.** *Under the assumptions of Lemma 3.13, assume that additionally  $M \in \mathcal{B}(d, b_0)$ . Given  $\varepsilon > 0$ , there exists  $b \in \mathbb{N}$  and block-upper triangular  $U_\varepsilon^{-1} \in \mathcal{B}(d, b)$  which is additionally block-banded in the sense  $U_\varepsilon^{-1}(i, j) = 0$  for  $|i - j| > b$  such that*

$$\|U^{-1} - U_\varepsilon^{-1}\|_2 \leq \varepsilon.$$

*The approximation  $U_\varepsilon^{-1}$  is invertible with bounded inverse such that*

$$\sup_{\varepsilon > 0} (\|U_\varepsilon\|_2 + \|U_\varepsilon^{-1}\|_2) < \infty.$$

*Moreover, there exists block diagonal  $D_\varepsilon \in \mathcal{B}(d, b)$  which is bounded and satisfies (3.1) such that  $\|D - D_\varepsilon\|_2 \leq \varepsilon$ .*

*Proof.* Lemma 3.11 shows that there exist  $R, R_k \in \mathcal{B}(d, b)$  which are block-banded with bandwidth  $b = b(\varepsilon)$  such that  $\|M^{-1} - R\|_2 + \|M[k]^{-1} - R_k\|_2 \leq \varepsilon$  for all  $k \in \mathbb{N}$ . Choosing  $\varepsilon > 0$  sufficiently small, we also ensure that  $R$  and  $R_k$  are bounded and satisfy (3.1) with uniform constant.

Inspired by (3.8), we define a first approximation to  $U^{-1}$  by

$$(3.10) \quad T(i, j) := \begin{cases} 0 & \text{for } j < i \text{ or } i < j - b, \\ R_j(i, j) & \text{for } j - b \leq i \leq j. \end{cases}$$

This ensures that  $T \in \mathcal{B}(d, b)$  and that  $T$  is block-banded with bandwidth  $b$ . Additionally, we obtain

$$(3.11) \quad \sup_{i,j \in \mathbb{N}} \|T(i,j) - U^{-1}(i,j)\|_2 \leq \varepsilon.$$

We define an approximation to  $L$  (which is block-banded with bandwidth  $b_0$ ) by

$$S(i,j) := \begin{cases} 0 & \text{for } j < i - b_0 \text{ or } j > i, \\ I & \text{for } j = i, \\ (MT)(i,j) & \text{for } i - b_0 \leq j < i. \end{cases}$$

The definition and (3.11) imply

$$(3.12) \quad \|L(i,j) - S(i,j)\|_2 \leq \left\| \sum_{k=i-b_0}^{i+b_0} M(i,k)(U^{-1}(k,j) - T(k,j)) \right\|_2$$

$$(3.13) \quad \leq \sum_{k=i-b_0}^{i+b_0} \|M\|_2 \|U^{-1}(k,j) - T(k,j)\|_2 \lesssim \|M\|_2 b_0 \varepsilon.$$

Since both  $L$  and  $S$  are block-banded with bandwidth  $b_0$ , the result [20, Lemma 8.4] even shows

$$(3.14) \quad \|L - S\|_2 \lesssim \varepsilon,$$

where the hidden constant is independent of  $\varepsilon$ . Moreover, Lemma 3.6 shows that  $S \in \mathcal{B}(d, \tilde{b})$ , for some  $\tilde{b} \in \mathbb{N}$  which depends only on  $b_0$  and  $b$ .

Recall  $R$  from above with  $\|M^{-1} - R\|_2 \leq \varepsilon$ ,  $R \in \mathcal{B}(d, b)$ , and  $R$  is block-banded with bandwidth  $b$ . This allows us to define  $U_\varepsilon^{-1}$  by

$$U_\varepsilon^{-1} := RS.$$

We obtain from the definition and with (3.14)

$$(3.15) \quad \begin{aligned} \|U_\varepsilon^{-1} - U^{-1}\|_2 &\leq \|R(S - L)\|_2 + \|(R - M^{-1})L\|_2 \lesssim \|R\|_2 \varepsilon + \|L\|_2 \varepsilon \\ &\leq (C_{\inf-\sup} + \varepsilon + 1) \varepsilon, \end{aligned}$$

where the hidden constant does not depend on  $\varepsilon$ . Moreover, Lemma 3.6 shows (since  $S$  and  $R$  are block-banded), that  $U_\varepsilon^{-1} \in \mathcal{B}(d)$  with bandwidth depending on  $\varepsilon$ . Analogously, we see  $U_\varepsilon^{-1}$  is block-banded with bandwidth  $b_0 + b$ . Since  $U^{-1}$  is invertible with bounded inverse, choosing  $\varepsilon > 0$  sufficiently small ensures that  $U_\varepsilon^{-1}$  is invertible, with bounded inverse uniformly in  $\varepsilon$ .

Let  $\tilde{D}$  denote the block diagonal of  $U_\varepsilon^{-1}$ . Obviously,  $\tilde{D} \in \mathcal{B}(d)$  and (3.15) implies  $\|\tilde{D} - D^{-1}\|_2 \lesssim \varepsilon$ . Lemma 3.13 shows that  $D$  and  $D^{-1}$  are bounded, thus sufficiently small  $\varepsilon > 0$  guarantees the same for  $\tilde{D}$  and  $\tilde{D}^{-1}$ . Since  $\tilde{D}$  is block diagonal, this implies (3.1) for  $\tilde{D}$ . Hence, Lemma 3.11 ensures that there exists block diagonal  $D_\varepsilon \in \mathcal{B}(d)$  (with bandwidth depending only on  $\varepsilon > 0$ ), such that  $\|\tilde{D}^{-1} - D_\varepsilon\|_2 \leq \varepsilon$ . From this, we obtain

$$\begin{aligned} \|D - D_\varepsilon\|_2 &\leq \|\tilde{D}^{-1} - D_\varepsilon\|_2 + \|\tilde{D}^{-1} - D\|_2 \leq \varepsilon + \|\tilde{D}^{-1}\|_2 \|D\|_2 \|\tilde{D} - D^{-1}\|_2 \\ &\lesssim (1 + \|\tilde{D}^{-1}\|_2 \|D\|_2) \varepsilon. \end{aligned}$$

For sufficiently small  $\varepsilon > 0$ ,  $\|\tilde{D}^{-1}\|_2$  is bounded in terms of  $\|D\|_2$ . This ensures that the constant above does not depend on  $\varepsilon$  and thus concludes the proof.  $\square$

LEMMA 3.15. *Under the assumptions of Lemmas 3.13–3.14, there exists  $b \in \mathbb{N}$  and block-lower triangular  $L_\varepsilon^{-1} \in \mathcal{B}(d, b)$  with  $L_\varepsilon^{-1}(i, i) = I$ , which is additionally block-banded in the sense  $L_\varepsilon^{-1}(i, j) = 0$  for  $|i - j| > b$  such that*

$$\|L^{-1} - L_\varepsilon^{-1}\|_2 \leq \varepsilon.$$

*The approximation  $L_\varepsilon^{-1}$  is invertible such that  $\sup_{\varepsilon > 0} (\|L_\varepsilon\|_2 + \|L_\varepsilon^{-1}\|_2) < \infty$ .*

*Proof.* Recall that  $M^T$  satisfies all the assumptions of Lemmas 3.13–3.14. Let  $M^T = \tilde{L}\tilde{U}$ . We apply Lemma 3.14 to  $M^T$  to obtain an approximation  $\tilde{U}_\varepsilon^{-1} \in \mathcal{B}(d, b)$ , block-banded with bandwidth  $b$ , bounded with bounded inverse (uniformly in  $\varepsilon$ ) such that

$$\|\tilde{U}^{-1} - \tilde{U}_\varepsilon^{-1}\|_2 \leq \varepsilon.$$

The identity (3.9) shows  $L^{-1} = D\tilde{U}^{-T}$  and thus motivates the definition

$$L_\varepsilon^{-1}(i, j) := \begin{cases} (D_\varepsilon \tilde{U}_\varepsilon^{-T})(i, j), & i \neq j, \\ I, & i = j \end{cases}$$

with  $D_\varepsilon \in \mathcal{B}(d, b)$  from Lemma 3.14 applied to  $M$ . Lemma 3.6 shows  $L_\varepsilon^{-1} \in \mathcal{B}(d)$  and  $L_\varepsilon^{-1}$  is also block-banded with bandwidth  $b$ . We obtain with the approximation estimates from Lemma 3.13

$$\|L_\varepsilon^{-1} - L^{-1}\|_2 \leq \|D_\varepsilon \tilde{U}_\varepsilon^{-T} - D\tilde{U}^{-T}\|_2 + \sup_{i \in \mathbb{N}} \|(D_\varepsilon \tilde{U}_\varepsilon^{-T})(i, i) - I\|_2.$$

The first term on the right-hand side is bounded by

$$\begin{aligned} \|D_\varepsilon \tilde{U}_\varepsilon^{-T} - D\tilde{U}^{-T}\|_2 &\leq \|(D_\varepsilon - D)\tilde{U}_\varepsilon^{-T}\|_2 + \|D(\tilde{U}_\varepsilon^{-T} - \tilde{U}^{-T})\|_2 \\ &\leq \varepsilon(\|\tilde{U}^{-1}\|_2 + \varepsilon) + \|D\|_2 \varepsilon \lesssim \varepsilon. \end{aligned}$$

The second term satisfies

$$\sup_{i \in \mathbb{N}} \|(D_\varepsilon \tilde{U}_\varepsilon^{-T})(i, i) - I\|_2 \leq \|D_\varepsilon \tilde{U}_\varepsilon^{-T} - L\|_2 = \|D_\varepsilon \tilde{U}_\varepsilon^{-T} - D\tilde{U}^{-T}\|_2 \lesssim \varepsilon.$$

Choosing  $\varepsilon > 0$  sufficiently small ensures that  $L_\varepsilon^{-1}$  is invertible with bounded inverse uniformly in  $\varepsilon > 0$ . This concludes the proof.  $\square$

We are now in position to prove the main theorem of this section.

*Proof of Theorem 3.10.* To avoid confusion, we denote the  $LU$ -factorization of  $M$  from Lemma 3.13 by  $\tilde{L}$  and  $\tilde{U}$  with diagonal matrix  $\tilde{D}$ . With Lemmas 3.14–3.15, we set  $D := D_\varepsilon$  and  $L^{-1} := L_\varepsilon^{-1}$ . This ensures  $L^{-1}, D \in \mathcal{B}(d)$ , and that  $L^{-1}$  is block-banded. Moreover,  $D$  is bounded and satisfies (3.1). This motivates the definition

$$U^{-1}(i, j) := \begin{cases} (U_\varepsilon^{-1} D_\varepsilon)(i, j), & i \neq j, \\ I, & i = j. \end{cases}$$

Lemma 3.6 shows that  $U^{-1} \in \mathcal{B}(d)$  with bandwidth depending on  $\varepsilon$  and moreover  $U^{-1}$  is block-banded. We obtain

$$\begin{aligned} \|M^{-1} - U^{-1}D^{-1}L^{-1}\|_2 &\leq \|M^{-1} - U_\varepsilon^{-1}L_\varepsilon^{-1}\|_2 + \|(U_\varepsilon^{-1} - U^{-1}D_\varepsilon^{-1})L_\varepsilon^{-1}\|_2 \\ &\leq \|M^{-1} - U_\varepsilon^{-1}L_\varepsilon^{-1}\|_2 + \sup_{i \in \mathbb{N}} \|(U_\varepsilon^{-1}D_\varepsilon)(i, i) - I\|_2 \|L_\varepsilon^{-1}\|_2 \\ &\leq \|M^{-1} - U_\varepsilon^{-1}L_\varepsilon^{-1}\|_2 + \|U_\varepsilon^{-1}D_\varepsilon - \tilde{U}^{-1}\tilde{D}\|_2 \|L_\varepsilon^{-1}\|_2. \end{aligned}$$

The first term on the right-hand side can be bounded by use of Lemmas 3.14–3.15 by

$$\|M^{-1} - U_\varepsilon^{-1} L_\varepsilon^{-1}\|_2 \leq \|\tilde{U}^{-1}\|_2 \|\tilde{L}^{-1} - L_\varepsilon^{-1}\|_2 + \|\tilde{U}^{-1} - U_\varepsilon^{-1}\|_2 \|L_\varepsilon^{-1}\|_2 \lesssim \varepsilon,$$

where the hidden constant does not depend on  $\varepsilon > 0$ . The second term can be bounded in a similar fashion by

$$\begin{aligned} & \|U_\varepsilon^{-1} D_\varepsilon - \tilde{U}^{-1} \tilde{D}\|_2 \|L_\varepsilon^{-1}\|_2 \\ & \leq \|U_\varepsilon^{-1}\|_2 \|D_\varepsilon - \tilde{D}\|_2 \|L_\varepsilon^{-1}\|_2 + \|U_\varepsilon^{-1} - \tilde{U}^{-1}\|_2 \|\tilde{D}\|_2 \|L_\varepsilon^{-1}\|_2 \lesssim \varepsilon \end{aligned}$$

with an  $\varepsilon$ -independent hidden constant. Altogether we proved

$$\|M^{-1} - U^{-1} D^{-1} L^{-1}\|_2 \lesssim \varepsilon.$$

Moreover, there holds

$$\begin{aligned} \|M - LDU\|_2 & \leq \|M\|_2 \|LDU\|_2 \|M^{-1} - U^{-1} D^{-1} L^{-1}\|_2 \\ & \leq \|M\|_2 (\|M\|_2 + \|LDU - M\|_2) \|M^{-1} - U^{-1} D^{-1} L^{-1}\|_2 \\ & \lesssim \|M\|_2^2 \varepsilon + \|M\|_2 \|M - LDU\|_2 \varepsilon. \end{aligned}$$

Sufficiently small  $\varepsilon > 0$  shows

$$\|M - LDU\|_2 \lesssim \varepsilon,$$

where the hidden constant does not depend on  $\varepsilon$ . This concludes the proof.  $\square$

**4. Riesz bases.** This section constructs suitable Riesz bases of  $H_0^1(\Omega)$  and  $L_*^2(\Omega)$  for the velocity in Theorem 4.5 and for the pressure in Theorem 4.7 of the Stokes problem. To that end, we define an auxiliary sequence of uniform meshes.

**DEFINITION 4.1.** *We consider an auxiliary sequence  $(\hat{\mathcal{T}}_\ell)_{\ell \in \mathbb{N}}$  of uniform refinements such that  $\hat{\mathcal{T}}_0 = \mathcal{T}_0$  and*

$$\hat{\mathcal{T}}_{\ell+1} = \text{bisec5}^k(\hat{\mathcal{T}}_\ell, \hat{\mathcal{T}}_\ell),$$

*which means that each element of  $\hat{\mathcal{T}}_\ell$  is refined  $k$ -times with bisec5 to obtain  $\hat{\mathcal{T}}_{\ell+1}$ . There exist constants  $C_{\text{base}}, C_{\text{mesh}} > 1$  which depend on  $k$  and on  $\mathcal{T}_0$  such that*

$$C_{\text{base}}^{-1} C_{\text{mesh}}^{-\ell} \leq \text{diam}(T) \leq C_{\text{base}} C_{\text{mesh}}^{-\ell}$$

*for all  $T \in \hat{\mathcal{T}}_\ell$  and all  $\ell \in \mathbb{N}$ . We choose  $k$  sufficiently large such that  $C_{\text{mesh}} \geq (C_{\text{sz}} + 1)^4$ , where  $C_{\text{sz}}$  is defined in Lemmas 7.2 and 7.4 below.*

Moreover, we use the fact that  $(H^{3/4}(\Omega), H^1(\Omega), H^{5/4}(\Omega))$  form a Gelfand triple with  $H^1(\Omega)$  as its pivot space as is shown in the following lemma.

**LEMMA 4.2** (from [20]). *For  $0 < s < 1/2$ , the interpolation spaces  $H^{1-s}(\Omega)$  and  $H^{1+s}(\Omega)$  form a Gelfand triple in the sense  $H^{1+s}(\Omega) \subset H^1(\Omega) \subset H^{1-s}(\Omega)$ . This means that functionals of the form  $H^{1+s}(\Omega) \rightarrow \mathbb{R}$ ,  $v \mapsto (v, w)_{H^1(\Omega)}$ ,  $w \in H^1(\Omega)$  can be identified with a dense subspace of  $H^{1-s}(\Omega)$ .*

To construct the Riesz basis, we define a natural hierarchical basis on the spaces  $\mathcal{S}^p(\mathcal{T})$ .

**DEFINITION 4.3.** Given a triangulation  $\mathcal{T}$ , define the hat functions  $v_z \in \mathcal{S}^1(\mathcal{T})$  associated with a certain node  $z$  of  $\mathcal{T}$ . For an edge  $E$  of  $\mathcal{T}$  with endpoints  $z_1$  and  $z_2$ , define the edge bubble  $v_E := \alpha_E v_{z_1} v_{z_2}$  with  $\alpha_E > 0$  such that  $\|v_E\|_{L^\infty(\Omega)} = 1$ . Let  $\mathcal{S}_B^p(\mathcal{T})$  for  $p \in \{1, 2\}$  denote the set of all hat-, respectively, hat and edge bubble-functions defined on  $\mathcal{T}$  and let  $\mathcal{S}^p(\mathcal{T})$  define the linear span of  $\mathcal{S}_B^p(\mathcal{T})$ . For a refinement  $\widehat{\mathcal{T}}$  of  $\mathcal{T}$ , we denote by  $\mathcal{S}_B^p(\widehat{\mathcal{T}} \setminus \mathcal{T})$  all hat functions  $v_z$  associated with new nodes  $z \in \mathcal{N}(\widehat{\mathcal{T}}) \setminus \mathcal{N}(\mathcal{T})$  and (for  $p = 2$ ) also all edge bubble functions  $v_E$  associated with new edges  $E \in \mathcal{E}(\widehat{\mathcal{T}}) \setminus \mathcal{E}(\mathcal{T})$ . By  $\mathcal{S}_{B,0}^p(\mathcal{T})$  and  $\mathcal{S}_{B,0}^p(\widehat{\mathcal{T}} \setminus \mathcal{T})$ , we denote the bases which vanish on  $\partial\Omega$ .

The following theorem establishes the Riesz basis for  $H_0^1(\Omega)$  which contains the velocity variable of the Stokes problem.

**DEFINITION 4.4.** We define the basis  $\tilde{B}_0^1$  as an arbitrary basis of  $\mathcal{S}_0^2(\widehat{\mathcal{T}}_0)$ . For  $\ell \geq 1$ , define

$$\tilde{B}_\ell^1 := \left\{ \frac{v}{\|v\|_{H^1(\Omega)}} : v \in \left( \bigcup_{k \in \mathbb{N}} \mathcal{S}_{B,0}^2(\mathcal{T}_k \setminus \mathcal{T}_{k-1}) \right) \cap (\mathcal{S}^2(\widehat{\mathcal{T}}_\ell) \setminus \mathcal{S}^2(\widehat{\mathcal{T}}_{\ell-1})) \right\}.$$

To prove the following theorem, we aim to employ the multiscale decomposition result by Dahmen [17]. To that end, we require projection operators  $S_\ell^2: H^1(\Omega) \rightarrow \mathcal{S}^2(\widehat{\mathcal{T}}_\ell)$  which are uniformly  $H^1(\Omega)$ -bounded and satisfy  $S_\ell^2 S_k^2 = S_\ell^2$  for all  $\ell \leq k$ . Moreover, the following approximation estimates

$$\begin{aligned} \|(1 - S_\ell^2)u\|_{H^{3/4}(\Omega)} &\lesssim C_{\text{mesh}}^{-\ell/4} \|u\|_{H^1(\Omega)}, \\ \|(1 - S_\ell^2)u\|_{H^1(\Omega)} &\lesssim C_{\text{mesh}}^{-\ell/4} \|u\|_{H^{5/4}(\Omega)}, \end{aligned}$$

need to hold as well as uniform boundedness  $S_\ell^2: H^{3/4}(\Omega) \rightarrow H^{3/4}(\Omega)$ . An obvious choice for such an operator would be the Scott–Zhang projection from [39] (see Lemma 7.4 for details). However, the standard operator does not satisfy the required symmetry  $S_\ell^2 S_k^2 = S_\ell^2$  for  $\ell \leq k$ . Therefore, we introduce a modified Scott–Zhang operator  $S_\ell^2$  in Theorem 7.5 below which satisfies all these properties, and additionally coincides with standard Scott–Zhang operator  $J_\ell^2: H^1(\Omega) \rightarrow \mathcal{S}^2(\widehat{\mathcal{T}}_\ell)$  on the set  $\tilde{B}_{\ell+1}^1$  from Definition 4.4. This allows us to use the standard operators  $J_\ell^2$  for the construction of the sets  $B_\ell^1$  below and to use the modified operators  $S_\ell^2$  in the proof of the theorem.

**THEOREM 4.5.** With the spaces from Definition 4.4 and the Scott–Zhang operator  $J_\ell^2: H^1(\Omega) \rightarrow \mathcal{S}^2(\widehat{\mathcal{T}}_\ell)$  from Lemma 7.4, define  $B_0^1 := \tilde{B}_0^1$  and for  $\ell \geq 1$

$$B_\ell^1 := \{(1 - J_{\ell-1}^2)v_0 : v_0 \in \tilde{B}_\ell^1\}.$$

Define  $B^1 := \bigcup_{\ell \in \mathbb{N}} B_\ell^1$ . Then,  $B^1$  is a Riesz bases of  $\overline{\bigcup_{k \in \mathbb{N}} \mathcal{S}_0^2(\mathcal{T}_k)} \subseteq H_0^1(\Omega)$ , i.e.,

$$(4.1) \quad \left\| \sum_{v \in B^1} \alpha_v v \right\|_{H^1(\Omega)} \simeq \left( \sum_{v \in B^1} \alpha_v^2 \right)^{1/2}.$$

Moreover,  $\text{diam}(\text{supp}(v)) \simeq C_{\text{mesh}}^{-\ell}$  for all  $v \in B_\ell^1$  and there holds

$$(4.2) \quad \|v\|_{H^s(\text{supp}(v))} \simeq C_{\text{mesh}}^{-s\ell} \|v\|_{L^2(\text{supp}(v))} \quad \text{for all } v \in B_\ell^1 \text{ and all } 0 \leq s < 3/2.$$

*Proof.* Due to the range of  $S_\ell^2$ , standard inverse estimates prove

$$\begin{aligned}\|S_\ell^2 u\|_{H^{5/4}(\Omega)} &\lesssim C_{\text{mesh}}^{\ell/4} \|S_\ell^2 u\|_{H^1(\Omega)}, \\ \|S_\ell^2 u\|_{H^1(\Omega)} &\lesssim C_{\text{mesh}}^{\ell/4} \|S_\ell^2 u\|_{H^{3/4}(\Omega)},\end{aligned}$$

and obviously the ranges  $\mathcal{S}^2(\widehat{\mathcal{T}}_\ell)$  form a dense and nested sequence of subspaces of  $H^1(\Omega)$ . Lemma 4.2 confirms that  $H^{3/4}(\Omega)$  is the dual space of  $H^{5/4}(\Omega)$  with respect to the  $H^1(\Omega)$ -scalar product. Hence, we are in the position to apply [17, Theorems 3.1 and 3.2] to prove

$$(4.3) \quad \|u\|_{H^1(\Omega)}^2 \simeq \sum_{\ell=0}^{\infty} \|(S_\ell^2 - S_{\ell-1}^2)u\|_{H^1(\Omega)}^2,$$

where we define  $S_{-1}^2 := 0$ . The identity (7.22) implies that for  $v_0 \in \tilde{B}_k^1$  there holds (with  $J_k^2$  from Theorem 7.5 below)

$$(1 - J_{k-1}^2)v_0 = (1 - S_{k-1}^2)v_0$$

and Theorem 7.5 shows

$$\begin{aligned}(S_\ell^2 - S_{\ell-1}^2)(1 - S_{k-1}^2)v_0 &= S_\ell^2 v_0 - S_\ell^2 S_{k-1}^2 v_0 - S_{\ell-1}^2 v_0 + S_{\ell-1}^2 S_{k-1}^2 v_0 \\ &= \begin{cases} S_\ell^2 v_0 - S_{\ell-1}^2 v_0 = 0, & k < \ell, \\ (1 - S_{\ell-1}^2)v_0, & k = \ell, \\ 0, & k > \ell. \end{cases}\end{aligned}$$

Thus, writing  $w = \sum_{\ell \in \mathbb{N}_0} \sum_{v \in B_\ell^1} \alpha_v v$ , we get with (4.3)

$$(4.4) \quad \|w\|_{H^1(\Omega)}^2 \simeq \sum_{\ell \in \mathbb{N}_0} \|(S_\ell^2 - S_{\ell-1}^2)w\|_{H^1(\Omega)}^2 \simeq \sum_{\ell \in \mathbb{N}_0} \left\| \sum_{v \in B_\ell^1} \alpha_v v \right\|_{H^1(\Omega)}^2.$$

For  $v \in B_\ell^1$ , there exists a unique  $v_0 \in \tilde{B}_\ell^1$  such that  $v = (1 - J_{\ell-1}^2)v_0$ . For  $T \in \widehat{\mathcal{T}}_{\ell-1}$ , scaling arguments and stability of  $J_{\ell-1}$  show  $\|\cdot\|_{H^1(T)} \lesssim \|(1 - J_{\ell-1}^2)(\cdot)\|_{H^1(T)} \lesssim \|\cdot\|_{H^1(\cup \omega(T, \widehat{\mathcal{T}}_{\ell-1}))}$  on the finite dimensional space  $\text{span}\{v|_{\cup \omega(T, \widehat{\mathcal{T}}_{\ell-1})} : v \in \tilde{B}_\ell^1\}$  (with uniform constants). This implies

$$\|w\|_{H^1(\Omega)}^2 \simeq \sum_{\ell \in \mathbb{N}} \left\| (1 - J_{\ell-1}^2) \sum_{v \in B_\ell^1} \alpha_v v_0 \right\|_{H^1(\Omega)}^2 \simeq \sum_{\ell \in \mathbb{N}} \left\| \sum_{v \in B_\ell^1} \alpha_v v_0 \right\|_{H^1(\Omega)}^2.$$

Since for all  $T \in \widehat{\mathcal{T}}_{\ell-1}$ , the set  $\{v_0|_T : v_0 \in \tilde{B}_\ell^1\}$  is linear independent, a scaling argument shows

$$(4.5) \quad \left\| \sum_{v \in B_\ell^1} \alpha_v v_0 \right\|_{H^1(T)}^2 \simeq \sum_{\substack{v \in B_\ell^1 \\ v_0|_T \neq 0}} \alpha_v^2.$$

Another norm equivalence argument shows  $1 = \|v_0\|_{H^1(\Omega)} \simeq \|(1 - J_{\ell-1}^2)v_0\|_{H^1(\Omega)}$  for all  $v_0 \in \tilde{B}_\ell^1$  and hence

$$\|w\|_{H^1(\Omega)}^2 \simeq \sum_{\ell \in \mathbb{N}} \sum_{v \in B_\ell^1} \alpha_v^2 \|v\|_{H^1(\text{supp}(v))}^2.$$

Since  $B_0^1$  contains only finitely many linear independent functions, we also obtain

$$\|w\|_{H^1(\Omega)}^2 \simeq \sum_{\ell \in \mathbb{N}_0} \sum_{v \in B_\ell^1} \alpha_v^2 \|v\|_{H^1(\text{supp}(v))}^2 \simeq \sum_{v \in B^1} \alpha_v^2.$$

Altogether, the operator  $\iota: \ell_2(B^1) \rightarrow H^1(\Omega)$ ,  $\iota(\alpha) := \sum_{v \in B^1} \alpha_v v$  is bounded and has a bounded inverse on its closed range. Since  $J_\ell^2$  retains homogeneous boundary values (Lemma 7.4), the range of  $\iota$  is dense in  $\overline{\bigcup_{k \in \mathbb{N}} \mathcal{S}_0^2(\mathcal{T}_k)} \subseteq H^1(\Omega)$  and hence  $\iota$  is bijective. This shows that  $B^1$  is a Riesz basis of  $\overline{\bigcup_{k \in \mathbb{N}} \mathcal{S}_0^2(\mathcal{T}_k)} \subseteq H_0^1(\Omega)$ .

The scaling estimate (4.2) can be proved as follows: Let  $v = (1 - J_{\ell-1}^2)v_0 \in B_\ell^1$  for some  $v_0 \in \tilde{B}_\ell^1$  and let  $\omega := \text{supp}(v)$ . The approximation property and the projection property of  $J_{\ell-1}^2$  show

$$\|v\|_{L^2(\omega)} \lesssim C_{\text{mesh}}^{-s\ell} \|v\|_{H^s(\omega)}$$

for all  $0 \leq s < 3/2$ . The converse estimates  $\|v\|_{L^2(\omega)} \gtrsim C_{\text{mesh}}^{-s\ell} \|v\|_{H^s(\omega)}$  for  $0 \leq s < 3/2$  follow from standard inverse estimates. This concludes (4.2) and thus concludes the proof.  $\square$

The following definition and theorem construct a Riesz basis for the pressure variable of the Stokes problem in  $L_*^2(\Omega)$ .

**DEFINITION 4.6.** *We define the basis  $\tilde{B}_0^0$  as an arbitrary basis of  $\mathcal{S}_*^1(\widehat{\mathcal{T}}_0)$ . For  $\ell \geq 1$ , define*

$$\tilde{B}_\ell^0 := \left\{ \frac{v}{\|v\|_{L^2(\Omega)}} : v \in \left( \bigcup_{k \in \mathbb{N}} \mathcal{S}_B^1(\mathcal{T}_k \setminus \mathcal{T}_{k-1}) \right) \cap (\mathcal{S}^1(\widehat{\mathcal{T}}_\ell) \setminus \mathcal{S}^1(\widehat{\mathcal{T}}_{\ell-1})) \right\}.$$

**THEOREM 4.7.** *With the spaces from Definition 4.3 and the modified Scott–Zhang operator from Definition 7.1,  $J_{\ell-1}^1$ , define  $B_0^0 = \tilde{B}_0^0$  and for  $\ell \geq 1$*

$$B_\ell^0 := \{(1 - J_{\ell-1}^1)v_0 : v_0 \in \tilde{B}_\ell^0\}.$$

*Define  $B^0 := \bigcup_{\ell \in \mathbb{N}} B_\ell^0$ . Then,  $B^0$  is the Riesz basis of  $\overline{\bigcup_{k \in \mathbb{N}} \mathcal{S}_*^1(\mathcal{T}_k)} \subseteq L_*^2(\Omega)$ , i.e.,*

$$(4.6) \quad \left\| \sum_{v \in B^0} \alpha_v v \right\|_{L^2(\Omega)} \simeq \left( \sum_{v \in B^0} \alpha_v^2 \right)^{1/2}.$$

*Moreover,  $\text{diam}(\text{supp}(v)) \simeq C_{\text{mesh}}^{-\ell}$  for all  $v \in B_\ell^0$  and there holds for all  $v \in B_\ell^0$*

$$(4.7) \quad \begin{aligned} \|v\|_{H^s(\text{supp}(v))} &\simeq C_{\text{mesh}}^{-s\ell} \|v\|_{L^2(\text{supp}(v))} && \text{for all } 0 \leq s < 3/2, \\ \|v\|_{\tilde{H}^{-s}(\text{supp}(v))} &\simeq C_{\text{mesh}}^{s\ell} \|v\|_{L^2(\text{supp}(v))} && \text{for all } 0 \leq s \leq 1. \end{aligned}$$

*Proof.* The idea of the proof is identical to that of Theorem 4.5; we would like to use a Scott–Zhang projector but require some additional properties provided by the projection from Theorem 7.3. First, we note that stability of the  $J_{\ell-1}^1$  and scaling argument prove the estimate  $\|v\|_{L^2(\Omega)} \simeq 1$  uniformly for all  $v \in B^0$ . Moreover, Lemma 7.2 shows immediately that  $(1 - J_{\ell-1}^1)v_0 \in L_*^2(\Omega)$ . Again, we use [17] with the operators  $(S_\ell^1)_{\ell \in \mathbb{N}_0}$  from Theorem 7.3: The operators  $S_\ell^1$  are uniformly  $L^2(\Omega)$ -bounded and satisfy for all  $\ell \leq k$

$$(4.8) \quad S_\ell^1 S_k^1 = S_\ell^1.$$

Moreover, their ranges  $\mathcal{S}^1(\widehat{\mathcal{T}}_\ell)$  form a dense and nested sequence of subspaces of  $L^2(\Omega)$ . Theorem 7.3 confirms the approximation estimates

$$\begin{aligned}\|(1 - S_\ell^1)u\|_{\tilde{H}^{-1/4}(\Omega)} &\lesssim C_{\text{mesh}}^{-\ell/4} \|u\|_{L^2(\Omega)}, \\ \|(1 - S_\ell^1)u\|_{L^2(\Omega)} &\lesssim C_{\text{mesh}}^{-\ell/4} \|u\|_{H^{1/4}(\Omega)}\end{aligned}$$

as well as uniform boundedness  $S_\ell^1: \tilde{H}^{-1/4}(\Omega) \rightarrow \tilde{H}^{-1/4}(\Omega)$ . Standard inverse estimates prove

$$\begin{aligned}\|S_\ell^1 u\|_{H^{1/4}(\Omega)} &\lesssim C_{\text{mesh}}^{\ell/4} \|S_\ell^1 u\|_{L^2(\Omega)}, \\ \|S_\ell^1 u\|_{L^2(\Omega)} &\lesssim C_{\text{mesh}}^{\ell/4} \|S_\ell^1 u\|_{\tilde{H}^{-1/4}(\Omega)}.\end{aligned}$$

Therefore, we may apply [17, Theorems 3.1 and 3.2] to prove

$$(4.9) \quad \|u\|_{L^2(\Omega)}^2 \simeq \sum_{\ell=0}^{\infty} \|(S_\ell^1 - S_{\ell-1}^1)u\|_{H^1(\Omega)}^2,$$

where we define  $S_{-1}^1 := 0$ . The identity (7.12) implies that for  $v_0 \in \tilde{B}_\ell^0$  for  $\ell \geq 1$ , there holds

$$J_k^1 v_0 = S_k^1 v_0 \quad \text{for all } k \geq \ell - 1.$$

The identity (4.8) shows for  $v_0 \in \tilde{B}_k^0$

$$\begin{aligned}(S_\ell^1 - S_{\ell-1}^1)(1 - S_{k-1}^1)v_0 &= S_\ell^1 v_0 - S_\ell^1 S_{k-1}^1 v_0 - S_{\ell-1}^1 v_0 + S_{\ell-1}^1 S_{k-1}^1 v_0 \\ &= \begin{cases} S_\ell^1 v_0 - S_{\ell-1}^1 v_0 = 0, & k < \ell, \\ (1 - S_{\ell-1}^1)v_0 = (1 - J_{\ell-1}^1)v_0, & k = \ell, \\ 0, & k > \ell. \end{cases}\end{aligned}$$

As in the proof of Theorem 4.5, we obtain

$$\|w\|_{L^2(\Omega)}^2 \simeq \sum_{\ell \in \mathbb{N}_0} \sum_{v \in B_\ell^0} \alpha_v^2 \|v\|_{L^2(\text{supp}(v))}^2$$

for all  $w = \sum_{\ell \in \mathbb{N}_0} \sum_{v \in B_\ell^0} \alpha_v v$ . Therefore, the operator  $\iota: \ell_2(B^0) \rightarrow L_*^2(\Omega)$ ,  $\iota(\alpha) := \sum_{v \in B^0} \alpha_v v$  is bounded and has a bounded inverse on its closed range. Obviously, the range is dense in  $\overline{\bigcup_{k \in \mathbb{N}} \mathcal{S}_*^1(\mathcal{T}_k)} \subseteq L^2(\Omega)$  and hence  $\iota: \ell_2(B^0) \rightarrow \overline{\bigcup_{k \in \mathbb{N}} \mathcal{S}_*^1(\mathcal{T}_k)}$  is bijective. This concludes that  $B^0$  is a Riesz basis of  $\overline{\bigcup_{k \in \mathbb{N}} \mathcal{S}_*^1(\mathcal{T}_k)} \subseteq L_*^2(\Omega)$ .

The scaling estimates (4.7) can be proved as follows: Let  $v = (1 - J_{\ell-2}^1)v_0 \in B_\ell^0$  for some  $v_0 \in \tilde{B}_\ell^0$  and let  $\omega := \text{supp}(v)$ . By construction of  $J_{\ell-1}^1$ , we see that  $v = (1 - J_{\ell-1}^1)v_0$  has integral mean zero on its support. Therefore, a Poincaré estimate proves

$$\|v\|_{\tilde{H}^{-s}(\omega)} \lesssim C_{\text{mesh}}^{-s\ell} \|v\|_{L^2(\omega)}$$

for all  $0 \leq s \leq 1$ . The approximation property (7.3) and the projection property of  $J_{\ell-1}^1$  show

$$\|v\|_{L^2(\omega)} \lesssim C_{\text{mesh}}^{-s\ell} \|v\|_{H^s(\omega)}$$

for all  $0 \leq s < 3/2$ . The converse estimates  $\|w\|_{L^2(\omega)} \gtrsim C_{\text{mesh}}^{-s\ell} \|w\|_{H^s(\omega)}$  for  $0 \leq s <$

$3/2$  as well as  $\|v\|_{\tilde{H}^{-s}(\omega)} \gtrsim C_{\text{mesh}}^{-s\ell} \|v\|_{L^2(\omega)}$  for  $0 \leq s \leq 1$  follow from standard inverse estimates. This concludes the proof.  $\square$

LEMMA 4.8. *Let  $\mathcal{T} \in \mathbb{T}_{\text{grad}}$  and  $\ell \in \mathbb{N}$ . Let  $T \in \mathcal{T} \setminus \widehat{\mathcal{T}}_\ell$  such that  $\mathcal{T}|_T$  is a strictly finer local refinement of  $\widehat{\mathcal{T}}_\ell$ . Then,  $\mathcal{T}|_{\omega^{D_{\text{grad}}}(T, \widehat{\mathcal{T}}_\ell)}$  is a local refinement of  $\widehat{\mathcal{T}}_\ell$ .*

*Proof.* Let  $L \in \mathbb{N}$  denote the level of the elements in  $\widehat{\mathcal{T}}_\ell$ . By assumption, there holds  $\text{level}(T) > L$  and hence the assertion is a direct consequence of (2.8).  $\square$

LEMMA 4.9. *Given a mesh  $\mathcal{T}$  with (2.8) for  $D_{\text{grad}} \geq 3$ , there holds*

$$\mathcal{S}_0^2(\mathcal{T}) = \text{span}(B^1 \cap \mathcal{S}^2(\mathcal{T})) \quad \text{and} \quad \mathcal{S}_\star^1(\mathcal{T}) = \text{span}(B^0 \cap \mathcal{S}_\star^1(\mathcal{T})).$$

*Proof.* We note that  $\mathcal{S}_0^2(\mathcal{T}) = \text{span}\{v_0 \in \bigcup_{\ell=0}^\infty \widetilde{B}_\ell^1 : v_0 \in \mathcal{S}_0^2(\mathcal{T})\}$ . For each  $v_0 \in \mathcal{S}_0^2(\mathcal{T}) \cap \widetilde{B}_\ell^1$ , we note that  $J_{\ell-1}v_0$  is supported on  $\omega(\text{supp}(v_0), \widehat{\mathcal{T}}_{\ell-1})$ . Since  $v_0 \in \mathcal{S}_0^2(\mathcal{T}) \cap \widetilde{B}_\ell^1$  implies that  $\mathcal{T}|_{\text{supp}(v_0)}$  is a true local refinement of  $\widehat{\mathcal{T}}_{\ell-1}$ , there exists at least one  $T \in \mathcal{T} \setminus \widehat{\mathcal{T}}_{\ell-1}$  with  $T \subseteq \text{supp}(v_0)$ . By definition of  $v = (1 - J_{\ell-1}^2)v_0$ , we know that  $\text{supp}(v) \subseteq \bigcup \omega^2(T, \widehat{\mathcal{T}}_{\ell-1})$ . Lemma 4.8 proves that  $\mathcal{T}|_{\omega^2(T, \widehat{\mathcal{T}}_{\ell-1})}$  is a local refinement of  $\widehat{\mathcal{T}}_{\ell-1}$  and hence  $v \in \mathcal{S}_0^2(\mathcal{T})$ . This concludes  $\mathcal{S}_0^2(\mathcal{T}) = \text{span}\{v \in B^1 : v \in \mathcal{S}_0^2(\mathcal{T})\}$ .

The same argument works for  $\mathcal{S}_\star^1(\mathcal{T})$  and concludes the proof.  $\square$

**5. Proof of (A1)–(A4).** The abstract theory developed in the previous sections allows us to prove optimality of the adaptive algorithm for the stationary Stokes problem.

**5.1. Proof of (A1), (A2), and (A4).** The proofs of (A1), (A2), and (A4) will not be surprising to experts as they are well-known in the literature and can be found in [26].

*Proof of (A1) and (A2).* The proof follows from standard arguments as for the Poisson problem (see, e.g., [13]) and can be found in condensed form in [26, Lemma 3.2].  $\square$

*Proof of (A4).* The property (A4) can be found in [26, Lemma 3.1].  $\square$

**5.2. Proof of (A3).** The main innovation of this paper is the proof of general quasi-orthogonality (A3). The strategy of proof for this section is the following:

- Approximate the infinite stiffness matrix of  $a(\cdot, \cdot)$  with respect to the basis from section 4 by a banded matrix.
- Use the tools from section 3.2.1 to show that the banded matrix has an inversely bounded  $LU$ -factorization.
- Use Theorem 3.2 to show that general quasi-orthogonality (A3) holds.

To that end, we first need to show that we can restrict the problem to the space  $\tilde{\mathcal{X}} := \overline{\bigcup_{\ell \in \mathbb{N}_0} \mathcal{X}_\ell} \subseteq \mathcal{X}$ . To that end, the result [37, Lemma 4.2] shows that  $a(\cdot, \cdot)$  satisfies the inf-sup condition (2.5) even on  $\tilde{\mathcal{X}}$  and that  $(u, p) \in \tilde{\mathcal{X}}$ . To fit the problem into our abstract framework, we choose the following Riesz basis of  $\tilde{\mathcal{X}}$  from Theorems 4.5 and 4.7:

$$\mathcal{B} := \{(v, 0, 0) : v \in B^1\} \cup \{(0, v, 0) : v \in B^1\} \cup \{(0, 0, w) : w \in B^0\}.$$

We recall that  $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1} \subset \mathcal{X}$  are nested finite dimensional spaces generated by the adaptive algorithm described in section 2.2. We introduce the level function  $L(\mathbf{w}) = \ell$  for all  $\mathbf{w} \in (B_\ell^1)^2 \times B_\ell^0$  and also  $L(v) = \ell$  for all  $v \in B_\ell^1 \cup B_\ell^0$ . We order the functions

in  $\mathbf{B}$  such that  $\mathcal{X}_\ell = \text{span}\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{N_\ell}\}$  for particular  $N_\ell \in \mathbb{N}$  and all  $\ell \in \mathbb{N}$  (note that this is possible due to Lemma 4.9).

The proofs below will use several distance functions to capture the sparsity structure of the involved matrices. Those distance functions are defined in [20] and we recall them in the following.

**DEFINITION 5.1.** For  $B := B^0 \cup B^1$ , define the following functions:

- $\delta: B \times B \rightarrow \{0, 1\}$  is defined by  $\delta(v, w) = 1$  if  $v \neq w$  and  $\delta(v, w) = 0$  if  $v = w$ .
- $\delta_k: B \times B \rightarrow \mathbb{N}$  is defined by

$$\delta_k(v, w) := \min \left\{ n \in \mathbb{N} : \exists T_1, \dots, T_n \in \widehat{\mathcal{T}}_k, \text{mid}(v) \in T_1, \text{mid}(w) \in T_n, T_i \cap T_{i+1} \neq \emptyset, i = 1, \dots, n-1 \right\},$$

where  $\text{mid}(\cdot)$  denotes the barycenter of the support of the function. Note that this metric measures the distance with respect to the element size on  $\widehat{\mathcal{T}}_k$ .

- $d_1: B \times B \rightarrow \mathbb{N}$  is defined by

$$d_1(v, w) := \delta_{\min\{L(v), L(w)\}}(v, w).$$

- Given  $\beta > 0$ ,  $d_2: B \times B \rightarrow [0, \infty)$  is defined by

$$d_2(v, w) := \delta(v, w) + \beta|L(v) - L(w)| + \log(\delta(v, w) + d_1(v, w)).$$

Note that this metric combines the physical distance of the supports with the level distance. The logarithm is necessary to ensure the triangle inequality.

- Given  $\gamma > 0$ ,  $d_3: B \times B \rightarrow [0, \infty)$  is defined by

$$d_3(v, w) := \begin{cases} \gamma^{\max\{L(v), L(w)\}}, & L(v) \neq L(w), \\ \delta(v, w) + d_1(v, w) - 1, & L(v) = L(w). \end{cases}$$

Note that this metric is designed for block-diagonal matrices, where  $d_3(v, w) = \infty$  for  $L(v) \neq L(w)$ . The quantity  $\gamma^{\max\{L(v), L(w)\}}$  is just a sufficiently large placeholder to ensure the triangle inequality and avoid the use of  $\infty$ .

It is shown in [20, section 4] that for sufficiently large  $\beta, \gamma > 0$ ,  $d_2$  and  $d_3$  are metrics on  $\mathbf{B}$ . If  $\mathbf{B}$  is identified with  $\mathbb{N}$ ,  $d_3$  even satisfies (3.6).

Moreover, we will consider two different block structures (Definition 3.1). The first one results from the adaptive spaces  $\mathcal{X}_\ell$ ,  $\ell \in \mathbb{N}$ , which induce the adaptive block structure  $N_1, N_2, \dots$  as described above. The second block structure is induced by reordering  $\mathbf{B}$  such that  $L(\mathbf{w}_{\pi(i)}) \leq L(\mathbf{w}_{\pi(j)})$  for all  $i \leq j$  and some permutation  $\pi$ . This permutation defines a unique permutation matrix  $P \in \{0, 1\}^{\mathbb{N} \times \mathbb{N}}$  defined by  $P_{ij} = 1$  if and only if  $i = \pi(j)$ . This yields the uniform block structure  $n_1, n_2, \dots$  such that  $\{\mathbf{w}_{k_i} : i = n_r, \dots, n_{r+1} - 1\} = \{\mathbf{w} \in \mathbf{B} : L(\mathbf{w}) = r\}$ .

The following lemma states that under (2.8), an adaptive space  $\mathcal{X}_\ell$  which contains a basis function  $\mathbf{w} \in \mathbf{B}$  also contains lower level basis functions which are close to  $\mathbf{w}$ . For brevity of presentation of the following lemma, we identify  $\mathbf{w} \in \mathbf{B}$  with its unique nonzero component, e.g.,  $\mathbf{w} = (0, w, 0)$ .

**LEMMA 5.2.** Let  $\mathbf{v}, \mathbf{w} \in \mathbf{B}$  such that  $\mathbf{v} \in \mathcal{X}_\ell$  for some  $\ell \in \mathbb{N}_0$ . Let  $D_{\text{grad}} > 0$  and assume (2.8) for all  $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$ . Then,  $L(\mathbf{w}) < L(\mathbf{v})$  and  $\delta_{L(\mathbf{w})}(\mathbf{v}, \mathbf{w}) \leq D_{\text{grad}} - C_{\text{grad}}$  imply  $\mathbf{w} \in \mathcal{X}_\ell$ . The constant  $C_{\text{grad}} \geq 1$  depends only on  $\mathcal{T}_0$  and  $C_{\text{mesh}}$ .

*Proof.* Define

$$C := \sup_{k \in \mathbb{N}} \max_{w \in \bigcup_{j=k}^{\infty} (B_j^1 \cup B_j^0)} \#\{T \in \widehat{\mathcal{T}}_k : T \subseteq \text{supp}(w)\}.$$

By definition of  $\mathbf{B}$ , we see that  $C < \infty$  in terms of  $\mathcal{T}_0$  and  $C_{\text{mesh}}$ . Let  $\text{level}(\widehat{\mathcal{T}}_k)$  denote the uniform level of elements in  $\widehat{\mathcal{T}}_k$ ,  $k \in \mathbb{N}_0$ . We prove the statement by contradiction, i.e., we show that  $\mathbf{w} \notin \mathcal{X}_{\ell}$  and  $L(\mathbf{w}) < L(\mathbf{v})$  lead to a contradiction. To that end, find  $T, T' \in \mathcal{T}_{\ell}$  with  $T \cap \text{supp}(\mathbf{w}) \neq \emptyset$  and  $T' \cap \text{supp}(\mathbf{v}) \neq \emptyset$ . Since  $\mathbf{w} \notin \mathcal{X}_{\ell}$ , we may choose  $T$  such that  $\text{level}(T) < \text{level}(\widehat{\mathcal{T}}_{L(\mathbf{w})})$ . Assumption (2.8) implies that all  $T'' \in \omega^{D_{\text{grad}}}(T, \mathcal{T}_{\ell})$  satisfy  $\text{level}(T'') \leq \text{level}(T) + 1 \leq \text{level}(\widehat{\mathcal{T}}_{L(\mathbf{w})})$ . This means that  $\widehat{\mathcal{T}}_{L(\mathbf{w})}|_{\omega^{D_{\text{grad}}}(T, \mathcal{T}_{\ell})}$  is a local refinement of  $\mathcal{T}_{\ell}$ . With  $\chi_T$  denoting the function which is one on  $T$  and zero elsewhere, we get

$$\delta_{L(\mathbf{w})}(\chi_T, \chi_{T'}) \leq 2C + \delta_{L(\mathbf{w})}(\mathbf{v}, \mathbf{w}).$$

Thus,  $\delta_k(\mathbf{v}, \mathbf{w}) \leq D_{\text{grad}} - 2C$  implies  $\delta_{L(\mathbf{w})}(\chi_T, \chi_{T'}) \leq D_{\text{grad}}$ . Since we argued above that  $\widehat{\mathcal{T}}_{L(\mathbf{w})}|_{\omega^{D_{\text{grad}}}(T, \mathcal{T}_{\ell})}$  is a local refinement of  $\mathcal{T}_{\ell}$ ,  $\delta_{L(\mathbf{w})}(\chi_T, \chi_{T'}) \leq D_{\text{grad}}$  implies  $T' \in \omega^{D_{\text{grad}}}(T, \mathcal{T}_{\ell})$ . This shows  $\text{level}(T') \leq \text{level}(T) + 1 \leq \text{level}(\widehat{\mathcal{T}}_{L(\mathbf{w})})$  which contradicts  $\mathbf{v} \in \mathcal{X}_{\ell}$  and  $L(\mathbf{v}) > L(\mathbf{w})$ . This concludes the proof with  $C_{\text{grad}} := 2C$ .  $\square$

The next lemma shows that the adaptive and the uniform block structure are compatible with each other in the sense that being block triangular is (sometimes) invariant under permutation between the two structures.

LEMMA 5.3. *Let  $L \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  with  $L^{-1} \in \mathcal{B}(d_2, b)$  for some bandwidth  $b \in \mathbb{N}$  and let  $L$  be lower triangular with identity diagonal blocks  $L|_{\{n_r, \dots, n_{r+1}-1\} \times \{n_r, \dots, n_{r+1}-1\}} = I$  for all  $r \in \mathbb{N}$ . If all  $(\mathcal{T}_{\ell})_{\ell \in \mathbb{N}}$  satisfy (2.8) for some sufficiently large  $D_{\text{grad}}$  depending only on  $b$ , then  $PLP^T$  is block lower triangular with respect to the adaptive block structure  $N_1, N_2, \dots$ .*

*Proof.* We show the equivalent statement that  $(PLP^T)^{-1} = PL^{-1}P^T$  is block lower triangular. Under the assumptions on  $L^{-1}$ , the fact  $(L^{-1})_{\pi^{-1}(i)\pi^{-1}(j)} \neq 0$  implies either  $\pi^{-1}(i) = \pi^{-1}(j)$  and hence  $i = j$  or  $\pi^{-1}(i) \geq n_r > \pi^{-1}(j)$  for some  $r \in \mathbb{N}$  and hence  $L(\mathbf{w}_{\pi^{-1}(j)}) < L(\mathbf{w}_{\pi^{-1}(i)})$ . In the latter case,  $L^{-1} \in \mathcal{B}(d_2, b)$  shows additionally that  $d_2(\mathbf{w}_{\pi^{-1}(i)}, \mathbf{w}_{\pi^{-1}(j)}) \leq b$  and hence  $d_{L(\mathbf{w}_{\pi^{-1}(j)})}(\mathbf{w}_{\pi^{-1}(i)}, \mathbf{w}_{\pi^{-1}(j)}) \lesssim \exp(b)$ . Thus, for sufficiently large  $D_{\text{grad}}$ , Lemma 5.2 applies and shows  $\mathbf{w}_{\pi^{-1}(j)} \in \mathcal{X}_{\ell}$  for any  $\ell \in \mathbb{N}$  with  $\mathbf{w}_{\pi^{-1}(i)} \in \mathcal{X}_{\ell}$ . In other words, we showed that  $(PL^{-1}P^T)_{ij} \neq 0$  implies  $j \leq N_{\ell}$  for all  $\ell \in \mathbb{N}$  with  $i \leq N_{\ell}$  and hence  $PL^{-1}P^T$  is block lower triangular with respect to the adaptive block structure  $N_1, N_2, \dots$ . This concludes the proof.  $\square$

THEOREM 5.4. *Given  $\varepsilon > 0$  and under all previous assumptions, there exists  $D_{\text{grad}} > 0$  sufficiently large such that the solutions (2.7) of the stationary Stokes problem (2.1) satisfy general quasi-orthogonality (A3).*

*Proof.* Step 1: With the basis  $\mathbf{B}$ , define the matrix  $A \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  by

$$A_{ij} := a(\mathbf{w}_j, \mathbf{w}_i).$$

With the two different block structures defined above, we consider the permuted matrix

$$\tilde{A}_{ij} := A_{\pi(i)\pi(j)} \quad \text{or} \quad \tilde{A} = P^T AP.$$

Since  $\mathbf{B}$  is a Riesz basis, (2.5)–(2.6) of  $a(\cdot, \cdot)$  imply that  $A$  satisfies (3.1). By definition of  $\mathbf{B}$ , we observe that  $\{\mathbf{w}_{\pi(i)} : i = 1, \dots, n_{k+1} - 1\}$  spans the space

$$\left(\mathcal{S}_0^2(\widehat{\mathcal{T}}_k)^2 \times \mathcal{S}_\star^1(\widehat{\mathcal{T}}_k)\right) \cap \bigcup_{\ell \in \mathbb{N}} \left(\mathcal{S}_0^2(\mathcal{T}_\ell)^2 \times \mathcal{S}_\star^1(\mathcal{T}_\ell)\right) = \mathcal{S}_0^2(\mathcal{T})^2 \times \mathcal{S}_\star^1(\mathcal{T}),$$

where  $\mathcal{T} \in \mathbb{T}$  is the unique mesh which contains exactly the nodes and edges that appear in  $\widehat{\mathcal{T}}_k$  and additionally in at least one (and thus all the subsequent) of the adaptive meshes  $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$ . Therefore, [8] shows that (2.6) is satisfied for  $\mathcal{X}_\mathcal{T} := \mathcal{S}_0^2(\mathcal{T})^2 \times \mathcal{S}_\star^1(\mathcal{T})$  (for all  $r \in \mathbb{N}$ ) and hence (3.1) holds also for the uniform block structure on  $\tilde{A}$ .

*Step 2:* Lemmas 6.4–6.5 below show that there exists  $\tilde{A}^\varepsilon$  which is block-banded (with the uniform block structure  $n_1, n_2, \dots$ ) for some bandwidth  $b$  such that

$$\|\tilde{A} - \tilde{A}^\varepsilon\|_2 \leq \varepsilon.$$

Moreover, if we identify  $i \mapsto \mathbf{w}_i$ , there holds  $\tilde{A}^\varepsilon \in \mathcal{B}(d_2)$  with the metric  $d_2(\cdot, \cdot)$  from Definition 5.1. Choosing  $\varepsilon > 0$  sufficiently small, we ensure that  $\tilde{A}^\varepsilon$  satisfies (3.1) as well. Thus, Theorem 3.10 shows that there exists an approximate block- $LDU$ -factorization with  $\|\tilde{A} - LDU\|_2 \leq 2\varepsilon$ , i.e.,

$$(5.1) \quad \|A - (PLP^T)(PDP^T)(PUP^T)\|_2 \leq 2\varepsilon.$$

The factors  $L^{-1}, D, U^{-1} \in \mathcal{B}(d_2)$  are block-banded (with respect to the uniform block structure  $n_1, n_2, \dots$ ), bounded, and have identity diagonal blocks. Thus, Lemma 5.3 applies and shows that  $PLP^T$  and  $PUP^T$  are block triangular with respect to the adaptive block structure  $N_1, N_2, \dots$ .

*Step 3:* Since  $D$  is block diagonal and satisfies (3.1) (according to Theorem 3.10, with the uniform block structure  $n_1, n_2, \dots$ ), we also have  $PDP^T \in \mathcal{B}(d_3)$  for the metric  $d_3(\cdot, \cdot)$  from Definition 5.1 (again identifying  $i \mapsto \mathbf{w}_i$ ). This shows  $PDP^T \in \mathcal{J}(d_3)$ . Since  $A$  satisfies (3.1), the estimate (5.1) together with the fact that  $PLP^T$  and  $PUP^T$  are block triangular and inversely bounded (uniformly in  $\varepsilon$ ; see Theorem 3.10) shows that  $PDP^T$  satisfies (3.1) for the adaptive block structure  $N_1, N_2, \dots$  as long as  $\varepsilon > 0$  is sufficiently small. Thus, Theorem 3.9 implies the existence of a bounded block- $LU$ -factorization  $PDP^T = \tilde{L}\tilde{U}$  (with respect to the adaptive block structure) with inversely bounded factors  $\tilde{L}, \tilde{U} \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$ . Altogether, we found a matrix  $A^\varepsilon := P\tilde{A}^\varepsilon P^T$  with block- $LU$ -factorization  $A^\varepsilon = (PLP^T\tilde{L})(\tilde{U}PUP^T)$  with inversely bounded factors. Finally, Theorem 3.2 with  $\tilde{\mathcal{X}}$  instead of  $\mathcal{X}$  applies and concludes the proof.  $\square$

**6. Technical results and auxiliary lemmas.** This section covers some of the more technical aspects of the work.

**6.1. Almost bandedness of differential operator matrices.** In this section, we prove that the operator matrices coming from the Stokes problem in the Riesz basis are arbitrarily close to banded matrices in the sense of Definition 3.5. The results will not surprise anyone familiar with wavelet analysis but cannot be found in the literature directly as we use a very particular Riesz basis. The main proof technique is to exploit locality of the basis functions as well as strengthened Cauchy–Schwarz inequalities between the levels.

**DEFINITION 6.1.** Define the Sobolev–Slobodeckij seminorm

$$|v|_{H^s(\omega)}^2 := \int_\omega \int_\omega |v(x) - v(y)| / |x - y|^{2+2s} dx dy \quad \text{for } 0 < s < 1.$$

For  $s = \nu + r \in \mathbb{R}$  with  $\nu \in \mathbb{N}$  and  $s \in (0, 1)$ , define  $|\cdot|_{H^s(\omega)} := |\nabla^\nu(\cdot)|_{H^r(\Omega)}$ , where  $\nabla^\nu$  denotes the tensor of all partial derivatives of order  $\nu$ . As shown in [27],  $\|\cdot\|_{H^\nu(\omega)} + |\cdot|_{H^r(\omega)}$  is equivalent to the  $H^s$ -norm obtained via (real) interpolation. The norm equivalence constants depend only on the shape of  $\omega$ .

LEMMA 6.2. Let  $v \in H^s(\omega)$  for some  $0 \leq s < 1/2$  and some Lipschitz domain  $\omega \subseteq \Omega$ . Then, there holds for the extension of  $v$  to  $\Omega$  by zero

$$\|v\|_{H^s(\Omega)} \leq C_{\text{res}} \|v\|_{H^s(\omega)}.$$

The constant  $C_{\text{res}} > 0$  does depend only on the shape of  $\omega$ .

*Proof.* There holds

$$\begin{aligned} |v|_{H^s(\Omega)}^2 - |v|_{H^s(\omega)}^2 &= 2 \int_{\Omega \setminus \omega} \int_\omega \frac{|v(x) - v(y)|^2}{|x - y|^{2s+2}} dy dx \\ &= 2 \int_\omega |v(y)|^2 \int_{\Omega \setminus \omega} \frac{1}{|x - y|^{2s+2}} dx dy \\ &\lesssim \int_\omega |v(y)|^2 dy = \|v\|_{L^2(\text{supp}(v))}^2, \end{aligned}$$

where we used that, in polar coordinates centered at  $y$ , the inner integral reads

$$\int_{\Omega \setminus \omega} \frac{1}{|x - y|^{2s+2}} dx = \int_{\Omega \setminus \omega} r^{-2s} dr \lesssim 1.$$

This, along with norm equivalence discussed in Definition 6.1 concludes the proof.  $\square$

LEMMA 6.3. Let  $k \leq \ell \in \mathbb{N}$ , let  $v \in \mathcal{P}^2(\widehat{\mathcal{T}}_k)$  with  $\text{supp}(v) \subseteq \bigcup \omega^n(T, \widehat{\mathcal{T}}_k)$  for some  $T \in \widehat{\mathcal{T}}_k$ . Additionally, let  $w \in B_\ell^1 \cup B_\ell^0$ . Then, there holds for  $0 \leq s < 3/2$  if  $v \in H^s(\text{supp}(v) \cup \text{supp}(w))$  that

$$\|v\|_{H^s(\text{supp}(w))} \leq C_{\text{sc}} C_{\text{mesh}}^{k-\ell} \|v\|_{H^s(\text{supp}(v))}.$$

The constant  $C_{\text{sc}} > 0$  depends only on  $\mathcal{T}_0$  and  $n \in \mathbb{N}$ .

*Proof.* First, we prove the statement for  $s = 0$ . To that end, consider the affine transformations  $\phi_v, \phi_w: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  with  $\phi_v(\omega_v) = \text{supp}(v)$  and  $\phi_w(\omega_w) = \text{supp}(w)$ , where  $\omega_v$  and  $\omega_w$  have unit area and belong to a finite family of shapes depending only on  $\mathcal{T}_0$  and  $n \in \mathbb{N}$ . This is possible since, by assumption, the supports of  $v$  and  $w$  are contained within a generalized patch and newest-vertex bisection produces only finitely many patch shapes. We obtain

$$\begin{aligned} \|v\|_{L^2(\text{supp}(w))}^2 &= |\text{supp}(w)| \|v \circ \phi_w\|_{L^2(\omega_w)}^2 \leq |\text{supp}(w)| \|v \circ \phi_w\|_{L^\infty(\omega_w)}^2 \\ &\leq |\text{supp}(w)| \|v \circ \phi_w\|_{L^\infty(\omega_w)}^2, \end{aligned}$$

where we used the fact that  $\phi_v(\omega_v) = \text{supp}(v)$  in the last estimate. Since  $v \circ \phi_v$  belongs to a finite dimensional space which depends only on  $\omega_v$ , we further conclude

$$|\text{supp}(w)| \|v \circ \phi_w\|_{L^\infty(\omega_w)}^2 \lesssim |\text{supp}(w)| \|v \circ \phi_w\|_{L^2(\omega_w)}^2 \simeq \frac{|\text{supp}(w)|}{|\text{supp}(v)|} \|v\|_{L^2(\text{supp}(v))}^2.$$

Both estimates and the support size estimates in Theorems 4.5 and 4.7 conclude the statement for  $s = 0$ . By applying the case  $s = 0$  to  $\nabla v$  and  $w$ , we immediately prove the case  $s = 1$ . To get the intermediate cases, we employ the interpolation lemma for the operators  $T_w^s: \mathcal{P}^2(\omega^n(T, \widehat{\mathcal{T}}_k)) \cap H^s(\Omega) \rightarrow \mathcal{P}^2(\omega^n(T, \widehat{\mathcal{T}}_k) \cap \text{supp}(w)) \cap H^s(\Omega)$ . The above shows that  $T_w^1$  and  $T_w^0$  are uniformly bounded for all  $w \in B_\ell^1 \cup B_\ell^0$  with operator

norms  $C_{\text{sc}} C_{\text{mesh}}^{\ell-k}$ . The interpolation lemma implies boundedness for all  $T_w^s$ ,  $0 \leq s \leq 1$ , where the operator norms are independent of  $T$  or  $w$ . The cases  $1 < s < 3/2$  follow from the same argument applied to  $\nabla v$ .  $\square$

LEMMA 6.4. *Let  $M_{ij} := \langle \nabla v_i, \nabla v_j \rangle_\Omega$  for all  $i, j \in \mathbb{N}$  with  $v_i, v_j \in B^1$ . Given  $\varepsilon > 0$ , there exists  $M^\varepsilon \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  and a constant  $C_M > 0$  such that*

$$\|M^\varepsilon - M\|_2 \leq \varepsilon$$

as well as

$$(6.1) \quad (|L(v_i) - L(v_j)| > C_M \text{ or } d_1(v_i, v_j) > C_M) \implies M_{ij}^\varepsilon = 0.$$

*Proof.* Assume  $k := L(v_i) \leq \ell := L(v_j)$ . There holds

$$|\langle \nabla v_i, \nabla v_j \rangle_\Omega| \leq \|\nabla v_i\|_{H^{1/4}(\text{supp}(v_j))} \|\nabla v_j\|_{\tilde{H}^{-1/4}(\text{supp}(v_j))}.$$

Lemma 6.3 proves  $\|\nabla v_i\|_{H^{1/4}(\text{supp}(v_j))} \lesssim \|v_i\|_{H^{5/4}(\text{supp}(v_j))} \lesssim C_{\text{mesh}}^{-|k-\ell|} \|v_i\|_{H^{5/4}(\text{supp}(v_i))}$  as well as

$$\begin{aligned} \|\nabla v_j\|_{\tilde{H}^{-1}(\text{supp}(v_j))} &= \sup_{w \in H^1(\text{supp}(v_j))^2} \frac{\langle \nabla v_j, w \rangle_{\text{supp}(v_j)}}{\|w\|_{H^1(\text{supp}(v_j))}} \\ &= \sup_{w \in H^1(\text{supp}(v_j))^2} \frac{-\langle v_j, \text{div } w \rangle_{\text{supp}(v_j)}}{\|w\|_{H^1(\text{supp}(v_j))}} \\ &\leq \|v_j\|_{L^2(\text{supp}(v_j))}. \end{aligned}$$

Together with the obvious estimate  $\|\nabla v_j\|_{L^2(\text{supp}(v_j))} \leq \|v_j\|_{H^1(\text{supp}(v_j))}$ , interpolation concludes  $\|\nabla v_j\|_{\tilde{H}^{-1/4}(\text{supp}(v_j))} \leq \|v_j\|_{H^{3/4}(\text{supp}(v_j))}$ . The above, together with (4.2) shows

$$(6.2) \quad |\langle \nabla v_i, \nabla v_j \rangle_\Omega| \lesssim C_{\text{mesh}}^{-5/4|k-\ell|}.$$

Symmetry of the problem shows the above also for  $\ell \leq k$  and hence for all  $i, j \in \mathbb{N}$ . We restrict the index set by

$$I := \{(i, j) \in \mathbb{N}^2 : |L(v_i) - L(v_j)| \leq r\}$$

and define  $M_{ij}^\varepsilon := M_{ij}$  for all  $(i, j) \in I$  and zero elsewhere. Note that

$$\#\{v_j \in B_k^1 : M_{ij} \neq 0\} \lesssim C_{\text{mesh}}^{2(k-\min\{\ell, k\})}$$

and  $\#\{v_i \in B_\ell^1 : M_{ij} \neq 0\} \lesssim C_{\text{mesh}}^{2(\ell-\min\{k, \ell\})}$ . Estimate (6.2) and [20, Lemma 8.3] with  $q = C_{\text{mesh}}^{-2}$  show  $\|M - M^\varepsilon\|_2 \lesssim C_{\text{mesh}}^{-r/4}$ . The implication (6.1) follows by definition of  $I$ . Thus, we conclude the proof by choosing  $r \in \mathbb{N}$  sufficiently large.  $\square$

LEMMA 6.5. *Let  $M_{ij} := \langle \text{div}(v_i, 0), w_j \rangle_\Omega$  or  $M_{ij} := \langle \text{div}(0, v_i), w_j \rangle_\Omega$  for all  $i, j \in \mathbb{N}$  with  $v_i \in B^1$  and  $w_j \in B^0$ . Given  $\varepsilon > 0$ , there exists  $M^\varepsilon \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$  and a constant  $C_M > 0$  such that*

$$\|M^\varepsilon - M\|_2 \leq \varepsilon$$

as well as

$$(6.3) \quad (|L(v_i) - L(w_j)| > C_M \text{ or } d_1(v_i, w_j) > C_M) \implies M_{ij}^\varepsilon = 0.$$

*Proof.* Assume  $k := L(v_i) \geq \ell := L(w_j)$ . Then, there holds with (4.2)

$$\begin{aligned} |M_{ij}| &= |\langle (v_i, 0), \nabla w_j \rangle_\Omega| \leq \|v_i\|_{L^2(\text{supp}(v_i))} \|w_j\|_{H^1(\text{supp}(v_i))} \\ &\lesssim C_{\text{mesh}}^{-k} \|w_j\|_{H^1(\text{supp}(v_i))}. \end{aligned}$$

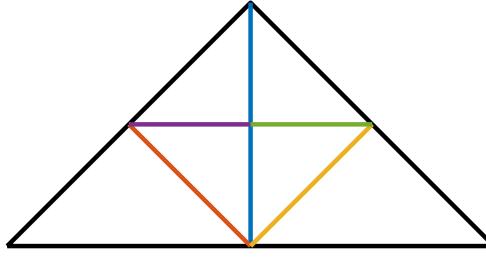


FIG. 7.1. The refinement pattern of bisec5 refinement.

From this, Lemma 6.3 together with (4.7) conclude

$$\|w_j\|_{H^1(\text{supp}(v_i))} \lesssim C_{\text{mesh}}^{-|k-\ell|} \|w_j\|_{H^1(\text{supp}(w_j))} \simeq C_{\text{mesh}}^{-|k-\ell|+\ell}.$$

Altogether, we obtain  $|M_{ij}| \lesssim C_{\text{mesh}}^{-2|k-\ell|}$ . For  $k < \ell$ , we have with Lemma 6.3 as well as (4.2) and (4.7) for some  $0 < s < 1/2$

$$\begin{aligned} |M_{ij}| &\leq \|v_i\|_{H^{1+s}(\text{supp}(w_j))} \|w_j\|_{H^{-s}(\text{supp}(w_j))} \\ &\lesssim C_{\text{mesh}}^{-s\ell} \|v_i\|_{H^{1+s}(\text{supp}(w_j))} \\ &\lesssim C_{\text{mesh}}^{-s\ell-|\ell-k|} \|v_i\|_{H^{1+s}(\text{supp}(v_i))} \lesssim C_{\text{mesh}}^{-(1+s)|\ell-k|}. \end{aligned}$$

We restrict the index set by

$$I := \{(i, j) \in \mathbb{N}^2 : |L(v_i) - L(v_j)| \leq r\}$$

and define  $M_{ij}^\varepsilon := M_{ij}$  for all  $(i, j) \in I$  and zero elsewhere. Note that

$$\#\{v_j \in B_k^1 : M_{ij} \neq 0\} \lesssim C_{\text{mesh}}^{2(k-\min\{\ell,k\})}$$

and  $\#\{v_i \in B_\ell^1 : M_{ij} \neq 0\} \lesssim C_{\text{mesh}}^{2(\ell-\min\{k,\ell\})}$ . The above estimates and [20, Lemma 8.3] with  $q = C_{\text{mesh}}^{-2}$  show  $\|M - M^\varepsilon\|_2 \lesssim C_{\text{mesh}}^{-sr}$ . The implication (6.1) follows by definition of  $I$ . Thus, we conclude the proof by choosing  $r \in \mathbb{N}$  sufficiently large.  $\square$

**7. A modified Scott–Zhang projection.** This section introduces the operators  $S_\ell^1$  and  $S_\ell^2$  used for constructing the Riesz basis in Theorems 4.5 and 4.7. First, we construct some slightly modified Scott–Zhang operator in Lemma 7.2 which is stable on  $L^2(\Omega)$  and satisfies some integral mean property. This is important as we use that operator on the pressure space  $L_*^2(\Omega)$  which contains functions with zero integral mean. In a second step in Theorems 7.3 and 7.5, we use the Scott–Zhang operators to construct the operators  $S_\ell^1$  and  $S_\ell^2$  which inherit all the nice properties from the Scott–Zhang operators but additionally are nested in the sense that  $S_\ell^i S_k^i = S_\ell^i$  for all  $i \in \{1, 2\}$  and all  $\ell \leq k \in \mathbb{N}$ .

In the following, we define a particular basis of  $\mathcal{S}^1(\mathcal{T})$  with a certain moment condition.

**DEFINITION 7.1.** Given a triangulation  $\mathcal{T}_c \in \mathbb{T}$  and it's bisec5 refinement  $\mathcal{T} \in \text{bisec5}(\mathcal{T}_c, \mathcal{T}_c)$  (see Figure 7.1) define for each  $T \in \mathcal{T}_c$  the unique interior node  $z_T \in \mathcal{N}(\mathcal{T})$  with  $z_T \in T \setminus \partial T$ . Define  $\mathcal{N}(\mathcal{T}, \mathcal{T}_c) := \{z_T : T \in \mathcal{T}_c\}$ . For  $z \in \mathcal{N}(\mathcal{T}) \setminus \mathcal{N}(\mathcal{T}, \mathcal{T}_c)$  and the corresponding hat function  $v_z \in \mathcal{S}_B^1(\mathcal{T})$ , define

$$v_{z,*} := v_z + \sum_{\substack{T \in \mathcal{T}_c \\ v_z|_T \neq 0}} \alpha_T v_{z_T},$$

where the  $\alpha_T \in \mathbb{R}$  are chosen such that  $\int_T v_{z,*} dx = 0$  for all  $T \in \mathcal{T}_c$ . Define the basis

$$\mathcal{S}_B^1(\mathcal{T}, \mathcal{T}_c) := \{v_{z,*} \in \mathcal{S}_B^1(\mathcal{T}) : z \in \mathcal{N}(\mathcal{T}) \setminus \mathcal{N}(\mathcal{T}, \mathcal{T}_c)\} \cup \{v_{z_T} \in \mathcal{S}_B^1(\mathcal{T}) : T \in \mathcal{T}_c\}$$

which satisfies  $\mathcal{S}^1(\mathcal{T}) = \text{span}(\mathcal{S}_B^1(\mathcal{T}, \mathcal{T}_c))$ . For each  $v \in \mathcal{S}_B^1(\mathcal{T}, \mathcal{T}_c)$  define some  $T_v \in \mathcal{T}$  with  $v|_{T_v} \neq 0$ . Note that for all  $T \in \mathcal{T}_c$ , the set  $\{v|_T : v \in \mathcal{S}_B^1(\mathcal{T}, \mathcal{T}_c)\} \setminus \{0\}$  is a basis of  $\mathcal{S}^1(\mathcal{T}|_T)$ . This allows us to define the dual basis functions  $v'_T$  of  $v|_T$  for all  $v \in \mathcal{S}_B^1(\mathcal{T}, \mathcal{T}_c)$  with  $v|_T \neq 0$ . Finally, we define  $v' := v'_{T_v}$  for all  $v \in \mathcal{S}_B^1(\mathcal{T}, \mathcal{T}_c)$ . We define the Scott–Zhang projection as

$$J_{\mathcal{T}, \mathcal{T}_c}^1 w := \sum_{v \in \mathcal{S}_B^1(\mathcal{T}, \mathcal{T}_c)} v \langle w, v' \rangle_{T_v}.$$

LEMMA 7.2. *The Scott–Zhang operator  $J_{\mathcal{T}, \mathcal{T}_c}^1 : L^2(\Omega) \rightarrow \mathcal{S}^1(\mathcal{T})$  defined in Definition 7.1 is a projection which satisfies for all  $0 \leq s \leq 1$  and all  $v \in H^s(\Omega)$*

$$(7.1) \quad \|J_{\mathcal{T}, \mathcal{T}_c}^1 v\|_{H^s(T)} \leq C_{sz} \|v\|_{H^s(\cup \omega(T, \mathcal{T}_c))},$$

$$(7.2) \quad \|J_{\mathcal{T}, \mathcal{T}_c}^1 v\|_{H^s(\Omega)} \leq C_{sz} \|v\|_{H^s(\Omega)},$$

as well as for all  $0 \leq r \leq s$

$$(7.3) \quad \|(1 - J_{\mathcal{T}, \mathcal{T}_c}^1)v\|_{H^r(T)} \leq C_{sz} \text{diam}(T)^{s-r} |v|_{H^s(\cup \omega(T, \mathcal{T}_c))},$$

$$(7.4) \quad \|(1 - J_{\mathcal{T}, \mathcal{T}_c}^1)v\|_{H^r(\Omega)} \leq C_{sz} \|h_{\mathcal{T}}^{1-r} \nabla v\|_{L^2(\Omega)}.$$

For  $0 < s < 1/2$  and  $v \in L^2(\Omega)$ ,  $w \in \tilde{H}^{-s}(\Omega)$ , there holds

$$(7.5) \quad \|J_{\mathcal{T}, \mathcal{T}_c}^1 w\|_{\tilde{H}^{-s}(\Omega)} \leq C_{sz} \|w\|_{\tilde{H}^{-s}(\Omega)},$$

$$(7.6) \quad \|(1 - J_{\mathcal{T}, \mathcal{T}_c}^1)v\|_{\tilde{H}^{-s}(\Omega)} \leq C_{sz} \|h_{\mathcal{T}}^s v\|_{L^2(\Omega)}.$$

There holds  $\int_T (1 - J_{\mathcal{T}, \mathcal{T}_c}^1)v dx = 0$  for all  $T \in \mathcal{T}_c$ . The constant  $C_{sz} > 0$  depends only on the shape regularity of  $\mathcal{T}$ , the fact that  $\mathcal{T}, \mathcal{T}_c$  are generated from  $\mathcal{T}_0$  by the newest vertex bisection and on a lower bound on  $s > -1/2$ . The function  $(J_{\mathcal{T}, \mathcal{T}_c}^1 v)|_T$  depends only on  $v|_{\cup \omega(T, \mathcal{T}_c)}$ .

*Proof.* The projection property follows as usual from the definition of the dual basis functions  $v'$ . For  $T \in \mathcal{T}_c$ ,  $(J_{\mathcal{T}} v)|_T$  depends only on  $v|_{T'}$  for all  $T' \in \mathcal{T}_c$  with  $T, T' \subseteq \text{supp}(w)$  for some  $w \in \mathcal{S}_B^1(\mathcal{T}, \mathcal{T}_c)$ . This implies  $T' \in \omega(T, \mathcal{T}_c)$ . Thus,  $(J_{\mathcal{T}} v)|_T$  depends only on  $v|_{\omega(T, \mathcal{T}_c)}$ .

The estimates (7.1)–(7.4) follow analogously to the proof of [3, section 3.2]. It remains to prove the estimates for  $-s$ . To that end, note that the dual function  $v'_{z_T}$  (where  $z_T$  was defined as the unique interior node of the element  $T \in \mathcal{T}_c$ ) is a constant on  $T$ . This follows from the fact that there holds  $\langle 1, v \rangle_T = 0$  for all  $v \in \mathcal{S}_B^1(\mathcal{T}, \mathcal{T}_c) \setminus \{v_{z_T} : T \in \mathcal{T}_c\}$  and  $\langle 1, v_{z_T} \rangle_T > 0$ . Uniqueness of the dual function implies that  $v'_{z_T}$  is some multiple of 1 and hence constant. This and the projection property imply for all  $T \in \mathcal{T}_c$ , all  $v \in L^2(\Omega)$ , and  $w := v + \sum_{T \in \mathcal{T}_c} \beta_T v_{z_T}$ ,

$$\begin{aligned} (1 - J_{\mathcal{T}, \mathcal{T}_c}^1)v &= (1 - J_{\mathcal{T}, \mathcal{T}_c}^1)w \\ &= w - \left( \sum_{T \in \mathcal{T}_c} v_{z_T} \langle w, v'_{z_T} \rangle_T + \sum_{v \in \mathcal{S}_B^1(\mathcal{T}, \mathcal{T}_c) \setminus \{v_{z_T} : T \in \mathcal{T}_c\}} v \langle w, v' \rangle_T \right). \end{aligned}$$

Choosing the  $\beta_T$  such that  $\int_T w \, dx = 0$  for all  $T \in \mathcal{T}_c$ , we obtain  $\langle w, v'_{z_T} \rangle_T = 0$  and hence

$$\int_T (1 - J_{\mathcal{T}, \mathcal{T}_c})v \, dx = 0 \quad \text{for all } T \in \mathcal{T}_c.$$

Moreover, we have for the  $L^2$ -orthogonal projection  $\Pi^1: L^2(\Omega) \rightarrow \mathcal{P}^1(\mathcal{T})$  the identity  $J_{\mathcal{T}, \mathcal{T}_c}^1 v = J_{\mathcal{T}, \mathcal{T}_c}^1 \Pi^1 v$ . For  $0 < s < 1/2$ , this together with a standard Poincaré estimate and an inverse estimate imply

$$\begin{aligned} \|(\Pi^1 - J_{\mathcal{T}, \mathcal{T}_c}^1)v\|_{\tilde{H}^{-s}(\Omega)} &= \|(1 - J_{\mathcal{T}, \mathcal{T}_c}^1)\Pi^1 v\|_{\tilde{H}^{-s}(\Omega)} \lesssim \|h_{\mathcal{T}_c}^s \Pi^1 v\|_{L^2(\Omega)} \\ &\lesssim \|h_{\mathcal{T}_c}/h_{\mathcal{T}}\|_{L^\infty(\Omega)} \|\Pi^1 v\|_{\tilde{H}^{-s}(\Omega)} \lesssim \|v\|_{\tilde{H}^{-s}(\Omega)}, \end{aligned}$$

where we used the continuity  $\Pi^1: \tilde{H}^{-s}(\Omega) \rightarrow \tilde{H}^{-s}(\Omega)$ . Together with the approximation properties of  $\Pi^1$ , this shows

$$\|J_{\mathcal{T}, \mathcal{T}_c}^1 v\|_{\tilde{H}^{-s}(\Omega)} \lesssim \|\Pi^1 v\|_{\tilde{H}^{-s}(\Omega)} + \|(\Pi^1 - J_{\mathcal{T}, \mathcal{T}_c}^1)v\|_{\tilde{H}^{-s}(\Omega)} \lesssim \|v\|_{\tilde{H}^{-s}(\Omega)}$$

as well as

$$\|(1 - J_{\mathcal{T}, \mathcal{T}_c}^1)v\|_{\tilde{H}^{-s}(\Omega)} \lesssim \|(1 - \Pi^1)v\|_{\tilde{H}^{-s}(\Omega)} + \|(\Pi^1 - J_{\mathcal{T}, \mathcal{T}_c}^1)v\|_{\tilde{H}^{-s}(\Omega)} \lesssim \|h_{\mathcal{T}_c}^s v\|_{L^2(\Omega)}$$

and concludes (7.5)–(7.6).  $\square$

For the next theorem, note that for all  $\ell \geq 1$ , there exists a coarsening  $\bar{\mathcal{T}}_\ell$  of  $\hat{\mathcal{T}}_\ell$  such that  $\hat{\mathcal{T}}_\ell = \text{bisecc5}(\bar{\mathcal{T}}_\ell, \bar{\mathcal{T}}_\ell)$ .

**THEOREM 7.3.** *Recall  $\hat{\mathcal{T}}_\ell$  and  $C_{\text{mesh}}$  from Definition 4.1. With  $J_\ell^1 := J_{\hat{\mathcal{T}}_\ell, \bar{\mathcal{T}}_\ell}^1$ , define*

$$S_\ell^1 v := \lim_{N \rightarrow \infty} (J_\ell^1 J_{\ell+1}^1 \dots J_{\ell+N}^1) v \in S_\star^1(\hat{\mathcal{T}}_\ell)$$

for all  $v \in L^2(\Omega)$  and all  $\ell \geq 1$ . Then, the operator  $S_\ell^1: L^2(\Omega) \rightarrow \mathcal{S}^1(\hat{\mathcal{T}}_\ell)$  is well-defined and satisfies

(7.7)

$$\|(1 - S_\ell^1)v\|_{\tilde{H}^{-s}(\Omega)} \leq C_S C_{\text{mesh}}^{-s\ell} \|v\|_{L^2(\Omega)} \quad \text{for all } v \in L^2(\Omega) \text{ and } 0 < s < 1/2,$$

(7.8)

$$\|(1 - S_\ell^1)v\|_{\tilde{H}^{-s}(\Omega)} \leq C_S C_{\text{mesh}}^{-(s-r)\ell} \|v\|_{\tilde{H}^{-r}(\Omega)} \quad \text{for all } v \in \tilde{H}^{-r}(\Omega), 0 < r \leq s - 1/4,$$

(7.9)

$$\|(1 - S_\ell^1)v\|_{L^2(\Omega)} \leq C_S C_{\text{mesh}}^{-\ell/4} \|v\|_{H^{1/4}(\Omega)}, \quad \text{for all } v \in H^{1/4}(\Omega),$$

(7.10)

$$\|S_\ell^1 v\|_{\tilde{H}^{-1/4}(\Omega)} \leq C_S \|v\|_{\tilde{H}^{-1/4}(\Omega)}, \quad \text{for all } v \in \tilde{H}^{-1/4}(\Omega),$$

$$(7.11) \quad \|S_\ell^1 v\|_{H^s(\Omega)} \leq C_S \|v\|_{H^s(\Omega)} \quad \text{for all } v \in H^s(\Omega) \text{ and } s \in \{0, 1/4\}.$$

Moreover, there holds  $S_\ell^1 S_k^1 = S_{\min\{\ell, k\}}^1$  for all  $\ell, k \in \mathbb{N}_0$  as well as

$$(7.12) \quad S_\ell^1 v = J_\ell^1 J_{\ell+1}^1 \dots J_{\ell+k}^1 v \quad \text{for all } v \in \mathcal{S}^1(\hat{\mathcal{T}}_{\ell+k})$$

for all  $\ell, k \in \mathbb{N}$ . The constant  $C_S > 0$  depends only on  $C_{\text{sz}}$  and  $C_{\text{mesh}}$ ,  $\mathcal{T}_0$ , and  $s$ .

*Proof.* For brevity of presentation, we also write  $H^{-s}(\Omega)$  for  $\tilde{H}^{-s}(\Omega)$  in this proof. From Lemma 7.2, we obtain for  $-1/2 < r \leq 1$  and  $s = 1$

$$(7.13) \quad \|(1 - J_\ell^1)v\|_{H^r(\Omega)} \lesssim C_{\text{mesh}}^{1-r} \|v\|_{H^s(\Omega)} \quad \text{for all } v \in H^s(\Omega).$$

Moreover, from (7.2), we even get  $\|(1 - J_\ell^1)v\|_{H^r(\Omega)} \lesssim \|v\|_{H^r(\Omega)}$  for all  $0 \leq r \leq 1$ . Interpolation arguments prove that (7.13) holds for all  $s$  with  $r \leq s \leq 1$ . Let  $u \in H^\mu(\Omega)$ . Since  $J_\ell^1 : H^\nu(\Omega) \rightarrow H^\nu(\Omega)$  is continuous for  $-1/2 < \nu \leq 1$  (see Lemma 7.2), there holds for  $-1/2 < \nu \leq 1$  and  $\nu + 1/4 \leq \mu \leq 1$  with  $\mu \geq 0$  and by use of (7.13) that

$$\begin{aligned} (7.14) \quad & \|(1 - (J_\ell^1 J_{\ell+1}^1 \dots J_{\ell+N}^1))u\|_{H^\nu(\Omega)} \\ & \leq \|(1 - J_\ell^1)u\|_{H^\nu(\Omega)} + \sum_{k=0}^{N-1} \|((J_\ell^1 \dots J_{\ell+k}^1) - (J_\ell^1 \dots J_{\ell+k+1}^1))u\|_{H^\nu(\Omega)} \\ & \leq \|(1 - J_\ell^1)u\|_{H^\nu(\Omega)} + \sum_{k=0}^{N-1} \|(J_\ell^1 \dots J_{\ell+k}^1)\|_{H^\nu(\Omega) \rightarrow H^\nu(\Omega)} \|(1 - J_{\ell+k+1}^1)u\|_{H^\nu(\Omega)} \\ & \lesssim \sum_{k=0}^N C_{\text{sz}}^k C_{\text{mesh}}^{-(\mu-\nu)(\ell+k)} \|u\|_{H^\mu(\Omega)} \lesssim C_{\text{mesh}}^{-(\mu-\nu)\ell} \|u\|_{H^\mu(\Omega)}, \end{aligned}$$

where we used  $C_{\text{mesh}}^{(\mu-\nu)} \geq C_{\text{mesh}}^{1/4} > C_{\text{sz}}$  from Definition 4.1. Hence, (7.14) for  $(\nu_1, \mu_1) = (0, 1/4)$  and  $(\nu_2, \mu_2) = (1/4, 1/2)$  implies for  $M \leq N$

$$\begin{aligned} & \|((J_\ell^1 J_{\ell+1}^1 \dots J_{\ell+M}^1) - (J_\ell^1 J_{\ell+1}^1 \dots J_{\ell+N}^1))u\|_{L^2(\Omega)} \\ & = \|(J_\ell^1 \dots J_{\ell+M}^1)((1 - (J_{\ell+M+1}^1 J_{\ell+1}^1 \dots J_{\ell+N}^1))u\|_{L^2(\Omega)} \\ & \leq \|((1 - (J_{\ell+M+1}^1 J_{\ell+1}^1 \dots J_{\ell+N}^1))u\|_{L^2(\Omega)} \\ & \quad + \|(1 - (J_\ell^1 \dots J_{\ell+M}^1))((1 - (J_{\ell+M+1}^1 J_{\ell+1}^1 \dots J_{\ell+N}^1))u\|_{L^2(\Omega)} \\ & \lesssim \|((1 - (J_{\ell+M+1}^1 J_{\ell+1}^1 \dots J_{\ell+N}^1))u\|_{H^{1/4}(\Omega)} \\ & \lesssim C_{\text{mesh}}^{-(\ell+M)/4} \|u\|_{H^{1/2}(\Omega)}. \end{aligned}$$

This shows, that  $(J_\ell^1 J_{\ell+1}^1 \dots J_{\ell+N}^1)u$  is a Cauchy sequence in  $\mathcal{S}^1(\widehat{\mathcal{T}}_\ell)$  with respect to the  $L^2(\Omega)$ -norm as  $N \rightarrow \infty$ . Thus, for  $u \in H^{1/2}(\Omega)$ , the limit  $S_\ell^1 u \in \mathcal{S}^1(\widehat{\mathcal{T}}_\ell)$  exists. The estimates (7.7)–(7.9) follow from (7.14), the convergence  $\lim_{N \rightarrow \infty} (J_\ell^1 \dots J_{\ell+N}^1)v = S_\ell v$  in  $L^2(\Omega)$ , and density arguments.

To see (7.10)–(7.11), we apply inverse estimates for  $s \in \{-1/4, 0, 1/4\}$  as well as Lemma 7.2 and (7.7)–(7.8) to see

$$\begin{aligned} \|S_\ell^1 u\|_{H^s(\Omega)} & \lesssim \|(S_\ell^1 - J_\ell^1)u\|_{H^s(\Omega)} + \|u\|_{H^s(\Omega)} \\ & \lesssim C_{\text{mesh}}^{\ell(2/5+s)} \|(S_\ell^1 - J_\ell^1)u\|_{\tilde{H}^{-2/5}(\Omega)} + \|u\|_{H^s(\Omega)} \\ & \lesssim C_{\text{mesh}}^{\ell(2/5+s)} (\|(1 - S_\ell^1)u\|_{\tilde{H}^{-2/5}(\Omega)} + \|(1 - J_\ell^1)u\|_{\tilde{H}^{-2/5}(\Omega)}) + \|u\|_{H^s(\Omega)} \\ & \lesssim \|u\|_{H^s(\Omega)}. \end{aligned}$$

The proof of  $S_k S_\ell^1 u = S_k u$ , (7.12) and the projection property follow analogously to [20, Theorem 5.6].  $\square$

LEMMA 7.4. Define the Scott–Zhang operator  $J_{\mathcal{T}}^2: H^1(\Omega) \rightarrow \mathcal{S}^2(\mathcal{T})$  from [39].  $J_{\mathcal{T}}^2$  is a projection which satisfies for all  $1/2 < s < 3/2$  and all  $v \in H^s(\Omega)$

$$(7.15) \quad \|J_{\mathcal{T}}^2 v\|_{H^s(T)} \leq C_{sz} \|v\|_{H^s(\cup\omega(T,\mathcal{T}))},$$

$$(7.16) \quad \|J_{\mathcal{T}}^2 v\|_{H^s(\Omega)} \leq C_{sz} \|v\|_{H^s(\Omega)},$$

as well as for all  $0 \leq r \leq s$

$$(7.17) \quad \|(1 - J_{\mathcal{T}}^2)v\|_{H^r(T)} \leq C_{sz} \operatorname{diam}(T)^{s-r} |v|_{H^s(\cup\omega(T,\mathcal{T}))},$$

$$(7.18) \quad \|(1 - J_{\mathcal{T}}^2)v\|_{H^r(\Omega)} \leq C_{sz} \|h_{\mathcal{T}}^{1-r} \nabla v\|_{L^2(\Omega)}.$$

The constant  $C_{sz} > 0$  depends only on the shape regularity of  $\mathcal{T}$ , the fact that  $\mathcal{T}$  is generated from  $\mathcal{T}_0$  by newest vertex bisection, and on a lower bound on  $s > 1/2$ . The function  $(J_{\mathcal{T}}^2 v)|_T$  depends only on  $v|_{\cup\omega(T,\mathcal{T})}$  and  $v|_{\partial\Omega} = 0$  implies  $(J_{\mathcal{T}}^2 v)|_{\partial\Omega} = 0$ .

THEOREM 7.5. With  $J_{\ell}^2 := J_{\widehat{\mathcal{T}}_{\ell}}^2: H^1(\Omega) \rightarrow \mathcal{S}^2(\widehat{\mathcal{T}}_{\ell})$  denoting the standard Scott–Zhang projection from Lemma 7.4, define

$$S_{\ell}^2 v := \lim_{N \rightarrow \infty} (J_{\ell}^2 J_{\ell+1}^2 \dots J_{\ell+N}^2) v \in \mathcal{S}^2(\widehat{\mathcal{T}}_{\ell})$$

for all  $v \in H^1(\Omega)$ . Then, the operator  $S_{\ell}^2: H^1(\Omega) \rightarrow \mathcal{S}^1(\widehat{\mathcal{T}}_{\ell})$  is well-defined and satisfies

(7.19)

$$\|(1 - S_{\ell}^2)v\|_{H^{1-s}(\Omega)} \leq C_S C_{\text{mesh}}^{-s\ell} \|v\|_{H^1(\Omega)} \quad \text{for all } v \in H^1(\Omega) \text{ and } 0 < s < 1/2,$$

$$(7.20) \quad \|(1 - S_{\ell}^2)v\|_{H^1(\Omega)} \leq C_S C_{\text{mesh}}^{-\ell/4} \|v\|_{H^{5/4}(\Omega)} \quad \text{for all } v \in H^1(\Omega),$$

$$(7.21) \quad \|S_{\ell}^2 v\|_{H^s(\Omega)} \leq C_S \|v\|_{H^s(\Omega)} \quad \text{for all } v \in H_{\star}^s(\Omega) \text{ and } s \in \{3/4, 1, 5/4\}.$$

Moreover, there holds  $S_{\ell}^2 S_k^2 = S_{\min\{\ell,k\}}^2$  for all  $\ell, k \in \mathbb{N}_0$  as well as

$$(7.22) \quad S_{\ell}^2 v = J_{\ell}^2 J_{\ell+1}^2 \dots J_{\ell+k}^2 v \quad \text{for all } v \in \mathcal{S}_{\star}^1(\widehat{\mathcal{T}}_{\ell+k})$$

for all  $\ell, k \in \mathbb{N}$ . The constant  $C_S > 0$  depends only on  $C_{sz}$  and  $C_{\text{mesh}}$ ,  $\mathcal{T}_0$ , and  $s$ .

*Proof.* The proof follows analogously to [20, Theorem 5.6] and is therefore omitted.  $\square$

## REFERENCES

- [1] K. T. ANDREWS AND J. D. WARD, *LU-factorization of order bounded operators on Banach sequence spaces*, J. Approx. Theory, 48 (1986), pp. 169–180, [https://doi.org/10.1016/0021-9045\(86\)90001-8](https://doi.org/10.1016/0021-9045(86)90001-8).
- [2] M. AURADA, M. FEISCHL, T. FÜHRER, M. KARKULIK, AND D. PRAETORIUS, *Efficiency and optimality of some weighted-residual error estimator for adaptive 2D boundary element methods*, J. Comput. Appl. Math., 255 (2014), pp. 481–501.
- [3] M. AURADA, M. FEISCHL, T. FÜHRER, M. KARKULIK, AND D. PRAETORIUS, *Energy norm based error estimators for adaptive BEM for hypersingular integral equations*, Appl. Numer. Math., 95 (2015), pp. 15–35, <https://doi.org/10.1016/j.apnum.2013.12.004>.
- [4] E. BÄNSCH, P. MORIN, AND R. H. NOCHETTO, *An adaptive Uzawa FEM for the Stokes problem: Convergence without the inf-sup condition*, SIAM J. Numer. Anal., 40 (2002), pp. 1207–1229, <https://doi.org/10.1137/S0036142901392134>.
- [5] R. BECKER AND S. MAO, *Quasi-optimality of adaptive nonconforming finite element methods for the Stokes equations*, SIAM J. Numer. Anal., 49 (2011), pp. 970–991, <https://doi.org/10.1137/100802967>.

- [6] R. BECKER, S. MAO, AND Z. SHI, *A convergent nonconforming adaptive finite element method with quasi-optimal complexity*, SIAM J. Numer. Anal., 47 (2010), pp. 4639–4659.
- [7] P. BINEV, W. DAHΜEN, AND R. DEVORE, *Adaptive finite element methods with convergence rates*, Numer. Math., 97 (2004), pp. 219–268.
- [8] D. BOFFI, *Stability of higher order triangular Hood-Taylor methods for the stationary Stokes equations*, Math. Models Methods Appl. Sci., 4 (1994), pp. 223–235, <https://doi.org/10.1142/S0218202594000133>.
- [9] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math. 15, Springer, New York, 1991.
- [10] C. CARSTENSEN, M. FEISCHL, M. PAGE, AND D. PRAETORIUS, *Axioms of adaptivity*, Comput. Math. Appl., 67 (2014), pp. 1195–1253, <https://doi.org/10.1016/j.camwa.2013.12.003>.
- [11] C. CARSTENSEN, D. PETERSEIM, AND H. RABUS, *Optimal adaptive nonconforming FEM for the Stokes problem*, Numer. Math., 123 (2013), pp. 291–308, <https://doi.org/10.1007/s00211-012-0490-8>.
- [12] C. CARSTENSEN AND H. RABUS, *The adaptive nonconforming FEM for the pure displacement problem in linear elasticity is optimal and robust*, SIAM J. Numer. Anal., 50 (2012), pp. 1264–1283.
- [13] J. M. CASCON, C. KREUZER, R. H. NOCHETTO, AND K. G. SIEBERT, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal., 46 (2008), pp. 2524–2550.
- [14] J. M. CASCON AND R. H. NOCHETTO, *Quasioptimal cardinality of AFEM driven by nonresidual estimators*, IMA J. Numer. Anal., 32 (2012), pp. 1–29.
- [15] L. CHEN, M. HOLST, AND J. XU, *Convergence and optimality of adaptive mixed finite element methods*, Math. Comp., 78 (2009), pp. 35–53.
- [16] S. DAHLKE, R. HOCHMUTH, AND K. URBAN, *Convergent adaptive wavelet methods for the Stokes problem*, in Multigrid Methods, VI (Gent, 1999), Lect. Notes Comput. Sci. Eng. 14, Springer, Berlin, 2000, pp. 66–72, [https://doi.org/10.1007/978-3-642-58312-4\\_8](https://doi.org/10.1007/978-3-642-58312-4_8).
- [17] W. DAHΜEN, *Stability of multiscale transformations*, J. Fourier Anal. Appl., 2 (1996), pp. 341–361.
- [18] A. DEMLOW AND R. STEVENSON, *Convergence and quasi-optimality of an adaptive finite element method for controlling  $L_2$  errors*, Numer. Math., 117 (2011), pp. 185–218, <https://doi.org/10.1007/s00211-010-0349-9>.
- [19] L. DIENING, C. KREUZER, AND R. STEVENSON, *Instance optimality of the adaptive maximum strategy*, Found. Comput. Math., 16 (2016), pp. 33–68, <https://doi.org/10.1007/s10208-014-9236-6>.
- [20] M. FEISCHL, *Optimal Adaptivity for Non-Symmetric FEM/BEM Coupling*, preprint, arXiv:1710.06082, 2017.
- [21] M. FEISCHL, T. FÜHRER, M. KARKULIK, J. MELENK, AND D. PRAETORIUS, *Quasi-optimal convergence rates for adaptive boundary element methods with data approximation, part II: Hypersingular integral equation*, Electron. Trans. Numer. Anal., 44 (2015), pp. 153–176.
- [22] M. FEISCHL, T. FÜHRER, M. KARKULIK, J. M. MELENK, AND D. PRAETORIUS, *Quasi-optimal convergence rates for adaptive boundary element methods with data approximation, part I: Weakly-singular integral equation*, Calcolo, 51 (2014), pp. 531–562, <https://doi.org/10.1007/s10092-013-0100-x>.
- [23] M. FEISCHL, T. FÜHRER, AND D. PRAETORIUS, *Adaptive FEM with optimal convergence rates for a certain class of nonsymmetric and possibly nonlinear problems*, SIAM J. Numer. Anal., 52 (2014), pp. 601–625, <https://doi.org/10.1137/120897225>.
- [24] M. FEISCHL, M. KARKULIK, J. M. MELENK, AND D. PRAETORIUS, *Quasi-optimal convergence rate for an adaptive boundary element method*, SIAM J. Numer. Anal., 51 (2013), pp. 1327–1348.
- [25] T. GANTUMUR, *Adaptive boundary element methods with convergence rates*, Numer. Math., 124 (2013), pp. 471–516.
- [26] T. GANTUMUR, *On the Convergence Theory of Adaptive Mixed Finite Element Methods for the Stokes Problem*, preprint, arXiv:1403.0895, 2014.
- [27] N. HEUER, *On the equivalence of fractional-order Sobolev semi-norms*, J. Math. Anal. Appl., 417 (2014), pp. 505–518, <https://doi.org/10.1016/j.jmaa.2014.03.047>.
- [28] J. HU AND J. XU, *Convergence and optimality of the adaptive nonconforming linear element method for the Stokes problem*, J. Sci. Comput., 55 (2013), pp. 125–148, <https://doi.org/10.1007/s10915-012-9625-4>.
- [29] S. JAFFARD, *Propriétés des matrices “bien localisées” près de leur diagonale et quelques applications*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 7 (1990), pp. 461–476, <http://eudml.org/doc/78234>.

- [30] H. J. GUO AND X. Y. FENG, *Convergence and complexity of arbitrary order adaptive mixed element methods for the Poisson equation*, Sci. China Math., 55 (2012), pp. 1083–1098.
- [31] M. KARKULIK, D. PAVLICEK, AND D. PRAETORIUS, *On 2D newest vertex bisection: Optimality of mesh-closure and  $H^1$ -stability of  $L_2$ -projection*, Constr. Approx., 38 (2013), pp. 213–234, <https://doi.org/10.1007/s00365-013-9192-4>.
- [32] Y. KONDRAKYUK AND R. STEVENSON, *An optimal adaptive finite element method for the Stokes problem*, SIAM J. Numer. Anal., 46 (2008), pp. 747–775, <https://doi.org/10.1137/06066566X>.
- [33] C. KREUZER AND M. SCHEDENSACK, *Instance optimal Crouzeix-Raviart adaptive finite element methods for the Poisson and Stokes problems*, IMA J. Numer. Anal., 36 (2016), pp. 593–617, <https://doi.org/10.1093/imanum/drv019>.
- [34] C. KREUZER AND K. G. SIEBERT, *Decay rates of adaptive finite elements with Dörfler marking*, Numer. Math., 117 (2011), pp. 679–716.
- [35] I. KRISHTAL, T. STROHMER, AND T. WERTZ, *Localization of matrix factorizations*, Found. Comput. Math., 15 (2015), pp. 931–951, <https://doi.org/10.1007/s10208-014-9196-x>.
- [36] S. MAO, X. ZHAO, AND Z. SHI, *Convergence of a standard adaptive nonconforming finite element method with optimal complexity*, Appl. Numer. Math., 60 (2010), pp. 673–688.
- [37] P. MORIN, K. G. SIEBERT, AND A. VEESER, *A basic convergence result for conforming adaptive finite elements*, Math. Models Methods Appl. Sci., 18 (2008), pp. 707–737.
- [38] H. RABUS, *A natural adaptive nonconforming FEM of quasi-optimal complexity*, Comput. Methods Appl. Math., 10 (2010), pp. 315–325.
- [39] L. R. SCOTT AND S. ZHANG, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
- [40] K. G. SIEBERT, *A convergence proof for adaptive finite elements without lower bound*, IMA J. Numer. Anal., 31 (2011), pp. 947–970, <https://doi.org/10.1093/imanum/drq001>.
- [41] R. STEVENSON, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., 7 (2007), pp. 245–269.
- [42] R. STEVENSON, *The completion of locally refined simplicial partitions created by bisection*, Math. Comp., 77 (2008), pp. 227–241.
- [43] R. VERFÜRTH, *Error estimates for a mixed finite element approximation of the Stokes equations*, RAIRO Anal. Numér., 18 (1984), pp. 175–182, <https://doi.org/10.1051/m2an/1984180201751>.