EXISTENCE OF $\mathcal{H}$-MATRIX APPROXIMANTS TO THE INVERSES OF BEM MATRICES: THE SIMPLE-LAYER OPERATOR

MARKUS FAUSTMANN, JENS MARKUS MELENK, AND DIRK PRAETORIUS

Dedicated to Wolfgang Hackbusch on the occasion of his 65th birthday

ABSTRACT. We consider the question of approximating the inverse $\mathbf{W} = \mathbf{V}^{-1}$ of the Galerkin stiffness matrix $\mathbf{V}$ obtained by discretizing the simple-layer operator V with piecewise constant functions. The block partitioning of $\mathbf{W}$ is assumed to satisfy any one of several standard admissibility criteria that are employed in connection with clustering algorithms to approximate the discrete BEM operator $\mathbf{V}$. We show that $\mathbf{W}$ can be approximated by block-wise low-rank matrices such that the error decays exponentially in the block rank employed. Similar exponential approximability results are shown for the Cholesky factorization of $\mathbf{V}$.

1. INTRODUCTION

The system matrices arising in the boundary element method (BEM) such as the matrix $\mathbf{V}$ for the classical simple-layer operator V are fully populated. The classical BEM is therefore often deemed inefficient with respect to memory requirements and, in turn, fast iterative solvers. These reservations can be met with various compression techniques that have been developed in the past to store the BEM matrices and realize the matrix-vector multiplication with log-linear (or even linear) complexity. We mention here multipole expansions, [Rok85, GR97], panel clustering, [NH88, HN89, HS93, Sau92], wavelet compression techniques, [Rat98, Rat01, Sch98, vPSS97, Tau03, TW03], the mosaic-skeleton method, [Tyr00], the adaptive cross approximation (ACA) method, [Beb00], and the hybrid cross approximation (HCA), [BG05]. Many of these data-sparse methods can be understood as specific instances of $\mathcal{H}$-matrices, which were introduced and analyzed in [Hac99, GH03, Gra01, Hac09] as blockwise low rank matrices. $\mathcal{H}$-matrices come with the additional feature that they permit an (approximate) arithmetic with log-linear complexity. In particular, this arithmetic includes the (approximate) inversion of matrices. Thus, the $\mathcal{H}$-matrix arithmetic can provide an approximation to the inverse, the LU factorization, or the Cholesky decomposition. However, the accuracy of this approximate inverse or factorization depends on various parameters including the rank of the blocks and is, for the matrices arising from BEM, mathematically not fully understood. In the present paper we show that for a block structure typically employed in the context of $\mathcal{H}$-matrices, the inverse $\mathbf{W} = \mathbf{V}^{-1}$ of the discretization of the simple-layer operator V can be approximated from the set

of blockwise rank-r matrices at an exponential rate in the block rank. While this result does not fully analyze the accuracy of the $\mathcal{H}$-matrix inversion algorithms, it shows that inversion algorithms within the $\mathcal{H}$-matrix framework could work. It thus gives some mathematical underpinning to the success of the $\mathcal{H}$-matrix calculus when employed to compute (approximate) inverses of BEM-matrices, which is observed numerically, for example, by Bebendorf [Beb05] and Grasedyck [Gra01, Gra05].

Quickly after the introduction of the $\mathcal{H}$-matrix arithmetic, also $\mathcal{H}$-factorizations mimicking the classical LU- and Cholesky decompositions were proposed, [Lin04, Beb05]. Again, numerical evidence indicates their great usefulness for example, for black box preconditioning in iterative solvers, [Beb05, Gra05, GHK08, LBG06, GKLB08].

The class of $\mathcal{H}$-matrices is not the only one for which inversion and factorizations of system matrices arising in the discretization of differential and integral operators have been devised. Closely related to the concept of $\mathcal{H}$-matrices and its arithmetic are "hierarchically semiseparable matrices", [Xia13, XCGL09, LGWX12] and the idea of "recursive skeletonization", [HG12, GGMR09, HY13a]; for discretizations of PDEs, we mention [HY13a, GM13, SY12, Mar09], and particular applications to boundary integral equations are [MR05, CMZ13, HY13b]. These factorization algorithms aim to exploit that some off-diagonal blocks of certain Schur complements are low rank. Following [Beb07, GKLB09, CDGS10], we rigorously establish that the off-diagonal blocks of certain Schur complements can be approximated by low-rank matrices. We exploit this fact to show that the Cholesky decomposition of $\mathbf{V}$ can be approximated at an exponential rate in the block rank in the $\mathcal{H}$-matrix format.

Hitherto, the mathematical analysis of approximability of the inverse of system matrices in the $\mathcal{H}$-matrix format has focused on the setting of the finite element method (FEM). The first result in this direction is due to [BH03]. Generalizations to elliptic systems [Sch06] and approximations in the framework of $\mathcal{H}^2$-matrices [Bör10a, Bör10b] are also analyzed. Our recent works [FMP12, FMP13] differ from the above mentioned references for $\mathcal{H}$-matrices for FEM-matrices in several ways. Among the differences, we highlight that, as in present paper, [FMP12, FMP13] work in a fully discrete setting in contrast to the earlier technique of approximating on the continuous level and then projecting into discrete spaces. This technique avoids the projection error associated with the transition from the continuous level to the discrete one, and leads to exponential convergence in the block rank.

In the present paper, we focus on the lowest-order discretization of the simple-layer operator associated with the Laplace operator. However, our arguments are based on rather general properties of elliptic operators so that similar assertions can be shown for higher order discretizations and the hypersingular integral equation, which can be found in the forthcoming work [FMP14]. Moreover, we expect that our approach can cover the case of elliptic systems amenable to a treatment with the BEM such as the Lamé system.

The paper is structured as follows. In Section 2, we present the main results for $\mathcal{H}$-matrices. Mathematically, the core of the paper is Section 3, where we investigate the question of how well solutions of the discrete system can be approximated locally from low dimensional spaces. These results are transferred to the matrix level in Sections 4 and 5 to show the approximability result for $\mathbf{V}^{-1}$ and for Cholesky decompositions, respectively. Section 6 is concerned with various extensions of

our approximation result: We show a similar compression result for the Poincaré-Steklov operator and we show that $\mathbf{V}^{-1}$ can be approximated at an exponential rate in the format of $\mathcal{H}^2$-matrices, [HKS00, Bör10a, Bör10b].

2. MAIN RESULT

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded Lipschitz domain such that $\Gamma := \partial\Omega$ is polygonal (for $d = 2$) or polyhedral (for $d = 3$). We consider the simple-layer operator $V \in L(H^{-1/2}(\Gamma), H^{1/2}(\Gamma))$ associated with the Laplacian given by

$$V\phi(x) = \int_\Gamma G(x - y)\phi(y)ds_y, \quad x \in \Gamma,$$

where $G(x) = -\frac{1}{2\pi}\log|x|$ for $d = 2$ and $G(x) = \frac{1}{4\pi}\frac{1}{|x|}$ for $d = 3$ is the fundamental solution of the Laplacian. The simple-layer operator is an elliptic isomorphism for $d = 3$ and for $d = 2$ provided $\text{diam}(\Omega) < 1$, which can be assumed by scaling. We refer the reader to the monographs [McL00, HW08, SS11, Ste08] for a detailed discussion of the pertinent properties of boundary integral operators such as the simple-layer operator studied here.

We assume that Γ is triangulated by a *quasiuniform* mesh $\mathcal{T}_h = \{T_1, \ldots, T_N\}$ of mesh width $h := \max_{T_j \in \mathcal{T}_h} \text{diam}(T_j)$. The elements $T_j \in \mathcal{T}_h$ are open line segments ($d = 2$) or triangles ($d = 3$). Additionally, we assume that the mesh $\mathcal{T}_h$ is regular in the sense of Ciarlet and γ-shape regular in the sense that for $d = 2$ the quotient of the diameters of neighboring elements is bounded by γ and for $d = 3$ we have $\text{diam}(T_j) \leq \gamma |T_j|^{1/2}$ for all $T_j \in \mathcal{T}_h$. In the following, the notation $\lesssim$ abbreviates $\leq$ up to a constant $C > 0$ which depends only on Ω, the dimension d, and the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$. Moreover, we use $\simeq$ to indicate that both estimates $\lesssim$ and $\gtrsim$ hold.

We consider the lowest-order Galerkin discretization of V by piecewise constant functions in $S^{0,0}(\mathcal{T}_h) := \{u \in L^2(\Gamma) : u|_{T_j} \text{ is constant} \, \forall T_j \in \mathcal{T}_h\}$. Throughout, we will work with the basis $\mathcal{B}_h := \{\chi_j : j = 1, \ldots, N\}$ of the space $S^{0,0}(\mathcal{T}_h)$, where χ_j is the characteristic function associated with $T_j \in \mathcal{T}_h$. With the isomorphism $\Phi : \mathbb{R}^N \to S^{0,0}(\mathcal{T}_h)$, $\mathbf{x} \mapsto \sum_{j=1}^N \mathbf{x}_j \chi_j$, we note

$$(2.1) \qquad h^{d/2} \|\mathbf{x}\|_2 \lesssim \|\Phi(\mathbf{x})\|_{L^2(\Gamma)} \lesssim h^{d/2} \|\mathbf{x}\|_2 \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

With the basis $\mathcal{B}_h$, the Galerkin discretization of V leads to a symmetric and positive definite matrix $\mathbf{V} \in \mathbb{R}^{N \times N}$, where

$$(2.2) \qquad \mathbf{V}_{jk} = \langle V\chi_k, \chi_j \rangle = \int_{T_j}\int_{T_k} G(x - y)ds_y ds_x, \quad \chi_j, \chi_k \in \mathcal{B}_h,$$

and $\langle \cdot, \cdot \rangle$ denotes the $L^2(\Gamma)$-scalar product.

In the following, we study the approximability of the inverse BEM matrix $\mathbf{W} = \mathbf{V}^{-1}$ by some blockwise low-rank matrix $\mathbf{W}_\mathcal{H}$. First, we need to define the underlying block structure, which is based on the concept of "admissibility", introduced in the following definition.

Definition 2.1 (bounding boxes and η-admissibility). A *cluster* τ is a subset of the index set $\mathcal{I} = \{1, \ldots, N\}$. For a cluster $\tau \subset \mathcal{I}$, we say that $B_{R_\tau} \subset \mathbb{R}^d$ is a *bounding box* if:

    (i) B_{R_τ} is a hyper cube with side length R_τ,

(ii) $\overline{T_i} \subset B_{R_\tau}$ for all $i \in \tau$,

For $\eta > 0$, a pair of clusters (τ, σ) with $\tau, \sigma \subset \mathcal{I}$ is η-*admissible* if there exist bounding boxes B_{R_τ}, B_{R_σ} satisfying (i)–(ii) such that

$$(2.3) \qquad \min\{\mathrm{diam}(B_{R_\tau}), \mathrm{diam}(B_{R_\sigma})\} \leq \eta\, \mathrm{dist}(B_{R_\tau}, B_{R_\sigma}).$$

Remark 2.2. Since the operator V is symmetric, we are able to use the admissibility condition (2.3) instead of the stronger admissibility condition

$$(2.4) \qquad \max\{\mathrm{diam}(B_{R_\tau}), \mathrm{diam}(B_{R_\sigma})\} \leq \eta\, \mathrm{dist}(B_{R_\tau}, B_{R_\sigma}),$$

which is often encountered in clustering algorithms. This follows from the fact that Proposition 3.1 only needs an admissibility criterion of the form $\mathrm{diam}(B_{R_\tau}) \leq \eta\, \mathrm{dist}(B_{R_\tau}, B_{R_\sigma})$. Due to the symmetry of V, deriving a block approximation for the block $\tau \times \sigma$ is equivalent to deriving an approximation for the block $\sigma \times \tau$. Therefore, we can interchange roles of the boxes B_{R_τ} and B_{R_σ}, and as a consequence the weaker admissibility condition (2.3) is sufficient.

Definition 2.3 (blockwise rank-r matrices). Let P be a partition of $\mathcal{I} \times \mathcal{I}$ and $\eta > 0$. A matrix $\mathbf{W}_{\mathcal{H}} \in \mathbb{R}^{N \times N}$ is said to be a *blockwise rank-r matrix*, if for every η-admissible cluster pair $(\tau, \sigma) \in P$, the block $\mathbf{W}_{\mathcal{H}}|_{\tau \times \sigma}$ is a rank-r matrix, i.e., it has the form $\mathbf{W}_{\mathcal{H}}|_{\tau \times \sigma} = \mathbf{X}_{\tau\sigma}\mathbf{Y}_{\tau\sigma}^T$ with $\mathbf{X}_{\tau\sigma} \in \mathbb{R}^{|\tau| \times r}$ and $\mathbf{Y}_{\tau\sigma} \in \mathbb{R}^{|\sigma| \times r}$. Here and below, $|\sigma|$ denotes the cardinality of a finite set σ.

2.1. **Approximation of $\mathbf{V}^{-1}$.** The following Theorem 2.4 shows that admissible matrix blocks of $\mathbf{V}^{-1}$ can be approximated by rank-r matrices and the error converges exponentially in the block rank.

Theorem 2.4. *Fix the admissibility parameter $\eta > 0$ and $q \in (0,1)$. Let the cluster pair (τ, σ) be η-admissible. Then, for every $k \in \mathbb{N}$, there are matrices $\mathbf{X}_{\tau\sigma} \in \mathbb{R}^{|\tau| \times r}$, $\mathbf{Y}_{\tau\sigma} \in \mathbb{R}^{|\sigma| \times r}$ of rank $r \leq C_{\mathrm{dim}}(2 + \eta)^d q^{-d} k^{d+1}$ such that*

$$(2.5) \qquad \left\| \mathbf{V}^{-1}|_{\tau \times \sigma} - \mathbf{X}_{\tau\sigma}\mathbf{Y}_{\tau\sigma}^T \right\|_2 \leq C_{\mathrm{apx}} N^{(d+2)/(d-1)} q^k.$$

The constants $C_{\mathrm{apx}}, C_{\mathrm{dim}} > 0$ depend only on Ω, d, and the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$.

The approximation estimates for the individual blocks can be combined to assess the approximability of $\mathbf{V}^{-1}$ by blockwise rank-r matrices. Particularly satisfactory estimates are obtained if the blockwise rank-r matrices have additional structure. To that end, we introduce the following definitions.

Definition 2.5 (cluster tree). A *cluster tree* with *leaf size* $n_{\mathrm{leaf}} \in \mathbb{N}$ is a binary tree $\mathbb{T}_{\mathcal{I}}$ with root $\mathcal{I}$ such that for each cluster $\tau \in \mathbb{T}_{\mathcal{I}}$ the following dichotomy holds: either τ is a leaf of the tree and $|\tau| \leq n_{\mathrm{leaf}}$, or there exist so called sons $\tau', \tau'' \in \mathbb{T}_{\mathcal{I}}$, which are disjoint subsets of τ with $\tau = \tau' \cup \tau''$. The *level function* level $: \mathbb{T}_{\mathcal{I}} \to \mathbb{N}_0$ is inductively defined by $\mathrm{level}(\mathcal{I}) = 0$ and $\mathrm{level}(\tau') := \mathrm{level}(\tau) + 1$ for τ' a son of τ. The *depth* of a cluster tree is $\mathrm{depth}(\mathbb{T}_{\mathcal{I}}) := \max_{\tau \in \mathbb{T}_{\mathcal{I}}} \mathrm{level}(\tau)$.

Definition 2.6 (far field, near field, and sparsity constant). A partition P of $\mathcal{I} \times \mathcal{I}$ is said to be based on the cluster tree $\mathbb{T}_{\mathcal{I}}$, if $P \subset \mathbb{T}_{\mathcal{I}} \times \mathbb{T}_{\mathcal{I}}$. For such a partition P and fixed admissibility parameter $\eta > 0$, we define the *far field* and the *near field* as

$$(2.6) \qquad P_{\mathrm{far}} := \{(\tau, \sigma) \in P \,:\, (\tau, \sigma) \text{ is } \eta\text{-admissible}\}, \quad P_{\mathrm{near}} := P \backslash P_{\mathrm{far}}.$$

The *sparsity constant* C_{sp}, introduced in [HK00a, HK00b, Gra01], of such a partition is defined by

(2.7)
$$C_{\mathrm{sp}} := \max\left\{\max_{\tau \in \mathbb{T}_{\mathcal{I}}} |\{\sigma \in \mathbb{T}_{\mathcal{I}} \,:\, \tau \times \sigma \in P_{\mathrm{far}}\}|\,, \max_{\sigma \in \mathbb{T}_{\mathcal{I}}} |\{\tau \in \mathbb{T}_{\mathcal{I}} \,:\, \tau \times \sigma \in P_{\mathrm{far}}\}|\right\}.$$

The following Theorem 2.7 shows that the matrix $\mathbf{V}^{-1}$ can be approximated by blockwise rank-r matrices at an exponential rate in the block rank r:

Theorem 2.7. *Fix the admissibility parameter $\eta > 0$. Let a partition P of $\mathcal{I} \times \mathcal{I}$ be based on a cluster tree $\mathbb{T}_{\mathcal{I}}$. Then, there is a blockwise rank-r matrix $\mathbf{W}_{\mathcal{H}}$ such that*

(2.8)
$$\left\|\mathbf{V}^{-1} - \mathbf{W}_{\mathcal{H}}\right\|_2 \le C_{\mathrm{apx}} C_{\mathrm{sp}} N^{(d+2)/(d-1)} \mathrm{depth}(\mathbb{T}_{\mathcal{I}}) e^{-br^{1/(d+1)}}.$$

The constant C_{apx} depends only on Ω, d, and the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$, while the constant $b > 0$ additionally depends on η.

Remark 2.8. For quasiuniform meshes with $\mathcal{O}(N)$ elements, typical clustering strategies such as the "geometric clustering" described in [Hac09] lead to fairly balanced cluster trees $\mathbb{T}_{\mathcal{I}}$ of depth $\mathcal{O}(\log N)$ and a sparsity constant C_{sp} that is bounded uniformly in N. We refer to [HK00a, HK00b, Gra01, Hac09] for the fact that the memory requirement to store $\mathbf{W}_{\mathcal{H}}$ is $\mathcal{O}\big((r + n_{\mathrm{leaf}})N \log N\big)$.

Remark 2.9. Using $h \simeq N^{-1/(d-1)}$ and $\frac{1}{\|\mathbf{V}^{-1}\|_2} \le \|\mathbf{V}\|_2 \lesssim h^{(d-1)/2} \simeq N^{-1/2}$, where the last estimate can be found, e.g, in [Ste08, Lemma 12.6], we get a bound for the relative error

(2.9)
$$\frac{\left\|\mathbf{V}^{-1} - \mathbf{W}_{\mathcal{H}}\right\|_2}{\|\mathbf{V}^{-1}\|_2} \lesssim C_{\mathrm{apx}} C_{\mathrm{sp}} N^{(d+5)/(2d-2)} \mathrm{depth}(\mathbb{T}_{\mathcal{I}}) e^{-br^{1/(d+1)}}.$$

Remark 2.10. The approximation result of Theorem 2.7 is formulated in the spectral norm. In fact, inspection of the proof of Theorem 2.4 shows that we prove an approximation result in the weighted L^2-operator norm $\|\cdot\|_{\sqrt{h}L^2 \to \frac{1}{\sqrt{h}}L^2}$. Other norms such as the Frobenius-norm are possible, and the estimates change only by some powers of h. For the Frobenius norm, we can for instance employ the estimate $\|\mathbf{A}\|_2 \le \|\mathbf{A}\|_F \le \sqrt{N}\,\|\mathbf{A}\|_2$ for $\mathbf{A} \in \mathbb{R}^{N \times N}$.

2.2. **$\mathcal{H}$-Cholesky decomposition of V.** *LU-* and Cholesky decompositions are well-established tools of numerical linear algebra. Properties of these factorizations depend on the choice of the ordering of the unknowns. For the $\mathcal{H}$-Cholesky decomposition of Theorem 2.11 below we assume that the unknowns are organized in a binary cluster tree $\mathbb{T}_{\mathcal{I}}$. This induces an ordering of the unknowns by requiring that the unknowns of one of the sons are numbered first and those of the other son later; the precise numbering for the leaves is immaterial for our purposes. This induced ordering of the unknowns allows us to speak of *block lower triangular* matrices, if the block partition P is based on the cluster tree $\mathbb{T}_{\mathcal{I}}$. With this notation, we have the following factorization result:

Theorem 2.11. *Let $\mathbf{V} = \mathbf{C}\mathbf{C}^T$ be the Cholesky decomposition. Let a partition P of $\mathcal{I} \times \mathcal{I}$ be based on a cluster tree $\mathbb{T}_{\mathcal{I}}$. Then, there exist a block lower triangular, blockwise rank-r matrix $\mathbf{C}_{\mathcal{H}}$ such that*

(i) $\dfrac{\|\mathbf{C} - \mathbf{C}_{\mathcal{H}}\|_2}{\|\mathbf{C}\|_2} \le C_{\mathrm{chol}} N^{\frac{3}{2d-2}} \mathrm{depth}(\mathbb{T}_{\mathcal{I}}) e^{-br^{1/(d+1)}}$

(ii) $\dfrac{\left\|\mathbf{V} - \mathbf{C}_{\mathcal{H}}\mathbf{C}_{\mathcal{H}}{}^T\right\|_2}{\|\mathbf{V}\|_2} \leq 2C_{\mathrm{chol}}N^{\frac{3}{2d-2}}\mathrm{depth}(\mathbb{T}_{\mathcal{I}})e^{-br^{1/(d+1)}}$

$$+ C_{\mathrm{chol}}^2 N^{\frac{3}{d-1}}\mathrm{depth}(\mathbb{T}_{\mathcal{I}})^2 e^{-2br^{1/(d+1)}},$$

where $C_{\mathrm{chol}} = C_{\mathrm{sp}}C_{\mathrm{sc}}\sqrt{\kappa_2(\mathbf{V})}$, with the sparsity constant C_{sp} of (2.7), the spectral condition number $\kappa_2(\mathbf{V}) := \|\mathbf{V}\|_2 \left\|\mathbf{V}^{-1}\right\|_2$, and a constant C_{sc} depending only on Ω, d, the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$, and η.

3. Local approximation from low dimensional spaces

For a given function $f \in L^2(\Gamma)$, we consider the boundary integral equation

$$V\phi = f \quad \text{on } \Gamma.$$

Here, we may consider the simple-layer operator $V \in L(H^{-1}(\Gamma), L^2(\Gamma))$ as a mapping from $H^{-1}(\Gamma)$ to $L^2(\Gamma)$, see, e.g., [SS11]. The discrete variational problem is to find $\phi_h \in S^{0,0}(\mathcal{T}_h)$ such that

$$(3.1) \qquad \langle V\phi_h, \psi_h \rangle = \langle f, \psi_h \rangle \qquad \forall \psi_h \in S^{0,0}(\mathcal{T}_h).$$

With $\mathbf{V}$ from (2.2) and $\mathbf{b} \in \mathbb{R}^N$ defined by $\mathbf{b}_j = \langle f, \chi_j \rangle$, the variational problem (3.1) is equivalent to solving the linear system

$$(3.2) \qquad \mathbf{V}\mathbf{x} = \mathbf{b}.$$

By ellipticity of the simple-layer operator, both problems (3.1) and (3.2) have a unique solution. The solution $\mathbf{x} \in \mathbb{R}^N$ is linked to (3.1) via $\phi_h = \sum_{j=1}^{N} \mathbf{x}_j \chi_j$.

In the following, we repeatedly employ the $L^2(\Gamma)$-orthogonal projection $\Pi^{L^2} : L^2(\Gamma) \to S^{0,0}(\mathcal{T}_h)$ defined by

$$(3.3) \qquad \left\langle \Pi^{L^2}v, \psi_h \right\rangle = \langle v, \psi_h \rangle \quad \forall \psi_h \in S^{0,0}(\mathcal{T}_h).$$

The question of approximating the matrix block $\mathbf{V}^{-1}|_{\tau\times\sigma} \approx \mathbf{X}_{\tau\sigma}\mathbf{Y}_{\tau\sigma}^T$ can be rephrased in terms of functions and function spaces as the question of how well $\phi_h|_{B_{R_\tau}}$ can be approximated from low dimensional spaces for (arbitrary) data $f \in L^2(\Gamma)$ with $\mathrm{supp}\, f \subset B_{R_\sigma} \cap \Gamma$. The present section is devoted to the proof of such an approximation result formulated in the following Proposition 3.1.

Proposition 3.1. *Let (τ, σ) be a cluster pair with bounding boxes B_{R_τ}, B_{R_σ}. Assume $\eta\,\mathrm{dist}(B_{R_\tau}, B_{R_\sigma}) \geq \mathrm{diam}(B_{R_\tau})$ for some $\eta > 0$, and $R_\tau \leq 2\,\mathrm{diam}(\Omega)$. Fix $q \in (0,1)$. Then, for each $k \in \mathbb{N}$ there exists a subspace W_k of $S^{0,0}(\mathcal{T}_h)$ with $\dim W_k \leq C_{\mathrm{dim}}(2+\eta)^d q^{-d}k^{d+1}$ such that for arbitrary $f \in L^2(\Gamma)$ with $\mathrm{supp}\, f \subset B_{R_\sigma} \cap \Gamma$, the solution ϕ_h of (3.1) satisfies*

$$(3.4) \quad \min_{w \in W_k} \|\phi_h - w\|_{L^2(B_{R_\tau}\cap\Gamma)} \leq C_{\mathrm{box}}h^{-2}q^k\|\Pi^{L^2}f\|_{L^2(\Gamma)} \leq C_{\mathrm{box}}h^{-2}q^k\|f\|_{L^2(\Gamma)}.$$

The constants $C_{\mathrm{dim}}, C_{\mathrm{box}} > 0$ depend only on Ω, d, and the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$.

The proof of Proposition 3.1 will be given at the end of this section and relies on several observations. First, the potential

$$u(x) := \widetilde{V}\phi_h(x) = \int_\Gamma G(x - y)\phi_h(y)ds_y, \qquad x \in \mathbb{R}^d \setminus \Gamma,$$

generated by the solution ϕ_h of (3.1) is harmonic on Ω as well as on $\Omega^c := \mathbb{R}^d \setminus \overline{\Omega}$ and satisfies the jump conditions

$$[\gamma_0 u] := \gamma_0^{\text{ext}} u - \gamma_0^{\text{int}} u = 0 \in H^{1/2}(\Gamma),$$

(3.5) $$[\partial_n u] := \gamma_1^{\text{ext}} u - \gamma_1^{\text{int}} u = -\phi_h \in H^{-1/2}(\Gamma).$$

Here, $\gamma_0^{\text{ext}}, \gamma_0^{\text{int}}$ denote the exterior and interior trace operator and $\gamma_1^{\text{ext}}, \gamma_1^{\text{int}}$ the exterior and interior conormal derivative, see, e.g., [SS11]. Hence, the potential u is in a space of piecewise harmonic functions, and the jump of the normal derivative is piecewise constant on the mesh $\mathcal{T}_h$. These properties will characterize the spaces $\mathcal{H}_h(D)$ to be introduced below. The second observation is an orthogonality condition. For f with $\operatorname{supp} f \subset B_{R_\sigma} \cap \Gamma$ equation (3.1) implies

(3.6) $$\langle u, \psi_h \rangle = \langle f, \psi_h \rangle = 0 \quad \forall \psi_h \in S^{0,0}(\mathcal{T}_h) \text{ with } \operatorname{supp} \psi_h \subset \Gamma \setminus B_{R_\sigma}.$$

With the admissibility condition $\eta \operatorname{dist}(B_{R_\tau}, B_{R_\sigma}) \ge \min\{\operatorname{diam}(B_{R_\tau}), \operatorname{diam}(B_{R_\sigma})\} > 0$, this leads to the orthogonality condition

(3.7) $$\langle u, \psi_h \rangle = 0 \quad \forall \psi_h \in S^{0,0}(\mathcal{T}_h) \text{ with } \operatorname{supp} \psi_h \subset B_{R_\tau} \cap \Gamma,$$

i.e., on $B_{R_\tau} \cap \Gamma$ the potential u is orthogonal to piecewise constants.

   With these observations we are able to prove a Caccioppoli-type estimate (Lemma 3.9) for piecewise harmonic functions satisfying the orthogonality (3.7). Then, a low dimensional approximation result (Lemma 3.10) derived by Scott-Zhang interpolation of the Galerkin solution ϕ_h, can be iterated as in [BH03, Bör10a], which finally leads to exponential convergence (Lemma 3.11).

3.1. **Properties of piecewise polynomial spaces.** For an edge/face $T \subset \Gamma$ with affine parametrization ξ and $p \ge 0$, we let $P_p(T)$ be the space $P_p(T) := \{\pi \circ \xi|_T : \pi \in P_p(\mathbb{R}^{d-1})\}$ of polynomials of degree p. Moreover, we define $S^{p,1}(\mathcal{T}_h) := \{v \in C(\Gamma) : v|_T \in P_p(T) \ \forall T \in \mathcal{T}_h\}$ to be the space of all $\mathcal{T}_h$-piecewise polynomials of degree p that are continuous on Γ.

   Throughout this section we make use of the Scott-Zhang projection

$$I_h : H^1(\Gamma) \to S^{1,1}(\mathcal{T}_h)$$

introduced in [SZ90]. We note that the Scott-Zhang projection in [SZ90] is only defined for functions in $H^1(\Gamma)$, but by averaging only over triangles (and not over faces) it may also be well defined for functions in $L^2(\Gamma)$.

   By

$$\omega_T := \operatorname{interior}\left(\bigcup \{\overline{T'} \in \mathcal{T}_h : T \cap T' \neq \emptyset\}\right),$$

we denote the element patch of T, which contains T and all elements $T' \in \mathcal{T}_h$ that share a node with T. The operator I_h has well-known local approximation properties for H^ℓ-functions

(3.8) $$\|v - I_h v\|^2_{H^m(T)} \le C h^{2(\ell - m)} |v|^2_{H^\ell(\omega_T)}, \quad 0 \le m \le 1, \ m \le \ell \le 1.$$

The constant $C > 0$ depends only on the γ-shape regularity of $\mathcal{T}_h$ and the dimension d. With the triangle inequality this also implies the L^2-stability estimate

(3.9) $$\|I_h v\|_{L^2(T)} \le C \|v\|_{L^2(\omega_T)}.$$

   The following lemma constructs a stable operator on $L^2(\Gamma)$ that features additional orthogonality properties:

Lemma 3.2. *There exists a linear operator $\mathcal{J}_h : L^2(\Gamma) \to S^{d,1}(\mathcal{T}_h)$ with the following properties for all $v \in H^1(\Gamma)$*

   (i) $\|\mathcal{J}_h v\|_{L^2(T)} \leq C \|v\|_{L^2(\omega_T)} \quad \forall T \in \mathcal{T}_h$;
   (ii) $\|\nabla \mathcal{J}_h v\|_{L^2(T)} \leq C \|\nabla v\|_{L^2(\omega_T)} \quad \forall T \in \mathcal{T}_h$;
   (iii) $\langle v - \mathcal{J}_h v, \psi \rangle = 0 \quad \forall \psi \in S^{0,0}(\mathcal{T}_h)$;
   (iv) $\|v - \mathcal{J}_h v\|_{L^2(T)} \leq Ch \|\nabla v\|_{L^2(\omega_T)} \quad \forall T \in \mathcal{T}_h$.

The constant $C > 0$ depends only on d and the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$.

Proof. Let $b_T \in S^{d,1}(\mathcal{T}_h)$ be the element bubble function for each $T \in \mathcal{T}_h$, which is the product of the d hat-functions associated with T and scaled such that $\|b_T\|_\infty = 1$. Denote by χ_T the characteristic function of T. With the Scott-Zhang projection I_h, we define

$$\mathcal{J}_h v := I_h v + \sum_{T \in \mathcal{T}_h} b_T \frac{\langle v - I_h v, \chi_T \rangle}{\int_T b_T}.$$

For $T' \in \mathcal{T}_h$ we have

$$\langle v - \mathcal{J}_h v, \chi_{T'} \rangle = \left\langle v - I_h v - \sum_{T \in \mathcal{T}_h} b_T \frac{\langle v - I_h v, \chi_T \rangle}{\int_T b_T}, \chi_{T'} \right\rangle$$

$$= \langle v - I_h v, \chi_{T'} \rangle - \frac{\langle b_{T'}, \chi_{T'} \rangle}{\int_{T'} b_{T'}} \langle v - I_h v, \chi_{T'} \rangle = 0,$$

which proves (iii). The Cauchy-Schwarz inequality and the approximation property (3.8) of I_h imply

$$\|v - \mathcal{J}_h v\|_{L^2(T)} \leq \|v - I_h v\|_{L^2(T)} + \frac{\|b_T\|_{L^2(T)}}{|\int_T b_T|} |T|^{1/2} \|v - I_h v\|_{L^2(T)}$$

$$\lesssim \|v - I_h v\|_{L^2(T)} \lesssim h \|\nabla v\|_{L^2(\omega_T)}.$$

This proves (iv). The first assertion (i) follows with the same argument due to the L^2-stability of the Scott-Zhang projection I_h (3.9). Finally, we get

$$\|\nabla(v - \mathcal{J}_h v)\|_{L^2(T)} \leq \|\nabla(v - I_h v)\|_{L^2(T)} + \frac{\|\nabla b_T\|_{L^2(T)}}{|\int_T b_T|} |T|^{1/2} \|v - I_h v\|_{L^2(T)}$$

$$\lesssim \|\nabla(v - I_h v)\|_{L^2(T)} + \frac{1}{h} \|v - I_h v\|_{L^2(T)} \lesssim \|\nabla v\|_{L^2(\omega_T)},$$

and the triangle inequality finishes the proof of (ii). $\square$

The following inverse inequalities also holds for locally refined K-meshes, but we will only require it for the quasiuniform mesh $\mathcal{T}_h$ at hand.

Lemma 3.3 ([DFG$^+$01, Thm 4.1, Thm 4.7]). *There is a constant $C > 0$ depending only on Γ, the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$, and the polynomial degree p such that*

$$(3.10) \qquad \|v_h\|_{H^{1/2}(\Gamma)} \leq Ch^{-1/2} \|v_h\|_{L^2(\Gamma)} \qquad \forall v_h \in S^{p,1}(\mathcal{T}_h),$$

$$(3.11) \qquad \|v_h\|_{L^2(\Gamma)} \leq Ch^{-1/2} \|v_h\|_{H^{-1/2}(\Gamma)} \qquad \forall v_h \in S^{p,0}(\mathcal{T}_h).$$

3.2. The spaces $\mathcal{H}_h(D)$ and $\mathcal{H}_{h,0}(D,\Gamma_\rho)$ of piecewise harmonic functions.

For an index set $\rho \subset \mathcal{I}$, let $\Gamma_\rho \subset \Gamma$ be a relative open polygonal manifold, consisting of the union of elements in $\mathcal{T}_h$ associated with the elements in ρ, i.e.,

$$(3.12) \qquad \Gamma_\rho = \text{interior}\left(\bigcup_{j\in\rho} \overline{T_j}\right).$$

Let D be a domain and set $D^- := D \cap \Omega$ and $D^+ := D \cap \overline{\Omega}^c$. A function $v \in H^1(D^+ \cup D^-)$ is called piecewise harmonic, if

$$\int_{D^\pm} \nabla v \cdot \nabla \varphi \, dx = 0 \quad \forall \varphi \in C_0^\infty(D^\pm).$$

Remark 3.4. For a piecewise harmonic function $v \in H^1(D^+ \cup D^-)$, we can define the jump of the normal derivative $[\partial_n v]|_{D\cap\Gamma}$ on $D \cap \Gamma$ as the functional

$$(3.13) \qquad \langle[\partial_n v]|_{D\cap\Gamma}, \varphi\rangle := \int_{D^+\cup D^-} \nabla v \cdot \nabla \varphi dx \quad \forall \varphi \in H_0^1(D).$$

We note that the value $\langle[\partial_n v]|_{D\cap\Gamma}, \varphi\rangle$ depends only on $\varphi|_{D\cap\Gamma}$ in the sense that $\langle[\partial_n v]|_{D\cap\Gamma}, \varphi\rangle = 0$ for all $\varphi \in C_0^\infty(D)$ with $\varphi|_{D\cap\Gamma} = 0$. Moreover, if $[\partial_n v]|_{D\cap\Gamma}$ is a function in $L^2(D\cap\Gamma)$, it is unique. The definition (3.13) is consistent with (3.5) in the following sense: For the potential $V\phi_h$ with $\phi_h \in S^{0,0}(\mathcal{T}_h)$, we have the jump condition $[\partial_n V\phi_h]|_{D\cap\Gamma} = -\phi_h|_{D\cap\Gamma}$.

The space of piecewise harmonic functions on D with piecewise constant jump of the normal derivative is defined by

$$\mathcal{H}_h(D) := \{v \in H^1(D^+ \cup D^-): \ v \text{ is piecewise harmonic,}$$
$$\exists \widetilde{v} \in S^{0,0}(\mathcal{T}_h) \text{ s.t. } [\partial_n v]|_{D\cap\Gamma} = \widetilde{v}|_{D\cap\Gamma}\}.$$

The potential $u = \widetilde{V}\phi_h$ for the problem (3.1) indeed satisfies $u \in \mathcal{H}_h(D) \cap H^1(D)$ for any bounded domain D. Moreover, for a bounding box B_{R_σ} with (2.3), the potential u additionally satisfies the orthogonality condition (3.7). These observations are captured by the following space $\mathcal{H}_{h,0}(D,\Gamma_\rho)$:

$$\mathcal{H}_{h,0}(D,\Gamma_\rho) := \mathcal{H}_h(D) \cap \{v \in H^1(D): \ \text{supp}[\partial_n v]|_{D\cap\Gamma} \subset \overline{\Gamma_\rho},$$
$$(3.14) \qquad\qquad \langle v, \varphi\rangle = 0 \ \forall \varphi \in S^{0,0}(\mathcal{T}_h) \text{ with supp}\,\varphi \subset D \cap \overline{\Gamma_\rho}\}.$$

For the proof of Proposition 3.1 and subsequently of Theorem 2.4 and Theorem 2.7, we will only need the case $\Gamma_\rho = \Gamma$. The general case of the screen problem $\Gamma_\rho \subsetneq \Gamma$ will only be required for our analysis of the $\mathcal{H}$-Cholesky decomposition in Section 5.

The following lemma shows that this space is a closed subspace of $H^1(D \setminus \Gamma)$; later, this property will allow us to consider the orthogonal projection from $H^1(D\setminus\Gamma)$ onto $\mathcal{H}_h(D)$ and from $H^1(D)$ onto $\mathcal{H}_{h,0}(D,\Gamma_\rho)$.

Lemma 3.5. *The space $\mathcal{H}_h(D)$ is a closed subspace of $H^1(D \setminus \Gamma)$, and $\mathcal{H}_{h,0}(D,\Gamma_\rho)$ is a closed subspace of $H^1(D)$.*

Proof. We first show that $\mathcal{H}_h(D)$ is a closed subspace of $H^1(D \setminus \Gamma)$. Let $(v^j)_{j\in\mathbb{N}} \subset \mathcal{H}_h(D)$ be a sequence converging to $v \in H^1(D\setminus\Gamma)$. For $\varphi \in C_0^\infty(D^\pm)$, we have

$$\langle\nabla v, \nabla\varphi\rangle_{L^2(D^\pm)} = \lim_{j\to\infty} \langle\nabla v^j, \nabla\varphi\rangle_{L^2(D^\pm)} = 0.$$

Hence, v is piecewise harmonic on $D^+ \cup D^-$.

Pick $\varphi \in C_0^\infty(D) \setminus \{0\}$ with $\operatorname{supp} \varphi \cap \Gamma \subset T \in \mathcal{T}_h$. Then

$$[\partial_n v^j]|_T \langle 1, \varphi \rangle_{L^2(T)} = \langle [\partial_n v^j], \varphi \rangle_{L^2(T)} = \langle \nabla v^j, \nabla \varphi \rangle_{L^2(D \setminus \Gamma)} \overset{j \to \infty}{\longrightarrow} \langle \nabla v, \nabla \varphi \rangle_{L^2(D \setminus \Gamma)}$$

shows that the piecewise constant function $[\partial_n v^j]$ converges elementwise. Hence, the sequence $\left([\partial_n v^j] \right)_{j \in \mathbb{N}}$ converges pointwise to a piecewise constant function $\widetilde{v}$. This piecewise constant limit $\tilde{v}$ coincides with the jump of the normal derivative $[\partial_n v] \in H^{-1/2}(\Gamma)$ as the following calculation for arbitrary $\varphi \in C_0^\infty(D)$ shows:

$$\langle \widetilde{v}, \varphi \rangle_{L^2(D \cap \Gamma)} = \lim_{j \to \infty} \langle [\partial_n v^j], \varphi \rangle_{L^2(D \cap \Gamma)} = \langle \nabla v, \nabla \varphi \rangle_{L^2(D \setminus \Gamma)} = \langle [\partial_n v], \varphi \rangle_{L^2(D \cap \Gamma)}.$$

Finally, $\mathcal{H}_{h,0}(D, \Gamma_\rho)$ is a closed subspace of $H^1(D)$, since $\mathcal{H}_h(D)$ is a closed subspace of $H^1(D \setminus \Gamma)$ and the intersection of closed spaces is again closed. $\qquad \square$

We will derive an approximation of the Galerkin solution ϕ_h by approximating the potential $u = \widetilde{V} \phi_h$. In view of the relation $\phi_h = -[\partial_n u]$ we have to control the jump of the normal derivative by a norm of u. Lemma 3.8 below provides such an estimate, which may be seen as an inverse estimate, since $[\partial_n u]$ is a discrete function. For its proof, we need the following Lemma 3.7 as well as the definition of "concentric boxes".

Definition 3.6. Two (open) boxes B_R, $B_{R'}$ are said to be concentric boxes with side lengths R and R', if they have the same barycenter and B_R can be obtained by a stretching of $B_{R'}$ by the factor R/R' taking their common barycenter as the origin.

The following Lemma 3.7 is quite classical, and we include its short proof for the reader's convenience.

Lemma 3.7.   (i) For $R \leq 1$ denote by $S_R := \{ x \in \mathbb{R}^d : \operatorname{dist}(x, \Gamma) < R \}$ the tubular neighborhood of Γ of width R. Then, there is a constant $C > 0$ that depends only on Γ such that

$$\|v\|_{L^2(S_R)} \leq C \left[\sqrt{R} \|\gamma_0^{\mathrm{int}} v\|_{L^2(\Gamma)} + R \|\nabla v\|_{L^2(S_R)} \right] \qquad \forall v \in H^1(S_R).$$

(ii) Let δ, $R > 0$. Let B_R and $B_{(1+\delta)R}$ be two concentric boxes with side lengths R and $(1+\delta)R$. Then, there is a constant $C > 0$, which depends only on the dimension d, such that for all $v \in H^1(B_{(1+\delta)R})$, we have

$$\|v\|_{L^2(B_{(1+\delta)R} \setminus B_R)}^2 \leq C \delta R \left(\frac{1}{(1+\delta)R} \|v\|_{L^2(B_{(1+\delta)R})}^2 + (1+\delta)R \|\nabla v\|_{L^2(B_{(1+\delta)R})}^2 \right).$$

Proof. ad (i): For smooth univariate functions v the fundamental theorem of calculus yields $v(x) = v(0) + \int_0^x v'(t) \, dt$. Hence, the Young inequality and the Cauchy-Schwarz inequality yield $v^2(x) \leq 2v^2(0) + 2R \|v'\|_{L^2(0,R)}^2$. Integration over the interval $(0, R)$ gives

$$\|v\|_{L^2(0,R)}^2 \leq 2R v^2(0) + 2R^2 \|v'\|_{L^2(0,R)}^2.$$

This 1D result implies the desired estimate by using (locally) boundary fitted coordinates.

ad (ii): We start with the 1D Gagliardo-Nirenberg inequality for the interval $I = (0,1)$: $\|v\|_{L^\infty(I)}^2 \leq C \|v\|_{L^2(I)} \|v\|_{H^1(I)} \leq C \|v\|_{L^2(I)}^2 + C \|v\|_{L^2(I)} \|v'\|_{L^2(I)}$. A

scaling argument then yields

$$\|v\|^2_{L^\infty(0,(1+\delta)R)} \lesssim \frac{1}{(1+\delta)R}\|v\|^2_{L^2(0,(1+\delta)R)} + \|v\|_{L^2(0,(1+\delta)R)}\|v'\|_{L^2(0,(1+\delta)R)}$$
$$\lesssim \frac{1}{(1+\delta)R}\|v\|^2_{L^2(0,(1+\delta)R)} + (1+\delta)R\|v'\|^2_{L^2(0,(1+\delta)R)}.$$

We may assume that $B_{(1+\delta)R} = (0,(1+\delta)R)^d$. Then, this 1D estimate implies

$$\|v\|^2_{L^2((0,\delta R)\times(0,(1+\delta)R)^{d-1})} \lesssim \frac{\delta R}{(1+\delta)R}\|v\|^2_{L^2(B_{(1+\delta)R})}$$
$$+ \delta R(1+\delta)R\|\nabla v\|^2_{L^2(B_{(1+\delta)R})}.$$

By arguing similarly for the remaining parts of $B_{(1+\delta)R} \setminus B_R$, we get the desired result. $\square$

Lemma 3.8. *Let $\delta \in (0,1)$, $R > 0$ be such that $\frac{h}{R} \leq \frac{\delta}{4}$. Let B_R, $B_{(1+\delta)R}$ be two concentric boxes of side lengths R and $(1+\delta)R$. Then, there exists a constant $C > 0$ depending only on Ω, d, and the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$, such that for all $v \in \mathcal{H}_h(B_{(1+\delta)R})$*

$$(3.15) \qquad \|[\partial_n v]\|_{L^2(B_{(1+\delta/2)R}\cap\Gamma)} \leq Ch^{-1/2}\|\nabla v\|_{L^2(B_{(1+\delta)R})}.$$

Proof. We prove (3.15) in two steps, the first step being the proof of the auxiliary estimate (3.16) below. The second step shows (3.15) with the aid of (3.16) and a simple covering argument.

Step 1: We show the following assertion: If $\frac{h}{r} \leq \frac{\varepsilon}{4}$ for $r, \varepsilon > 0$, then there exists a constant $C > 0$ depending only on the shape regularity constant γ, the domain Ω, and d such that for all $v \in \mathcal{H}_h(B_{(1+\varepsilon)r})$

$$(3.16) \qquad \|[\partial_n v]\|_{L^2(B_r\cap\Gamma)} \leq Ch^{-1/2}\sqrt{1 + \frac{1}{\varepsilon}}\|\nabla v\|_{L^2(B_{(1+\varepsilon)r})}.$$

To see this, let $\mathcal{E}^{\text{int}} : H^{1/2}(\Gamma) \to H^1(\Omega)$ and $\mathcal{E}^{\text{ext}} : H^{1/2}(\Gamma) \to H^1(\overline{\Omega}^c)$ be (bounded, linear) lifting operators for Ω and Ω^c (cf. [Neč67, Thm. 5.7]). Then, introduce the (bounded, linear) lifting $\mathcal{L} : H^{1/2}(\Gamma) \to H^1(\mathbb{R}^d)$ by

$$\mathcal{L}w := \begin{cases} \mathcal{E}^{\text{int}}w & \text{in } \overline{\Omega}, \\ \mathcal{E}^{\text{ext}}w & \text{in } \overline{\Omega}^c. \end{cases}$$

Lemma 3.2 provides an operator $\mathcal{J}_h : L^2(\Gamma) \to S^{d,1}(\mathcal{T}_h) \subset H^1(\Gamma)$. Furthermore, $w - \mathcal{J}_h w$ is orthogonal to piecewise constant functions so that

$$(3.17) \qquad \|[\partial_n v]\|_{L^2(B_r\cap\Gamma)} = \sup_{\substack{w\in L^2(\Gamma)\\ \text{supp } w \subset B_r}} \frac{\langle[\partial_n v], w\rangle}{\|w\|_{L^2(\Gamma)}} = \sup_{\substack{w\in L^2(\Gamma)\\ \text{supp } w \subset B_r}} \frac{\langle[\partial_n v], \mathcal{J}_h w\rangle}{\|w\|_{L^2(\Gamma)}}.$$

Note that the construction of $\mathcal{J}_h$ implies $\text{supp}\,\mathcal{J}_h w \subset B_{(1+\varepsilon/2)r} \cap \Gamma$. Let η be a smooth cut-off function with $0 \leq \eta \leq 1$, $\eta \equiv 1$ on $B_{(1+\varepsilon/2)r}$, and $\text{supp}\,\eta \subset B_{(1+\varepsilon)r}$

and $\|\nabla\eta\|_{L^\infty(B_{(1+\varepsilon)r})} \lesssim \frac{1}{\varepsilon r}$. With the lifting $\mathcal{L}(\mathcal{J}_h w)$, we can estimate

$$
\begin{aligned}
\langle[\partial_n v], \mathcal{J}_h w\rangle_{L^2(B_{r+2h}\cap\Gamma)} &= \langle[\partial_n v], \gamma_0^{\mathrm{int}}\eta\mathcal{L}\mathcal{J}_h w\rangle_{L^2(B_{r+2h}\cap\Gamma)} \\
&= \int_{B_{(1+\varepsilon)r}} \nabla v \cdot \nabla(\eta\mathcal{L}\mathcal{J}_h w)dx \\
&\leq \|\nabla v\|_{L^2(B_{(1+\varepsilon)r})} \|\nabla(\eta\mathcal{L}\mathcal{J}_h w)\|_{L^2(B_{(1+\varepsilon)r})}.
\end{aligned}
$$

(3.18)

We have to estimate $\eta\mathcal{L}\mathcal{J}_h w$ further. Noting that $\eta \equiv 1$ on B_r, we use the product rule to estimate

(3.19) $\|\nabla(\eta\mathcal{L}\mathcal{J}_h w)\|_{L^2(B_{(1+\varepsilon)r})} \lesssim \frac{1}{\varepsilon r}\|\mathcal{L}\mathcal{J}_h w\|_{L^2(B_{(1+\varepsilon)r}\setminus B_r)} + \|\nabla\mathcal{L}\mathcal{J}_h w\|_{L^2(B_{(1+\varepsilon)r})}.$

The continuity of the lifting $\mathcal{L}: H^{1/2}(\Gamma) \to H^1(\mathbb{R}^d)$ and the inverse estimate (3.10) of Lemma 3.3 give

$$
\begin{aligned}
\|\nabla\mathcal{L}\mathcal{J}_h w\|_{L^2(S_{\sqrt{d}(1+\varepsilon)r})} &\leq \|\nabla\mathcal{L}\mathcal{J}_h w\|_{L^2(\mathbb{R}^d)} \\
&\lesssim \|\mathcal{J}_h w\|_{H^{1/2}(\Gamma)} \lesssim \frac{1}{\sqrt{h}}\|\mathcal{J}_h w\|_{L^2(\Gamma)},
\end{aligned}
$$

(3.20)

which is the key step for the treatment of the second term in (3.19). Let us now turn to the first term in (3.19). Using Lemma 3.7, (ii) and then Lemma 3.7, (i) with the observation $B_{(1+\varepsilon)r} \subset S_{\sqrt{d}(1+\varepsilon)r}$, we get

$$
\begin{aligned}
\frac{1}{\varepsilon r}\|\mathcal{L}\mathcal{J}_h w\|_{L^2(B_{(1+\varepsilon)r}\setminus B_r)} &\lesssim \frac{1}{\sqrt{\varepsilon r}\sqrt{(1+\varepsilon)r}}\|\mathcal{L}\mathcal{J}_h w\|_{L^2(B_{(1+\varepsilon)r})} \\
&\quad + \frac{\sqrt{(1+\varepsilon)r}}{\sqrt{\varepsilon r}}\|\nabla\mathcal{L}\mathcal{J}_h w\|_{L^2(B_{(1+\varepsilon)r})} \\
&\lesssim \frac{1}{\sqrt{\varepsilon r}}\|\mathcal{L}\mathcal{J}_h w\|_{L^2(\Gamma)} + \frac{\sqrt{(1+\varepsilon)r}}{\sqrt{\varepsilon r}}\|\nabla\mathcal{L}\mathcal{J}_h w\|_{L^2(S_{\sqrt{d}(1+\varepsilon)r})}.
\end{aligned}
$$

We have $\mathcal{L}\mathcal{J}_h w|_\Gamma = \mathcal{J}_h w$, and (3.20) leads to

(3.21) $\frac{1}{\varepsilon r}\|\mathcal{L}\mathcal{J}_h w\|_{L^2(B_{(1+\varepsilon)r}\setminus B_r)} \lesssim \frac{1}{\sqrt{\varepsilon r}}\|\mathcal{J}_h w\|_{L^2(\Gamma)} + h^{-1/2}\sqrt{1+\varepsilon^{-1}}\|\mathcal{J}_h w\|_{L^2(\Gamma)}.$

We note that $\varepsilon r \geq 4h$ so that $1/\sqrt{\varepsilon r} \lesssim h^{-1/2}$. Inserting (3.20) and (3.21) in (3.19) and using the $L^2(\Gamma)$-stability of $\mathcal{J}_h$ given by Lemma 3.2, we obtain

$$
\|\nabla(\eta\mathcal{L}\mathcal{J}_h w)\|_{L^2(B_{(1+\varepsilon)r})} \lesssim h^{-1/2}\sqrt{1+\varepsilon^{-1}}\|\mathcal{J}_h w\|_{L^2(\Gamma)} \lesssim h^{-1/2}\sqrt{1+\varepsilon^{-1}}\|w\|_{L^2(\Gamma)}.
$$

Finally, inserting this bound into (3.18) and then into (3.17) allows us to conclude the proof of (3.16).

Step 2: The bound (3.15) is shown with the aid of (3.16) and a covering argument. We may assume that $B_R = (0, R)^d$. Set $r = \delta R$. Let $n \in \mathbb{N}$ be given by $n = \lceil R/r\rceil$. Let x_i, $i = 1, \ldots, (n+1)^d =: N$ be the points of a regular grid in the closed box $\overline{B_R}$ with spacing R/n. For $i = 1, \ldots, N$ consider the boxes $B_i := x_i + (-r/2, r/2)^d$ as well as the scaled boxes $\widehat{B}_i := x_i + (-r, r)^d$, $i = 1, \ldots, N$. The essential properties of these boxes are: first, the boxes B_i, $i = 1, \ldots, N$ cover $B_{(1+\delta/2)R}$; secondly, the scaled boxes $\widehat{B}_i$, $i = 1, \ldots, N$ are contained in $B_{(1+\delta)R}$; thirdly, and most importantly, they have a finite overlap property (with an overlap constant that depends solely on the spatial dimension d, since the ratio of the spacing R/n and the side length r satisfies $r/(R/n) = (1/\delta)/\lceil 1/\delta\rceil \in [1/2, 1]$ for

the case $\delta \in (0,1)$ under consideration here). Observing $\frac{h}{r} = \frac{h}{\delta R} \leq \frac{1}{4}$ due to our assumption $\frac{h}{R} \leq \frac{\delta}{4}$, the estimate (3.16) implies for each i

$$\|[\partial_n v]\|_{L^2(B_i \cap \Gamma)} \leq Ch^{-1/2}\|\nabla v\|_{L^2(\widehat{B}_i)}.$$

The desired estimate (3.15) follows from the covering and overlap properties.   □

For a box B_R with side length R, we introduce the norm

$$\|v\|_{h,R}^2 := \left(\frac{h}{R}\right)^2 \|\nabla v\|_{L^2(B_R)}^2 + \frac{1}{R^2}\|v\|_{L^2(B_R)}^2,$$

which is, for fixed h, equivalent to the H^1-norm.

Similarly as in [BH03, Bör10a], a main part of the proof is a Caccioppoli-type inequality, which is, for functions in $\mathcal{H}_{h,0}(B_{(1+\delta)R}, \Gamma_\rho)$, stated in the following lemma.

Lemma 3.9. *Let $\delta \in (0,1)$ and $\frac{h}{R} \leq \frac{\delta}{16}$ and let $\Gamma_\rho \subset \Gamma$ be of the form (3.12). Let B_R, $B_{(1+\delta)R}$ be two concentric boxes. Then, there exists a constant $C > 0$ depending only on Ω, d, and the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$, such that for $v \in \mathcal{H}_{h,0}(B_{(1+\delta)R}, \Gamma_\rho)$*

$$(3.22) \qquad \|\nabla v\|_{L^2(B_R)} \leq C\frac{1+\delta}{\delta}\|v\|_{h,(1+\delta)R}.$$

Proof. The proof of (3.22) is done in two steps.
Step 1: We show that for $\varepsilon > 0$ with $\frac{h}{R} \leq \frac{\varepsilon}{8}$, the estimate

$$(3.23) \qquad \|\nabla v\|_{L^2(B_R)}^2 \lesssim \frac{h}{\varepsilon R}\|\nabla v\|_{L^2(B_{(1+\varepsilon)R})}^2 + \frac{1}{(\varepsilon R)^2}\|v\|_{L^2(B_{(1+\varepsilon)R})}^2.$$

holds. To see this, let η be a smooth cut-off function with supp $\eta \subset B_{(1+\varepsilon/4)R}$ and $\eta \equiv 1$ on B_R, and $\|\nabla \eta\|_{L^\infty(B_{(1+\varepsilon)R})} \lesssim \frac{1}{\varepsilon R}$. We will need a second smooth cut-off function $\widetilde{\eta}$ with supp $\widetilde{\eta} \subset B_{(1+\varepsilon)R}$ and $\widetilde{\eta} \equiv 1$ on $B_{(1+\varepsilon/2)R}$ and $\|\nabla \widetilde{\eta}\|_{L^\infty(B_{(1+\varepsilon)R})} \lesssim \frac{1}{\varepsilon R}$. Since h is the maximal element diameter, $8h \leq \varepsilon R$ implies $T \subset B_{(1+\varepsilon/2)R}$ for all $T \in \mathcal{T}_h$ with $T \cap \text{supp}\, \eta \neq \emptyset$. Integration by parts, the fact that v is piecewise harmonic and $\text{supp}([\partial_n v]|_{B_{(1+\varepsilon)R} \cap \Gamma}) \subset \overline{\Gamma_\rho}$ lead to

$$\|\nabla(\eta v)\|_{L^2(B_{(1+\varepsilon)R})}^2 = \int_{B_{(1+\varepsilon)R}} \nabla(\eta v) \cdot \nabla(\eta v)\, dx$$

$$= \int_{B_{(1+\varepsilon)R}} \nabla v \cdot \nabla(\eta^2 v) + v^2 |\nabla \eta|^2\, dx$$

$$= \int_{\Gamma \cap B_{(1+\varepsilon)R}} \eta^2 [\partial_n v] v\, ds_x + \int_{B_{(1+\varepsilon)R}} v^2 |\nabla \eta|^2\, dx$$

$$(3.24) \qquad = \int_\Gamma \eta^2 [\partial_n v] v\, ds_x + \int_{B_{(1+\varepsilon)R}} v^2 |\nabla \eta|^2\, dx,$$

where in the last step we used the support property supp $\eta \subset B_{(1+\varepsilon/4)R}$ to extend the function $\eta^2 [\partial_n v] v$, which is defined on $\Gamma \cap B_{(1+\varepsilon)R}$, by zero to the whole set Γ. We first focus on the surface integral in (3.24). With the $L^2(\Gamma)$-orthogonal projection Π^{L^2} onto $S^{0,0}(\mathcal{T}_h)$ from (3.3), we get by definition of the space $\mathcal{H}_{h,0}(B_{(1+\varepsilon)R}, \Gamma_\rho)$

that $\operatorname{supp}\Pi^{L^2}(\eta^2[\partial_n v]) \subset \overline{\Gamma_\rho} \cap B_{(1+\varepsilon)R}$. Therefore, we can use the orthogonality (3.14) satisfied by v to get

$$
\begin{aligned}
\langle \eta^2[\partial_n v], v\rangle &= \langle \eta^2[\partial_n v] - \Pi^{L^2}(\eta^2[\partial_n v]), v\rangle \\
&= \langle \eta^2[\partial_n v] - \Pi^{L^2}(\eta^2[\partial_n v]), \widetilde\eta^2 v\rangle \\
&= \langle \eta^2[\partial_n v] - \Pi^{L^2}(\eta^2[\partial_n v]), \widetilde\eta^2 v - \Pi^{L^2}(\widetilde\eta^2 v)\rangle,
\end{aligned}
$$

(3.25)

where we were able to insert the cut-off function $\widetilde\eta$ since $\widetilde\eta \equiv 1$ on $\operatorname{supp}(\eta^2[\partial_n v] - \Pi^{L^2}(\eta^2[\partial_n v])) \subset \overline{(B_{(1+\varepsilon/2)R} \setminus B_{R-2h}) \cap \Gamma_\rho}$. With these observations in hand, we estimate

$$
\begin{aligned}
\|\eta^2[\partial_n v] - \Pi^{L^2}(\eta^2[\partial_n v])\|_{L^2(\Gamma)}^2 &\lesssim \sum_{T \in \mathcal{T}_h} h^2 \|\nabla_\Gamma(\eta^2[\partial_n v])\|_{L^2(T)}^2 \\
&\lesssim C\frac{h^2}{(\varepsilon R)^2} \|[\partial_n v]\|_{L^2((B_{(1+\varepsilon/2)R}) \cap \Gamma_\rho)}^2.
\end{aligned}
$$

The standard approximation property $\|\widetilde\eta^2 v - \Pi^{L^2}(\widetilde\eta^2 v)\|_{L^2(\Gamma)} \lesssim h^{1/2}\|\widetilde\eta^2 v\|_{H^{1/2}(\Gamma)}$ and the bound (3.15) of Lemma 3.8 as well as the trace inequality for Γ give

$$
\begin{aligned}
\left|\langle \eta^2[\partial_n v], v\rangle\right| &\lesssim \frac{h}{\varepsilon R}\|[\partial_n v]\|_{L^2((B_{(1+\varepsilon/2)R}) \cap \Gamma_\rho)} h^{1/2}\|\widetilde\eta^2 v\|_{H^{1/2}(\Gamma)} \\
&\lesssim \frac{h}{\varepsilon R}\|\nabla v\|_{L^2(B_{(1+\varepsilon)R})}\|\widetilde\eta^2 v\|_{H^{1/2}(\Gamma)} \lesssim \frac{h}{\varepsilon R}\|\nabla v\|_{L^2(B_{(1+\varepsilon)R})}\|\widetilde\eta^2 v\|_{H^1(\Omega)}.
\end{aligned}
$$

(3.26)

The properties of $\widetilde\eta$ imply $\|\widetilde\eta^2 v\|_{H^1(\Omega)} \lesssim \|\nabla v\|_{L^2(B_{(1+\varepsilon)R})} + (\varepsilon R)^{-1}\|v\|_{L^2(B_{(1+\varepsilon)R})}$. Inserting this into (3.26) and the result into (3.24) yields

$$
\begin{aligned}
\|\nabla(\eta v)\|_{L^2(B_{(1+\varepsilon)R})}^2 &\lesssim \frac{h}{\varepsilon R}\|\nabla v\|_{L^2(B_{(1+\varepsilon)R})}^2 + \frac{h}{(\varepsilon R)^2}\|\nabla v\|_{L^2(B_{(1+\varepsilon)R})}\|v\|_{L^2(B_{(1+\varepsilon)R})} \\
&\quad + \frac{1}{(\varepsilon R)^2}\|v\|_{L^2(B_{(1+\varepsilon)R})}^2 \\
&\lesssim \frac{h}{\varepsilon R}\|\nabla v\|_{L^2(B_{(1+\varepsilon)R})}^2 + \frac{1}{(\varepsilon R)^2}\|v\|_{L^2(B_{(1+\varepsilon)R})}^2,
\end{aligned}
$$

where we employed an appropriate Young inequality in the last step and $h/(\varepsilon R) \leq 1$. This implies (3.23).

Step 2: Starting from estimate (3.23) with $\varepsilon = \frac{\delta}{2}$, we use (3.23) again with $\varepsilon = \frac{\delta}{2+\delta}$ and $\widetilde R = (1+\delta/2)R$. Since $\left(1+\frac{\delta}{2}\right)\left(1+\frac{\delta}{2+\delta}\right) = 1+\delta$ and $\frac{h}{R} \leq \frac{\delta}{16}$ implies $\frac{h}{R} \leq \frac{\varepsilon}{8}$, we arrive at

$$
\begin{aligned}
\|\nabla v\|_{L^2(B_R)}^2 &\lesssim \frac{h}{\delta R}\|\nabla v\|_{L^2(B_{(1+\delta/2)R})}^2 + \frac{1}{(\delta R)^2}\|v\|_{L^2(B_{(1+\delta/2)R})}^2 \\
&\lesssim \left(\frac{h}{\delta R}\right)^2\|\nabla v\|_{L^2(B_{(1+\delta)R})}^2 + \left(\frac{h}{(\delta R)^3} + \frac{1}{(\delta R)^2}\right)\|v\|_{L^2(B_{(1+\delta)R})}^2,
\end{aligned}
$$

and with $h/(\delta R) \leq 1$ we conclude the proof. $\qquad\square$

3.3. **Low-dimensional approximation in** $\mathcal{H}_{h,0}(D, \Gamma_\rho)$. Since $\mathcal{H}_{h,0}(B_R, \Gamma_\rho) \subset H^1(B_R)$ is a closed subspace by Lemma 3.5, the orthogonal projection $\Pi_{h,R} : (H^1(B_R), \|\cdot\|_{h,R}) \to (\mathcal{H}_{h,0}(B_R, \Gamma_\rho), \|\cdot\|_{h,R})$ is well-defined.

Lemma 3.10. *Let $\delta \in (0,1)$, $R > 0$ such that $\frac{h}{R} \leq \frac{\delta}{16}$ and B_R, $B_{(1+\delta)R}$, $B_{(1+2\delta)R}$ be concentric boxes. Let $\Gamma_\rho \subset \Gamma$ be of the form (3.12) and $v \in \mathcal{H}_{h,0}(B_{(1+2\delta)R}, \Gamma_\rho)$. Let $\mathcal{K}_H$ be an (infinite) γ-shape regular triangulation of $\mathbb{R}^d$ of mesh width H and assume $\frac{H}{R} \leq \frac{\delta}{4}$ for the corresponding mesh width H. Let $I_H : H^1(\mathbb{R}^d) \to S^{1,1}(\mathcal{K}_H)$ be the Scott-Zhang projection. Then, there exists a constant $C_{\mathrm{app}} > 0$ that depends only on Ω, d, and γ, such that*

(i) $\big(v - \Pi_{h,R} I_H v\big)|_{B_R} \in \mathcal{H}_{h,0}(B_R, \Gamma_\rho)$;

(ii) $\|v - \Pi_{h,R} I_H v\|_{h,R} \leq C_{\mathrm{app}} \frac{1+2\delta}{\delta} \left(\frac{h}{R} + \frac{H}{R}\right) \|v\|_{h,(1+2\delta)R}$;

(iii) $\dim W \leq C_{\mathrm{app}} \left(\frac{(1+2\delta)R}{H}\right)^d$, *where* $W := \Pi_{h,R} I_H \mathcal{H}_{h,0}(B_{(1+2\delta)R}, \Gamma_\rho)$.

Proof. Since $v \in \mathcal{H}_{h,0}(B_{(1+2\delta)R}, \Gamma_\rho)$ implies $v \in \mathcal{H}_{h,0}(B_R, \Gamma_\rho)$, we have that $\Pi_{h,R}(v|_{B_R}) = v|_{B_R}$, which proves (i).

The assumption $\frac{H}{R} \leq \frac{\delta}{4}$ implies $\bigcup\{K \in \mathcal{K}_H : K \cap B_R \neq \emptyset\} \subset B_{(1+\delta)R}$. Then, the locality and approximation properties (3.8) of the Scott-Zhang projection I_H yield

$$\frac{1}{H} \|v - I_H v\|_{L^2(B_R)} + \|\nabla(v - I_H v)\|_{L^2(B_R)} \lesssim \|\nabla v\|_{L^2(B_{(1+\delta)R})}.$$

We apply Lemma 3.9 with $\widetilde{R} = (1+\delta)R$ and $\widetilde{\delta} = \frac{\delta}{1+\delta}$. Note that $(1+\widetilde{\delta})\widetilde{R} = (1+2\delta)R$. Moreover, $16h \leq \delta R = \widetilde{\delta}\widetilde{R}$ implies $\frac{h}{R} \leq \frac{\widetilde{\delta}}{16}$. Therefore, we get

$$\|v - \Pi_{h,R} I_H v\|_{h,R}^2 = \|\Pi_{h,R}(v - I_H v)\|_{h,R}^2 \leq \|v - I_H v\|_{h,R}^2$$

$$= \left(\frac{h}{R}\right)^2 \|\nabla(v - I_H v)\|_{L^2(B_R)}^2 + \frac{1}{R^2} \|v - I_H v\|_{L^2(B_R)}^2$$

$$\lesssim \frac{h^2}{R^2} \|\nabla v\|_{L^2(B_{(1+\delta)R})}^2 + \frac{H^2}{R^2} \|\nabla v\|_{L^2(B_{(1+\delta)R})}^2$$

$$\lesssim \left(\frac{1+\delta}{\delta}\left(\frac{h}{R} + \frac{H}{R}\right)\right)^2 \|v\|_{h,(1+2\delta)R}^2,$$

which concludes the proof of (ii). Finally, the statement (iii) follows from the fact that $\dim I_H(\mathcal{H}_{h,0}(B_{(1+2\delta)R}, \Gamma_\rho)) \lesssim ((1+2\delta)R/H)^d$. $\square$

The property (i) of Lemma 3.10 can be used to iterate the approximation result (ii) on suitable concentric boxes. This will allow us to construct a subspace of $\mathcal{H}_{h,0}(B_{(1+\kappa)R}, \Gamma_\rho)$ for $\kappa \in (0,1)$ with the capability to approximate at an exponential rate.

Lemma 3.11. *Let C_{app} be the constant of Lemma 3.10. Let $q, \kappa \in (0,1)$, $R > 0$, $k \in \mathbb{N}$ and $\Gamma_\rho \subset \Gamma$ be of the form (3.12). Assume*

$$(3.27) \qquad \frac{h}{R} \leq \frac{\kappa q}{64k \max\{1, C_{\mathrm{app}}\}}.$$

Then, there exists a finite dimensional subspace $\widehat{W}_k$ of $\mathcal{H}_{h,0}(B_{(1+\kappa)R}, \Gamma_\rho)$ with dimension

$$\dim \widehat{W}_k \leq C_{\dim} \left(\frac{1 + \kappa^{-1}}{q}\right)^d k^{d+1}$$

such that for every $v \in \mathcal{H}_{h,0}(B_{(1+\kappa)R}, \Gamma_\rho)$

$$(3.28) \quad \min_{\widehat{w} \in \widehat{W}_k} \sqrt{h} \, \|[\partial_n v] - [\partial_n \widehat{w}]\|_{L^2(B_R \cap \Gamma_\rho)} \leq \min_{\widehat{w} \in \widehat{W}_k} \sqrt{h} \, \|[\partial_n v] - [\partial_n \widehat{w}]\|_{L^2(B_R \cap \Gamma)}$$

$$\leq C_{\text{low}} \frac{R}{h} \min_{\widehat{w} \in \widehat{W}_k} \|v - \widehat{w}\|_{h,(1+\kappa/2)R} \leq C_{\text{low}} \frac{R}{h} q^k \|v\|_{h,(1+\kappa)R}.$$

The constants C_{dim}, $C_{\text{low}} > 0$ depend only on Ω, d, and the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$.

Proof. Let B_R and $B_{(1+\delta_j)R}$, with $\delta_j := \kappa(1 - \frac{j}{2k})$ for $j = 0, \ldots, k$ be concentric boxes. We have $\kappa = \delta_0 > \delta_1 > \cdots > \delta_k = \frac{\kappa}{2}$. In the following, we iterate the approximation result of Lemma 3.10 on the boxes $B_{(1+\delta_j)R}$. Choosing $H = \frac{\kappa q R}{64k \max\{C_{\text{app}}, 1\}}$, we have $h \leq H$. We apply Lemma 3.10 with $\widetilde{R}_j = (1+\delta_j)R$ and $\widetilde{\delta}_j = \frac{\kappa}{4k(1+\delta_j)} < \frac{1}{4}$. Note that $\delta_{j-1} = \delta_j + \frac{\kappa}{2k}$ gives $(1+\delta_{j-1})R = (1+2\widetilde{\delta}_j)\widetilde{R}_j$ and our choice of H implies $\frac{H}{\widetilde{R}_j} \leq \frac{\widetilde{\delta}_j}{4}$. Hence, for $j = 1$, Lemma 3.10 provides an approximation w_1 in a subspace W_1 of $\mathcal{H}_{h,0}(B_{(1+\delta_1)R}, \Gamma_\rho)$ with $\dim W_1 \leq C_{\text{app}} \left(\frac{(1+2\widetilde{\delta}_1)\widetilde{R}_1}{H} \right)^d = C_{\text{app}} \left(\frac{(1+\kappa)R}{H} \right)^d$, satisfying

$$\|v - w_1\|_{h,(1+\delta_1)R} = \|v - w_1\|_{h,\widetilde{R}_1} \leq C_{\text{app}} \frac{1+2\widetilde{\delta}_1}{\widetilde{\delta}_1} \left(\frac{h}{\widetilde{R}_1} + \frac{H}{\widetilde{R}_1} \right) \|v\|_{h,(1+2\widetilde{\delta}_1)\widetilde{R}_1}$$

$$\leq 2C_{\text{app}} \frac{H}{(1+\delta_1)R} \frac{1+2\widetilde{\delta}_1}{\widetilde{\delta}_1} \|v\|_{h,(1+\delta_0)R}$$

$$= 8C_{\text{app}} \frac{kH}{\kappa R}(1+2\widetilde{\delta}_1) \|v\|_{h,(1+\kappa)R} \leq q \|v\|_{h,(1+\kappa)R}.$$

Since $v|_{B_{(1+\delta_1)R}} - w_1 \in \mathcal{H}_{h,0}(B_{(1+\delta_1)R}, \Gamma_\rho)$, Lemma 3.10 can be applied to $v - w_1$, and provides an approximation w_2 of $v - w_1$ in a subspace W_2 of $\mathcal{H}_{h,0}(B_{(1+\delta_2)R}, \Gamma_\rho)$ with $\dim W_2 \leq C_{\text{app}} \left(\frac{(1+\kappa)R}{H} \right)^d$. Arguing as for $j = 1$, we get

$$\|v - w_1 - w_2\|_{h,(1+\delta_2)R} \leq q \|v - w_1\|_{h,(1+\delta_1)R} \leq q^2 \|v\|_{h,(1+\kappa)R}.$$

Continuing this process $k-2$ times, one obtains an approximation $\widehat{w} := \sum_{j=1}^k w_i$ in the space $\widehat{W}_k := \sum_{j=1}^k W_j$ of dimension $\dim \widehat{W}_k \leq C_{\text{app}} k \left(\frac{(1+\kappa)R}{H} \right)^d \leq C_{\text{dim}}((1+\kappa^{-1})q^{-1})^d k^{d+1}$ with

$$\|v - \widehat{w}\|_{h,(1+\kappa/2)R} = \|v - \widehat{w}\|_{h,(1+\delta_k)R} \leq q^k \|v\|_{h,(1+\kappa)R}.$$

Finally, since (3.27) ensures $h/R \leq \kappa/8$, we may use Lemma 3.8 to estimate

$$\sqrt{h} \|[\partial_n v] - [\partial_n \widehat{w}]\|_{L^2(B_R \cap \Gamma)} \leq C \|\nabla(v - \widehat{w})\|_{L^2(B_{(1+\kappa/2)R})} \leq C \frac{R}{h} \|v - \widehat{w}\|_{h,(1+\kappa/2)R}$$

to conclude the argument. $\square$

Now we are able to prove the main result of this section, Proposition 3.1.

Proof of Proposition 3.1. Due to the admissibility condition and the definition of bounding boxes, we have $\text{dist}(B_{R_\tau}, B_{R_\sigma}) \geq \eta^{-1} \text{diam}(B_{R_\tau}) = \eta^{-1}\sqrt{d}R_\tau$. The

choice $\kappa = \frac{1}{1+\eta}$ implies

$$\operatorname{dist}(B_{(1+\kappa)R_\tau}, B_{R_\sigma}) \geq \operatorname{dist}(B_{R_\tau}, B_{R_\sigma}) - \kappa R_\tau \sqrt{d}$$
$$\geq \sqrt{d}R_\tau(\eta^{-1} - \kappa) = \sqrt{d}R_\tau \left(\frac{1}{\eta} - \frac{1}{1+\eta}\right) > 0.$$

The potential $u = \widetilde{V}\phi_h$ with the Galerkin solution $\phi_h \in S^{0,0}(\mathcal{T}_h)$ of (3.1) then satisfies $u \in \mathcal{H}_{h,0}(B_{(1+\kappa)R_\tau}, \Gamma)$. The inverse estimate $\sqrt{h}\,\|\phi_h\|_{L^2(\Gamma)} \lesssim \|\phi_h\|_{H^{-1/2}(\Gamma)}$ of (3.11) and the ellipticity of the simple-layer operator as well as the discrete boundary integral equation (3.1) provide

$$\|\phi_h\|_{H^{-1/2}(\Gamma)}^2 \lesssim \langle V\phi_h, \phi_h \rangle = \langle f, \phi_h \rangle = \left\langle \Pi^{L^2} f, \phi_h \right\rangle \lesssim \left\|\Pi^{L^2} f\right\|_{L^2(\Gamma)} \|\phi_h\|_{L^2(\Gamma)}$$
$$\lesssim h^{-1/2} \left\|\Pi^{L^2} f\right\|_{L^2(\Gamma)} \|\phi_h\|_{H^{-1/2}(\Gamma)}.$$

Then, the boundedness of $\widetilde{V} : H^{-1/2}(\Gamma) \to H^1_{\mathrm{loc}}(\mathbb{R}^d)$ and $\frac{h}{R_\tau} < 1$ lead to

$$\left\|\!\left\|\widetilde{V}\phi_h\right\|\!\right\|_{h,R_\tau(1+\kappa)} \leq 2\left(1 + \frac{1}{R_\tau}\right)\left\|\widetilde{V}\phi_h\right\|_{H^1(B_{2R_\tau})} \lesssim \left(1 + \frac{1}{R_\tau}\right)\|\phi_h\|_{H^{-1/2}(\Gamma)}$$
$$\lesssim \left(1 + \frac{1}{R_\tau}\right) h^{-1/2} \left\|\Pi^{L^2} f\right\|_{L^2(\Gamma)}.$$

After these preparations, we are in a position to define the space W_k, for which we distinguish two cases.

Case 1: The condition (3.27) is satisfied with $R = R_\tau$.

With the space $\widehat{W}_k$ provided by Lemma 3.11 we set $W_k := \{[\partial_n \widehat{w}] : \widehat{w} \in \widehat{W}_k\}$. Then, Lemma 3.11 and $R_\tau \leq 2\operatorname{diam}(\Omega)$ lead to

$$\min_{w \in W_k} \|\phi_h - w\|_{L^2(B_{R_\tau} \cap \Gamma)} \lesssim \frac{R_\tau}{h^{3/2}} q^k \left\|\!\left\|\widetilde{V}\phi_h\right\|\!\right\|_{h,(1+\kappa)R_\tau} \lesssim (R_\tau + 1)h^{-2}q^k \left\|\Pi^{L^2} f\right\|_{L^2(\Gamma)}$$
$$\lesssim h^{-2}q^k \left\|\Pi^{L^2} f\right\|_{L^2(\Gamma)},$$

and the dimension of W_k is bounded by

$$\dim W_k \leq C_{\dim}\left(\frac{1+\kappa^{-1}}{q}\right)^d k^{d+1} = C_{\dim}(2+\eta)^d q^{-d} k^{d+1}.$$

Case 2: The condition (3.27) is not satisfied.

Then, we select $W_k := \left\{w|_{B_{R_\tau} \cap \Gamma} : w \in S^{0,0}(\mathcal{T}_h)\right\}$ and the minimum in (3.4) is obviously zero. By the choice of κ and $\frac{h}{R} > \frac{\kappa q}{64k \max\{1, C_{\mathrm{app}}\}}$, the dimension of W_k is bounded by

$$\dim W_k \lesssim \left(\frac{R_\tau}{h}\right)^{d-1} \lesssim \left(\frac{64k \max\{C_{\mathrm{app}}, 1\}}{\kappa q}\right)^{d-1} \simeq \left((1+\eta)q^{-1}k\right)^{d-1}$$
$$\lesssim (2+\eta)^d q^{-d} k^{d+1}.$$

This concludes the proof of the first inequality in (3.4). The second inequality in (3.4) follows from the $L^2(\Gamma)$-stability of the $L^2(\Gamma)$-orthogonal projection. $\qquad\square$

4. Proof of the approximation results for $\mathbf{V}^{-1}$

In this section, the approximation result given in Proposition 3.1 is used to construct a low-rank approximation of a matrix block $\mathbf{V}^{-1}|_{\tau \times \sigma}$ and in turn an $\mathcal{H}$-matrix approximation of $\mathbf{V}^{-1}$. This is achieved with local variants of the isomorphism (2.1), and our arguments follow the lines of [Bör10a, Theorem 2].

Proof of Theorem 2.4. If $C_{\dim}(2 + \eta)^d q^{-d} k^{d+1} \geq \min(|\tau|, |\sigma|)$, we use the exact matrix block $\mathbf{X}_{\tau\sigma} = \mathbf{V}^{-1}|_{\tau \times \sigma}$ and $\mathbf{Y}_{\tau\sigma} = \mathbf{I} \in \mathbb{R}^{|\sigma| \times |\sigma|}$.

If $C_{\dim}(2 + \eta)^d q^{-d} k^{d+1} < \min(|\tau|, |\sigma|)$, we employ the approximation result of Proposition 3.1 in the following way. For the cluster $\tau \subset \mathcal{I}$, we define $\mathbb{R}^\tau := \left\{ \mathbf{x} \in \mathbb{R}^N : x_i = 0 \, \forall i \notin \tau \right\}$ and the mappings

$$\Phi_\tau : \mathbb{R}^\tau \to S^{0,0}(\mathcal{T}_h), \quad \mathbf{x} \mapsto \sum_{j \in \tau} x_j \chi_j, \quad \text{and} \quad \Lambda_\tau : L^2(\Gamma) \to \mathbb{R}^{|\tau|}, \quad w \mapsto (\overline{w}_i)_{i \in \tau},$$

where $\overline{w}_i$ denotes the mean value on the element $T_i \in \mathcal{T}_h$. Hence, for a piecewise constant function the mapping Λ_τ returns the constant value on each element corresponding to the cluster τ. Moreover, $\Phi_\tau \Lambda_\tau$ is the restriction of the L^2-projection onto $S^{0,0}(\mathcal{T}_h)$ to $\Gamma_\tau := \text{interior}\left(\bigcup_{i \in \tau} \overline{T_i}\right) \subset B_{R_\tau}$. Thus, in particular, for a piecewise constant function $\widetilde{\phi} \in S^{0,0}(\mathcal{T}_h)$ we get $\Phi_\tau(\Lambda_\tau \widetilde{\phi}) = \widetilde{\phi}|_{\Gamma_\tau}$. For $\mathbf{x} \in \mathbb{R}^\tau$, (2.1) implies

$$C h^{d/2} \|\mathbf{x}\|_2 \leq \|\Phi_\tau(\mathbf{x})\|_{L^2(\Gamma)} \leq \widetilde{C} h^{d/2} \|\mathbf{x}\|_2, \quad \mathbf{x} \in \mathbb{R}^\tau.$$

The adjoint $\Lambda_{\mathcal{I}}^* : \mathbb{R}^N \to L^2(\Gamma)' \simeq L^2(\Gamma), \mathbf{b} \mapsto \sum_{i \in \mathcal{I}} b_i(u \mapsto \overline{u}_i)$ of $\Lambda_{\mathcal{I}}$ satisfies, because of (2.1) and the L^2-stability of $\Phi_{\mathcal{I}} \Lambda_{\mathcal{I}}$,

$$\|\Lambda_{\mathcal{I}}^* \mathbf{b}\|_{L^2(\Gamma)} = \sup_{w \in L^2(\Gamma)} \frac{\langle \mathbf{b}, \Lambda_{\mathcal{I}} w \rangle_2}{\|w\|_{L^2(\Gamma)}} \lesssim \|\mathbf{b}\|_2 \sup_{w \in L^2(\Gamma)} \frac{h^{-d/2} \|\Phi_{\mathcal{I}} \Lambda_{\mathcal{I}} w\|_{L^2(\Gamma)}}{\|w\|_{L^2(\Gamma)}}$$
$$\lesssim h^{-d/2} \|\mathbf{b}\|_2.$$

Let $\mathbf{b} \in \mathbb{R}^N$. Defining $f := \Lambda_{\mathcal{I}}^* \mathbf{b}|_\sigma$, we get $b_i = \langle f, \chi_i \rangle$ for $i \in \sigma$, and $\text{supp} f \subset B_{R_\sigma} \cap \Gamma$. Proposition 3.1 provides a finite dimensional space W_k and an element $w \in W_k$ that is a good approximation to the Galerkin solution $\phi_h|_{B_{R_\tau} \cap \Gamma}$. It is important to note that the space W_k is constructed independently of the function f; it depends only on the cluster pair (τ, σ). The estimate (2.1), the approximation result from Proposition 3.1, and $\left\|\Pi^{L^2} f\right\|_{L^2(\Gamma)} = \|\Lambda_{\mathcal{I}}^* \mathbf{b}\|_{L^2(\Gamma)} \lesssim h^{-d/2} \|\mathbf{b}\|_2$ imply

$$\|\Lambda_\tau \phi_h - \Lambda_\tau w\|_2 \lesssim h^{-d/2} \|\Phi_\tau(\Lambda_\tau \phi_h - \Lambda_\tau w)\|_{L^2(\Gamma)} \lesssim h^{-d/2} \|\phi_h - w\|_{L^2(B_{R_\tau} \cap \Gamma)}$$
$$\lesssim h^{-d/2-2} q^k \left\|\Pi^{L^2} f\right\|_{L^2(\Gamma)} \lesssim h^{-(d+2)} q^k \|\mathbf{b}\|_2.$$

In order to translate this approximation result to the matrix level, let

$$\mathcal{W} := \{\Lambda_\tau w \, : \, w \in W_k\}.$$

Let the columns of $\mathbf{X}_{\tau\sigma}$ be an orthogonal basis of the space $\mathcal{W}$. Then, the rank of $\mathbf{X}_{\tau\sigma}$ is bounded by $C_{\dim}(2+\eta)^d q^{-d} k^{d+1}$. Since $\mathbf{X}_{\tau\sigma} \mathbf{X}_{\tau\sigma}^T$ is the orthogonal projection from $\mathbb{R}^N$ onto $\mathcal{W}$, we get that $z := \mathbf{X}_{\tau\sigma} \mathbf{X}_{\tau\sigma}^T \Lambda_\tau \phi_h$ is the best approximation of $\Lambda_\tau \phi_h$ in $\mathcal{W}$ and arrive at

$$(4.1) \quad \|\Lambda_\tau \phi_h - z\|_2 \leq \|\Lambda_\tau \phi_h - \Lambda_\tau w\|_2 \lesssim h^{-(d+2)} q^k \|\mathbf{b}\|_2 \simeq N^{(d+2)/(d-1)} q^k \|\mathbf{b}\|_2.$$

Noting that $\Lambda_\tau \phi_h = \mathbf{V}^{-1}|_{\tau \times \sigma} \mathbf{b}|_\sigma$, if we define $\mathbf{Y}_{\tau,\sigma} := \mathbf{V}^{-1}|_{\tau \times \sigma}^T \mathbf{X}_{\tau\sigma}$, we thus get $z = \mathbf{X}_{\tau\sigma} \mathbf{Y}_{\tau\sigma}^T \mathbf{b}|_\sigma$. The bound (4.1) expresses

$$(4.2) \qquad \left\| \left(\mathbf{V}^{-1}|_{\tau \times \sigma} - \mathbf{X}_{\tau\sigma} \mathbf{Y}_{\tau\sigma}^T \right) \mathbf{b}|_\sigma \right\|_2 \lesssim N^{(d+2)/(d-1)} q^k \left\| \mathbf{b} \right\|_2.$$

The space W_k depends only on the cluster pair (τ, σ) and the estimate (4.2) is valid for any $\mathbf{b}$. This concludes the proof. $\qquad \square$

The following lemma gives an estimate for the global spectral norm by the local spectral norms.

Lemma 4.1 ([Gra01],[Hac09, Lemma 6.5.8]). *Let $\mathbf{M} \in \mathbb{R}^{N \times N}$ and P be a partitioning of $\mathcal{I} \times \mathcal{I}$. Then,*

$$\|\mathbf{M}\|_2 \le C_{\mathrm{sp}} \left(\sum_{\ell=0}^{\infty} \max\{\|\mathbf{M}|_{\tau \times \sigma}\|_2 : (\tau, \sigma) \in P, \mathrm{level}(\tau) = \ell\} \right),$$

where the sparsity constant C_{sp} is defined in (2.7).

Now we are able to prove our main result, Theorem 2.7.

Proof of Theorem 2.7. Theorem 2.4 provides matrices $\mathbf{X}_{\tau\sigma} \in \mathbb{R}^{|\tau| \times r}$, $\mathbf{Y}_{\tau\sigma} \in \mathbb{R}^{|\sigma| \times r}$, so we can define the $\mathcal{H}$-matrix $\mathbf{V}_{\mathcal{H}}$ by

$$\mathbf{W}_{\mathcal{H}} = \begin{cases} \mathbf{X}_{\tau\sigma} \mathbf{Y}_{\tau\sigma}^T & \text{if } (\tau, \sigma) \in P_{\mathrm{far}}, \\ \mathbf{V}^{-1}|_{\tau \times \sigma} & \text{otherwise.} \end{cases}$$

On each admissible block $(\tau, \sigma) \in P_{\mathrm{far}}$ we can use the blockwise estimate of Theorem 2.4 and get

$$\left\| (\mathbf{V}^{-1} - \mathbf{W}_{\mathcal{H}})|_{\tau \times \sigma} \right\|_2 \le C_{\mathrm{apx}} N^{(d+2)/(d-1)} q^k.$$

On inadmissible blocks, the error is zero by definition. Therefore, Lemma 4.1 leads to

$$\left\| \mathbf{V}^{-1} - \mathbf{W}_{\mathcal{H}} \right\|_2 \le C_{\mathrm{sp}} \left(\sum_{\ell=0}^{\infty} \max\{ \left\| (\mathbf{V}^{-1} - \mathbf{V}_{\mathcal{H}})|_{\tau \times \sigma} \right\|_2 : (\tau, \sigma) \in P, \mathrm{level}(\tau) = \ell \} \right)$$

$$\le C_{\mathrm{apx}} C_{\mathrm{sp}} N^{(d+2)/(d-1)} q^k \, \mathrm{depth}(\mathbb{T}_{\mathcal{I}}).$$

With $r = C_{\mathrm{dim}}(2 + \eta)^d q^{-d} k^{d+1}$, defining $b = -\frac{\ln(q)}{C_{\mathrm{dim}}^{1/(d+1)}} q^{d/(d+1)} (2+\eta)^{-d/(1+d)} > 0$ leads to $q^k = e^{-br^{1/(d+1)}}$, and hence

$$\left\| \mathbf{V}^{-1} - \mathbf{W}_{\mathcal{H}} \right\|_2 \le C_{\mathrm{apx}} C_{\mathrm{sp}} N^{(d+2)/(d-1)} \mathrm{depth}(\mathbb{T}_{\mathcal{I}}) e^{-br^{1/(d+1)}},$$

which concludes the proof. $\qquad \square$

5. $\mathcal{H}$-CHOLESKY DECOMPOSITION: PROOF OF THEOREM 2.11

The aim of this section is the proof of Theorem 2.11. Our procedure follows [Beb07, GKLB09, FMP13] and is based on showing that the off-diagonal block of certain Schur complements can be approximated by low-rank matrices. The analysis of these Schur complement matrices in Section 5.1 is therefore the main contribution of the section.

The matrix $\mathbf{V}$ is symmetric and positive definite and therefore has a (classical) Cholesky-decomposition $\mathbf{V} = \mathbf{C}\mathbf{C}^T$, where $\mathbf{C}$ is a lower triangular matrix. Moreover, the existence of the Cholesky decomposition does not depend on the numbering of the degrees of freedom, i.e., for every other numbering of the basis

functions there is a Cholesky decomposition as well (see, e.g., [HJ13, Cor. 3.5.6]). The existence of the Cholesky decomposition implies the invertibility of the matrix $\mathbf{V}|_{\rho \times \rho}$ for any $n \leq N$ and index set $\rho := \{1, \ldots, n\}$ (see, e.g., [HJ13, Cor. 3.5.6]).

The first step is the approximation of appropriate Schur complements.

5.1. **Schur complements.** For a cluster pair (τ, σ), we define the index set $\rho := \{i \in \mathcal{I} : i < \min(\tau \cup \sigma)\}$ and the Schur complement

$$(5.1) \qquad \mathbf{S}(\tau, \sigma) = \mathbf{V}|_{\tau \times \sigma} - \mathbf{V}|_{\tau \times \rho}(\mathbf{V}|_{\rho \times \rho})^{-1}\mathbf{V}|_{\rho \times \sigma}.$$

One way to approximate the Schur complement is to use the $\mathcal{H}$-arithmetic. As stated in [GKLB09, Theorem 15], this results in a low-rank approximation to $\mathbf{S}(\tau, \sigma)$ of rank $C_{\mathrm{id}}C_{\mathrm{sp}}(\mathrm{depth}(\mathbb{T}_{\mathcal{I}}) + 1)^2 r$, where the idempotency constant C_{id} is defined in [GH03], and r is the block rank used for the approximation of the inverse matrix $\mathbf{V}^{-1}$. In the following Theorem 5.2, we provide a low-rank approximation by using a different approach, which uses the techniques developed in Section 3 and gives a better bound in terms of the rank of the approximation, i.e., a rank of Cr is sufficient to obtain the same accuracy. This approach relies on interpreting Schur complements as BEM matrices from certain constrained spaces.

The key step is Theorem 5.2 below. For its proof, we need a degenerate approximation of the Green's function $G(\cdot, \cdot)$. This is a classical result that underlies the log-linear matrix-vector multiplication in BEM and can be achieved by multipole expansions [Rok85, GR97], Taylor expansions [NH88, HN89, Sau92, HS93] or by interpolation (see, e.g., [SS11, Sec. 7.1.3.1]). The following lemma recalls a variant of such a degenerate approximation that is obtained with Chebyshev interpolation:

Lemma 5.1. *Let $\widetilde{\eta} > 0$ and fix $\eta' \in (0, 2\widetilde{\eta})$. Then, for every hyper cube $B_Y \subset \mathbb{R}^d$, $d \in \{2, 3\}$ and closed $D_X \subset \mathbb{R}^d$ with $\mathrm{dist}(B_Y, D_X) \geq \widetilde{\eta}\,\mathrm{diam}(B_Y)$ the following is true: For every $r \in \mathbb{N}$ there exist functions $g_{1,i}$, $g_{2,i}$, $i = 1, \ldots, r$, such that*

$$\left\| G(x, \cdot) - \sum_{i=1}^{r} g_{1,i}(x)g_{2,i}(\cdot) \right\|_{L^\infty(B_Y)} \leq C\frac{(1 + 1/\widetilde{\eta})}{\mathrm{dist}(\{x\}, B_Y)^{d-2}}(1 + \eta')^{-r^{1/d}} \quad \forall x \in D_X,$$

for a constant C that depends solely on the choice of $\eta' \in (0, 2\widetilde{\eta})$.

Proof. Let $I_k^y : C(\overline{B_Y}) \to \mathcal{Q}_k$ be the tensor product interpolation operator of degree k defined on $C(\overline{B_Y})$ and mapping into the space $\mathcal{Q}_k$ of polynomial of degree k in each variable. Note that $\dim \mathcal{Q}_k = (k+1)^d =: r$. The approximation $G_r(x, y) := \sum_{i=1}^{r} g_{1,i}(x)g_{2,i}(y)$ is then taken to be $G_r(x, \cdot) := I_k^y G(x, \cdot)$. The stated error bound follows from estimates for Chebyshev interpolation. We note that the Green's function for the Laplacian is asymptotically smooth (see [Hac09, Definition 4.2.5] with constant $c_{\mathrm{as}}(\nu) = C\nu!$). Tensorial interpolation in the form given in [BG04] allows us to estimate

$$\|G(x, \cdot) - I_k^y G(x, \cdot)\|_{L^\infty(B_Y)} \lesssim \frac{1}{\mathrm{dist}(\{x\}, B_Y)^{d-2}}\left(1 + \frac{\mathrm{diam}(B_Y)}{\mathrm{dist}(B_Y, \{x\})}\right)\Lambda_k^d r^{1/d}$$

$$\cdot \left(1 + \frac{2\,\mathrm{dist}(B_Y, \{x\})}{\mathrm{diam}(B_Y)}\right)^{-r^{1/d}},$$

where $\Lambda_k \leq 1 + \frac{2}{\pi}\ln(k+1)$ is the Lebesgue constant of Chebyshev interpolation, cf. [Riv74]. The observation $\mathrm{dist}(\{x\}, B_Y) \geq \mathrm{dist}(B_Y, D_X) \geq \widetilde{\eta}\,\mathrm{diam}(B_Y)$ and the choice $\eta' < 2\widetilde{\eta}$ imply the claimed estimate. $\qquad\square$

Theorem 5.2. *Let (τ, σ) be an η-admissible cluster pair, set $\rho := \{i \in \mathcal{I} : i < \min(\tau \cup \sigma)\}$, and let the Schur complement $\mathbf{S}(\tau, \sigma)$ be defined in (5.1). Then, there exists a rank-r matrix $\mathbf{S}_r(\tau, \sigma)$ such that*

$$\|\mathbf{S}(\tau, \sigma) - \mathbf{S}_r(\tau, \sigma)\|_2 \le C_{\mathrm{sc}} N^{3/(2d-2)} e^{-br^{1/(d+1)}} \|\mathbf{V}\|_2,$$

where the constants $C_{\mathrm{sc}}, b > 0$ depend only on Ω, d, the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$, and η.

Proof. Let B_{R_τ}, B_{R_σ} be bounding boxes for the clusters τ, σ satisfying (2.3). We define $\Gamma_\rho = \mathrm{interior}\left(\bigcup_{i \in \rho} \mathrm{supp}\,\psi_i\right) \subset \Gamma$. First, we recall that the Schur complement matrix $\mathbf{S}(\tau, \sigma)$ can be understood in terms of an orthogonalization with respect to the degrees of freedom in ρ. More precisely, a direct calculation (see for the essentials, e.g., [Bre99]) shows for $\phi \in \mathbb{R}^{|\tau|}$, $\psi \in \mathbb{R}^{|\sigma|}$ the representation

$$(5.2) \qquad \phi^T \mathbf{S}(\tau, \sigma)\psi = \left\langle V\widetilde{\phi}, \psi \right\rangle,$$

with the following relation between the functions ψ, $\widetilde{\phi}$ and the vectors ψ, ϕ, respectively: $\psi = \sum_{j=1}^{|\sigma|} \psi_j \chi_{j_\sigma}$, where the index j_σ denotes the j-th basis function corresponding to the cluster σ, and the function $\widetilde{\phi} \in S^{0,0}(\mathcal{T}_h)$ is defined by $\widetilde{\phi} = \phi + \phi_\rho$ with $\phi = \sum_{j=1}^{|\tau|} \phi_j \chi_{j_\tau}$ and $\mathrm{supp}\,\phi_\rho \subset \overline{\Gamma}_\rho$ such that

$$(5.3) \qquad \left\langle V\widetilde{\phi}, \widehat{\psi} \right\rangle = 0 \quad \forall \widehat{\psi} \in S^{0,0}(\mathcal{T}_h) \text{ with } \mathrm{supp}\,\widehat{\psi} \subset \overline{\Gamma}_\rho.$$

Our low-rank approximation of the Schur complement matrix $\mathbf{S}(\tau, \sigma)$ will have two ingredients: first, based on the the techniques of Section 3 we exploit the orthogonality (5.3) to construct a low-dimensional space $\widehat{W}_k$ from which for any ϕ, the corresponding function $\widetilde{\phi}$ can be approximated well. Second, we exploit that the function ψ in (5.2) is supported by Γ_σ, and we will use Lemma 5.1.

Let $\delta = \frac{1}{1+\eta}$ and B_{R_σ}, $B_{(1+\delta)R_\sigma}$ be concentric boxes. The symmetry of V leads to

$$(5.4) \qquad \begin{aligned} \left\langle V\widetilde{\phi}, \psi \right\rangle &= \left\langle \widetilde{\phi}, V\psi \right\rangle \\ &= \left\langle \widetilde{\phi}, V\psi \right\rangle_{L^2(B_{(1+\delta)R_\sigma} \cap \Gamma_\rho)} + \left\langle \widetilde{\phi}, V\psi \right\rangle_{L^2(\Gamma \setminus B_{(1+\delta)R_\sigma})}. \end{aligned}$$

First, we treat the first term on the right-hand side of (5.4). The choice of δ and the admissibility condition (2.3), where we can assume $\min\{\mathrm{diam}(B_{R_\tau}), \mathrm{diam}(B_{R_\sigma})\} = \sqrt{d}R_\sigma$ due to the symmetry $\mathbf{S}(\tau, \sigma) = \mathbf{S}(\sigma, \tau)^T$, imply

$$\mathrm{dist}(B_{(1+2\delta)R_\sigma}, B_{R_\tau}) \ge \mathrm{dist}(B_{R_\sigma}, B_{R_\tau}) - \sqrt{d}\delta R_\sigma \ge \sqrt{d}R_\sigma(\eta^{-1} - \delta) > 0.$$

Therefore, we have $\widetilde{\phi}|_{B_{(1+2\delta)R_\sigma} \cap \Gamma_\rho} = \phi_\rho|_{B_{(1+2\delta)R_\sigma} \cap \Gamma_\rho}$ and the orthogonality (5.3) holds on the box $B_{(1+2\delta)R_\sigma}$. Thus, by definition of $\mathcal{H}_{h,0}$, we have that the potential $\widetilde{V\phi} \in \mathcal{H}_{h,0}(B_{(1+2\delta)R_\sigma}, \Gamma_\rho)$.

As a consequence, Lemma 3.11 can be applied to $\widetilde{V\phi}$ with $R := (1 + \delta)R_\sigma$ and $\kappa := \frac{1}{2+\eta} = \frac{\delta}{1+\delta}$. Note that $(1+\kappa)(1+\delta) = 1 + 2\delta$ and $1 + \kappa^{-1} = 3 + \eta$. Hence, we get a low dimensional space $\widehat{W}_k$ of dimension $\dim \widehat{W}_k \le C_{\mathrm{dim}}(3+\eta)^d q^{-d} k^{d+1} =: r$,

and the best approximation $\widehat{\phi} = \Pi_{\widehat{W}_k} \widetilde{\phi}$ to $\widetilde{\phi}$ from the space $\widehat{W}_k$ satisfies

$$\left\| \widetilde{\phi} - \widehat{\phi} \right\|_{L^2(B_{(1+\delta)R_\sigma} \cap \Gamma_\rho)} \lesssim R_\sigma h^{-3/2} q^k \left\| \left\| \widetilde{V}\widetilde{\phi} \right\| \right\|_{h,(1+2\delta)R_\sigma}$$

$$\lesssim h^{-3/2} e^{-b_1 r^{1/(d+1)}} \left\| \widetilde{\phi} \right\|_{H^{-1/2}(\Gamma)},$$

where we defined $b_1 := -\frac{\ln(q)}{C_{\dim}^{1/(d+1)}} q^{d/(d+1)} (3 + \eta)^{-d/(1+d)} > 0$ to obtain $q^k = e^{-b_1 r^{1/(d+1)}}$. Therefore, we get

$$(5.5) \quad \left| \left\langle \widetilde{\phi} - \widehat{\phi}, V\psi \right\rangle_{L^2(B_{(1+\delta)R_\sigma} \cap \Gamma_\rho)} \right| \lesssim h^{-3/2} e^{-b_1 r^{1/(d+1)}} \left\| \widetilde{\phi} \right\|_{H^{-1/2}(\Gamma)} \|V\psi\|_{L^2(\Gamma)}.$$

The ellipticity of V, $\operatorname{supp}(\widetilde{\phi} - \phi) = \operatorname{supp}\phi_\rho \subset \overline{\Gamma_\rho}$, and the orthogonality (5.3) yield

$$\left\| \widetilde{\phi} - \phi \right\|^2_{H^{-1/2}(\Gamma)} \lesssim \left\langle V(\widetilde{\phi} - \phi), \widetilde{\phi} - \phi \right\rangle = -\left\langle V\phi, \widetilde{\phi} - \phi \right\rangle$$

$$(5.6) \qquad\qquad \lesssim \|V\phi\|_{H^{1/2}(\Gamma)} \left\| \widetilde{\phi} - \phi \right\|_{H^{-1/2}(\Gamma)} \lesssim \|\phi\|_{L^2(\Gamma)} \left\| \widetilde{\phi} - \phi \right\|_{H^{-1/2}(\Gamma)}.$$

Thus, with the triangle inequality, (5.6), and the stability of $V : L^2(\Gamma) \to H^1(\Gamma)$, we can estimate (5.5) by

$$\left| \left\langle \widetilde{\phi} - \widehat{\phi}, V\psi \right\rangle_{L^2(B_{(1+\delta)R_\sigma} \cap \Gamma_\rho)} \right| \lesssim h^{-3/2} e^{-br^{1/(d+1)}} \left(\left\| \widetilde{\phi} - \phi \right\|_{H^{-1/2}(\Gamma)} \|V\psi\|_{L^2(\Gamma)} \right.$$

$$+ \|\phi\|_{H^{-1/2}(\Gamma)} \|V\psi\|_{L^2(\Gamma)} \Big)$$

$$\lesssim h^{-3/2} e^{-br^{1/(d+1)}} \|\phi\|_{L^2(\Gamma)} \|\psi\|_{L^2(\Gamma)}.$$

For the second term in (5.4), we exploit the asymptotic smoothness of the Green's function $G(\cdot, \cdot)$: Lemma 5.1 can be applied with $B_Y = B_{R_\sigma}$ and $D_X = \Gamma \backslash B_{(1+\delta)R_\sigma}$, where the choice of δ implies

$$(5.7) \qquad\qquad \operatorname{dist}(B_Y, D_X) \geq \frac{1}{2\sqrt{d}(1+\eta)} \operatorname{diam}(B_Y).$$

Therefore, we get an approximation $G_r(x,y) = \sum_{i=1}^r g_{1,i}(x) g_{2,i}(y)$ such that

$$(5.8) \quad \|G(x,\cdot) - G_r(x,\cdot)\|_{L^\infty(B_{R_\sigma})} \lesssim \frac{1}{\operatorname{dist}(\{x\}, B_{R_\sigma})^{d-2}} e^{-b_2 r^{1/d}} \ \forall x \in \Gamma \backslash B_{(1+\delta)R_\sigma};$$

here, the constant $b_2 > 0$ depends only on d and η. As a consequence of (5.7) and (5.8), the rank-r operator V_r given by $V_r\psi(x) := \int_{B_{R_\sigma} \cap \Gamma} G_r(x,y)\psi(y)ds_y$ satisfies

$$\left| \left\langle \widetilde{\phi}, (V - V_r)\psi \right\rangle_{L^2(\Gamma \backslash B_{(1+\delta)R_\sigma})} \right|$$

$$= \left| \int_{\Gamma \backslash B_{(1+\delta)R_\sigma}} \widetilde{\phi}(x) \int_{B_{R_\sigma} \cap \Gamma} (G(x,y) - G_r(x,y))\psi(y)ds_y ds_x \right|$$

$$\lesssim \left\| \widetilde{\phi} \right\|_{L^2(\Gamma)} \sqrt{\operatorname{meas}(\Gamma \cap B_{R_\sigma})} \left\| G - \widetilde{G}_r \right\|_{L^\infty\left((\Gamma \backslash B_{(1+\delta)R_\sigma}) \times (B_{R_\sigma} \cap \Gamma)\right)} \|\psi\|_{L^2(\Gamma)}$$

$$\lesssim h^{-1/2} \delta^{2-d} R_\sigma^{(3-d)/2} e^{-b_2 r^{1/d}} \left\| \widetilde{\phi} \right\|_{H^{-1/2}(\Gamma)} \|\psi\|_{L^2(\Gamma)}$$

$$\lesssim h^{-1/2} e^{-b_2 r^{1/d}} \|\phi\|_{L^2(\Gamma)} \|\psi\|_{L^2(\Gamma)},$$

where the last two inequalities follow from the inverse estimate Lemma 3.3, the stability estimate (5.6) for the mapping $\phi \mapsto \widetilde{\phi}$, the assumption $d \leq 3$ as well as $R_\sigma \leq \eta \operatorname{diam}(\Omega)$, and the choice $\delta = \frac{1}{1+\eta}$. Here, the hidden constant additionally depends on η. Therefore, we get

$$\left| \left\langle V\widetilde{\phi}, \psi \right\rangle - \left\langle \widehat{\phi}, V\psi \right\rangle_{L^2(B_{(1+\delta)R_\sigma} \cap \Gamma_\rho)} - \left\langle \widetilde{\phi}, V_r\psi \right\rangle_{L^2(\Gamma \backslash B_{(1+\delta)R_\sigma})} \right|$$
$$\lesssim h^{-3/2} e^{-br^{1/(d+1)}} \|\phi\|_{L^2(\Gamma)} \|\psi\|_{L^2(\Gamma)},$$

with $b := \min\{b_1, b_2\}$. Since the mapping $(\phi, \psi) \mapsto \left\langle \widehat{\phi}, V\psi \right\rangle_{L^2(B_{(1+\delta)R_\sigma} \cap \Gamma_\rho)} + \left\langle \widetilde{\phi}, V_r\psi \right\rangle_{L^2(\Gamma \backslash B_{(1+\delta)R_\sigma})}$ defines a bounded bilinear form on $L^2(\Gamma)$, there exists a linear operator $\widehat{V} : L^2(\Gamma) \to L^2(\Gamma)$ such that

$$\left\langle \widehat{\phi}, V\psi \right\rangle_{L^2(B_{(1+\delta)R_\sigma} \cap \Gamma_\rho)} + \left\langle \widetilde{\phi}, V_r\psi \right\rangle_{L^2(\Gamma \backslash B_{(1+\delta)R_\sigma})} = \left\langle \widehat{V}\phi, \psi \right\rangle,$$

and the dimension of the range of $\widehat{V}$ is bounded by $2r$. Therefore, we get a matrix $\mathbf{S}_r(\tau, \sigma)$ of rank $2r$ such that

$$\|\mathbf{S}(\tau, \sigma) - \mathbf{S}_r(\tau, \sigma)\|_2 = \sup_{\boldsymbol{\phi} \in \mathbb{R}^{|\tau|}, \boldsymbol{\psi} \in \mathbb{R}^{|\sigma|}} \frac{\left| \boldsymbol{\phi}^T (\mathbf{S}(\tau, \sigma) - \mathbf{S}_r(\tau, \sigma)) \boldsymbol{\psi} \right|}{\|\boldsymbol{\phi}\|_2 \|\boldsymbol{\psi}\|_2}$$
$$\leq C h^{d-3/2} e^{-br^{1/(d+1)}},$$

where we have used (2.1). The estimate $\frac{1}{\|\mathbf{V}\|_2} \lesssim h^{-d}$ from [Ste08, Lemma 12.6] and $h \simeq N^{-1/(d-1)}$ finish the proof. $\qquad\square$

As a direct consequence of the representation (5.2) and the results from Section 3, we can get a blockwise rank-r approximation of the inverse of the Schur complement $\mathbf{S}(\tau, \tau)$. For the existence of the inverse $\mathbf{S}(\tau, \tau)^{-1}$, we refer to the next subsection. For a given right-hand side $f \in L^2(\Gamma)$, (5.2) implies that solving $\mathbf{S}(\tau, \tau)\boldsymbol{\phi} = \mathbf{f}$ with $\mathbf{f} \in \mathbb{R}^{|\tau|}$ defined by $\mathbf{f}_i = \langle f, \chi_{i_\tau} \rangle$, is equivalent to solving $a(\widetilde{\phi}, \psi) = \langle f, \psi \rangle$ for all $\psi \in S^{0,0}(\mathcal{T}_h)$ with $\operatorname{supp}\psi \subset \overline{\Gamma_\tau}$. Let $\tau_1 \times \sigma_1 \subset \tau \times \tau$ be an η-admissible subblock. For $f \in L^2(\Gamma)$ with $\operatorname{supp} f \subset B_{R_{\sigma_1}} \cap \Gamma$, the support properties as well as the admissibility condition (2.3) for the cluster pair (τ_1, σ_1) imply the orthogonality

$$a(\widetilde{\phi}, \psi) = 0 \quad \forall \psi \in S^{0,0}(\mathcal{T}_h) \text{ with } \operatorname{supp}\psi \subset B_{R_{\tau_1}} \cap \overline{\Gamma_\tau}.$$

Therefore, we have $\widetilde{V}\widetilde{\phi} \in \mathcal{H}_{h,0}(B_{R_{\tau_1}}, \Gamma_\tau)$, and Lemma 3.11 provides an approximation to $\widetilde{\phi}$ on $B_{R_{\tau_1}} \cap \Gamma_\tau$. Then, a rank-$r$ factorization of the subblock $\mathbf{S}(\tau, \tau)^{-1}|_{\tau_1 \times \sigma_1}$ can be constructed as in Section 4, which is summarized in the following theorem.

Theorem 5.3. *Let $\tau \subset \mathcal{I}$, $\rho := \{i \in \mathcal{I} : i < \min(\tau)\}$, $\tau_1 \times \sigma_1 \subset \tau \times \tau$ be η-admissible, and let the Schur complement $\mathbf{S}(\tau, \tau)$ be defined in (5.1). Then, there exist rank-r matrices $\mathbf{X}_{\tau_1 \sigma_1} \in \mathbb{R}^{|\tau_1| \times r}$, $\mathbf{Y}_{\tau_1 \sigma_1} \in \mathbb{R}^{|\sigma_1| \times r}$ such that*

$$(5.9) \qquad \left\| \mathbf{S}(\tau, \tau)^{-1}|_{\tau_1 \times \sigma_1} - \mathbf{X}_{\tau_1 \sigma_1} \mathbf{Y}_{\tau_1 \sigma_1}^T \right\|_2 \leq C_{\text{apx}} N^{(d+2)/(d-1)} e^{-br^{1/(d+1)}}.$$

The constants C_{apx} depends only on Ω, d, and the γ-shape regularity of the quasi-uniform triangulation $\mathcal{T}_h$, and the constant $b > 0$ additionally depends on η.

5.2. **Existence of $\mathcal{H}$-Cholesky decomposition: conclusion of the proof of Theorem 2.11.** In this subsection, we will use the approximation of the Schur complement from the previous section to prove the existence of an (approximate) $\mathcal{H}$-Cholesky decomposition. We start with a hierarchical relation of the Schur complements $\mathbf{S}(\tau, \tau)$.

The Schur complements $\mathbf{S}(\tau, \tau)$ for a block $\tau \in \mathbb{T}_\mathcal{I}$ can be derived from the Schur complements of its sons τ_1, τ_2 by

$$\mathbf{S}(\tau, \tau) = \begin{pmatrix} \mathbf{S}(\tau_1, \tau_1) & \mathbf{S}(\tau_1, \tau_2) \\ \mathbf{S}(\tau_2, \tau_1) & \mathbf{S}(\tau_2, \tau_2) + \mathbf{S}(\tau_2, \tau_1)\mathbf{S}(\tau_1, \tau_1)^{-1}\mathbf{S}(\tau_1, \tau_2) \end{pmatrix},$$

A proof of this relation can be found in [Beb07, Lemma 3.1]. One should note that the proof does not use any properties of the matrix $\mathbf{V}$ other than invertibility and existence of a Cholesky decomposition. Moreover, we have by definition of $\mathbf{S}(\tau, \tau)$ that $\mathbf{S}(\mathcal{I}, \mathcal{I}) = \mathbf{V}$.

If τ is a leaf, we get the Cholesky decomposition of $\mathbf{S}(\tau, \tau)$ by the classical Cholesky decomposition, which exists since $\mathbf{V}$ has a Cholesky decomposition. If τ is not a leaf, we use the hierarchical relation of the Schur complements to define a Cholesky decomposition of the Schur complement $\mathbf{S}(\tau, \tau)$ by

$$(5.10) \qquad \mathbf{C}(\tau) := \begin{pmatrix} \mathbf{C}(\tau_1) & 0 \\ \mathbf{S}(\tau_2, \tau_1)(\mathbf{C}(\tau_1)^T)^{-1} & \mathbf{C}(\tau_2) \end{pmatrix},$$

with $\mathbf{S}(\tau_1, \tau_1) = \mathbf{C}(\tau_1)\mathbf{C}(\tau_1)^T$, $\mathbf{S}(\tau_2, \tau_2) = \mathbf{C}(\tau_2)\mathbf{C}(\tau_2)^T$ and indeed get $\mathbf{S}(\tau, \tau) = \mathbf{C}(\tau)\mathbf{C}(\tau)^T$. Moreover, the uniqueness of the Cholesky decomposition of $\mathbf{V}$ implies that due to $\mathbf{C}\mathbf{C}^T = \mathbf{V} = \mathbf{S}(\mathcal{I}, \mathcal{I}) = \mathbf{C}(\mathcal{I})\mathbf{C}(\mathcal{I})^T$, we have $\mathbf{C} = \mathbf{C}(\mathcal{I})$.

The existence of the inverse $\mathbf{C}(\tau_1)^{-1}$ follows from the representation (5.10) by induction over the levels, since on a leaf the existence is clear and the matrices $\mathbf{C}(\tau)$ are block triangular matrices. Consequently, the inverse of $\mathbf{S}(\tau, \tau)$ exists.

Moreover, as shown in [GKLB09, Lemma 22] in the context of LU-factorizations instead of Cholesky decompositions, the restriction of the lower triangular part $\mathbf{S}(\tau_2, \tau_1)(\mathbf{C}(\tau_1)^T)^{-1}$ of the matrix $\mathbf{C}(\tau)$ to a subblock $\tau_2' \times \tau_1'$ with τ_i' a son of τ_i satisfies

$$(5.11) \qquad \left(\mathbf{S}(\tau_2, \tau_1)(\mathbf{C}(\tau_1)^T)^{-1}\right)\big|_{\tau_2' \times \tau_1'} = \mathbf{S}(\tau_2', \tau_1')(\mathbf{C}(\tau_1')^T)^{-1}.$$

The following lemma shows that the spectral norm of the inverse $\mathbf{C}(\tau)^{-1}$ can be bounded by the norm of the inverse $\mathbf{C}(\mathcal{I})^{-1}$.

Lemma 5.4. *For $\tau \in \mathbb{T}_\mathcal{I}$, let $\mathbf{C}(\tau)$ be given by (5.10). Then,*

$$\max_{\tau \in \mathbb{T}_\mathcal{I}} \left\|\mathbf{C}(\tau)^{-1}\right\|_2 = \left\|\mathbf{C}(\mathcal{I})^{-1}\right\|_2.$$

Proof. With the block structure of (5.10), we get the inverse

$$\mathbf{C}(\tau)^{-1} = \begin{pmatrix} \mathbf{C}(\tau_1)^{-1} & 0 \\ -\mathbf{C}(\tau_2)^{-1}\mathbf{S}(\tau_2, \tau_1)(\mathbf{C}(\tau_1)^T)^{-1}\mathbf{C}(\tau_1)^{-1} & \mathbf{C}(\tau_2)^{-1} \end{pmatrix}.$$

So, we get by choosing $\mathbf{x}$ such that $\mathbf{x}_i = 0$ for $i \in \tau_1$ that

$$\left\|\mathbf{C}(\tau)^{-1}\right\|_2 = \sup_{\mathbf{x} \in \mathbb{R}^{|\tau|}, \|x\|_2 = 1} \left\|\mathbf{C}(\tau)^{-1}\mathbf{x}\right\|_2$$

$$\geq \sup_{\mathbf{x} \in \mathbb{R}^{|\tau_2|}, \|x\|_2 = 1} \left\|\mathbf{C}(\tau_2)^{-1}\mathbf{x}\right\|_2 = \left\|\mathbf{C}(\tau_2)^{-1}\right\|_2.$$

The same argument for $\left(\mathbf{C}(\tau)^{-1}\right)^T$ leads to

$$\left\|\mathbf{C}(\tau)^{-1}\right\|_2 = \left\|\left(\mathbf{C}(\tau)^{-1}\right)^T\right\|_2 \geq \left\|\mathbf{C}(\tau_1)^{-1}\right\|_2.$$

Thus, we have $\left\|\mathbf{C}(\tau)^{-1}\right\|_2 \geq \max_{i=1,2}\left\|\mathbf{C}(\tau_i)^{-1}\right\|_2$ and as a consequence $\max_{\tau\in\mathbb{T}_\mathcal{I}}\left\|\mathbf{C}(\tau)^{-1}\right\|_2 = \left\|\mathbf{C}(\mathcal{I})^{-1}\right\|_2.$ □

We are now in position to prove Theorem 2.11:

Proof of Theorem 2.11. In the following, we show that every admissible subblock $\tau \times \sigma$ of $\mathbf{C}(\mathcal{I})$, recursively defined by (5.10), has a rank-r approximation. Since an admissible block of the lower triangular part of $\mathbf{C}(\mathcal{I})$ has to be a subblock of a matrix $\mathbf{C}(\tau')$ for some $\tau' \in \mathbb{T}_\mathcal{I}$, we get in view of (5.11) that $\mathbf{C}(\mathcal{I})|_{\tau\times\sigma} = \mathbf{S}(\tau,\sigma)(\mathbf{C}(\sigma)^T)^{-1}$. Theorem 5.2 provides a rank-r approximation $\mathbf{S}_r(\tau,\sigma)$ to $\mathbf{S}(\tau,\sigma)$. Therefore, we can estimate

$$\left\|\mathbf{C}(\mathcal{I})|_{\tau\times\sigma} - \mathbf{S}_r(\tau,\sigma)(\mathbf{C}(\sigma)^T)^{-1}\right\|_2 = \left\|\left(\mathbf{S}(\tau,\sigma) - \mathbf{S}_r(\tau,\sigma)\right)(\mathbf{C}(\sigma)^T)^{-1}\right\|_2$$
$$\leq C_{\mathrm{sc}}N^{3/(2d-2)}e^{-br^{1/(d+1)}}\left\|(\mathbf{C}(\sigma)^T)^{-1}\right\|_2\|\mathbf{V}\|_2.$$

Since $\mathbf{S}_r(\tau,\sigma)(\mathbf{C}(\sigma)^T)^{-1}$ is a rank-r matrix for each η-admissible cluster pair (τ,σ), we immediately get an $\mathcal{H}$-matrix approximation $\mathbf{C}_\mathcal{H}$ of the Cholesky factor $\mathbf{C}(\mathcal{I}) = \mathbf{C}$. With Lemma 4.1 and Lemma 5.4, we get

$$\|\mathbf{C} - \mathbf{C}_\mathcal{H}\|_2 \leq C_{\mathrm{sc}}C_{\mathrm{sp}}N^{3/(2d-2)}\operatorname{depth}(\mathbb{T}_\mathcal{I})e^{-br^{1/(d+1)}}\left\|\mathbf{C}^{-1}\right\|_2\|\mathbf{V}\|_2,$$

and with $\|\mathbf{V}\|_2 = \|\mathbf{C}\|_2^2$, we conclude the proof of (i).

Since $\mathbf{V} = \mathbf{C}\mathbf{C}^T$, the triangle inequality finally leads to

$$\left\|\mathbf{V} - \mathbf{C}_\mathcal{H}\mathbf{C}_\mathcal{H}^T\right\|_2 \leq \|\mathbf{C} - \mathbf{C}_\mathcal{H}\|_2\left\|\mathbf{C}^T\right\|_2 + \left\|\mathbf{C}^T - \mathbf{C}_\mathcal{H}^T\right\|_2\|\mathbf{C}\|_2$$
$$+ \|\mathbf{C} - \mathbf{C}_\mathcal{H}\|_2\left\|\mathbf{C}^T - \mathbf{C}_\mathcal{H}^T\right\|_2$$
$$\leq 2C_{\mathrm{sc}}C_{\mathrm{sp}}\kappa_2(\mathbf{C})\operatorname{depth}(\mathbb{T}_\mathcal{I})N^{3/(2d-2)}e^{-br^{1/(d+1)}}\|\mathbf{V}\|_2$$
$$+ \kappa_2(\mathbf{C})^2 C_{\mathrm{sc}}^2 C_{\mathrm{sp}}^2 \operatorname{depth}(\mathbb{T}_\mathcal{I})^2 N^{3/(d-1)}e^{-2br^{1/(d+1)}}\frac{\|\mathbf{V}\|_2^2}{\|\mathbf{C}\|_2^2},$$

and the equality $\kappa_2(\mathbf{V}) = \kappa_2(\mathbf{C})^2$ finishes the proof. □

6. EXTENSIONS

6.1. **The Poincaré-Steklov operator.** The interior Poincaré-Steklov operator S^{int} is defined as $S^{\mathrm{int}} := V^{-1}\left(\frac{1}{2}I + K\right) : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$, where K denotes the double-layer operator. In a similar way, the exterior Poincaré-Steklov operator S^{ext} is given by $S^{\mathrm{ext}} = -V^{-1}(\frac{1}{2}I - K)$.

The discrete Poincaré-Steklov operators are given by $\mathbf{S}^{\mathrm{int}} = \mathbf{V}^{-1}\left(\frac{1}{2}\mathbf{M} + \mathbf{K}\right)$ and $\mathbf{S}^{\mathrm{ext}} = -\mathbf{V}^{-1}\left(\frac{1}{2}\mathbf{M} - \mathbf{K}\right)$, where $\mathbf{K}$ is the stiffness matrices corresponding to K, and $\mathbf{M}$ is the mass matrix. Here, we consider piecewise affine basis functions for the the discretizations $\mathbf{K}, \mathbf{M}$. We illustrate the use of the $\mathcal{H}$-arithmetic to derive $\mathcal{H}$-matrix approximations to the discrete Poincaré-Steklov operators, which is stated in the following corollary of Theorem 2.7.

Corollary 6.1. *Fix $\eta > 0$. Let a partition P of $\mathcal{I} \times \mathcal{I}$ be based on a cluster tree $\mathbb{T}_{\mathcal{I}}$ created by the geometric clustering algorithm from [Hac09, Section 5.4.2]. Let $\mathbf{S} \in \{\mathbf{S}^{\mathrm{int}}, \mathbf{S}^{\mathrm{ext}}\}$. Then, there is a blockwise rank-r matrix $\mathbf{S}_{\mathcal{H}}$ such that*

$$(6.1) \qquad \|\mathbf{S} - \mathbf{S}_{\mathcal{H}}\|_2 \leq C_{\mathrm{PS}} N^{(d+2)/(d-1)} \log N \exp\left(-b\left(\frac{r}{\log N + 1}\right)^{1/(d+1)}\right).$$

The constant $C_{\mathrm{PS}} > 0$ depends only on Ω, d, and the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$, and the constant $b > 0$ depends additionally on η.

Proof. By $\mathcal{H}$-arithmetic, [GH03, Theorem 2.24], the rank of the multiplication of $\mathcal{H}$-matrices increases by a factor of $C_{\mathrm{id}} C_{\mathrm{sp}}(\mathrm{depth}(\mathbb{T}_{\mathcal{I}}) + 1)$, where the appearing idempotency constant C_{id} is defined in [GH03] and can be bounded uniformly in N for geometrically balanced cluster trees. $\qquad\square$

Remark 6.2. In computations, due to stability reasons, usually the symmetric formulation of the Poincaré-Steklov operator $S^{\mathrm{int}} := W + \left(\frac{1}{2}I + K'\right) V^{-1} \left(\frac{1}{2}I + K\right)$ with the hypersingular integral operator $W : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$ and the adjoint double-layer potential K' is used, see, e.g., [Ste08]. Using $\mathcal{H}$-arithmetics, an approximation for this representation can be derived as well, but leads to an additional logarithmic factor in the exponential in (6.1). Since we are only interested in an existence result, the non symmetric formulation is sufficient for our purpose.

6.2. **$\mathcal{H}^2$-approximation.** In this section, we briefly describe how an approximation in the more refined framework of $\mathcal{H}^2$-matrices can be derived. The main advantage of $\mathcal{H}^2$-matrices compared to $\mathcal{H}$-matrices is that the storage complexity as well as the complexity of the matrix-vector multiplication is $\mathcal{O}(rN)$, i.e., linear in the degrees of freedom instead of the logarithmic-linear complexity of $\mathcal{O}(rN \log N)$ for $\mathcal{H}$-matrices.

In fact, [Bör10a] proves that an $\mathcal{H}^2$-approximation can be derived from blockwise estimates as in Theorem 2.4. Therefore, this can be done in the same way, and this section follows the lines of [Bör10a].

$\mathcal{H}^2$-matrices are based on nested cluster bases, which are defined in the following.

Definition 6.3 (Nested cluster basis)**.** A family of matrices $(\mathbf{U}_\tau)_{\tau \in \mathbb{T}_I}, \mathbf{U}_\tau \in \mathbb{R}^{|\tau| \times r}$ is said to be a nested cluster basis, if there exists a family of transfer matrices $(\mathbf{T}_\tau)_{\tau \in \mathbb{T}_{\mathcal{I}}}, \mathbf{T}_\tau \in \mathbb{R}^{r \times r}$ such that $\mathbf{U}_\tau|_{\tau'} = \mathbf{U}_{\tau'}\mathbf{T}_{\tau'}$ $\forall \tau \in \mathbb{T}_{\mathcal{I}}, \tau' \in \mathrm{sons}(\tau)$. If the matrices $\mathbf{U}_\tau$ are orthogonal for all $\tau \in \mathbb{T}_{\mathcal{I}}$, $(\mathbf{U}_\tau)_{\tau \in \mathbb{T}_{\mathcal{I}}}$ is said to be a nested orthogonal cluster basis.

Definition 6.4 (Descendants, predecessors, block row)**.** The set of descendants of $\tau \in \mathbb{T}_{\mathcal{I}}$ is recursively defined by

$$\mathrm{sons}^*(\tau) := \begin{cases} \{\tau\} & \text{if } \mathrm{sons}(\tau) = \emptyset, \\ \{\tau\} \cup \bigcup_{\tau' \in \mathrm{sons}(\tau)} \mathrm{sons}^*(\tau') & \text{otherwise.} \end{cases}$$

The set of predecessors is given by

$$\mathrm{pred}(\tau) := \{\tau^+ \in \mathbb{T}_{\mathcal{I}} \ : \ \tau \in \mathrm{sons}^*(\tau^+)\}.$$

Further, the block row is defined by

$$\mathrm{row}^*(\tau) := \{\sigma \in \mathbb{T}_{\mathcal{I}} \ : \ \exists \tau^+ \in \mathrm{pred}(\tau) : (\tau^+, \sigma) \in \mathrm{P}_{\mathrm{far}}\}.$$

Definition 6.5 ($\mathcal{H}^2$-matrix). Let the partition P of $\mathcal{I} \times \mathcal{I}$ be based on the cluster tree $\mathbb{T}_\mathcal{I}$ and $\eta > 0$. Let $(\mathbf{S}_\tau)_{\tau \in \mathbb{T}_\mathcal{I}}$ and $(\mathbf{U}_\sigma)_{\sigma \in \mathbb{T}_\mathcal{I}}$ be nested cluster bases. A matrix $\mathbf{W}_{\mathcal{H}^2}$ is said to be an $\mathcal{H}^2$-matrix, if for each η-admissible cluster pair $(\tau, \sigma) \in P_{\text{far}}$, there is a coupling matrix $\mathbf{M}_{\tau\sigma} \in \mathbb{R}^{r \times r}$ such that $\mathbf{W}_{\mathcal{H}^2}|_{\tau \times \sigma} = \mathbf{S}_\tau \mathbf{M}_{\tau\sigma} \mathbf{U}_\sigma^T$.

We refer to [HB02] for the storage complexity of $\mathcal{O}(rN)$ for the family of transfer matrices $(\mathbf{T}_\tau)_{\tau \in \mathbb{T}_\mathcal{I}}$ and coupling matrices $(\mathbf{M}_{\tau\sigma})_{(\tau,\sigma) \in P_{\text{far}}}$. Since the cluster basis only needs to be stored for the leaf clusters, we get a storage requirement of $\mathcal{O}(rN)$ for the cluster bases $(\mathbf{S}_\tau)_{\tau \in \mathbb{T}_\mathcal{I}}$ and $(\mathbf{U}_\tau)_{\tau \in \mathbb{T}_\mathcal{I}}$, and therefore a total storage requirement of $\mathcal{O}(rN)$ for $\mathcal{H}^2$-matrices.

The following theorem shows that the matrix $\mathbf{V}^{-1}$ can be approximated by an $\mathcal{H}^2$-matrix and that the error converges exponentially in the block rank.

Theorem 6.6. *Fix the admissibility parameter $\eta > 0$. Let a partition P of $\mathcal{I} \times \mathcal{I}$ be based on a cluster tree $\mathbb{T}_\mathcal{I}$. Then, there is a blockwise rank-r $\mathcal{H}^2$-matrix $\mathbf{W}_{\mathcal{H}^2}$ such that*

$$\left\| \mathbf{V}^{-1} - \mathbf{W}_{\mathcal{H}^2} \right\|_2 \leq C_{\text{H2}} N^{(d+2)/(d-1)} \text{depth}(\mathbb{T}_\mathcal{I}) \sqrt{|\mathbb{T}_\mathcal{I}|} e^{-br^{1/(d+1)}}.$$

The constant $C_{\text{H2}} > 0$ depends only on Ω, d, and the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$, where the constant $b > 0$ additionally depends on η.

Proof. We define the total cluster basis $M_\tau := \bigcup \{ \sigma : \sigma \in \text{row}^*(\tau) \}$. By definition of $\text{row}^*(\tau)$, there exists a cluster $\tau^+ \in \text{pred}(\tau)$ such that $(\tau^+, \sigma) \in P_{\text{far}}$ for all $\sigma \in \text{row}^*(\tau)$. Let $B_{R_\tau}, B_{R_{\tau^+}}, B_{R_\sigma}$ be bounding boxes for the clusters τ, τ^+, σ, and we assume that $B_{R_\tau} \subset B_{R_{\tau^+}}$ for $\tau \in \text{sons}^*(\tau^+)$. Then, we have

$$\text{diam}(B_{R_\tau}) \leq \text{diam}(B_{R_{\tau^+}}) \leq \eta \, \text{dist}(B_{R_{\tau^+}}, B_{R_\sigma})$$
$$\leq \eta \, \text{dist}(B_{R_\tau}, B_{R_\sigma}) \qquad \forall \sigma \in \text{row}^*(\tau).$$

Therefore, Theorem 2.4 can be used to derive a low-rank approximation of the matrix block $\mathbf{V}^{-1}|_{\tau \times M_\tau}$, i.e., it provides matrices $\mathbf{X}_{\tau M_\tau} \in \mathbb{R}^{|\tau| \times r}, \mathbf{Y}_{\tau M_\tau} \in \mathbb{R}^{|M_\tau| \times r}$ with $r = C_{\dim}(2 + \eta)^d q^{-d} k^{d+1}$, such that

$$\left\| \mathbf{V}^{-1}|_{\tau \times M_\tau} - \mathbf{X}_{\tau M_\tau} \mathbf{Y}_{\tau M_\tau}^T \right\|_2 \lesssim N^{(d+2)/(d-1)} q^k \simeq N^{(d+2)/(d-1)} e^{-br^{1/(d+1)}},$$

where the constant $b > 0$ depends only on Ω, d, the γ-shape regularity of the quasiuniform triangulation $\mathcal{T}_h$ and η. Then, [Bör10b, Corollary 6.18] states that there exists a nested orthogonal cluster basis $(\mathbf{U}_\tau)_{\tau \in \mathbb{T}_\mathcal{I}} \in \mathbb{R}^{|\tau| \times r}$ of rank $r \leq C_{\dim}(2 + \eta)^d q^{-d} k^{d+1}$ such that for each $(\tau, \sigma) \in P_{\text{far}}$ we have

$$\left\| \mathbf{V}^{-1}|_{\tau \times \sigma} - \mathbf{U}_\tau \mathbf{U}_\tau^T \mathbf{V}^{-1}|_{\tau \times \sigma} \right\|_2^2 \lesssim \sum_{\tau \in \mathbb{T}_\mathcal{I}} \left\| \mathbf{V}^{-1}|_{\tau \times M_\tau} - \mathbf{X}_{\tau M_\tau} \mathbf{Y}_{\tau M_\tau}^T \right\|_2^2$$
$$\lesssim N^{(2d+4)/(d-1)} |\mathbb{T}_\mathcal{I}| e^{-2br^{1/(d+1)}}.$$

With $\mathbf{M}_{\tau\sigma} := \mathbf{U}_\tau^T \mathbf{V}^{-1}|_{\tau \times \sigma} \mathbf{U}_\sigma$, the matrix

$$(6.2) \qquad \mathbf{W}_{\mathcal{H}^2} = \begin{cases} \mathbf{U}_\tau \mathbf{M}_{\tau\sigma} \mathbf{U}_\sigma^T, & \text{for } (\tau, \sigma) \in P_{\text{far}}, \\ \mathbf{V}^{-1}|_{\tau \times \sigma}, & \text{otherwise} \end{cases}$$

is then the desired $\mathcal{H}^2$-matrix approximating $\mathbf{V}^{-1}$. The symmetry of $\mathbf{V}^{-1}$, i.e. $\mathbf{V}^{-1} = \mathbf{V}^{-T}$, and $\left\|\mathbf{A}^T\right\|_2 = \|\mathbf{A}\|_2$ for $\mathbf{A} \in \mathbb{R}^{|\tau| \times |\sigma|}$ imply

$$
\begin{aligned}
\left\|\mathbf{V}^{-1}|_{\tau \times \sigma} - \mathbf{W}_{\mathcal{H}^2}|_{\tau \times \sigma}\right\|_2 &\leq \left\|\mathbf{V}^{-1}|_{\tau \times \sigma} - \mathbf{U}_\tau \mathbf{U}_\tau^T \mathbf{V}^{-1}|_{\tau \times \sigma}\right\|_2 \\
&\quad + \left\|\mathbf{U}_\tau \mathbf{U}_\tau^T (\mathbf{V}^{-1}|_{\tau \times \sigma} - \mathbf{V}^{-1}|_{\tau \times \sigma} \mathbf{U}_\sigma \mathbf{U}_\sigma^T)\right\|_2 \\
&\leq \left\|\mathbf{V}^{-1}|_{\tau \times \sigma} - \mathbf{U}_\tau \mathbf{U}_\tau^T \mathbf{V}^{-1}|_{\tau \times \sigma}\right\|_2 \\
&\quad + \left\|\mathbf{V}^{-1}|_{\sigma \times \tau} - \mathbf{U}_\sigma \mathbf{U}_\sigma^T \mathbf{V}^{-1}|_{\sigma \times \tau}\right\|_2 \\
&\lesssim N^{(d+2)/(d-1)}\sqrt{|\mathbb{T}_\mathcal{I}|}e^{-br^{1/(d+1)}}.
\end{aligned}
$$

Finally, Lemma 4.1 finishes the proof. $\square$

For a quasiuniform mesh and typical clustering strategies, there holds $|\mathbb{T}_\mathcal{I}| \simeq N$ and $\text{depth}(\mathbb{T}_\mathcal{I}) \simeq \log N$. Therefore, the rank r of the $\mathcal{H}^2$-matrix needed to obtain an accuracy ε in Theorem 6.6 is given by $r \simeq \left|\ln \varepsilon - (\frac{d+2}{d-1} + \frac{1}{2})\log N - \log \log N\right|^{d+1} \simeq (C_1 + C_2 \log N)^{d+1}$. In comparison, for the $\mathcal{H}$-matrix approximation in Theorem 2.7 a smaller rank of $r \simeq \left|\ln \varepsilon - \frac{d+2}{d-1}\log N - \log \log N\right|^{d+1} \simeq (C_1 + C_3 \log N)^{d+1}$ is sufficient. However, since the storage requirement for $\mathcal{H}^2$-matrices is given by $\mathcal{O}(rN)$ in comparison to the storage requirement for $\mathcal{H}$-matrices of $\mathcal{O}(rN \log N)$, we observe that in terms of storage, the $\mathcal{H}^2$-approximation leads to better results.

We refer to [Bör10a] for numerical examples concerning this comparison for FEM matrices.

7. Numerical Examples

In this section, we present some numerical examples in two and three dimensions to illustrate our theoretical estimates derived in the previous sections. Further numerical examples for the $\mathcal{H}$-matrix approximation of inverse BEM matrices and black-box preconditioning with an $\mathcal{H}$-LU decomposition can be found in [Gra01, Beb05, Gra05, Bör10b].

With the choice $\eta = 2$ for the admissibility parameter in (2.3), the clustering is done by the standard geometric clustering algorithm, i.e., by choosing axis parallel bounding boxes of minimal volume and splitting these bounding boxes in half across the largest face until they are admissible or contain less degrees of freedom than n_leaf, which we choose as $n_\text{leaf} = 25$ for our computations. An approximation $\mathbf{W}_\mathcal{H}$ to the inverse Galerkin matrix $\mathbf{V}^{-1}$ is computed by using a truncated singular value decomposition of the exact inverse. Throughout, we use the C-library HLib [BG99].

7.1. **2D-Example.** We consider the L-shaped domain $\Omega = (0,1) \times (0, \frac{1}{2}) \cup (0, \frac{1}{2}) \times [\frac{1}{2}, 1)$ and the lowest-order discretization of the simple-layer potential

$$\mathbf{V}_{jk} = \langle V \chi_k, \chi_j \rangle$$

from (2.2).

In Figure 1, we compare the decrease of the upper bound $\|\mathbf{I} - \mathbf{V}\mathbf{W}_\mathcal{H}\|_2$ of the relative error with the increase in the block-rank for a fixed number $N = 16,384$ of degrees of freedom, where the largest block of $\mathbf{W}_\mathcal{H}$ has a size of $2,048^2$. Moreover, Table 1 shows the storage requirement for the computed $\mathcal{H}$-matrix approximation and the percentage of memory needed compared to the full representation. As

theoretically expected, we observe a linear growth in the rank r for the storage requirements. Moreover, we remark that $\|\mathbf{W}_{\mathcal{H}}\|_2 = 1.22 \cdot 10^8$.

r	Storage $\mathbf{W}_{\mathcal{H}}$ (MB)	Compressed to (%)
2	24.5	1.2
3	32.6	1.6
5	49.0	2.4
7	65.3	3.2
9	81.6	4.0
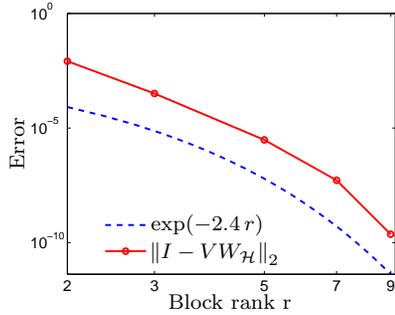
FIGURE 1. Exponential convergence in block rank

TABLE 1. Storage (MB) for $\mathbf{W}_{\mathcal{H}}$

We observe exponential convergence in the block rank, where the convergence rate is $\exp(-br)$, which is even faster than the rate of $\exp(-br^{1/3})$ guaranteed by Theorem 2.7.

7.2. **3D-Example.** We consider the crankshaft generated by NETGEN [Sch97] visualized in Figure 2.

FIGURE 2. Crankshaft domain

In Figure 3, we compare the decrease of $\|\mathbf{I} - \mathbf{VW}_{\mathcal{H}}\|_2$ with the increase in the block-rank for a fixed number $N = 27,968$ of degrees of freedom, where the largest block of $\mathbf{W}_{\mathcal{H}}$ has a size of $3,496^2$. Table 2 shows the storage requirement for the matrix $\mathbf{W}_{\mathcal{H}}$ and the compression rates. We mention that $\|\mathbf{W}_{\mathcal{H}}\|_2 = 41.3$.

Comparing the results with our theoretical bound from Theorem 2.7, we empirically observe a rate of $\exp(-br^{1/2})$ instead of $\exp(-br^{1/4})$.

REFERENCES

[Beb00]   M. Bebendorf, *Approximation of boundary element matrices*, Numer. Math. **86** (2000), no. 4, 565–589.

[Beb05]   ———, *Hierarchical LU decomposition-based preconditioners for BEM*, Computing **74** (2005), no. 3, 225–247.

[Beb07]   ———, *Why finite element discretizations can be factored by triangular hierarchical matrices*, SIAM J. Numer. Anal. **45** (2007), no. 4, 1472–1494.

[BG99]   S. Börm and L. Grasedyck, *H-Lib - a library for $\mathcal{H}$- and $\mathcal{H}^2$-matrices*, available at http://www.hlib.org, 1999.
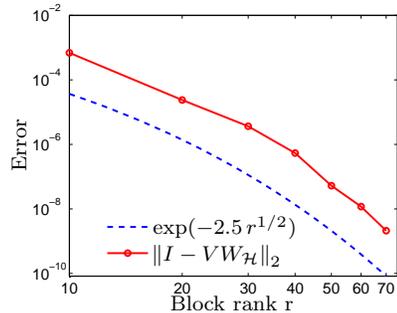
FIGURE 3. Exponential convergence in block rank

r	Storage $\mathbf{W}_{\mathcal{H}}$ (MB)	Compressed to (%)
10	1298	21.8
20	1465	24.6
30	1633	27.4
40	1800	30.2
50	1968	33.0
60	2135	35.8
70	2303	38.6

TABLE 2. Storage (MB) for $\mathbf{W}_{\mathcal{H}}$

[BG04] _____, *Low-rank approximation of integral operators by interpolation*, Computing **72** (2004), no. 3-4, 325–332.

[BG05] _____, *Hybrid cross approximation of integral operators*, Numer. Math. **101** (2005), no. 2, 221–249.

[BH03] M. Bebendorf and W. Hackbusch, *Existence of $\mathcal{H}$-matrix approximants to the inverse FE-matrix of elliptic operators with L^∞-coefficients*, Numer. Math. **95** (2003), no. 1, 1–28.

[Bör10a] S. Börm, *Approximation of solution operators of elliptic partial differential equations by $\mathcal{H}$- and $\mathcal{H}^2$-matrices*, Numer. Math. **115** (2010), no. 2, 165–193.

[Bör10b] _____, *Efficient numerical methods for non-local operators*, EMS Tracts in Mathematics, vol. 14, European Mathematical Society (EMS), Zürich, 2010.

[Bre99] S. C. Brenner, *The condition number of the Schur complement in domain decomposition*, Numer. Math. **83** (1999), no. 2, 187–203.

[CDGS10] S. Chandrasekaran, P. Dewilde, M. Gu, and N. Somasunderam, *On the numerical rank of the off-diagonal blocks of Schur complements of discretized elliptic PDEs*, SIAM J. Matrix Anal. Appl. **31** (2010), no. 5, 2261–2290.

[CMZ13] E. Corona, P.G. Martinsson, and D. Zorin, *An $O(N)$ direct solver for integral equations on the plane*, Tech. report, 2013, to appear in Advances in Computational Harmonic Analysis, `arXiv:1303.5466 [math.NA]`.

[DFG+01] W. Dahmen, B. Faermann, I.G. Graham, W. Hackbusch, and S.A. Sauter, *Inverse inequalities on non-quasiuniform meshes and application to the mortar element method*, Math. Comp. **73** (2001), 1107–1138.

[FMP12] M. Faustmann, J. M. Melenk, and D. Praetorius, *A new proof for existence of $\mathcal{H}$-matrix approximants to the inverse of FEM matrices: the Dirichlet problem for the Laplacian*, ASC Report 51/2012, Institute for Analysis and Scientific Computing, Vienna University of Technology, Wien (2012).

[FMP13] _____, *$\mathcal{H}$-matrix approximability of the inverse of FEM matrices*, ASC Report 20/2013, Institute for Analysis and Scientific Computing, Vienna University of Technology, Wien (2013).

[FMP14] _____, *Existence of $\mathcal{H}$-matrix approximants to the inverses of BEM matrices: the hyper singular integral operator*, work in progress (2014).

[GGMR09] L. Greengard, D. Gueyffier, P.G. Martinsson, and V. Rokhlin, *Fast direct solvers for integral equations in complex three-dimensional domains*, Acta Numer. **18** (2009), 243–275.

[GH03] L. Grasedyck and W. Hackbusch, *Construction and arithmetics of $\mathcal{H}$-matrices*, Computing **70** (2003), no. 4, 295–334.

[GHK08] L. Grasedyck, W. Hackbusch, and R. Kriemann, *Performance of $\mathcal{H}$-LU preconditioning for sparse matrices*, Comput. Methods Appl. Math. **8** (2008), no. 4, 336–349.

[GKLB08] L. Grasedyck, R. Kriemann, and S. Le Borne, *Parallel black box $\mathcal{H}$-LU preconditioning for elliptic boundary value problems*, Comput. Vis. Sci. **11** (2008), no. 4-6, 273–291.

[GKLB09] _____, *Domain decomposition based 𝓗-LU preconditioning*, Numer. Math. **112** (2009), no. 4, 565–600.

[GM13]   A. Gillman and P. Martinsson, *A direct solver with $O(N)$ complexity for variable coefficient elliptic PDEs discretized via a high-order composite spectral collocation method*, Tech. report, 2013, to appear in SIAM J. Sci. Comput., `arXiv:1302.5995 [math.NA]`.

[GR97]   L. Greengard and V. Rokhlin, *A new version of the fast multipole method for the Laplace in three dimensions*, Acta Numerica 1997, Cambridge University Press, 1997, pp. 229–269.

[Gra01]   L. Grasedyck, *Theorie und Anwendungen Hierarchischer Matrizen*, doctoral thesis, Kiel, 2001.

[Gra05]   _____, *Adaptive recompression of 𝓗-matrices for BEM*, Computing **74** (2005), no. 3, 205–223.

[Hac99]   W. Hackbusch, *A sparse matrix arithmetic based on 𝓗-matrices. Introduction to 𝓗-matrices*, Computing **62** (1999), no. 2, 89–108.

[Hac09]   _____, *Hierarchische Matrizen: Algorithmen und Analysis*, Springer, 2009.

[HB02]   W. Hackbusch and S. Börm, *Data-sparse approximation by adaptive $\mathcal{H}^2$-matrices*, Computing **69** (2002), no. 1, 1–35.

[HG12]   K.L. Ho and L. Greengard, *A fast direct solver for structured linear systems by recursive skeletonization*, SIAM J. Sci. Comput. **34** (2012), no. 5, A2507–A2532.

[HJ13]   R.A. Horn and Ch.R. Johnson, *Matrix analysis*, second ed., Cambridge University Press, Cambridge, 2013.

[HK00a]   W. Hackbusch and B. N. Khoromskij, *A sparse 𝓗-matrix arithmetic: general complexity estimates*, J. Comput. Appl. Math. **125** (2000), no. 1-2, 479–501.

[HK00b]   _____, *A sparse 𝓗-matrix arithmetic. II. Application to multi-dimensional problems*, Computing **64** (2000), no. 1, 21–47.

[HKS00]   W. Hackbusch, B. Khoromskij, and S.A. Sauter, *On $\mathcal{H}^2$-matrices*, Lectures on Applied Mathematics (2000), 9–29.

[HN89]   W. Hackbusch and Z.P. Nowak, *On the fast matrix multiplication in the boundary element method by panel clustering*, Numer. Math. **54** (1989), 463–491.

[HS93]   W. Hackbusch and S.A. Sauter, *On the efficient use of the Galerkin method to solve Fredholm integral equations*, Proceedings of ISNA '92—International Symposium on Numerical Analysis, Part I (Prague, 1992), vol. 38, 1993, pp. 301–322.

[HW08]   G.C. Hsiao and W.L. Wendland, *Boundary integral equations*, Applied Mathematical Sciences, vol. 164, Springer-Verlag, Berlin, 2008.

[HY13a]   K.L. Ho and L. Ying, *Hierarchical interpolative factorization for elliptic operators: differential equations*, Tech. report, 2013, `arXiv:1307.2895 [math.NA]`.

[HY13b]   _____, *Hierarchical interpolative factorization for elliptic operators: integral equations*, Tech. report, 2013, `arXiv:1307.2666 [math.NA]`.

[LBG06]   S. Le Borne and L. Grasedyck, *𝓗-matrix preconditioners in convection-dominated problems*, SIAM J. Matrix Anal. Appl. **27** (2006), no. 4, 1172–1183 (electronic).

[LGWX12] S. Li, M. Gu, C.J. Wu, and J. Xia, *New efficient and robust HSS Cholesky factorization of SPD matrices*, SIAM J. Matrix Anal. Appl. **33** (2012), no. 3, 886–904.

[Lin04]   M. Lintner, *The eigenvalue problem for the 2D Laplacian in 𝓗- matrix arithmetic and application to the heat and wave equation*, Computing **72** (2004), no. 3-4, 293–323.

[Mar09]   P.G. Martinsson, *A fast direct solver for a class of elliptic partial differential equations*, J. Sci. Comput. **38** (2009), no. 3, 316–330. MR 2475654 (2010c:65041)

[McL00]   W. McLean, *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, 2000.

[MR05]   P.G. Martinsson and V. Rokhlin, *A fast direct solver for boundary integral equations in two dimensions*, J. Comput. Phys. **205** (2005), no. 1, 1–23.

[Neč67]   J. Nečas, *Les méthodes directes en théorie des équations elliptiques*, Masson et Cie, Éditeurs, Paris, 1967.

[NH88]   Z.P. Novak and W. Hackbusch, *O cloznosti metoda panelej*, Vycislitel'nye prozessy i sistemy, Nauka, Moscow, 1988, pp. 233–244.

[Rat98]   A. Rathsfeld, *A wavelet algorithm for the boundary element solution of a geodetic boundary value problem*, Comput. Methods Appl. Mech. Engrg. **157** (1998), no. 3-4,

267–287, Seventh Conference on Numerical Methods and Computational Mechanics in Science and Engineering (NMCM 96) (Miskolc).

[Rat01]   ———, *On a hierarchical three-point basis in the space of piecewise linear functions over smooth surfaces*, Problems and methods in mathematical physics (Chemnitz, 1999), Oper. Theory Adv. Appl., vol. 121, Birkhäuser, Basel, 2001, pp. 442–470.

[Riv74]   Th.J. Rivlin, *The Chebyshev polynomials*, Wiley-Interscience, New York, 1974.

[Rok85]   V. Rokhlin, *Rapid solution of integral equations of classical potential theory*, J. Comput. Phys. **60** (1985), 187–207.

[Sau92]   S.A. Sauter, *Über die effiziente Verwendung des Galerkinverfahrens zur Lösung Fredholmscher Integralgleichungen*, Ph.D. thesis, Universität Kiel, 1992.

[Sch97]   J. Schöberl, *NETGEN - An advancing front 2D/3D-mesh generator based on abstract rules*, Comput.Visual.Sci (1997), no. 1, 41–52.

[Sch98]   R. Schneider, *Multiskalen- und Wavelet-Matrixkompression: Analysisbasierte Methoden zur effizienten Lösung großer vollbesetzter Gleichungssysteme*, Advances in Numerical Mathematics, Teubner, 1998.

[Sch06]   R. Schreittmiller, *Zur Approximation der Lösungen elliptischer Systeme partieller Differentialgleichungen mittels Finiter Elemente und $\mathcal{H}$-Matrizen*, Ph.D. thesis, Technische Universität München, 2006.

[SS11]   S.A. Sauter and Ch. Schwab, *Boundary element methods*, Springer Series in Computational Mathematics, vol. 39, Springer-Verlag, Berlin, 2011.

[Ste08]   O. Steinbach, *Numerical approximation methods for elliptic boundary value problems*, Springer, New York, 2008.

[SY12]   P.G. Schmitz and L. Ying, *A fast direct solver for elliptic problems on general meshes in 2D*, J. Comput. Phys. **231** (2012), no. 4, 1314–1338. MR 2876456

[SZ90]   L.R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp. **54** (1990), no. 190, 483–493.

[Tau03]   J. Tausch, *Sparse BEM for potential theory and Stokes flow using variable order wavelets*, Comput. Mech. **32** (2003), no. 4-6, 312–318.

[TW03]   J. Tausch and J. White, *Multiscale bases for the sparse representation of boundary integral operators on complex geometry*, SIAM J. Sci. Comput. **24** (2003), no. 5, 1610–1629.

[Tyr00]   E.E. Tyrtyshnikov, *Incomplete cross approximation in the mosaic-skeleton method*, Computing **64** (2000), no. 4, 367–380, International GAMM-Workshop on Multigrid Methods (Bonn, 1998).

[vPSS97]   T. von Petersdorff, Ch. Schwab, and R. Schneider, *Multiwavelets for second-kind integral equations*, SIAM J. Numer. Anal. **34** (1997), no. 6, 2212–2227.

[XCGL09]   J. Xia, S. Chandrasekaran, M. Gu, and X.S. Li, *Superfast multifrontal method for large structured linear systems of equations*, SIAM J. Matrix Anal. Appl. **31** (2009), no. 3, 1382–1411.

[Xia13]   J. Xia, *Efficient structured multifrontal factorization for general large sparse matrices*, SIAM J. Sci. Comput. **35** (2013), no. 2, A832–A860.

Institute for Analysis and Scientific Computing (Inst. E 101), Vienna University of Technology, Wiedner Hauptstrae 8-10, 1040 Wien, Austria
  E-mail address: `markus.faustmann tuwien.ac.at`

Institute for Analysis and Scientific Computing (Inst. E 101), Vienna University of Technology, Wiedner Hauptstrae 8-10, 1040 Wien, Austria
  E-mail address: `melenk tuwien.ac.at`

Institute for Analysis and Scientific Computing (Inst. E 101), Vienna University of Technology, Wiedner Hauptstrae 8-10, 1040 Wien, Austria
  E-mail address: `dirk.praetorius tuwien.ac.at`