ASC Report No. 51/2012

A new proof for existence of $\mathcal{H}$-matrix approximants to the inverse of FEM matrices: the Dirichlet problem for the Laplacian

M. Faustmann, J.M. Melenk, and D. Praetorius

Institute for Analysis and Scientific Computing
Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-05-6

Most recent ASC Reports

50/2012 *W. Auzinger, O. Koch, M. Thalhammer*
Defect-based local error estimators for splitting methods, with application to Schrödinger equations
Part II. Higher-order methods for linear problems

49/2012 *M. Feischl, T. Führer, M. Karkulik, J.M. Melenk and D. Praetorius*
Novel Inverse Estimates for Non-Local Operators

48/2012 *M. Feischl, T. Führer, M. Karkulik, and D. Praetorius*
Quasi-Optimal Adaptive BEM

47/2012 *M. Feischl, T. Führer, M. Karkulik, J.M. Melenk, and D. Praetorius*
FEM-BEM Couplings without Stabilization

46/2012 *P. Dörsek, J.M. Melenk*
Symmetry-free, p-robust equilibrated error indication for the hp-version of the FEM in almost incompressible linear elasticity

45/2012 *P. Dörsek, J.M. Melenk*
A Numerical Study of Averaging Error Indicators in p-FEM

44/2012 *A. Jüngel, I. Stelzer*
Existence analysis of Maxwell-Stefan systems for multicomponent mixtures

43/2012 *M. Feischl, T. Führer, and D. Praetorius*
Adaptive FEM with optimal convergence rates for a certain class of non-symmetric and possibly non-linear problems

42/2012 *L. Banăs, M. Page, D. Praetorius*
A general integrator for the Landau-Lifshitz-Gilbert equation

41/2012 *I. Rachunková, A. Spielauer, S. Staněk, E. Weinmüller*
Positive solutions of nonlinear Dirichlet BVPs in ODEs with time and space singularities

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

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

ISBN 978-3-902627-05-6

© Alle Rechte vorbehalten. Nachdruck nur mit Genehmigung des Autors.

ASC TU WIEN

Chapter 1
A new proof for existence of $\mathcal{H}$-matrix approximants to the inverse of FEM matrices: the Dirichlet problem for the Laplacian

Markus Faustmann, Jens M. Melenk, and Dirk Praetorius

Abstract We study the question of approximability of the inverse of the FEM stiffness matrix for the Laplace problem with Dirichlet boundary conditions by blockwise low rank matrices such as those given by the $\mathcal{H}$-matrix format introduced in [Hac99]. We show that exponential convergence in the local block rank r can be achieved. Unlike prior works [BH03, Bör10a], our analysis avoids any *a priori* coupling $r = \mathcal{O}(|\log h|)$ of r and the mesh width h. Moreover, the techniques developed can be used to analyze other boundary conditions as well.

1.1 Introduction

The format of $\mathcal{H}$-matrices was introduced in [Hac99] as blockwise low-rank matrices that permit storage, application, and even a full (approximate) arithmetic with log-linear complexity. This data-sparse format is well suited to represent exactly sparse matrices arising from discretizations of differential operators and to represent at high accuracy matrices stemming from discretizations of many integral operators, for example, those appearing in boundary integral equation methods.

The inverse of the finite element (FEM) stiffness matrix corresponding to the Dirichlet problem for elliptic operators with bounded coefficients can be approximated in the format of $\mathcal{H}$-matrices with an error that decays exponentially in the block rank employed. This was first observed numerically in [Gra01]. Using properties of the continuous Green's function, [BH03] proves this exponential decay in the block rank up to the discretization error. The work [Bör10a] improves on the result [BH03] in several ways, in particular, by proving a corresponding approximation result in the framework of $\mathcal{H}^2$-matrices. Whereas the analysis of [BH03, Bör10a] is based on the solution operator on the continuous level (e.g., by studying the Green's function), the present approach works on the discrete level.

Institut für Analysis und Scientific Computing, Technische Universität Wien, A–1040 Wien

The exponential approximability in the block rank shown here, is therefore not limited by the discretization error. Moreover, in [BH03, Bör10a] the block rank r and the mesh width h are coupled by $r \sim |\log h|$, which is not needed in our case, since we prove an error estimate explicit in both r and h. We mention that the result presented here can be generalized in various ways. First, [FMP12b] shows that with similar techniques other boundary conditions such as Neumann boundary conditions can be treated, which was not done in [BH03, Bör10a]. Second, [FMP12a] illustrates that approximation results for the inverses of discretizations of first kind boundary integral operators can be obtained with the techniques employed here.

1.2 Main results

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded polygonal (for $d = 2$) or polyhedral (for $d = 3$) Lipschitz domain with boundary $\Gamma := \partial\Omega$. We consider the bilinear form $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ associated with the Poisson problem and given by

$$a(u,v) := \langle \nabla u, \nabla v \rangle \,, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the $L^2(\Omega)$-scalar product. For its discretization, we assume that Ω is triangulated by a quasiuniform mesh $\mathcal{T}_h = \{T_1, \ldots, T_{N_\mathcal{T}}\}$ of mesh width $h := \max_{T_j \in \mathcal{T}_h} \operatorname{diam}(T_j)$. The elements $T_j \in \mathcal{T}_h$ are triangles ($d = 2$) or tetrahedra ($d = 3$), and we assume that $\mathcal{T}_h$ is regular in the sense of Ciarlet. The nodes are denoted by $x_i \in \mathcal{N}_h$, for $i = 1, \ldots, N_\mathcal{N}$. Moreover, the mesh $\mathcal{T}_h$ is assumed to be γ-shape regular in the sense of $\operatorname{diam}(T_j) \leq \gamma\, |T_j|^{1/d}$ 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 γ-shape regularity of $\mathcal{T}_h$. Moreover, we use $\simeq$ to abbreviate that both estimates $\lesssim$ and $\gtrsim$ hold.

For the sake of definiteness, we consider the lowest order Galerkin discretization of the bilinear form $a(\cdot,\cdot)$ by piecewise affine functions in $S_0^{1,1}(\mathcal{T}_h) := S^{1,1}(\mathcal{T}_h) \cap H_0^1(\Omega)$ with $S^{1,1}(\mathcal{T}_h) = \{u \in C(\Omega) \,:\, u|_{T_j} \in \mathcal{P}_1, \forall T_j \in \mathcal{T}_h\}$, taking as the basis of $S_0^{1,1}(\mathcal{T}_h)$ the classical hat-functions associated with the interior nodes of the triangulation. This basis is denoted by $\mathcal{B}_h := \{\psi_j \,:\, j = 1, \ldots, N\}$.

The Galerkin discretization of (1.1) results in a symmetric, positive definite matrix $\mathbf{A} \in \mathbb{R}^{N\times N}$ with

$$\mathbf{A}_{jk} = \langle \nabla\psi_j, \nabla\psi_k \rangle \,, \quad \psi_j, \psi_k \in \mathcal{B}_h.$$

Our goal is to derive an $\mathcal{H}$-matrix approximation $\mathbf{B}_\mathcal{H}$ of the inverse matrix $\mathbf{B} = \mathbf{A}^{-1}$. An $\mathcal{H}$-matrix $\mathbf{B}_\mathcal{H}$ is a blockwise low rank matrix based on the concept of "admissibility", which we now introduce:

Definition 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 box with diameter R_τ,
(ii) supp $\psi_i \subset B_{R_\tau}$ for all $i \in \tau$.

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

$$\max\{R_\tau, R_\sigma\} \leq \eta \operatorname{dist}(B_{R_\tau}, B_{R_\sigma}). \tag{1.2}$$

Definition 2 (blockwise rank-r-matrices). Let P be a partition of $\mathcal{I} \times \mathcal{I}$. A matrix $\mathbf{B}_\mathcal{H} \in \mathbb{R}^{N \times N}$ is said to be a *blockwise rank-r matrix*, if for every η-admissible cluster pair $(\tau, \sigma) \in P$, the block $\mathbf{B}_\mathcal{H}|_{\tau \times \sigma}$ is a rank-r-matrix, i.e., it has the form $\mathbf{B}_\mathcal{H}|_{\tau \times \sigma} = \mathbf{X}_{\tau\sigma}\mathbf{Y}_{\tau\sigma}^T$ 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.

The following theorems are the main results of this paper. Theorem 1 shows that admissible blocks can be approximated by rank-r-matrices:

Theorem 1. *Fix $\eta > 0$. Let the cluster pair (τ, σ) be η-admissible. Fix $q \in (0, 1)$. Then, for $k \in \mathbb{N}$ 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 $r \leq C_{\dim} q^{-d} k^{d+1}$ with*

$$\left\| \mathbf{A}^{-1}|_{\tau \times \sigma} - \mathbf{X}_{\tau\sigma}\mathbf{Y}_{\tau\sigma}^T \right\|_2 \leq C_{\mathrm{apx}}(1 + \eta) h^{-d} q^k. \tag{1.3}$$

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

The approximations for the individual blocks can be combined to gauge the approximability of $\mathbf{A}^{-1}$ 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 3 (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 τ', $\tau'' \in \mathbb{T}_\mathcal{I}$, 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 $\operatorname{level}(\mathcal{I}) = 0$ and $\operatorname{level}(\tau') := \operatorname{level}(\tau) + 1$ for τ' a son of τ. The *depth* of a cluster tree is $\operatorname{depth}(\mathbb{T}_\mathcal{I}) := \max_{\tau \in \mathbb{T}_\mathcal{I}} \operatorname{level}(\tau)$.

Definition 4 (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 $\eta > 0$, we define the *far field* and the *near field* as

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

The *sparsity constant* C_{sp} of such a partition is defined by

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

The following Theorem 2 shows that the matrix $\mathbf{A}^{-1}$ can be approximated by blockwise rank-r-matrices at an exponential rate in the block rank r:

Theorem 2. *Fix $\eta > 0$. 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 $\mathbf{B}_{\mathcal{H}}$ such that*

$$\left\|\mathbf{A}^{-1} - \mathbf{B}_{\mathcal{H}}\right\|_2 \leq C_{\text{apx}} C_{\text{sp}} (1+\eta) N \operatorname{depth}(\mathbb{T}_{\mathcal{I}}) e^{-br^{1/(d+1)}}. \tag{1.4}$$

The constants $C_{\text{apx}}, b > 0$ depend only on Ω, d, and γ-shape regularity of $\mathcal{T}_h$.

Remark 1. Typical clustering strategies such as the "geometric clustering" described in [Hac09] and applied to quasiuniform meshes with $\mathcal{O}(N)$ elements lead to fairly balanced cluster trees $\mathbb{T}_{\mathcal{I}}$ of depth $\mathcal{O}(\log N)$ and feature a sparsity constant C_{sp} that is bounded uniformly in N. We refer to [Hac09] for the fact that the memory requirement to store $\mathbf{B}_{\mathcal{H}}$ is $\mathcal{O}\big((r + n_{\text{leaf}}) N \log N\big)$.

Remark 2. Using $h \simeq N^{-1/d}$ and $\frac{1}{\|\mathbf{A}^{-1}\|_2} \leq \|\mathbf{A}\|_2 \lesssim h^{d/2-1} \simeq N^{-(d-2)/(2d)}$, we get a bound for the relative error

$$\frac{\left\|\mathbf{A}^{-1} - \mathbf{B}_{\mathcal{H}}\right\|_2}{\|\mathbf{A}^{-1}\|_2} \lesssim C_{\text{apx}} C_{\text{sp}} (1+\eta) N^{(d+2)/(2d)} \operatorname{depth}(\mathbb{T}_{\mathcal{I}}) e^{-br^{1/(d+1)}}. \tag{1.5}$$

1.3 Approximation of Galerkin solution on admissible blocks

In terms of functions and function spaces, the question of approximating $\mathbf{A}^{-1}|_{\tau\times\sigma} \approx \mathbf{X}_{\tau\sigma}\mathbf{Y}_{\tau\sigma}^T$ by a low-rank factorization can be phrased as one of how well one can approximate the solution ϕ_h from low-dimensional spaces on B_{R_τ} for data supported by B_{R_σ}. In order to study this question, we consider the question of finding $\phi_h \in S_0^{1,1}(\mathcal{T}_h)$ such that

$$a(\phi_h, \psi_h) = \langle \nabla\phi_h, \nabla\psi_h \rangle = \langle f, \psi_h \rangle \qquad \forall \psi_h \in S_0^{1,1}(\mathcal{T}_h) \tag{1.6}$$

with $\operatorname{supp}(f) \subset B_{R_\sigma}$. By coercivity of $a(\cdot,\cdot)$, the solution ϕ_h is well-defined. In the following, we extend the Galerkin solution by zero outside of Ω and denote this extension by ϕ_h as well. Due to the boundary conditions, this extension belongs to $H^1(B_{R_\tau})$. For η-admissible cluster pairs (τ,σ), the restriction of the solution ϕ_h to B_{R_τ} can be approximated from a low-dimensional space. The heart of the matter is stated in the following:

Proposition 1. *Fix $\eta > 0$. Let the cluster pair (τ,σ) be η-admissible. Fix $q \in (0,1)$. Then, for each $k \in \mathbb{N}$ there exists a sequence V_k of spaces with $\dim V_k \leq C_{\dim} q^{-d} k^{d+1}$ such that for arbitrary f with $\operatorname{supp}(f) \subset B_{R_\sigma} \cap \Omega$, the solution ϕ_h of* (1.6) *satisfies*

$$\min_{v \in V_k} \|\phi_h - v\|_{L^2(B_{R_\tau})} \leq C_{\text{box}} (1+\eta) q^k \|f\|_{L^2(B_{R_\sigma}\cap\Omega)}. \tag{1.7}$$

The constant $C_{\rm dim} > 0$ *depends only on* Ω, d, *and* γ*-shape regularity of* $\mathcal{T}_h$, *while* $C_{\rm box} > 0$ *depends only on* Ω.

The proof of Proposition 1 will be given at the end of this section. The basic steps are as follows: First, one observes that $\mathrm{supp}(f) \subset B_{R_\sigma} \cap \Omega$ as well as the admissibility condition $\mathrm{dist}(B_{R_\tau}, B_{R_\sigma}) \geq \eta^{-1} \max\{\mathrm{diam}(B_{R_\tau}), \mathrm{diam}(B_{R_\sigma})\} > 0$ imply the orthogonality condition

$$\langle \nabla \phi_h, \nabla \psi_h \rangle = \langle f, \psi_h \rangle_{L^2(B_{R_\sigma} \cap \Omega)} = 0 \;\; \forall \psi_h \in S_0^{1,1}(\mathcal{T}_h), \mathrm{supp}(\psi_h) \subset B_{R_\tau} \quad (1.8)$$

i.e. ϕ_h is discrete harmonic on B_{R_τ}. Second, this observation will allow us to prove a Caccioppoli-type estimate (Lemma 1) in which stronger norms of ϕ_h are estimated by weaker norms of ϕ_h on slightly enlarged regions. Third, we proceed as in [BH03, Bör10a] by iterating an approximation result (Lemma 2) derived from the Scott-Zhang interpolation of the Galerkin solution ϕ_h. This iteration argument accounts for the exponential convergence (Lemma 3).

1.3.1 The space $\mathcal{H}_h(D)$ of discrete harmonic functions

Let $D \subset \mathbb{R}^d$ be a domain. A function $u \in H^1(D)$ is called discrete harmonic on $D \cap \Omega$, if

$$\int_{D \cap \Omega} \nabla u \cdot \nabla \varphi_h \, dx = 0 \quad \forall \varphi_h \in S_0^{1,1}(\mathcal{T}_h), \quad \mathrm{supp}(\varphi_h) \subset D \cap \Omega. \qquad (1.9)$$

For domains D, we need a space of functions that are piecewise affine and discrete harmonic on $D \cap \Omega$:

$$\begin{aligned}\mathcal{H}_h(D) := \{u \in H^1(D) \colon \exists \widetilde{u} \in L^2(\mathbb{R}^d) \text{ s.t. } u|_D = \widetilde{u}|_D, \quad \widetilde{u}|_\Omega \in S_0^{1,1}(\mathcal{T}_h),\\ \mathrm{supp}(\widetilde{u}) \subset \overline{\Omega}, \quad u \text{ is discrete harmonic on } D \cap \Omega\}.\end{aligned}$$

Clearly, the finite dimensional space $\mathcal{H}_h(D)$ is a closed subspace of $H^1(D)$, and we have $\phi_h \in \mathcal{H}_h(B_{R_\tau})$ for the solution ϕ_h of (1.6) with $\mathrm{supp}(f) \subset B_{R_\sigma}$ and bounding boxes B_{R_τ}, B_{R_σ} which satisfy the η-admissibility criterion (1.2).

A main tool in our proofs is the Scott-Zhang projection

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

introduced in [SZ90], which preserves homogeneous Dirichlet boundary conditions, i.e. it maps $H_0^1(\Omega)$ to $S_0^{1,1}(\mathcal{T}_h)$. By $\omega_T := \bigcup \{T' \in \mathcal{T}_h \;:\; T \cap T' \neq \emptyset\}$, we denote the element patch of T, which contains T and all elements $T' \in \mathcal{T}_h$ that have a common node with T. Then, J_h has some local approximation property for $\mathcal{T}_h$-piecewise H^ℓ-functions $u \in H^\ell_{\rm pw}(\Omega)$

$$\|u - J_h u\|^2_{H^m(T)} \leq Ch^{2(\ell-m)} \sum_{T' \subset \omega_T} |u|^2_{H^\ell(T')}\,,\ 0 \leq m \leq 1,\ m \leq \ell \leq 2. \quad (1.10)$$

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

For a box B_R with diameter R, we introduce the norm

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

which is, for fixed h, equivalent to the H^1-norm. The following lemma states a Caccioppoli-type estimate for functions in $\mathcal{H}_h(B_{(1+\delta)R})$), where $B_{(1+\delta)R}$ is a box of diameter $(1+\delta)R$ with the same center as the box B_R.

Lemma 1. *Let $\delta > 0$ and $\frac{h}{R} \leq \frac{\delta}{2}$. Let $u \in \mathcal{H}_h(B_{(1+\delta)R})$ for a box $B_{(1+\delta)R}$ of diameter $(1+\delta)R$. Then, there exists a constant $C > 0$ which depends only on γ and d, such that*

$$\|\nabla u\|_{L^2(B_R)} \leq C\frac{1+\delta}{\delta}|||u|||_{h,(1+\delta)R}\,. \quad (1.11)$$

Proof. Let η be a smooth cut-off function with $\operatorname{supp}(\eta) \subset B_{(1+\delta/2)R}$, $\eta \equiv 1$ on B_R, and $\|\nabla\eta\|_{L^\infty(B_R)} \lesssim \frac{1}{\delta R}$, $\|D^2\eta\|_{L^\infty(B_R)} \lesssim \frac{1}{\delta^2 R^2}$. Recall that h is the maximal element width and $2h \leq \delta R$. Therefore, $T \subseteq B_{(1+\delta)R}$ for all $T \in \mathcal{T}_h$ with $T \cap \operatorname{supp}(\eta) \neq \emptyset$. With the abbreviate notation $B := B_{(1+\delta)R}$, we have

$$\|\nabla u\|_{L^2(B_R)} \leq \|\eta\nabla u\|^2_{L^2(B)} = \int_B \nabla u \cdot \nabla(\eta^2 u) - 2\eta u \nabla\eta \cdot \nabla u\, dx.$$

By locality of the Scott-Zhang projection $J_h : H^1(\Omega) \to S^{1,1}(\mathcal{T}_h)$, we observe $\operatorname{supp}(J_h(\eta^2 u)) \subset B$. The orthogonality relation (1.9) implies

$$\begin{aligned}\left|\int_B \nabla u \cdot \nabla(\eta^2 u)dx\right| &= \left|\int_B \nabla u \cdot \nabla(\eta^2 u - J_h(\eta^2 u))dx\right| \\ &\leq \|\nabla u\|_{L^2(B)} \left\|\nabla(\eta^2 u - J_h(\eta^2 u))\right\|_{L^2(B)}.\end{aligned}$$

For the last term, we use the approximation property (1.10) and obtain

$$\begin{aligned}\left\|\nabla(\eta^2 u - J_h(\eta^2 u))\right\|^2_{L^2(B)} &\lesssim h^2 \sum_{\substack{T\in\mathcal{T}_h \\ T\subseteq B}} \left\|D^2(\eta^2 u)\right\|^2_{L^2(T)} \lesssim h^2 \left\|D^2(\eta^2 u)\right\|^2_{L^2(B)} \\ &\lesssim h^2 \Big(\left\|D^2\eta\right\|_{L^\infty(B)} \|\eta u\|_{L^2(B)} + \|\nabla\eta\|_{L^\infty(B)} \|\eta\nabla u\|_{L^2(B)} \\ &\qquad + \|D\eta\|^2_{L^\infty(B)} \|u\|_{L^2(B)}\Big)^2 \\ &\lesssim \left(\frac{h}{\delta^2 R^2}\|u\|_{L^2(B)} + \frac{h}{\delta R}\|\eta\nabla u\|_{L^2(B)}\right)^2.\end{aligned}$$

Finally, we combine these estimates and use the Young inequality to see

$$
\begin{aligned}
\|\eta\nabla u\|_{L^2(B)}^2 &\lesssim \frac{h}{\delta R}\|\nabla u\|_{L^2(B)}\left(\frac{1}{\delta R}\|u\|_{L^2(B)}+\|\eta\nabla u\|_{L^2(B)}\right)\\
&\quad+\frac{1}{\delta R}\|u\|_{L^2(B)}\|\eta\nabla u\|_{L^2(B)}\\
&\leq C\frac{h^2}{\delta^2R^2}\|\nabla u\|_{L^2(B)}^2+C\frac{1}{\delta^2R^2}\|u\|_{L^2(B)}^2+\frac{1}{2}\|\eta\nabla u\|_{L^2(B)}^2 .
\end{aligned}
$$

Moving the term $\frac{1}{2}\|\eta\nabla u\|_{L^2(B)}^2$ to the left-hand side, we conclude the proof. □

1.3.2 Low-dimensional approximation in $\mathcal{H}_h(D)$

Let $\varPi_{h,R} : (H^1(B_R), |||\cdot|||_{h,R}) \to (\mathcal{H}_h(B_R), |||\cdot|||_{h,R})$ be the orthogonal projection, which is well-defined since $\mathcal{H}_h(B_R) \subset H^1(B_R)$ is a closed subspace.

Lemma 2. *Let $\delta > 0$ and $u \in \mathcal{H}_h(B_{(1+2\delta)R})$. Assume $\frac{h}{R} \leq \frac{\delta}{2}$. Let $\mathcal{K}_H$ be an (infinite) γ-shape regular triangulation of $\mathbb{R}^d$ and assume $\frac{H}{R} \leq \frac{\delta}{2}$ for the corresponding mesh width H. Let $J_H : H^1(\mathbb{R}^d) \to S^{1,1}(\mathcal{K}_H)$ be the Scott-Zhang projection. Then, there exists a constant $C > 0$ which depends only on Ω, d, and γ, such that*

(i) $(u - \varPi_{h,R}J_H u)|_{B_R} \in \mathcal{H}_h(B_R)$

(ii) $|||u - \varPi_{h,R}J_H u|||_{h,R} \leq C\frac{1+2\delta}{\delta}\left(\frac{h}{R}+\frac{H}{R}\right)|||u|||_{h,(1+2\delta)R}$

(iii) $\dim W \leq C\left(\frac{(1+2\delta)R}{H}\right)^d$*, where* $W := \varPi_{h,R}J_H\mathcal{H}_h(B_{(1+2\delta)R})$.

Proof. The statement (iii) follows from the fact that $\dim J_H(\mathcal{H}_h(B_{(1+2\delta)R})) \simeq ((1+2\delta)R/H)^d$. For $u \in \mathcal{H}_h(B_{(1+2\delta)R})$, we have $u \in \mathcal{H}_h(B_R)$ as well and hence $\varPi_{h,R}(u|_{B_R}) = u|_{B_R}$, which gives (i). It remains to prove (ii): The assumption $\frac{H}{R} \leq \frac{\delta}{2}$ implies $\bigcup\{K \in \mathcal{K}_H \ : \ \omega_K \cap B_R \neq \emptyset\} \subseteq B_{(1+\delta)R}$. The locality and the approximation properties (1.10) of J_H yield

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

We apply Lemma 1 with $\widetilde{R} = (1+\delta)R$ and $\widetilde{\delta} = \frac{\delta}{1+\delta}$. Note that $(1+\widetilde{\delta})\widetilde{R} = (1+2\delta)R$, and $\frac{h}{\widetilde{R}} \leq \frac{\widetilde{\delta}}{2}$ follows from $2h \leq \delta R = \widetilde{\delta}\widetilde{R}$. Hence, we obtain

$$
\begin{aligned}
|||u - \varPi_{h,R}J_H u|||_{h,R}^2 &= |||\varPi_{h,R}(u - J_H u)|||_{h,R}^2 \leq |||u - J_H u|||_{h,R}^2\\
&= \left(\frac{h}{R}\right)^2\|\nabla(u - J_H u)\|_{L^2(B_R)}^2 + \frac{1}{R^2}\|u - J_H u\|_{L^2(B_R)}^2\\
&\lesssim \frac{h^2}{R^2}\|\nabla u\|_{L^2(B_{(1+\delta)R})}^2 + \frac{H^2}{R^2}\|\nabla u\|_{L^2(B_{(1+\delta)R})}^2\\
&\leq \left(C\frac{1+2\delta}{\delta}\left(\frac{h}{R}+\frac{H}{R}\right)\right)^2 |||u|||_{h,(1+2\delta)R}^2 ,
\end{aligned}
$$

which concludes the proof. □

We have not specified the gridsize H of $\mathcal{K}_H$ to be used in Lemma 2. We will use $h < H$ below. Otherwise, we can choose a grid $\mathcal{K}_H$ such that $\mathcal{K}_H|_\Omega$ is a refinement of $\mathcal{T}_h$ and so the piecewise affine approximation constructed above is equal to the Galerkin solution. In terms of matrix blocks, this corresponds to the case that the ranks of the matrix blocks are comparable to the blocksize. Hence, in the estimate above the term $\frac{h}{R} \leq \frac{H}{R}$ can be dropped.

Lemma 3. *Let $q, \kappa \in (0,1)$, $k \in \mathbb{N}$. Then, there exists a finite dimensional subspace V_k of $\mathcal{H}_h(B_R)$ with dimension $\dim V_k \leq C_{\dim} q^{-d} k^{d+1}$ such that for every $u \in \mathcal{H}_h(B_{(1+\kappa)R})$ it holds*

$$\min_{v \in V_k} \|u - v\|_{h,R} \leq q^k \|u\|_{h,(1+\kappa)R}. \tag{1.12}$$

The constant $C_{\dim} > 0$ depends only on Ω, d, and γ-shape regularity of $\mathcal{T}_h$.

Proof. We iterate the approximation result of Lemma 2 on boxes $B_{(1+\delta_j)R}$, with $\delta_j := \kappa \frac{k-j}{k}$ for $j = 0, \dots, k$. We note that $\kappa = \delta_0 > \delta_1 > \dots > \delta_k = 0$. We choose $H = \frac{qR}{8k \max\{C,1\}}$, where C is the constant in Lemma 2.

If $h \geq H$, then we select $V_k = \mathcal{H}_h(B_R)$. Due to the choice of H we have $\dim V_k \lesssim \left(\frac{R}{h}\right)^d \lesssim k \left(\frac{R}{H}\right)^d \simeq C_{\dim} q^{-d} k^{d+1}$.

If $h < H$, we apply Lemma 2 with $\widetilde{R} = (1+\delta_j)R$ and $\widetilde{\delta_j} = \frac{1}{2k(1+\delta_j)} < \frac{1}{2}$. Note that $\delta_{j-1} = \delta_j + \frac{1}{k}$ gives $(1+\delta_{j-1})R = (1+2\widetilde{\delta_j})\widetilde{R}$. The assumption $\frac{H}{\widetilde{R}} \leq \frac{1}{4k(1+\delta_j)} = \frac{\widetilde{\delta_j}}{2}$ is fulfilled due to our choice of H. For $j = 1$, Lemma 2 provides an approximation w_1 in a subspace W_1 of $\mathcal{H}_h(B_{(1+\delta_1)R})$ with $\dim W_1 \leq C\left(\frac{(1+\kappa)R}{H}\right)^d$ such that

$$\begin{aligned}\|u - w_1\|_{h,(1+\delta_1)R} &\leq 2C \frac{H}{(1+\delta_1)R} \frac{1+2\widetilde{\delta_1}}{\widetilde{\delta_1}} \|u\|_{h,(1+\delta_0)R} \\ &= 4C\frac{kH}{R}(1+2\widetilde{\delta_1}) \|u\|_{h,(1+\kappa)R} \leq q \|u\|_{h,(1+\kappa)R}.\end{aligned}$$

Since $u - w_1 \in \mathcal{H}_h(B_{(1+\delta_1)R})$, we can use Lemma 2 again and get an approximation w_2 of $u - w_1$ in a subspace W_2 of $\mathcal{H}_h(B_{(1+\delta_1)R})$ with $\dim W_2 \leq C\left(\frac{(1+\kappa)R}{H}\right)^d$. Arguing as for $j = 1$, we get

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

Continuing this process $k-2$ times leads to an approximation $v := \sum_{j=1}^k w_i$ in the space $V_k := \sum_{j=1}^k W_j$ of dimension $\dim V_k \leq Ck\left(\frac{(1+\kappa)R}{H}\right)^d = C_{\dim} q^{-d} k^{d+1}$.
□

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

Proof (of Proposition 1). As mentioned in the beginning of this section, for ϕ_h the assumptions of Lemma 3 are satisfied if we fix $\kappa \in (0,1)$ small enough such that $\operatorname{dist}(B_{(1+\kappa)R_\tau}, B_{R_\sigma}) > 0$. The Poincaré inequality implies

$$\|\phi_h\|^2_{H^1(\Omega)} \lesssim \|\nabla\phi_h\|^2_{L^2(\Omega)} = \langle f, \phi_h\rangle \lesssim \|f\|_{L^2(B_{R_\sigma}\cap\Omega)} \|\phi_h\|_{H^1(\Omega)}.$$

Furthermore, with $\frac{h}{R_\tau} < 1$, we get

$$\begin{aligned}\|\!|\phi_h|\!\|_{h,(1+\kappa)R_\tau} &\lesssim \left(1+\frac{1}{R_\tau}\right)\|\phi_h\|_{H^1(B_{2R_\tau})} \leq \left(1+\frac{1}{R_\tau}\right)\|\phi_h\|_{H^1(\Omega)} \\ &\lesssim \left(1+\frac{1}{R_\tau}\right)\|f\|_{L^2(B_{R_\sigma}\cap\Omega)},\end{aligned}$$

and we have a bound on the right-hand side of (1.12). Finally, the admissibility condition leads to

$$\begin{aligned}\min_{v\in V_k}\|\phi_h - v\|_{L^2(B_{R_\tau})} \leq R_\tau \min_{v\in V_k}\|\!|\phi_h - v|\!\|_{h,R_\tau} &\lesssim (R_\tau+1)q^k\|f\|_{L^2(B_{R_\sigma}\cap\Omega)} \\ &\lesssim (\eta+1)\operatorname{diam}(\Omega)q^k\|f\|_{L^2(B_{R_\sigma}\cap\Omega)},\end{aligned}$$

which concludes the proof. □

1.4 Proof of main results

We use the approximation of ϕ_h from the low dimensional spaces, given in Proposition 1, to construct a blockwise low-rank approximation and consequently an $\mathcal{H}$-matrix approximation of the inverse FEM-matrix. The remaining steps of the proof of Theorem 1 follow the lines of [Bör10a]. Therefore, we only sketch the proof.

Proof (of Theorem 1). If $C_{\dim}q^{-d}k^{d+1} \geq \min(|\tau|,|\sigma|)$, we use the exact matrix block $\mathbf{X}_{\tau\sigma} = \mathbf{A}^{-1}|_{\tau\times\sigma}$ and $\mathbf{Y}_{\tau\sigma} = I \in \mathbb{R}^{|\sigma|\times|\sigma|}$. If $C_{\dim}q^{-d}k^{d+1} < \min(|\tau|,|\sigma|)$, let $\lambda_i : L^2(\Omega) \to \mathbb{R}$ be continuous linear functionals satisfying $\lambda_i(\psi_j) = \delta_{ij}$. We define the mappings

$$\Lambda_\tau : L^2(\Omega) \to \mathbb{R}^{|\tau|}, v \mapsto (\lambda_i(v))_{i\in\tau} \text{ and } \mathcal{J}_\tau : \mathbb{R}^{|\tau|} \to S_0^{1,1}(\mathcal{T}_h),\ \mathbf{x} \mapsto \sum_{j\in\tau} x_j\psi_j.$$

Let V_k be the finite dimensional subspace from Proposition 1. We define $\mathbf{X}_{\tau\sigma}$ as an orthogonal basis of the space $\mathcal{V} := \{\Lambda_\tau v \ :\ v \in V_k\}$ and $\mathbf{Y}_{\tau\sigma} := \mathbf{A}^{-1}|^T_{\tau\times\sigma}\mathbf{X}_{\tau\sigma}$. Then, the rank of $\mathbf{X}_{\tau\sigma}, \mathbf{Y}_{\tau\sigma}$ is bounded by $\dim V_k \leq C_{\dim}q^{-d}k^{d+1}$. The error estimate follows from combining the estimate in Proposition 1 with the stability estimate $h^{d/2}\|\mathbf{x}\|_2 \lesssim \|\mathcal{J}_\tau\mathbf{x}\|_{L^2(\Omega)} \lesssim h^{d/2}\|\mathbf{x}\|_2$, see [Bör10a] for details. □

Now, the estimates on each block can be put together to prove our main result.

Proof (of Theorem 2). Theorem 1 provides matrices $\mathbf{X}_{\tau\sigma} \in \mathbb{R}^{|\tau|\times r}$, $\mathbf{Y}_{\tau\sigma} \in \mathbb{R}^{|\sigma|\times r}$, and we define the $\mathcal{H}$-matrix $\mathbf{B}_{\mathcal{H}}$ by

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

On each admissible block $(\tau,\sigma) \in P_{\text{far}}$, we use the blockwise estimate of Theorem 1. On the other blocks, the error is zero by definition. Now, an estimate for the global spectral norm by the local spectral norms from e.g. [Gra01, Hac09] leads to

$$\begin{aligned} \left\|\mathbf{A}^{-1} - \mathbf{B}_{\mathcal{H}}\right\|_2 &\leq C_{\text{sp}} \left(\sum_{\ell=0}^{\infty} \max\left\{ \left\|(\mathbf{A}^{-1} - \mathbf{B}_{\mathcal{H}})|_{\tau\times\sigma}\right\|_2 : (\tau,\sigma) \in P, \text{level}(\tau) = \ell \right\} \right) \\ &\leq C_{\text{sp}} C_{\text{apx}} (1+\eta) h^{-d} q^k \text{depth}(\mathbb{T}_I). \end{aligned}$$

Defining $b = -\frac{\ln(q)}{C_{\text{dim}}^{1/(d+1)}} q^{d/(d+1)} > 0$, we obtain $q^k = e^{-br^{1/(d+1)}}$ and hence

$$\left\|\mathbf{A}^{-1} - \mathbf{B}_{\mathcal{H}}\right\|_2 \lesssim C_{\text{apx}} C_{\text{sp}} (1+\eta) N \text{depth}(\mathbb{T}_{\mathcal{I}}) e^{-br^{1/(d+1)}},$$

which concludes the proof. □

References

[Beb08] M. Bebendorf, *Hierarchical Matrices*, Lecture Notes in Computational Science and Engineering, vol. 63, Springer, Berlin, 2008.

[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.

[FMP12a] M. Faustmann, J.M. Melenk, and D. Praetorius, *Existence of $\mathcal{H}$-matrix approximants to the inverse of bem matrices: the simple layer potential*, work in progress (2012).

[FMP12b] ———, *A new proof for existence of $\mathcal{H}$-matrix approximants to the inverse of FEM matrices*, work in progress (2012).

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

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

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

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

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