Computers and Mathematics with Applications 130 (2023) 21–40

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Exponential meshes and $\mathcal{H}$-matrices ☆

Niklas Angleitner, Markus Faustmann *, Jens Markus Melenk

Institute of Analysis and Scientific Computing (Inst. E 101), Technische Universität Wien, Wiedner Hauptstrasse 8-10, 1040 Wien, Austria

ARTICLE INFO

Keywords:
FEM
H-matrices
Approximability
Non-uniform meshes

ABSTRACT

In [1], we proved that the inverse of the stiffness matrix of an h-version finite element method (FEM) applied to scalar second order elliptic boundary value problems can be approximated at an exponential rate in the block rank by $\mathcal{H}$-matrices. Here, we improve on this result in multiple ways: (1) The class of meshes is significantly enlarged and includes certain exponentially graded meshes. (2) The dependence on the polynomial degree p of the discrete ansatz space is made explicit in our analysis. (3) The bound for the approximation error is sharpened, and (4) the proof is simplified.

1. Introduction

$\mathcal{H}$-matrices were introduced by W. Hackbusch in [33] as a data-sparse matrix format of blockwise low-rank matrices. A particular feature of the $\mathcal{H}$-matrix format is that it comes with an arithmetic that includes the approximate addition, multiplication, and inversion in logarithmic-linear complexity; we refer to [34,10,30,32,6] for a detailed discussion of the algorithmic aspects of $\mathcal{H}$-matrices. A large class of matrices can be represented or at least approximated well in the $\mathcal{H}$-matrix format. Discretizations of differential equations can typically be presented exactly, and the matrices from the discretization of integral operators with so-called asymptotically smooth kernels, which forms a large class of practically relevant integral operators, can be approximated with an error that is exponentially small in the block rank. Given that $\mathcal{H}$-matrices come with an approximate arithmetic, it is important to understand, for which matrices that can be approximated well in that format, also their inverses can be approximated well. It is the purpose of the present paper to study this question for matrices arising from Galerkin discretizations of second order elliptic equations on strongly graded meshes.

The question of $\mathcal{H}$-matrix approximability of the inverses of matrices arising in the finite element method (FEM) has attracted some attention in the past. The first results [7] for scalar elliptic problem and [9] for the time-harmonic Maxwell system showed the existence of locally separable approximations of the Green's function and inferred from that the approximability of the inverses of the FEM matrices by $\mathcal{H}$-matrices via a final projection step. This approach generalizes to certain classes of pseudodifferential operators [20], and results in exponential convergence in the block rank up a multiple of the final projection error. A fully discrete approach, which avoids the final projection steps and leads to exponential convergence in the block rank, was taken in [24,27] in a FEM setting on quasi-uniform meshes and for the boundary element method (BEM) in [25,26,28]. The generalization of [24] to non-uniform meshes was achieved in [1] for low order FEM on certain classes of meshes that includes algebraically graded meshes. In the present work, we generalize [1] in several directions: first, we admit a larger class of meshes that includes certain shape-regular meshes that are graded exponentially towards a lower-dimensional manifold. In particular, we can show exponential approximability in the block rank for the inverses of FEM matrices arising in variants of the boundary concentrated FEM, [36]. Second, our analysis is explicit in the polynomial degree p. For our p-explicit analysis, we develop polynomial-preserving lifting and polynomial projection operators on simplices in arbitrary spatial dimension. Such operators, generalizing the projection-based operators of [17,18,14,19,38], which were restricted to spatial dimensions $d \in \{1,2,3\}$, are of independent interest. Third, on a more technical level, we remove the condition of [1, Def. 2.4] on the relation between the minimal and the maximal mesh size (see Section 3.1 for details).

In general, we follow the notation from [1], which will also for the sake of readability be introduced in the following. We write $a \lesssim b$, if there exists a constant $C > 0$, such that $a \leq Cb$. The constant C may depend on the space dimension d, the computational domain Ω, the PDE coefficients

☆ The research of JMM was supported by the Austrian Science Fund (FWF) by the special research program *Taming complexity in PDE systems* (grant SFB F65).

* Corresponding author.
E-mail addresses: niklas.angleitner@tuwien.ac.at (N. Angleitner), markus.faustmann@tuwien.ac.at (M. Faustmann), melenk@tuwien.ac.at (J.M. Melenk).

https://doi.org/10.1016/j.camwa.2022.11.011
Received 25 March 2022; Received in revised form 10 November 2022; Accepted 13 November 2022
Available online 28 November 2022

a_1, a_2, a_3, the shape-regularity constant σ_{shp}, the admissibility constant σ_{adm} or the sparsity constant σ_{sparse}. However, it may *not* depend on the polynomial degree p.

2. Main results

2.1. The model problem

We investigate the following *model problem*: Let $d \geq 1$ and $\Omega \subseteq \mathbb{R}^d$ be a bounded polyhedral Lipschitz domain. Furthermore, let $a_1 \in L^\infty(\Omega, \mathbb{R}^{d\times d})$, $a_2 \in L^\infty(\Omega, \mathbb{R}^d)$, and $a_3 \in L^\infty(\Omega, \mathbb{R})$ be given coefficient functions and $f \in L^2(\Omega)$ be a given right-hand side. We seek a weak solution $u \in H_0^1(\Omega)$ to the following equations:

$$\begin{aligned} -\operatorname{div}(a_1 \nabla u) + a_2 \cdot \nabla u + a_3 u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

We assume that a_1 is coercive in the sense $\langle a_1(x)y, y\rangle \geq \alpha_1 \|y\|_2^2$ for all $x \in \Omega$, $y \in \mathbb{R}^d$ and some constant $\alpha_1 > \sigma_{\text{Pcr}}^2(\|a_2\|_{L^\infty(\Omega)} + \|a_3\|_{L^\infty(\Omega)}) \geq 0$. Here, $\sigma_{\text{Pcr}} > 0$ denotes the constant in the Poincaré inequality $\|\cdot\|_{H^1(\Omega)} \leq \sigma_{\text{Pcr}} |\cdot|_{H^1(\Omega)}$ on $H_0^1(\Omega)$.

Definition 2.1. We introduce the following bilinear form:

$$\forall u, v \in H_0^1(\Omega): \qquad a(u,v) := \langle a_1 \nabla u, \nabla v\rangle_{L^2(\Omega)} + \langle a_2 \cdot \nabla u, v\rangle_{L^2(\Omega)} + \langle a_3 u, v\rangle_{L^2(\Omega)}.$$

The weak formulation of the *model problem* reads as follows: Find $u \in H_0^1(\Omega)$ such that

$$\forall v \in H_0^1(\Omega): \qquad a(u,v) = \langle f, v\rangle_{L^2(\Omega)}.$$

The assumptions on the PDE coefficients imply that the bilinear form $a(\cdot,\cdot)$ is continuous and coercive. In particular, the well-known Lax-Milgram Lemma yields the existence of a unique solution $u \in H_0^1(\Omega)$.

2.2. The spline spaces

For the discretization of the model problem, we introduce the well-known spline spaces $\mathbb{S}_0^{p,1}(\mathcal{T}) \subseteq H_0^1(\Omega)$, where $\mathcal{T}$ is a *mesh* on Ω and $p \geq 1$ is a prescribed polynomial degree.

Definition 2.2. A finite set $\mathcal{T} \subseteq \operatorname{Pow}(\Omega)$ is a *mesh*, if there exists an open simplex $\hat{T} \subseteq \mathbb{R}^d$ (the *reference element*) such that every *element* $T \in \mathcal{T}$ is of the form $T = F_T(\hat{T})$, where $F_T : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is an affine diffeomorphism. Furthermore, the elements must be pairwise disjoint, i.e., $|T \cap S| = 0$ for all $T \neq S \in \mathcal{T}$, and constitute a partition of Ω, i.e., $\bigcup_{T\in\mathcal{T}} \overline{T} = \overline{\Omega}$. Finally, a mesh must be regular in the sense of [15], i.e., for any two simplices T, $T' \in \mathcal{T}$, the intersection $\overline{T} \cap \overline{T'}$ is either empty or, for a $j \in \{0, \dots, d\}$, the closure of a j-dimensional subsimplex of both T and T'. Here, j-dimensional subsimplices of an element $T \in \mathcal{T}$ are the push-forwards under the element maps F_T of j-dimensional subsimplices of the reference simplex $\hat{T}$.

For every element $T \in \mathcal{T}$, we define the *patch* $\mathcal{T}(T) := \{S \in \mathcal{T} \mid \overline{S} \cap \overline{T} \neq \emptyset\}$ as the set of all elements that touch T. To measure the size of an element $T \in \mathcal{T}$, we introduce the local *mesh width* $h_T := \sup_{x,y\in T} \|y - x\|_2$.

Our main results are based on local approximations, which induces the need to work with subsets of the triangulation $\mathcal{B} \subseteq \mathcal{T}$. For every such collection of elements $\mathcal{B} \subseteq \mathcal{T}$, we introduce the local maximal and minimal element sizes by $h_{\max,\mathcal{B}} := \max_{T\in\mathcal{B}} h_T$ and $h_{\min,\mathcal{B}} := \min_{T\in\mathcal{B}} h_T$. In the case $\mathcal{B} = \mathcal{T}$, we abbreviate $h_{\max} := h_{\max,\mathcal{T}}$ and $h_{\min} := h_{\min,\mathcal{T}}$ for the global maximal and minimal mesh widths.

For every $T \in \mathcal{T}$, we denote the center of the largest inscribable ball by $x_T \in T$ (the *incenter*). Here, $\operatorname{Ball}_2(x,r) := \{y \in \mathbb{R}^d \mid \|y - x\|_2 < r\}$ is the open ball with radius $r > 0$, centered around $x \in \mathbb{R}^d$. The following assumption on shape-regularity guarantees that simplices do not degenerate, i.e., the angles stay uniformly bounded away from zero.

Assumption 2.3 *(Shape regularity).* We assume that $\mathcal{T}$ is part of a *shape-regular* family of meshes, i.e., there exists a constant $\sigma_{\text{shp}} \geq 1$ such that

$$\forall T \in \mathcal{T}: \qquad \operatorname{Ball}_2(x_T, \sigma_{\text{shp}}^{-1} h_T) \subseteq T \subseteq \bigcup \mathcal{T}(T) \subseteq \operatorname{Ball}_2(x_T, \sigma_{\text{shp}} h_T).$$

Let us next give a formal definition of the spline spaces used in the discrete formulation.

Definition 2.4 *(Spline space $\mathbb{S}_0^{p,1}(\mathcal{T})$).* We set

$$\mathbb{S}_0^{p,1}(\mathcal{T}) := \{v \in H_0^1(\Omega) \mid \forall T \in \mathcal{T} : v \circ F_T \in \mathbb{P}^p(\hat{T})\},$$

where $\mathbb{P}^p(\hat{T}) := \operatorname{span}\{\hat{T} \ni x \mapsto x^q \mid \|q\|_1 \leq p\}$ denotes the usual space of polynomials of (total) degree p on the reference element $\hat{T}$. Similarly, we set

$$\mathbb{S}^{p,0}(\mathcal{T}) := \{v \in L^2(\Omega) \mid \forall T \in \mathcal{T} : v \circ F_T \in \mathbb{P}^p(\hat{T})\}.$$

Remark 2.5. Note that the polynomial degree p is the same for all elements of the mesh $\mathcal{T}$. In contrast, in the hp-version of the FEM, usually a polynomial degree distribution $\{p_T \mid T \in \mathcal{T}\}$ is prescribed. In this context, p may be regarded as the maximum of these values. The analysis of a general polynomial degree distribution is beyond the scope of the present work, and we focus on the uniform polynomial degree distribution.

As our main result will be formulated in terms of matrices, a choice of a basis of the discrete space has to be made. In the following, we state assumptions on bases of $\mathbb{S}_0^{p,1}(\mathcal{T})$ that are essential for our analysis. A key observation is that assumptions (like locality) only have to be made for a dual basis. We refer to Example 2.7 below for constructions of such dual systems.

Definition 2.6 *(Local dual functions λ_m).* Let $N := N(\mathcal{T}, p) := \dim \mathbb{S}_0^{p,1}(\mathcal{T})$ and let $\{\varphi_1, \dots, \varphi_N\} \subseteq \mathbb{S}_0^{p,1}(\mathcal{T})$ be a basis. We say that the basis *allows for a system of local dual functions*, if there exist functions $\{\lambda_1, \dots, \lambda_N\} \subseteq L^2(\Omega)$ with the following properties:

(1) Duality: For all $n, m \in \{1, \dots, N\}$, there holds $\langle \varphi_n, \lambda_m \rangle_{L^2(\Omega)} = \delta_{nm}$ (Kronecker delta).
(2) Stability: There exist constants $C_{\text{stab}}, \sigma_{\text{stab}} > 0$ such that

$$\forall \boldsymbol{x} \in \mathbb{R}^N : \qquad \left\| \sum_{m=1}^{N} \boldsymbol{x}_m \lambda_m \right\|_{L^2(\Omega)} \le C_{\text{stab}} p^{\sigma_{\text{stab}}} h_{\min}^{-d/2} \|\boldsymbol{x}\|_2.$$

(3) Locality and overlap: For every $n \in \{1, \dots, N\}$, there exists a characteristic element $T_n \in \mathcal{T}$ such that $\operatorname{supp}(\lambda_n) \subseteq \bigcup \mathcal{T}(T_n)$. For all $T \in \mathcal{T}$, there holds the bound $\#\{n \mid T_n = T\} \le \binom{p+d}{d}$.

Example 2.7. Typically, a finite element basis $\{\varphi_1, \dots, \varphi_N\} \subseteq \mathbb{S}_0^{p,1}(\mathcal{T})$ is constructed from a predefined basis of *shape functions* $\{\hat{\varphi}_i \mid i = 1, \dots, \binom{d+p}{d}\} \subseteq \mathbb{P}^p(\hat{T})$ on the reference element $\hat{T} \subseteq \mathbb{R}^d$. Following [1, Sec. 3.3], we can then build the dual functions $\{\lambda_1, \dots, \lambda_N\} \subseteq L^2(\Omega)$ from the *dual shape functions* $\{\hat{\lambda}_j \mid j = 1, \dots, \binom{d+p}{d}\} \subseteq \mathbb{P}^p(\hat{T})$, which are defined via the conditions $\langle \hat{\varphi}_i, \hat{\lambda}_j \rangle_{L^2(\hat{T})} = \delta_{ij}$. However, since we want to include the case $p \to \infty$ in our analysis, the standard Lagrange basis has to be replaced with a basis with good stability properties in p.

In $d = 2$ space dimensions, for example, we can pick the shape functions $\hat{\varphi}_i$ from [29, Def. 2.4]. It was shown in [29, Lem. 4.4] that the corresponding coordinate mapping $\hat{\Phi} c := \sum_i c_i \hat{\varphi}_i$ exhibits the stability bounds $p^{-3}\|c\|_2 \lesssim \|\hat{\Phi} c\|_{L^2(\hat{T})} \lesssim \|c\|_2$ for all $c \in \mathbb{R}^{(p+2)(p+1)/2}$. In particular, using the Euclidean unit vectors $e_i \in \mathbb{R}^{(p+2)(p+1)/2}$, we also get a stability bound for the dual shape functions $\hat{\lambda}_j$:

$$\|\hat{\lambda}_j\|_{L^2(\hat{T})}^2 = \langle \hat{\Phi}\hat{\Phi}^{-1}\hat{\lambda}_j, \hat{\lambda}_j \rangle_{L^2(\hat{T})} = \sum_i \langle \hat{\Phi}^{-1}\hat{\lambda}_j, e_i \rangle_2 \langle \hat{\varphi}_i, \hat{\lambda}_j \rangle_{L^2(\hat{T})} = \langle \hat{\Phi}^{-1}\hat{\lambda}_j, e_j \rangle_2 \le \|\hat{\Phi}^{-1}\hat{\lambda}_j\|_2 \lesssim p^3 \|\hat{\lambda}_j\|_{L^2(\hat{T})},$$

so that $\|\hat{\lambda}_j\|_{L^2(\hat{T})} \lesssim p^3$. Together with a scaling argument, this gives $\|\lambda_j\|_{L^2(\Omega)} \lesssim p^3 h_{\min}^{-d/2}$. Consequently, we find that the stability bound in Definition 2.6 is satisfied with $\sigma_{\text{stab}} = d/2 + 3 = 4$ (see also [1, Lem. 3.6]).

Finally, let us motivate the assumption $\#\{n \mid T_n = T\} \le \binom{p+d}{d}$ from Definition 2.6: The previously mentioned construction in [1, Sec. 3.3] guarantees that not only $\operatorname{supp}(\lambda_n) \subseteq \bigcup \mathcal{T}(T_n)$, but even $\operatorname{supp}(\lambda_n) = T_n$. Furthermore, owing to item (1) in Definition 2.6, the system $\{\lambda_1, \dots, \lambda_N\} \subseteq L^2(\Omega)$ is linearly independent. Then, given an arbitrary element $T \in \mathcal{T}$, the system $\{\lambda_n|_T \mid n \in \{1, \dots, N\} \text{ with } T_n = T\} \subseteq \mathbb{P}^p(T)$ must be linearly independent as well. It follows that $\#\{n \mid T_n = T\} \le \dim(\mathbb{P}^p(T)) = \binom{d+p}{d}$, i.e., the overlap condition is fulfilled.

As mentioned above, the (local) supports of the dual functions play an import role in our analysis. It pays off to introduce notation for those as well as for unions of supports of dual functions corresponding to some index set.

Definition 2.8 *(Ω_n; support of dual functions λ_n).* For all $n \in \{1, \dots, N\}$ and all $I \subseteq \{1, \dots, N\}$, we set

$$\Omega_n := \operatorname{supp}(\lambda_n) \subseteq \mathbb{R}^d, \qquad \Omega_I := \bigcup_{n \in I} \Omega_n \subseteq \mathbb{R}^d.$$

2.3. The system matrix

Now that the spline spaces $\mathbb{S}_0^{p,1}(\mathcal{T}) \subseteq H_0^1(\Omega)$ are at our disposal, the *discrete model problem* reads as follows: For given $f \in L^2(\Omega)$, find $u \in \mathbb{S}_0^{p,1}(\mathcal{T})$ such that

$$\forall v \in \mathbb{S}_0^{p,1}(\mathcal{T}) : \qquad a(u, v) = \langle f, v \rangle_{L^2(\Omega)}.$$

Again, existence and uniqueness of a solution $u \in \mathbb{S}_0^{p,1}(\mathcal{T})$ follow from the Lax-Milgram Lemma.

As usual, given a basis of the ansatz space, the discrete model problem can be rephrased as an equivalent linear system of equations. The bilinear form $a(\cdot,\cdot)$ from Definition 2.1 and the basis functions $\varphi_n \in \mathbb{S}_0^{p,1}(\mathcal{T})$ from Definition 2.6 compose the governing system matrix.

Definition 2.9 *(Stiffness matrix).* We define the system matrix

$$\boldsymbol{A} := (a(\varphi_n, \varphi_m))_{m,n=1}^N \in \mathbb{R}^{N \times N}.$$

Note that the unique solvability of the discrete model problem already ensures that the matrix $\boldsymbol{A}$ is invertible.

2.4. Hierarchical matrices

In this section, we provide the basic definitions from the theory of hierarchical matrices. We slightly divert from [1, Sec. 2.5] and use the formulation from our previous work on radial basis functions, [2, Sec. 2.4]. As will be discussed later in Section 3.1, a formulation in terms of axes-parallel boxes $B \subseteq \mathbb{R}^d$ rather than collections of elements $\mathcal{B} \subseteq \mathcal{T}$ has certain advantages. An extensive discussion of hierarchical matrices can be found, e.g., in the books [34,6,10].

Definition 2.10 *(Box).* A subset $B \subseteq \mathbb{R}^d$ is called *(axes-parallel) box*, if it has the form $B = \bigtimes_{i=1}^d (a_i, b_i)$ with $a_i < b_i$.

For the next definition, we remind the reader of the subsets $\Omega_I \subseteq \mathbb{R}^d$, introduced in Definition 2.8. Furthermore, we use the usual definition of Euclidean diameter and distance of subsets $B, B_1, B_2 \subseteq \mathbb{R}^d$, i.e.,

$$\operatorname{diam}_2(B) := \sup_{x,y\in B} \|x-y\|_2, \qquad \operatorname{dist}_2(B_1,B_2) := \inf_{x\in B_1, y\in B_2} \|x-y\|_2.$$

Hierarchical matrices are based on blockwise decompositions of a matrix, where approximations usually are only employed for blocks corresponding to well-separated index sets, which is quantified by a so-called *admissibility condition*. In contrast to the classical literature, e.g., [34], which formulates admissibility in terms of the supports of the shape functions φ_n, we formulate the admissibility condition in terms of the supports of the dual functions.

Definition 2.11. Let $\sigma_{\mathrm{small}}, \sigma_{\mathrm{adm}} > 0$. A tuple (I,J) with $I, J \subseteq \{1,\dots,N\}$ is called *small*, if there holds $\min\{\#I, \#J\} \le \sigma_{\mathrm{small}}$. It is called *admissible*, if there exist boxes $B_I, B_J \subseteq \mathbb{R}^d$ such that $\Omega_I \subseteq B_I$, $\Omega_J \subseteq B_J$ and

$$\operatorname{diam}_2(B_I) \le \sigma_{\mathrm{adm}} \operatorname{dist}_2(B_I, B_J).$$

A set $\mathbb{P}$ of tuples (I,J) with $I, J \subseteq \{1,\dots,N\}$ is called *sparse block partition*, if the following assumptions are satisfied:

(1) The system $\{I \times J \mid (I,J) \in \mathbb{P}\}$ forms a partition of $\{1,\dots,N\} \times \{1,\dots,N\}$.
(2) There holds $\mathbb{P} = \mathbb{P}_{\mathrm{small}} \cup \mathbb{P}_{\mathrm{adm}}$, where every $(I,J) \in \mathbb{P}_{\mathrm{small}}$ is small and every $(I,J) \in \mathbb{P}_{\mathrm{adm}}$ is admissible.
(3) For all $\boldsymbol{B} \in \mathbb{R}^{N\times N}$, there holds the bound

$$\|\boldsymbol{B}\|_2 \lesssim \ln(h_{\min}^{-d}) \max_{(I,J)\in\mathbb{P}} \|\boldsymbol{B}|_{I\times J}\|_2.$$

We emphasize that for our analysis it is only crucial to have a partition of $I \times J$ at hand satisfying these three conditions. In the $\mathcal{H}$-matrix literature, usually constructions of such partitions are presented using hierarchical tree structures. Subsequently, item (3) is actually shown rather than assumed using properties of the construction, see e.g. [34, Lem. 6.5.8]. In this sense, we may argue that *sparsity* is a hidden assumption in item (3), but, by assuming property (3) rather than showing it for a specific construction, we may formulate a main result that may also be valid for partitions constructed differently than, e.g., in [34]. We make a brief comment on the existence of a block partition satisfying items (1)–(3) in the following.

A sparse block partition $\mathbb{P}$ can be constructed, e.g., using the *geometrically balanced clustering strategy* from [31]. In fact, recall from item (3) of Definition 2.6 that, at any given point $x \in \mathbb{R}^d$, no more than $\binom{p+d}{d}$ of the sets Ω_n can overlap. Then, assuming that the clustering parameter σ_{small} is chosen such that $\binom{p+d}{d} \le \sigma_{\mathrm{small}} \lesssim \binom{p+d}{d}$, the authors of [31] derived the following bounds for the *block cluster tree* $\mathbb{T}_{N\times N}$:

$$C_{\mathrm{sparse}}(\mathbb{T}_{N\times N}) \lesssim 1, \qquad \operatorname{depth}(\mathbb{T}_{N\times N}) \lesssim \ln(h_{\min}^{-d}).$$

(See, e.g., [31] or [34] for a precise definition of these fundamental quantities.) The asserted bound in item (3) of Definition 2.11 then follows readily from [34, Lem. 6.5.8].

Definition 2.12. Let $\mathbb{P}$ be a sparse block partition and $r \in \mathbb{N}$ a given *block rank bound*. We define the set of $\mathcal{H}$*-matrices* by

$$\mathcal{H}(\mathbb{P}, r) := \{\boldsymbol{B} \in \mathbb{R}^{N\times N} \mid \forall (I,J) \in \mathbb{P}_{\mathrm{adm}} : \exists \boldsymbol{X} \in \mathbb{R}^{I\times r}, \boldsymbol{Y} \in \mathbb{R}^{J\times r} : \boldsymbol{B}|_{I\times J} = \boldsymbol{X}\boldsymbol{Y}^T\}.$$

Finally, according to [34, Lem. 6.3.6], the memory requirements to store an $\mathcal{H}$-matrix $\boldsymbol{B} \in \mathcal{H}(\mathbb{P}, r)$ can be bounded by

$$N_{\mathrm{memory}} \le C_{\mathrm{sparse}}(\mathbb{T}_{N\times N})(r + \sigma_{\mathrm{small}})\operatorname{depth}(\mathbb{T}_{N\times N})N \lesssim (r + p^d)\ln(h_{\min}^{-d})N.$$

Since $\boldsymbol{B}$ shall serve as an approximation for the N^2 entries of the matrix $\boldsymbol{A}^{-1} \in \mathbb{R}^{N\times N}$, this approach requires bounds on r, p and $h_{\min}$ in terms of $N = \dim \mathbb{S}_0^{p,1}(\mathcal{T}) \approx p^d \#\mathcal{T}$. For example, if the mesh $\mathcal{T}$ is such that

$$1 \lesssim (\#\mathcal{T})^{\sigma_{\mathrm{card}}} h_{\min}^d \tag{2.1}$$

for some constant $\sigma_{\mathrm{card}} \ge 1$, then we end up with the following bound:

$$N_{\mathrm{memory}} \lesssim (r + p^d)\ln(N/p^d)N \le (r + p^d)\ln(N)N.$$

2.5. The main result

The following theorem is the main result of the present work. Roughly speaking, it states that inverses of FEM matrices can be approximated at an exponential rate in the block rank by hierarchical matrices.

Theorem 2.13. *Let $a(\cdot,\cdot)$ be the elliptic bilinear form from Definition 2.1, let $\mathcal{T} \subseteq \operatorname{Pow}(\Omega)$ be a mesh as in Definition 2.2, and let $p \ge 1$ be an arbitrary integer. Let $\{\varphi_1, \dots, \varphi_N\} \subseteq \mathbb{S}_0^{p,1}(\mathcal{T})$ be a basis that allows for a system of local dual functions (see Definition 2.6) and denote the corresponding stability constant by $\sigma_{\mathrm{stab}} > 0$. Furthermore, let $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ be the Galerkin stiffness matrix from Definition 2.9, and let $\mathbb{P}$ be a sparse block partition as in Definition 2.11. Finally, let $\sigma_{\mathrm{red}} \ge 2$ be the constant from Lemma 3.9 further below. Then, there exists a constant $\sigma_{\mathrm{exp}} = C(d, \Omega, a, \sigma_{\mathrm{shp}}, \sigma_{\mathrm{adm}}) > 0$ such that the following holds true: For every block rank bound $r \in \mathbb{N}$, there exists an $\mathcal{H}$-matrix $\boldsymbol{B} \in \mathcal{H}(\mathbb{P}, r)$ such that*

$$\|\boldsymbol{A}^{-1} - \boldsymbol{B}\|_2 \lesssim p^{2\sigma_{\mathrm{stab}}} \ln(h_{\min}^{-d}) h_{\min}^{-d} \exp(-\sigma_{\mathrm{exp}} r^{1/(d+1)} p^{-\sigma_{\mathrm{red}}}).$$

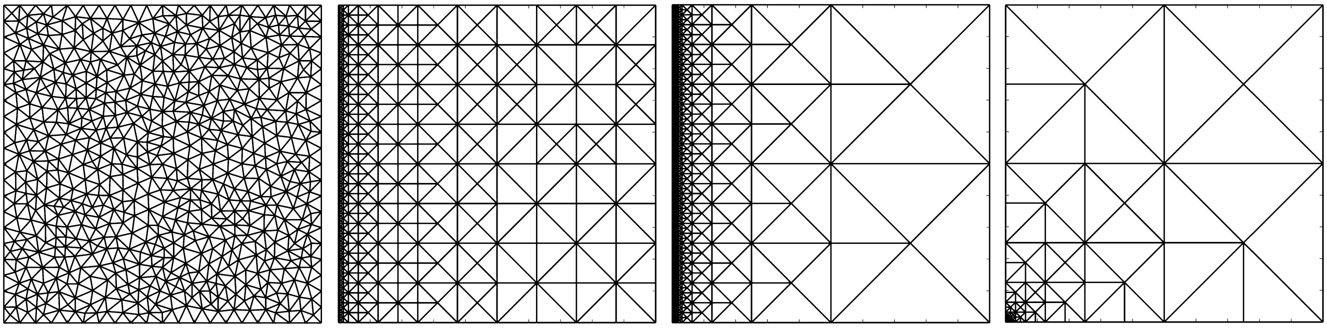

Fig. 1. From left to right: Uniform, algebraically graded towards edge, exponentially graded towards edge, exponentially graded towards corner. Corollary 2.14 covers the first, second and third type, but *not* the last one.

Note that, apart from shape-regularity, this result needs no further assumptions on the mesh $\mathcal{T}$. However, the previous discussion about the storage complexity of $\mathcal{H}$-matrices suggests that we might as well assume Eq. (2.1). In this case, we immediately get the following corollary:

Corollary 2.14. *Assume that the mesh $\mathcal{T}$ satisfies Eq.* (2.1)*, for some constant $\sigma_{\mathrm{card}} \geq 1$. Then, Theorem 2.13 holds verbatim with a bound*

$$\|\boldsymbol{A}^{-1} - \boldsymbol{B}\|_2 \lesssim p^{2\sigma_{\mathrm{stab}} - d\sigma_{\mathrm{card}}} \ln(N) N^{\sigma_{\mathrm{card}}} \exp(-\sigma_{\exp} r^{1/(d+1)} p^{-\sigma_{\mathrm{red}}}).$$

The assumption Eq. (2.1) is satisfied for a wide variety of meshes including uniform, algebraically graded and even some exponentially graded ones (cf. Fig. 1). Given parameters $H > 0$ and $\alpha \in [1, \infty]$, and a subset $\Gamma \subseteq \overline{\Omega}$, we say that a mesh $\mathcal{T}$ is *graded towards* Γ, if there holds the relationship $h_T \approx \mathrm{dist}_2(x_T, \Gamma)^{1-1/\alpha} H$ for all $T \in \mathcal{T}$. The case $\alpha = 1$ is called *uniform*, the case $\alpha \in (1, \infty)$ is an *algebraic grading* and the case $\alpha = \infty$ represents *exponential grading*. If $\alpha \in [1, \infty)$, then Eq. (2.1) is satisfied with $\sigma_{\mathrm{card}} = \alpha$. In the case $\alpha = \infty$, however, the relationship need not necessarily be fulfilled.

Remark 2.15. One possible application of exponentially graded meshes can be found in the context of the *boundary concentrated FEM*, e.g., [36] and [35]. This method is similar to the boundary element method (BEM) in that most mesh elements are near the boundary of Ω. However, we mention that Theorem 2.13 is not directly applicable to this method, because [36] replaces the (constant-degree) spline spaces $\mathbb{S}_0^{p,1}(\mathcal{T})$ from Definition 2.4 with variable degree spline spaces $\mathbb{S}_0^{\boldsymbol{p},1}(\mathcal{T})$, $\boldsymbol{p} = \{p_T \mid T \in \mathcal{T}\}$.

Remark 2.16. In contrast to our previous work, [1, Thm. 2.15], where, for $p = 1$, we established a bound as in Corollary 2.14 but with the exponential replaced by $\exp(-\sigma_{\exp} r^{1/(d\sigma_{\mathrm{card}}+1)})$, we observe that the constant σ_{card} from Corollary 2.14 does not enter the argument of the exponential in the error bound any more. In particular, the rate of convergence (as $r \to \infty$) does not deteriorate for meshes with stronger grading. This behavior is in accordance with the radial basis function setting, observed in [2, Thm. 2.18].

3. Proof of main result

3.1. Overview

In the following, we provide an informal step-by-step description of the proof of our main result. As the main steps are very similar to our previous work, [1], we additionally point out the changes made here to obtain the improved main result.

We start by mentioning that the main result of our previous work, [1, Thm. 2.15], was applicable to a class of meshes with *locally bounded cardinality*, which included uniform and algebraically graded meshes, but excluded exponential grading. More precisely, a mesh $\mathcal{T}$ has locally bounded cardinality, if there exists a constant $\sigma_{\mathrm{card}} \geq 1$ such that

$$h_{\max}^{\sigma_{\mathrm{card}}} \lesssim h_{\min} \quad \text{and} \tag{3.1a}$$

$$\forall \mathcal{B} \subseteq \mathcal{T}: \quad \#\mathcal{B} \lesssim (1 + \mathrm{diam}_{\mathcal{T}}(\mathcal{B})/h_{\max,\mathcal{B}})^{d\sigma_{\mathrm{card}}}. \tag{3.1b}$$

In order to obtain the main result for more general meshes, we need to avoid these conditions. Condition (3.1a) could easily be replaced by the assumption $1 \lesssim N^{\sigma_{\mathrm{card}}} h_{\min}^d$ from Corollary 2.14. However, we may as well avoid it altogether. In fact, whenever $h_{\min}$ appears during the subsequent proof, we simply leave it as is and refrain from replacing it with any potential lower bound. Consequently, the error bound in Theorem 2.13 is formulated in terms of $h_{\min}$, rather than $h_{\max}$ or N. The requirement (3.1b) is much harder to remove and requires an updated construction, which we outline in the following.

(1) We start with a reformulation of the problem of approximating a matrix blockwise to an approximation of functions locally (Lemma 3.2). In this algebraic argument the model problem/mesh only enters by means of the stability bound on the dual basis (see Definition 2.6).
(2) As we now deal with local approximation problems, we need to provide a way to localize functions to boxes. This is done with suitable continuous *cut-off* functions $\kappa_B^\delta \in C^\infty(\overline{\Omega})$ that localize to an enlarged box B^δ and induce a so-called *cut-off operator* $K_B^\delta : H^1(\Omega) \longrightarrow H^1(\Omega)$.
(3) We collect crucial properties of functions that can be approximated in a space $\mathbb{S}_{\mathrm{harm}}(B) \subset H^1(\Omega)$, which we call a space of *locally discrete, harmonic functions* (see Definition 3.7 below). Functions in this space admit a discrete Caccioppoli inequality (Lemma 3.12) that allows one to estimate the H^1-seminorm on a box by the L^2-norm on an enlarged box.

(4) Using the discrete Caccioppoli inequality, we can construct an operator called *single-step coarsening operator* $Q_B^\delta : \mathbb{S}_{\mathrm{harm}}(B^\delta) \longrightarrow \mathbb{S}_{\mathrm{harm}}(B)$ that does some low-rank coarse grid interpolation (Theorem 3.14).
(5) Finally, exponentially convergent approximations can be obtained by iterative application of single-step coarsening operators (Theorem 3.15).

In order to remove condition (3.1b), steps (2)–(4) are different than our previous works. Most notably, [1] worked with *discrete* cut-off functions $\kappa_{\mathcal{B}}^\delta \in \mathbb{S}^{1,1}(\mathcal{T})$. In order to fulfill typical requirements for those ($\kappa_{\mathcal{B}}^\delta|_{\mathcal{B}} \equiv 1$ and $\mathrm{supp}_{\mathcal{T}}(\kappa_{\mathcal{B}}^\delta) \subseteq \mathcal{B}^\delta$) one has to make the technical assumption $\delta \gtrsim h_{\max,\mathcal{B}}$ linking the inflation parameter δ and the maximal mesh-size on $\mathcal{B}$.

In a natural way, this now induced two cases $\delta \gtrsim h_{\max,\mathcal{B}}$ and $\delta \lesssim h_{\max,\mathcal{B}}$ that are treated differently. In the case $\delta \gtrsim h_{\max,\mathcal{B}}$, we used a uniform mesh $\mathcal{S}$ of meshsize δ for re-interpolation, producing an approximant with roughly $\mathcal{O}(\delta^{-d})$ degrees of freedom. In the remaining case $\delta \lesssim h_{\max,\mathcal{B}}$, however, re-interpolation was not necessary. In fact, due to the assumption of locally bounded cardinality, the function $u \in \mathbb{S}_{\mathrm{harm}}(\mathcal{B}^\delta)$ to be approximated had less than $\mathcal{O}(\delta^{-d\sigma_{\mathrm{card}}})$ degrees of freedom anyway. Now, in the case of an exponentially graded mesh $\mathcal{T}$, the input u might have significantly more than $\mathcal{O}(\delta^{-d\sigma_{\mathrm{card}}})$ degrees of freedom. For example, we could refine one of the elements in $\mathcal{B}$ arbitrarily often without ever affecting $h_{\max,\mathcal{B}}$, essentially raising the dimension above any fixed power of δ^{-1}.

The main idea of our revised proof is to eliminate all occurrences of the technical assumption $\delta \gtrsim h_{\max,\mathcal{B}}$, so that the uniform mesh $\mathcal{S}$ can be used in all cases, regardless of the relative sizes of δ and $h_{\max,\mathcal{B}}$. Thus, we revert to the original idea of [24] of using axes-parallel boxes $B \subseteq \mathbb{R}^d$ instead of element clusters $\mathcal{B} \subseteq \mathcal{T}$ and *smooth* cut-off functions, which can easily be constructed, even if $\delta \ll h_{\max,\mathcal{B}}$. Consequently, the cut-off operator in step (2) above now has different mapping properties than before and, as we also require it to map into $\mathbb{S}_{\mathrm{harm}}(B)$ for input in $\mathbb{S}_{\mathrm{harm}}(B)$, the definition of $\mathbb{S}_{\mathrm{harm}}(B)$ (which before was a subspace of $\mathbb{S}_0^{p,1}(\mathcal{T})$) in step (3) needs to be changed. Finally, we need to show the discrete Caccioppoli inequality for the updated spaces $\mathbb{S}_{\mathrm{harm}}(B) \subseteq H_0^1(\Omega)$. It turns out that the assumption $\delta \gtrsim h_{\max,\mathcal{B}}$ can be dropped there as well, because the discrete Caccioppoli inequality reduces to a standard polynomial inverse inequality on large elements.

3.2. Reduction from matrix level to function space level

Definition 3.1. Let $a : H_0^1(\Omega) \times H_0^1(\Omega) \longrightarrow \mathbb{R}$ be the bilinear form from Definition 2.1. For every $f \in L^2(\Omega)$, denote by $S_{\mathcal{T}} f \in \mathbb{S}_0^{p,1}(\mathcal{T})$ the unique function satisfying the following variational equality:

$$\forall v \in \mathbb{S}_0^{p,1}(\mathcal{T}) : \qquad a(S_{\mathcal{T}} f, v) = \langle f, v \rangle_{L^2(\Omega)}.$$

The linear mapping $S_{\mathcal{T}} : L^2(\Omega) \longrightarrow \mathbb{S}_0^{p,1}(\mathcal{T})$ is called the *discrete solution operator*.

Note that existence and uniqueness of $S_{\mathcal{T}} f$ are provided by the Lax-Milgram Lemma. Additionally, there holds the a priori bound $\|S_{\mathcal{T}} f\|_{H^1(\Omega)} \lesssim \|f\|_{L^2(\Omega)}$.

According to Definition 2.12 and the asserted stability bound in Definition 2.11, the task of approximating the whole matrix $\boldsymbol{A}^{-1}$ by an $\mathcal{H}$-matrix $\boldsymbol{B} \in \mathcal{H}(\mathbb{P}, r)$ reduces to the one of approximating the admissible blocks $\boldsymbol{A}^{-1}|_{I\times J}$ by means of matrices $\boldsymbol{X} \in \mathbb{R}^{I\times r}$ and $\boldsymbol{Y} \in \mathbb{R}^{J\times r}$. In the following lemma, we transfer this matrix approximation problem to a problem of approximating functions. The proof is essentially given in [1, Lem. 3.11, Lem. 3.13], but for the sake of completeness we include the crucial steps here.

Lemma 3.2. *Let $I, J \subseteq \{1, \dots, N\}$, and let $V \subseteq L^2(\Omega)$ be a finite-dimensional subspace. Then, there exist an integer $r \leq \dim V$ and matrices $\boldsymbol{X} \in \mathbb{R}^{I\times r}$ and $\boldsymbol{Y} \in \mathbb{R}^{J\times r}$ such that there holds the following error bound:*

$$\|\boldsymbol{A}^{-1}|_{I\times J} - \boldsymbol{X}\boldsymbol{Y}^T\|_2 \lesssim p^{2\sigma_{\mathrm{stab}}} h_{\min}^{-d} \sup_{\substack{f\in V:\\ \mathrm{supp}(f)\subseteq \Omega_J}} \inf_{v\in V} \frac{\|S_{\mathcal{T}} f - v\|_{L^2(\Omega_I)}}{\|f\|_{L^2(\Omega)}}.$$

Proof. Using the coordinate mapping of the dual basis $\Lambda : \begin{cases} \mathbb{R}^N & \longrightarrow & L^2(\Omega) \\ \boldsymbol{x} & \longmapsto & \sum_{n=1}^N \boldsymbol{x}_n \lambda_n \end{cases}$ and its Hilbert space transposed Λ^T, one can directly compute a representation formula for the inverse matrix as

$$\forall \boldsymbol{f} \in \mathbb{R}^N : \qquad \boldsymbol{A}^{-1}\boldsymbol{f} = \Lambda^T S_{\mathcal{T}} \Lambda \boldsymbol{f}. \tag{3.2}$$

Defining $\boldsymbol{V} := (\Lambda^T V)|_I \subseteq \mathbb{R}^I$ gives $r := \dim \boldsymbol{V} \leq \dim V$. Let the columns of the matrix $\boldsymbol{X} \in \mathbb{R}^{I\times r}$ be an $l^2(I)$-orthonormal basis of $\boldsymbol{V}$. Then, the product $\boldsymbol{X}\boldsymbol{X}^T \in \mathbb{R}^{I\times I}$ represents the $l^2(I)$-orthogonal projection from $\mathbb{R}^I$ onto $\boldsymbol{V}$. Finally, set $\boldsymbol{Y} := (\boldsymbol{A}^{-1}|_{I\times J})^T \boldsymbol{X} \in \mathbb{R}^{J\times r}$. With the representation formula (3.2), this gives, for every $\boldsymbol{f} \in \mathbb{R}^N$ with $\mathrm{supp}(\boldsymbol{f}) \subseteq J$, the bound

$$\begin{aligned} \|(\boldsymbol{A}^{-1}|_{I\times J} - \boldsymbol{X}\boldsymbol{Y}^T)\boldsymbol{f}|_J\|_{l^2(I)} &= \|(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^T)(\boldsymbol{A}^{-1}\boldsymbol{f})|_I\|_{l^2(I)} = \inf_{\boldsymbol{v}\in\boldsymbol{V}} \|(\boldsymbol{A}^{-1}\boldsymbol{f})|_I - \boldsymbol{v}\|_{l^2(I)} \\ &= \inf_{v\in V} \|\Lambda^T(S_{\mathcal{T}}\Lambda\boldsymbol{f} - v)\|_{l^2(I)} \leq \|\Lambda\| \cdot \inf_{v\in V} \|S_{\mathcal{T}}\Lambda\boldsymbol{f} - v\|_{L^2(\mathcal{T}(I))}. \end{aligned}$$

Division by $\|\boldsymbol{f}\|_{l^2(J)}$, substituting $f := \Lambda\boldsymbol{f} \in L^2(\Omega)$ and using $\|\boldsymbol{f}\|_{l^2(J)}^{-1} \leq \|\Lambda\| \|f\|_{L^2(\Omega)}^{-1}$ gives the bound

$$\|\boldsymbol{A}^{-1}|_{I\times J} - \boldsymbol{X}\boldsymbol{Y}^T\|_2 \leq \|\Lambda\|^2 \sup_f \inf_v \|S_{\mathcal{T}} f - v\|_{L^2(\Omega_I)} / \|f\|_{L^2(\Omega)}.$$

Using the asserted stability bound from Definition 2.6 gives the stated estimate. □

3.3. The cut-off operator

As we have to construct approximations on boxes corresponding to admissible index tuples, we have to introduce a means of localization, which will be done using cut-off functions on some enlarged boxes defined in the following.

Definition 3.3. Let $B = \bigtimes_{i=1}^{d}(a_i, b_i)$ with $a_i < b_i$ be a box as in Definition 2.10. For every $\delta \geq 0$, we introduce the *inflated box* $B^\delta := \bigtimes_{i=1}^{d}(-\delta + a_i, b_i + \delta) \subseteq \mathbb{R}^d$.

Note that B^δ is again a box. In particular, we can iterate $(B^\delta)^\delta = B^{2\delta}$, $((B^\delta)^\delta)^\delta = B^{3\delta}$, et cetera.

Lemma 3.4. *Let $B \subseteq \mathbb{R}^d$ be a box and $\delta > 0$. Then, there exists a* smooth cut-off function κ_B^δ *with the following properties:*

$$\kappa_B^\delta \in C^\infty(\overline{\Omega}), \qquad \operatorname{supp}(\kappa_B^\delta) \subseteq \Omega \cap B^\delta, \qquad \kappa_B^\delta|_{\Omega\cap B} \equiv 1, \qquad 0 \leq \kappa_B^\delta \leq 1, \qquad \forall l \in \mathbb{N}_0 : |\kappa_B^\delta|_{W^{l,\infty}(\Omega)} \lesssim \delta^{-l}.$$

Proof. Write $B = \bigtimes_{i=1}^{d}(a_i, b_i)$ and pick a univariate function $g \in C^\infty(\mathbb{R})$ with $0 \leq g \leq 1$, $g|_{(-\infty,0]} \equiv 1$ and $g|_{[1/2,\infty)} \equiv 0$. Then, the function $\kappa_B^\delta(x) := \prod_{i=1}^{d} g((a_i - x_i)/\delta) g((x_i - b_i)/\delta)$, $x \in \Omega$, is a valid choice. $\square$

Since κ_B^δ is a smooth function, the corresponding cut-off operator only maps into $H^1(\Omega)$ even for discrete input functions (in contrast to the definitions made in our previous work [1, Def. 3.23]).

Definition 3.5. Let $B \subseteq \mathbb{R}^d$ be a box and $\delta > 0$. Denote by $\kappa_B^\delta \in C^\infty(\overline{\Omega})$ the smooth cut-off function from Lemma 3.4. We define the *cut-off operator*

$$K_B^\delta : \begin{cases} H^1(\Omega) & \longrightarrow & H^1(\Omega) \\ v & \longmapsto & \kappa_B^\delta v \end{cases}.$$

Let us summarize the key properties of this operator:

Lemma 3.6. *Let $B \subseteq \mathbb{R}^d$ be a box and $\delta > 0$. For all $v \in H^1(\Omega)$, there holds the cut-off property $\operatorname{supp}(K_B^\delta v) \subseteq \Omega \cap B^\delta$ and the local projection property $(K_B^\delta v)|_{\Omega\cap B} = v|_{\Omega\cap B}$. If $v \in H_0^1(\Omega)$, then $K_B^\delta v \in H_0^1(\Omega)$ as well. Finally, for all $v \in H^1(\Omega)$, there holds the stability estimate*

$$\sum_{l=0}^{1} \delta^{2l} |K_B^\delta v|^2_{H^l(\Omega)} \lesssim \sum_{l=0}^{1} \delta^{2l} |v|^2_{H^l(\Omega\cap B^\delta)}.$$

Proof. The stability estimate follows from Leibniz' product rule for derivatives and the relation $|\kappa_B^\delta|_{W^{l,\infty}(\Omega)} \lesssim \delta^{-l}$, $l \in \mathbb{N}_0$. The remaining properties are immediate consequences of Lemma 3.4. $\square$

3.4. The spaces of discrete and harmonic functions

In this section, we introduce the space $\mathbb{S}_{\text{harm}}(B)$ of functions that are discrete and harmonic on some subset $B \subseteq \mathbb{R}^d$. As we actually want to have that $K_B^\delta : \mathbb{S}_{\text{harm}}(B) \to \mathbb{S}_{\text{harm}}(B)$, the space $\mathbb{S}_{\text{harm}}(B)$ now has to be infinite-dimensional.

Definition 3.7. Let $B \subseteq \mathbb{R}^d$. A function $u \in H_0^1(\Omega)$ is called ...

(1) ... *discrete on* B, if there exists a function $\tilde{u} \in \mathbb{S}_0^{p,1}(\mathcal{T})$ such that $u|_{\Omega\cap B} = \tilde{u}|_{\Omega\cap B}$.
(2) ... *harmonic on* B, if $a(u, v) = 0$ for all $v \in \mathbb{S}_0^{p,1}(\mathcal{T})$ with $\operatorname{supp}(v) \subseteq B$.

We define the space of *discrete and harmonic functions*,

$$\mathbb{S}_{\text{harm}}(B) := \{u \in H_0^1(\Omega) \,|\, u \text{ is discrete and harmonic on } B\} \subseteq H_0^1(\Omega).$$

Note that $\mathbb{S}_{\text{harm}}(B)$ consists of global functions $u : \Omega \longrightarrow \mathbb{R}$ that merely happen to have some additional properties on the subset $\Omega \cap B$. Furthermore, we emphasize that $\mathbb{S}_{\text{harm}}(B)$ is an infinite-dimensional space, in general.

The next lemma summarizes the relevant properties of these spaces. Recall the definition of the discrete solution operator $S_{\mathcal{T}} : L^2(\Omega) \longrightarrow \mathbb{S}_0^{p,1}(\mathcal{T})$ from Definition 3.1 and the cut-off operator $K_B^\delta : H^1(\Omega) \longrightarrow H^1(\Omega)$ from Definition 3.5.

Lemma 3.8.

(1) The subspace $\mathbb{S}_{\text{harm}}(B) \subseteq H_0^1(\Omega)$ is closed for open $B \subseteq \mathbb{R}^d$.
(2) For all $B \subseteq B^+ \subseteq \mathbb{R}^d$, there holds $\mathbb{S}_{\text{harm}}(B^+) \subseteq \mathbb{S}_{\text{harm}}(B)$.
(3) For all $B, D \subseteq \mathbb{R}^d$ with $B \cap D = \emptyset$ and all $f \in L^2(\Omega)$ with $\operatorname{supp}(f) \subseteq D$, there holds $S_{\mathcal{T}} f \in \mathbb{S}_{\text{harm}}(B)$.
(4) For all boxes $B \subseteq \mathbb{R}^d$, $\delta > 0$ and $u \in \mathbb{S}_{\text{harm}}(B)$, there holds $K_B^\delta u \in \mathbb{S}_{\text{harm}}(B)$.

Proof. We only show closedness. Since $\Omega \cap B \subseteq \mathbb{R}^d$ is open, the Sobolev space $H^1(\Omega \cap B)$ is well-defined. The subset $Z := \{\tilde{u}|_{\Omega\cap B} \,|\, \tilde{u} \in \mathbb{S}_0^{p,1}(\mathcal{T})\} \subseteq H^1(\Omega \cap B)$ is a finite-dimensional subspace and thus closed. Note that any given function $u \in H_0^1(\Omega)$ is discrete on B (in the sense of Definition 3.7), if and only if $u|_{\Omega\cap B} \in Z$.

Let $(u_n)_{n\in\mathbb{N}} \subseteq \mathbb{S}_{\text{harm}}(B)$ and $u \in H_0^1(\Omega)$ with $\|u - u_n\|_{H^1(\Omega)} \xrightarrow{n} 0$. In particular, for every $n \in \mathbb{N}$, we know that $u_n|_{\Omega\cap B} \in Z$ and that $a(u_n, v) = 0$ for all $v \in \mathbb{S}_0^{p,1}(\mathcal{T})$ with $\operatorname{supp}(v) \subseteq B$. The trivial bound $\|u - u_n\|_{H^1(\Omega\cap B)} \leq \|u - u_n\|_{H^1(\Omega)} \xrightarrow{n} 0$ and the closedness of Z immediately yield $u|_{\Omega\cap B} \in Z$, meaning that u is discrete on B. Finally, for all $v \in \mathbb{S}_0^{p,1}(\mathcal{T})$ with $\operatorname{supp}(v) \subseteq B$, we have

$$|a(u,v)| = |a(u-u_n,v)| \lesssim \|u-u_n\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)} \xrightarrow{n} 0,$$

indicating that u is harmonic on B. This concludes the proof of closedness. □

In the remainder of this section, we develop an improved version of the discrete Caccioppoli inequality from [1, Lem. 3.27]. This time we are interested in large polynomial degrees $p \to \infty$ as well. Therefore, we need to revisit our previous proof and keep track of p. Since the elementwise Lagrange interpolant $I^p_{\mathcal{T}} : C^0_{\mathrm{pw}}(\mathcal{T}) \longrightarrow \mathbb{S}^{p,0}(\mathcal{T})$ is not suitable for large p, we employ an alternative operator:

Lemma 3.9. *There exists a linear operator $J^p_{\mathcal{T}} : \mathbb{S}^{p+2,0}(\mathcal{T}) \longrightarrow \mathbb{S}^{p,0}(\mathcal{T})$ with the following properties:*

(1) Continuity and boundary values*: For all $v \in \mathbb{S}^{p+2,1}_0(\mathcal{T})$, there holds $J^p_{\mathcal{T}} v \in \mathbb{S}^{p,1}_0(\mathcal{T})$.*
(2) Supports*: For $v \in \mathbb{S}^{p+2,0}(\mathcal{T})$, there holds $\operatorname{supp}(J^p_{\mathcal{T}} v) \subseteq \operatorname{supp}(v)$.*
(3) Error bound*: Let $\sigma_{\mathrm{red}} := d(d+1)/4+2$. Then, for all $\kappa \in \mathbb{S}^{1,0}(\mathcal{T})$, all $u \in \mathbb{S}^{p,0}(\mathcal{T})$ and all $T \in \mathcal{T}$, there holds the error bound*

$$\sum_{l=0}^{1} h_T^l |(\mathrm{id} - J^p_{\mathcal{T}})(\kappa^2 u)|_{H^l(T)} \lesssim p^{\sigma_{\mathrm{red}}} h_T |\kappa^2|_{W^{1,\infty}(T)} \|u\|_{L^2(T)}.$$

For the sake of readability, we postpone the lengthy proof of Lemma 3.9 to Section 4 further below. Furthermore, we mention that the value of the constant σ_{red} is not optimal. (The subscript “red” is reminiscent of the fact that the operator $J^p_{\mathcal{T}}$ reduces the polynomial degree of its input.)

For the subsequent revision of [1, Lem. 3.27], we remind the reader of our definition of *inflated clusters*:

$$\forall \mathcal{B} \subseteq \mathcal{T} : \forall \delta > 0 : \qquad \mathcal{B}^\delta := \{T \in \mathcal{T} \mid \exists S \in \mathcal{B} : \|x_T - x_S\|_2 \le \delta\}.$$

Lemma 3.10. *Let $\mathcal{B} \subseteq \mathcal{T}$ be a collection of elements and $\delta > 0$ be a parameter satisfying $4\sigma^3_{\mathrm{shp}} h_{\max,\mathcal{B}} \le \delta \lesssim 1$. Let $u \in \mathbb{S}^{p,1}_0(\mathcal{T})$ be a function that satisfies $a(u,v) = 0$ for all $v \in \mathbb{S}^{p,1}_0(\mathcal{T})$ with $\operatorname{supp}(v) \subseteq \bigcup \mathcal{B}^\delta$. Then, with the constant $\sigma_{\mathrm{red}} \ge 2$ from Lemma 3.9, there holds the Caccioppoli inequality*

$$\delta |u|_{H^1(B)} \lesssim p^{\sigma_{\mathrm{red}}} \|u\|_{L^2(\mathcal{B}^\delta)}.$$

Proof. According to [1, Lem. 3.18], the assumption $4\sigma^3_{\mathrm{shp}} h_{\max,\mathcal{B}} \le \delta \lesssim 1$ allows us to construct a *discrete cut-off function* κ with the following properties:

$$\kappa \in \mathbb{S}^{1,1}(\mathcal{T}), \qquad \operatorname{supp}(\kappa) \subseteq \bigcup \mathcal{B}^\delta, \qquad \kappa|_{\bigcup \mathcal{B}} \equiv 1, \qquad 0 \le \kappa \le 1, \qquad \forall l \in \{0,1\} : |\kappa|_{W^{l,\infty}(\Omega)} \lesssim \delta^{-l}.$$

Let $u \in \mathbb{S}^{p,1}_0(\mathcal{T})$ be as in the statement of the lemma. We consider the function $v := J^p_{\mathcal{T}}(\kappa^2 u)$, where $J^p_{\mathcal{T}} : \mathbb{S}^{p+2,0}(\mathcal{T}) \longrightarrow \mathbb{S}^{p,0}(\mathcal{T})$ denotes the approximation operator from Lemma 3.9. Since $\kappa^2 u \in \mathbb{S}^{p+2,1}_0(\mathcal{T})$, we know that $v \in \mathbb{S}^{p,1}_0(\mathcal{T})$ and that $\operatorname{supp}(v) \subseteq \operatorname{supp}(\kappa^2 u) \subseteq \bigcup \mathcal{B}^\delta$. In particular, v is a viable test function, and we obtain the identity $a(u,v) = 0$. Using the constant $\sigma_{\mathrm{red}} \ge 2$ defined in Lemma 3.9, we compute

$$\begin{aligned}
a(u,\kappa^2 u) &= a(u,\kappa^2 u - v) = a(u,(\mathrm{id} - J^p_{\mathcal{T}})(\kappa^2 u)) \\
&\overset{\text{Definition 2.1}}{\lesssim} \sum_{T \in \mathcal{B}^\delta} \|u\|_{H^1(T)} \|(\mathrm{id} - J^p_{\mathcal{T}})(\kappa^2 u)\|_{H^1(T)} \overset{\text{Lemma 3.9}}{\lesssim} p^{\sigma_{\mathrm{red}}} \sum_{T \in \mathcal{B}^\delta} |\kappa^2|_{W^{1,\infty}(T)} \|u\|_{H^1(T)} \|u\|_{L^2(T)} \\
&\lesssim p^{\sigma_{\mathrm{red}}} \delta^{-1} \sum_{T \in \mathcal{B}^\delta} \|\kappa\|_{L^\infty(T)} \|u\|_{H^1(T)} \|u\|_{L^2(T)}.
\end{aligned}$$

Using Taylor's Theorem, we may estimate $|\kappa(y)| \le \min_{x \in T} |\kappa(x)| + h_T |\kappa|_{W^{1,\infty}(T)}$ for all $y \in T$. Together with a polynomial inverse inequality, see e.g. [21, Cor. 4.2], and the relation $\delta \lesssim 1 \le p$, we find that

$$\begin{aligned}
\|\kappa\|_{L^\infty(T)} \|u\|_{H^1(T)} &\lesssim \|\kappa\|_{L^\infty(T)} \|u\|_{L^2(T)} + \left(\min_{x \in T} |\kappa(x)| + h_T \delta^{-1}\right) |u|_{H^1(T)} \lesssim (1 + p^2 \delta^{-1}) \|u\|_{L^2(T)} + \min_{x \in T} |\kappa(x)| |u|_{H^1(T)} \\
&\lesssim p^2 \delta^{-1} \|u\|_{L^2(T)} + \|\kappa \nabla u\|_{L^2(T)},
\end{aligned}$$

which, using a weighted Young's inequality with parameter $\varepsilon > 0$, leads us to the following bound:

$$\begin{aligned}
a(u,\kappa^2 u) &\lesssim p^{\sigma_{\mathrm{red}}} \delta^{-1} \sum_{T \in \mathcal{B}^\delta} \|\kappa\|_{L^\infty(T)} \|u\|_{H^1(T)} \|u\|_{L^2(T)} \lesssim p^{\sigma_{\mathrm{red}}} \delta^{-1} \Big(\sum_{T \in \mathcal{B}^\delta} p^2 \delta^{-1} \|u\|^2_{L^2(T)} + \sum_{T \in \mathcal{B}^\delta} \|\kappa \nabla u\|_{L^2(T)} \|u\|_{L^2(T)} \Big) \\
&\lesssim p^{\sigma_{\mathrm{red}}+2} \delta^{-2} \|u\|^2_{L^2(\mathcal{B}^\delta)} + \sum_{T \in \mathcal{B}^\delta} \varepsilon \|\kappa \nabla u\|^2_{L^2(T)} + C_\varepsilon p^{2\sigma_{\mathrm{red}}} \delta^{-2} \|u\|^2_{L^2(T)} \\
&\overset{\sigma_{\mathrm{red}} \ge 2}{\lesssim} C_\varepsilon p^{2\sigma_{\mathrm{red}}} \delta^{-2} \|u\|^2_{L^2(\mathcal{B}^\delta)} + \varepsilon \|\kappa \nabla u\|^2_{L^2(\Omega)}.
\end{aligned}$$

On the other hand, we can use the definition of $a(\cdot,\cdot)$ from Definition 2.1 to expand the term $a(u,\kappa^2 u)$. One of the summands is amenable to the coercivity of the PDE coefficient a_1:

$$\begin{aligned}
\|\kappa \nabla u\|^2_{L^2(\Omega)} &\lesssim \langle a_1 \kappa \nabla u, \kappa \nabla u \rangle_{L^2(\Omega)} \\
&= a(u,\kappa^2 u) - 2\langle a_1 \kappa \nabla u, u \nabla \kappa \rangle_{L^2(\Omega)} - \langle a_2 \cdot \nabla u, \kappa^2 u \rangle_{L^2(\Omega)} - \langle a_3 u, \kappa^2 u \rangle_{L^2(\Omega)} \\
&\lesssim C_\varepsilon p^{2\sigma_{\mathrm{red}}} \delta^{-2} \|u\|^2_{L^2(\mathcal{B}^\delta)} + \varepsilon \|\kappa \nabla u\|^2_{L^2(\Omega)} + \|\kappa \nabla u\|_{L^2(\Omega)} \left(\|u \nabla \kappa\|_{L^2(\Omega)} + \|\kappa u\|_{L^2(\Omega)} \right) + \|\kappa u\|^2_{L^2(\Omega)}
\end{aligned}$$

$$\overset{\delta \lesssim 1 \le p}{\lesssim} C_\varepsilon p^{2\sigma_{\mathrm{red}}} \delta^{-2} \|u\|^2_{L^2(B^\delta)} + \varepsilon \|\kappa \nabla u\|^2_{L^2(\Omega)},$$

where the last estimate follows using Young's inequality and the properties of the cut-off function κ. Since the Young parameter ε can be chosen arbitrarily small, we may absorb the last summand in the left-hand side of the overall inequality. Finally, since $\kappa \equiv 1$ on $\mathcal{B}$, we obtain the desired Caccioppoli inequality:

$$|u|_{H^1(\mathcal{B})} \le \|\kappa \nabla u\|_{L^2(\Omega)} \lesssim p^{\sigma_{\mathrm{red}}} \delta^{-1} \|u\|_{L^2(B^\delta)}. \quad \square$$

We close this section with the promised improvement of the discrete Caccioppoli inequality. This time, it will be phrased in terms of the new spaces $\mathbb{S}_{\mathrm{harm}}(B)$ and $\mathbb{S}_{\mathrm{harm}}(B^\delta)$, where $B \subseteq \mathbb{R}^d$ is an axes-parallel box, $\delta > 0$ is a given parameter, and $B^\delta \subseteq \mathbb{R}^d$ is the inflated box in the sense of Definition 3.3. Most importantly, *no* lower bound on δ is assumed.

The basic idea is to decompose the set of elements touching the inner box B into two groups, based on the relative size of h_T and δ. The first group contains the elements that are small in relation to δ and can therefore be treated with Lemma 3.10. The second group contains the larger elements (relative to δ), and we can use an inverse inequality to derive the desired bound. However, since the larger elements might not be fully contained in the outer box B^δ, we have to break them up into smaller pieces first.

Lemma 3.11. *Denote by $\sigma_{\mathrm{shp}} \ge 1$ the shape-regularity constant from Assumption 2.3. Let $T \in \mathcal{T}$ and $\delta > 0$ be such that $16\sigma^4_{\mathrm{shp}} h_T > \delta$. Then, there exists a mesh $\mathcal{S} \subseteq \mathrm{Pow}(T)$ with the following properties:*

(1) For all $S, \tilde{S} \in \mathcal{T}$ with $S \ne \tilde{S}$, there holds $S \cap \tilde{S} = \emptyset$. Furthermore, $\overline{\bigcup \mathcal{S}} = \overline{T}$.
(2) There hold the bounds $h_{\max,\mathcal{S}} \le \delta \le C(d, \sigma_{\mathrm{shp}}) h_{\min,\mathcal{S}}$.
(3) The mesh $\mathcal{S}$ is shape-regular in the sense of Assumption 2.3 with a constant $\widetilde{\sigma}_{\mathrm{shp}} = C(d, \sigma_{\mathrm{shp}})$.

Proof. Denote by $\hat{T} \subseteq \mathbb{R}^d$ the reference simplex, let $M \in \mathbb{N}$ and set $J := \{1, \dots, M^d\}$. In [22], it was shown that $\hat{T}$ can be partitioned into M^d simplices $\hat{S}_1, \dots, \hat{S}_{M^d} \subseteq \hat{T}$ of at most $d!/2$ congruence classes, such that $|\hat{S}_j| = M^{-d} |\hat{T}|$. Since the number of congruence classes is uniformly bounded (independent of M), one can then show that $C^{-1} \le M h_{\hat{S}_j} \le C$, for some constant $C = C(d) \ge 1$.

Now, denote by $F_T : \hat{T} \longrightarrow T$ the affine diffeomorphism from Definition 2.2. Without proof, we mention that $\|\nabla F_T\|_2 \le \hat{r}^{-1} h_T$, where $\hat{r} > 0$ is the radius of the largest ball that can be inscribed into $\hat{T}$. Similarly, exploiting the shape regularity of the mesh $\mathcal{T}$ (cf. Assumption 2.3), there holds $\|\nabla(F_T^{-1})\|_2 \le (\sigma^{-1}_{\mathrm{shp}} h_T)^{-1} h_{\hat{T}} \lesssim h_T^{-1}$. Then, using the ceiling function $\lceil \cdot \rceil$, we choose

$$M := \lceil C \hat{r}^{-1} h_T \delta^{-1} \rceil \in \mathbb{N}$$

and argue that the system $\mathcal{S} := \{F_T(\hat{S}_j) \mid j \in J\}$ has the desired properties: Item (1) follows from the fact that the simplices $\hat{S}_j$ partition $\hat{T}$ and item (3) follows from the uniform bound on the number of congruence classes. Finally, to see item (2), we compute

$$h_{\max,\mathcal{S}} = \max_{j \in J} h_{F_T(\hat{S}_j)} \le \max_{j \in J} \|\nabla F_T\|_2 h_{\hat{S}_j} \le C \hat{r}^{-1} h_T M^{-1} \overset{\mathrm{Def.}\, M}{\le} \delta.$$

An analogous computation involving the inverse mapping $F_T^{-1} : T \longrightarrow \hat{T}$ reveals the bound $\min_{j \in J} h_{\hat{S}_j} \lesssim h_T^{-1} h_{\min,\mathcal{S}}$. Furthermore, we invoke the assumption $16\sigma^4_{\mathrm{shp}} h_T > \delta$ to conclude that $M \le C \hat{r}^{-1} h_T \delta^{-1} + 1 \lesssim h_T \delta^{-1}$. Combining both, we end up with the lower bound

$$h_T^{-1} \delta \lesssim C^{-1} M^{-1} \le \min_{j \in J} h_{\hat{S}_j} \lesssim h_T^{-1} h_{\min,\mathcal{S}},$$

which readily yields $\delta \lesssim h_{\min,\mathcal{S}}$. This concludes the proof. $\square$

Now that we know how to break up an element $T \in \mathcal{T}$ into smaller pieces, we present the updated proof of the discrete Caccioppoli inequality.

Lemma 3.12. *Let $B \subseteq \mathbb{R}^d$ be a box and $\delta > 0$ with $\delta \lesssim 1$. Denote by $\sigma_{\mathrm{red}} \ge 2$ the constant from Lemma 3.9. Then, for all $u \in \mathbb{S}_{\mathrm{harm}}(B^\delta)$, there holds the* Caccioppoli inequality

$$\delta |u|_{H^1(\Omega \cap B)} \lesssim p^{\sigma_{\mathrm{red}}} \|u\|_{L^2(\Omega \cap B^\delta)}.$$

Proof. First, we apply Lemma 3.10 to the collection $\mathcal{B} := \{T \in \mathcal{T} \mid \overline{T} \cap \overline{B} \ne \emptyset, 16\sigma^4_{\mathrm{shp}} h_T \le \delta\}$ and the parameter $\varepsilon := \delta/(4\sigma_{\mathrm{shp}}) > 0$. It is not difficult to see that $4\sigma^3_{\mathrm{shp}} h_{\max,\mathcal{B}} \le \varepsilon \lesssim 1$, meaning that ε is indeed a valid parameter choice. Next, let us demonstrate that $\bigcup \mathcal{B}^\varepsilon \subseteq \Omega \cap B^\delta$: Given $T \in \mathcal{B}^\varepsilon$, we know that there exists an element $S \in \mathcal{B}$ such that $\|x_T - x_S\|_2 \le \varepsilon$. Since S touches B, we can use two triangle inequalities to derive the inclusion $T \subseteq B^{h_{\max,\mathcal{B}} + \varepsilon + h_{\max,\mathcal{B}^\varepsilon}}$. From [1, Lem. 3.15], we know that $h_{\max,\mathcal{B}^\varepsilon} \le h_{\max,\mathcal{B}} + \sigma_{\mathrm{shp}} \varepsilon$, which implies $h_{\max,\mathcal{B}} + \varepsilon + h_{\max,\mathcal{B}^\varepsilon} \le 2\sigma_{\mathrm{shp}}(h_{\max,\mathcal{B}} + \varepsilon) \le \delta$ and ultimately $T \subseteq B^\delta$. Since $T \in \mathcal{B}^\varepsilon$ was arbitrary, we conclude that indeed $\bigcup \mathcal{B}^\varepsilon \subseteq \Omega \cap B^\delta$. Now, for every $u \in \mathbb{S}_{\mathrm{harm}}(B^\delta)$, we know from Definition 3.7 that there exists a function $\tilde{u} \in \mathbb{S}^{p,1}_0(\mathcal{T})$ such that $u|_{\Omega \cap B^\delta} = \tilde{u}|_{\Omega \cap B^\delta}$. Additionally, for all $v \in \mathbb{S}^{p,1}_0(\mathcal{T})$ with $\mathrm{supp}(v) \subseteq B^\delta$, we know that $a(u, v) = 0$. In particular, $a(\tilde{u}, v) = a(u, v) = 0$ as well, because v restricts the effective integration domain to $\Omega \cap B^\delta$, where u and $\tilde{u}$ coincide. In other words, we are allowed to apply Lemma 3.10 to the function $\tilde{u}$:

$$\delta |u|_{H^1(\mathcal{B})} \eqsim \varepsilon |\tilde{u}|_{H^1(\mathcal{B})} \lesssim p^{\sigma_{\mathrm{red}}} \|\tilde{u}\|_{L^2(\mathcal{B}^\varepsilon)} \le p^{\sigma_{\mathrm{red}}} \|\tilde{u}\|_{L^2(\Omega \cap B^\delta)} = p^{\sigma_{\mathrm{red}}} \|u\|_{L^2(\Omega \cap B^\delta)}.$$

Second, consider an element $T \in \mathcal{T}$ with $\overline{T} \cap \overline{B} \ne \emptyset$ and $16\sigma^4_{\mathrm{shp}} h_T > \delta$. Using Lemma 3.11, we can find a uniform mesh $\mathcal{S}$ such that $\overline{\bigcup \mathcal{S}} = \overline{T}$ and $h_{\max,\mathcal{S}} \le \delta \lesssim h_{\min,\mathcal{S}}$. Now consider the elements $\mathcal{S}_B := \{S \in \mathcal{S} \mid \overline{S} \cap \overline{B} \ne \emptyset\}$. Exploiting $\overline{\bigcup \mathcal{S}} = \overline{T}$, it is not difficult to show that $T \cap B \subseteq \overline{\bigcup \mathcal{S}_B}$. Furthermore, since

$h_{\max,\mathcal{S}} \le \delta$, an elementary geometric argument proves that $\bigcup \mathcal{S}_B \subseteq T \cap B^\delta$. Now, for every $u \in \mathbb{S}_{\mathrm{harm}}(B^\delta)$, we know from Definition 3.7 that there exists a function $\tilde{u} \in \mathbb{S}_0^{p,1}(\mathcal{T})$ such that $u|_{\Omega\cap B^\delta} = \tilde{u}|_{\Omega\cap B^\delta}$. Then, using the inverse inequality $h_S|\tilde{u}|_{H^1(S)} \lesssim p^2\|\tilde{u}\|_{L^2(S)}$, $S \in \mathcal{S}_B$, see e.g. [21], we get

$$\delta^2|u|^2_{H^1(T\cap B)} \lesssim h^2_{\min,\mathcal{S}}|\tilde{u}|^2_{H^1(\bigcup \mathcal{S}_B)} \le \sum_{S\in\mathcal{S}_B} h_S^2|\tilde{u}|^2_{H^1(S)} \lesssim p^4 \sum_{S\in\mathcal{S}_B} \|\tilde{u}\|^2_{L^2(S)} = p^4\|\tilde{u}\|^2_{L^2(\bigcup \mathcal{S}_B)} \le p^4\|u\|^2_{L^2(T\cap B^\delta)}.$$

Note that the implicit constant from the inverse inequality only depends on the shape regularity constant $\widetilde{\sigma}_{\mathrm{shp}} = C(d, \sigma_{\mathrm{shp}})$ from Lemma 3.11.

Finally, for every $u \in \mathbb{S}_{\mathrm{harm}}(B^\delta)$, we put the estimates for both groups of elements together:

$$\delta^2|u|^2_{H^1(\Omega\cap B)} = \delta^2 \sum_{\substack{T\in\mathcal{T}:\\ T\cap B\neq\emptyset}} |u|^2_{H^1(T\cap B)} \le \delta^2|u|^2_{H^1(B)} + \sum_{\substack{T\in\mathcal{T}:\\ T\cap B\neq\emptyset,\\ 16\sigma^4_{\mathrm{shp}}h_T>\delta}} \delta^2|u|^2_{H^1(T\cap B)} \lesssim (p^{2\sigma_{\mathrm{red}}} + p^4)\|u\|^2_{L^2(\Omega\cap B^\delta)}.$$

Noting $\sigma_{\mathrm{red}} \ge 2$, this finishes the proof. $\square$

3.5. The low-rank approximation operator

In our previous construction of the single-step coarsening operator $Q_B^\delta : \mathbb{S}_{\mathrm{harm}}(\mathcal{B}^\delta) \longrightarrow \mathbb{S}_{\mathrm{harm}}(\mathcal{B})$, [1, Thm. 3.31], we used the orthogonal projection $\Pi_{\mathcal{S}}^p : L^2(\Omega) \longrightarrow \mathbb{S}^{p,0}(\mathcal{S})$ on a uniform mesh $\mathcal{S}$ to reduce the overall rank. The output was subsequently fed into the orthogonal projection $P_{\mathcal{B}} : L^2(\Omega) \longrightarrow \mathbb{S}_{\mathrm{harm}}(\mathcal{B})$ in order to generate an element of $\mathbb{S}_{\mathrm{harm}}(\mathcal{B})$ again. The existence of $P_{\mathcal{B}}$ hinged on the fact that $\mathbb{S}_{\mathrm{harm}}(\mathcal{B})$ was finite-dimensional and thus a closed subspace of $L^2(\Omega)$. However, according to Lemma 3.8, the updated spaces $\mathbb{S}_{\mathrm{harm}}(B)$ from Definition 3.7 are closed subspaces of $H^1(\Omega)$, rather than $L^2(\Omega)$. Therefore, we now have to use the orthogonal projection $P_B : H^1(\Omega) \longrightarrow \mathbb{S}_{\mathrm{harm}}(B)$, and a replacement $\Pi_H : H^1(\Omega) \longrightarrow H^1(\Omega)$ for the orthogonal projection $\Pi_{\mathcal{S}}^p : L^2(\Omega) \longrightarrow \mathbb{S}^{p,0}(\mathcal{S})$ is needed.

Lemma 3.13. *Let $H > 0$ be a free parameter. Then, there exists a* low-rank approximation operator

$$\Pi_H : H^1(\Omega) \longrightarrow H^1(\Omega)$$

with the following properties:

(1) Local rank: *For all boxes $B \subseteq \mathbb{R}^d$, there holds the dimension bound*

$$\dim\{\Pi_H v \mid v \in H^1(\Omega) \text{ with } \mathrm{supp}(v) \subseteq B\} \lesssim (1 + \mathrm{diam}_2(B)/H)^d.$$

(2) Error bound: *For all $v \in H^1(\Omega)$, there holds the global error bound*

$$\sum_{l=0}^{1} H^{2l}|v - \Pi_H v|^2_{H^l(\Omega)} \lesssim H^2|v|^2_{H^1(\Omega)}.$$

Proof. Using successive refinements of an arbitrary initial mesh, we can construct a uniform mesh $\mathcal{S} \subseteq \mathrm{Pow}(\Omega)$ with $h_{\max,\mathcal{S}} \le H \lesssim h_{\min,\mathcal{S}}$. Denote by $\mathcal{N}$ the set of nodes and by $\{g_N \mid N \in \mathcal{N}\} \subseteq \mathbb{S}^{1,1}(\mathcal{S})$ the corresponding basis of hat functions. We choose the classical *Clément operator* $\Pi_H : L^2(\Omega) \longrightarrow \mathbb{S}^{1,1}(\mathcal{S})$ from [16], which maps any given input $v \in L^2(\Omega)$ to the linear combination $\Pi_H v := \sum_{N\in\mathcal{N}} v_N g_N$, where $v_N \in \mathbb{R}$ is the mean value of v over the support of g_N. While the error bound is common knowledge (e.g., [16, Thm. 1]), the dimension bound amounts to counting the number of mesh elements lying inside the slightly inflated box B^H:

$$\#\{S \in \mathcal{S} \mid S \subseteq B^H\} \le H^{-d} \sum_{S\subseteq B^H} h_S^d \lesssim H^{-d} \sum_{S\subseteq B^H} |S| \le H^{-d}|B^H| \lesssim (1 + \mathrm{diam}_2(B)/H)^d. \quad \square$$

3.6. The coarsening operators

At this point, we present the construction of the single-step coarsening operator used for low dimensional approximation. The proof is less complicated than in our previous work, because the tedious case analysis for the parameter δ has become obsolete.

Theorem 3.14. *Let $B \subseteq \mathbb{R}^d$ be a box and $\delta > 0$ be a free parameter with $\delta \lesssim 1$. Denote by $\sigma_{\mathrm{red}} \ge 2$ the constant from Lemma 3.9. Then, there exists a linear* single-step coarsening operator

$$Q_B^\delta : \mathbb{S}_{\mathrm{harm}}(B^\delta) \longrightarrow \mathbb{S}_{\mathrm{harm}}(B)$$

with the following properties:

(1) Rank bound: *The rank is bounded by*

$$\mathrm{rank}(Q_B^\delta) \lesssim p^{d\sigma_{\mathrm{red}}}(1 + \mathrm{diam}_2(B)/\delta)^d.$$

(2) Approximation error: *For all $u \in \mathbb{S}_{\mathrm{harm}}(B^\delta)$, there holds the error bound*

$$\|u - Q_B^\delta u\|_{L^2(\Omega\cap B)} \le \frac{1}{2}\|u\|_{L^2(\Omega\cap B^\delta)}.$$

Proof. Denote by $K_B^{\delta/2} : H^1(\Omega) \longrightarrow H^1(\Omega)$ the cut-off operator from Definition 3.5. Next, let $H > 0$ and denote by $\Pi_H : H^1(\Omega) \longrightarrow H^1(\Omega)$ the low-rank approximation operator from Lemma 3.13. Furthermore, since $\mathbb{S}_{\text{harm}}(B) \subseteq H^1(\Omega)$ is a closed subspace (cf. Lemma 3.8), we may introduce the orthogonal projection $P_B : H^1(\Omega) \longrightarrow \mathbb{S}_{\text{harm}}(B)$ with respect to the equivalent norm $\sum_{l=0}^{1} H^{2l}|\cdot|^2_{H^l(\Omega)} = \|\cdot\|^2_{L^2(\Omega)} + H^2|\cdot|^2_{H^1(\Omega)}$. We then define the combined operator

$$Q_B^\delta := P_B \Pi_H K_B^{\delta/2} : \mathbb{S}_{\text{harm}}(B^\delta) \longrightarrow \mathbb{S}_{\text{harm}}(B).$$

First, we establish the error bound: Let $u \in \mathbb{S}_{\text{harm}}(B^\delta)$. From Lemma 3.8 we know that $u \in \mathbb{S}_{\text{harm}}(B)$ and that $K_B^{\delta/2}u \in \mathbb{S}_{\text{harm}}(B)$. Since P_B is a projection onto $\mathbb{S}_{\text{harm}}(B)$, it follows that $P_B K_B^{\delta/2}u = K_B^{\delta/2}u$. Then, using Lemma 3.6, we get the identity $u|_{\Omega\cap B} = (K_B^{\delta/2}u)|_{\Omega\cap B} = (P_B K_B^{\delta/2}u)|_{\Omega\cap B}$. We compute

$$\begin{aligned}
\sum_{l=0}^{1} H^{2l}|u - Q_B^\delta u|^2_{H^l(\Omega\cap B)} &= \sum_{l=0}^{1} H^{2l}|P_B K_B^{\delta/2}u - P_B\Pi_H K_B^{\delta/2}u|^2_{H^l(\Omega\cap B)} = \sum_{l=0}^{1} H^{2l}|P_B(\mathrm{id} - \Pi_H)(K_B^{\delta/2}u)|^2_{H^l(\Omega)} \\
&\le \sum_{l=0}^{1} H^{2l}|(\mathrm{id} - \Pi_H)(K_B^{\delta/2}u)|^2_{H^l(\Omega)} \overset{\text{Lemma 3.13}}{\lesssim} H^2|K_B^{\delta/2}u|^2_{H^1(\Omega)} \\
&\overset{\text{Lemma 3.6}}{\lesssim} (H/\delta)^2 \sum_{l=0}^{1} \delta^{2l}|u|^2_{H^l(\Omega\cap B^{\delta/2})} \overset{\text{Lemma 3.12}}{\lesssim} p^{2\sigma_{\text{red}}}(H/\delta)^2\|u\|^2_{L^2(\Omega\cap B^\delta)}.
\end{aligned}$$

Now, denote the implicit cumulative constant by $C > 0$. Then, by choosing $H := \delta/(2\sqrt{C}p^{\sigma_{\text{red}}}) > 0$, we get the desired factor $1/2$.

Finally, the rank bound can be seen as follows:

$$\begin{aligned}
\mathrm{rank}(Q_B^\delta) &= \dim\{P_B\Pi_H K_B^{\delta/2}u \mid u \in \mathbb{S}_{\text{harm}}(B^\delta)\} \overset{\text{Lemma 3.6}}{\le} \dim\{\Pi_H v \mid v \in H^1(\Omega) \text{ with } \mathrm{supp}(v) \subseteq B^{\delta/2}\} \\
&\overset{\text{Lemma 3.13}}{\lesssim} (1 + \mathrm{diam}_2(B^{\delta/2})/H)^d \underset{\delta\approx Hp^{\sigma_{\text{red}}}}{\lesssim} p^{d\sigma_{\text{red}}}(1 + \mathrm{diam}_2(B)/\delta)^d.
\end{aligned}$$

This concludes the proof. □

Now that the new version of the *single*-step coarsening operator Q_B^δ is established, the *multi*-step coarsening operator $Q_B^{\delta,L}$ can be constructed by iterating the updated single-step operator.

Theorem 3.15. *Let $B \subseteq \mathbb{R}^d$ be a box and $\delta > 0$ be a free parameter with $\delta \lesssim 1$. Furthermore, let $L \in \mathbb{N}$. Denote by $\sigma_{\text{red}} \ge 2$ the constant from Lemma 3.9. Then, there exists a linear* multi-step coarsening operator

$$Q_B^{\delta,L} : \mathbb{S}_{\text{harm}}(B^{\delta L}) \longrightarrow \mathbb{S}_{\text{harm}}(B)$$

with the following properties:

(1) Rank bound: *The rank is bounded by*

$$\mathrm{rank}(Q_B^{\delta,L}) \lesssim p^{d\sigma_{\text{red}}}(L + \mathrm{diam}_2(B)/\delta)^{d+1}.$$

(2) Approximation error: *For all $u \in \mathbb{S}_{\text{harm}}(B^{\delta L})$, there holds the error bound*

$$\|u - Q_B^{\delta,L}u\|_{L^2(\Omega\cap B)} \le 2^{-L}\|u\|_{L^2(\Omega\cap B^{\delta L})}.$$

Proof. For the box $B \subseteq \mathbb{R}^d$, $\delta > 0$, and $L \in \mathbb{N}$, we define the nested concentric boxes $B_l := B^{\delta l}$, $l \in \{0, \dots, L\}$. Using the corresponding single-step coarsening operators $Q_l := Q_{B_l}^\delta : \mathbb{S}_{\text{harm}}(B_{l+1}) \longrightarrow \mathbb{S}_{\text{harm}}(B_l)$ from Theorem 3.14, we make the following definition:

$$\forall u \in \mathbb{S}_{\text{harm}}(B^{\delta L}) : \qquad Q_B^{\delta,L}u := u - (\mathrm{id} - Q_0)\circ \dots \circ(\mathrm{id} - Q_{L-1})(u) \in \mathbb{S}_{\text{harm}}(B).$$

This gives rise to the rank bound

$$\begin{aligned}
\mathrm{rank}(Q_B^{\delta,L}) &\le \sum_{l=0}^{L-1} \mathrm{rank}(Q_l) \overset{\text{Theorem 3.14}}{\lesssim} \sum_{l=0}^{L-1} p^{d\sigma_{\text{red}}}(1 + \delta^{-1}\mathrm{diam}_2(B_l))^d \\
&\lesssim p^{d\sigma_{\text{red}}} \sum_{l=0}^{L-1}(1 + l + \delta^{-1}\mathrm{diam}_2(B))^d \le p^{d\sigma_{\text{red}}} \sum_{l=0}^{L-1}(L + \delta^{-1}\mathrm{diam}_2(B))^d \\
&\le p^{d\sigma_{\text{red}}} L(L + \delta^{-1}\mathrm{diam}_2(B))^d \le p^{d\sigma_{\text{red}}}(L + \delta^{-1}\mathrm{diam}_2(B))^{d+1}.
\end{aligned}$$

Finally, the definition of $Q_B^{\delta,L}$ directly provides the error bound: For every $u \in \mathbb{S}_{\text{harm}}(B^{\delta L})$, iteration of Theorem 3.14 gives

$$\|u - Q_B^{\delta,L}u\|_{L^2(\Omega\cap B)} = \|(\mathrm{id} - Q_0)\circ \dots \circ(\mathrm{id} - Q_{L-1})(u)\|_{L^2(\Omega\cap B)} \le 2^{-L}\|u\|_{L^2(B^{\delta L})},$$

which finishes the proof. □

3.7. Putting everything together

In this section, we finally prove our main result, Theorem 2.13. Recall from Section 3.2 that we need to approximate the admissible blocks $\boldsymbol{A}^{-1}|_{I\times J}$ by low-rank matrices in order to get an $\mathcal{H}$-matrix approximation to the full matrix $\boldsymbol{A}^{-1}$. Then, Lemma 3.2 translated the problem into the realm of function spaces, implying that a suitable subspace $V \subseteq L^2(\Omega)$ needs to be constructed. Here, the range of the multi-step coarsening operator $Q_B^{\delta,L}$ does the trick:

Theorem 3.16. *Let $B, D \subseteq \mathbb{R}^d$ be two boxes with $0 < \mathrm{diam}_2(B) \le \sigma_{\mathrm{adm}} \mathrm{dist}_2(B, D)$. Furthermore, let $L \in \mathbb{N}$. Denote by $\sigma_{\mathrm{red}} \ge 2$ the constant from Lemma 3.9. Then, there exists a subspace*

$$V_{B,D,L} \subseteq L^2(\Omega)$$

with the following properties:

(1) Dimension bound: *There holds the dimension bound*

$$\dim V_{B,D,L} \lesssim p^{d\sigma_{\mathrm{red}}} L^{d+1}.$$

(2) Approximation property: *For every $f \in L^2(\Omega)$ with $\mathrm{supp}(f) \subseteq D$, there holds the error bound*

$$\inf_{v \in V_{B,D,L}} \|S_{\mathcal{T}} f - v\|_{L^2(\Omega \cap B)} \lesssim 2^{-L} \|f\|_{L^2(\Omega)}.$$

Proof. Let $B, D \subseteq \mathbb{R}^d$ and $L \in \mathbb{N}$ as above. Set $\delta := \mathrm{diam}_2(B)/(2\sqrt{d}\sigma_{\mathrm{adm}} L) > 0$ and denote by $Q_B^{\delta,L} : \mathbb{S}_{\mathrm{harm}}(B^{\delta L}) \longrightarrow \mathbb{S}_{\mathrm{harm}}(B)$ the multi-step coarsening operator from Theorem 3.15. We choose the space

$$V_{B,D,L} := \mathrm{ran}(Q_B^{\delta,L}) \subseteq \mathbb{S}_{\mathrm{harm}}(B) \subseteq L^2(\Omega).$$

Using Theorem 3.15 and the definition of δ, we can bound the dimension as follows:

$$\dim V_{B,D,L} = \mathrm{rank}(Q_B^{\delta,L}) \lesssim p^{d\sigma_{\mathrm{red}}} (L + \mathrm{diam}_2(B)/\delta)^{d+1} \lesssim p^{d\sigma_{\mathrm{red}}} L^{d+1}.$$

Finally, let $f \in L^2(\Omega)$ with $\mathrm{supp}(f) \subseteq D$. In order to show that the error bound from Theorem 3.15 is applicable to the function $S_{\mathcal{T}} f \in \mathbb{S}_0^{p,1}(\mathcal{T})$, we first need to establish the fact that $S_{\mathcal{T}} f \in \mathbb{S}_{\mathrm{harm}}(B^{\delta L})$. According to Lemma 3.8, it suffices to prove that the sets $B^{\delta L}$ and D are disjoint. To that end, we choose a point $z \in \overline{B^{\delta L}}$ with $\mathrm{dist}_2(B^{\delta L}, D) = \mathrm{dist}_2(z, D)$. Then, $\mathrm{dist}_2(B, D) \le \mathrm{dist}_2(B, z) + \mathrm{dist}_2(z, D) \le \sqrt{d}\delta L + \mathrm{dist}_2(B^{\delta L}, D)$. Combined with the definition of δ and the admissibility condition, this yields

$$\mathrm{dist}_2(B^{\delta L}, D) \ge \mathrm{dist}_2(B, D) - \sqrt{d}\delta L = \mathrm{dist}_2(B, D) - \mathrm{diam}_2(B)/(2\sigma_{\mathrm{adm}}) \ge \mathrm{diam}_2(B)/(2\sigma_{\mathrm{adm}}) > 0.$$

Lemma 3.8 implies $S_{\mathcal{T}} f \in \mathbb{S}_{\mathrm{harm}}(B^{\delta L})$, so that $Q_B^{\delta,L}(S_{\mathcal{T}} f) \in V_{B,D,L}$. Hence, the error bound from Theorem 3.15 is applicable to the function $S_{\mathcal{T}} f$. Using the *a priori* stability bound of the discrete solution operator $S_{\mathcal{T}}$ (cf. Definition 3.1), we then estimate

$$\inf_{v \in V_{B,D,L}} \|S_{\mathcal{T}} f - v\|_{L^2(\Omega \cap B)} \le \|S_{\mathcal{T}} f - Q_B^{\delta,L}(S_{\mathcal{T}} f)\|_{L^2(\Omega \cap B)} \le 2^{-L} \|S_{\mathcal{T}} f\|_{L^2(\Omega \cap B^{\delta L})} \lesssim 2^{-L} \|f\|_{L^2(\Omega)}.$$

This concludes the proof. □

Finally, we have everything we need to derive our main result:

Proof of Theorem 2.13. Let $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ be the system matrix from Definition 2.9 and $r \in \mathbb{N}$ a given block rank bound. We define the asserted $\mathcal{H}$-matrix approximant $\boldsymbol{B} \in \mathbb{R}^{N \times N}$ in a block-wise fashion:

First, consider an admissible block $(I, J) \in \mathbb{P}_{\mathrm{adm}}$. From Definition 2.11 we know that there exist boxes $B_I, B_J \subseteq \mathbb{R}^d$ with $\Omega_I \subseteq B_I$, $\Omega_J \subseteq B_J$ and $\mathrm{diam}_2(B_I) \le \sigma_{\mathrm{adm}} \mathrm{dist}_2(B_I, B_J)$. In particular, $\mathrm{diam}_2(B_I) \ge \mathrm{diam}_2(\Omega_I) > 0$, so that Theorem 3.16 is applicable to B_I and B_J. Now, denote by $C > 0$ the implicit constant from the dimension bound in Theorem 3.16. We set $\sigma_{\exp} := \ln(2)/C^{1/(d+1)} > 0$ and $L := \lfloor (r/C)^{1/(d+1)} p^{-\sigma_{\mathrm{red}}} \rfloor \in \mathbb{N}$. Then, Theorem 3.16 provides a subspace $V_{I,J,r} \subseteq V$. We apply Lemma 3.2 to this subspace and get an integer $\tilde{r} \le \dim V_{I,J,r}$ and matrices $\boldsymbol{X}_{I,J,r} \in \mathbb{R}^{I \times \tilde{r}}$ and $\boldsymbol{Y}_{I,J,r} \in \mathbb{R}^{J \times \tilde{r}}$. We set

$$\boldsymbol{B}|_{I \times J} := \boldsymbol{X}_{I,J,r} (\boldsymbol{Y}_{I,J,r})^T.$$

Second, for every small block $(I, J) \in \mathbb{P}_{\mathrm{small}}$, we make the trivial choice

$$\boldsymbol{B}|_{I \times J} := \boldsymbol{A}^{-1}|_{I \times J}.$$

By Definition 2.12, we have $\boldsymbol{B} \in \mathcal{H}(\mathbb{P}, \tilde{r})$ with a block rank bound

$$\tilde{r} \le \dim V_{I,J,r} \overset{\text{Def. } C}{\le} C p^{d\sigma_{\mathrm{red}}} L^{d+1} \le C (p^{\sigma_{\mathrm{red}}} L)^{d+1} \overset{\text{Def. } L}{\le} r.$$

As for the error, we get

$$\begin{aligned}
\|\boldsymbol{A}^{-1} - \boldsymbol{B}\|_2 &\overset{\text{Definition 2.11}}{\lesssim} \ln(h_{\min}^{-d}) \max_{(I,J) \in \mathbb{P}_{\mathrm{adm}}} \|\boldsymbol{A}^{-1}|_{I \times J} - \boldsymbol{X}_{I,J,r} (\boldsymbol{Y}_{I,J,r})^T\|_2 \\
&\overset{\text{Lemma 3.2}}{\lesssim} p^{2\sigma_{\mathrm{stab}}} \ln(h_{\min}^{-d}) h_{\min}^{-d} \max_{(I,J) \in \mathbb{P}_{\mathrm{adm}}} \sup_{\substack{f \in V_{I,J,r}: \\ \mathrm{supp}(f) \subseteq \Omega_J}} \inf_{v \in V_{I,J,r}} \frac{\|S_{\mathcal{T}} f - v\|_{L^2(\Omega_I)}}{\|f\|_{L^2(\Omega)}} \\
&\overset{\text{Theorem 3.16}}{\lesssim} p^{2\sigma_{\mathrm{stab}}} \ln(h_{\min}^{-d}) h_{\min}^{-d} 2^{-L}
\end{aligned}$$

$$\overset{\text{Def. } L}{\lesssim} \quad p^{2\sigma_{\text{stab}}} \ln(h_{\min}^{-d}) h_{\min}^{-d} \exp(-\ln(2)(r/C)^{1/(d+1)} p^{-\sigma_{\text{red}}})$$
$$\overset{\text{Def. } \sigma_{\exp}}{=} \quad p^{2\sigma_{\text{stab}}} \ln(h_{\min}^{-d}) h_{\min}^{-d} \exp(-\sigma_{\exp} r^{1/(d+1)} p^{-\sigma_{\text{red}}}),$$

which finishes the proof. □

4. Polynomial preserving lifting from the boundary and an elementwise defined projection

In this section, we provide the proof of Lemma 3.9, i.e., we devise an approximation operator $J_{\mathcal{T}}^p : \mathbb{S}^{p+2,0}(\mathcal{T}) \longrightarrow \mathbb{S}^{p,0}(\mathcal{T})$ that is defined in a elementwise fashion and preserves global continuity. In other words, we need to approximate a spline $u \in \mathbb{S}_0^{p+2,1}(\mathcal{T})$ of degree $p+2$ by a spline $\tilde{u} \in \mathbb{S}_0^{p,1}(\mathcal{T})$ of degree p in a way that is stable in p.

The subsequent construction of $J_{\mathcal{T}}^p$ generalizes [38, Lem. 4.1, Def. 2.5, Def. 2.1] from $d \in \{1,2,3\}$ to arbitrary spatial dimensions $d \geq 1$. In [38], $J_{\mathcal{T}}^p$ was defined in a piecewise manner by means of an operator $\hat{J}^p : H^{(d+1)/2}(\hat{T}) \longrightarrow \mathbb{P}^p(\hat{T})$ on the reference simplex $\hat{T} \subseteq \mathbb{R}^d$. While the results from [38] produce the optimal powers of p, the proofs are rather involved due to the nonlocality of the pertinent fractional Sobolev norms. In the present paper, we only need the specific case of *polynomial* inputs $f \in \mathbb{P}^{p+2}(\hat{T})$. Therefore, using inverse inequalities, we may work with the much simpler norms $\|\cdot\|_{L^2(\hat{T})}$ and $\|\cdot\|_{L^2(\partial\hat{T})}$ at the expense of powers of p. The definition of the operator $\hat{J}^p$ from [38] can easily be generalized to arbitrary space dimension $d \geq 1$. However, in order to derive error estimates, a polynomial preserving lifting operator has to be used. The literature on polynomial preserving liftings is extensive (e.g., [13,3,4,37,8,5]), but many authors focus on the special cases $d \in \{2,3\}$ and stability estimates are usually phrased in terms of the norm $\|\cdot\|_{H^{1/2}(\partial\hat{T})}$. In the sequel, we present a polynomial preserving lifting for arbitrary space dimension $d \geq 1$ that seems to have been overlooked in the pertinent literature. As usual, we first devise a lifting from one of $\hat{T}$'s hyperplanes $\hat{\Gamma} \subseteq \partial\hat{T}$ into its interior (cf. Lemma 4.3). Then, in Lemma 4.4, we combine the liftings of all such $\hat{\Gamma}$.

To get things going, let $d \geq 1$ as before and consider the *reference d-simplex* $\hat{T} := \hat{T}^d := \{x \in [0,1]^d \mid \|x\|_1 \leq 1\} \subseteq \mathbb{R}^d$. (In this section, we use the closed version in order to ease notation.) We denote by $\mathcal{N}(\hat{T}) := \{0, e_1, \dots, e_d\}$ its set of nodes, $0 \in \mathbb{R}^d$ being the origin and $e_i \in \mathbb{R}^d$ being the i-th Euclidean unit vector. In order to describe the boundary $\partial\hat{T}$ efficiently, let us introduce *k-simplices*. The definition uses the notion of convex hulls, $\text{conv}(\Omega) := \{(1-t)x + ty \mid x, y \in \Omega, t \in [0,1]\}$ for all $\Omega \subseteq \mathbb{R}^d$.

Definition 4.1. Let $k \in \{0, \dots, d\}$. A subset $\hat{\Sigma} \subseteq \hat{T}$ is called *k-simplex*, if there exist $k+1$ distinct nodes $\hat{N}_0, \dots, \hat{N}_k \in \mathcal{N}(\hat{T})$ such that $\hat{\Sigma} = \text{conv}\{\hat{N}_0, \dots, \hat{N}_k\}$.

(Again, we think of k-simplices $\hat{\Sigma}$ as being closed.) Note that $\hat{\Sigma} \subseteq \partial\hat{T}$, if $k \leq d-1$, and $\hat{\Sigma} = \hat{T}$, if $k = d$. Recall that any k-simplex $\hat{\Sigma} = \text{conv}\{\hat{N}_0, \dots, \hat{N}_k\}$ is isomorphic to the reference k-simplex $\hat{T}^k = \{t \in [0,1]^k \mid \|t\|_1 \leq 1\} \subseteq \mathbb{R}^k$. In fact, consider the affine parametrization $\sigma : \hat{T}^k \longrightarrow \hat{\Sigma}$, $\sigma(t) := \hat{N}_0 + \sum_{i=1}^k t_i(\hat{N}_i - \hat{N}_0)$. Then, there holds the representation

$$\hat{\Sigma} = \text{conv}\{\hat{N}_0, \dots, \hat{N}_k\} = \{\sigma(t) \mid t \in \hat{T}^k\}. \tag{4.1}$$

Next, let us introduce some function spaces.

Definition 4.2. Let $k \in \{0, \dots, d\}$ and consider a k-simplex $\hat{\Sigma} \subseteq \hat{T}$ along with an affine parametrization $\sigma : \hat{T}^k \longrightarrow \hat{\Sigma}$. We define the spaces

$$\begin{aligned}
\mathbb{P}^p(\hat{\Sigma}) &:= \{f : \hat{\Sigma} \longrightarrow \mathbb{R} \mid f \circ \sigma \in \mathbb{P}^p(\hat{T}^k)\}, \\
\mathbb{P}^p(\partial\hat{T}) &:= \{f \in C^0(\partial\hat{T}) \mid \forall (d-1)\text{-simplices } \hat{\Gamma} \subseteq \hat{T} : f|_{\hat{\Gamma}} \in \mathbb{P}^p(\hat{\Gamma})\}.
\end{aligned}$$

Before we construct the lifting operator in an arbitrary space dimension $d \geq 1$, let us first look at the case $d = 3$, i.e., $\hat{T} = \{(x_1, x_2, x_3) \mid x_i \in [0,1], x_1 + x_2 + x_3 \leq 1\}$. We enumerate the nodes as $\hat{N}_0 := (0,0,0)$, $\hat{N}_1 := (1,0,0)$, $\hat{N}_2 := (0,1,0)$ and $\hat{N}_3 := (0,0,1)$. Now, looking at Fig. 2, our goal is to find a lifting from the bottom face $\hat{\Gamma} := \text{conv}\{\hat{N}_0, \hat{N}_1, \hat{N}_2\} = \{x \in \hat{T} \mid x_3 = 0\}$ upwards, into the x_3-dimension.

Consider given boundary data $f \in C^0(\hat{\Gamma})$. Given any point $y \in \hat{\Gamma}$, the basic idea is to propagate the value $f(y)$ along the line segment from y to $\hat{N}_3$. To be more precise, let us denote the result of the lifting process by $\hat{L}_{\hat{\Gamma}} f \in C^0(\hat{T})$. In order to define the value $(\hat{L}_{\hat{\Gamma}} f)(x)$, for any given point $x \in \hat{T} \backslash \{\hat{N}_3\}$, we proceed as follows: First, we cast a ray from the top node $\hat{N}_3$ through the given point x and compute the intersection point with the bottom face $\hat{\Gamma}$. In fact, this intersection point is given by $P(x) := (x_1/(1-x_3), x_2/(1-x_3), 0)$. Then, we set

$$(\hat{L}_{\hat{\Gamma}} f)(x) := (1-x_3)^p f(P(x)) = (1-x_3)^p f\left(\frac{x_1}{1-x_3}, \frac{x_2}{1-x_3}, 0\right).$$

The purpose of the prefactor $(1-x_3)^p$ is to guarantee that $\hat{L}_{\hat{\Gamma}} f \in \mathbb{P}^p(\hat{T})$, whenever $f \in \mathbb{P}^p(\hat{\Gamma})$, by "undoing" the division by $(1-x_3)$ inside the argument of f. In fact, if we plug in $f(x) = \sum_{|\alpha| \leq p} f_\alpha x^\alpha$, we can see that

$$(\hat{L}_{\hat{\Gamma}} f)(x) = (1-x_3)^p f\left(\frac{x_1}{1-x_3}, \frac{x_2}{1-x_3}, 0\right) = \sum_{|\alpha| \leq p} f_\alpha (1-x_3)^{p-|\alpha|} x_1^{\alpha_1} x_2^{\alpha_2} 0^{\alpha_3} \in \mathbb{P}^p(\hat{T}).$$

Note that, since f is bounded, we get the added benefit of $\lim_{x \to \hat{N}_3} (\hat{L}_{\hat{\Gamma}} f)(x) = 0$. Finally, the prefactor satisfies $(1-x_3)^p = 1$, for all $x \in \hat{\Gamma}$, so that $(\hat{L}_{\hat{\Gamma}} f)|_{\hat{\Gamma}} = f$.

Now let us have a look at what happens at the edges and faces connecting $\hat{\Gamma}$ with $\hat{N}_3$. Assume, for example, that the input $f \in C^0(\hat{\Gamma})$ vanishes at one of $\hat{\Gamma}$'s nodes, say, $f(\hat{N}_0) = 0$. Then it is immediately clear from the picture in Fig. 2 that $(\hat{L}_{\hat{\Gamma}} f)(x) = 0$ for all x on the edge $\text{conv}\{\hat{N}_0, \hat{N}_3\}$. (More rigorously, if $x = (0,0,x_3)$, then $(\hat{L}_{\hat{\Gamma}} f)(x) = (1-x_3)^p f(0,0,0) = 0$.) As another example, consider the case of an input $g \in C^0(\hat{\Gamma})$ that vanishes on one of $\hat{\Gamma}$'s edges, say, $g = 0$ on $\text{conv}\{\hat{N}_0, \hat{N}_1\}$. Then, again by Fig. 2, we can see that $(\hat{L}_{\hat{\Gamma}} g)(x) = 0$ for all x on the face $\text{conv}\{\hat{N}_0, \hat{N}_1, \hat{N}_3\}$. As a mnemonic, we may say that the operator $\hat{L}_{\hat{\Gamma}}$ lifts zeros on k-simplices to zeros on $(k+1)$-simplices.

This concludes our introductory example in $d = 3$ space dimensions and we are now ready to treat the general case $d \geq 1$.

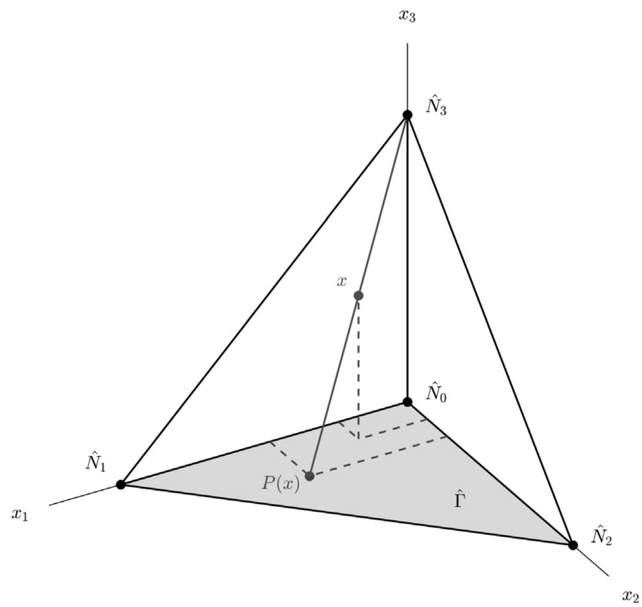

Fig. 2. The lifting operator in the case $d = 3$.

Lemma 4.3. *Let $\hat\Gamma \subseteq \hat T$ be a $(d-1)$-simplex, say, $\hat\Gamma = \operatorname{conv}\{\hat N_0, \dots, \hat N_{d-1}\}$. Denote the remaining node of $\hat T$ by $\hat N_d \in \mathcal{N}(\hat T)$. Then, there exists a lifting operator $\hat L_{\hat\Gamma} : C^0(\hat\Gamma) \longrightarrow C^0(\hat T)$ with the following properties:*

(1) For every $f \in C^0(\hat\Gamma)$, there holds $(\hat L_{\hat\Gamma} f)|_{\hat\Gamma} = f$.
(2) For every $f \in \mathbb{P}^p(\hat\Gamma)$, there holds $\hat L_{\hat\Gamma} f \in \mathbb{P}^p(\hat T)$.
(3) For every $f \in C^0(\hat\Gamma)$, there holds $(\hat L_{\hat\Gamma} f)(\hat N_d) = 0$.
(4) Let $k \in \{0, \dots, d-1\}$ and let $\hat\Sigma \subseteq \hat T$ be a k-simplex with $\hat\Sigma \subseteq \hat\Gamma$. Furthermore, consider the $(k+1)$-simplex $\hat\Sigma^+ := \operatorname{conv}(\hat\Sigma \cup \{\hat N_d\}) \subseteq \hat T$. Then, for every $f \in C^0(\hat\Gamma)$ with $f|_{\hat\Sigma} = 0$, there holds $(\hat L_{\hat\Gamma} f)|_{\hat\Sigma^+} = 0$.
(5) For all $f \in C^0(\hat\Gamma)$, there holds the stability bound

$$\|\hat L_{\hat\Gamma} f\|_{L^2(\hat T)} \lesssim p^{-1/2} \|f\|_{L^2(\hat\Gamma)}.$$

(In the case $d = 1$, we interpret $\|f\|_{L^2(\hat\Gamma)} = \|f\|_{l^2(\hat\Gamma)}$.)

Proof. First, since the vectors $\{\hat N_1 - \hat N_0, \dots, \hat N_d - \hat N_0\} \subseteq \mathbb{R}^d$ form a basis, we may pick a normal vector $n \in \mathbb{R}^d$ of $\hat\Gamma$ such that $\langle \hat N_1 - \hat N_0, n\rangle = \dots = \langle \hat N_{d-1} - \hat N_0, n\rangle = 0$ and $\langle \hat N_d - \hat N_0, n\rangle = 1$. Note that n can be used to write $\hat\Gamma$ in the normal form

$$\hat\Gamma = \{x \in \hat T \mid \langle x - \hat N_0, n\rangle = 0\}. \tag{4.2}$$

For every $x \in \hat T \backslash \{\hat N_d\}$, the line passing through $\hat N_d$ and x is given by $\{\hat N_d + s(x - \hat N_d) \mid s \in \mathbb{R}\}$. Using the normal form (4.2), the intersection point with $\hat\Gamma$ can easily be computed:

$$P(x) := \hat N_d + \langle \hat N_d - x, n\rangle^{-1}(x - \hat N_d) \in \hat\Gamma.$$

Let us verify that indeed $\langle \hat N_d - x, n\rangle \neq 0$: Since $x \in \hat T \backslash \{\hat N_d\}$, we know from Eq. (4.1) that it can be written in the form $x = \hat N_0 + \sum_{i=1}^d t_i(\hat N_i - \hat N_0)$, where $t_d \in [0,1)$. But then $\langle \hat N_d - x, n\rangle = \langle \hat N_d - \hat N_0, n\rangle - \sum_{i=1}^d t_i \langle \hat N_i - \hat N_0, n\rangle = 1 - t_d > 0$.

Now, for every $f \in C^0(\hat\Gamma)$, consider the lifting $\hat L_{\hat\Gamma} f$ defined as follows:

$$\forall x \in \hat T \backslash \{\hat N_d\}: \qquad \begin{array}{rcl} (\hat L_{\hat\Gamma} f)(x) & := & \langle \hat N_d - x, n\rangle^p f(P(x)), \\ (\hat L_{\hat\Gamma} f)(\hat N_d) & := & 0. \end{array}$$

Since f is bounded, it is easy to check that $\hat L_{\hat\Gamma} f \in C^0(\hat T)$. Furthermore, we have $(\hat L_{\hat\Gamma} f)|_{\hat\Gamma} = f$, which follows from the identities $\langle \hat N_d - x, n\rangle = 1$ and $P(x) = x$, for all $x \in \hat\Gamma$.

Next, consider the case $f \in \mathbb{P}^p(\hat\Gamma)$. Expanding $f(x) = \sum_{|\alpha| \le p} f_\alpha x^\alpha$, we can see that $\hat L_{\hat\Gamma} f$ is a polynomial as well:

$$\begin{aligned} \forall x \in \hat T: \qquad (\hat L_{\hat\Gamma} f)(x) &= \langle \hat N_d - x, n\rangle^p \sum_{|\alpha| \le p} f_\alpha (\hat N_d + \langle \hat N_d - x, n\rangle^{-1}(x - \hat N_d))^\alpha \\ &= \sum_{|\alpha| \le p} f_\alpha \langle \hat N_d - x, n\rangle^{p - |\alpha|} (\langle \hat N_d - x, n\rangle \hat N_d + x - \hat N_d)^\alpha \in \mathbb{P}^p(\hat T). \end{aligned}$$

Now, let $k \in \{0, \dots, d-1\}$ and consider a k-simplex $\hat\Sigma \subseteq \hat T$ with $\hat\Sigma \subseteq \hat\Gamma$. Let $\hat\Sigma^+ := \operatorname{conv}(\hat\Sigma \cup \{\hat N_d\})$ and consider a function $f \in C^0(\hat\Gamma)$ with $f|_{\hat\Sigma} = 0$. In order to prove the identity $(\hat L_{\hat\Gamma} f)|_{\hat\Sigma^+} = 0$, let $x \in \hat\Sigma^+$ be given. In the case $x = \hat N_d$, we immediately get $(\hat L_{\hat\Gamma} f)(x) = 0$ from the definition of $\hat L_{\hat\Gamma} f$. In the non-trivial case $x \neq \hat N_d$, we know that there exist $y \in \hat\Sigma$ and $t_d \in [0,1]$ such that $x = (1 - t_d)y + t_d \hat N_d$. Since $y \in \hat\Sigma \subseteq \hat\Gamma$, we have $P(x) = P(y) = y$, so that $(\hat L_{\hat\Gamma} f)(x) = \langle \hat N_d - x, n\rangle^p f(y) = 0$.

Finally, as for the stability bound, we only prove the non-trivial case $d \geq 2$. To this end, we consider the following parametrizations of $\hat\Gamma$ and $\hat T$:

$$\gamma : \begin{cases} \hat T^{d-1} & \longrightarrow & \hat\Gamma \\ t & \longmapsto & \hat N_0 + \sum_{i=1}^{d-1} t_i(\hat N_i - \hat N_0) \end{cases}, \qquad \tau : \begin{cases} \hat T^{d-1}\times[0,1] & \longrightarrow & \hat T \\ (t,t_d) & \longmapsto & (1-t_d)\gamma(t) + t_d\hat N_d \end{cases}.$$

There hold the identities

$$\sqrt{\det((\nabla\gamma)(t)^T(\nabla\gamma)(t))} = C_1, \qquad |\det((\nabla\tau)(t,t_d))| = (1-t_d)^{d-1}C_2,$$

where $C_1 := \sqrt{\det((\langle \hat N_i - \hat N_0, \hat N_j - \hat N_0\rangle)_{i,j=1}^{d-1})}$ and $C_2 := |\det((\hat N_1 - \hat N_0|\dots|\hat N_d - \hat N_0))|$. Exploiting the relations $\langle \hat N_d - \tau(t,t_d), n\rangle = 1 - t_d$ and $P(\tau(t,t_d)) = P(\gamma(t)) = \gamma(t)$, we compute, for all $f \in C^0(\hat\Gamma)$,

$$\begin{aligned}
\|\hat L_{\hat\Gamma} f\|_{L^2(\hat T)}^2 &= \int_{\hat T} \langle \hat N_d - x, n\rangle^{2p} f(P(x))^2 \,\mathrm{d}x \\
&= \int_0^1 \int_{\hat T^{d-1}} \langle \hat N_d - \tau(t,t_d), n\rangle^{2p} f(P(\tau(t,t_d)))^2 |\det((\nabla\tau)(t,t_d))| \,\mathrm{d}t\,\mathrm{d}t_d \\
&= \frac{C_2}{C_1}\left(\int_0^1 (1-t_d)^{2p+d-1}\,\mathrm{d}t_d\right)\left(\int_{\hat T^{d-1}} f(\gamma(t))^2 \sqrt{\det((\nabla\gamma)(t)^T(\nabla\gamma)(t))}\,\mathrm{d}t\right) \\
&= \frac{C_2}{C_1}(2p+d)^{-1}\|f\|_{L^2(\hat\Gamma)}^2.
\end{aligned}$$

This finishes the proof. □

Now that the lifting operators for all $(d-1)$-simplices $\hat\Gamma \subseteq \partial\hat T$ are available, we can combine them.

Lemma 4.4. *There exists a* lifting operator $\hat L : C^0(\partial\hat T) \longrightarrow C^0(\hat T)$ *with the following properties:*

(1) For every $f \in C^0(\partial\hat T)$, there holds $(\hat L f)|_{\partial\hat T} = f$.
(2) For every $f \in \mathbb{P}^p(\partial\hat T)$, there holds $\hat L f \in \mathbb{P}^p(\hat T)$.
(3) For all $f \in \mathbb{P}^p(\partial\hat T)$, there holds the stability estimate

$$\|\hat L f\|_{L^2(\hat T)} \lesssim p^{(d-2)/2}\|f\|_{L^2(\partial\hat T)}.$$

(In the case $d = 1$, we interpret $\|f\|_{L^2(\partial\hat T)} = \|f\|_{l^2(\partial\hat T)}$.)

Proof. For all $(d-1)$-simplices $\hat\Gamma \subseteq \hat T$, denote by $\hat L_{\hat\Gamma} : C^0(\hat\Gamma) \longrightarrow C^0(\hat T)$ the corresponding lifting operators from Lemma 4.3. In the following we write $\sum_{\hat\Gamma} \hat L_{\hat\Gamma} f$ to indicate that the operators corresponding to $(d-1)$-subsimplices $\hat\Gamma \subseteq \hat T$ are added. First, we define an auxiliary operator $\hat M : C^0(\partial\hat T) \longrightarrow C^0(\hat T)$: For every $f \in C^0(\partial\hat T)$, we set $\hat M f := \sum_{\hat\Gamma} \hat L_{\hat\Gamma} f \in C^0(\hat T)$, where $\hat L_{\hat\Gamma} f$ is meant as an abbreviation for $\hat L_{\hat\Gamma}(f|_{\hat\Gamma})$. Before we construct the alleged operator $\hat L$ from $\hat M$, let us first present the relevant properties of $\hat M$.

Clearly, if $f \in \mathbb{P}^p(\partial\hat T)$, then $\hat M f \in \mathbb{P}^p(\hat T)$ by item (2) of Lemma 4.3.

Let $f \in C^0(\partial\hat T)$ and $\hat N \in \mathcal{N}(\hat T)$ be given. For each $(d-1)$-simplex $\hat\Gamma \subseteq \hat T$, we distinguish between two cases: First, if $\hat N \in \hat\Gamma$, then $(\hat L_{\hat\Gamma} f)(\hat N) = f(\hat N)$ by item (1) of Lemma 4.3. Second, if $\hat N \notin \hat\Gamma$, then item (3) of Lemma 4.3 immediately tells us that $(\hat L_{\hat\Gamma} f)(\hat N) = 0$. Since the total number of $(d-1)$-simplices $\hat\Gamma \subseteq \hat T$ is given by $d+1$, and since only one of them falls into the second category, we end up with the following identity:

$$(\hat M f)(\hat N) = \sum_{\hat\Gamma:\,\hat N\in\hat\Gamma} (\hat L_{\hat\Gamma} f)(\hat N) + \sum_{\hat\Gamma:\,\hat N\notin\hat\Gamma} (\hat L_{\hat\Gamma} f)(\hat N) = d\cdot f(\hat N). \tag{4.3}$$

Next, let $k \in \{0,\dots,d-1\}$ and let $f \in C^0(\partial\hat T)$ be such that $f|_{\hat\Sigma} = 0$ for all k-simplices $\hat\Sigma \subseteq \hat T$. Furthermore, let $\hat\Sigma^+ \subseteq \hat T$ be an arbitrary $(k+1)$-simplex. Considering a $(d-1)$-simplex $\hat\Gamma \subseteq \hat T$, we distinguish between two cases again: First, if $\hat\Sigma^+ \subseteq \hat\Gamma$, then $(\hat L_{\hat\Gamma} f)|_{\hat\Sigma^+} = f|_{\hat\Sigma^+}$ by item (1) of Lemma 4.3. Second, if $\hat\Sigma^+ \not\subseteq \hat\Gamma$, then there must hold $\hat\Sigma^+ = \mathrm{conv}(\hat\Sigma \cup \{\hat N\})$, where $\hat\Sigma \subseteq \hat T$ is a k-simplex with $\hat\Sigma \subseteq \hat\Gamma$ and where $\hat N \in \mathcal{N}(\hat T)\backslash\hat\Gamma$. Since $f|_{\hat\Sigma} = 0$ by assumption, we obtain from item (4) of Lemma 4.3 that there must hold $(\hat L_{\hat\Gamma} f)|_{\hat\Sigma^+} = 0$. We mention that the first case occurs $d-k-1$ times, since $\hat\Sigma^+$ occupies $k+2$ nodes, so that the remaining node in $\mathcal{N}(\hat T)\backslash\hat\Gamma$ must be one of the $(d+1)-(k+2) = d-k-1$ unoccupied nodes of $\hat T$. Altogether, it follows that

$$(\hat M f)|_{\hat\Sigma^+} = \sum_{\hat\Gamma:\,\hat\Sigma^+\subseteq\hat\Gamma} (\hat L_{\hat\Gamma} f)|_{\hat\Sigma^+} + \sum_{\hat\Gamma:\,\hat\Sigma^+\not\subseteq\hat\Gamma} (\hat L_{\hat\Gamma} f)|_{\hat\Sigma^+} = (d-k-1)f|_{\hat\Sigma^+}. \tag{4.4}$$

Finally, for all $f \in C^0(\partial\hat T)$, we may use item (5) of Lemma 4.3 to derive a stability bound for the operator $\hat M$:

$$\|\hat M f\|_{L^2(\hat T)} \leq \sum_{\hat\Gamma} \|\hat L_{\hat\Gamma} f\|_{L^2(\hat T)} \lesssim p^{-1/2}\|f\|_{L^2(\partial\hat T)}. \tag{4.5}$$

Our presentation of the auxiliary operator $\hat M$ is now finished and we proceed to construct $\hat L$ from $\hat M$. Let $f \in C^0(\partial\hat T)$ be given. We use the restriction operator $\hat R := (\cdot)|_{\partial\hat T}$ and the coefficients $c_k := (d-k)^{-1}$, $k \in \{0,\dots,d-1\}$, to define an auxiliary function:

$$\tilde f := (\mathrm{id} - c_{d-1}\hat R\hat M)\cdots(\mathrm{id} - c_1\hat R\hat M)(\mathrm{id} - c_0\hat R\hat M)f \in C^0(\partial\hat T).$$

From Eq. (4.3) we know that the function $(\mathrm{id} - c_0 \hat{R}\hat{M}) f \in C^0(\partial\hat{T})$ vanishes on all 0-simplices of $\hat{T}$. Then, using Eq. (4.4), we may conclude that $(\mathrm{id} - c_1 \hat{R}\hat{M})(\mathrm{id} - c_0 \hat{R}\hat{M}) f \in C^0(\partial\hat{T})$ vanishes on all 1-simplices of $\hat{T}$. Proceeding forwards with Eq. (4.4), we find that $\tilde{f}$ must vanishes on all $(d-1)$-simplices of $\hat{T}$. However, since the $(d-1)$-simplices make up all of $\partial\hat{T}$, we have $\tilde{f} = 0 \in C^0(\partial\hat{T})$. Expanding $\tilde{f}$, we find that, for certain coefficients $\tilde{c}_k \in \mathbb{R}$,

$$0 = \tilde{f} = f - \sum_{k=1}^{d} \tilde{c}_k (\hat{R}\hat{M})^k f = f - \hat{R}\hat{M} \sum_{k=0}^{d-1} \tilde{c}_{k+1} (\hat{R}\hat{M})^k f.$$

Now, define

$$\hat{L} f := \hat{M} \sum_{k=0}^{d-1} \tilde{c}_{k+1} (\hat{R}\hat{M})^k f \in C^0(\hat{T}).$$

Clearly, $(\hat{L} f)|_{\partial\hat{T}} = \hat{R}\hat{L} f = f$. Furthermore, since $\hat{R}$ and $\hat{M}$ preserve polynomials, so does $\hat{L}$. Finally, let us derive a bound for $\hat{L} f$ in the case of a polynomial input, $f \in \mathbb{P}^p(\partial\hat{T})$. Since the powers $(\hat{R}\hat{M})^k f$ are polynomials as well, it pays off to have a look at $\hat{R}\hat{M} g$, where $g \in \mathbb{P}^p(\partial\hat{T})$. Using a multiplicative trace inequality, [12, Thm. 1.6.6], and an inverse inequality, [21, Cor. 4.2], we compute

$$\|\hat{R}\hat{M} g\|_{L^2(\partial\hat{T})} = \|\hat{M} g\|_{L^2(\partial\hat{T})} \lesssim \|\hat{M} g\|_{L^2(\hat{T})}^{1/2} \|\hat{M} g\|_{H^1(\hat{T})}^{1/2} \lesssim p \|\hat{M} g\|_{L^2(\hat{T})} \overset{\text{Eq. (4.5)}}{\lesssim} p^{1/2} \|g\|_{L^2(\partial\hat{T})}.$$

We conclude the proof with the bound for $\hat{L} f$:

$$\|\hat{L} f\|_{L^2(\hat{T})} \overset{\text{Eq. (4.5)}}{\lesssim} p^{-1/2} \sum_{k=0}^{d-1} \|(\hat{R}\hat{M})^k f\|_{L^2(\partial\hat{T})} \lesssim p^{-1/2} \sum_{k=0}^{d-1} p^{k/2} \|f\|_{L^2(\partial\hat{T})} \lesssim p^{(d-2)/2} \|f\|_{L^2(\partial\hat{T})}. \quad \square$$

The next lemma generalizes the results from [38] to arbitrary space dimensions $d \geq 1$. The approach taken here is slightly different from [38], since we define the operator $\hat{J}^p$ by induction on d. Furthermore, as was pointed out at the beginning of this section, it suffices to consider polynomial inputs $f \in \mathbb{P}^{p+2}(\hat{T})$.

Lemma 4.5. *There exists a linear operator $\hat{J}^p : \mathbb{P}^{p+2}(\hat{T}) \longrightarrow \mathbb{P}^p(\hat{T})$ with the following properties:*

(1) For all $k \in \{0, \ldots, d\}$, all k-simplices $\hat{\Sigma} \subseteq \hat{T}$ and all $f \in \mathbb{P}^{p+2}(\hat{T})$, the quantity $(\hat{J}^p f)|_{\hat{\Sigma}}$ is uniquely determined by $f|_{\hat{\Sigma}}$.
(2) $\hat{J}^p$ is a projection, i.e., $\hat{J}^p f = f$ for all $f \in \mathbb{P}^p(\hat{T})$.
(3) For all $f \in \mathbb{P}^{p+2}(\hat{T})$, there hold the following stability and error bounds:

$$\|\hat{J}^p f\|_{L^2(\hat{T})} \lesssim p^{d(d+1)/4} \|f\|_{L^2(\hat{T})},$$
$$\|f - \hat{J}^p f\|_{L^2(\hat{T})} \lesssim p^{d(d+1)/4} \inf_{g \in \mathbb{P}^p(\hat{T})} \|f - g\|_{L^2(\hat{T})}.$$

Proof. We construct the operator $\hat{J}^p$ via induction on the space dimension $d \geq 1$ and write $\hat{J}_1^p, \hat{J}_2^p, \ldots, \hat{J}_d^p$ for the corresponding operators. As part of the induction argument, we prove item (1), item (2) and the stability bound from item (3). Finally, the error bound is *not* part of the induction, since it follows readily from the projection property and the stability bound.

The case $d = 1$: Denote by $\hat{L} : C^0(\partial\hat{T}) \longrightarrow C^0(\hat{T})$ the polynomial preserving lifting operator from Lemma 4.4. Note that, since $d = 1$, we have $\hat{L} f \in \mathbb{P}^p(\hat{T})$ for *all* $f \in C^0(\partial\hat{T})$. Let $\mathbb{P}_0^p(\hat{T}) := \{f \in \mathbb{P}^p(\hat{T}) \mid f|_{\partial\hat{T}} = 0\}$ and denote by $\hat{P} : L^2(\hat{T}) \longrightarrow \mathbb{P}_0^p(\hat{T})$ the orthogonal projection. We define

$$\forall f \in \mathbb{P}^{p+2}(\hat{T}) : \qquad \hat{J}_1^p f := \hat{L}(f|_{\partial\hat{T}}) + \hat{P}(f - \hat{L}(f|_{\partial\hat{T}})) \in \mathbb{P}^p(\hat{T}).$$

The identity $(\hat{J}_1^p f)|_{\partial\hat{T}} = f|_{\partial\hat{T}}$, for all $f \in \mathbb{P}^{p+2}(\hat{T})$, proves item (1). Since $\hat{P}$ is a projection, so is $\hat{J}_1^p$. Using a multiplicative trace inequality, [12, Thm. 1.6.6], and an inverse inequality, [21, Cor. 4.2], we obtain, for all $f \in \mathbb{P}^{p+2}(\hat{T})$, the stability bound

$$\begin{aligned}\|\hat{J}_1^p f\|_{L^2(\hat{T})} &\leq \|\hat{L}(f|_{\partial\hat{T}})\|_{L^2(\hat{T})} + \|\hat{P}(f - \hat{L}(f|_{\partial\hat{T}}))\|_{L^2(\hat{T})} \lesssim \|\hat{L}(f|_{\partial\hat{T}})\|_{L^2(\hat{T})} + \|f\|_{L^2(\hat{T})} \\ &\overset{\text{Lemma 4.4}}{\lesssim} p^{-1/2} \|f\|_{l^2(\partial\hat{T})} + \|f\|_{L^2(\hat{T})} \lesssim p^{-1/2} \|f\|_{L^2(\hat{T})}^{1/2} \|f\|_{H^1(\hat{T})}^{1/2} + \|f\|_{L^2(\hat{T})} \lesssim p^{1/2} \|f\|_{L^2(\hat{T})}.\end{aligned}$$

The step $d - 1 \mapsto d$: Assume that an operator $\hat{J}_{d-1}^p : \mathbb{P}^{p+2}(\hat{T}^{d-1}) \longrightarrow \mathbb{P}^p(\hat{T}^{d-1})$ satisfying items (1), (2) and the stability bound from (3) is well-defined. Furthermore, let us denote the $(d-1)$-subsimplices of $\hat{T}$ by $\hat{\Gamma}_0, \ldots, \hat{\Gamma}_d$ and fix affine parametrizations $\gamma_i : \hat{T}^{d-1} \longrightarrow \hat{\Gamma}_i$. In order to construct the operator $\hat{J}_d^p : \mathbb{P}^{p+2}(\hat{T}^d) \longrightarrow \mathbb{P}^p(\hat{T}^d)$ from $\hat{J}_{d-1}^p$, we proceed roughly as follows: Given $f \in \mathbb{P}^{p+2}(\hat{T})$, we can use $\hat{J}_{d-1}^p$ to find a polynomial $g \in \mathbb{P}^p(\partial\hat{T})$ with $g \approx f|_{\partial\hat{T}}$. Then, using the polynomial preserving lifting operator $\hat{L} : C^0(\partial\hat{T}) \longrightarrow C^0(\hat{T})$ from Lemma 4.4, we introduce the quantity $G := \hat{L} g \in \mathbb{P}^p(\hat{T})$. Clearly, $G|_{\partial\hat{T}} \approx f|_{\partial\hat{T}}$, but *not* necessarily $G \approx f$ in all of $\hat{T}$. However, if $\hat{P} : L^2(\hat{T}) \longrightarrow \mathbb{P}_0^p(\hat{T})$ denotes the orthogonal projection onto the space of homogeneous polynomials $\mathbb{P}_0^p(\hat{T})$, then indeed $\hat{J}_d^p f := G + \hat{P}(f - G) \approx G + (f - G) = f$ on $\hat{T}$.

In order to work out the details, let $f \in \mathbb{P}^{p+2}(\hat{T})$ be given. Then, for each $i \in \{0, \ldots, d\}$, $f \circ \gamma_i \in \mathbb{P}^{p+2}(\hat{T}^{d-1})$, so that the polynomial $\hat{J}_{d-1}^p(f \circ \gamma_i) \in \mathbb{P}^p(\hat{T}^{d-1})$ is well-defined by the induction hypothesis. We define boundary data $g : \partial\hat{T} \longrightarrow \mathbb{R}$ in a piecewise manner:

$$\forall i \in \{0, \ldots, d\} : \qquad g|_{\hat{\Gamma}_i} := \hat{J}_{d-1}^p(f \circ \gamma_i) \circ \gamma_i^{-1} \in \mathbb{P}^p(\hat{\Gamma}_i).$$

(Note that γ_i is injective and thus invertible on its range, $\hat{\Gamma}_i$.)

We argue that $g \in \mathbb{P}^p(\partial\hat{T})$: Consider the boundary $\hat{\Sigma}_{ij} := \hat{\Gamma}_i \cap \hat{\Gamma}_j \subseteq \hat{T}$ between any two $(d-1)$-simplices $\hat{\Gamma}_i, \hat{\Gamma}_j$ and note that $\hat{\Sigma}_{ij}$ is a $(d-2)$-simplex. Then, the pre-images $\gamma_i^{-1}(\hat{\Sigma}_{ij}), \gamma_j^{-1}(\hat{\Sigma}_{ij}) \subseteq \hat{T}^{d-1}$ are $(d-2)$-simplices as well. Using item (1) of the induction hypothesis, we know that

$(\hat{J}^p_{d-1}(f\circ\gamma_i))|_{\gamma_i^{-1}(\hat{\Sigma}_{ij})}$ is uniquely determined by $(f\circ\gamma_i)|_{\gamma_i^{-1}(\hat{\Sigma}_{ij})}$ and that $(\hat{J}^p_{d-1}(f\circ\gamma_j))|_{\gamma_j^{-1}(\hat{\Sigma}_{ij})}$ is uniquely determined by $(f\circ\gamma_j)|_{\gamma_j^{-1}(\hat{\Sigma}_{ij})}$. However, since $(f\circ\gamma_i)|_{\gamma_i^{-1}(\hat{\Sigma}_{ij})} = f|_{\hat{\Sigma}_{ij}} = (f\circ\gamma_j)|_{\gamma_j^{-1}(\hat{\Sigma}_{ij})}$, there must hold $(\hat{J}^p_{d-1}(f\circ\gamma_i))|_{\gamma_i^{-1}(\hat{\Sigma}_{ij})} = (\hat{J}^p_{d-1}(f\circ\gamma_j))|_{\gamma_j^{-1}(\hat{\Sigma}_{ij})}$. It follows that $g|_{\hat{\Gamma}_i}|_{\hat{\Sigma}_{ij}} = g|_{\hat{\Gamma}_j}|_{\hat{\Sigma}_{ij}}$, i.e., that $g\in C^0(\partial\hat{T})$. According to Definition 4.2, it follows that $g\in\mathbb{P}^p(\partial\hat{T})$.

Furthermore, to get a stability estimate for g, we can use item (3) of the induction hypothesis:

$$\|g\|_{L^2(\partial\hat{T})} \lesssim \sum_{i=0}^{d} \|\hat{J}^p_{d-1}(f\circ\gamma_i)\circ\gamma_i^{-1}\|_{L^2(\hat{\Gamma}_i)} \lesssim \sum_{i=0}^{d} \|\hat{J}^p_{d-1}(f\circ\gamma_i)\|_{L^2(\hat{T}^{d-1})} \overset{(3)}{\lesssim} p^{(d-1)d/4}\sum_{i=0}^{d}\|f\circ\gamma_i\|_{L^2(\hat{T}^{d-1})} \lesssim p^{(d-1)d/4}\|f\|_{L^2(\partial\hat{T})}.$$

We proceed as stated above and lift g from $\partial\hat{T}$ into $\hat{T}$. From Lemma 4.4, we know that the function $G := \hat{L}g$ satisfies

$$G|_{\partial\hat{T}} = g, \qquad G\in\mathbb{P}^p(\hat{T}), \qquad \|G\|_{L^2(\hat{T})} \lesssim p^{(d-2)/2}\|g\|_{L^2(\partial\hat{T})}.$$

Now, recalling that $\hat{P}: L^2(\hat{T}) \longrightarrow \mathbb{P}^p_0(\hat{T})$ denotes the orthogonal projection, consider the function

$$\hat{J}^p_d f := G + \hat{P}(f-G) \in \mathbb{P}^p(\hat{T}).$$

Clearly, the mapping $f\mapsto \hat{J}^p_d f$ defines a linear operator $\hat{J}^p_d : \mathbb{P}^{p+2}(\hat{T}) \longrightarrow \mathbb{P}^p(\hat{T})$. To prove item (1), let $k\in\{0,\dots,d\}$ and consider a k-simplex $\hat{\Sigma}\subseteq\hat{T}$. If $k=d$, the statement becomes trivial. If $k\le d-1$, then there exists a $(d-1)$-simplex $\hat{\Gamma}_i\subseteq\hat{T}$ such that $\hat{\Sigma}\subseteq\hat{\Gamma}_i\subseteq\partial\hat{T}$. Since $\hat{P}(f-G)$ vanishes on $\partial\hat{T}$, we find that

$$(\hat{J}^p_d f)|_{\hat{\Sigma}} = G|_{\hat{\Sigma}} = g|_{\hat{\Sigma}} = \hat{J}^p_{d-1}(f\circ\gamma_i)\circ(\gamma_i^{-1}|_{\hat{\Sigma}}).$$

Item (1) of the induction hypothesis tells us that this function is uniquely determined by $(f\circ\gamma_i)|_{\gamma_i^{-1}(\hat{\Sigma})}$, i.e., by $f|_{\hat{\Sigma}}$.

As for the projection property of $\hat{J}^p_d$, consider an input $f\in\mathbb{P}^p(\hat{T})$. Then, $g = f|_{\partial\hat{T}}\in\mathbb{P}^p(\partial\hat{T})$, since $\hat{J}^p_{d-1}$ is a projection by the induction hypothesis. It follows that $f-G\in\mathbb{P}^p_0(\hat{T})$ so that $\hat{J}^p_d f = G + \hat{P}(f-G) = G + (f-G) = f$.

Finally, for all $f\in\mathbb{P}^{p+2}(\hat{T})$, a multiplicative trace inequality, [12], and an inverse inequality, [21, Cor. 4.2], give us the desired stability estimate:

$$\begin{aligned}
\|\hat{J}^p_d f\|_{L^2(\hat{T})} &\le \|G\|_{L^2(\hat{T})} + \|\hat{P}(f-G)\|_{L^2(\hat{T})} &&\le \|G\|_{L^2(\hat{T})} + \|f-G\|_{L^2(\hat{T})}\\
&\lesssim \|f\|_{L^2(\hat{T})} + \|G\|_{L^2(\hat{T})} &&\lesssim \|f\|_{L^2(\hat{T})} + p^{(d-2)/2}\|g\|_{L^2(\partial\hat{T})}\\
&\lesssim \|f\|_{L^2(\hat{T})} + p^{(d-2)/2+(d-1)d/4}\|f\|_{L^2(\partial\hat{T})} &&\lesssim \|f\|_{L^2(\hat{T})} + p^{(d-2)/2+(d-1)d/4}\|f\|^{1/2}_{L^2(\hat{T})}\|f\|^{1/2}_{H^1(\hat{T})}\\
&\lesssim p^{(d-2)/2+(d-1)d/4+1}\|f\|_{L^2(\hat{T})} &&= p^{d(d+1)/4}\|f\|_{L^2(\hat{T})}.
\end{aligned}$$

This finishes the proof. $\square$

We close this section with the delayed proof of Lemma 3.9.

Proof of Lemma 3.9. Denote by $\hat{J}^p : \mathbb{P}^{p+2}(\hat{T}) \longrightarrow \mathbb{P}^p(\hat{T})$ the operator from Lemma 4.5. We define the asserted operator $J^p_{\mathcal{T}} : \mathbb{S}^{p+2,0}(\mathcal{T}) \longrightarrow \mathbb{S}^{p,0}(\mathcal{T})$ in an elementwise fashion: For every $v\in\mathbb{S}^{p+2,0}(\mathcal{T})$ and every element $T\in\mathcal{T}$, we set

$$(J^p_{\mathcal{T}}v)|_T := \hat{J}^p(v\circ F_T)\circ F_T^{-1}.$$

(Recall from Definition 2.2 that $F_T : \hat{T}\longrightarrow T$ is the affine transformation between $\hat{T}$ and T.)

The preservation of continuity and boundary values follows from item (1) in Lemma 4.5. The preservation of supports is obvious from the elementwise definition. Finally, to see the error bound, let $\kappa\in\mathbb{P}^1(\hat{T})$ and $u\in\mathbb{P}^p(\hat{T})$. Then, using an inverse inequality once again, we obtain

$$\begin{aligned}
\|(\mathrm{id}-\hat{J}^p)(\kappa^2 u)\|_{H^1(\hat{T})} &\lesssim p^2\|(\mathrm{id}-\hat{J}^p)(\kappa^2 u)\|_{L^2(\hat{T})} &&\overset{\text{Lemma 4.5}}{\lesssim} p^{d(d+1)/4+2}\inf_{g\in\mathbb{P}^p(\hat{T})}\|\kappa^2 u - g\|_{L^2(\hat{T})}\\
&\le p^{d(d+1)/4+2}\|\kappa^2 u - \kappa(0)^2 u\|_{L^2(\hat{T})} &&\overset{\text{Taylor}}{\lesssim} p^{d(d+1)/4+2}|\kappa^2|_{W^{1,\infty}(\hat{T})}\|u\|_{L^2(\hat{T})}.
\end{aligned}$$

Item (3) of Lemma 3.9 then follows with the standard scaling relation $h_T^l|v|_{W^{l,q}(T)} \approx h_T^{d/q}|v\circ F_T|_{W^{l,q}(\hat{T})}$. In fact, if $\hat{\kappa} := \kappa\circ F_T$ and $\hat{u} := u\circ F_T$ denote the pull-backs of κ and u, then

$$\sum_{l=0}^{1} h_T^l|(\mathrm{id}-J^p_{\mathcal{T}})(\kappa^2 u)|_{H^l(T)} \lesssim h_T^{d/2}\|(\mathrm{id}-\hat{J}^p)(\hat{\kappa}^2\hat{u})\|_{H^1(\hat{T})} \lesssim p^{d(d+1)/4+2}h_T^{d/2}|\hat{\kappa}^2|_{W^{1,\infty}(\hat{T})}\|\hat{u}\|_{L^2(\hat{T})} \approx p^{d(d+1)/4+2}h_T|\kappa^2|_{W^{1,\infty}(T)}\|u\|_{L^2(T)}.$$

This concludes the proof of Lemma 3.9. $\square$

5. Numerical results

In this final section, we illustrate the validity of Theorem 2.13 with two numerical examples in $d=2$ space dimensions. On the unit square $\Omega := (0,1)\times(0,1)\subseteq\mathbb{R}^2$, we solve the Poisson equation with homogeneous Dirichlet boundary conditions

$$-\Delta u = 1 \qquad \text{in } \Omega, \qquad u = 0 \qquad \text{on } \partial\Omega.$$

For comparisons, the domain Ω is triangulated with meshes $\mathcal{T}$ with algebraic grading ($\alpha=4$) and exponential grading ($\alpha=\infty$) towards the left edge $\Gamma := \{0\}\times[0,1]$ (cf. Section 2.5). Each element $T\in\mathcal{T}$ satisfies $h_T \approx \mathrm{dist}_2(x_T,\Gamma)^{1-1/\alpha}H$, where $H=0.5$. Moreover, we also consider the mesh $\mathcal{T}$ graded exponentially towards the origin as depicted in Fig. 1 (right).

We mention that, for the case of quasi-uniform meshes, numerical results can e.g. be found in [7,24], while our previous work [1] also includes numerical examples on algebraically graded meshes.

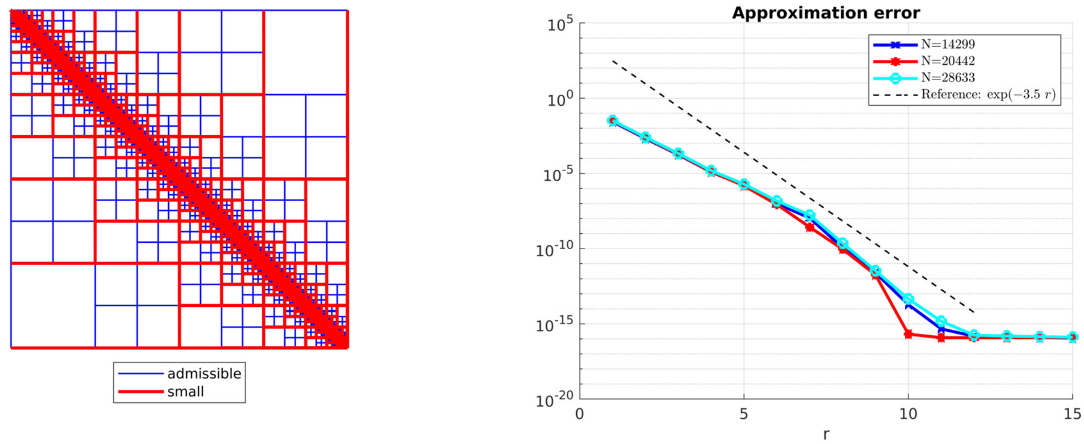

Fig. 3. Example on mesh graded exponentially towards edge $x = 0$. Left: The block partition $\mathbb{P}$. Right: Empirical approximation errors.

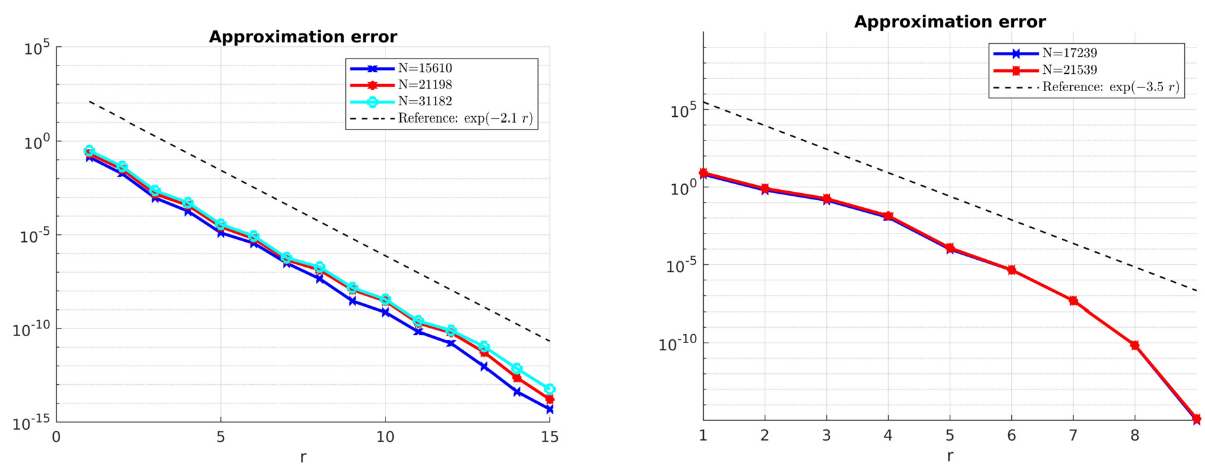

Fig. 4. Left: Approximation errors on mesh graded algebraically towards edge $x = 0$. Right: Approximation errors on mesh graded exponentially towards the origin.

We start with the special case $p = 1$. The system matrix $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ is assembled in MATLAB and explicitly inverted using MATLAB's built-in inversion routine `inv(...)`. A block decomposition is computed using the *geometrically balanced clustering* algorithm from [34] with choices $\sigma_{\mathrm{adm}} = 1$ and $\sigma_{\mathrm{small}} = 10$. Then, for each rank bound $r \in \{1, \ldots, 15\}$, an approximation $\boldsymbol{B}_r \in \mathcal{H}(\mathbb{P}, r)$ to $\boldsymbol{A}^{-1}$ is computed via blockwise truncated singular values decompositions. Denoting by σ_{r+1} the $(r+1)$-largest singular value of $\boldsymbol{A}^{-1}$, this procedure gives rise to the *computable* error bound

$$\|\boldsymbol{A}^{-1} - \boldsymbol{B}_r\|_2 \lesssim \mathrm{depth}(\mathbb{T}_{N\times N}) \cdot \max_{(I,J)\in\mathbb{P}} \sigma_{r+1}(\boldsymbol{A}^{-1}|_{I\times J}).$$

The right-hand image in Fig. 3 depicts a comparison between three different problem sizes of roughly $N \approx 14.300$, $N \approx 20.400$ and $N \approx 28.600$ degrees of freedom. The error appears to decay as $O(\exp(-3.5r))$, which is even better than our theoretical prediction $O(\exp(-\sigma_{\exp} r^{1/3}))$ from Theorem 2.13, which is also observed in previous works on approximation of inverses, [7,24,1], and indicates that the bounds in Theorem 2.13 are not sharp.

Fig. 4 shows the same exponential convergence behavior for algebraically graded meshes (with grading exponent $\alpha = 4$) for different problem sizes, which is also covered by our theory, as well as for a mesh graded exponentially towards the origin (which is not covered by our main result). While this may indicate that our main result might also be valid for more general meshes, we want to point out that employing the standard clustering algorithm to compute the block partition might not work as intended in this case. As degrees of freedom are strongly concentrated around the origin, the geometrically balanced clustering algorithm may produce almost exclusively non-admissible blocks. In order to derive a meaningful approximation, we therefore changed the admissibility parameter to $\sigma_{\mathrm{adm}} = 5$ in the right plot of Fig. 4. Nonetheless, only less than 10% of blocks were admissible.

Fig. 5 shows the scaling of the storage memory requirement of the matrix $\boldsymbol{B}_r$, where $r = 5$ is fixed, with increasing matrix size N for the mesh graded exponentially towards the edge $x = 0$ (left picture) and to the origin (right picture). As expected, the case of exponential grading towards the edge, which is covered by our theory, produces a nice complexity of $\mathcal{O}(N \log N)$, whereas the issue in the clustering algorithm (which actually produces trees of depth of $\mathcal{O}(N)$) for the case of exponential grading to origin leads to a bad scaling of $\mathcal{O}(N^2)$.

Finally, with the previously defined mesh that is graded exponentially towards the edge $x = 0$, we also compute an example with higher polynomial degrees on each element. We employ a combination of the finite element code `NGSolve` (which is capable of higher order polynomials), [39], and the C++ $\mathcal{H}$-matrix library, [11]. Hereby, both codes are coupled using a code also employed in [23]. We use polynomial degrees $p = 5$ (which leads to a problem size of $N = 5791$), $p = 6$ (which leads to a problem size of $N = 17053$) and $p = 7$ (which leads to a problem size of $N = 46915$). The $\mathcal{H}$-matrix approximations are computed using $\mathcal{H}$-Cholesky decompositions and then inverting the Cholesky factors using $\mathcal{H}$-matrix

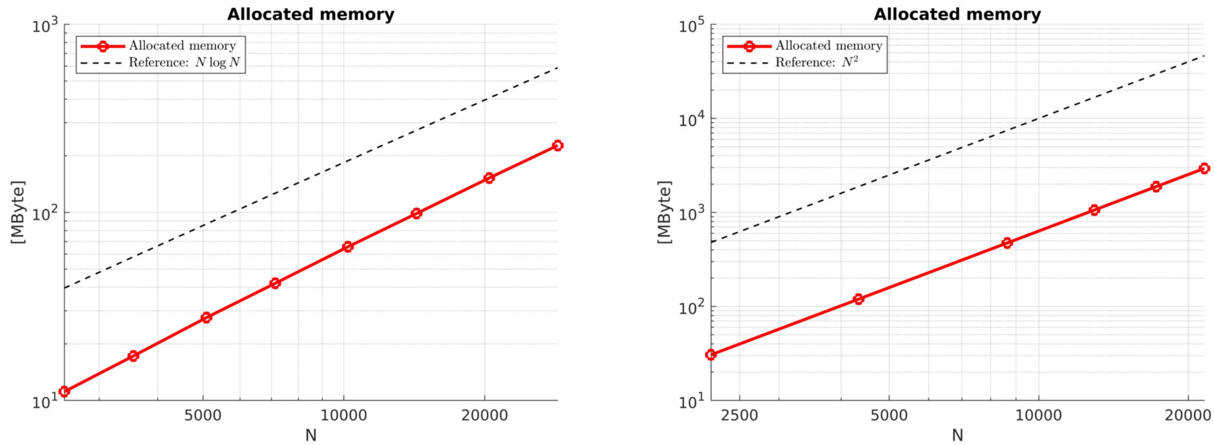

Fig. 5. Memory requirements to store the $\mathcal{H}$-matrix approximations. Left: Mesh graded exponentially towards edge $x = 0$. Right: Mesh graded exponentially towards origin.

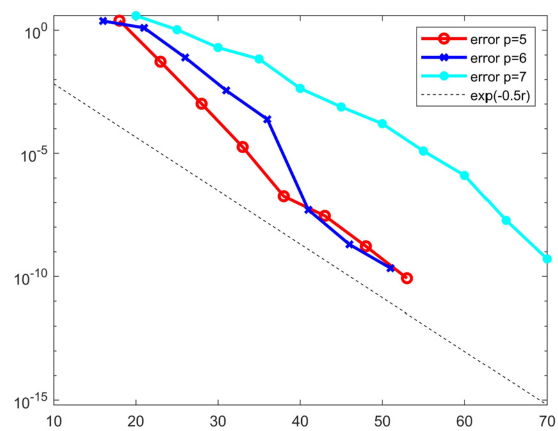

Fig. 6. Exponential convergence of $\mathcal{H}$-matrix approximations for $p = 5, 6, 7$.

arithmetic. In order to avoid computing the full inverse matrix, we compute the error measure $\|\boldsymbol{I} - (\boldsymbol{C}_{\mathcal{H}}\boldsymbol{C}_{\mathcal{H}}^T)^{-1}\boldsymbol{A}\|_2$, which is an upper bound for the relative error.

Fig. 6 shows exponential convergence of the error measure as predicted by our main result.

Data availability

Data will be made available on request.

References

[1] N. Angleitner, M. Faustmann, J.M. Melenk, Approximating inverse FEM matrices on non-uniform meshes with $\mathcal{H}$-matrices, Calcolo 58 (3) (2021) 31, 36.
[2] N. Angleitner, M. Faustmann, J.M. Melenk, $\mathcal{H}$-inverses for RBF interpolation, arXiv e-prints, arXiv:2109.05763, 2021.
[3] I. Babuška, A. Craig, J. Mandel, J. Pitkäranta, Efficient preconditioning for the p version finite element method in two dimensions, SIAM J. Numer. Anal. 28 (3) (1991) 624–661.
[4] C. Bernardi, M. Dauge, Y. Maday, Trace liftings which preserve polynomials, C. R. Acad. Sci. Paris, Ser. I 315 (1992) 333–338.
[5] C. Bernardi, M. Dauge, Y. Maday, Polynomials in the Sobolev world (version 2), Tech. Report 14, IRMAR, 2007, https://hal.archives-ouvertes.fr/hal-00153795.
[6] M. Bebendorf, Hierarchical Matrices, Lecture Notes in Computational Science and Engineering, vol. 63, Springer, Berlin, 2008.
[7] M. Bebendorf, W. Hackbusch, Existence of $\mathcal{H}$-matrix approximants to the inverse FE-matrix of elliptic operators with L^∞-coefficients, Numer. Math. 95 (1) (2003) 1–28.
[8] C. Bernardi, Y. Maday, Spectral methods, in: P.G. Ciarlet, J.L. Lions (Eds.), Handbook of Numerical Analysis, vol. 5, North Holland, Amsterdam, 1997.
[9] M. Bebendorf, J. Ostrowski, Parallel hierarchical matrix preconditioners for the curl-curl operator, J. Comput. Math. (2009) 624–641.
[10] S. Börm, Efficient Numerical Methods for Non-local Operators, EMS Tracts in Mathematics, vol. 14, European Mathematical Society (EMS), Zürich, 2010.
[11] S. Börm, $\mathcal{H}$2LIB software library, University of Kiel, http://www.h2lib.org, 2021.
[12] S.C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, Texts in Applied Mathematics, vol. 15, Springer-Verlag, New York, 2002.
[13] I. Babuška, B.A. Szabo, I.N. Katz, The p-version of the finite element method, SIAM J. Numer. Anal. 18 (3) (1981) 515–545.
[14] W. Cao, L. Demkowicz, Optimal error estimate of a projection based interpolation for the p-version approximation in three dimensions, Comput. Math. Appl. 50 (3–4) (2005) 359–366.
[15] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, Studies in Mathematics and Its Applications, vol. 4, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978.
[16] Ph. Clément, Approximation by finite element functions using local regularization, Rev. Fr. Autom. Inform. Rech. Opér., Anal. Numér. 9 (R-2) (1975) 77–84.
[17] L. Demkowicz, I. Babuška, p interpolation error estimates for edge finite elements of variable order in two dimensions, SIAM J. Numer. Anal. 41 (4) (2003) 1195–1208.
[18] L. Demkowicz, A. Buffa, H^1, $H(\mathrm{curl})$ and $H(\mathrm{div})$-conforming projection-based interpolation in three dimensions. Quasi-optimal p-interpolation estimates, Comput. Methods Appl. Mech. Eng. 194 (2–5) (2005) 267–296.
[19] L. Demkowicz, Polynomial exact sequences and projection-based interpolation with applications to Maxwell's equations, in: D. Boffi, F. Brezzi, L. Demkowicz, L.F. Durán, R. Falk, M. Fortin (Eds.), Mixed Finite Elements, Compatibility Conditions, and Applications, in: Lecture Notes in Mathematics, vol. 1939, Springer Verlag, 2008.

[20] J. Dölz, H. Harbrecht, Ch. Schwab, Covariance regularity and $\mathcal{H}$-matrix approximation for rough random fields, Numer. Math. 135 (4) (2017) 1045–1071.
[21] Z. Ditzian, Multivariate Bernstein and Markov inequalities, J. Approx. Theory 70 (3) (1992) 273–283.
[22] H. Edelsbrunner, D.R. Grayson, Edgewise subdivision of a simplex, in: ACM Symposium on Computational Geometry, Miami, FL, 1999, vol. 24, 2000, pp. 707–719.
[23] C. Erath, L. Mascotto, J.M. Melenk, I. Perugia, A. Rieder, Mortar coupling of hp-discontinuous Galerkin and boundary element methods for the Helmholtz equation, J. Sci. Comput. 92 (1) (2022) 2, 41.
[24] M. Faustmann, J.M. Melenk, D. Praetorius, H-matrix approximability of the inverses of FEM matrices, Numer. Math. 131 (4) (2015) 615–642.
[25] M. Faustmann, J.M. Melenk, D. Praetorius, Existence of $\mathcal{H}$-matrix approximants to the inverse of BEM matrices: the simple-layer operator, Math. Comput. 85 (2016) 119–152.
[26] M. Faustmann, J.M. Melenk, D. Praetorius, Existence of $\mathcal{H}$-matrix approximants to the inverse of BEM matrices: the hyper-singular integral operator, IMA J. Numer. Anal. 37 (3) (2017) 1211–1244.
[27] M. Faustmann, J.M. Melenk, M. Parvizi, $\mathcal{H}$-matrix approximability of inverses of FEM matrices for the time-harmonic Maxwell equations, Adv. Comput. Math. 48 (2022) 59.
[28] M. Faustmann, J.M. Melenk, M. Parvizi, Caccioppoli-type estimates and $\mathcal{H}$-matrix approximations to inverses for FEM-BEM couplings, Numer. Math. 150 (3) (2022) 849–892.
[29] T. Führer, J.M. Melenk, D. Praetorius, A. Rieder, Optimal additive Schwarz methods for the hp-BEM: the hypersingular integral operator in 3D on locally refined meshes, Comput. Math. Appl. 70 (7) (2015) 1583–1605.
[30] L. Grasedyck, W. Hackbusch, Construction and arithmetics of $\mathcal{H}$-matrices, Computing 70 (4) (2003) 295–334.
[31] L. Grasedyck, W. Hackbusch, S. Le Borne, Adaptive geometrically balanced clustering of $\mathcal{H}$-matrices, Computing 73 (1) (2004) 1–23.
[32] L. Grasedyck, Theorie und Anwendungen Hierarchischer Matrizen, Ph.D. thesis, Universität Kiel, 2001.
[33] W. Hackbusch, A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Introduction to $\mathcal{H}$-matrices, Computing 62 (2) (1999) 89–108.
[34] W. Hackbusch, Hierarchical Matrices: Algorithms and Analysis, Springer Series in Computational Mathematics, vol. 49, Springer, Heidelberg, 2015.
[35] B.N. Khoromskij, J.M. Melenk, An efficient direct solver for the boundary concentrated FEM in 2D, Computing 69 (2) (2002) 91–117.
[36] B.N. Khoromskij, J.M. Melenk, Boundary concentrated finite element methods, SIAM J. Numer. Anal. 41 (1) (2003) 1–36.
[37] R. Muñoz Sola, Polynomial liftings on a tetrahedron and applications to the h-p version of the finite element method in three dimensions, SIAM J. Numer. Anal. 34 (1) (1997) 282–314.
[38] J.M. Melenk, C. Rojik, On commuting p-version projection-based interpolation on tetrahedra, Math. Comput. 89 (321) (2020) 45–87.
[39] J. Schöberl, NGSolve, Available at https://ngsolve.org/, 2022.