Numer. Math. (2005) 99: 605–643
Digital Object Identifier (DOI) 10.1007/s00211-004-0564-3

Numerische Mathematik

Approximation of Integral Operators by Variable-Order Interpolation

Steffen Börm[1], **Maike Löhndorf**[1], **Jens M. Melenk**[2]

[1] Max-Planck-Institut für Mathematik in den Naturwissenschaften, Inselstrasse 22–26, 04103 Leipzig, Germany; e-mail: {sbo, mal}@mis.mpg.de
[2] Department of Mathematics, The University of Reading, P.O. Box 220, Whiteknights RG6 6AX, United Kingdom; e-mail: j.m.melenk@reading.ac.uk

Received September 26, 2002 / Revised version received October 31, 2003
Published online: January 17, 2005 –

Summary. We employ a data-sparse, recursive matrix representation, so-called $\mathcal{H}^2$-matrices, for the efficient treatment of discretized integral operators. We obtain this format using local tensor product interpolants of the kernel function and replacing high-order approximations with piecewise lower-order ones. The scheme has optimal, i.e., linear, complexity in the memory requirement and time for the matrix-vector multiplication. We present an error analysis for integral operators of order zero. In particular, we show that the optimal convergence $\mathcal{O}(h)$ is retained for the classical double layer potential discretized with piecewise constant functions.

1 Introduction and data-sparse representation

Integral operators arise in several applications, e.g., when partial differential equations that are originally posed on domains are reformulated as equations on the domain boundary or when transparent boundary conditions have to be modeled. Their efficient numerical treatment is non-trivial since, typically, the stiffness matrix $\mathbf{K} \in \mathbb{R}^{N\times N}$ corresponding to the discretized version of the operator is densely populated; thus setting it up and performing a matrix-vector multiplication is an $\mathcal{O}(N^2)$ operation.

Different schemes have been developed to reduce the complexity from $\mathcal{O}(N^2)$ to an almost linear complexity: for very simple domains and translation-invariant integral kernels, the fast Fourier transformation can be used to diagonalize $\mathbf{K}$. For a large class of domains, wavelet bases have been developed, [7], which permit efficient matrix compression by identifying and dropping the "insignificant" matrix entries.

A second very successful technique is based on approximating the integral kernel by degenerate kernels, which leads to an efficient block representation of the corresponding matrix. This approach is the foundation of the panel-clustering technique [16], multi-pole expansion methods [12,1], and hierarchical matrices ($\mathcal{H}$-matrices) [13,14,2] as well as the closely related mosaic-skeleton matrices [26].

An efficient specialization of $\mathcal{H}$-matrices are $\mathcal{H}^2$-matrices [15,3]. The algorithm presented and analyzed in the present paper stores the matrix $\mathbf{K}$ in this form. This $\mathcal{H}^2$-format is achieved by polynomial interpolation of the kernel function and successive re-interpolation of this polynomial on son clusters similar to the way [9,3] proceed.

The complexity of the typical algorithms for compressing integral operators is $\mathcal{O}(N \log^q N)$ for some $q > 0$. To reduce the complexity to $\mathcal{O}(N)$, variable-order approaches have recently been introduced in the panel-clustering context [22,23,17] and for wavelet-based compression techniques [25]. The basic idea is to employ lower approximation on the many small matrix blocks and only large approximation order on the few large ones. Our approach generalizes the techniques described in [22,23] and introduces an additional algorithmic simplification in that it is based on (Chebyshev) interpolation instead of Taylor expansion of the kernel function κ. This circumvents the need for a priori analytical treatment and explicit knowledge of derivatives of κ. The idea to interpolate the kernel function can be found, for example, in [5] and in [24,18], where a precise error analysis for the case of Chebyshev interpolation is presented.

The paper is organized as follows: in the present section we introduce the basic concepts of $\mathcal{H}^2$-matrices (Section 1.5) and the idea of kernel approximation by interpolation and successive re-interpolation of this polynomial on sub-domains (Section 1.6). The proof that the storage requirement and the cost of the matrix-vector multiplications are linear in the problem size is provided in Section 2. The analysis of our method requires the understanding of iterated polynomial interpolation schemes. Section 3 is devoted to the fundamental one-dimensional results: stability (cf. Theorem 3.11) and approximation (cf. Theorem 3.15). By tensor product arguments, these results are extended to the higher-dimensional case in Section 4 and used to prove optimal-order convergence of the approximate matrix in Theorem 4.6. The practical applicability of our scheme is demonstrated in Section 5, where we observe the optimal-order convergence $\mathcal{O}(h)$ for the two- and three-dimensional double layer potential discretized by piecewise constant finite element functions.

1.1 Model problem

For a kernel function $\kappa : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ on a bounded domain $\Gamma \subset \mathbb{R}^d$ we consider an integral operator $\mathcal{K} : L^2(\Gamma) \to L^2(\Gamma)$ of the form

$$\mathcal{K}[u](x) = \int_\Gamma \kappa(x, y)u(y)\,dy. \tag{1.1}$$

Concerning the integral kernel $\kappa : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, we assume furthermore that it is *asymptotically smooth*, i.e.,

$$\left|\partial_x^\alpha \partial_y^\beta \kappa(x, y)\right| \leq C_{\text{asymp}}(\alpha+\beta)!c_0^{|\alpha+\beta|}\,\|x-y\|^{-g-|\alpha|-|\beta|} \tag{1.2}$$

holds for all $x, y \in \mathbb{R}^d$, $x \neq y$, for all multi-indices $\alpha, \beta \in \mathbb{N}_0^d$ and some suitable $C_{\text{asymp}}, c_0 \geq 0$ and $g \in [0, d-1[$.

Remark 1.1 Our analysis in Section 4 will permit a more general setting to cover the case of boundary integral operators such as the double layer potential. Then Γ will also be allowed to be a d_Γ-dimensional manifold, and we consider the integral operator

$$\mathcal{K}[u](x) = \int_\Gamma \kappa(x, y)\omega(x, y)u(y)\,d\sigma(y),$$

where κ still satisfies (1.2) while ω is a separable function defined merely on $\Gamma \times \Gamma$ that can reduce the order of the singularity of κ. □

In order to discretize $\mathcal{K}$ by Galerkin's method, we introduce the corresponding bilinear form

$$a(u, v) := \int_\Gamma u(x) \int_\Gamma v(y)\kappa(x, y)\,dy\,dx \tag{1.3}$$

and fix an N-dimensional finite element space $V_N \subseteq L^2(\Gamma)$ spanned by basis functions Ψ_i, $i \in I$. Clearly $\#I = N$ holds. A typical example of basis functions is given by the standard piecewise polynomial finite element shape functions. Galerkin's method then leads to the stiffness matrix $\mathbf{K} \in \mathbb{R}^{N\times N}$ with entries

$$\mathbf{K}_{ij} = a(\Psi_i, \Psi_j) = \int_\Gamma \Psi_i(x) \int_\Gamma \kappa(x, y)\Psi_j(y)\,dy\,dx. \tag{1.4}$$

In this paper, we will construct a data-sparse approximation $\tilde{\mathbf{K}}$ of this stiffness matrix $\mathbf{K}$ and analyze its accuracy. This data-sparse approximation $\tilde{\mathbf{K}}$ will have the following form: for a partition P of the index set $I \times I$, a matrix block $\mathbf{K}|_{\tau\times\sigma}$ corresponding to the indices included in $\tau \times \sigma \in P$ is approximated by a low-rank matrix of the form

$$\tilde{\mathbf{K}}|_{\tau\times\sigma} = \mathbf{V}^\tau \mathbf{S}^{\tau,\sigma} \mathbf{W}^{\sigma\top} \tag{1.5}$$

if the block satisfies a so-called *admissibility condition*. We note the special structure of the approximation (1.5): the *cluster bases* $\mathbf{V}^\tau$, $\mathbf{W}^\sigma$ depend only on the *clusters* τ, σ and not on the product $\tau \times \sigma$; the connection between the

clusters τ and σ is described only by the *coefficient matrix* $\mathbf{S}^{\tau,\sigma}$. The cluster bases and the coefficient matrices are constructed by a procedure based on interpolating the kernel κ. The next section therefore introduces notation associated with interpolation that will be required in the following.

1.2 Notation

1.2.1 Interpolation A set $\tau \subset I$ of indices is called a *cluster*. For a cluster τ we define

$$\Gamma_\tau := \bigcup_{i\in\tau} \operatorname{supp}(\Psi_i),$$

and choose a closed axis-parallel box

$$B_\tau := J_1^\tau \times J_2^\tau \times \cdots \times J_d^\tau \supset \Gamma_\tau. \tag{1.6}$$

We consider interpolation schemes on such boxes. For each coordinate axis we choose a polynomial degree $k_j^\tau \in \mathbb{N}_0$, $j = 1, \ldots, d$, and collect these values in the degree vector $k^\tau \in \mathbb{N}_0^d$. Next, we choose (pairwise disjoint) interpolation points $x_{j,k'}^\tau \in J_j^\tau$, $k' = 0, \ldots, k_j^\tau$, for each of the closed intervals J_j^τ, $j = 1, \ldots, d$. These one-dimensional interpolation points are combined into d-dimensional interpolation points in the standard way by tensorization: for the set of relevant multi-indices

$$K_\tau := \prod_{j=1}^{d} \{0, \ldots, k_j^\tau\} \tag{1.7}$$

the interpolation points x_ν^τ and the Lagrange interpolation polynomials $\mathcal{L}_\nu^\tau$ are given by

$$x_\nu^\tau := (x_{1,\nu_1}^\tau, x_{2,\nu_2}^\tau, \ldots, x_{d,\nu_d}^\tau) \in B_\tau, \quad \nu \in K_\tau,$$

$$\mathcal{L}_\nu^\tau(x) := \prod_{j=1}^{d} \prod_{l=0, l\neq\nu_j}^{k_j^\tau} \frac{x_j - x_{j,l}^\tau}{x_{j,\nu_j}^\tau - x_{j,l}^\tau}.$$

The interpolation operator $\mathcal{I}_\tau$ associated with B_τ is defined in the standard way by

$$\mathcal{I}_\tau[u] := \sum_{\mu\in K_\tau} u(x_\mu^\tau)\mathcal{L}_\mu^\tau. \tag{1.8}$$

1.2.2 Admissible blocks A block $\tau \times \sigma \subset I \times I$ is called a *far-field block* (or: an *admissible block*) if, for a given $\eta > 0$, the *admissibility condition*

$$\operatorname{diam}(B_\tau \times B_\sigma) \leq 2\eta \operatorname{dist}(B_\tau, B_\sigma) \tag{1.9}$$

is satisfied. Otherwise, the block $\tau \times \sigma$ is called a *near-field block*. In our compression scheme, the matrix is represented exactly on near-field blocks, i.e.,

the approximation $\widetilde{\mathbf{K}}$ satisfies $\tilde{\mathbf{K}}|_{\tau\times\sigma} = \mathbf{K}|_{\tau\times\sigma}$. On far-field blocks $\tau \times \sigma$, the approximation $\widetilde{\mathbf{K}}|_{\tau\times\sigma}$ is obtained by replacing the kernel κ by an interpolant.

1.3 Kernel approximation by degenerate kernels: the constant-order case

To explain our construction of the cluster bases $\mathbf{V}^\tau$, $\mathbf{W}^\sigma$ and the coefficient matrices $\mathbf{S}^{\tau,\sigma}$ in (1.5) and to clarify the meaning of the admissibility condition (1.9), we now consider the simpler case of constant approximation order; that is, the degree vectors k^τ of Section 1.2.1 are the same for all clusters. For simplicity, we assume them to be of the form $k^\tau = (k, k, \ldots, k) \in \mathbb{N}_0^d$.

The asymptotic smoothness of the kernel function κ guarantees that the blow-up of κ and its derivatives is controlled as $x - y \to 0$. This property permits us to approximate $\mathbf{K}$ accurately by block-wise low-rank matrices as we now elaborate.

Let P be a partition of $I \times I$. On near-field blocks $\tau \times \sigma$, we take, as stated above, $\tilde{\mathbf{K}}|_{\tau\times\sigma} = \mathbf{K}|_{\tau\times\sigma}$. For far-field blocks $\tau \times \sigma$ we approximate $\kappa|_{B_\tau \times B_\sigma}$ by its interpolating polynomial

$$\tilde{\kappa}_{\tau\times\sigma}(x, y) := (\mathcal{I}_\tau \otimes \mathcal{I}_\sigma)[\kappa](x, y) = \sum_{\nu\in K_\tau} \sum_{\mu\in K_\sigma} \kappa(x_\nu^\tau, x_\mu^\sigma)\mathcal{L}_\nu^\tau(x)\mathcal{L}_\mu^\sigma(y)$$

and define the matrix approximation $\tilde{\mathbf{K}}|_{\tau\times\sigma}$ by (1.5) with cluster bases $\mathbf{V}^\tau$, $\mathbf{W}^\sigma$ and coefficient matrix $\mathbf{S}^{\tau,\sigma}$ given by

$$\mathbf{V}_{i\nu}^\tau := \int_\Gamma \mathcal{L}_\nu^\tau(x)\Psi_i(x)\,dx, \quad \mathbf{W}_{j\mu}^\sigma := \int_\Gamma \mathcal{L}_\mu^\sigma(y)\Psi_j(y)\,dy \quad \text{and} \tag{1.10}$$

$$\mathbf{S}_{\nu\mu}^{\tau,\sigma} := \kappa(x_\nu^\tau, x_\mu^\sigma). \tag{1.11}$$

Typical k-th order interpolation schemes such as the tensor product Chebyshev interpolation lead to an error estimate of the form

$$\begin{aligned}&\|\kappa - \tilde{\kappa}_{\tau\times\sigma}\|_{\infty, B_\tau\times B_\sigma}\\ &\leq C\frac{c_1^{k+1}}{(k+1)!}\operatorname{diam}(B_\tau \times B_\sigma)^{k+1} \sup_{|\alpha+\beta|\leq k+1} \|\partial_x^\alpha \partial_y^\beta \kappa\|_{\infty, B_\tau\times B_\sigma},\end{aligned}$$

where C, $c_1 > 0$ are suitable constants. Combining this estimate with the asymptotic smoothness of κ and the admissibility condition (1.9), we get

$$\begin{aligned}
&\|\kappa - \tilde{\kappa}_{\tau\times\sigma}\|_{\infty,B_\tau\times B_\sigma}\\
&\overset{(1.2)}{\leq} C(c_0c_1)^{k+1}\left(\frac{\operatorname{diam}(B_\tau\times B_\sigma)}{\operatorname{dist}(B_\tau,B_\sigma)}\right)^{k+1}\operatorname{dist}(B_\tau,B_\sigma)^{-g}\\
&\overset{(1.9)}{\leq} C(2c_0c_1\eta)^{k+1}\operatorname{dist}(B_\tau,B_\sigma)^{-g};
\end{aligned}$$

thus we obtain *exponential convergence* in the polynomial degree k if the parameter η in the admissiblity condition is sufficiently small. The error $\|\mathbf{K}-\tilde{\mathbf{K}}\|_2$ can then be decreased by increasing the approximation order k.

1.4 Cluster tree and block partition

Having established the admissibility condition (1.9) as the key property of far-field blocks we now turn to an algorithm for selecting them. To do this efficiently, we will not consider all subsets $\tau, \sigma \subseteq I$ of the finite index set, which would lead to a complexity of 2^N, but organize the subsets hierarchically.

A tree $\mathcal{T}_I$ whose nodes are collected in the set T_I is called a *cluster tree* if all its nodes are subsets of I, if its root is the index set I and if each non-leaf node τ has at least two sons and is their disjoint union (cf. [10,11] for algorithms for the construction of cluster trees). The nodes of the cluster tree are called *clusters*, and the set of leaves is denoted by L_I. The set of sons of a cluster τ is denoted by $\operatorname{sons}(\tau)$. For clusters $\tau \neq I$, we denote furthermore by $\operatorname{father}(\tau)$ the unique cluster σ with $\tau \in \operatorname{sons}(\sigma)$.

Given a cluster tree, we construct a partition P of $I \times I$ consisting of admissible blocks and blocks corresponding to leaves of the cluster tree by recursive subdivision: If a block is admissible, it is added to the partition. If a block is not admissible and if it can be split into sub-blocks, the procedure is applied recursively to these sub-blocks. Otherwise, the block is added to the near-field. Calling the following procedure with $\tau = \sigma = I$ and $P = \emptyset$ yields the desired partition:

Algorithm 1.2 (Block partition)

```
procedure divide(τ, σ, var P);
begin
  {Admissible block}
  if τ × σ is admissible then P := P ∪ {τ × σ}
  {Non-admissible leaf block}
  else if τ and σ is a leaf then P := P ∪ {τ × σ}
  {Check sub-blocks}
  else if τ is a leaf then for all σ' ∈ sons σ do divide(τ, σ', P)
  else if σ is a leaf then for all τ' ∈ sons τ do divide(τ', σ, P)
  else for all τ' ∈ sons τ and σ' ∈ sons σ do divide(τ', σ', P)
end
```

For typical cluster trees, this algorithm has complexity $\mathcal{O}(N)$ and creates a partition P with $\mathcal{O}(N)$ admissible blocks and $\mathcal{O}(N)$ leaf blocks (cf. [10, 11]).

1.5 Nested bases and $\mathcal{H}^2$-matrices

In the case of constant approximation order discussed in Section 1.3 the cluster bases $\mathbf{V}^\tau$ (and $\mathbf{W}^\sigma$) defined in (1.10) have a special structure. For a cluster τ' and its father $\tau := \operatorname{father}(\tau')$ we have

$$\mathcal{L}_\nu^\tau = \mathcal{I}_{\tau'}\mathcal{L}_\nu^\tau = \sum_{\nu' \in K_{\tau'}} \mathcal{L}_\nu^\tau(x_{\nu'}^{\tau'})\mathcal{L}_{\nu'}^{\tau'} = \sum_{\nu' \in K_{\tau'}} \mathbf{B}_{\nu'\nu}^{\tau',\tau}\mathcal{L}_{\nu'}^{\tau'} \tag{1.12}$$

with

$$\mathbf{B}_{\nu'\nu}^{\tau',\tau} := \mathcal{L}_\nu^\tau(x_{\nu'}^{\tau'}); \tag{1.13}$$

therefore

$$\begin{aligned}\mathbf{V}_{i\nu}^\tau &= \int_\Gamma (\mathcal{I}_{\tau'}\mathcal{L}_\nu^\tau)(x)\Psi_i(x)\,dx = \sum_{\nu' \in K_{\tau'}} \mathbf{B}_{\nu'\nu}^{\tau',\tau} \int_\Gamma \mathcal{L}_{\nu'}^{\tau'}(x)\Psi_i(x)\,dx \\ &= \sum_{\nu' \in K_{\tau'}} \mathbf{B}_{\nu',\nu}^{\tau',\tau} V_{i\nu'}^{\tau'} = (\mathbf{V}^{\tau'}\mathbf{B}^{\tau',\tau})_{i\nu}\end{aligned} \tag{1.14}$$

holds for all $i \in \tau'$ and $\nu \in K_\tau$. A completely analogous formula holds for the matrices $\mathbf{W}^\sigma$. This relation makes use of the *nestedness* of the cluster bases; i.e., the cluster bases of father clusters can be represented in terms of the son clusters. This property can be exploited algorithmically: it suffices to store the matrices $\mathbf{V}^\tau$, $\mathbf{W}^\tau$ for leaves τ and the transfer matrices $\mathbf{B}^{\tau',\operatorname{father}(\tau')}$ for all clusters $\tau' \in T_I \setminus \{I\}$.

Definition 1.3 ($\mathcal{H}^2$-matrix) *Let $P = P_{\text{near}} \dot\cup P_{\text{far}}$ be a block partition, and let $\mathbf{V} = (\mathbf{V}^\tau)_{\tau \in T_I}$ and $\mathbf{W} = (\mathbf{W}^\sigma)_{\sigma \in T_I}$ be nested cluster bases.*

A matrix $\mathbf{M} \in \mathbb{R}^{I \times I}$ is called an $\mathcal{H}^2$-matrix *(cf. [15]) based on P, $\mathbf{V}$ and $\mathbf{W}$ if, for each $\tau \times \sigma \in P_{\text{far}}$, there exists a matrix $\mathbf{S}^{\tau,\sigma} \in \mathbb{R}^{K_\tau \times K_\sigma}$ with*

$$\mathbf{M}|_{\tau\times\sigma} = \mathbf{V}^\tau \mathbf{S}^{\tau,\sigma}(\mathbf{W}^\sigma)^\top.$$

$\mathbf{V}$ is called the row cluster basis *and $\mathbf{W}$ is called the* column cluster basis *of $\mathbf{M}$.*

1.6 Variable-order approximation

In Section 1.3 we interpolated the kernel function κ by polynomials of a fixed degree k on bounding boxes $B_\tau \times B_\sigma$ corresponding to admissible blocks $\tau \times \sigma$. We now wish to exploit the possibilities inherent in the general interpolation scheme introduced in Section 1.2.1; that is, we wish to assign

different polynomial degrees to different clusters. This freedom to choose the approximation order permits us to find more efficient low-rank approximations of admissible matrix blocks. While—as we saw in Section 1.5—nestedness of the cluster bases $\mathbf{V}^\tau$, $\mathbf{W}^\tau$ is automatically guaranteed in the case of constant approximation order k, special care has to be taken to ensure this desirable property in the case of variable approximation order; we discuss in the present section how to realize this algorithmically.

1.6.1 $\mathcal{H}^2$-matrix approximation by variable-order interpolation We assume that a cluster tree $\mathcal{T}_I$ is given and that a polynomial degree vector $k^\tau \in \mathbb{N}_0^d$ is associated with each cluster $\tau \in T_I$. Suitable strategies for choosing the degrees will be investigated in Section 4 and in Section 5.

As in the case of constant approximation order the approximation $\tilde{\mathbf{K}}|_{\tau\times\sigma}$ for a far-field block $\tau \times \sigma$ is sought in the form (1.5), where the coefficient matrix $\mathbf{S}^{\tau,\sigma}$ is again given by (1.11), i.e.,

$$\mathbf{S}^{\tau,\sigma}_{\mu,\nu} = \kappa(x^\tau_\nu, x^\sigma_\mu), \quad \nu \in K_\tau, \quad \mu \in K_\sigma.$$

In order to define the cluster bases $\mathbf{V}^\tau$ with the desired nestedness property, we construct them recursively using (1.14) and a suitable definition of the transfer matrices $\mathbf{B}^{\tau_0,\tau_1}$ in the following way (the construction for the $\mathbf{W}^\tau$ is completely analogous):

- For leaf clusters $\tau_0 \in L_I$, the definition is retained, i.e.,

$$\mathbf{V}^{\tau_0}_{i\nu} = \int_\Gamma \mathcal{L}^{\tau_0}_\nu(x)\Psi_i(x)\,dx \quad \text{for } i \in \tau_0 \text{ and } \nu \in K_{\tau_0}.$$

- For τ_1 with $\operatorname{sons}(\tau_1) \neq \emptyset$ we define $\mathbf{V}^{\tau_1}$ as follows: for each son $\tau_0 \in \operatorname{sons}(\tau_1)$ the entries $\mathbf{V}^{\tau_1}_{i\nu}$, $i \in \tau_0$, $\nu \in K_{\tau_1}$, of $\mathbf{V}^{\tau_1}$ are given by

$$\mathbf{V}^{\tau_1}_{i\nu} := \sum_{\nu'\in K_{\tau_0}} \mathbf{V}^{\tau_0}_{i\nu'}\mathbf{B}^{\tau_0,\tau_1}_{\nu'\nu},$$

  where the transfer matrix $\mathbf{B}^{\tau_0,\tau_1}$ is *defined* as

$$\mathbf{B}^{\tau_0,\tau_1}_{\nu'\nu} := \mathcal{L}^{\tau_1}_\nu(x^{\tau_0}_{\nu'}), \qquad \nu' \in K_{\tau_0}, \nu \in K_{\tau_1}.$$

The recursively defined cluster bases $\mathbf{V}^\tau$ are nested by construction. To obtain a better understanding of these cluster bases $\mathbf{V}^\tau$, we observe that for $\tau_0 \in \operatorname{sons}(\tau_1)$ and $i \in \tau_0$, $\nu \in K_{\tau_1}$, the entries of $\mathbf{V}^{\tau_1}_{i,\nu}$ can be written as

$$\begin{aligned}\mathbf{V}^{\tau_1}_{i\nu} &= \sum_{\nu'\in K_{\tau_0}} \mathbf{V}^{\tau_0}_{i\nu'}\mathbf{B}^{\tau_0,\tau_1}_{\nu'\nu} = \int_\Gamma \left(\sum_{\nu'\in K_{\tau_0}} \mathcal{L}^{\tau_1}_\nu(x^{\tau_0}_{\nu'})\mathcal{L}^{\tau_0}_{\nu'}\right)(x)\Psi_i(x)\,dx \\ &= \int_\Gamma \mathcal{I}_{\tau_0}[\mathcal{L}^{\tau_1}_\nu](x)\Psi_i(x)\,dx.\end{aligned}$$

Proceeding by induction, we find

$$\mathbf{V}_{i\nu}^{\tau_n} = \int_\Gamma (\mathcal{I}_{\tau_0} \circ \mathcal{I}_{\tau_1} \circ \cdots \circ \mathcal{I}_{\tau_{n-1}})[\mathcal{L}_\nu^{\tau_n}](x)\Psi_i(x)\,dx \tag{1.15}$$

for $i \in \tau_0$, $\nu \in K_{\tau_n}$ and a cluster sequence $\tau_0 \subseteq \tau_1 \subseteq \cdots \subseteq \tau_n$ with $\tau_{j-1} \in \operatorname{sons}(\tau_j)$ for $j \in \{1, \ldots, n\}$.

As in the constant order case the Lagrange polynomial $\mathcal{L}_\nu^\tau$ is replaced with an interpolant. However, whereas previously we had the equality (1.12), we now have to take into account that $\mathcal{I}_{\tau'}[\mathcal{L}_\nu^\tau] \neq \mathcal{L}_\nu^\tau$ in general. In fact, a main aspect of the present paper is the analysis of the stability and approximation properties of the iterated interpolation operator that appears in the representation formula (1.15).

2 Algorithms and Complexity

In this section we show that setting up the $\mathcal{H}^2$*-matrix* $\tilde{\mathbf{K}}$ and performing a matrix-vector multiplication is in $\mathcal{O}(N)$. Since all estimates in this context follow the same pattern we establish some notation:

Definition 2.1 (Tree levels) *We define the function*

$$\operatorname{level} : T_I \to \mathbb{N}_0$$

inductively by $\operatorname{level}(I) := 0$ *and* $\operatorname{level}(\tau) := \operatorname{level}(\operatorname{father}(\tau)) + 1$ *for* $\tau \in T_I \setminus \{I\}$. *Let* $\ell_{\max} := \max \operatorname{level}(T_I)$. *The families* $(T_{I,\ell})_{\ell=0}^{\ell_{\max}}$ *and* $(L_{I,\ell})_{\ell=0}^{\ell_{\max}}$ *defined by*

$$T_{I,\ell} := \{\tau \in T_I \ : \ \operatorname{level}(\tau) = \ell\}, \quad L_{I,\ell} := \{\tau \in L_I \ : \ \operatorname{level}(\tau) = \ell\}$$

are partitions of the nodes T_I *and the set of leaves* L_I.

We split the partition P into *farfield* and *nearfield* blocks:

$$P_{\text{far}} := \{\tau \times \sigma \in P \ : \ \tau \times \sigma \text{ is admissible}\} \quad \text{and} \quad P_{\text{near}} := P \setminus P_{\text{far}}.$$

Notation 2.2 *The number of elementary arithmetic operations for a computation* op*, such as matrix-vector multiplication or setting up leaf basis matrices, is denoted by* $\mathcal{W}_{\text{op}}$. *Let* $\mathcal{N}_{\text{op}}$ *be the set of "nodes" (i.e., clusters or blocks) involved in the operation* op *and* $\mathcal{N}_{\text{op},\ell}$ *the nodes on level* ℓ. $\mathcal{W}_{\text{op}}(n)$ *is the work required for* $n \in \mathcal{N}_{\text{op}}$ *and* $\mathcal{W}_{\text{op}}(\ell) := \max_{n \in \mathcal{N}_{\text{op},\ell}} \mathcal{W}_{\text{op}}(n)$.

Lemma 2.3 *If there exists a polynomial π with positive coefficients and if there are $C > 1, c \in \mathbb{R}$ with*

$$\mathcal{W}_{\mathrm{op}}(\ell) \leq \pi(\ell_{\max} - \ell) \quad \text{and} \quad \#\mathcal{N}_{\mathrm{op},\ell} \leq cNC^{\ell-\ell_{\max}} \tag{2.1}$$

for the operation op*, then it is of linear complexity, i.e., there is C^* with*

$$\mathcal{W}_{\mathrm{op}} \leq C^* N.$$

Proof We have

$$\begin{aligned}\mathcal{W}_{\mathrm{op}} &= \sum_{n\in\mathcal{N}_{\mathrm{op}}} \mathcal{W}_{\mathrm{op}}(n) = \sum_{\ell=0}^{\ell_{\max}} \sum_{n\in\mathcal{N}_{\mathrm{op},\ell}} \mathcal{W}_{\mathrm{op}}(n) \leq \sum_{\ell=0}^{\ell_{\max}} \mathcal{W}_{\mathrm{op}}(\ell)\#\mathcal{N}_{\mathrm{op},\ell} \\ &\leq Nc \sum_{\ell=0}^{\ell_{\max}} \pi(\ell_{\max} - \ell) C^{\ell-\ell_{\max}} \leq Nc \sum_{\ell'=0}^{\infty} \pi(\ell') C^{-\ell'}.\end{aligned}$$

Since $C > 1$ the infinite sum is obviously bounded. □

To fulfill the conditions (2.1) we need the following assumptions on the cluster tree, the block partition, and the rank distribution.

Assumption 2.4 *The block partition P is* sparse *in the sense of [10,11], i.e., there exists a constant C_{sp} such that*

$$\max_{\tau\in T_I} \#\{\sigma \in T_I \,:\, \tau \times \sigma \in P \vee \sigma \times \tau \in P\} \leq C_{\mathrm{sp}}. \tag{2.2}$$

For standard situations with quasi-uniform meshes, this estimate has been established in [10,11].

Definition 2.5 (Nearly balanced trees) *A cluster tree $\mathcal{T}_I$ is called* nearly balanced*, if there are constants $C_{\mathrm{sons}} \in \mathbb{R}_{>1}$ and $C_{\mathrm{bal}} \in \mathbb{R}_{>0}$ such that*

$$\#T_{I,\ell} \leq C_{\mathrm{bal}} N / C_{\mathrm{sons}}^{\ell_{\max}-\ell} \qquad \text{for all } \ell \in \{0, \ldots, \ell_{\max}\}.$$

Example 2.6 A balanced binary tree is also nearly balanced with $C_{\mathrm{sons}} = 2$ and $C_{\mathrm{bal}} = 1$.

Assumption 2.7 *$\mathcal{T}_I$ is nearly balanced and there are polynomials π_{order} and π_{bal} with positive coefficients satisfying*

$$\begin{aligned}|k^\tau|_\infty &\leq \pi_{\mathrm{order}}(\ell_{\max} - \mathrm{level}(\tau)) \quad \text{for all } \tau \in T_I, \\ \#\tau &\leq \pi_{\mathrm{bal}}(\ell_{\max} - \mathrm{level}(\tau)) \quad \text{for all } \tau \in L_I.\end{aligned}$$

For quasi-uniform grids and standard clustering strategies, this assumption can be expected to hold for the rank distribution computed with Algorithm 5.1.

Remark 2.8 Let Assumption 2.7 be valid. Then for clusters $\tau \in T_I$ we have for the index set K_τ of interpolation points (cf. (1.7))

$$\#K_\tau \leq \pi^d_{\text{order}}(\ell_{\max} - \text{level}(\tau)),$$

which means that all terms with sums and products of $\#K_\tau$, $|k^\tau|_\infty$ for $\tau \in T_I$ and terms including $\#K_\tau$, $|k^\tau|_\infty$, $\#\tau$ for all $\tau \in L_I$ can be bounded by a polynomial in $\ell_{\max} - \text{level}(\tau)$ with positive coefficients.

Since $\mathcal{T}_I$ is nearly balanced, $\#\mathcal{N}_{\text{op},\ell} \leq cNC^{\ell-\ell_{\max}}$ holds for all $\mathcal{N}_{\text{op}} \subseteq T_I$. For $\mathcal{N}_{\text{op}} \subseteq P$ this estimate follows from Assumption 2.4 since $\#\{\tau \times \sigma \in P : \tau \in T_{I,\ell}\} \leq C_{\text{sp}}\#T_{I,\ell} \leq C_{\text{sp}}C_{\text{bal}}N/C_{\text{sons}}^{\ell_{\max}-\ell}$. □

2.1 Matrix-vector multiplication

The following algorithm (cf. [15, Section 3.2]) computes the matrix-vector product for an $\mathcal{H}^2$-Matrix $\tilde{\mathbf{K}}$ for a given vector $x \in \mathbb{R}^I$:

Algorithm 2.9 (Matrix-vector multiplication)

```
procedure FastMVM(x, var y);
begin
  for σ ∈ T_I \ L_I do x_σ := 0;                          {forward transformation}
  for σ ∈ L_I do x_σ := W^{σT} x_{|σ};
  for ℓ = ℓ_max downto 1 do
    for σ ∈ T_{I,ℓ} do x_{father(σ)} := x_{father(σ)} + (B^{σ,father(σ)})^T x_σ;
  for τ ∈ T_I do y_τ := 0;                                {multiplication}
  for τ × σ ∈ P_far do y_τ := y_τ + S^{τ,σ} x_σ;
  for ℓ = 1 to ℓ_max do                                   {backward transformation}
    for τ ∈ T_{I,ℓ} do y_τ := y_τ + B^{τ,father(τ)} x_{father(τ)};
  for τ ∈ L_I do y_{|τ} := V^τ x_τ;
  for τ × σ ∈ P \ P_far do y_{|τ} := y_{|τ} + K̃_{|τ×σ} x_{|σ}   {Near-field}
end
```

The far-field part splits into three steps. First, the products of the input vector with the matrices $\mathbf{W}^{\sigma T}$ are computed by recursively applying equation (1.14). Then, the resulting coefficients are multiplied by $\mathbf{S}^{\tau,\sigma}$. In the last step, the result is transformed back from the coefficients of the cluster bases into the standard basis.

Lemma 2.10 (Far-field matrix-vector multiplication) *The complexity for the far-field matrix-vector multiplication is $\mathcal{O}(N)$ if the Assumptions 2.7 and 2.4 are valid.*

Proof There is a constant C_m such that the matrix-vector multiplication with an $n_1 \times n_2$ matrix requires $C_m n_1 n_2$ operations. Concerning the forward and

backward transformation we see that $\mathcal{N}_{\text{trans}} \subseteq T_I$. $\mathcal{W}_{\text{trans}}(\tau)$ involves the multiplication with a $\#\tau \times \#K_\tau$ matrix for leaves and with a $\#K_{\text{father}(\tau)} \times \#K_\tau$ matrix for all nodes except I, i.e.,

$$\mathcal{W}_{\text{trans}}(\tau) \leq C_m \, \#K_\tau (\#\tau \; + \#K_{\text{father}(\tau)}) \quad \text{for all } \tau \in L_I$$
$$\mathcal{W}_{\text{trans}}(\tau) \leq C_m \, \#K_\tau \#K_{\text{father}(\tau)} \quad \text{for all } \tau \in T_I \setminus \{L_I \cup \{I\}\}.$$

Remark 2.8 together with Lemma 2.3 gives linear complexity for both transformations. With $\mathcal{N}_{\text{mul}} := P_{\text{far}}$ we get linear complexity for the multiplication part by Lemma 2.3 and Remark 2.8 observing that $\mathcal{W}_{\text{mul}}(\tau \times \sigma) \leq C_m \#K_\tau \#K_\sigma$. □

2.2 Far-field computation

Setting up the far-field is performed by the following algorithm:

Algorithm 2.11 (Far-field computation)
`procedure` *FarFieldSetup*;
`begin`
  `for` $\sigma \in L_I$ *compute* $\mathbf{V}^\sigma$, $\mathbf{W}^\sigma$; {*cf.* (1.10)}
  `for` $\sigma \in T_I \setminus \{I\}$ *compute* $\mathbf{B}^{\sigma,\text{father}(\sigma)}$; {*cf.* (1.14)}
  `for` $\tau \times \sigma \in P_{\text{far}}$ *compute* $\mathbf{S}^{\tau,\sigma}$ {*cf.* (1.11)}
`end`

Lemma 2.12 (Far-field computation) *If the Assumptions 2.7 and 2.4 hold, if the basis functions Ψ_i are piecewise polynomial functions and if the manifold Γ is piecewise affine (e.g., the surface of a polygon), then Algorithm 2.11 has linear complexity.*

Proof The node sets required in Algorithm 2.11 fulfill (2.1) by Remark 2.8. Since calculating $\mathbf{B}^{\sigma,\text{father}(\sigma)}$, $\mathbf{V}^\sigma$ (and $\mathbf{W}^\sigma$) involves Lagrange polynomials, we note that $\mathcal{L}_\iota^\sigma$, for $\sigma \in \mathcal{T}_I$ and $\iota \in K_\sigma$, is a polynomial of order $|k^\sigma|_1$. It can be evaluated with work bounded by $C_{\text{eval}}|k^\sigma|_1 \leq C_{\text{eval}} d |k^\sigma|_\infty$ for some $C_{\text{eval}} \in \mathbb{R}$, and its product with a piecewise polynomial finite element basis function $\mathcal{L}_\iota^\sigma \Psi_i$ for $i \in \sigma$ can be integrated exactly with Gaussian quadrature with $\mathcal{O}(|k^\tau|_1^{d_\Gamma})$ quadrature points. The matrix $\mathbf{B}^{\sigma,\text{father}(\sigma)} = (\mathcal{L}_\nu^{\text{father}(\sigma)}(x_{\nu'}^\sigma))_{\nu' \in K_\sigma, \nu \in K_{\text{father}(\sigma)}}$ can be computed with work bounded by $C_{\text{eval}} d |k^\sigma|_\infty \#K_\sigma \#K_{\text{father}(\sigma)}$. For leaves $\sigma \in L_I$ the setup of the matrix $\mathbf{V}^\sigma = \left(\int_\Gamma \mathcal{L}_\iota^\sigma(x)\Psi_i(x)dx\right)_{i \in \sigma, \iota \in K_\sigma}$ takes $\mathcal{O}(|k^\sigma|_1^{d_\Gamma+1} \#\sigma \, \#K_\sigma)$ operations. Computing the coefficient matrix $\mathbf{S}^{\tau,\sigma} = \left(\kappa(x_\iota^\tau, x_\lambda^\sigma)\right)_{\iota \in K_\tau, \lambda \in K_\sigma}$ requires $\mathcal{O}(\#K_\tau \, \#K_\sigma)$ operations. Remark 2.8 together with Lemma 2.3 concludes the proof. □

In order to ensure that the Gaussian quadrature is exact, we have assumed that the surface Γ is polygonal. For general curved surfaces, Gaussian quadrature entails additional quadrature errors, whose analysis is beyond the scope the present paper.

3 Stability and approximation properties of one-dimensional iterated polynomial interpolation

To construct a nested basis we re-interpolate successively the polynomials of the expansion of the kernel function with a lower degree on son clusters down to the leaves of the cluster tree. We prove that this iterated interpolation process is stable and analyze its approximation properties under the assumption that the Lebesgue constant of the underlying interpolation process grows at most polynomially. This assumption is typically satisfied, in particular for Chebyshev, Gauß-Legendre, or Gauß-Lobatto interpolation.

We first focus on the one-dimensional case; the multivariate case will then be obtained by tensor product arguments. In one step of the interpolation process a polynomial on a large domain is interpolated on a significantly smaller one by polynomials of lower degree. In the first part we show how accurately this can be done depending on the polynomial degrees and the size of the domains.

In the next part we prove stability (Theorem 3.11) and approximation properties (Theorem 3.15) for an iterated interpolation process associated with a hierarchy of clusters, in which the degree of approximation decreases as the size of the clusters gets smaller.

3.1 Polynomial approximation in one dimension

3.1.1 Notation and interpolation operators

Notation 3.1 *We will make use of the following notation:*

1. **(Affine Scaling)** *Let $J = [a, b] \subset \mathbb{R}$ be an interval with $a < b$. The affine scaling function*

$$\Phi_J : \mathbb{C} \to \mathbb{C}, \quad z \mapsto ((1+z)b + (1-z)a)/2$$

   maps $[-1, 1]$ into J and is invertible.
2. **(Regularity Ellipses)** *For $\rho \in \mathbb{R}_{>1}$ we denote by $\mathcal{E}_\rho \subset \mathbb{C}$ the ellipse with foci ± 1 and semi axes of lengths $\frac{1}{2}(\rho + 1/\rho)$, $\frac{1}{2}(\rho - 1/\rho)$, i.e.,*

$$\mathcal{E}_\rho := \left\{ z \in \mathbb{C} \,:\, \frac{4(\operatorname{Re} z)^2}{(\rho + 1/\rho)^2} + \frac{4(\operatorname{Im} z)^2}{(\rho - 1/\rho)^2} = 1 \right\}.$$

   The interior is denoted by $\operatorname{int}(\mathcal{E}_\rho)$.

We will consider polynomial interpolation operators that satisfy the following standard assumptions:

Assumption 3.2 (Interpolation scheme) *The family* $(\mathcal{I}_k)_{k\in\mathbb{N}_0}$ *of polynomial interpolation operators of degree* k

$$\mathcal{I}_k : C([-1,1]) \to \mathcal{P}_k := \operatorname{span}\{x^i \mid 0 \le i \le k\}$$

consists of projections, i.e., for each $k \in \mathbb{N}_0$ *they satisfy*

$$\mathcal{I}_k u = u \quad \text{for all } u \in \mathcal{P}_k. \tag{3.1}$$

Additionally, for constants C_Λ, $\lambda \in \mathbb{R}_{>0}$, *the Lebesgue constants* Λ_k *satisfy*

$$\Lambda_k := \sup_{\substack{u\in C([-1,1])\\ u\neq 0}} \frac{\|\mathcal{I}_k u\|_{\infty,[-1,1]}}{\|u\|_{\infty,[-1,1]}} \le C_\Lambda (k+1)^\lambda \quad \text{for all } k \in \mathbb{N}_0. \tag{3.2}$$

For an arbitrary closed interval J, we define the scaled interpolation operators

$$\mathcal{I}_{k,J} : C(J) \to \mathcal{P}_k,\ u \mapsto (\mathcal{I}_k(u \circ \Phi_J)) \circ \Phi_J^{-1},$$

which also inherit the properties (3.2), (3.1), i.e.,

$$\mathcal{I}_{k,J} u = u \qquad \text{for all } u \in \mathcal{P}_k, \tag{3.3}$$

$$\|\mathcal{I}_{k,J} u\|_{\infty,J} \le \Lambda_k \|u\|_{\infty,J} \qquad \text{for all } u \in C(J). \tag{3.4}$$

Example 3.3 The Chebyshev interpolation operator, on which our numerical experiments in Section 5.2 are based, satisfies the estimate (3.2) with $\lambda = 1$ and $C_\Lambda = 1$, since by [21] $\Lambda_k \le \frac{2}{\pi}\ln(k+1) + 1$. □

Concerning polynomial approximation of analytic functions, we have the following well-known results:

Lemma 3.4 *Let* $\rho > 1$. *Let* $u \in L^\infty(\operatorname{int}(\mathcal{E}_\rho))$ *be holomorphic on* $\operatorname{int}(\mathcal{E}_\rho)$. *Then*

$$\min_{v\in\mathcal{P}_k} \|u - v\|_{\infty,[-1,1]} \le \frac{2\rho}{\rho - 1} \rho^{-(k+1)} \|u\|_{\infty,\operatorname{int}(\mathcal{E}_\rho)}.$$

Proof See, e.g., [8, Chap. 7, Sec. 8, eq. (8.7)]. □

3.1.2 Approximation of polynomials by polynomials of lower degree

Lemma 3.5 *Let* $[a,b] \subseteq [-1,1]$, *let* $p, k \in \mathbb{N}_0$ *with* $k \le p$. *For every* $u \in \mathcal{P}_p$ *and* $\rho \in \mathbb{R}_{\ge 2}$, $\sigma \in \mathbb{R}_{>1}$ *satisfying*

$$\Phi_{[a,b]}(\operatorname{int}(\mathcal{E}_\rho)) \subseteq \operatorname{int}(\mathcal{E}_\sigma),$$

we have

$$\min_{v\in\mathcal{P}_k} \|u - v\|_{\infty,[a,b]} \le 4\rho^{-k-1}\sigma^p \|u\|_{\infty,[-1,1]}. \tag{3.5}$$

Proof Due to the maximum principle for holomorphic functions, we have

$$\|u\|_{\infty,\Phi_{[a,b]}(\mathcal{E}_\rho)} \leq \|u\|_{\infty,\mathcal{E}_\sigma}.$$

Since $u \in \mathcal{P}_p$, Bernstein's estimate for polynomials [8, Thm. 2.2] implies

$$\|u\|_{\infty,\mathcal{E}_\sigma} \leq \sigma^p \|u\|_{\infty,[-1,1]}.$$

Let $\hat{u} := u \circ \Phi_{[a,b]}^{-1} \in \mathcal{P}_p$. Due to Lemma 3.4, there is a $\hat{v} \in \mathcal{P}_k$ satisfying

$$\|\hat{u} - \hat{v}\|_{\infty,[-1,1]} \leq \frac{2\rho}{\rho - 1} \rho^{-k-1} \|\hat{u}\|_{\infty,\mathcal{E}_\rho}.$$

By defining $v := \hat{v} \circ \Phi_{[a,b]} \in \mathcal{P}_k$, we find

$$\begin{aligned}\|u - v\|_{\infty,[a,b]} &= \|\hat{u} - \hat{v}\|_{\infty,[-1,1]} \leq \frac{2\rho}{\rho - 1} \rho^{-k-1} \|\hat{u}\|_{\infty,\mathcal{E}_\rho} \\ &= \frac{2\rho}{\rho - 1} \rho^{-k-1} \|u\|_{\infty,\Phi_{[a,b]}(\mathcal{E}_\rho)} \leq \frac{2\rho}{\rho - 1} \rho^{-k-1} \|u\|_{\infty,\mathcal{E}_\sigma} \\ &\leq \frac{2\rho}{\rho - 1} \rho^{-k-1} \sigma^p \|u\|_{\infty,[-1,1]},\end{aligned}$$

concluding our proof since the assumption $\rho \geq 2$ implies $\frac{2\rho}{\rho-1} \leq 4$. □

Theorem 3.6 *For all intervals* $[a, b] \subseteq [-1, 1]$*, all* $p \in \mathbb{N}$*,* $k \in \mathbb{N}_0$ *with* $k \leq p$*, and all* $u \in \mathcal{P}_p$

$$\min_{v \in \mathcal{P}_k} \|u - v\|_{\infty,[a,b]} \leq 4e(4p)^{p-(k+1)} \left(\frac{b-a}{2}\right)^{k+1} \|u\|_{\infty,[-1,1]}.$$

Proof Let $\delta := (b - a)/2$. We introduce

$$\sigma : \mathbb{R}_{\geq 1} \to \mathbb{R}_{\geq 1}, \quad \rho \mapsto 3 + \delta(\rho + 1/\rho),$$

and find that for all $\rho \in \mathbb{R}_{>0}$

$$\begin{aligned}\sigma(\rho) - 1/\sigma(\rho) &= 3 + \delta(\rho + 1/\rho) - \frac{1}{3 + \delta(\rho + 1/\rho)} \\ &\geq 2 + \delta(\rho + 1/\rho) \geq |a + b| + \delta(\rho + 1/\rho).\end{aligned}$$

Hence, upon writing $B(z, r) \subset \mathbb{C}$ for the ball of radius $r > 0$ about $z \in \mathbb{C}$,

$$\begin{aligned}\Phi_{[a,b]}(\mathrm{int}(\mathcal{E}_\rho)) &\subseteq B\left(0, \frac{|a + b| + \delta(\rho + 1/\rho)}{2}\right) \\ &\subseteq B\left(0, \frac{\sigma(\rho) - 1/\sigma(\rho)}{2}\right) \subseteq \mathrm{int}(\mathcal{E}_{\sigma(\rho)}).\end{aligned}$$

Thus, we can apply Lemma 3.5 to the pair $(\rho, \sigma(\rho))$ and find a $v \in \mathcal{P}_k$ with

$$\|u - v\|_{\infty,[a,b]} \leq 4\rho^{p-k-1} \left(\frac{\sigma(\rho)}{\rho}\right)^p \|u\|_{\infty,[-1,1]}. \tag{3.6}$$

The polynomial v depends on the parameter $\rho \in \mathbb{R}_{\geq 2}$, which is still at our disposal. For a good choice, we calculate

$$\begin{aligned}\left(\frac{\sigma(\rho)}{\rho}\right)^p &= \left(\frac{3+\delta(\rho+1/\rho)}{\rho}\right)^p = \left(\delta + \frac{3}{\rho} + \frac{\delta}{\rho^2}\right)^p \leq \left(\delta + \frac{4}{\rho}\right)^p \\ &= \delta^p \exp\left(p \log\left(1 + \frac{4}{\delta\rho}\right)\right) \leq \delta^p \exp\left(p\frac{4}{\delta\rho}\right),\end{aligned}$$

so that by choosing $\rho := 4p/\delta \geq 4p/2 \geq 2$, we find

$$\left(\frac{\sigma(\rho)}{\rho}\right)^p \leq \delta^p e.$$

Therefore (3.6) takes the form

$$\begin{aligned}\|u - v\|_{\infty,[a,b]} &\leq 4e\rho^{p-k-1}\delta^p \|u\|_{\infty,[-1,1]} \\ &\leq 4e(4p)^{p-k-1}\delta^{k+1}\|u\|_{\infty,[-1,1]},\end{aligned}$$

which is the desired result. □

A consequence of Theorem 3.6 is the following corollary.

Corollary 3.7 *Let $J' \subseteq J$ be closed intervals. Then for each $u \in \mathcal{P}_p$ and each $k \in \mathbb{N}_0$, there exists a $v \in \mathcal{P}_k$ satisfying*

$$\|u - v\|_{\infty,J'} \leq 4e(4p)^{p-k-1} \left(\frac{|J'|}{|J|}\right)^{k+1} \|u\|_{\infty,J}. \tag{3.7}$$

The best approximation result Corollary 3.7 allows us to quantify the interpolation error:

Corollary 3.8 *Let $J' \subseteq J$ be closed intervals and let $(\mathcal{I}_l)_{l\in\mathbb{N}_0}$ be a family of interpolation operators satisfying Assumption 3.2. For $p, k \in \mathbb{N}_0$ with $k \leq p$ and for all $u \in \mathcal{P}_p$ we have*

$$\|u - \mathcal{I}_{k,J'}u\|_{\infty,J'} \leq 4e(1+\Lambda_k)(4p)^{p-k-1} \left(\frac{|J'|}{|J|}\right)^{k+1} \|u\|_{\infty,J}, \tag{3.8}$$

$$\|\mathcal{I}_{k,J'}u\|_{\infty,J'} \leq \left(1 + 4e(1+\Lambda_k)(4p)^{p-k-1} \left(\frac{|J'|}{|J|}\right)^{k+1}\right) \|u\|_{\infty,J}. \tag{3.9}$$

Proof Let $v \in \mathcal{P}_k$ be the approximation to u given by Corollary 3.7. Then

$$\begin{aligned}
&\|u - \mathcal{I}_{k,J'}u\|_{\infty,J'} \\
&= \|u - v - (\mathcal{I}_{k,J'}u - v)\|_{\infty,J'} \leq \|u - v\|_{\infty,J'} + \|\mathcal{I}_{k,J'}(u - v)\|_{\infty,J'} \\
&\leq (1 + \Lambda_k)\|u - v\|_{\infty,J'} \leq (1 + \Lambda_k)4e(4p)^{p-k-1}\left(\frac{|J'|}{|J|}\right)^{k+1}\|u\|_{\infty,J}.
\end{aligned}$$

The triangle equality finally establishes (3.9). □

3.2 Iterated interpolation with polynomials in one dimension

The results of the last section allow us to estimate the error introduced when interpolating polynomials of higher degree by polynomials of lower degree on smaller domains.

Now we are going to consider the iteration of this procedure, namely, the re-interpolation of a higher-order polynomial on a *sequence* of sub-domains by polynomials of lower order. We will prove that this process is stable as long as the decrease in degree is slow enough compared to the reduction of the size of the sub-domains.

Definition 3.9 (Interval chain) *A sequence $\mathfrak{C} = (J^i)_{i=0}^n$ of closed, bounded intervals with $J^0 \subseteq J^1 \subseteq \cdots \subseteq J^n$ is called an* interval chain. *It is $\bar{q}$*-regular *for $\bar{q} \in]0, 1[$ if for all $i \in \{1, \ldots, n\}$ there holds*

$$\frac{|J^{i-1}|}{|J^i|} \leq \bar{q}. \tag{3.10}$$

Each interval chain gives rise to an interpolation operator:

Definition 3.10 (Iterated interpolation) *Given an interval chain $\mathfrak{C}$ and α, $\beta \in \mathbb{N}_0$, we set $k_i := \beta + \alpha i$ and define the induced* iterated interpolation operator *by*

$$\mathcal{I}_{\mathfrak{C}} := \mathcal{I}_{k_0,J^0} \circ \mathcal{I}_{k_1,J^1} \circ \cdots \circ \mathcal{I}_{k_n,J^n}. \tag{3.11}$$

Theorem 3.11 (Stability) *Let $\mathfrak{C} = (J^i)_{i=0}^n$ be a $\bar{q}$-regular interval chain and $\mathcal{I}_{\mathfrak{C}}$ be the induced iterated interpolation operator. Then there exists $C_s > 0$ depending only on α, $\bar{q}$, and the constants C_Λ, λ of* (3.2) *such that*

$$\|\mathcal{I}_{k_0,J^0} \circ \cdots \circ \mathcal{I}_{k_m,J^m} w\|_{\infty,J^0} \leq C_s \Lambda_{\beta+\alpha m}\|w\|_{\infty,J^m} \tag{3.12}$$

for all $w \in C(J^m)$ and all $m \in \{1, \ldots, n\}$. In particular

$$\|\mathcal{I}_{\mathfrak{C}} w\|_{\infty,J^0} \leq C_s \Lambda_{\beta+\alpha n}\|w\|_{\infty,J^n} \quad \text{for all } w \in C(J^n).$$

Proof The case $\alpha = 0$ is trivial, so we only consider $\alpha \geq 1$. Let $m \in \{1, \ldots, n\}$. By Corollary 3.8, we have

$$\|\mathcal{I}_{k_0,J^0} \circ \cdots \circ \mathcal{I}_{k_{m-1},J^{m-1}} u\|_{\infty,J^0} \leq F \|u\|_{\infty,J^m}$$

for all $u \in \mathcal{P}_{\beta+\alpha m}$ with

$$F := \prod_{j=0}^{m-1} \left(1 + 4e(1 + \Lambda_{\beta+\alpha j})(4(\beta + \alpha(j+1)))^{\alpha-1} \bar{q}^{\beta+\alpha j+1}\right)$$

Since $\ln(1+x) \leq x$ for $x \geq 0$, we find

$$\begin{aligned}\ln F &\leq 4e \sum_{j=0}^{m-1} (1 + \Lambda_{\beta+\alpha j})(4(\beta + \alpha j + \alpha))^{\alpha-1} \bar{q}^{\beta+\alpha j+1} \\ &\leq 4e \sum_{j=0}^{\infty} (1 + \Lambda_j)(4(j + \alpha - 1))^{\alpha-1} \bar{q}^{j} \leq 4e\tilde{C}_{\bar{q}},\end{aligned}$$

where $\tilde{C}_{\bar{q}}$ depends on $\bar{q}$, α, and the constants C_Λ, λ of (3.2). Therefore

$$\|\mathcal{I}_{k_0,J^0} \circ \cdots \circ \mathcal{I}_{k_{m-1},J^{m-1}} u\|_{\infty,J^0} \leq C_s \|u\|_{\infty,J^m} \tag{3.13}$$

holds for all $u \in \mathcal{P}_{\beta+\alpha m}$ with

$$C_s := \exp\left(4e\tilde{C}_{\bar{q}}\right).$$

For $w \in C(J^m)$ we get by combining (3.13) with (3.4)

$$\begin{aligned}\|\mathcal{I}_{k_0,J^0} \circ \cdots \circ \mathcal{I}_{k_m,J^m} w\|_{\infty,J^0} &\leq C_s \|\mathcal{I}_{k_m,J^m} w\|_{\infty,J^m} \\ &\leq C_s \Lambda_{\beta+\alpha m} \|w\|_{\infty,J^m}.\end{aligned}$$

This is the desired estimate. □

For the approximation properties of $\mathcal{I}_{\mathfrak{C}}$, we have the following estimate:

Lemma 3.12 (Approximation properties of iterated interpolation) *Let $(J^i)_{i=0}^n$ be a $\bar{q}$-regular interval chain and let $\mathcal{I}_{\mathfrak{C}}$ be the corresponding iterated interpolation operator. Then we have for all $u \in C(J^n)$*

$$\|u - \mathcal{I}_{\mathfrak{C}} u\|_{\infty,J^0} \leq C_\Lambda C_s \sum_{j=0}^{n} (\beta + \alpha j + 1)^{\lambda} \|u - \mathcal{I}_{k_j,J^j} u\|_{\infty,J^j}.$$

Proof For each $j \in \{0, \dots, n\}$, we define

$$u_j := \mathcal{I}_{k_0, J^0} \circ \cdots \circ \mathcal{I}_{k_j, J^j} u.$$

We have $u - \mathcal{I}_{\mathfrak{C}} u = u - u_n$ and therefore by Theorem 3.11

$$\begin{aligned}
\|u - \mathcal{I}_{\mathfrak{C}} u\|_{\infty, J^0} &\leq \|u - u_0\|_{\infty, J^0} + \sum_{j=0}^{n-1} \|u_j - u_{j+1}\|_{\infty, J^0} \\
&= \|u - u_0\|_{\infty, J^0} + \sum_{j=0}^{n-1} \|\mathcal{I}_{k_0, J^0} \circ \cdots \circ \mathcal{I}_{k_j, J^j} (u - \mathcal{I}_{k_{j+1}, J^{j+1}} u)\|_{\infty, J^0} \\
&\leq \|u - \mathcal{I}_{k_0, J^0} u\|_{\infty, J^0} + C_s \sum_{j=0}^{n-1} \Lambda_{\beta+\alpha j} \|u - \mathcal{I}_{k_{j+1}, J^{j+1}} u\|_{\infty, J^{j+1}} \\
&\leq C_\Lambda C_s \sum_{j=0}^{n} (\beta + \alpha j + 1)^\lambda \|u - \mathcal{I}_{k_j, J^j} u\|_{\infty, J^j},
\end{aligned}$$

which concludes the argument. □

3.3 Polynomial approximation of analytic functions

3.3.1 Approximation on single intervals We now address the question of polynomial approximating univariate functions u that are analytic on a closed interval $J \subseteq \mathbb{R}$.

Lemma 3.13 (Approximation of analytic functions) *Let $J \subseteq \mathbb{R}$ be a closed, bounded interval and assume that a function u satisfies for some C_u, $\gamma_u \geq 0$*

$$\|u^{(n)}\|_{\infty, J} \leq C_u \gamma_u^n n! \quad \text{for all } n \in \mathbb{N}_0. \tag{3.14}$$

Then we have for all $k \in \mathbb{N}_0$

$$\min_{v \in \mathcal{P}_k} \|u - v\|_{\infty, J} \leq C_u 4e(1 + \gamma_u |J|)(k+1) \left(1 + \frac{2}{\gamma_u |J|}\right)^{-(k+1)}. \tag{3.15}$$

Proof Recall that the affine bijection Φ_J maps the reference interval $[-1, 1]$ onto J. Defining $\hat{u} := u \circ \Phi_J$ and observing

$$\min_{v \in \mathcal{P}_k} \|u - v\|_{\infty, J} = \min_{v \in \mathcal{P}_k} \|\hat{u} - v\|_{\infty, [-1,1]},$$

we may restrict our attention to the polynomial approximation of $\hat{u}$. Since $\Phi_J' = \frac{|J|}{2}$, we get

$$\|\hat{u}^{(n)}\|_{\infty, [-1,1]} \leq C_u \hat{\gamma}_u^{-n} n! \quad \text{for all } n \in \mathbb{N}_0, \tag{3.16}$$

where we abbreviated

$$\hat{\gamma}_u := \frac{2}{\gamma_u |J|}. \tag{3.17}$$

This bound implies that the Taylor series of $\hat{u}$ at a point $\xi \in]-1,1[$ converges on the ball $B(\xi, \hat{\gamma}_u) \subset \mathbb{C}$ with center ξ and radius $\hat{\gamma}_u$; thus, $\hat{u}$ has a holomorphic extension to the set

$$G_{\hat{\gamma}_u} := \cup_{\xi \in]-1,1[} B(\xi, \hat{\gamma}_u).$$

Lemma 3.14 below then implies $\mathcal{E}_{\hat{\rho}} \subset G_{\hat{\gamma}_u}$ for

$$\hat{\rho} := 1 + \hat{\gamma}_u > 1. \tag{3.18}$$

For an ultimate application of Lemma 3.4, we need to bound $\hat{u}$. To that end let $\varepsilon > 0$, which will be chosen more precisely below. On the subset $G_{\hat{\gamma}_u/(1+\varepsilon)} \subset G_{\hat{\gamma}_u}$ we obtain by Taylor expansion around the points $\xi \in]-1,1[$ in view of the estimate (3.16)

$$\|\hat{u}\|_{\infty, G_{\hat{\gamma}_u/(1+\varepsilon)}} \leq C_u \sum_{n=0}^{\infty} (1+\varepsilon)^{-n} = C_u \frac{1+\varepsilon}{\varepsilon}.$$

Since for $\hat{\rho}_\varepsilon := 1 + \hat{\gamma}_u/(1+\varepsilon)$ we have $\operatorname{int}(\mathcal{E}_{\hat{\rho}_\varepsilon}) \subset G_{\hat{\gamma}_u/(1+\varepsilon)}$, we obtain from Lemma 3.4 the following approximation result:

$$\begin{aligned}\min_{v \in \mathcal{P}_k} \|\hat{u} - v\|_{\infty,[-1,1]} &\leq \frac{2\hat{\rho}_\varepsilon}{\hat{\rho}_\varepsilon - 1} \hat{\rho}_\varepsilon^{-(k+1)} \|\hat{u}\|_{\infty, \mathcal{E}_{\hat{\rho}_\varepsilon}} \\ &\leq C_u \frac{2\hat{\rho}_\varepsilon}{\hat{\rho}_\varepsilon - 1} \hat{\rho}_\varepsilon^{-(k+1)} \frac{1+\varepsilon}{\varepsilon}.\end{aligned}$$

We choose $\varepsilon = \frac{1}{k+1}$ and bound $1+\varepsilon \leq 2$, $\hat{\rho}_\varepsilon/(\hat{\rho}_\varepsilon - 1) \leq 1 + 2/\hat{\gamma}_u$. It remains to bound $\hat{\rho}_\varepsilon^{-(k+1)}$:

$$\begin{aligned}\hat{\rho}_\varepsilon^{-(k+1)} &= \left(1 + \frac{\hat{\gamma}_u}{1+\varepsilon}\right)^{-(k+1)} = \frac{(1+\epsilon)^{k+1}}{(1+\hat{\gamma}_u)^{k+1}} \left(\frac{1+\hat{\gamma}_u}{1+\hat{\gamma}_u+\epsilon}\right)^{k+1} \\ &\leq \hat{\rho}^{-(k+1)} \left(1 + \frac{1}{k+1}\right)^{k+1} \leq \hat{\rho}^{-(k+1)} e.\end{aligned}$$

Combining the above estimates, we obtain

$$\min_{v \in \mathcal{P}_k} \|\hat{u} - v\|_{\infty,[-1,1]} \leq C_u 4e \left(1 + \frac{2}{\hat{\gamma}_u}\right) (k+1) \hat{\rho}^{-(k+1)},$$

which is the desired result in view of the definition of $\hat{\rho}$ and $\hat{\gamma}_u$ in (3.17), (3.18). □

Lemma 3.14 (Covering of $\mathcal{E}_\rho$) *Let $\gamma > 0$ and $\rho = \gamma + \sqrt{1+\gamma^2}$. Then $\operatorname{int}(\mathcal{E}_\rho) \subset \cup_{\xi \in]-1,1[} B(\xi, \gamma)$. In particular, for $\hat{\rho} := 1 + \gamma < \rho$, we find $\operatorname{int} \mathcal{E}_{\hat{\rho}} \subset \operatorname{int} \mathcal{E}_\rho \subset \cup_{\xi \in]-1,1[} B(\xi, \gamma)$.*

Proof This follows from elementary computations (cf. [4] for details). □

3.3.2 Approximation by iterated interpolation We now turn to the question of approximating analytic functions by iterated interpolation.

Theorem 3.15 (Iterated approximation) *Let $(J^i)_{i=0}^n$ be a $\bar{q}$-regular interval chain, let $\mathcal{I}_{\mathfrak{C}}$ be the induced iterated interpolation operator for the rank parameters $\alpha, \beta \in \mathbb{N}_0$ (see Definition 3.10). Let $u \in C^\infty(J^n)$ satisfy*

$$\|u^{(\nu)}\|_{\infty,J^n} \leq C_u \gamma_u^\nu \nu! \quad \text{for all } \nu \in \mathbb{N}_0. \tag{3.19}$$

Then the iterated interpolation $\mathcal{I}_{\mathfrak{C}}$ satisfies

$$\|u - \mathcal{I}_{\mathfrak{C}} u\|_{\infty,J^0} \leq C_{1d}(\alpha,\beta,n)\,(1+\gamma)^{-(\beta+1)-\min\{(\beta+1)\bar{\alpha},\alpha\}n},$$

where the parameters $\gamma > 0$ and $\bar{\alpha}$ are given by

$$\gamma = \min\left\{\frac{2}{\gamma_u |J^n|}, \frac{1}{\bar{q}}\right\}, \qquad \bar{\alpha} = C_q \frac{2}{2+\gamma_u |J^n| \bar{q}} > 0,$$

for a constant $C_q \in \mathbb{R}_{>0}$ that depends only on $\bar{q}$. The function C_{1d} grows polynomially in n and linearly in $|J^n|$ and is given by

$$C_{1d}(\alpha,\beta,n) = C_s C_u 8e C_\Lambda^2 (1+\gamma_u|J^n|)(n+1)(\beta+1+\alpha n)^{2\lambda+1}.$$

Proof As a first step, we will prove the following slightly stronger statement:

(3.20)

$$\begin{aligned}&\|u - \mathcal{I}_{\mathfrak{C}} u\|_{\infty,J^0}\\ &\leq C(\alpha,\beta,n) \max\left\{\left(\frac{\gamma_u |J^n| \bar{q}^n}{2+\gamma_u |J^n| \bar{q}^n}\right)^{\beta+1}, \left(\frac{\gamma_u |J^n|}{2+\gamma_u |J^n|}\right)^{\beta+1+\alpha n}\right\}.\end{aligned}$$

From Lemma 3.12 we get

$$\|u - \mathcal{I}_{\mathfrak{C}} u\|_{\infty,J^0} \leq C_\Lambda C_s \sum_{i=0}^n (\beta+\alpha i+1)^\lambda \|u - \mathcal{I}_{\beta+\alpha i, J^i} u\|_{\infty,J^i}.$$

Furthermore, we have the standard estimate

$$\|u - \mathcal{I}_{\beta+\alpha i, J^i} u\|_{\infty,J^i} \leq (1+\Lambda_{\beta+\alpha i}) \min_{v \in \mathcal{P}_{\beta+\alpha i}} \|u - v\|_{\infty,J^i}.$$

Hence, using $1 \leq \Lambda_{\beta+\alpha i}$, we arrive at

$$\|u - \mathcal{I}_{\mathfrak{C}} u\|_{\infty,J^0} \leq 2C_\Lambda^2 C_s \sum_{i=0}^n (\beta+\alpha i+1)^{2\lambda} \min_{v \in \mathcal{P}_{\beta+\alpha i}} \|u - v\|_{\infty,J^i}.$$

The assumptions on the decay of the lengths of the intervals in the cluster chain imply

$$|J^i| \leq \bar{q}^{n-i} |J^n|.$$

Hence, the approximation result Lemma 3.13 implies

$$\min_{v\in\mathcal{P}_{\beta+\alpha i}} \|u-v\|_{\infty,J^i} \leq C_u 4e(1+\gamma_u|J_n|)(\beta+1+\alpha i)\left(1+\frac{2}{\gamma_u|J_n|\overline{q}^{n-i}}\right)^{-(\beta+1+\alpha i)}$$

for $i\in\{0,\ldots,n\}$, leading to

$$\|u-\mathcal{I}_{\mathfrak{C}}u\|_{\infty,J^0} \leq C_\Lambda^2 C_s C_u 8e(1+\gamma_u|J^n|)\left(\sum_{k=0}^{n}(\beta+1+\alpha k)^{2\lambda+1}\right) \times \max_{i\in\{0,\ldots,n\}}\left(1+\frac{2}{\gamma_u|J^n|\overline{q}^{n-i}}\right)^{-(\beta+1+\alpha i)}.$$

The sum can be bounded by $(n+1)(\beta+1+\alpha n)^{2\lambda+1}$. The maximum can be bounded by Lemma 3.16 below by

$$\max\left\{\left(1+\frac{2}{\gamma_u|J^n|\overline{q}^n}\right)^{-(\beta+1)},\left(1+\frac{2}{\gamma_u|J^n|}\right)^{-(\beta+1+\alpha n)}\right\}.$$

This yields the bound (3.20). To obtain the bound stated in the Theorem we apply Lemma 3.17 to the first term of the maximum in (3.20) with $\delta = 2/(\gamma_u|J^n|)$ and $q=\bar{q}$. □

We conclude this subsection with two lemmas used in the last proof.

Lemma 3.16 *Let C, s, $\alpha>0$, $\beta\in\mathbb{R}$. Then for $n>0$ the function*

$$f:[0,n]\to\mathbb{R},\qquad x\mapsto -(\beta+\alpha x)\ln(1+Ce^{s(n-x)})$$

attains its maximum at one of the endpoints, i.e.,

$$\max_{x\in[0,n]} f(x) = \max\{f(0),f(n)\}.$$

Proof We can demonstrate that each local extremum in $]0,n[$ is a minimum by using elementary analysis (cf. [4] for details). □

Lemma 3.17 *Let $\delta>0$, $q\in]0,1[$, $\beta\geq -1$. Then, upon setting*

$$\zeta := \min\{\delta,1/q\}>0,$$
$$\bar{\alpha} := \frac{1-q}{\ln(1+1/q)}\frac{\delta/q}{1+\delta/q}>0 \tag{3.21}$$

we have

$$\left(1+\frac{\delta}{q^n}\right)^{-(\beta+1)} \leq (1+\zeta)^{-(\beta+1)(1+\bar{\alpha}n)} \quad \text{for all } n\in\mathbb{N}_0. \tag{3.22}$$

Proof We take the logarithm of both sides and show that the estimate holds for $n=0$, $n\to\infty$ and all inner extremal points (cf. [4] for details). □

4 Global error bounds

Having established stability and approximation estimates for the one-dimensional case in Section 3, we can turn to the multi-dimensional domains used in the construction in Section 1.6.

4.1 Multi-dimensional error estimate

Definition 4.1 (Descendants) *For all $\tau, \sigma \in T_I$ we define*

$$\begin{aligned}\operatorname{sons}^*(\tau) &:= \{\tau_0 \in L_I \ : \ \text{there exists a cluster chain } \tau_0 \subseteq \cdots \subseteq \tau_n = \tau \\ &\qquad \text{with } \tau_{i-1} \in \operatorname{sons}(\tau_i) \text{ for } i \in \{1, \ldots, n\}\}, \\ \operatorname{sons}^*(\tau \times \sigma) &:= \{\tau_0 \times \sigma_0 \ : \ \tau_0 \in \operatorname{sons}^*(\tau), \sigma_0 \in \operatorname{sons}^*(\sigma)\}.\end{aligned}$$

We associate a closed bounding box

$$B_\tau = J_1^\tau \times \cdots \times J_d^\tau \tag{4.1}$$

with each cluster $\tau \in T_I$ (cf. (1.6)). On these boxes, we can combine the one-dimensional interpolation operators $\mathcal{I}_{k,J}$ introduced in Assumption 3.2 to define the cluster interpolation operators (cf. (1.8)) by

$$\mathcal{I}_\tau := \mathcal{I}_{k_1^\tau, J_1^\tau} \otimes \cdots \otimes \mathcal{I}_{k_d^\tau, J_d^\tau}.$$

Let $\tau \times \sigma \in P_{\text{far}}$ and $\tau_0 \times \sigma_0 \in \operatorname{sons}^*(\tau \times \sigma)$ with cluster chains $\tau_0 \subseteq \cdots \subseteq \tau_n$ and $\sigma_0 \subseteq \cdots \subseteq \sigma_m$. Combining (1.5), (1.11), and (1.15), we see that for $i \in \tau_0$ and $j \in \sigma_0$, we have the representation

$$\begin{aligned}\tilde{\mathbf{K}}_{ij} &= \int_\Gamma \int_\Gamma \tilde{\kappa}^{\tau\times\sigma}_{\tau_0\times\sigma_0}(x, y)\Psi_i(x)\Psi_j(y)\, dx\, dy, \\ \tilde{\kappa}^{\tau\times\sigma}_{\tau_0\times\sigma_0} &:= \left(\mathcal{I}_{\tau_0} \circ \cdots \circ \mathcal{I}_{\tau_n}\right) \otimes \left(\mathcal{I}_{\sigma_0} \circ \cdots \circ \mathcal{I}_{\sigma_m}\right) \kappa.\end{aligned}$$

To derive error bounds for the approximation $\tilde{\mathbf{K}}$, we have to analyze the $2d$-dimensional iterated interpolation operator

$$\mathcal{I}^{\tau\times\sigma}_{\tau_0\times\sigma_0} := \left(\mathcal{I}_{\tau_0} \circ \cdots \circ \mathcal{I}_{\tau_n}\right) \otimes \left(\mathcal{I}_{\sigma_0} \circ \cdots \circ \mathcal{I}_{\sigma_m}\right).$$

Since interpolation operators acting on different variables commute, we can rearrange the operators in the following way:

$$\begin{aligned}&\mathcal{I}_{\tau_0} \circ \cdots \circ \mathcal{I}_{\tau_n} \\ &\quad = (\mathcal{I}_{k_1^{\tau_0}, J_1^{\tau_0}} \circ \cdots \circ \mathcal{I}_{k_1^{\tau_n}, J_1^{\tau_n}}) \otimes \cdots \otimes (\mathcal{I}_{k_d^{\tau_0}, J_d^{\tau_0}} \circ \cdots \circ \mathcal{I}_{k_d^{\tau_n}, J_d^{\tau_n}}) \quad \text{and} \\ &\mathcal{I}_{\sigma_0} \circ \cdots \circ \mathcal{I}_{\sigma_m} \\ &\quad = (\mathcal{I}_{k_1^{\sigma_0}, J_1^{\sigma_0}} \circ \cdots \circ \mathcal{I}_{k_1^{\sigma_n}, J_1^{\sigma_m}}) \otimes \cdots \otimes (\mathcal{I}_{k_d^{\sigma_0}, J_d^{\sigma_0}} \circ \cdots \circ \mathcal{I}_{k_d^{\sigma_n}, J_d^{\sigma_m}}).\end{aligned}$$

Upon introducing the one-dimensional iterated interpolation operators

$$\mathcal{I}_{\mathfrak{C},j} := \mathcal{I}_{k_j^{\tau_0},J_j^{\tau_0}} \circ \cdots \circ \mathcal{I}_{k_j^{\tau_n},J_j^{\tau_n}}, \quad \mathcal{I}_{\mathfrak{C},d+j} := \mathcal{I}_{k_j^{\sigma_0},J_j^{\sigma_0}} \circ \cdots \circ \mathcal{I}_{k_j^{\sigma_m},J_j^{\sigma_m}}$$

for $j \in \{1, \ldots, d\}$, we can write

$$\mathcal{I}_{\tau_0\times\sigma_0}^{\tau\times\sigma} = (\mathcal{I}_{\tau_0} \circ \cdots \circ \mathcal{I}_{\tau_n}) \otimes (\mathcal{I}_{\sigma_0} \circ \cdots \circ \mathcal{I}_{\sigma_m}) = \mathcal{I}_{\mathfrak{C},1} \otimes \cdots \otimes \mathcal{I}_{\mathfrak{C},2d},$$

i.e., we now have to analyze only one-dimensional iterated interpolations.

In order to keep the presentation simple, we will consider only the case of regular bounding boxes. Analogous to Definition 3.9 we require:

Assumption 4.2 (Regular clusters) *Let $\bar{q} \in]0, 1[$. We assume that $B_{\tau'} \subseteq B_\tau$ and*

$$|J_j^{\tau'}| \leq \bar{q}|J_j^\tau|$$

(cf. (4.1)) hold for all $j \in \{1, \ldots, d\}$, all $\tau \in T_I$, and all $\tau' \in \operatorname{sons}(\tau)$.

For quasi-uniform meshes, this assumption can be satisfied if the correct clustering strategy is used (cf. [11]).

Assumption 4.2 ensures that all cluster chains appearing in our proof will be $\bar{q}$-regular, so we can use the simple rank distribution

$$k^\tau := \beta + \alpha(\ell_{\max} - \operatorname{level}(\tau)) \tag{4.2}$$

without sacrificing stability.

Assumption 4.3 (Local homogeneity) *We assume that the block partition P is* locally homogeneous*, i.e., that there is a $\delta \in \mathbb{N}_0$ such that*

$$|\operatorname{level}(\tau) - \operatorname{level}(\sigma)| \leq \delta$$

holds for all $\tau \times \sigma \in P_{\text{far}}$.

This assumption is satisfied for locally quasi-uniform meshes and allows us to simplify the analysis significantly by characterizing blocks $\tau \times \sigma \in P_{\text{far}}$ by the average of the level of τ and that of σ. Using these averages, the following approximation result can be proven:

Theorem 4.4 (Local error estimate) *Let $p \in]0, 1[$ and $C_{\text{in}} \in \mathbb{R}_{>0}$. Let Assumptions 4.2 and 4.3 hold. Then there are $\alpha, \beta \in \mathbb{N}_0$ (defining the rank distribution due to (4.2)) such that*

$$\|\kappa - \mathcal{I}_{\tau_0\times\sigma_0}^{\tau\times\sigma}\kappa\|_{\infty,B_\tau\times B_\sigma} \leq \frac{C_{\text{in}} p^{\ell_{\max}-\ell}}{\operatorname{dist}(B_\tau, B_\sigma)^g}$$

holds for all $\tau \times \sigma \in P_{\text{far}}$, where $\ell := (\operatorname{level}(\tau) + \operatorname{level}(\sigma))/2$.

Proof We will first derive general error estimates for the operator $\mathcal{I}^{\tau\times\sigma}_{\tau_0\times\sigma_0}$. Let $\alpha, \beta \in \mathbb{N}_0$. Due to (4.2) and Assumption 4.2, we can use Theorem 3.11 to find

$$\begin{aligned}
&\left\| \left(\bigotimes_{i=1}^{j-1} \mathrm{Id} \otimes \mathcal{I}_{\mathfrak{C},j} \otimes \bigotimes_{i=j+1}^{2d} \mathrm{Id} \right) u \right\|_{\infty, B_{\tau_0}\times B_{\sigma_0}} \\
&\quad \leq C_s C_\Lambda (k^{\tau_0} + 1 + \alpha(\mathrm{level}(\tau_0) - \mathrm{level}(\tau)))^\lambda \|u\|_{\infty, B_\tau \times B_\sigma} \\
&\quad \leq C_s C_\Lambda (\beta + 1 + \alpha(\ell_{\max} - \mathrm{level}(\tau)))^\lambda \|u\|_{\infty, B_\tau \times B_\sigma} \qquad (4.3)
\end{aligned}$$

for all $u \in C(B_\tau \times B_\sigma)$ and all $j \in \{1, \ldots, d\}$, i.e., the tensorized interpolation operator is stable. We observe

$$\begin{aligned}
\ell_{\max} - \mathrm{level}(\tau) &= \ell_{\max} - \mathrm{level}(\tau)/2 - \mathrm{level}(\tau)/2 \\
&\quad + \mathrm{level}(\sigma)/2 - \mathrm{level}(\sigma)/2 \leq \ell_{\max} - \ell + \delta/2,
\end{aligned}$$

which implies

$$\begin{aligned}
(\beta + 1 + \alpha(\ell_{\max} - \mathrm{level}(\tau)))^\lambda &\leq (\beta + 1 + \alpha(\ell_{\max} - \ell) + \alpha\delta/2)^\lambda \\
&\leq (\alpha\delta/2 + 1)^\lambda (\beta + 2 + \alpha(\ell_{\max} - \ell))^\lambda,
\end{aligned}$$

so the stability estimate (4.3) takes the form

$$\begin{aligned}
&\left\| \left(\bigotimes_{i=1}^{j-1} \mathrm{Id} \otimes \mathcal{I}_{\mathfrak{C},j} \otimes \bigotimes_{i=j+1}^{2d} \mathrm{Id} \right) u \right\|_{\infty, B_{\tau_0}\times B_{\sigma_0}} \\
&\quad \leq C_s C_\Lambda (k^{\tau_0} + 1 + \alpha(\mathrm{level}(\tau_0) - \mathrm{level}(\tau)))^\lambda \|u\|_{\infty, B_\tau \times B_\sigma} \\
&\quad \leq C_s C_\Lambda (\alpha\delta/2 + 1)^\lambda (\beta + 2 + \alpha(\ell_{\max} - \ell))^\lambda \|u\|_{\infty, B_\tau \times B_\sigma} \qquad (4.4)
\end{aligned}$$

for all $u \in C(B_\tau \times B_\sigma)$ and all $j \in \{1, \ldots, 2d\}$.

In order to be able to find approximation results, i.e., to apply Theorem 3.15, we have to find constants C_u, γ_u satisfying (3.19). Due to the asymptotic smoothness (1.2), we have

$$\|\partial_j^\nu \kappa\|_{\infty, B_\tau \times B_\sigma} \leq C_{\mathrm{asymp}} \nu! \frac{c_0^\nu}{\mathrm{dist}(B_\tau, B_\sigma)^{g+\nu}} = C_u \gamma_u^\nu \nu!$$

for all $j \in \{1, \ldots, 2d\}$ and $\nu \in \mathbb{N}_0$, where

$$C_u := \frac{C_{\mathrm{asymp}}}{\mathrm{dist}(B_\tau, B_\sigma)^g}, \quad \gamma_u := \frac{c_0}{\mathrm{dist}(B_\tau, B_\sigma)}.$$

Since $\tau \times \sigma$ is admissible, we have $\mathrm{dist}(B_\tau, B_\sigma) \geq \mathrm{diam}(B_\tau \times B_\sigma)/(2\eta)$, which implies

$$\gamma_u |J_j^\tau| \leq \frac{2\eta c_0 |J_j^\tau|}{\mathrm{diam}(B_\tau \times B_\sigma)} \leq \frac{2\eta c_0 |J_j^\tau|}{|J_j^\tau|} = 2\eta c_0,$$

$$\gamma_u |J_j^\sigma| \leq \frac{2\eta c_0 |J_j^\sigma|}{\operatorname{diam}(B_\tau \times B_\sigma)} \leq \frac{2\eta c_0 |J_j^\sigma|}{|J_j^\sigma|} = 2\eta c_0$$

for all $j \in \{1, \ldots, d\}$. We introduce

$$\begin{aligned} C_{\mathrm{ap}}(\alpha, \beta, \tau) := 8eC_\Lambda^2 C_s C_{\mathrm{asymp}}(1 + 2\eta c_0)(\ell_{\max} - \operatorname{level}(\tau) + 1) \\ \times (\beta + 1 + \alpha(\ell_{\max} - \operatorname{level}(\tau)))^{2\lambda+1} \end{aligned}$$

and find for C_{1d} of Theorem 3.15

$$C_{1d}(\alpha, k^{\tau_0}, \operatorname{level}(\tau_0) - \operatorname{level}(\tau)) \leq \frac{C_{\mathrm{ap}}(\alpha, \beta, \tau)}{\operatorname{dist}(B_\tau, B_\sigma)^g},$$

since $k^{\tau_0} = \beta + \alpha(\ell_{\max} - \operatorname{level}(\tau_0))$ holds. Therefore, Theorem 3.15 implies

$$\begin{aligned} &\left\| \left(\bigotimes_{i=1}^{j-1} \mathrm{Id} \otimes (\mathrm{Id} - \mathcal{I}_{\mathfrak{C},j}) \otimes \bigotimes_{i=j+1}^{2d} \mathrm{Id} \right) \kappa \right\|_{\infty, B_{\tau_0} \times B_{\sigma_0}} \\ &\quad \leq \frac{C_{\mathrm{ap}}(\alpha, \beta, \tau)}{\operatorname{dist}(B_\tau, B_\sigma)^g} \left(\frac{1}{1+\gamma} \right)^{\beta+1} \\ &\qquad \times \left(\frac{1}{1+\gamma} \right)^{\alpha(\ell_{\max} - \operatorname{level}(\tau_0)) + \min\{(\beta+1)\bar{\alpha}, \alpha\}(\operatorname{level}(\tau_0) - \operatorname{level}(\tau))} \\ &\quad \leq \frac{C_{\mathrm{ap}}(\alpha, \beta, \tau)}{\operatorname{dist}(B_\tau, B_\sigma)^g} \left(\frac{1}{1+\gamma} \right)^{\beta+1} \left(\frac{1}{1+\gamma} \right)^{\min\{(\beta+1)\bar{\alpha}, \alpha\}(\ell_{\max} - \operatorname{level}(\tau))} \end{aligned} \tag{4.5}$$

for all $j \in \{1, \ldots, 2d\}$, where

$$\gamma = \min\left\{ \frac{1}{\eta c_0}, \frac{1}{\bar{q}} \right\} \quad \text{and} \quad \bar{\alpha} = \frac{C_q}{1 + c_0 \eta \bar{q}}.$$

In order to symmetrize the estimate (4.5), we observe

$$\ell_{\max} - \operatorname{level}(\tau) \geq \ell_{\max} - \ell - \delta/2,$$

which implies

$$\begin{aligned} &\left(\frac{1}{1+\gamma} \right)^{\min\{(\beta+1)\bar{\alpha}, \alpha\}(\ell_{\max} - \operatorname{level}(\tau))} \\ &\leq \left(\frac{1}{1+\gamma} \right)^{\min\{(\beta+1)\bar{\alpha}, \alpha\}(\ell_{\max} - \ell - \delta/2)} \end{aligned} \tag{4.6}$$

and

$$\begin{aligned} C_{\mathrm{ap}}(\alpha, \beta, \tau) &\leq 8eC_\Lambda^2 C_s C_{\mathrm{asymp}}(1 + 2\eta c_0)(\ell_{\max} - \ell + 1 + \delta/2) \\ &\quad \times (\beta + 1 + \alpha(\ell_{\max} - \ell + \delta/2))^{2\lambda+1} =: \tilde{C}_{\mathrm{ap}}(\alpha, \beta, \ell_{\max} - \ell). \end{aligned} \tag{4.7}$$

Combining (4.5) with (4.6) yields

$$\begin{aligned}(4.8)\quad &\left\| \left(\bigotimes_{i=1}^{j-1} \mathrm{Id} \otimes (\mathrm{Id} - \mathcal{I}_{\mathfrak{C},j}) \otimes \bigotimes_{i=j+1}^{2d} \mathrm{Id} \right) \kappa \right\|_{\infty, B_{\tau_0} \times B_{\sigma_0}} \\ &\leq \frac{\tilde{C}_{\mathrm{ap}}(\alpha, \beta, \ell_{\max} - \ell)}{\operatorname{dist}(B_\tau, B_\sigma)^g} \left(\frac{1}{1+\gamma} \right)^{\beta+1} \left(\frac{1}{1+\gamma} \right)^{\min\{(\beta+1)\bar{\alpha}, \alpha\}(\ell_{\max} - \ell - \delta/2)}\end{aligned}$$

for all $j \in \{1, \ldots, 2d\}$. Due to

$$\kappa - \mathcal{I}^{\tau \times \sigma}_{\tau_0 \times \sigma_0} \kappa = \sum_{j=1}^{2d} \left(\bigotimes_{i=1}^{j-1} \mathcal{I}_{\mathfrak{C},i} \right) \otimes (\mathrm{Id} - \mathcal{I}_{\mathfrak{C},j}) \otimes \left(\bigotimes_{i=j+1}^{2d} \mathrm{Id} \right) \kappa,$$

we can combine (4.4) and (4.9) to prove

$$\begin{aligned}&\| \kappa - \mathcal{I}^{\tau \times \sigma}_{\tau_0 \times \sigma_0} \kappa \|_{\infty, B_{\tau_0} \times B_{\sigma_0}} \\ &\quad \leq \frac{2d(\alpha\delta/2+1)^{\lambda(2d-1)} C_\Lambda^{2d-1} (1+\gamma)^{\delta/2} (\beta + 2 + \alpha(\ell_{\max} - \ell))^{\lambda(2d-1)}}{\operatorname{dist}(B_\tau, B_\sigma)^g} \\ &\quad \times C_s^{2d-1} \tilde{C}_{\mathrm{ap}}(\alpha, \beta, \ell_{\max} - \ell) \left(\frac{1}{1+\gamma} \right)^{\beta+1} \left(\frac{1}{1+\gamma} \right)^{\min\{(\beta+1)\bar{\alpha}, \alpha\}(\ell_{\max} - \ell)}\end{aligned}$$

by standard tensor arguments. We can see that this expression can be split into four factors: The (rather lengthy) stability term

$$\begin{aligned}\tilde{C}(\alpha, \beta, \ell_{\max} - \ell) &:= 2d C_s^{2d-1} (\alpha\delta/2+1)^{\lambda(2d-1)} C_\Lambda^{2d-1} (1+\gamma)^{\delta/2} \\ &\quad \times (\beta + 1 + \alpha(\ell_{\max} - \ell))^{\lambda(2d-1)} \tilde{C}_{\mathrm{ap}}(\alpha, \beta, \ell_{\max} - \ell),\end{aligned}$$

which depends only polynomially on β and $\ell_{\max} - \ell$, the singularity-related term $\operatorname{dist}(B_\tau, B_\sigma)^{-g}$, which does not depend on α, β or ℓ, the exponential term $(1+\gamma)^{-(\beta+1)}$, which decreases exponentially in β, and the level-dependent term

$$\left(\frac{1}{1+\gamma} \right)^{\min\{(\beta+1)\bar{\alpha}, \alpha\}(\ell_{\max} - \ell)},$$

which decreases exponentially in $\ell_{\max} - \ell$.

Let $\epsilon \in \mathbb{R}_{>0}$. Since $\bar{\alpha} > 0$, we can find α and β such that

$$\left(\frac{1}{1+\gamma} \right)^{\min\{(\beta+1)\bar{\alpha}, \alpha\}} \leq p^{1+\epsilon}$$

holds. The stability term $\tilde{C}$ depends only polynomially on $\ell_{\max} - \ell$, so we can find a function $C_1(\alpha, \beta)$ that depends only polynomially on β such that

$$\tilde{C}(\alpha, \beta, \ell_{\max} - \ell) p^{\epsilon(\ell_{\max} - \ell)} \leq C_1(\alpha, \beta)$$

holds for all values of $\ell_{\max} - \ell$. Therefore we can ensure

$$C_1(\alpha, \beta) \left(\frac{1}{1+\gamma} \right)^{\beta+1} \leq C_{\text{in}}$$

by choosing β sufficiently large. For these values of α and β, we have

$$\begin{aligned}
&\| \kappa - \mathcal{I}^{\tau \times \sigma}_{\tau_0 \times \sigma_0} \kappa \|_{\infty, B_{\tau_0} \times B_{\sigma_0}} \\
&\leq \frac{\tilde{C}(\alpha, \beta, \ell_{\max} - \ell)}{\operatorname{dist}(B_\tau, B_\sigma)^g} p^{(1+\epsilon)(\ell_{\max} - \ell)} \left(\frac{1}{1+\gamma} \right)^{\beta+1} \\
&\leq \frac{C_1(\alpha, \beta)}{\operatorname{dist}(B_\tau, B_\sigma)^g} p^{\ell_{\max} - \ell} \left(\frac{1}{1+\gamma} \right)^{\beta+1} \\
&\leq \frac{C_{\text{in}} p^{\ell_{\max} - \ell}}{\operatorname{dist}(B_\tau, B_\sigma)^g},
\end{aligned}$$

which is the desired result. □

It is important to note that we can reach *any* constant C_{in} and *any* decay rate $p \in]0, 1[$ by choosing α and β appropriately.

4.2 Global error estimate

Assumption 4.5 (Basis functions) *We assume that the basis functions are* local*, i.e., that there is a constant $C_{\text{ov}} \in \mathbb{R}_{>0}$ satisfying*

$$\sum_{i \in X} |\operatorname{supp}(\Psi_i)| \leq C_{\text{ov}} \left| \bigcup_{i \in X} \operatorname{supp}(\Psi_i) \right|$$

*for all $X \subseteq I$. We also assume that the finite element basis is L^2-*stable*, i.e., that there is a constant $C_{\text{fe}} \in \mathbb{R}_{>0}$ satisfying*

$$C_{\text{fe}}^{-1} \left\| \sum_{i \in I} \hat{u}_i \Psi_i \right\|_{L^2}^2 \leq \sum_{i \in I} \| \hat{u}_i \Psi_i \|_{L^2}^2 \leq C_{\text{fe}} \left\| \sum_{i \in I} \hat{u}_i \Psi_i \right\|_{L^2}^2$$

for all coefficient vectors $\hat{u} = (\hat{u}_i)_{i \in I} \in \mathbb{R}^I$.

The locality condition holds if each element of the triangulation is contained in the supports of not more than C_{ov} basis functions. For piecewise constant basis functions, we have $C_{\mathrm{ov}} = 1$, for piecewise linear basis functions on triangles, we have $C_{\mathrm{ov}} = 3$.

The L^2-stability condition implies a spectral equivalence between the mass matrix and its diagonal. It can be shown to hold for relatively general triangulations (cf. [6]).

For all $u \in V_N$, the vector $\hat{u} \in \mathbb{R}^I$ with

$$u = \sum_{i \in I} \hat{u}_i \Psi_i$$

is called the *coefficient vector* corresponding to u, and for all $\tau \in T_I$, we define the restriction

$$u_\tau := \sum_{i \in \tau} \hat{u}_i \Psi_i.$$

The approximated matrix $\tilde{\mathbf{K}}$ corresponds to the bilinear form

$$\tilde{a}(u,v) := \sum_{\tau\times\sigma\in P_{\mathrm{far}}} \sum_{\substack{\tau_0\times\sigma_0\\ \in \mathrm{sons}^*(\tau\times\sigma)}} \int_\Gamma u_{\tau_0}(x) \int_\Gamma v_{\sigma_0}(y) \mathcal{I}^{\tau\times\sigma}_{\tau_0\times\sigma_0}[\kappa](x,y)\,dy\,dx + \sum_{\tau\times\sigma\in P_{\mathrm{near}}} \int_\Gamma u_\tau(x) \int_\Gamma v_\sigma(y)\kappa(x,y)\,dy\,dx.$$

Theorem 4.6 (Global error estimate) *Let Assumptions 2.4, 4.2, 4.3 and 4.5 hold. Let $p_1 \in]0,1[$ and $C_{\mathrm{gl}} \in \mathbb{R}_{>0}$. Then there exist rank parameters $\alpha, \beta \in \mathbb{N}_0$ such that*

$$|a(u,v) - \tilde{a}(u,v)| \leq C_{\mathrm{gl}} \xi \|u\|_{L^2} \|v\|_{L^2}$$

holds for all $u, v \in V_N$ with

$$\xi := \max\left\{ \frac{|\Gamma_\tau| p_1^{\ell_{\max} - \mathrm{level}(\tau)}}{\mathrm{diam}(B_\tau)^g} \;:\; \tau \in T_I \right\}.$$

Proof Let $p_2 \in]0,1[$. We choose $\alpha, \beta \in \mathbb{N}_0$ as in Theorem 4.4 for $p := p_1 p_2$ and

$$C_{\mathrm{in}} := \frac{C_{\mathrm{gl}}(1-p_2)}{C_{\mathrm{sp}} C_{\mathrm{fe}}^2 C_{\mathrm{ov}} 2^{g/2} \eta^g}.$$

Let $u, v \in V_N$. Due to the definition of the kernel approximation, we have

$$a(u,v) - \tilde{a}(u,v) = \sum_{\tau\times\sigma\in P_{\mathrm{far}}} \sum_{\substack{\tau_0\times\sigma_0\\ \in \mathrm{sons}^*(\tau\times\sigma)}} \int_\Gamma u_{\tau_0}(x) \int_\Gamma v_{\sigma_0}(y)(\kappa - \mathcal{I}^{\tau\times\sigma}_{\tau_0\times\sigma_0}\kappa)(x,y)\,dy\,dx.$$

We will first consider the inner sum. For $\tau \times \sigma \in P_{\mathrm{far}}$ and $\ell := (\mathrm{level}(\tau) + \mathrm{level}(\sigma))/2$, we use the Cauchy-Schwarz inequality and Theorem 4.4 in order to get

$$
\begin{aligned}
(4.9)\qquad & \left| \sum_{\substack{\tau_0\times\sigma_0\\ \in \mathrm{sons}^*(\tau\times\sigma)}} \int_\Gamma u_{\tau_0}(x) \int_\Gamma v_{\sigma_0}(y)(\kappa - \mathcal{I}^{\tau\times\sigma}_{\tau_0\times\sigma_0}\kappa)(x,y)\,dy\,dx \right| \\
& \leq \sum_{\substack{\tau_0\times\sigma_0\\ \in \mathrm{sons}^*(\tau\times\sigma)}} \|u_{\tau_0}\|_{L^1} \|v_{\sigma_0}\|_{L^1} \|\kappa - \mathcal{I}^{\tau\times\sigma}_{\tau_0\times\sigma_0}\kappa\|_{\infty, B_{\tau_0}\times B_{\sigma_0}} \\
& \leq \sum_{\substack{\tau_0\times\sigma_0\\ \in \mathrm{sons}^*(\tau\times\sigma)}} |\Gamma_{\tau_0}|^{1/2} |\Gamma_{\sigma_0}|^{1/2} \|u_{\tau_0}\|_{L^2} \|v_{\sigma_0}\|_{L^2} \frac{C_{\mathrm{in}} p^{\ell_{\max}-\ell}}{\mathrm{dist}(B_\tau, B_\sigma)^g} \\
& \leq \frac{C_{\mathrm{in}} p^{\ell_{\max}-\ell}}{\mathrm{dist}(B_\tau, B_\sigma)^g} \left(\sum_{\substack{\tau_0\times\sigma_0\\ \in \mathrm{sons}^*(\tau\times\sigma)}} |\Gamma_{\tau_0}|\,|\Gamma_{\sigma_0}| \right)^{1/2} \\
& \quad \times \left(\sum_{\substack{\tau_0\times\sigma_0\\ \in \mathrm{sons}^*(\tau\times\sigma)}} \|u_{\tau_0}\|_{L^2}^2 \|v_{\sigma_0}\|_{L^2}^2 \right)^{1/2} \\
& \leq \frac{C_{\mathrm{ov}} C_{\mathrm{fe}} C_{\mathrm{in}} p^{\ell_{\max}-\ell}}{\mathrm{dist}(B_\tau, B_\sigma)^g} |\Gamma_\tau|^{1/2} |\Gamma_\sigma|^{1/2} \left(\sum_{i\in\tau} \|u_i \Psi_i\|_{L^2}^2 \sum_{j\in\sigma} \|v_j \Psi_j\|_{L^2}^2 \right)^{1/2}.
\end{aligned}
$$

Since $\tau \times \sigma \in P_{\mathrm{far}}$ holds, the admissibility condition (1.9) implies

$$
\frac{1}{\mathrm{dist}(B_\tau, B_\sigma)^g} \leq \frac{(2\eta)^g}{\mathrm{diam}(B_\tau \times B_\sigma)^g},
$$

and we can use

$$
\begin{aligned}
\mathrm{diam}(B_\tau \times B_\sigma) &= (\mathrm{diam}(B_\tau)^2 + \mathrm{diam}(B_\sigma)^2)^{1/2} \\
&\geq (2\,\mathrm{diam}(B_\tau)\,\mathrm{diam}(B_\sigma))^{1/2}
\end{aligned}
$$

in order to prove

$$
\frac{1}{\mathrm{dist}(B_\tau, B_\sigma)^g} \leq \frac{2^{g/2}\eta^g}{\mathrm{diam}(B_\tau)^{g/2}\,\mathrm{diam}(B_\sigma)^{g/2}},
$$

which implies

$$
\frac{p_1^{\ell_{\max}-\ell} |\Gamma_\tau|^{1/2} |\Gamma_\sigma|^{1/2}}{\mathrm{dist}(B_\tau, B_\sigma)^g} \leq 2^{g/2}\eta^g \xi.
$$

Now we can turn to the outer sum:

$$\begin{aligned}
&|a(u,v) - \tilde{a}(u,v)| \\
&\quad \leq C_{\mathrm{ov}} C_{\mathrm{fe}} C_{\mathrm{in}} 2^{g/2} \eta^g \xi \sum_{\tau\times\sigma\in P_{\mathrm{far}}} p_2^{(\ell_{\max}-\mathrm{level}(\tau))/2} p_2^{(\ell_{\max}-\mathrm{level}(\sigma))/2} \\
&\qquad \times \left(\sum_{i\in\tau} \|u_i \Psi_i\|_{L^2}^2 \right)^{1/2} \left(\sum_{j\in\sigma} \|v_j \Psi_j\|_{L^2}^2 \right)^{1/2} \\
&\quad \leq C_{\mathrm{ov}} C_{\mathrm{fe}} C_{\mathrm{in}} 2^{g/2} \eta^g \xi \left(\sum_{\tau\times\sigma\in P_{\mathrm{far}}} p_2^{\ell_{\max}-\mathrm{level}(\tau)} \sum_{i\in\tau} \|u_i \Psi_i\|_{L^2}^2 \right)^{1/2} \\
&\qquad \times \left(\sum_{\tau\times\sigma\in P_{\mathrm{far}}} p_2^{\ell_{\max}-\mathrm{level}(\sigma)} \sum_{j\in\sigma} \|v_j \Psi_j\|_{L^2}^2 \right)^{1/2} \\
&\quad \leq C_{\mathrm{ov}} C_{\mathrm{fe}} C_{\mathrm{sp}} C_{\mathrm{in}} 2^{g/2} \eta^g \xi \left(\sum_{\tau\in T_I} p_2^{\ell_{\max}-\mathrm{level}(\tau)} \sum_{i\in\tau} \|u_i \Psi_i\|_{L^2}^2 \right)^{1/2} \\
&\qquad \times \left(\sum_{\sigma\in T_I} p_2^{\ell_{\max}-\mathrm{level}(\sigma)} \sum_{j\in\sigma} \|v_j \Psi_j\|_{L^2}^2 \right)^{1/2}.
\end{aligned}$$

Since $p_2 \in]0,1[$ holds, we find

$$\begin{aligned}
&\sum_{\tau\in T_I} p_2^{\ell_{\max}-\mathrm{level}(\tau)} \sum_{i\in\tau} \|u_i \Psi_i\|_{L^2}^2 \\
&\quad = \sum_{\ell=0}^{\ell_{\max}} p_2^{\ell_{\max}-\ell} \sum_{\tau\in T_{I,\ell}} \sum_{i\in\tau} \|u_i \Psi_i\|_{L^2}^2 \\
&\quad \leq \sum_{\ell=0}^{\ell_{\max}} p_2^{\ell_{\max}-\ell} \sum_{i\in I} \|u_i \Psi_i\|_{L^2}^2 \\
&\quad \leq C_{\mathrm{fe}} \|u\|_{L^2}^2 \sum_{\ell=0}^{\ell_{\max}} p_2^{\ell_{\max}-\ell} \leq \frac{C_{\mathrm{fe}}}{1-p_2} \|u\|_{L^2}^2
\end{aligned}$$

and can conclude

$$\begin{aligned}
|a(u,v) - \tilde{a}(u,v)| &\leq \frac{C_{\mathrm{sp}} C_{\mathrm{ov}} C_{\mathrm{fe}}^2 C_{\mathrm{in}}}{1-p_2} 2^{g/2} \eta^g \xi \|u\|_{L^2} \|v\|_{L^2} \\
&= C_{\mathrm{gl}} \xi \|u\|_{L^2} \|v\|_{L^2},
\end{aligned}$$

which is the desired result. □

In order to be able to bound ξ in Theorem 4.6 we need additional assumptions:

Assumption 4.7 *We assume that the support of the clusters grows regularly, i.e., that there is a constant $C_{\mathrm{gr}} \in \mathbb{R}_{>0}$ satifying*

$$|\Gamma_\tau| \leq C_{\mathrm{gr}} |\Gamma_{\tau'}|$$

for all $\tau \in T_I$ and $\tau' \in \operatorname{sons}(\tau)$.

In a sense, this assumption corresponds to a balancing of the clusters: Whenever a cluster τ is split into its sons, we have to ensure that their supports are of similar size.

Assumption 4.8 *Let d_Γ be the dimension of the submanifold Γ. We assume that the finite element mesh is quasi-uniform with grid parameter $h \in \mathbb{R}_{>0}$ and that the leaves of the cluster tree are small enough, i.e., that there is a constant $C_{\mathrm{lf}} \in \mathbb{R}_{>0}$ such that*

$$|\Gamma_\tau| \leq C_{\mathrm{lf}}^{d_\Gamma} h^{d_\Gamma} \quad \text{and} \quad \operatorname{diam}(B_\tau) \geq C_{\mathrm{lf}}^{-1} h$$

holds for all leaves $\tau \in L_I$.

This assumption can be satisfied by choosing the correct stopping criterion for the clustering algorithm: If we keep splitting clusters as long as possible, leaf clusters will contain only one finite element, i.e., $|\Gamma_\tau| \sim h^{d_\Gamma}$, and optimal bounding boxes will satisfy $\operatorname{diam}(B_\tau) = \operatorname{diam}(\Gamma_\tau)$.

Corollary 4.9 *Let Assumptions 2.4, 4.2, 4.3, 4.5, 4.7 and 4.8 hold. Let $C_{\mathrm{apx}} \in \mathbb{R}_{>0}$. Then there are rank parameters $\alpha, \beta \in \mathbb{N}_0$ such that*

$$|a(u,v) - \tilde{a}(u,v)| \leq C_{\mathrm{apx}} h^{d_\Gamma - g} \|u\|_{L^2} \|v\|_{L^2}$$

holds for all $u, v \in V_N$.

Proof Let $\tau \in T_I$. Let $\tau^* \in \operatorname{sons}^*(\tau)$. Due to Assumptions 4.7 and 4.2, we have

$$\frac{|\Gamma_\tau|}{\operatorname{diam}(B_\tau)^g} \leq C_{\mathrm{gr}}^{\operatorname{level}(\tau^*) - \operatorname{level}(\tau)} \frac{|\Gamma_{\tau^*}|}{\operatorname{diam}(B_{\tau^*})^g}$$

and can use Assumption 4.8 to find

$$\frac{|\Gamma_\tau|}{\operatorname{diam}(B_\tau)^g} \leq C_{\mathrm{lf}}^{d_\Gamma + g} h^{d_\Gamma - g} C_{\mathrm{gr}}^{\operatorname{level}(\tau^*) - \operatorname{level}(\tau)}.$$

This suggests that $p_1 := C_{\mathrm{gr}}^{-1}$ is the correct choice, since it guarantees

$$\begin{aligned} \frac{|\Gamma_\tau| p_1^{\ell_{\max} - \operatorname{level}(\tau)}}{\operatorname{diam}(B_\tau)^g} &\leq C_{\mathrm{lf}}^{d_\Gamma + g} h^{d_\Gamma - g} C_{\mathrm{gr}}^{\operatorname{level}(\tau^*) - \operatorname{level}(\tau)} p_1^{\ell_{\max} - \operatorname{level}(\tau)} \\ &\leq C_{\mathrm{lf}}^{d_\Gamma + g} h^{d_\Gamma - g}, \end{aligned}$$

for all $\tau \in T_I$, i.e. $\xi \leq C_{\mathrm{lf}}^{d_\Gamma + g} h^{d_\Gamma - g}$. Applying Theorem 4.6 to $C_{\mathrm{gl}} := C_{\mathrm{apx}} / C_{\mathrm{lf}}^{d_\Gamma + g}$ proves the desired result. □

4.3 Approximation with separable weights

In many practical applications the kernel function κ is not asymptotically smooth (cf. (1.2)), e.g., if it consists of a globally defined singular part and a factor that only exists on the submanifold.

Let us consider the kernel function

$$\kappa_\Gamma(x, y) := \kappa(x, y)\omega(x, y) \tag{4.10}$$

consisting of an asymptotically smooth kernel function κ and a separable function

$$\omega : \Gamma \times \Gamma \to \mathbb{R}.$$

We cannot interpolate κ_Γ directly, since it is only defined on Γ. Therefore we interpolate only κ (cf. Subsection 1.3) and multiply the interpolant by ω to define the approximation

$$\tilde{\kappa}_{\Gamma,\tau\times\sigma}(x, y) := \tilde{\kappa}_{\tau\times\sigma}(x, y)\omega(x, y). \tag{4.11}$$

Since ω is separable, it can be written in the form

$$\omega(x, y) = \sum_{\iota=1}^{s} \omega_\iota^1(x)\omega_\iota^2(y),$$

and the approximation $\tilde{\kappa}_{\Gamma,\tau\times\sigma}$ takes the form

$$\tilde{\kappa}_{\Gamma,\tau\times\sigma}(x, y) = \sum_{\iota=1}^{s} \sum_{\nu\in K_\tau} \sum_{\mu\in K_\sigma} \kappa(x_\nu^\tau, x_\mu^\sigma)\omega_\iota^1(x)\mathcal{L}_\nu^\tau(x)\omega_\iota^2(y)\mathcal{L}_\mu^\sigma(y).$$

We can treat the additional factors by replacing the standard Lagrange polynomials in (1.10) by the functions $\omega_\iota^1\mathcal{L}_\nu^\tau$ and $\omega_\iota^2\mathcal{L}_\mu^\sigma$: For each $\iota \in \{1, \ldots, s\}$, the matrices $\mathbf{V}^\tau$ and $\mathbf{W}^\sigma$ are replaced by

$$\mathbf{V}_{i\nu}^{\tau,\iota} := \int_\Gamma \omega_\iota^1(x)\mathcal{L}_\nu^\tau(x)\Psi_i(x)\,dx, \quad \mathbf{W}_{j\mu}^{\sigma,\iota} := \int_\Gamma \omega_\iota^2(y)\mathcal{L}_\mu^\sigma(y)\Psi_j(y)\,dy,$$

and the coefficient matrices $\mathbf{S}^{\tau,\sigma}$ and the transfer matrices $\mathbf{B}^{\tau',\tau}$ remain unchanged.

If the function ω satisfies

$$|\omega(x, y)| \leq C_{\mathrm{wg}}\|x - y\|^r$$

for $r \in \mathbb{R}_{>0}$, this property can be used to mitigate the singularity of κ: Let $\tau \times \sigma \in P_{\mathrm{far}}$. Due to admissibility, we have

$$\max\{\operatorname{diam}(B_\tau), \operatorname{diam}(B_\sigma)\} \leq 2\eta \operatorname{dist}(B_\tau, B_\sigma),$$

and this implies

$$\|x - y\| \leq \operatorname{diam}(B_\tau) + \operatorname{dist}(B_\tau, B_\sigma) + \operatorname{diam}(B_\sigma) \leq (1 + 4\eta) \operatorname{dist}(B_\tau, B_\sigma)$$

for all $x \in B_\tau$, $y \in B_\sigma$, i.e.,

$$\|\omega\|_{\infty, B_\tau \times B_\sigma} \leq C_{\mathrm{wg}} (1 + 4\eta)^r \operatorname{dist}(B_\tau, B_\sigma)^r.$$

We can use this estimate in step (4.9) of the proof of Theorem 4.6 to replace the singularity order g of κ by the reduced singularity order $g - r$ of κ_Γ and get the following result:

Corollary 4.10 *Let Assumptions 2.4, 4.2, 4.3, 4.5, 4.7 and 4.8 hold. Let $C_{\mathrm{apx}} \in \mathbb{R}_{>0}$. Then there are rank parameters $\alpha, \beta \in \mathbb{N}_0$ such that*

$$|a(u, v) - \tilde{a}(u, v)| \leq C_{\mathrm{apx}} h^{d_\Gamma + r - g} \|u\|_{L^2} \|v\|_{L^2}$$

holds for all $u, v \in V_N$, where in the definition of a and $\tilde{a}$ the functions κ and $\tilde{\kappa}_{\tau \times \sigma}$ are replaced with κ_Γ and $\tilde{\kappa}_{\Gamma, \tau \times \sigma}$ (cf. (4.10)*,* (4.11)*).*

*Example 4.11 (*Double layer potential, $d = 3$*)* The classical double layer potential has the form

$$\kappa_\Gamma(x, y) := \frac{1}{4\pi} \frac{\langle n_y, x - y \rangle}{\|x - y\|^3}$$

for $x, y \in \Gamma$. We set

$$\kappa(x, y) := \frac{1}{4\pi} \frac{1}{\|x - y\|^3} \quad \text{and} \quad \omega(x, y) := \langle n_y, x - y \rangle$$

and observe that κ is asymptotically smooth with $g = 3$, while ω can be written as

$$\omega(x, y) = \sum_{i=1}^{d} n_{y,i} x_i - \langle n_y, y \rangle,$$

i.e., is separable with $s = d + 1$.

On smooth curves or surfaces Γ, there is a constant $C_{\mathrm{wg}} \in \mathbb{R}_{>0}$ such that

$$|\omega(x, y)| \leq C_{\mathrm{wg}} \|x - y\|^2$$

for all $x, y \in \Gamma$, so the effective singularity order of κ_Γ is only 1, and for each $C_{\mathrm{apx}} \in \mathbb{R}_{>0}$ we can find parameters $\alpha, \beta \in \mathbb{N}_0$ such that

$$|a(u, v) - \tilde{a}(u, v)| \leq C_{\mathrm{apx}} C_{\mathrm{wg}} (1 + 4\eta)^2 h \|u\|_{L^2} \|v\|_{L^2} \tag{4.12}$$

holds for all $u, v \in V_N$. The estimate (4.12) can also be obtained in the two-dimensional case. For non-smooth surfaces, it is possible to prove a similar estimate for a modified rank distribution [20].

Remark 4.12 The approximation result Corollary 4.10 can be applied, for example, to the three-dimensional double layer potential. It is known that the double layer potential $\mathcal{K} : L^2(\Gamma) \to L^2(\Gamma)$ is compact for smooth surfaces Γ. Then, if for $\lambda \neq 0$ the operator $\mathcal{B}_\lambda := \lambda \mathrm{Id} + \mathcal{K}$ is injective, it is in fact bijective and there exist $h_0, \gamma > 0$ such that for $h \leq h_0$ the discretized operator defined on V_N is L^2-stable, i.e.,

$$\gamma \leq \inf_{0 \neq u \in V_N} \sup_{0 \neq V_N} \frac{\langle \mathcal{B}_\lambda u, v \rangle}{\|u\|_{L^2(\Gamma)} \|v\|_{L^2(\Gamma)}}. \tag{4.13}$$

This inf-sup condition also guarantees quasi-optimality of the Galerkin BEM for solving the integral equation $\mathcal{B}_\lambda u = f \in L^2(\Gamma)$. Replacing the discretized version of $\mathcal{K}$ with the $\mathcal{H}^2$-matrix approximation $\widetilde{\mathcal{K}}$ introduces additional errors. Recalling from Example 4.11 that $d_\Gamma + r - g = 1$, we can infer from (4.13) that the perturbed system is still stable. An application of the Strang lemma then allows us to infer for the total approximation error

$$\|u - u_N\|_{L^2(\Gamma)} \leq C \inf_{v \in V_V} \|u - v\|_{L^2(\Gamma)} + Ch.$$

We remark that the storage requirement for the $\mathcal{H}^2$-matrix $\widetilde{\mathbf{K}}_N$ is $O(N)$; the work to compute the near field entries and to compute the leaf contributions (see the discussion following Lemma 2.12) to sufficient accuracy could entail additional logarithmic factors.

5 Implementational issues and numerical examples

In the preceding section we presented a convergence analysis for the approximation of integral operators by a variable order interpolation procedure. The purpose of the present section is twofold: first, we discuss the way in which the approximation orders are distributed in our implementation since it differs slightly from the procedure analyzed in the preceding section; second, we illustrate the performance of our algorithm by applying it to the two-dimensional and three-dimensional double layer potential.

5.1 Implementation

Our algorithmic realization is based on the following construction of the cluster tree and the distribution of the approximation orders:

The cluster tree is organized as a binary tree; the bounding boxes of the two son clusters are obtained by splitting the longest extent of the father bounding box and afterwards shrinking the son bounding boxes as much as possible.

While this construction is straightforward and easily implemented, the resulting clusters usually do not satisfy the Assumption 4.2. Therefore it is necessary to replace the simple rank distribution (4.2) by a more sophisticated anisotropic construction. Algorithm 5.1 below assigns ranks to the clusters of the cluster tree in a bottom up fashion. The leaves are assigned the (isotropic) degree $\beta \in \mathbb{N}_0$. One key requirement of the rank distribution is that the iterated interpolation procedure be stable. While going up in the cluster tree this stability is ensured by increasing the approximation order only if the extents of the bounding boxes of father and son clusters differ significantly. The parameter $\bar{q} \in]0, 1[$ quantifies the notion of "significant" and the increase of the degree is controlled by a parameter $\alpha \in \mathbb{N}$. Since anisotropic polynomial degree distributions are sought, the above considerations are done for each coordinate $j \in \{1, \ldots, d\}$ separately.

Algorithm 5.1 (Degree distribution)
`procedure` *DegreeDistribution*(τ)*;*

```
begin
  if sons(τ) = ∅ then for j ∈ {1, ..., d} do k_j^τ := β

  else begin
      for τ' ∈ sons(τ) do begin

        DegreeDistribution(τ');
        for j ∈ {1, ..., d} do begin
          q := |J_j^τ'|/|J_j^τ|;
          if q ≤ q̄ then k̃_j^τ' := k_j^τ' + α⌊log_2(q̄/q)⌋
              else k̃_j^τ' := k_j^τ'
        end
      end;
      for j ∈ {1, ..., d} do k_j^τ := max{k̃_j^τ' | τ' ∈ sons(τ)}
  end
end
```

The approximation order of the clusters is anisotropic (i.e., different polynomial degrees for the different directions), which is particularly useful in conjunction with highly anisotropic bounding boxes: if parts of the geometry are flat, the bounding boxes may degenerate to (effectively) $(d-1)$-dimensional bounding boxes and the approximation order in the degenerate direction may be kept low.

An extensive investigation of different rank distribution schemes can be found in [19].

5.2 Numerical examples

We illustrate the performance of our variable order interpolation scheme by applying it to the classical double layer potential in two and three dimensions. In both examples, the trial spaces V_N are taken as the space of piecewise constant functions on quasi-uniform meshes. The interpolation operator $\mathcal{I}_k$ employed is the Chebyshev interpolation, which satisfies the assumptions set out in (3.2). Our numerical experiments were performed on a SUN SunFire 6800.

In view of (4.12) we expect convergence $\mathcal{O}(h)$ for suitable choices of the rank distribution in the operator norm (i.e., the matrix is viewed as an operator $V_N \to L^2(\Gamma)$, where the space V_N is endowed with the $L^2(\Gamma)$-norm); from the results of Section 2 we expect $\mathcal{O}(N)$ complexity for storage, time for a matrix-vector multiplication, and time to set up the far field.

Example 5.2 We consider the double layer potential on a two-dimensional ellipse with semi axes $a = 2$ and $b = 1$. The block partition is generated with $\eta = 2.5$, and we chose $\beta = 0$, $\alpha = 1$, and $\bar{q} = 0.6$ as parameters for the order distribution. The results of the numerical example are given in Table 1. In the column "error" of Table 1 we present the spectral norm of the error matrix $\mathbf{K} - \tilde{\mathbf{K}}$. Since the spectral norm $\|\mathbf{K}\|_2$ of the exact matrix scales like $\mathcal{O}(h)$, the optimal convergence $\mathcal{O}(h)$ of (4.12) translates into an $\mathcal{O}(h^2)$ behavior for the spectral norm of the error matrix, which is indeed visible. These errors were computed by comparing the $\mathcal{H}^2$-matrix for $\beta = 0$ with that for $\beta = 2$; a comparison with the exact matrix, which is possible for values of $N \leq 32768$, showed the approximation corresponding to $\beta = 2$ to be sufficiently accurate to estimate the error.

The columns "mem/N", "build/N" and "MVM/N" give the memory requirements (in MB), the time to set up the far field (in seconds) and the time for a matrix-vector multiplication (in seconds, only the far field part) divided by the number N of degrees of freedom. These quantities are all bounded, which fits our complexity estimates.

For comparison purposes, the last column contains the number of near-field entries of the matrix divided by N, which is bounded, implying that the memory requirement and the time to the perform the near-field contribution of the matrix-vector multiplication can also be performed with complexity $\mathcal{O}(N)$.

Example 5.3 This numerical experiment is the three-dimensional analog of the preceding Example 5.2. We consider the double layer potential on an ellipsoid with semi axes of lengths $a = 3$, $b = 2$, $c = 1$; the numerical results are reported in Table 2. As in Example 5.2, the block partition is generated with $\eta = 2.5$ and the approximation order distribution is done with parameters $\beta = 0$, $\alpha = 1$, $\bar{q} = 0.6$. The column "error" in Table 2 contains

Table 1. Two-dimensional double layer potential on ellipse (cf. Example 5.2)

N	error	rate	mem/N	build/N	MVM/N	near
32	2.75_{-2}	—	2.25_{-3}	—	—	6
64	4.63_{-3}	0.17	3.03_{-3}	1.56_{-4}	5.68_{-5}	6
128	1.12_{-3}	0.24	4.01_{-3}	7.81_{-5}	6.39_{-5}	6
256	3.35_{-4}	0.30	5.03_{-3}	1.56_{-4}	9.23_{-5}	6
512	1.00_{-4}	0.30	5.96_{-3}	1.56_{-4}	1.17_{-4}	6
1024	2.84_{-5}	0.28	6.76_{-3}	1.76_{-4}	1.29_{-4}	6
2048	7.68_{-6}	0.27	7.39_{-3}	1.86_{-4}	1.39_{-4}	6
4096	2.02_{-6}	0.26	7.86_{-3}	2.03_{-4}	1.48_{-4}	6
8192	5.22_{-7}	0.26	8.20_{-3}	2.14_{-4}	1.56_{-4}	6
16384	1.33_{-7}	0.26	8.44_{-3}	2.19_{-4}	1.62_{-4}	6
32768	3.38_{-8}	0.25	8.61_{-3}	2.24_{-4}	1.67_{-4}	6
65536	8.51_{-9}	0.25	8.72_{-3}	2.28_{-4}	1.71_{-4}	6
131072	2.14_{-9}	0.25	8.79_{-3}	2.29_{-4}	1.74_{-4}	6
262144	5.36_{-10}	0.25	8.84_{-3}	2.30_{-4}	1.78_{-4}	6

Table 2. Three-dimensional double layer potential on ellipsoid (cf. Example 5.3)

N	error	rate	mem/N	build/N	MVM/N	near
512	2.28_{-3}	—	2.12_{-4}	5.86_{-5}	5.86_{-6}	320
2048	1.44_{-3}	0.63	6.00_{-4}	7.81_{-5}	7.81_{-5}	312
8192	2.95_{-4}	0.20	1.16_{-3}	9.28_{-5}	1.44_{-5}	316
32768	5.07_{-5}	0.17	1.80_{-3}	1.14_{-4}	2.15_{-5}	323
131072	7.13_{-6}	0.14	2.16_{-3}	1.33_{-4}	2.52_{-5}	319
524288	1.00_{-6}	0.14	2.80_{-3}	1.90_{-4}	3.07_{-5}	322
2097152	—	—	2.67_{-3}	2.55_{-4}	2.98_{-5}	321
8388608	—	—	2.67_{-3}	2.72_{-4}	—	321

again the spectral norm of the error matrix; since the spectral norm of the exact matrix scales like $\mathcal{O}(h^2)$, the expected relative error $\mathcal{O}(h)$ results in an expected convergence $\mathcal{O}(h^3)$, i.e., a rate 0.125, which is close to the observed 0.14. The error was evaluated by comparing the $\mathcal{H}^2$-matrix for $\beta = 0$ with that for $\beta = 2$.

References

1. Amini, S., Profit, A.T.J.: Multi-level fast multipole solution of the scattering problem. Engineering Analysis with Boundary Elements. **27**, 547–564 (2003)
2. Börm, S., Grasedyck L., Hackbusch, W.: Introduction to hierarchical matrices with applications. Engineering Analysis with Boundary Elements. **27**, 405–422 (2003)
3. Börm, S., Hackbusch, W.: $\mathcal{H}^2$-matrix approximation of integral operators by interpolation. Applied Numerical Mathematics. **43**, 129–143 (2002)

4. Börm, S., Löhndorf, M., Melenk, J.M.: Approximation of integral operators by variable-order interpolation. Technical Report 82, Max Planck Institute for Mathematics in the Sciences, 2002. Extended preprint.
5. Brandt A., Lubrecht, A.A.: Multilevel matrix multiplication and fast solution of integral equations. J. Comput. Phys. **90**, 348–370 1990
6. Dahmen, W., Faermann, B., Graham, I.G., Hackbusch, W., Sauter, S.A.: Inverse inequalities on non-quasiuniform meshes and applications to the mortar element method. Math. Comp. **73**, 1107–1138 (2004)
7. Dahmen, W., Schneider, R.: Wavelets on manifolds I: Construction and domain decomposition. SIAM Journal of Mathematical Analysis. **31**, 184–230 (1999)
8. DeVore, R.A., Lorentz, G.G.: Constructive Approximation. Springer-Verlag, 1993
9. Giebermann, K.: Multilevel approximation of boundary integral operators. Computing. **67**, 183–207 (2001)
10. Grasedyck, L.: Theorie und Anwendungen Hierarchischer Matrizen. PhD thesis, Universität Kiel, 2001
11. Grasedyck, L., Hackbusch, W.: Construction and arithmetics of $\mathcal{H}$-matrices. Computing. **70**(4), 295–334 (2003)
12. Greengard L., Rokhlin V.: A fast algorithm for particle simulations. Journal of Computational Physics. **73**, 325–348 (1987)
13. Hackbusch, W.: A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices. Computing. **62**, 89–108 (1999)
14. Hackbusch, W., Khoromskij, B.: A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part II: Application to multi-dimensional problems. Computing. **64**, 21–47 (2000)
15. Hackbusch, W., Khoromskij, B., Sauter, S.: On $\mathcal{H}^2$-matrices. In H. Bungartz, R. Hoppe, and C. Zenger, editors, Lectures on Applied Mathematics. Springer-Verlag, Berlin, 2000 pp. 9–29
16. Hackbusch, W., Nowak, Z.P.: On the fast matrix multiplication in the boundary element method by panel clustering. Numerische Mathematik. **54**, 463–491 (1989)
17. Krezebek, N., Sauter, S.: Fast cluster techniques for BEM. Technical Report 04, Institut für Mathematik, University of Zurich, 2002
18. Lage, C., Schmidlin, G., Schwab, C.: Rapid solution of first kind boundary integral equations in $\mathbb{R}^3$. Engineering Analysis with Boundary Elements. **27**, 469–490 (2003)
19. Löhndorf, M.: Effiziente Behandlung von Integraloperatoren mit $\mathcal{H}^2$-Matrizen variabler Ordnung. PhD thesis, Universität Leipzig, 2003
20. Löhndorf, M., Melenk, J.M.: Approximation of integral operators by variable-order approximation. part II: Non-smooth domains. In preparation.
21. Rivlin, T.J.: The Chebyshev Polynomials. Wiley-Interscience, New York, 1984
22. Sauter, S.: Variable order panel clustering (extended version). Technical Report 52, Max-Planck-Institut für Mathematik, Leipzig, Germany, 1999
23. Sauter, S.: Variable order panel clustering. Computing. **64**, 223–261 (2000)
24. Schmidlin, G.: Fast Solution Algorithm for Integral Equations in $\mathbb{R}^3$. PhD thesis, ETH Zürich, 2003
25. Tausch, J.: A variable order wavelet method for the sparse representation of layer potentials in the non-standard form. Journal of Numerical Mathematics. **12** (3), 233–254 (2004)
26. Tyrtyshnikov, E.: Mosaic-skeleton approximation. Calcolo. **33**, 47–57 (1996)