Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Blended kernel approximation in the $\mathcal{H}$-matrix techniques

by

Wolfgang Hackbusch and Boris N. Khoromskij

Preprint no.: 66 2000

Blended Kernel Approximation in the $\mathcal{H}$-Matrix Techniques

Wolfgang Hackbusch and Boris N. Khoromskij
Max-Planck-Institut *Mathematik in den Naturwissenschaften*,
Inselstr. 22-26, D-04103 Leipzig, Germany

Abstract

Several types of $\mathcal{H}$-matrices were shown to provide a data-sparse approximation of nonlocal (integral) operators in FEM and BEM applications [5]-[10]. The general construction is applied to the operators with asymptotically smooth kernel function (which is not necessary given explicitly) provided that the Galerkin ansatz space has a hierarchical structure.

The new class of $\mathcal{H}$-matrices is based on so-called *blended FE and polynomial approximations* of the kernel function and leads to matrix blocks with a tensor-product of block-Toeplitz (block-circulant) and rank-k matrices. This implies the translation (rotation) invariance of the kernel combined with the corresponding tensor-product grids. The approach is devoted to the fast evaluation of volume/boundary integral operators with possibly non-smooth kernels defined on canonical[1] domains/manifolds in the FEM/BEM applications. In particular, we provide the error and complexity analysis for blended expansions to the Helmholtz kernel.

AMS Subject Classification: 65F05, 65F50, 65F30

1 Introduction

The task of a data-sparse approximation to nonlocal (integral) operators is closely related to a number of important problems in applied mathematics, physics, astrophysics, quantum chemistry and biology. In particular, for the efficient evaluation of particle interactions, of integral operators in BEM/FEM as well as Green's functions in elliptic problems, we are interested in the sparse approximation of arising integral mappings on spatial domains or manifolds in $\mathbb{R}^d$, $d = 2, 3$.

The standard $N \times N$ matrix arithmetic (e.g., inversion of sparse matrices in FEM and evaluation of the fully populated stiffness matrices in BEM) has the algebraic complexity of $O(N^2) - O(N^3)$ operations. Several approaches have been accomplished towards creating a fast *matrix-vector multiplication* $\mathbf{A} * \mathbf{x}$, $\mathbf{x} \in \mathbb{R}^N$, where the dense matrix $\mathbf{A}$ approximates an integral operator with asymptotically smooth kernel $s(x, y)$: panel clustering [11], wavelet approximation [1], multipole-based methods [3], mosaic-skeleton approximation [14], use of block Toeplitz/circulant matrices [13, 15, 16], etc. A new concept for the efficient implementation of rather general matrix operations is based on the $\mathcal{H}$-matrix techniques [5], which allows a fast and accurate matrix-vector and matrix-matrix product, computing the inverse and the matrix exponential etc., with linear-logarithmic cost for rather general applications [5]-[10].

The data-sparsity of general $\mathcal{H}$-matrices is determined by the hierarchical block partitioning P_2 of the corresponding product index-set on the one hand, and by the particular choice of matrix-blocks on the other hand. A systematic approach for optimising the data-sparsity of $\mathcal{H}$-matrices is based on imposing the *structural properties* of (a) the kernel function $s(x, y)$, (b) the approximation ansatz space and (c) the block-partitioning of $\mathcal{H}$-matrix itself. In this paper, we shall simplify the general situation and consider only the mutual influence of structures specified in topics (a) and (b), i.e., we assume that the hierarchical block decomposition P_2 of a matrix is fixed. Therefore, the partitioning P_2 will only effect our global error and complexity estimates due to the summation over the cluster tree. We shall distinguish the following structural properties of the kernel:

(a1) asymptotical smoothness or analyticity of the kernel function $s(x, y)$, $x \neq y$,
(a2) $\mathcal{L}$-harmonicity of $s(x, y)$, i.e., $\mathcal{L}_x s(x, y) = \mathcal{L}_y s(x, y) = 0$ for $x \neq y$,
(a3) translation invariance $s(x, y) = S(x - y)$,

[1] Here and in the following, we call domains *canonical* if they are obtained by translation or rotation of one of their parts (e.g., parallelepiped, cylinder, sphere, etc.)

and of the Galerkin ansatz spaces

(b1) piece-wise constant/linear elements on quasi-uniform or locally refined grids,

(b2) piece-wise constant/linear elements on tensor-product grids.

Using *Taylor or other polynomial interpolants* of the kernel $s(x,y)$, we arrive at expansions with a number of terms equal to $k = O(\log^d N)$, which are related to the spectral FE approximation of the kernel on each admissible product domain. However, for special classes of discretisations/kernels, a further improvement of data-sparsity is possible on the base of the so-called *wire-basket* and/or *blended FE and polynomial* expansions of the kernels which leads to the complexity $O(N)$ and $O(N \log N)$ for the memory requirements and for the matrix-vector multiplication, respectively.

The *wire-basket expansions* [9] are based on the reduction of domain integrals to the product of corresponding boundaries due to the basic property $\mathcal{L}S(z) = \delta_0(z)$, $z \neq 0$, fulfilled by the fundamental solution of the underlying elliptic operator $\mathcal{L}$. This method reduces the block rank down to $k = O(\log^{d-1} N)$. The memory requirement for the corresponding $n_b \times n_b$ matrix block is then estimated by $O(\log^{d-1} N n_b^{1-1/d} + n_b)$ (see [9]). The matrix-vector product costs $O(N \log N)$ operations and the memory requirements are $O(N)$. It is worth noting that the wire-basket approximations provide an extension of the familiar *multipole expansions* to general elliptic operators with constant coefficients.

The *blended FE and polynomial* approximations to be investigated in this paper provide new opportunities for the robust treatment of problems with, e.g., oscillatory kernels. The analysis of such $\mathcal{H}$-matrix techniques including the case of kernels arising from wave problems (e.g., Helmholtz equation) will be presented in §2 and §4. They may be applied to volume integral calculation within the wire-basket expansions as well as in BEM (or coupling of FEM and BEM) on special surfaces.

Further essential reduction of the computational complexity is based on the concept of variable order expansions [10] in the spirit of the *hp* version of the FE approximation to the kernel. This modification provides asymptotically optimal consistency error estimates in the L^2 setting and leads to linear complexity. The analysis for the case of volume integral operators will be given in §3.

2 Description of the Matrix Classes

2.1 $\mathcal{H}$-Matrices in FEM and BEM: Brief Survey

Suppose we are given a second order elliptic operator

$$\mathcal{L} := -\sum_{j,k=1}^{d} \partial_j a_{jk} \partial_k + \sum_{j=1}^{d} b_j \partial_j + c_0 \qquad (\partial_j := \frac{\partial}{\partial x_j}) \tag{2.1}$$

with real constant coefficients a_{jk}, b_j and c_0. In BEM applications, we are interested in solving the homogeneous equation $\mathcal{L}u = 0$ in Ω with certain boundary conditions by using the equivalent boundary integral equation on $\Sigma = \partial\Omega$. We consider the h-version of the Galerkin FE method for approximating the continuous integral operators $A : V \to V'$ defined in the Sobolev space $V = H^r(\Sigma)$,

$$(Au)(x) = \int_\Sigma s(x,y)u(y)dy \qquad (x \in \Sigma), \tag{2.2}$$

with $s(x,y) = S(x-y)$, where S is the fundamental solution satisfying $\mathcal{L}S = \delta_0$, or with s replaced by a suitable directional derivatives Ds of s. In FEM applications, as well as in calculations of particle interactions, we are interested in approximating volume integral operators or the corresponding discrete sums.

We distinguish two different cases:

(A) Σ is a bounded $(d-1)$-dimensional manifold (BEM applications);

(B) Σ is a polyhedron in $\mathbb{R}^d$, $d = 2, 3$ (FEM applications).

In BEM techniques, we are interested in the fast solving of the variational boundary integral equation

$$\langle(\lambda I + A)u, v\rangle = \langle f, v\rangle \qquad \text{for all } v \in V := H^r(\Sigma),\ r \leq 1, \tag{2.3}$$

where $\lambda \in \mathbb{R}$ is a given parameter. To calculate particular solutions of elliptic PDEs, we are looking for a fast computation of volume integrals like in (2.2) with $\Sigma = \Omega$ (say with $A : L^2(\Omega) \to H^2(\Omega)$, $\Omega \in \mathbb{R}^d$).

Similar problems arise in the evaluation of particle interactions. A number of applications in FEM is related to the accurate data-sparse approximation of the inverse operator $A := \mathcal{L}^{-1}$ (say with $V = H^{-1}(\Omega),\ V' = H_0^1(\Omega)$, $\Omega \in \mathbb{R}^d$) based on its classical integral representation,

$$\mathcal{L}^{-1}u = \int_\Omega G(x,y)u(y)dy \qquad (u \in L^2(\Omega)), \tag{2.4}$$

where the Green's function $G(x,y)$ solves the equation

$$\begin{array}{rcll} \mathcal{L}G(x,y) & = & \delta(x-y) & \text{for } x,y \in \Omega, \\ G(x,y) & = & 0 & \text{for } x \in \partial\Omega,\ y \in \Omega. \end{array} \tag{2.5}$$

Together with an adjoint system of equations in the second variable y, equation (2.5) allows to prove the smoothness of $G(x,y)$ which implies the existence of the $\mathcal{H}$-matrix approximation of $\mathcal{L}^{-1}$. The latter can be computed using the *formatted matrix arithmetic* described in [5, 6].

The construction of $\mathcal{H}$-matrices is based on the low rank representation of the matrix blocks associated with an admissible block partitioning. Given the Galerkin ansatz space $V_h \subset V$, let I be the index set of unknowns (e.g., the FE-nodal points) and $T(I)$ be the hierarchical cluster tree [5]. For each $i \in I$, the support of the corresponding basis function φ_i is denoted by $X(i) := \mathrm{supp}(\varphi_i)$ and for each cluster $\tau \in T(I)$ we define its geometric image by $X(\tau) = \bigcup_{i\in\tau} X(i)$.

In a canonical way (cf. [6]), a block-cluster tree $T(I \times I)$ can be constructed from $T(I)$, where all vertices $b \in T(I\times I)$ are of the form $b = \tau \times \sigma$ with $\tau, \sigma \in T(I)$. Given a matrix $M \in \mathbb{R}^{I\times I}$, the block-matrix corresponding to $b \in T(I \times I)$ is denoted by $M^b = (m_{ij})_{(i,j)\in b}$. An *admissible block partitioning* P_2 of $I \times I$ is a set of disjoint blocks $b \in T(I\times I)$, satisfying the admissibility condition,

$$\min\{\mathrm{diam}\, X(\sigma), \mathrm{diam}\, X(\tau)\} \le 2\eta\ \ \mathrm{dist}(X(\sigma), X(\tau)), \tag{2.6}$$

$(\sigma,\tau) \in P_2$, $\eta < 1$, whose union equals $I \times I$. Let a block partitioning P_2 and $k_b \ll N = \dim V_h$ be given for each $b \in P_2$. The set of real $\mathcal{H}$-matrices induced by P_2 and k_b is

$$\mathcal{M}_{\mathcal{H},k}(I\times I, P_2) := \{M \in \mathbb{R}^{I\times I} : \text{for all } b \in P_2 \text{ there holds } \mathrm{rank}(M^b) \le k_b\}. \tag{2.7}$$

The linear-logarithmic complexity $O(N \log^q N)$ of $\mathcal{H}$-matrix arithmetic was proven in [5], [6] and [7]. Let $\mathbf{A} := \langle A\varphi_i, \varphi_j\rangle_{i,j=1}^N$ be the exact Galerkin matrix. With the splitting $P_2 = P_{far} \cup P_{near}$, where $P_{far} := \{\sigma\times\tau \in P_2 : \mathrm{dist}(X(\tau), X(\sigma)) > 0\}$, the reduction in operation count is due to the replacement of the full matrix blocks $\mathbf{A}^{\tau\times\sigma}$, $\tau\times\sigma \in P_{far}$, by their low-rank approximations

$$\mathbf{A}_{\mathcal{H}}^{\tau\times\sigma} := \sum_{\alpha=1}^{k} \mathbf{a}_\alpha^\tau \cdot (\mathbf{c}_\alpha^\sigma)^\top, \qquad \mathbf{a}_\alpha^\tau := \{\langle a_\alpha(x), \varphi_j\rangle\}_{j\in\tau} \in \mathbb{R}^\tau,\ \mathbf{c}_\alpha^\sigma := \{\langle c_\alpha(y), \varphi_j\rangle\}_{j\in\sigma} \in \mathbb{R}^\sigma,$$

where the coefficient vectors $\mathbf{a}_\alpha^\tau$, $\mathbf{c}_\alpha^\sigma$ are obtained from a *separable expansion*

$$s_{\tau,\sigma}(x,y) = \sum_{\alpha=1}^{k} a_\alpha(x)c_\alpha(y) \tag{2.8}$$

with $k \ll N = \dim V_h$ approximating the kernel $s(x,y)$. This implies the following storage and matrix-vector multiplication cost

$$\mathcal{N}_{st}\left(\mathbf{A}_{\mathcal{H}}^{\tau\times\sigma}\right) = k(\#\tau + \#\sigma), \qquad \mathcal{N}_{MV}\left(\mathbf{A}_{\mathcal{H}}^{\tau\times\sigma}\right) = 2k(\#\tau + \#\sigma), \tag{2.9}$$

where $k = O(\log^\beta N)$ with $\beta = d$ or $\beta = d-1$.

Similar to the construction of *uniform $\mathcal{H}$-matrices* in [5], our blended expansions defined below are based on the representation of $\#\tau \times \#\sigma$ rank-k matrix blocks spanned by a *fixed* basis.

Let $n = \#\tau = \#\sigma$ for the moment. With given linear spaces $\mathcal{V}_a := \mathrm{span}\{\mathbf{a}_i\}_{i=1}^k$ and $\mathcal{V}_c := \mathrm{span}\{\mathbf{c}_j\}_{j=1}^k$, where $\mathbf{a}_i, \mathbf{c}_j \in \mathbb{R}^n$, any $M \in \mathcal{R}_{k,n} := \mathcal{V}_a \otimes \mathcal{V}_c = \mathrm{span}\{(\mathbf{a}_i \cdot \mathbf{c}_j^\intercal)\}$ has the representation

$$M = \sum_{i,j=1}^{k} \zeta_{ij}(\mathbf{a}_i \cdot \mathbf{c}_j^\intercal) \qquad (\zeta_{ij} \in \mathbb{R}). \tag{2.10}$$

In the context of uniform $\mathcal{H}$-matrices, the construction (2.10) will be applied in the next section to the family $\{A_\alpha\}$ of hierarchical matrices, where each matrix block $A_\alpha^{\tau\times\sigma}$, $\tau\times\sigma \in P_{far}$, has the form (2.10) and is represented by the coefficients $\zeta_{ij} = \zeta_{\alpha,ij}^{\tau\times\sigma}$.

2.2 Block-Toeplitz and Block-circulant Matrices

In this section we recall the standard definitions of block-Toeplitz and block-circulant matrices. The standard *Toeplitz matrix* $M = \{t_{ij}\}_{i,j=1}^n$ is defined by the fact that the entries t_{ij} depend only on $i-j$. Using $t_{i-j} := t_{ij}$, we introduce the notation

$$M = \{t_{i-j}\}_{i,j=1}^n =: Toepl\{t_{-n+1}, \dots, t_0, \dots t_{n-1}\} \in \mathbb{R}^{n\times n}$$

We say that the $mn \times mn$ matrix $M = \{T_{ij}\}_{i,j=1}^n$ with $m\times m$ blocks T_{ij} has a *two-level $n\times n$ block-Toeplitz structure*, if $T_{ij} = T_{i-j}$ depends on $i-j$ only. For the block-Toeplitz matrix we use the notation

$$M = BToepl\{T_{-n+1}, \dots, T_0, \dots, T_{n-1}\} \quad \text{with } T_q \in \mathbb{R}^{m\times m} \quad \text{for } q = -n+1, \dots, 0, \dots, n-1,$$

In the following definition, we introduce the class $\mathcal{M}_{T_{\mathbf{n}},q}$ ($\mathbf{n} \in \mathbb{N}^q$) of *block-Toeplitz matrices* with q-level block structure.

Definition 2.1 *For $q \in \mathbb{N}$ we define the matrix class $\mathcal{M}_{T_{\mathbf{n}},q}$ recursively. For $q = 1$, $M = \{t_{ij}\}_{i,j=1}^n \in \mathcal{M}_{T_n,1}$ denotes the standard Toeplitz matrix, i.e., $M = Toepl\{t_{-n+1}, \dots, t_0, \dots t_{n-1}\} \in \mathbb{R}^{n\times n}$.*
Given $q \in \mathbb{N}$, $q \geq 2$, and $\mathbf{n} = (n_1, \dots, n_q) \in \mathbb{N}^q$, a matrix M belongs to $\mathcal{M}_{T_{\mathbf{n}},q} \subset \mathbb{R}^{|\mathbf{n}|\times|\mathbf{n}|}$ if

$$M = BToepl\{T_{-n_1+1}, \dots, T_0, \dots, T_{n_1-1}\} \qquad \text{with } T_j \in \mathcal{M}_{T_{\mathbf{n}'},q-1} \text{ for } |j| \leq n_1 - 1,$$

where $\mathbf{n}' := (n_2, \dots, n_q) \in \mathbb{N}^{q-1}$ and $|\mathbf{n}| := n_1 n_2 \cdots n_q$.

The favourable feature of the matrix class $\mathcal{M}_{T_{\mathbf{n}},q}$ is due to the following well-known property.

Proposition 2.2 *For any $M \in \mathcal{M}_{T_{\mathbf{n}},q}$, the matrix-vector multiplication and the memory requirements have the complexity*

$$\mathcal{N}_{MV}(M) = c_{FFT} 2^{q+1} |\mathbf{n}| \log |\mathbf{n}|, \qquad \mathcal{N}_{st}(M) = 2^q |\mathbf{n}|,$$

where c_{FFT} is the constant characterising the FFT cost. In the symmetric case ($T_i = T_{-i}^\top$), the storage reduces to $\mathcal{N}_{st}(M) = |\mathbf{n}|$.

We recall the definition of circulant matrices.

Definition 2.3 *An $n\times n$ matrix $\mathcal{C}$ is called* circulant *if it has the representation*

$$\mathcal{C} = circ(c_1, \dots, c_n) := \begin{pmatrix} c_1 & c_2 & \dots & c_n \\ c_n & c_1 & \dots & c_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_2 & \dots & c_n & c_1 \end{pmatrix}, \qquad c_i \in \mathbb{C}.$$

The set of all $n \times n$ circulant matrices is closed with respect to addition and multiplication. Any circulant matrix $\mathcal{C}$ is associated with the polynomial $p_c(z) := c_1 + c_2 + \dots + c_n z^{n-1}$ ($z \in \mathbb{C}$) and has a diagonal representation in the Fourier basis,

$$\mathcal{C} = F_n^T \Lambda_c F_n \quad \text{with } \Lambda_c = \operatorname{diag}\{p_c(1), \dots, p_c(\omega^{n-1})\}, \qquad \omega = e^{i\pi/n}.$$

The eigenvector corresponding to the eigenvalue $p_c(\omega^{j-1})$ is given by jth column of F_n, i.e.,

$$\vec{\omega}_j = \frac{1}{\sqrt{n}} (\omega^{(k-1)(j-1)})_{k=1,\dots,n}.$$

The matrix class $\mathcal{M}_{C_{\mathbf{n}},q}$ of q-level block-circulant matrices is introduced similarly to $\mathcal{M}_{T_{\mathbf{n}},q}$ in Definition 2.1.

Matrices from $\mathcal{M}_{T_{\mathbf{n}},q}$ arise from the FE approximation of integral operators on uniform tensor product grids with translation invariant kernels $s(x,y) = S(x-y)$. Similarly, the circulant variant $\mathcal{M}_{C_{\mathbf{n}},q}$ arises from approximations by means of uniform tensor product grids in cylindrical coordinates, provided that the kernel is rotationally invariant.

The attractive idea is to combine matrices from $\mathcal{M}_{T_{\mathbf{n}},q}$ (or $\mathcal{M}_{C_{\mathbf{n}},q}$) and the uniform $\mathcal{H}$-matrices (in particular, $\mathcal{R}_{k,n}$-matrices). This combination will be the subject of the next subsection.

2.3 $\mathcal{H}$-Matrices of Blended Type

Blended approximations of the matrix blocks $\mathbf{A}^{\tau\times\sigma}$ are of a mixed type in the sense that they are based on a tensor product of rank-k matrices and Toeplitz (or circulant) matrices. In what follows we use the *tensor product* $A \otimes B \in \mathbb{R}^{nm\times nm}$ of two matrices $A \in \mathbb{R}^{n\times n}$ and $B = \{b_{ij}\} \in \mathbb{R}^{m\times m}$ defined by $A \otimes B := \{B_{ij}\}_{i,j=1}^{m}$ with block matrices $B_{ij} := b_{ij}A$. Simple examples of blended formats are

$$A \otimes B \quad \text{and} \quad B \otimes A \qquad (A \in \mathcal{R}_{k,n},\ B \in \mathcal{M}_{T_m,1}). \tag{2.11}$$

Matrices of the type $A \otimes B$ and $B \otimes A$ are interesting objects since they allow basic matrix operations of almost linear complexity in the sense that

$$\mathcal{N}_{st}(A \otimes B) = \mathcal{N}_{st}(B \otimes A) = O(kn + m), \qquad \mathcal{N}_{MV}(A \otimes B) = \mathcal{N}_{MV}(B \otimes A) = O(knm\log m). \tag{2.12}$$

Moreover, in the following we use rank-k matrices with *fixed* basis which lead to sublinear memory needs and also enable a faster matrix-vector multiplication. The following definition generalises the class of rank-k matrices spanned by a fixed basis (see (2.10)).

Definition 2.4 *Given $q \in \mathbb{N}$, $\mathbf{m} \in \mathbb{N}^q$, $k, n \in \mathbb{N}$, and $\mathcal{R}_{k,n} = \operatorname{span}\{(\mathbf{a}_i \cdot \mathbf{c}_j^\top)\}_{i,j=1}^{k}$ with $\mathbf{a}_i, \mathbf{c}_j \in \mathbb{R}^n$, a matrix $M \in \mathbb{R}^{n|\mathbf{m}|\times n|\mathbf{m}|}$ belongs to $\mathcal{M}_{T_\mathbf{m}\otimes\mathcal{R}_{k,n}}$ if*

$$M = \sum_{i,j=1}^{k} T_{ij} \otimes (\mathbf{a}_i \cdot \mathbf{c}_j^\top), \qquad T_{ij} \in \mathcal{M}_{T_\mathbf{m},q}.$$

A matrix M belongs to $\mathcal{M}_{\mathcal{R}_{k,n}\otimes T_\mathbf{m}}$ if

$$M = \sum_{i,j=1}^{k} (\mathbf{a}_i \cdot \mathbf{c}_j^\top) \otimes T_{ij}, \qquad T_{ij} \in \mathcal{M}_{T_\mathbf{m},q}.$$

The following statement proves the complexity of the above defined matrix classes. First we note that the vector-spaces $\mathcal{M}_{T_\mathbf{m}\otimes\mathcal{R}_{k,n}}$ and $\mathcal{M}_{\mathcal{R}_{k,n}\otimes T_\mathbf{m}}$ are isomorphic. If $A \otimes B \in \mathcal{M}_{T_\mathbf{m}\otimes\mathcal{R}_{k,n}}$, there is a permutation matrix[2] Π such that $\Pi \cdot (A \otimes B) \cdot \Pi^\top = B \otimes A \in \mathcal{M}_{\mathcal{R}_{k,n}\otimes T_\mathbf{m}}$. Hence, the matrix-vector costs $\mathcal{N}_{MV}(M)$ are the same for $M \in \mathcal{M}_{T_\mathbf{m}\otimes\mathcal{R}_{k,n}}$ and $M \in \mathcal{M}_{\mathcal{R}_{k,n}\otimes T_\mathbf{m}}$. The same holds for the storage $\mathcal{N}_{st}(M)$.

Lemma 2.5 *For any $M \in \mathcal{M}_{T_\mathbf{m}\otimes\mathcal{R}_{k,n}}$ or $M \in \mathcal{M}_{\mathcal{R}_{k,n}\otimes T_\mathbf{m}}$ with $\mathbf{m} \in \mathbb{N}^q$ there holds*

$$\mathcal{N}_{st}(M) = 2^q k^2 |\mathbf{m}| + 2kn, \qquad \mathcal{N}_{MV}(M) = 4kn|\mathbf{m}| + 4c_{FFT}k^2|\mathbf{m}|\log|\mathbf{m}| + k^2|\mathbf{m}|. \tag{2.13}$$

The summand $2kn$ in $\mathcal{N}_{st}(M)$ is due to the storage of the vectors $\mathbf{a}_i, \mathbf{c}_j$. Since these vectors are fixed, they can be stored once for all and any further matrix needs only a storage of $\mathcal{N}_{st}(M) = 2^q k^2|\mathbf{m}|$.

Proof. The cost of $2kn$ for storing $\mathbf{a}_i, \mathbf{c}_j \in \mathbb{R}^n$ is already mentioned. The k^2 matrices T_{ij} require a storage of $2^q k^2|\mathbf{m}|$ due to Proposition 2.2.

For the proof of $\mathcal{N}_{MV}(M)$ we describe the multiplication algorithm $\mathbf{x} \mapsto \mathbf{y} = M * \mathbf{x}$ in detail, where $M \in \mathcal{M}_{\mathcal{R}_{k,n}\otimes T_\mathbf{m}}$. We employ the block structure $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_{|\mathbf{m}|} \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_{|\mathbf{m}|} \end{pmatrix}$ with $x_\alpha, y_\alpha \in \mathbb{R}^n$ for $\alpha = 1, \ldots, |\mathbf{m}|$.

In **Step 1**, we compute the vector $\mathbf{g}^j = (\langle \mathbf{c}_j, x_\alpha \rangle)_{\alpha=1,\ldots,|\mathbf{m}|} \in \mathbb{R}^{|\mathbf{m}|}$ of scalar products. Since j varies in $\{1, \ldots, k\}$, the total cost of Step 1 is $2n|\mathbf{m}|k$.

In **Step 2**, the q-level block-Toeplitz matrices $T_{ij} \in \mathcal{M}_{T_\mathbf{m},q}$ are applied to $\mathbf{g}^j$: $\mathbf{h}^i := \sum_{j=1}^{|\mathbf{m}|} T_{ij}\mathbf{g}^j$ $(1 \le i \le k)$. Due to Proposition 2.2, the cost amounts to $4c_{FFT}k^2|\mathbf{m}|\log|\mathbf{m}| + k^2|\mathbf{m}|$, where the latter term corresponds to the summation.

In **Step 3**, the resulting vector $\mathbf{y} = M * \mathbf{x}$ is computed by means of $y_\alpha := \sum_{i=1}^{k} \mathbf{h}_\alpha^i * \mathbf{a}_i$ $(\alpha = 1, \ldots, |\mathbf{m}|)$, which requires $2n|\mathbf{m}|k$ operations. Summing over Steps 1-3, we obtain the result for $\mathcal{N}_{MV}(M)$. ∎

[2] Π is defined by $(\Pi x)_i = x_{\pi(i)}$ for $1 \le i \le n|\mathbf{m}|$, where π is the permutation $\pi : i = \alpha + (\beta - 1)n \mapsto \beta + (\alpha - 1)|\mathbf{m}|$ $(1 \le \alpha \le n,\ 1 \le \beta \le |\mathbf{m}|)$.

Matrices from $\mathcal{M}_{\mathcal{R}_{k,n}\otimes T_{\mathbf{m}}}$ may be applied for the FE approximation of kernels in the 3D BEM (see Example 2.12 below).

Now we define blended matrix formats based on block-circulant matrices. The index set I is assumed to have the product form $I = I_c \times I_0$, where $I_c = \{1, \dots m\}$ is responsible for the circulant part. We introduce the hierarchical tree $T(I_0)$ of depth L. As in [6] we introduce the level subsets $T(\ell) := \{\tau \in T(I_0) : \tau$ belongs to level $\ell\}$. Also the partitioning P_2 splits into the level sets $P_2^{(\ell)} := \{\tau \times \sigma \in P_2 : \tau, \sigma \in T(\ell)\}$. We recall that $\ell = 0$ corresponds to the biggest cluster $I_0 \times I_0$ (root of the tree), while $\ell = L$ corresponds to the leaves (1×1 blocks). For each block $b = \tau \times \sigma \in P_2^{(\ell)}$, the corresponding matrix-block belongs to the vector space $\mathcal{R}(k, \tau \times \sigma) := \mathcal{V}_\tau \otimes \mathcal{V}_\sigma = \operatorname{span}\{\mathbf{a}_i^\tau \cdot (\mathbf{c}_j^\sigma)^\top\}_{i,j=1}^{k_\ell}$, where the rank k_ℓ depends on the level only. As in [5], we denote this vector space of uniform $\mathcal{H}$-matrices by $\mathcal{U}_{\mathcal{H},k}(I_0 \times I_0, P_2, \mathcal{R})$.

Definition 2.6 *Let $I = I_0 \times I_c$, where $I_c = \{1, \dots m\}$, $n := \#I_0$. For a given mapping $k : \{0, \dots, L\} \to \mathbb{N}$, the set $\mathcal{U}_{\mathcal{H},k}(I_0 \times I_0, P_2, \mathcal{R})$ of uniform $\mathcal{H}$-matrices is defined as described above. Then a matrix $M \in \mathbb{R}^{nm \times nm}$ belongs to $\mathcal{M}_{\mathcal{U}_{\mathcal{H},k}\otimes C_m}$ if*

$$M = Bcirc\{A_1, \dots A_m\} \qquad \text{for certain } A_p \in \mathcal{U}_{\mathcal{H},k} \subset \mathbb{R}^{n\times n},\ p \in I_c.$$

Now we propose the algorithm for fast matrix-vector multiplication with matrices from $\mathcal{M}_{\mathcal{U}_{\mathcal{H},k}\otimes C_m}$ based on the simultaneous use of circulant and $\mathcal{H}$-matrix structures.

Algorithm 2.7 *Given the matrices $A_1, \dots, A_m \in \mathcal{U}_{\mathcal{H},k}$, with blocks specified by*

$$A_p^{\tau\times\sigma} = \sum\nolimits_{i,j=1}^{k_\ell} \zeta_{p,ij}^{\tau\times\sigma}\, \mathbf{a}_i^\tau \cdot (\mathbf{c}_j^\sigma)^\top \qquad (p = 1, \dots, m),$$

and the vector $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, *we have to compute* $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} := M\mathbf{x}$, *where M is the block-circulant matrix given by Definition 2.6. All vector blocks x_α, y_α belong to $\mathbb{R}^{I_0}$, while $\mathbf{x}, \mathbf{y} \in \mathbb{R}^I = \mathbb{R}^{I_c \times I_0}$.*

Step 1. *Compute the set of vectors formed by scalar products*

$$\mathbf{x}_j^\sigma := \left(\langle \mathbf{c}_j^\sigma, x_{1\sigma}\rangle, \dots, \langle \mathbf{c}_j^\sigma, x_{m\sigma}\rangle\right)^\top \in \mathbb{R}^m \qquad \text{for all } 0 \le \ell \le L,\ \sigma \in T(\ell),\ 1 \le j \le k_\ell,$$

where $x_{\alpha\tau}$ denotes the block vectors $(x_{\alpha,i})_{i\in\tau}$.

Step 2. *(a) Multiply the $m \times m$ circulant matrices $circ\{\zeta_{1,ij}^{\tau\times\sigma}, \dots, \zeta_{m,ij}^{\tau\times\sigma}\}$ by the vectors $\mathbf{x}_j^\sigma$,*

$$\mathbf{z}_{ij}^{\tau\times\sigma} := circ\{\zeta_{1,ij}^{\tau\times\sigma}, \dots, \zeta_{m,ij}^{\tau\times\sigma}\} \cdot \mathbf{x}_j^\sigma \in \mathbb{R}^m \qquad \text{for all } 0 \le \ell \le L,\ \tau\times\sigma \in P_2^{(\ell)},\ 1 \le i, j \le k_\ell,$$

and (b) form the sums

$$\mathbf{z}_i^\tau := \sum_{\sigma:\tau\times\sigma\in P_2} \sum_{j=1}^{k_\ell} \mathbf{z}_{ij}^{\tau\times\sigma} \in \mathbb{R}^m \qquad \text{for all } 0 \le \ell \le L,\ \tau \in T(\ell),\ 1 \le i \le k_\ell.$$

Step 3. *The summation of the intermediate results*

$$y''_{\beta\tau} := \sum_{i=1}^{k_\ell} \mathbf{a}_i^\tau \cdot (\mathbf{z}_i^\tau)_\beta \in \mathbb{R}^\tau \qquad \text{for all } 0 \le \ell \le L,\ \tau \in T(\ell),\ \beta = 1, \dots, m,$$

starts at the leaves of the tree $T(I_0)$, where $y'_{\beta\tau} := y''_{\beta\tau}$. Then

$$y'_{\beta\tau} := y''_{\beta\tau} + (y'_{\beta\tau'})_{\tau' \text{ son of } \tau} \in \mathbb{R}^\tau \qquad \text{for all } \tau \in T(I_0),\ \beta = 1, \dots, m,$$

ends at the root $I_0 \in T(I_0)$, where $y_\beta = y'_{\beta I_0}$ $(\beta = 1, \dots, m)$ represents the final result.

The following statement establishes the complexity of Algorithm 2.7. Note that $N = nm$ is the dimension of the problem.

Lemma 2.8 *Under the assumptions made in [6] on the construction of the partitioning, $\#P_2^{(\ell)} \le c_P n 2^{\ell-L}$ holds for a certain constant c_P. Furthermore, we assume $L = \log n$ (with $n = \#I_0$), which holds for a uniform tree $T(I_0)$; otherwise, further constants must be inserted. Concerning k_ℓ we assume $k_L \le k_{L-1} \le \ldots \le k_0$. Then for any $M \in \mathcal{M}_{\mathcal{U}_{\mathcal{H},k}\otimes C_m}$ the storage requirements and the vector-matrix multiplication cost are*

$$\mathcal{N}_{st}(M) \le n \left[2k_0 \log n + c_P m \sum_{\ell=0}^{L} k_\ell^2 2^{\ell-L} \right], \tag{2.14}$$

$$\mathcal{N}_{MV}(M) \le nm \left[(1 + 4k_0 \log n) + c_P c_{FFT} (1 + \log m) \sum_{\ell=0}^{L} k_\ell^2 2^{\ell-L} \right]. \tag{2.15}$$

An interesting choice of k_ℓ is $k_\ell = k_L + (L-\ell)\delta$ ($\delta \ge 0$). Then $\sum_{\ell=0}^{L} k_\ell^2 2^{\ell-L} = O(k_L^2)$ depends on the smallest value k_L, not on the maximal one k_0.

Proof. The storage for vectors $\{\mathbf{a}_i^\tau, \mathbf{c}_j^\sigma\}$ for all $\tau, \sigma \in T(I_0)$ is estimated by

$$2 \sum_{\ell=0}^{L} k_\ell \sum_{\tau \in T(\ell)} \#\tau = 2n \sum_{\ell=0}^{L} k_\ell \le 2k_0 n \log n,$$

where we use that $L = \log n$ in the regular case, while a further constant appears in more general cases. The storage of the coefficients $\{\zeta_{p,ij}^{\tau\times\sigma}\}$ for all $1 \le p \le m$, $\tau \times \sigma \in P_2^\ell$, $0 \le \ell \le L$, is bounded by

$$m \sum_{\ell=0}^{L} k_\ell^2 \# P_2^\ell = c_P n m \sum_{\ell=0}^{L} k_\ell^2 2^{\ell-L}.$$

This proves (2.14).

The cost of Step 1 is $2m \sum_{\ell=0}^{L} k_\ell \sum_{\sigma \in T(\ell)} \#\sigma$ operations, since one scalar product needs $2\#\sigma$ operations. As $\sum_{\sigma\in T(\ell)} \#\sigma = n$, we arrive at $2nm \sum_{\ell=0}^{L} k_\ell \le 2k_0 nm \log n$.

One multiplication by the circulant matrix in Step 2a costs $c_{FFT}\, m \log m$. Hence, the resulting costs are $c_{FFT}\, m \log m \sum_{\ell=0}^{L} k_\ell^2 \# P_2^{(\ell)} = c_P c_{FFT} nm \log m \sum_{\ell=0}^{L} k_\ell^2 2^{\ell-L}$ operations. The summation in Step 2b requires $m \sum_{\ell=0}^{L} k_\ell^2 \# P_2^{(\ell)} \le c_P nm \sum_{\ell=0}^{L} k_\ell^2 2^{\ell-L}$. Thus, the total cost of Step 2 is $c_P nm(1+\log m) \sum_{\ell=0}^{L} k_\ell^2 2^{\ell-L}$

In Step 3, the computation of $y''_{\beta\tau}$ needs $m \sum_{\ell=0}^{L} 2k_\ell \sum_{\tau\in T(\ell)} \#\tau = 2nm \sum_{\ell=0}^{L} k_\ell \le 2nmk_0 \log n$ operations. The summation over the tree involves $m(\#T(I_0) - n) < nm$ additions. Together, we have $nm(1 + 2k_0 \log n)$ operations. These partial costs add up to (2.15). ∎

Corollary 2.9 *In addition to the situation from above, assume that the data are uniform in the sense that the subsets $\tau, \tau' \in T(\ell)$ differ only by a shift. Furthermore, the vectors $\mathbf{a}_i^\tau, \mathbf{a}_i^{\tau'}$ are equal for $\tau, \tau' \in T(\ell)$ and, similarly, $\mathbf{c}_i^\sigma = \mathbf{c}_i^{\sigma'}$ for $\sigma, \sigma' \in T(\ell)$. Then the storage requirements are reduced to*

$$\mathcal{N}_{st}(M) \le 2k_0 \log n + c_P nm \sum_{\ell=0}^{L} k_\ell^2 2^{\ell-L}.$$

The family of matrices $\mathcal{M}_{\mathcal{U}_{\mathcal{H},k}\otimes T_m}$ is defined and analysed in a similar way. We shall use the following definition.

Definition 2.10 *Let $I = I_0 \times I_c$, where $I_c = \{1, \ldots m\}$, $n := \#I_0$. For a given mapping $k : \{0, \ldots, L\} \to \mathbb{N}$, the set $\mathcal{U}_{\mathcal{H},k}(I_0 \times I_0, P_2, \mathcal{R})$ of uniform $\mathcal{H}$-matrices is defined as above. Then a matrix $M \in \mathbb{R}^{I\times I}$ belongs to $\mathcal{M}_{\mathcal{U}_{\mathcal{H},k}\otimes T_m}$ if*

$$M = BToepl\{T_{-m+1}, \ldots, T_0, \ldots, T_{m-1}\} \qquad \text{for certain } T_p \in \mathcal{U}_{\mathcal{H},k} \subset \mathbb{R}^{n\times n},\ p = -m+1, ..., m-1.$$

The new class of $\mathcal{H}$-matrices under consideration provides a reduction of the computational rank of the matrix blocks in the case of oscillatory kernels in 3D. Among with the standard BEM, the construction may be directly applied in coupled FEM-BEM techniques, where the boundary integral operator is defined on a coupling (auxiliary) surface (e.g., the cylinder, parallelepiped or their combinations). Typical applications include scattering problems with a bounded obstacle in $\mathbb{R}^3$. In the case of Helmholtz' kernel, this leads to a method with linear-logarithmic memory requirements $O(N \log N)$ for the range of wave-numbers $\kappa \le O(h^{-1}) = O(N^{1/d})$, where h is the mesh parameter (cf. §4).

2.4 Examples of Blended Approximations

The first example corresponds to the circulant version $\mathcal{M}_{\mathcal{U}_{\mathcal{H},k}\otimes C_m}$ of blended matrix-formats in the BEM application.

Example 2.11 *3D BEM on a rotational surface.*

Assume that the (single-layer) kernel of the boundary integral operator A is given by a translation invariant function $s(x,y) = S(x-y)$ $(x,y \in \mathbb{R}^3)$. Let Γ be a rotational surface $\Gamma = \Gamma_z \times [0, 2\pi] \subset \mathbb{R}^3$ obtained by means of an arc Γ_z. We seek $u \in V := H^{-1/2}(\Gamma)$ such that

$$\langle(\lambda E + A)u, v\rangle = \langle f, v\rangle \qquad \text{for all } v \in V. \tag{2.16}$$

Consider a tensor-product ansatz space $V_h = V_z \times W_\varphi$ of piecewise constant finite elements associated with the tensor product of a uniform grid in $\varphi \in [0, 2\pi]$ (with mesh-size $h_\varphi = \frac{2\pi}{n_\varphi}$) and a quasi-uniform mesh on Γ_z. To be more specific, let $V_z := \mathrm{span}\{\psi_i\}_{i\in I_z}$ with $I_z := \{i = 1, \ldots, n_z\}$ and $W_\varphi := \mathrm{span}\{\varphi_j\}_{j=1}^{n_\varphi}$. Define $V_j := V_z \otimes \varphi_j = \mathrm{span}\{v_i^j\}_{i\in I_z}$, where $v_i^j := \psi_i \otimes \varphi_j$. Note that the entries $\langle A v_i^j, v_k^l\rangle$ of the exact stiffness matrix A_h depend on $i, k \in I_z$ and on the difference $j - l$ modulo n_φ. Hence, $A_h = Bcirc\{A_1, ..., A_{n_\varphi}\}$ has the block structure of $\mathcal{M}_{C_{n_\varphi},1}$ specified similar to Definitions 2.1 and 2.3, where the generating blocks $A_l = \langle A v_i^1, v_k^l\rangle_{i,k=1}^{n_z}$ $(l = 1, \ldots, n_\varphi)$ correspond to the product spaces $V_1 \times V_l$. Approximating these blocks by $\mathcal{H}$-matrices from $\mathcal{U}_{\mathcal{H},k}(I_z \times I_z, P_2, \mathcal{R})$, we arrive at the desired construction. The linear-logarithmic complexity bound from Lemma 2.8 holds with $m = n_\varphi$ and $n = n_z$.

In the next example, the blocks are partly of type $\mathcal{M}_{T_{\mathbf{n}},2}$ and partly of type $\mathcal{M}_{\mathcal{U}_{\mathcal{H},k}\otimes T_n}$.

Example 2.12 *3D BEM on the surface of parallelepiped.*

Consider the integral equation (2.16) on the surface $\Gamma = \partial\Omega$ of $\Omega = (0,1)\times(0,1)\times(0,a)$. We use the translation invariance and the invariance with respect to a rotation by $\pi/2$ around the axis $\{(\frac{1}{2}, \frac{1}{2}, z) : z \in \mathbb{R}\}$. Let $\mathcal{T} = \{\tau_k\}$ be a *uniform quadrangulation* with the mesh-sizes $h_x = h_y = 1/n$ and $h_z = a/m$. Consider the boundary element approximation of the integral operator in (2.16) with respect to the Galerkin space V_h of piecewise constant functions $\{\psi_\alpha\}_{\alpha\in I}$ associated to $\mathcal{T}$. Denote the square facets of Γ by Γ_1 (top) and Γ_2 (bottom), while the remaining four facets are numbered clockwise: $\Gamma_3, \ldots, \Gamma_6$. The exact stiffness matrix has the block form

$$A_h = \{A_{ij}\}_{i,j=1}^6, \qquad A_{ij} = \left(\int_{\Gamma_i\times\Gamma_j} s(x,y)\psi_\alpha(x)\psi_\beta(y)dxdy\right)_{\alpha\in I_i,\,\beta\in I_j}, \tag{2.17}$$

where $\alpha \in I_i$ are those indices corresponding to ψ_α with support in Γ_i. For instance, $I_1 = I_x \times I_y \times \{1\}$ and $I_3 = I_x \times \{1\} \times I_z$ holds, where the one-dimensional index sets are $I_x = I_y = \{1, \ldots, n\}$ and $I_z = \{1, \ldots, m\}$. Introduce the index sets

$$\begin{aligned}
J_{T_1} &:= \{(1,1), (2,2), (2,1), (1,2)\},\\
J_{T_2} &:= \{(3,3), (4,4), (5,5), (6,6), (3,5), (5,3), (4,6), (6,4)\},\\
J_{B_1} &:= \{(1,i), (i,1), (2,i), (i,2) : \; i = 3, \ldots, 6\},\\
J_{B_2} &:= \{(3,4), (4,3), (4,5), (5,4), (5,6), (6,5), (6,3), (3,6)\},
\end{aligned}$$

such that $J_{T_1} \cup J_{T_2} \cup J_{B_1} \cup J_{B_2} = J \times J$, where $J := \{1, \ldots, 6\}$. The translation invariance with respect to two directions implies the two-level $n \times n$ block-Toeplitz structure

$$A_{ij} \in \mathcal{M}_{T_{\mathbf{n}},2} \qquad \text{with } \mathbf{n} = (n,n) \;\text{ for } (i,j) \in J_{T_1}, \tag{2.18}$$

$$A_{ij} \in \mathcal{M}_{T_{\mathbf{n}},2} \qquad \text{with } \mathbf{n} = (n,m) \text{ for } (i,j) \in J_{T_2}. \tag{2.19}$$

Note that each block A_{ij}, $(i,j) \in J_{B_1}$, is obtained from A_{13} by a permutation. The same is true concerning A_{ij}, $(i,j) \in J_{B_2}$, and the reference matrix A_{34}. The matrices A_{13}, A_{34} are approximated by matrices[3] $B_{13}, B_{34} \in \mathcal{M}_{\mathcal{U}_{\mathcal{H},k}\otimes T_n}$ of the blended format using the translation invariance with respect to one direction (the intersection of Γ_1 and Γ_3 in the case of B_{13}); e.g., B_{13} equals

$$B_{13} = BToepl\{T_{1-n}, \ldots, T_0, \ldots, T_{n-1}\} \qquad \text{with } T_i \in \mathcal{U}_{\mathcal{H},k}(I_y \times I_z, P_2) \subset \mathbb{R}^{n\times m},$$

[3] So far, we have mentioned only $\mathcal{H}$-matrices of square format, but the generalisation to a rectangular format is straight-forward.

such that $T_i = T_{-i}^\top$, $i = 1, ..., n-1$. The index sets are $I_x = I_y = \{1, \ldots, n\}$ and $I_z = \{1, \ldots, m\}$. The index i of T_i corresponds to the x-shift between $\{(i_x, i_y) : i_y \in I_y\}$ and $\{(i_x + i, i_z) : i_z \in I_z\}$, where $i_x, i_x + i \in I_x$. The location of admissible clusters in the $\mathcal{H}$-matrix approximation is depicted in Figure 1b (case of $a = 1$).

Lemma 2.13 *Consider $A_h = \{A_{ij}\}_{i,j=1}^6$ from (2.17). $B_h := \{B_{ij}\}_{i,j=1}^6$ is obtained from A_h by approximating all blocks A_{ij}, $(i,j) \in J_{B_1} \cup J_{B_2}$, by uniform $\mathcal{H}$-matrices $B_{ij} \in \mathcal{M}_{\mathcal{U}_{\mathcal{H},k} \otimes T_n}$ based on a uniform hierarchical trees $T(I_x) = T(I_y)$ and $T(I_z)$ of depth L. Further, $B_{ij} := A_{ij} \in \mathcal{M}_{T_{\mathbf{n}},2}$ holds for $(i,j) \in J_{T_1} \cup J_{T_2}$. We assume $n = 2^{L'} \le m = 2^L$. Then there holds*

$$\mathcal{N}_{st}(B_h) \le (n+m)\left[2\sum_{\ell=0}^{L} k_\ell^2 + \sum_{\ell=0}^{L} k_\ell 2^{-\ell} + 2n\right], \quad \mathcal{N}_{MV}(B_h) = O\left((k_0 + \sum_{\ell=0}^{L} k_\ell^2 2^{\ell-L}) nm \log m\right).$$

Proof. First, we consider $B_{13} = BToepl\{T_{1-n}, \ldots, T_0, \ldots, T_{n-1}\}$. Since clusters $\tau, \tau' \in T(I_x)$ on the same level have the same size and differ only by a shift, the basis vectors $\mathbf{a}_i^\tau = \mathbf{c}_i^\tau, \mathbf{a}_j^\sigma = \mathbf{c}_j^\sigma$ for $\tau \in T(I_x) = T(I_y)$ and $\sigma \in T(I_z)$ depend only on the level ℓ. Each $T_i \in \mathcal{U}_{\mathcal{H},k}$ leads to[4] $\mathcal{N}_{st}(B_{13}) \le 2n \sum_{\ell=0}^{L} k_\ell^2$. The basis vectors $\mathbf{a}_i^\tau, \mathbf{c}_j^\sigma$ need a storage size of $n \sum_{\ell=0}^{L'} k_\ell 2^{-\ell} + m \sum_{\ell=0}^{L} k_\ell 2^{-\ell} \le (n+m) \sum_{\ell=0}^{L} k_\ell 2^{-\ell}$.

A similar estimate holds for $\mathcal{N}_{st}(B_{34})$.

Among the matrices A_{ij}, $(i,j) \in J_{T_1} \cup J_{T_2}$, only A_{11}, A_{12}, A_{33}, A_{35} are essential, all other are obtained by permutations. The storage of each A_{ij} amounts to n^2 for $(i,j) \in J_{T_1}$ and nm for $(i,j) \in J_{T_2}$ due to the symmetry (cf. Proposition 2.2)

The second assertion is shown similarly to the proof of Lemma 2.8. ∎

Note that the dominating term in $\mathcal{N}_{st}(B_h)$ is $2(n+m)n$, which is smaller than the total dimension $N := 2n^2 + 4nm$ of this problem.

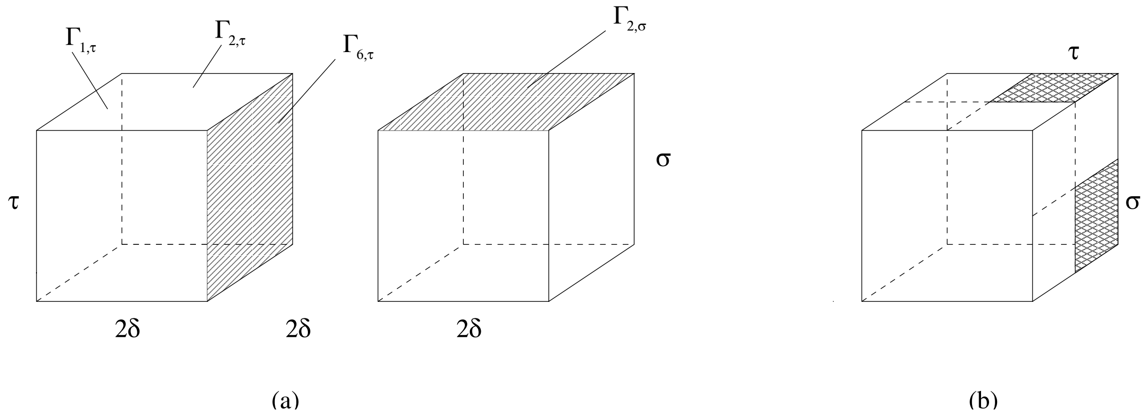

Figure 1: Location of clusters τ and σ in FEM (a) and BEM (b) applications

The next example illustrates an application of multi-level Toeplitz matrices to the approximation volume integrals on a product domain.

Example 2.14 *3D FEM (approximation of volume integrals by block-Toeplitz matrices).*

We consider the FEM application (i.e., the case (**B**)) for the translation invariant kernel $s(x,y) = S(x-y)$ with $d = 3$. We are interested in the data-sparse representation of the Galerkin stiffness matrix

$$A_h = \{A_{\alpha\beta}\}_{(\alpha,\beta)\in I\times I}, \qquad A_{\alpha\beta} = \left(\int_{\Omega\times\Omega} s(x,y)\psi_\alpha(x)\psi_\beta(y)dxdy\right)_{\alpha\in I,\, \beta\in I}, \tag{2.20}$$

where Ω is the unit cube and $\{\psi_\alpha\}$ is a family of piecewise constant FE functions with respect to the uniform $n \times n \times n$ grid. It is easy to see that $A_h \in \mathcal{M}_{T_{\mathbf{n}},3}$, $\mathbf{n} = (n,n,n)$, i.e., it is a three-level Toeplitz matrix (see Definition 2.1). Thus we obtain

$$\mathcal{N}_{MV}(A_h) = c_{FFT} 2^4 N \log N, \qquad \mathcal{N}_{st}(M) = 2^3 N,$$

[4]Note that the singularity of the kernel function is located at the common edge. Therefore, the admissible partitioning is of the $\mathcal{N}$-type explained in [5].

where $N = n^3$ (see Proposition 2.2). In particular, A_h may be considered as a block of the Galerkin stiffness matrix in the domain containing Ω.

The last example illustrates an application of a mixed Toeplitz and low rank representation of blocks from Definition 2.4.

Example 2.15 *The wire-basket expansions in the 3D FEM (reduction of volume integrals to the surface).*

We consider again the 3D FEM application for the translation invariant kernel $s(x,y) = S(x-y)$ now given by the fundamental solution of an elliptic operator with constant coefficients. We use piecewise constant finite elements on a tensor-product grid. To fix the idea, let us choose an admissible block $\tau \times \sigma$ of the size $n \times n \times n$ as shown in Figure 1a (the left cube is the domain $X(\tau)$, while $X(\sigma)$ is the right one). Following [9, Lemma 3.1], we apply the wire-basket expansions which include only the surface integrals approximated by a similar FE scheme as in Example 2.12. We recall that for $u = \mathcal{L}_x g_u$ and $v = \mathcal{L}_y g_v$, $g_u \in H_0^1(X(\tau))$, $g_v \in H_0^1(X(\sigma))$, there holds

$$\int_{X(\tau)\times X(\sigma)} s(x,y)u(x)v(y)dxdy = \int_{\partial X(\sigma)} \int_{\partial X(\tau)} s(x,y)\partial_\nu g_v(y)\partial_\nu g_u(x)dxdy$$

for all $\tau \times \sigma \in P_{far}$ (see [9] for more details). Further, we assume that the conormal derivatives $\partial_\nu g_v(y)$ and $\partial_\nu g_u(x)$ are already approximated by piecewise constant functions on the same grid as before. The corresponding matrix representation is obtained replacing the functions u, v by the respective (piecewise constant) basis functions for the uniform $n \times n$ grid on the facets $\Gamma_\tau = \partial X(\tau)$ and $\Gamma_\sigma = \partial X(\sigma)$. Hence, to approximate the exact stiffness matrix defined for the product surface $\Gamma_\tau \times \Gamma_\sigma$, we consider the blockwise blended rank-k and Toeplitz type representation of the form

$$\mathbf{B} =: \{B_{ij}\}_{i,j=1}^{6}, \qquad B_{ij} \in \mathbb{R}^{n^2 \times n^2},$$

with respect to the degrees of freedom located on the product pieces $\Gamma_{i,\tau} \times \Gamma_{j,\sigma}$, $i,j \in J = \{1,\dots,6\}$ (see Figure 1a). Consider the splitting of the product index set $J \times J = J_T \cup J_B$, where

$$J_T := \{(1,1),(1,6),(2,3),(2,5),(3,3),(3,4),(4,4),(4,3),(5,5),(5,2),(6,1),(6,6)\}.$$

Here J_T describes the pairs of parallel facets, while J_B corresponds to orthogonal ones. Now we specify the block representation by

$$B_{ij} \in \mathcal{M}_{T_{\mathbf{m}},2}, \quad (i,j) \in J_T, \quad \mathbf{m} = (n,n) \quad \text{and} \quad B_{ij} \in \mathcal{M}_{\mathcal{R}_k \otimes T_n}, \quad (i,j) \in J_B. \tag{2.21}$$

Note that the only the blocks B_{ij} for $(i,j) \in J_{Ts} := \{(1,1),(1,6),(2,2),(2,5),(6,1)\} \subset J_T$ and $(i,j) \in J_{Bs} := \{(1,2),(2,6),(2,1),(2,3)\} \subset J_B$ have to be stored. All other blocks B_{ij}, i.e., for $(i,j) \notin J_{Ts} \cup J_{Bs}$, are obtained from the former ones by simple permutations. The two-level Toeplitz blocks correspond to the parallel facets (case of $(i,j) \in J_T$), while matrices of blended type are related to orthogonal facets (case of $(i,j) \in J_B$).

Lemma 2.16 *Assume that* $\mathbf{B}$ *has the block structure (2.21) and let* $N_\Gamma = 6n^2$. *Then there holds*

$$\mathcal{N}_{st}(\mathbf{B}) = \frac{5}{3}N_\Gamma + 4(2kn + (2n-1)k^2),$$
$$\mathcal{N}_{MV}(\mathbf{B}) = 4c_{FFT}N_\Gamma \log N_\Gamma + 24(4c_{FFT}\log n + 1)k^2 n + 16kN_\Gamma.$$

Proof. The statement follows from Lemma 2.5 and Proposition 2.2 due to the relations

$$\mathcal{N}_{st}(\mathbf{B}) = 2n^2 \# J_{Ts} + (2kn + (2n-1)k^2)\# J_{Bs},$$
$$\mathcal{N}_{MV}(\mathbf{B}) = 2c_{FFT}n^2 \# J_T \log n^2 + [(4c_{FFT}\log n + 1)k^2 n + 4kn^2]\# J_B,$$

and taking into account that $\# J_T = 12$ and $\# J_B = 24$. This completes our proof. ■

In the 2D case, the advantage of blended approximations versus the multipole and the wire-basket expansion was illustrated in [9, Figure 4]. For $d = 3$, the superiority turns out to be even greater. The point is that the rank-k ansatz here corresponds to the separable approximation of a function of two variables, thus implying $k = O(L)$ instead of $k = O(L^2)$ for the product low rank approximation of the blocks B_{ij}. In particular, this improves the situation in the case of oscillatory kernels, where we reduce the block-rank from $k = O(\kappa^2 + \log^2 N)$ down to $k = O(\kappa \log N)$, where κ is the wave number (see §4).

To conclude this section, we emphasise that the $\mathcal{H}$-matrix formats $\mathcal{M}_{\mathcal{U}_{\mathcal{H},k}\otimes T_{\mathbf{m}}}$ and $\mathcal{M}_{\mathcal{U}_{\mathcal{H},k}\otimes C_m}$ based on the blended FE and polynomial expansions under consideration can be applied to the fast matrix-vector product only, but, in general, not to the formatted matrix-matrix multiplication and inversion.

3 Error and Complexity Analysis for Variable Order Expansions

Below, we present general consistency error estimates for hierarchical matrices using variable order kernel expansions depending on the behaviour of the local rank. We focus on the case of volume integral operators arising in FEM applications. In the case of constant rank $\mathcal{H}$-matrices the consistency error analysis was presented in [7]. Our analysis here extends the results from [10], where the case of variable order expansions in BEM for $d = 1$ has been considered first.

3.1 Error Analysis (Case (B))

The perturbation of the matrix A by $A_{\mathcal{H}} - A$ yields a perturbed discrete solution $u_{\mathcal{H}} \in V_h$ of the Galerkin equation

$$\langle (\lambda I + A_{\mathcal{H}})u_{\mathcal{H}}, v\rangle = \langle f, v\rangle \qquad \text{for all } v \in V_h. \tag{3.1}$$

The existence of accurate kernel approximations by separable functions on admissible geometrical blocks (cf. (3.6) below), will be based on the analyticity of the kernel for $x \neq y$:

Assumption 3.1 *For any $x_0, y_0 \in \Omega$, $x_0 \neq y_0$, the kernel function $s(x,y)$ is analytic with respect to x and y at least in the domain*

$$\{(x,y) \in \Omega \times \Omega : \ |x - x_0| + |y - y_0| < |x_0 - y_0|\}. \tag{3.2}$$

We consider local expansions of level-dependent variable degree $m = m_\ell$ defined by

$$m_\ell = a\, L^{1-q}(L-\ell)^q + b \qquad \text{with } 0 \le q \le 1,\ a > 0, b \ge 0. \tag{3.3}$$

Then the rank k_ℓ of block of level ℓ is defined by

$$k_\ell := \min\{2^{d(L-\ell)}, m_\ell^{d-1}\}. \tag{3.4}$$

The first argument $2^{d(L-\ell)}$ is the size of a block of level ℓ in the case of a regular tree, where each father has 2^d sons (cf. Assumption 3.2 below). The second argument is an upper bound of dim{polynomial in $\mathbb{R}^{d-1}$ of degree $< m_\ell$}.

In the following, $\lesssim$ means $\le$ up to a constant factor independent of the particular parameters involved.

Assumption 3.2 *For each $\tau \times \sigma \in P_2^{(\ell)}$ there holds*

$$|\tau| := \operatorname{meas}(X(\tau)) \lesssim 2^{-d\ell}, \qquad \#\tau \lesssim 2^{d(L-\ell)}, \qquad 2^{-\ell} \lesssim \operatorname{dist}(\tau,\sigma), \qquad \#P_2^{(\ell)} \lesssim 2^{d\ell}. \tag{3.5}$$

Assumption 3.2 is valid, e.g., for the case of quasi-uniform grids. Introduce

$$N_0 = \max\left\{ \max_{\tau \in T(I)} \sum_{\tau:\tau\times\sigma\in P_2} 1,\ \max_{\sigma\in T(I)} \sum_{\sigma:\tau\times\sigma\in P_2} 1 \right\}.$$

This is a constant since it is bounded by $\#P_2 \le const$ (cf. [5, Subsection 3.1])

Lemma 3.3 *Let $d \ge 2$, $\eta = 2^{-\alpha}$ with $\alpha > 0$ (cf. (2.6)). Moreover, the approximation of s by the separable expansion (2.8) is assumed to satisfy*

$$|s(x,y) - s_{\tau\sigma}(x,y)| \lesssim \eta^{m_\ell}\ell^{3-d}\operatorname{dist}(\tau,\sigma)^{2-d} \quad \textit{for all } \tau\times\sigma \in P_2^{(\ell)}, (x,y) \in X(\tau)\times X(\sigma) \tag{3.6}$$

where the order of expansion m_ℓ is defined by (3.3) with $a > 0$ satisfying $-\alpha a + 2 < 0$. Then, there holds

$$\|A - A_{\mathcal{H}}\|_{L^2(\Omega)\to L^2(\Omega)} \lesssim h^2 N_0 \delta(L,q), \tag{3.7}$$

where

$$\delta(L,q) = \sum_{j=1}^{L} (L-j)^{3-d} 2^{(-\alpha a+2)L^{1-q}j^q} \qquad (q \textit{ from } (3.3)). \tag{3.8}$$

The condition $-\alpha a+2<0$, which is equivalent to $\eta^a<1/4$, is not needed for the proof, but is essential for the behaviour of $\delta(L,q)$. Below we shall give an precise analysis of $\delta(L,q)$.

Proof. For the case of volume integrals our proof is a simple modification of arguments from [10, Theorem 4.13] related to the situation of boundary integral operators with $d=1$, $q=1$. Using Assumptions 3.2 and (3.6) we obtain[5] for all $u,v\in L^2(\Omega)$,

$$\begin{aligned}
|\langle (A-A_{\mathcal{H}})u,v\rangle| &= \left|\sum_{\ell=0}^{L-1}\sum_{\tau\times\sigma\in P_2^{(\ell)}}\int_{X(\tau)\times X(\sigma)} u(x)(s(x,y)-s_{\tau,\sigma}(x,y))v(y)dxdy\right| \\
&\lesssim \sum_{\ell=1}^{L-1}\ell^{3-d}\,2^{-\alpha m_\ell+(d-2)\ell}\sum_{\tau\times\sigma\in P_2^{(\ell)}}\sqrt{|\tau|\,|\sigma|}\,||u||_{0,\tau}||v||_{0,\sigma} \\
&\lesssim 2^{-2L}\sum_{\ell=0}^{L-1}\ell^{3-d}\,2^{-\alpha a L^{1-q}(L-\ell)^q+2(L-\ell)}\cdot\sqrt{\sum_{\tau\times\sigma\in P_2^{(\ell)}}||u||_{0,\tau}^2}\sqrt{\sum_{\tau\times\sigma\in P_2^{(\ell)}}||v||_{0,\sigma}^2} \\
&\lesssim h^2\sum_{\ell=0}^{L-1}\ell^{3-d}\,2^{(-\alpha a+2)L^{1-q}(L-\ell)^q}\cdot\sqrt{\sum_{\tau\in T(\ell)}||u||_{0,\tau}^2\sum_{\sigma:\tau\times\sigma\in P_2^{(\ell)}}1}\sqrt{\sum_{\sigma\in T(\ell)}||u||_{0,\sigma}^2\sum_{\tau:\tau\times\sigma\in P_2^{(\ell)}}1} \\
&\lesssim h^2 N_0\sum_{j=1}^{L}(L-j)^{3-d}\,2^{(-\alpha a+2)L^{1-q}j^q}\,||u||_0\,||v||_0,
\end{aligned}$$

which proves the statement. ■

Note that in the case of a constant order expansions, i.e., for $q=0$, we obtain exponential convergence,

$$||A-A_{\mathcal{H}}||_{L^2(\Omega)\to L^2(\Omega)}\lesssim N_0L^{4-d}2^{-\alpha a L}$$

for any $a>0$. Let us investigate more carefully the function $\delta(L,q)$ defined by (3.8). First, consider the case $d=3$. Introduce the parameter $\beta:=\log 2^{(\alpha a-2)L^{1-q}}>0$, i.e., $2^{(-\alpha a+2)L^{1-q}}=\mathrm{e}^{-\beta}$. Then we obtain the estimate

$$\delta(L,q)=\sum_{j=1}^{L}\mathrm{e}^{-\beta j^q}\leq\int_0^L\mathrm{e}^{-\beta z^q}dz.$$

The latter integral can be presented in terms of the incomplete gamma function

$$\gamma(\alpha,x)=\int_0^x\mathrm{e}^{-t}t^{\alpha-1}dt\qquad(\Re e\,\alpha>0).$$

In fact, substituting $t=\beta z^q$, we obtain $\int_0^L e^{-\beta z^q}dz=\frac{1}{q}\beta^{-\frac{1}{q}}\int_0^{\beta L^q}\mathrm{e}^{-t}t^{\frac{1}{q}-1}dt=\frac{1}{q}\beta^{-\frac{1}{q}}\gamma(\frac{1}{q},\beta L^q)$. To derive an explicit estimate, we further set $q=n^{-1}$ ($n\in\mathbb{N}$) and then apply the following special representation of $\gamma(\alpha,x)$ for the case $\alpha=n+1$ (cf. [2]), $\gamma(n+1,x)=n!\left[1-\mathrm{e}^{-x}\left(\sum\limits_{m=0}^{n}\frac{x^m}{m!}\right)\right]$, which leads to the bound

$$\begin{aligned}
\delta\left(L,\frac{1}{n}\right)&\leq n\beta^{-n}(n-1)!\left[1-\mathrm{e}^{-\beta L^{\frac{1}{n}}}\left(\sum_{m=0}^{n-1}\frac{(\beta L^{\frac{1}{n}})^m}{m!}\right)\right]\lesssim n!\beta^{-n}\mathrm{e}^{-\beta L^{\frac{1}{n}}}\sum_{m=n}^{\infty}\frac{(\beta L^{\frac{1}{n}})^m}{m!} \\
&\lesssim n!\,\beta^{-n}\,\mathrm{e}^{-\mu L}\frac{\beta^n}{n!}\,L\lesssim L\,\mathrm{e}^{-\mu L},
\end{aligned}$$

where, by definition, $\beta L^{\frac{1}{n}}=(\alpha a-2)\,L\log 2$, so that $\mu=(\alpha a-2)\log 2>0$. We then obtain exponential convergence with respect to L at least for $n>\mu\,L$. In the case $d=2$, one obtains an extra factor L in front of the above estimate.

[5] Here we assume that the supports $X(\tau)$ are disjoint; i.e., $\tau'\cap\tau''=\emptyset$ implies $X(\tau')$ and $X(\tau'')$ are at most touching. This holds for piecewise constant basis functions; otherwise, $X(\tau')$ and $X(\tau'')$ have an overlap and the proof becomes more technical. In the overlap different expansions are used, but all satisfy the estimate (3.6).

3.2 Remarks on the Complexity for Variable Order Expansions

The complexity of the variable order $\mathcal{H}$-matrices with m_ℓ given by (3.3) for $d = 2, 3$ depends on the representation of the matrix blocks. Using the representation (2.10) with fixed basis, we obtain

$$\mathbf{A}_{\mathcal{H}}^{\tau\times\sigma} = \sum_{i,j=1}^{k_\ell} a_{ij}(\mathbf{a}_i^\tau \cdot (\mathbf{c}_j^\sigma)^\top) \in \mathcal{V}_a^\tau \otimes \mathcal{V}_c^\sigma, \qquad \mathbf{a}_i^\tau \in \mathbb{R}^\tau, \; \mathbf{c}_j^\sigma \in \mathbb{R}^\sigma, \tag{3.9}$$

while for the wire-basket method (see [9]) there holds

$$\mathbf{A}_{\mathcal{H}}^{\tau\times\sigma} = \mathbf{S}_\tau^T \left(\sum_{i=1}^{k_\ell} \mathbf{a}_i^{\partial\tau} \cdot \mathbf{c}_i^{\partial\sigma\top} \right) \mathbf{S}_\sigma, \qquad \mathbf{a}_i^{\partial\tau} \in \mathbb{R}^{n_{\partial\tau}}, \; \mathbf{c}_i^{\partial\sigma} \in \mathbb{R}^{n_{\partial\sigma}} \tag{3.10}$$

with $k_\ell = O(m_\ell^{d-1})$. Here the matrices $\mathbf{S}_\tau, \mathbf{S}_\sigma$ stand for the Schur-complements associated with the local elliptic solution operator (see [9] for more details).

Lemma 3.4 *Under Assumption 3.2 we obtain for the block representation (3.9) the following storage estimate*

$$\mathcal{N}_{st}(\mathbf{A}_{\mathcal{H}}) \lesssim N_0 \sum_{\ell=0}^{L} k_\ell^2 2^{d\ell} \lesssim N_0 L^{2(1-q)(d-1)} N,$$

while for the representation (3.10) there holds

$$\mathcal{N}_{st}(\mathbf{A}_{\mathcal{H}}) \leq N_0 L^{(1-q)(d-1)} N + \max_\tau \mathcal{N}_{st}(\mathbf{S}_\tau).$$

Proof. Case (3.9) is proven in [10] and concerning case (3.10) we refer to [9]. ■

As a result, in both situations, we arrive at linear complexity with the choice $q = 1$ in (3.3). It is shown that the matrix-vector product has linear complexity as well (cf. [9, 10]).

3.3 Application to the Operator $A = \mathcal{L}^{-1}$

Let us consider the special case of $A = \mathcal{L}^{-1}$. For second order elliptic problems, A is a mapping $A : H^{-1} \to H^1$. The first important consequence of Lemma 3.3 is that for variable order expansions with $q = 1$ (implying $\delta(L, q) \approx 1/h$) the asymptotically optimal convergence of the order $O(h)$ is verified only for trial functions from $L^2(\Omega)$. On the other hand, exponential convergence in the operator norm $||\cdot||_{H^{-1}\to H^1}$ may be proven at least for $0 \leq q < O(L^{-1})$. Denote by $A_h : V_h \to V_h'$ (resp. $A_{\mathcal{H},h}$) the restriction of A (resp. $A_{\mathcal{H}}$) onto the Galerkin subspace $V_h \subset L^2(\Omega)$ defined by $\langle A_h u, v\rangle = \langle Au, v\rangle$ and $\langle A_{\mathcal{H},h}u, v\rangle = \langle A_{\mathcal{H}}Au, v\rangle$ for all $u, v \in V_h$. We summarise:

Corollary 3.5 *Suppose that the inverse inequality $||v||_{0,\Omega} \lesssim h^{-1}||v||_{-1,\Omega}$ is valid for any $u \in V_h$. Then there holds*

$$||A_h - A_{\mathcal{H},h}||_{H^{-1}(\Omega)\to H^1(\Omega)} \lesssim N_0\delta(L, q), \qquad q \in [0, 1]. \tag{3.11}$$

Proof. Estimate (3.7) and the inverse inequality imply

$$||(A_h - A_{\mathcal{H},h})u_h||_{H^1(\Omega)} = \sup_{v\in V_h} \frac{\langle (A_h - A_{\mathcal{H},h})u_h, v\rangle}{||v||_{-1,\Omega}} \lesssim hN_0\delta(L, q)||u_h||_0.$$

Finally, the repeated application of the inverse inequality now to the term $||u_h||_0$ implies (3.11). ■

Remark 3.6 *In the case $q = 1$ and $d = 2, 3$, we obtain the optimal error estimate for functions $u \in L^2(\Omega)$. The case $d = 1$ is not supported by Lemma 3.3. However, for $d = 1$ we do not need the variable order expansions because in this case the wire-basket expansion provides the exact approximation with the local rank of constant order $k_{const} = O(m^{d-1}) = O(1)$ which again leads to linear complexity.*

4 Applications to Oscillatory Kernels

Consider the 3D Galerkin BEM on, e.g., $\Gamma = \partial((0,1)^3)$ (see Figure 1b) with the Helmholtz kernel

$$s(x,y) = \frac{e^{i\kappa|x-y|}}{|x-y|} \qquad (x, y \in \mathbb{R}^3). \tag{4.1}$$

The blended $\mathcal{H}$-matrix approximation for canonical boundaries leads to robust methods of linear-logarithmic complexity in the range $\kappa \le O(h^{-1})$. In particular, for the Helmholtz kernel defined on the boundary shown in Figure 1b, we obtain a memory estimate $O(N(\log N)(\kappa + \log N))$ compared to $O(N(\kappa + \log N)^2)$, where the latter bound corresponds to methods based on the multipole or standard polynomial expansions.

Specifically, we consider the Galerkin stiffness matrix for piecewise constant basis functions. We use a uniform grid on each facet of the cube with $(2n)^2$ cells. The cluster tree starts at level 0 with the index set I (therefore $X(I) = \Gamma$). The clusters of level 1 are the subsets $I_1, \dots, I_6$ such that $X(I_i)$ are the 6 facets of the cube. Two clusters of level 2 are τ and σ shown in Figure 1b. Both clusters τ, σ contain n^2 indices. We order the indices of τ (and σ) with respect to blocks $\tau_1, \dots, \tau_n$ (and $\sigma_1, \dots, \sigma_n$) of n indices corresponding to strips *orthogonal* to $X(\sigma)$ (and $X(\tau)$, respectively). Thus the matrix block B corresponding to $\tau \times \sigma$ has the block-Toeplitz structure[6] BToepl$\{T_0 \dots, T_{n-1}\}$ for $n \times n$ matrices T_j. In particular, T_0 coincides with all diagonal blocks of B. The κ-dependence of the eigenvalues of T_0 is depicted in the bottom pictures of Figure 2. The eigenvalues of T_j, $j > 0$, are even smaller.

Next we consider another block structure of the indices: We order the indices of τ (and σ) with respect to blocks $\tau_1, \dots, \tau_n$ (and $\sigma_1, \dots, \sigma_n$) corresponding to strips *parallel* to $X(\sigma)$ (and $X(\tau)$, respectively), which gives rise to another block-Toeplitz structure. The upper pictures in Figure 2 show the eigenvalues for the $n \times n$ block matrix corresponding to the two parallel grid lines in $X(\tau) \times X(\sigma)$ with the smallest distance 1/2. The results show that the former choice is more advantageous.

In particular, for our choice $n = 256$, we may observe that the low-rank approximations of $n \times n$ subblocks by means of the rank $k_{subblock} \le 32$ matrices provides the accuracy 10^{-10} for the whole range of $\kappa \le n/4$. However, for the low-rank approximations of the corresponding full $n^2 \times n^2$ block we expect the squared rank of the order $k_{block} \sim k^2_{subblock} \sim (\kappa + \log n)^2$ which is far from practical usage. Our blended approximation essentially reduces this dramatic (quadratic) growth of the complexity with respect to the wave number κ.

4.1 Polynomial Approximation for the Helmholtz Kernel

The approximation by separable expansions is based on Assumption 3.1 and on the results from [9] on the polynomial interpolation by multivariate functions in the domain I_1^d, $d \ge 1$, where $I_1 := [-1, 1]$.

Definition 4.1 *A function $f \in C^\infty(I_1)$ has Bernstein's regularity ellipse $\mathcal{E}_H(I_1)$ if it admits an analytic extension to the closed ellipse $\mathcal{E}_H(I_1) \subset \mathbb{C}$ with foci at $z = \pm 1$ and the sum of semi-axes equal to $H > 1$.*

For multivariate functions $f = f(x_1, \dots, x_d) : I_1^d \subset \mathbb{R}^d \to \mathbb{R}$ we use the tensor product interpolant

$$\mathbf{I}_p f = I_p^1 \cdots I_p^d f \in \mathcal{P}_p[I_1^d],$$

where $I_p^i f$ denotes the interpolation polynomial of degree p with respect to the variable x_i $(i = 1, \dots, d)$ at the Chebyshev nodes. The interpolation points ξ_α $(\alpha = (i_1, \dots, i_d) \in \mathbb{N}_0^d)$ in I_1^d are obtained by the Cartesian product of the one-dimensional nodes,

$$\xi_\alpha := \left(\cos\frac{\pi i_1}{p}, \dots, \cos\frac{\pi i_d}{p}\right), \qquad i_j = 0, \dots, p, \quad j = 1, \dots, d.$$

Denote by X_{-i} the subset $X_{-i} := \{(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d) : x_j \in I_1\} \subset \mathbb{R}^{d-1}$ of $d-1$ spatial variables. The properties of Definition 4.1 allow an explicit description of the polynomial approximation error.

Proposition 4.2 *[9]Assume that for a given function $f \in C^\infty(I_1^d)$ there is an $H_0 > 1$ such that for any one-dimensional subset $[x_i, z_i] \in I_1$ with fixed coordinate-vector $z_i \in X_{-i}$ $(i = 1, \dots, d)$ there exists an analytic extension with respect to $x_i \in \mathcal{E}_{H_0}(I_i) \subset \mathbb{C}$. Then, for $1 < H < H_0$, there holds*

$$||f - \mathbf{I}_p f||_{L^\infty(I_1^d)} \le cp \log^{d-1} p \frac{H^{-p}}{H-1} M_H(f), \tag{4.2}$$

[6] Note that the matrix B_{62} in the FEM application from Example 2.15 has the same structure. $B_{62} \in \mathcal{M}_{\mathcal{R}_k \otimes T_n}$ corresponds to the facets $\Gamma_{6,\tau} \times \Gamma_{2,\sigma}$ from Figure 1a.

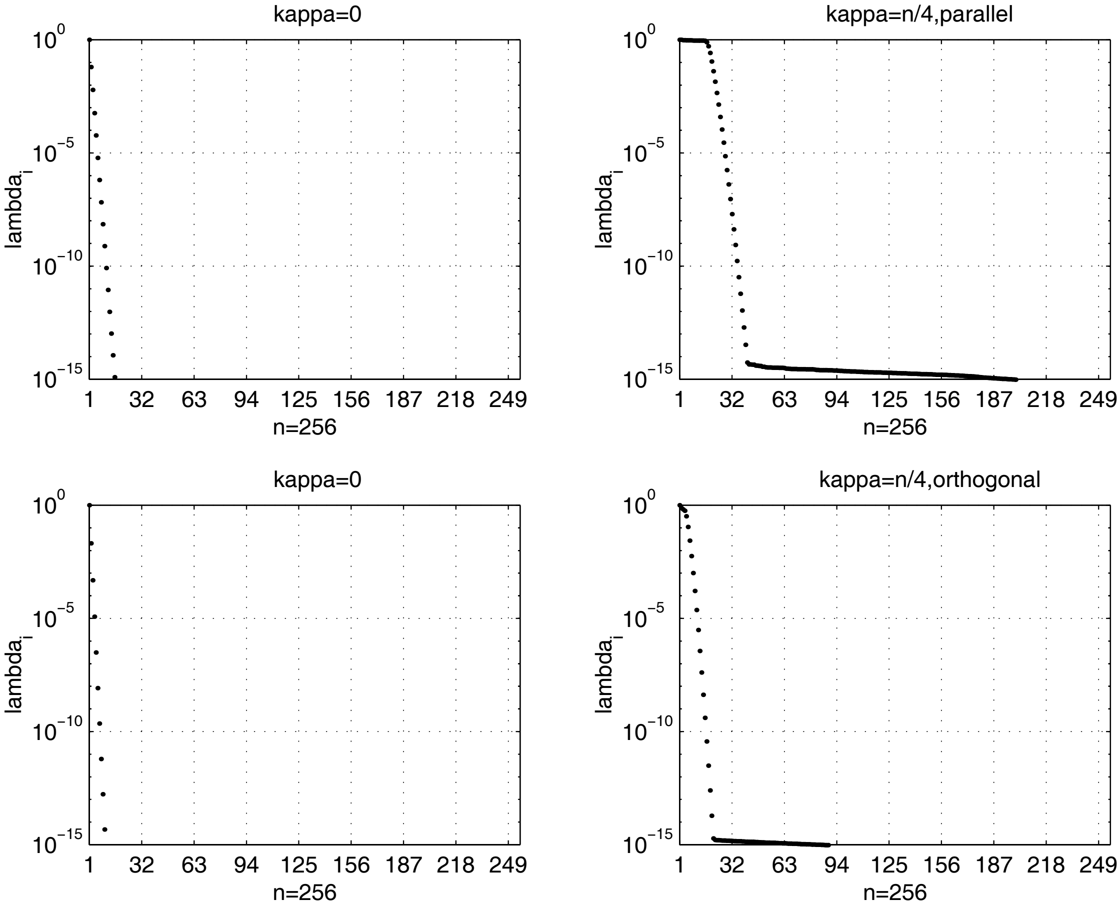

Figure 2: SVD corresponding to the nearest parallel and orthogonal facets.

where $M_H(f) = \max_{1\le j\le d}\{\max_{X_{-j}} \max_{x_j\in\mathcal{E}_H(I_j)} |f(x_1,\dots,x_d)|\}$.

Remark 4.3 *In the case of a scaled domain* I_δ^d, $I_\delta = [-\delta,\delta], \delta > 0$, *the parameter* H *in the error estimate (4.2) is to be substituted by* H/δ, *while* H_0 *then satisfies* $H_0 > \delta$.

In the BEM applications for $d \ge 2$, let $\tau\times\sigma \in P_2$ be a block satisfying the admissibility condition (2.6). To simplify the exposition, we assume that $X(\tau)$ and $X(\sigma)$ are $(d-1)$-dimensional cubes.

For example, in the 3D case of surface integrals, we assume that $X(\tau)$ is a rectangle with the boundary $\partial X(\tau) = \cup_{i=1}^4 \Gamma_\tau^i$ and $\partial X(\sigma) = \cup_{i=1}^4 \Gamma_\sigma^i$ with $|\Gamma_\sigma^i| = |\Gamma_\tau^i| = 2\delta$ (see Figure 3). Suppose that the edges Γ_σ^3 and Γ_τ^1 are parallel to the x_2-axis and satisfy $\mathrm{dist}(\Gamma_\sigma^3, \Gamma_\tau^1) = 2\delta$. this corresponds to the choice $\eta = \frac{1}{\sqrt{2}}$ in (2.6). The location of clusters $X(\tau), X(\sigma)$ depicted in Figure 2 may be considered similarly. Now we analyse the error of a kernel expansion on the product domain $X(\sigma)\times X(\tau) \in \mathbb{R}^4$.

Lemma 4.4 *Let* $s(x,y)$, $(x,y)\in X(\tau)\times X(\sigma)$, *be given by (4.1). Then for the tensor product interpolant by polynomials of degree* p *there holds*

$$||s(x,y) - \mathbf{I}_p s||_{L^\infty(X(\sigma)\times X(\tau))} \le c \log^3 p\, \frac{\exp(2\kappa\lambda\delta)}{2\delta(1-\lambda)} \left(2\lambda + \sqrt{4\lambda^2+1}\right)^{-p}, \tag{4.3}$$

for any $\lambda \in (0,1)$, *uniformly with respect to the block-size* n.

Proof. We apply Proposition 4.2 to the function $f := s(x,y)$ of four variables, i.e., $(x,y)\in I_\delta^2\times I_\delta^2$. Consider the particular choice of (x_i, z_i) in Proposition 4.2 identifying I_δ with Γ_τ^1, i.e., $x_i \in I_\delta := \Gamma_\tau^1$. The corresponding regularity ellipse $\mathcal{E}_{H_0}$ in the sense of Proposition 4.2 has $H_0 = a_0 + b_0$, where $a_0^2 = b_0^2 + \delta^2$, and the small

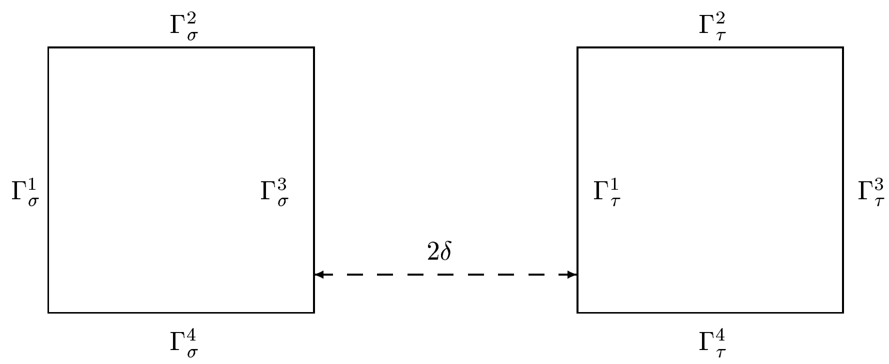

Figure 3: Location of geometrical clusters $X(\tau)$ and $X(\sigma)$.

semiaxis b_0 given by $b_0 = \lambda_0 \operatorname{dist}(X(\tau), X(\sigma)) = 2\lambda_0\delta$, with some $\lambda_0 < 1$ and also $a_0 = \sqrt{1+4\lambda_0^2}\,\delta$. With this choice we have

$$H_0 = \left(2\lambda_0 + \sqrt{1+4\lambda_0^2}\right)\delta. \tag{4.4}$$

Applying Proposition 4.2 with the scaling argument from Remark 4.3, we are lead to the representation

$$\left(\frac{H}{\delta}\right)^{-1} = \left(2\lambda + \sqrt{1+4\lambda^2}\right)^{-1}, \qquad (\lambda < \lambda_0 < 1),$$

corresponding to $b = 2\lambda\delta < b_0$. For the kernel given by (4.1), the constant $M_H(f)$ can be estimated by

$$M_H(f) \le c \max_{x\in\mathcal{E}_H(I_\delta),\, y\in\Gamma^3_\delta} |S(x-y)| \le c\frac{\exp(\kappa b)}{2\delta - b},$$

where S denotes the corresponding fundamental solution. Then the assertion follows. ■

4.2 Complexity for the Helmholtz Kernel

Applying the error estimate (4.3) in the situations from Examples 2.11 and 2.12 (see §2.4), we obtain $\delta = O(2^{-\ell})$ on level ℓ implying that the local rank can be estimated by

$$k_\ell = O(\log N + 2^{-\ell}\kappa) \quad \text{and} \quad k_\ell = O(a(L-\ell) + 2^{-\ell}\kappa),$$

in the case of constant and variable order approximations, respectively. Combining these bounds with Lemmata 2.8 and 2.13 leads to the desired complexity estimates in the case of the Helmholtz kernel. In particular, for a rotational surface and with constant order expansions, we obtain the estimate

$$\mathcal{N}_{st}(\mathbf{B}) = O(N + n_z L\,\kappa \log^2 N), \tag{4.5}$$

where $N = n_z n_\varphi$, while for the surface of parallelepiped there holds

$$\mathcal{N}_{st}(\mathbf{B}) = O(N + (n+m)(\kappa^2 + \log^3 N)),$$

with $N = 2n^2 + 4nm$. Now one may estimate the complexity in the practically interesting range of wave numbers $\kappa \le O(h^{-1}) = O(\sqrt{N})$.

Corollary 4.5 *In the situation of Example 2.11, the BEM stiffness matrix $\mathbf{B} \in \mathcal{M}_{\mathcal{M}_{\mathcal{H},k}\otimes C_m}$ of blended type yields the following complexity estimate*

$$\mathcal{N}_{MV}(\mathbf{B}) = O(N(L + \kappa + L\frac{\kappa^2}{n_z})\log N),$$

where $N = n_z n_\varphi$. The storage is estimated by 4.5.

Corresponding to Examples 2.12 and 2.14, the numerical calculations of the local rank versus the wave number (in the case of orthogonal edges) are presented in Figure 2 (bottom). The results are much better compared with the case of parallel facets, see Figure 2 (top), where instead of a low-rank approximation, we now adapt the exact stiffness matrix of the block-Toeplitz type.

To complete the discussion, we note that the blended approximations may be directly applied in the 3D BEM for special surfaces (e.g., rotational surface, boundary of parallelepiped or L-shaped domains, etc.). In particular, this is the case for coupled FEM-BEM methods for solving elliptic problems in unbounded domains since, in this situation, an auxiliary boundary can be chosen as a special surface.

References

[1] W. Dahmen, S. Prößdorf, and R. Schneider: *Wavelet approximation methods for pseudodifferential equations II: Matrix compression and fast solution*, Adv. Comput. Math. **1** (1993) 259-335.

[2] I.S. Gradshteyn, I.M. Ryzhik: *Tables of integrals, series and products*, Academic Press, 1994.

[3] L. Greengard and V. Rokhlin: *A new version of the fast multipole method for the Laplace equation in three dimensions.* Acta Numerica, **6**, 1997, 229-269.

[4] W. Hackbusch: *Elliptic differential equations: theory and numerical treatment.* Springer-Verlag, Berlin, 1992.

[5] W. Hackbusch: *A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices.* Computing **62** (1999) 89-108.

[6] W. Hackbusch and B.N. Khoromskij: *A sparse $\mathcal{H}$-matrix arithmetic. Part II: Application to multi-dimensional problems.* Computing **64** (2000) 21-47.

[7] W. Hackbusch and B. N. Khoromskij: *A sparse $\mathcal{H}$-matrix arithmetic: General complexity estimates.* In: Numerical Analysis in the 20th Century, vol. 6: Ordinary Differential and Integral Equations (J. Pryce, G. Van den Berghe, C. T. H. Baker, G. Monegato, eds.), J. Comp. Appl. Math., Elsevier, 2000.

[8] W. Hackbusch and B. N. Khoromskij: *$\mathcal{H}$-Matrix approximation on graded meshes.* The Mathematics of Finite Elements and Applications X, J.R. Whiteman (ed), Elsevier, Amsterdam, 2000, pp. 307-316.

[9] W. Hackbusch and B.N. Khoromskij. *Towards $\mathcal{H}$*-matrix approximation of the linear complexity. Preprint No. 15, MPI Leipzig, 2000; Siegfrid Prößdorf Memorial Volume (J. Elschner, I. Gohberg, B. Silbermann, eds.) Birkhäuser Verlag, 2000, to appear.

[10] W. Hackbusch, B.N. Khoromskij and S. Sauter: *On $\mathcal{H}^2$-matrices.* In: Lectures on Applied Mathematics (H.-J. Bungartz, R. Hoppe, C. Zenger, eds.), Springer Verlag, 2000, pp. 9-29.

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

[12] C. Lage and C. Schwab: *Advanced boundary element algorithms*, Preprint ETH Zürich, June, 1999.

[13] S. Rjasanow: *Effective algorithms with circulant-block matrices*, Linear Algebra Appl. **202** (1994) 55-69.

[14] E.E. Tyrtyshnikov: *Mosaic-skeleton approximations.* Calcolo **33** (1996) 47-57.

[15] E.P. Zhidkov and B.N. Khoromskij. *Some cost-effective algorithms using Toeplitz-type matrices.* In: Numerical Processes and Systems. Nauka, Moscow, 1988, 134-144, (in Russian).

[16] E.P. Zhidkov and B.N. Khoromskij. *Boundary integral equations on special surfaces and their applications.* Sov. J. Numerical Anal. Math. Modelling **2**, 6 (1988) 463-488.