Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

A Sparse $\mathcal{H}$-Matrix Arithmetic, Part II:
Application to Multi-Dimensional
Problems

by

Wolfgang Hackbusch and Boris N. Khoromskij

A Sparse $\mathcal{H}$-Matrix Arithmetic.
Part II: Application to Multi-Dimensional Problems

Wolfgang Hackbusch and Boris N. Khoromskij
Max-Planck-Institut *Mathematik in den Naturwissenschaften*
Inselstr. 22-26
D-04103 Leipzig
Germany
email: wh@mis.mpg.de

Abstract

The preceding Part I of this paper has introduced a class of matrices ($\mathcal{H}$-matrices) which are data-sparse and allow an approximate matrix arithmetic of almost linear complexity. The matrices discussed in Part I are able to approximate discrete integral operators in the case of one spatial dimension.

In the present Part II, the construction of $\mathcal{H}$-matrices is explained for FEM and BEM applications in two and three spatial dimensions. The orders of complexity of the various matrix operations are exactly the same as in Part I. In particular, it is shown that the applicability of $\mathcal{H}$-matrices does not require a regular mesh. We discuss quasi-uniform unstructured meshes and the case of composed surfaces as well.

AMS Subject Classification: 65F05, 65F30, 65F50
Key words: fast algorithms, hierarchical matrices, hierarchical block partitioning, sparse matrices, matrix inversion, BEM, FEM

1 Introduction

In Part I [6], the class of $\mathcal{H}$-matrices is introduced and it is shown that this technique provides an efficient tool for sparse hierarchical approximation to large and fully populated stiffness matrices arising in BEM (boundary element method) and FEM[1] applications. In particular, the storage, the matrix-vector multiplication and standard matrix operations like the (truncated) matrix-matrix product and matrix inversion of $\mathcal{H}$-matrices have a complexity between $O(n)$ and $O(n \log^2 n)$, where n is the problem size[2]. For example, the arithmetic of $\mathcal{H}$-matrices can be applied to Schur-complements of $\mathcal{H}$-matrices.

The construction of $\mathcal{H}$-matrices involves the same cluster tree of the underlying domain as the *panel clustering technique* (see [7] or [4, 8, 10]). The panel clustering matrix representation uses a row-wise clustering procedure and provides a matrix-vector multiplication of the complexity $O(n \log^{d+1} n)$ for boundary element problems posed in $\mathbb{R}^d$, $d = 2, 3$. However, these panel clustering matrices cannot cheaply be multiplied or inverted. For this purpose, the $\mathcal{H}$-matrices are based on a further *block*-cluster tree, which leads to a rather general block decomposition of the matrix. Such a block decomposition is discussed in Part I [6] for a regular one-dimensional mesh. Here we concentrate on the corresponding construction of $\mathcal{H}$-matrices for 2D and 3D applications. In particular, these matrices approximate dense matrices arising in 2D and 3D boundary element Galerkin/collocation methods. Our assumption (12) on the kernel is also typical for the reliability of wavelet approximation techniques (cf. [1, 2, 11]). However, different from wavelet applications, we do not require (global or piecewise) smoothness of the surface (normal direction).

The practical implementation of the $\mathcal{H}$-matrices is uniquely defined by the cluster tree and the choice of the far field condition (9). A general construction of the block-cluster trees is presented in §2 and §3. The reliability of $\mathcal{H}$-matrix approximations in BEM will be briefly discussed in §3.5. The complexity analysis is prepared in several steps. First, the case of a regular 2D tensor-product mesh is analysed, see §4. In a second step, quasi-uniform and shape-regular unstructured triangulations are admitted (§5.1). In this way,

---

[1] In the FEM case, the inverse matrix is the full one which needs a data-sparse representation.
[2] If rank-k matrices are used for the block matrices, the constant in $O(\dots)$ depends on k.

the results obtained for the tensor-product meshes are used for the construction of an asymptotically optimal approximation to the minimal admissible cluster tree on the given unstructured grid. More complicated case like manifolds composed from smooth patches are discussed in §5.2. Concerning the 3D case, we describe the $\mathcal{H}$-matrices for regular 3D meshes in §6. Generalisations to unstructured meshes are completely analogous to the 2D case. Although the same techniques can be applied also to any dimension $d > 3$, we omit this case.

The discussion of the $\mathcal{H}$-matrix arithmetic is to completed by further papers about the numerical performance, the applicability to adaptive (non-quasi-uniform) grids, and the analysis of anisotropic kernels etc.

2 The Cluster Tree

We recall the definitions of an $\mathcal{H}$-tree and of the particular partitionings introduced in [6]. Let I be a finite index set. Consider the vector space $\mathbb{K}^I$ consisting of vectors $v = (v_i)_{i \in I}$ over the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. The usual block partitioning of a vector is described by a (fixed) partitioning of I into disjoint subsets, i.e., $P = \{I_j : 1 \leq j \leq k\}$ with

$$I = \bigcup_{j=1}^{k} I_j. \tag{1}$$

In the following, we consider many different partitionings including (locally) coarse or fine partitionings. The set of these partitionings is hierarchically structured and is uniquely defined by the tree $T = T(I)$. The name $\mathcal{H}$-tree is due to its hierarchical structure. The exact description of T is given in Definition 2.1. Therein we use the notation

$$S(t) := \{s \in T : s \text{ is son of } t\} \qquad \text{for } t \in T \tag{2}$$

for the *sons* of a vertex. A *leaf* is characterised by $S(t) = \emptyset$ (or $\#S(t) = 0$).

Definition 2.1 *Let I be an index set. A tree T is called an $\mathcal{H}$-tree (based on I) if the following conditions hold:*
(i) $I \in T$.
(ii) If $t \in T$ is no leaf, $S(t)$ contains disjoint subsets of I and t is the union of its sons, i.e.,

$$t = \bigcup_{s \in S(t)} s. \tag{3}$$

We conclude that I is always the *root* of T and $t \subseteq I$ holds for all $t \in T$. Usually, the tree is constructed such that $\#S(t) \neq 1$, i.e., either t is a leaf or it has at least two sons (cf. Definition 2.1 in [6]). Note that $\#S(t) = 1$ and $s \in S(t)$ imply $s = t$ because of (3). However, in view of later theoretical constructions, we do not require $\#S(t) \neq 1$ in Definition 2.1.

Since the subsets $t \subseteq I$ which form the vertices of T, are called *clusters*, the $\mathcal{H}$-tree T is also named *cluster tree*.

As in [6], the set of all leaves is denoted by

$$\mathcal{L}(T) := \{t \in T : S(t) = \emptyset\}.$$

In the following, we restrict all partitionings to those which are built by sets contained in the tree T. For these we use the name $\mathcal{H}$-*partitioning* (or, T-*partitioning*, if we want to refer to the tree T), i.e., $P = \{I_j : 1 \leq j \leq k\}$ with (1) is a T-partitioning of I if $I_j \in T$ (or equivalently, $P \subset T$). The set of all such T-partitionings is denoted by $\mathcal{P}(T)$.

Remark 2.2 *(i) $\mathcal{P}(T) = \{\mathcal{L}(T') : T' \text{ is a subtree of } T \text{ and an } \mathcal{H}\text{-tree}\}$.*
(ii) There is a one-to-one mapping between T-partitionings and $\mathcal{H}$-subtrees T', given by $T' \mapsto P := \mathcal{L}(T') \in \mathcal{P}(T)$.

Proof. Part (i) follows from (ii). For the proof of (ii) let an $\mathcal{H}$-subtree T' be given. Then (3) ensures $I = \bigcup_{s \in \mathcal{L}(T')} s$; hence, $P := \mathcal{L}(T')$ is a T-partitioning. If a T-partitioning P is given, consider the subtree T' of T consisting of all $t \in T$ with $t \cap I_j = I_j$ or $t \cap I_j = \emptyset$ for all $I_j \in P$. ∎

Remark 2.2(ii) allows to define a partial ordering.

Definition 2.3 *If $P' = \mathcal{L}(T')$ and $P'' = \mathcal{L}(T'')$ are two T-partitionings, P' is called coarser (finer) than P'' if and only if $T' \supseteq T''$ ($T' \subseteq T''$).*

So far, we have admitted arbitrary subsets of I as vertices of the tree T. Later, each index $i \in I$ will carry a position $x_i \in \mathbb{R}^d$ (e.g., $d = 3$). We may also identify the index i with the position x_i. This allows to define a diameter

$$\operatorname{diam}(t) := \max_{i,j \in t} \|x_i - x_j\| \qquad \text{for } t \in T \tag{4}$$

using the Euclidean norm in $\mathbb{R}^d$. Since we want to have many indices in t with t possessing a diameter as small as possible, the naming *cluster* for $t \in T$ makes sense and leads to the name *cluster tree* for T. Further, we will need the distance of two clusters:

$$\operatorname{dist}(s,t) := \min_{i \in s, j \in t} \|x_i - x_j\| \qquad \text{for } s, t \in T. \tag{5}$$

Remark 2.4 *For Galerkin discretisations it is more reasonable to associate each index i with the support[3] $X_i \subset \mathbb{R}^d$ of the corresponding basis functions. We introduce the notation*

$$X(t) := \bigcup_{i \in t} X_i \qquad \text{for } t \in T. \tag{6}$$

In this case, the definitions (4) and (5) become

$$\operatorname{diam}(t) \quad = \quad \max_{x,y \in X(t)} \|x - y\| \qquad \text{for } t \in T, \tag{7}$$

$$\operatorname{dist}(s,t) \quad = \quad \min_{x \in X(t), \; y \in X(s)} \|x - y\| \qquad \text{for } s, t \in T. \tag{8}$$

The construction of the cluster tree T is the essential part of the $\mathcal{H}$-matrix construction. In §4 we describe the cluster tree for a particular two-dimensional grid. This gives rise to a more general cluster tree for general two-dimensional manifolds (§5). The three-dimensional case is discussed in §6. Other algorithms for generating the cluster tree can be considered as well (compare, e.g., Lage [9]).

3 The Block-Cluster Tree

While the vector components are indexed by $i \in I$, the matrix entries have indices from the index set $I \times I$. The block-cluster tree is nothing but the cluster tree for $I \times I$ instead of I. Its notation is $T_2 = T(I \times I)$, while we write $T_1 = T(I)$ for the previous cluster tree corresponding to I. We will describe a mapping $\tau : T_1 \mapsto T_2$ which constructs the block-cluster tree T_2 in a unique way from the cluster tree T_1 discussed above. Therefore, the block-cluster tree T_2 is fixed as soon as the cluster tree T_1 is defined. In any case, the vertices of T_2 belong to the Cartesian product $T \times T$.

The $\mathcal{H}$-matrices will be constructed on the basis of a particular (optimal) block partitioning $P_2 \subset T_2$.

3.1 Construction of T_2 from T

We describe two mappings $\tau : T \mapsto T_2$. The simpler one is given in

Construction 3.1 *Start with $I \times I \in T_2$ and define the sons of $b = (t_1, t_2) \in T_2$ (where $t_1, t_2 \in T$) recursively by*

- *(s_1, s_2) with $s_1 \in S(t_1)$, $s_2 \in S(t_2)$, provided these sons exist,*

- *(t_1, s_2) with $s_2 \in S(t_2)$, if $S(t_1) = \emptyset$ and $S(t_2) \neq \emptyset$,*

- *(s_1, t_2) with $s_1 \in S(t_1)$, if $S(t_1) \neq \emptyset$ and $S(t_2) = \emptyset$,*

- *$S_2(b) = \emptyset$ if $S(t_1) = \emptyset$ and $S(t_2) = \emptyset$.*

---

[3]Replace $X_i \subset \mathbb{R}^d$ by $X_i \subset \Gamma \subset \mathbb{R}^{d+1}$ in the case of a manifold Γ.

In the latter case, S_2 denotes the set function (2) for T_2 instead of $T_1 = T(I)$. We collect some trivial results in the next remark.

Remark 3.2 *a) The depth of the tree T_2 equals the depth of T.*
b) If all branches of T have the same length k, only the first and fourth cases of Construction 3.1 occur.
c) Assume the case of b). If T is a binary tree, then T_2 is a quadtree.

Due to Part c) of the remark, we might like to modify Construction 3.1. The second construction tries to ensure that the components $t_1, t_2 \in T$ in $b = (t_1, t_2) \in T_2$ do not have too different diameters, while the degree of the vertices $b = (t_1, t_2)$ equals the degree of either t_1 or $t_2 \in T$. This leads to the

Construction 3.3 *Start with $I \times I \in T_2$ and define the sons of $b = (t_1, t_2) \in T_2$ (where $t_1, t_2 \in T$) recursively by*

- *(s_1, t_2) with $s_1 \in S(t_1)$, if $S(t_2) = \emptyset$ or $\mathrm{diam}(t_1) \geq \mathrm{diam}(t_2)$, provided $S(t_1) \neq \emptyset$,*

- *(t_1, s_2) with $s_2 \in S(t_2)$, if $S(t_1) = \emptyset$ or $\mathrm{diam}(t_1) < \mathrm{diam}(t_2)$, provided $S(t_2) \neq \emptyset$,*

- *$S_2(b) = \emptyset$ if $S(t_1) = \emptyset$ and $S(t_2) = \emptyset$.*

Similarly, one can replace the goal $\mathrm{diam}(t_1) \approx \mathrm{diam}(t_2)$ by other options, e.g., $\#t_1 \approx \#t_2$ (i.e., t_1 and t_2 should contain a similar number of indices). In the latter case, a binary tree T_1 leads to a binary tree T_2.

It is an easy exercise to check that T_2 satisfies the conditions of Definition 2.1, i.e., in both cases we obtain an $\mathcal{H}$-partitioning of the product set $I \times I$.

3.2 Admissible Blocks, Admissible T_2-Partitionings

To guarantee a sufficient approximation, we need the *admissibility condition*

$$\min\{\mathrm{diam}(t_1), \mathrm{diam}(t_2)\} \leq 2\eta \, \mathrm{dist}\,(t_1, t_2) \tag{9}$$

for the block $b = (t_1, t_2)$. Here, $\eta < 1$ is a constant which will be fixed later (see, e.g., (16)).

Definition 3.4 *A block $b = (t_1, t_2) \in T_2$ is called* admissible *if either b is a leaf or (9) holds.*

In §2, we have introduced a T_2-partitioning of $I \times I$. It can be regarded as the set $\mathcal{L}(T')$, where T' is a subtree of T_2 with the properties $I \times I \in T'$ and (3) with respect to $I \times I$. Another name for the T_2-partitioning would be a *covering* of $I \times I$, since it is a subset $P = \{b_1, \dots, b_p\} \subset T_2$ of disjoint blocks with $\cup_{1 \leq i \leq p}\, b_i = I \times I$.

Definition 3.5 *A T_2-partitioning P of $I \times I$ is called* admissible, *if all blocks $t \in P$ are admissible.*

A trivial example for an admissible T_2-partitioning is $P = \mathcal{L}(T_2)$.

Remark 3.6 *Let $P' = \mathcal{L}(T')$ and $P'' = \mathcal{L}(T'')$ be two different admissible T_2-partitionings, then the intersection $T' \cap T''$ yields an admissible T_2-partitioning $P = \mathcal{L}(T' \cap T'')$, which is finer than P' and P'' in the sense of Definition 2.3. Furthermore, $\#P < \min\{\#P', \#P''\}$ holds for the number of blocks.*

Due to this remark, we can ask for the smallest admissible partitioning. This leads to

Definition 3.7 *The* minimal admissible *T_2-partitioning of $I \times I$ is the admissible T_2-partitioning with the minimal number of blocks.*

The minimal admissible T_2-partitioning can be obtained by a simple search in the tree T_2.

Algorithm 3.8 *The construction of the minimal admissible T_2-partitioning $P_{\min}$ of $I \times I$ is obtained as $P_{\min} := \Phi(\{I \times I\})$ with Φ from*

$$
\begin{aligned}
&function \ \Phi(P); \ comment \ P \subset T_2; \\
&begin \ P' := P; \\
&\quad for \ all \ vertices \ t \in P \ do \\
&\quad if \ t \ is \ not \ admissible \ then \ P' := (P' \backslash \{t\}) \cup \Phi(S_2(t)); \\
&\quad \Phi := P' \\
&end;
\end{aligned}
\tag{10}
$$

Proof. The proof that (10) yields the minimal admissible T_2-partitioning is based on the following observation: If $b = (t_1, t_2) \in T_2$ is admissible, then also all sons of t are admissible. This is due to the fact that the left-hand side in (9) weakly decreases if t_1 or t_2 are replaced by the (smaller) sons, while the right-hand side dist(t_1, t_2) weakly increases. ∎

Remark 3.9 *If one likes to replace the admissibility condition (9) by another condition Adm(b) (a Boolean-valued function), one should ensure that $Adm(b) \implies Adm(s)$ holds for all $s \in S_2(b)$.*

Remark 3.10 *The minimal admissible T_2-partitioning is coarser (cf. Definition 2.3) than any other admissible T_2-partitioning.*

3.3 Complexity Considerations

In [6] we described two particular partitionings $P_2 \subset T_2$. Similarly, we will describe a T_2-partitioning in §4. This partitioning is admissible (the minimality is not discussed but can be shown if η is of appropriate size). For the fixed partitioning, one can study the complexity of the various arithmetical operations.

In the general case, the T_2-partitioning is determined as the minimal admissible T_2-partitioning resulting from Algorithm 3.8. In order to ensure the desired complexity, it is sufficient to prove the complexity for *some* admissible T_2-partitioning. The existence of such a T_2-partitioning is sufficient, a constructive description is not needed. The proof uses Remark 3.10.

Lemma 3.11 *Assume that (i) the computational work increases if the T_2-partitioning becomes finer (cf. Definition 2.3), (ii) the complexity of some admissible T_2-partitioning is known. Then the complexity of the minimal admissible T_2-partitioning is at least as good.*

3.4 Hierarchical $\mathcal{H}$-Matrices

In the following definition, P_2 is a general T_2-partitioning of $I \times I$, although in practical applications we shall use only *admissible* T_2-partitionings P_2. Each $b \in P_2$ corresponds to a location of a matrix block. Given a matrix $M = (m_{ij})_{(i,j) \in I \times I} \in \mathbb{K}^{I \times I}$, the matrix block corresponding to b is denoted by $M^b = (m_{ij})_{(i,j) \in b}$.

Definition 3.12 *Let P_2 be a block partitioning of $I \times I$ and $k \in \mathbb{N}$. The underlying field of the vector space of matrices is $\mathbb{K}$. The set of $\mathcal{H}$-matrices induced by P_2 is*

$$\mathcal{M}_{\mathcal{H},k}(I \times I, P_2) := \{ M \in \mathbb{K}^{I \times I} : \text{ each block } M^b, b \in P_2, \text{ satisfies } \operatorname{rank}(M^b) \leq k \}. \tag{11}$$

We call a matrix A an *Rk-matrix* if $\operatorname{rank}(A) \leq k$. The properties of Rk- and, in particular, $R1$-matrices are discussed in [6].

Remark 3.13 *All considerations about $\mathcal{H}$-matrices do not refer to a special ordering of the unknowns. The index set is allowed to possess no ordering at all. Only if we visualise the block partitioning as in §4.3, we introduce a numbering of the blocks.*

Remark 3.14 *In Definition 3.12 the upper bound of the rank is assumed to be the same for all submatrices. One may consider variable bounds. Then k is a function $k : P_2 \to \mathbb{N}$ of the block and the inequality in (11) becomes $\operatorname{rank}(M^b) \leq k(b)$. For the sake of simplicity, we regard k as a constant for the rest of the paper.*

3.5 Approximation by $\mathcal{H}$-Matrices

The reliability of $\mathcal{H}$−matrices for the approximation of the integral operators

$$(Au)(x) = \int_\Sigma k(x,y)u(y)dy, \qquad x \in \Sigma,$$

is essentially based on smoothness properties of the kernel[4] $k(x,y)$. In the boundary element method, integral operators occur with $k(x,y)$ being Green's function associated with the partial differential equation under consideration or with $k(x,y)$ replaced by a suitable directional derivatives Dk of $k(x,y)$. Here Σ is either a bounded d-dimensional manifold (surface) $\Gamma \subset \mathbb{R}^{d+1}$ or a bounded domain Ω in $\mathbb{R}^d$, $d = 2, 3$. The single layer

---

[4]Note that the rank $k \in \mathbb{N}$ and the kernel function $k(x,y)$ are both written as k.

potential for the Laplace equation in $\mathbb{R}^3$ gives the familiar example $k(x,y) := \frac{1}{4\pi}|x-y|^{-1}$ for $x,y \in \Sigma$. The smoothness of $k(x,y)$ with respect to both x and y depends in a typical manner on the distance $|x-y|$. Note that both the panel clustering method and the $\mathcal{H}$-partitioning approach exploit only the approximation of $k(x,y)$ by a degenerate kernel (cf. [5, Definition 3.3.3]). This holds for $k(x,y)$ as well as for $\partial k(x,y)/\partial n(x)$ or $\partial k(x,y)/\partial n(y)$ (double layer kernel and its adjoint; cf. [5, (8.1.31a,b)]) even if the normal direction n is nonsmooth because of the non-smoothness of the surface Γ, since only the smoothness properties of the singularity function $k(x,y)$ are involved. More precisely, we assume that the singularity function $k(x,y)$ satisfies[5]

$$|\partial_x^\alpha \partial_y^\beta k(x,y)| \le c(|\alpha|,|\beta|)|x-y|^{-|\alpha|-|\beta|}|k(x,y)| \qquad \text{for all } \alpha,\beta \in \mathbb{N}_0^d, \ x,y \in \mathbb{R}^d, \tag{12}$$

where α,β are multi-indices with $|\alpha| = \alpha_1 + \ldots + \alpha_d$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Note that similar assumptions are usually required in the wavelet or multi-resolution technique (cf. [1, 2, 11]).

By Definition 3.12, $\mathcal{H}$-matrices consist locally (blockwise) of rank-k matrices. As in the panel clustering method, these low rank matrices can be constructed via a Taylor expansion[6] of $k(x,y)$. Let x,y vary in the respective sets $X(t_x)$ and $X(t_y)$ (cf. (6)) corresponding to the clusters $t_x, t_y \in T$ and assume without loss of generality that $\text{diam}(X(t_y)) \le \text{diam}(X(t_x))$. The optimal centre of expansion is the Chebyshev centre[7] y_* of $X(t_y)$, since then $\|y - y_*\| \le \frac{1}{2}\text{diam}(X(t_y))$ for all $y \in X(t_y)$. The Taylor expansion reads $k(x,y) = \widetilde{k}(x,y) + R$ with the polynomial

$$\widetilde{k}(x,y) = \sum_{|\nu|=0}^{m-1} \frac{1}{\nu!}(y_* - y)^\nu \frac{\partial^\nu k(x,y_*)}{\partial y^\nu} \tag{13}$$

and the remainder R, which can be estimated by

$$|R| = |k(x,y) - \widetilde{k}(x,y)| \le \frac{1}{m!}\|y_* - y\|^m \max_{\zeta \in X(t_y), \ |\gamma|=m} \left| \frac{\partial^\gamma k(x,\zeta)}{\partial \zeta^\gamma} \right|. \tag{14}$$

Lemma 3.15 *Assume (12) and (9) involving the sufficiently small parameter $\eta < 1$. Then for $m \ge 1$, the remainder (14) satisfies the estimate*

$$|k(x,y) - \widetilde{k}(x,y)| \le c(m)\eta^m |k(x,y)| \qquad \text{for } x \in X(t_x), \ y \in X(t_y). \tag{15}$$

Proof. The estimate $\|y - y_*\| \le \frac{1}{2}\text{diam}(X(t_y))$ for $y \in X(t_y)$ is already stated. All $x \in X(t_x)$ and $y, \zeta \in X(t_y)$ satisfy $\|x - \zeta\| \ge \text{dist}(X(t_x), X(t_y)) \ge \frac{1}{2\eta}\min\{\text{diam}(X(t_x)), \text{diam}(X(t_y))\} = \frac{1}{2\eta}\text{diam}(X(t_y)) \ge \frac{1}{\eta}\|y - y_*\|$. Therefore,

$$|R| = |k(x,y) - \widetilde{k}(x,y)| \le \frac{c(0,m)}{m!} \frac{\|y_* - y\|^m |k(x,\zeta)|}{\|x-\zeta\|^m} \le \frac{c(0,m)|k(x,\zeta)|}{m!}\eta^m$$

for some $\zeta \in X(t_y)$. In the upper estimate, we may choose ζ with $\max\{|k(x,\eta)| : \eta \in t_y\} = |k(x,\zeta)|$ and use $\|k(x,y)| - |k(x,\zeta)\| \le |k(x,y) - k(x,\zeta)| \le |y - \zeta| \left| \frac{\partial k(x,\tilde{\zeta})}{\partial y} \right| \le c(0,1)(|y - \zeta|/|x - \tilde{\zeta}|)|k(x,\tilde{\zeta})|$ for some $\tilde{\zeta}$ between y and ζ. Together with $|y - \zeta|/|x - \tilde{\zeta}| \le \frac{1}{2}\text{diam}\, X(t_y)/\text{dist}(X(t_x), X(t_y)) \le \eta$, we conclude that $|k(x,\zeta)| \le |k(x,y)|/(1 - c(0,1)\eta)$. Hence, (15) holds with $c(m) := c(0,m)/(m!(1 - c(0,1)\eta))$. ∎

Let $\widetilde{A}$ be the integral operator with $k(x,y)$ replaced by $\widetilde{k}(x,y)$, provided that $(t_x, t_y) \in T_2$ is an admissible block and no leaf (i.e., (9) holds). Construct the collocation or Galerkin system matrix from $\widetilde{A}$ instead of A. The perturbation of the matrix induced by $\widetilde{A} - A$ yields a perturbed discrete solution. The effect of this perturbation is studied in several papers on the panel clustering method (cf. [7], [10]). Perturbations in the case of a negative order operator A are considered in [3].

---

[5]Estimate (12) is a bit simplified. It covers most of the situations, e.g., the case of the singularity function $\frac{1}{4\pi}|x-y|^{-1}$ for $d = 3$. As soon as logarithmic terms appear (as for $d = 2$; $k(x,y) = \log(x-y)/2\pi$), one has to modify (12).

[6]This does not require that the practical implementation has to use the Taylor expansion. If the singular-value decomposition technique from [6] is applied, the estimates are at least as good as the particular ones for the Taylor expansion.

[7]Given a set X, the Chebyshev sphere is the minimal one containing X. Its centre is called the *Chebyshev centre*.

3.6 On the Choice of η and m

In order to obtain a small error, $\eta^m \le \varepsilon < 1$ must be ensured for a suitable ε. The rank k corresponding to the expansion (13) equals $k = \#\{\nu \in \mathbb{N}_0^d : 0 \le |\nu| \le m-1\} \le m^d$. We may fix m (and k) and choose $\eta \le \varepsilon^{1/m}$. On the other hand, η can be fixed while the polynomial degree m is chosen: $m \ge \log \varepsilon / \log \eta$. The first case reminds to the h-version of the FEM, while the latter corresponds to the p-version. The optimal choice is determined by the arising cost of the $\mathcal{H}$-matrix operations. In the case of quasi-uniform meshes, one may conclude from [7] (therein (3.9a)) that the number of admissible clusters $(t_1, t_2) \in P_2$ on each level ℓ may be estimated by $O(\eta^{-d} 2^{d\ell})$ (see also §4.1). This result implies that the leading term in the cost has a factor proportional to $\eta^{-d} m^d \sim \eta^{-d} k$. Hence, one has to minimise $\eta^{-d} m^d$ under the side condition $\eta^m = \varepsilon$. Allowing for simplicity real-valued m, the result is

$$m = |\log \varepsilon| \text{ and } \eta = 1/e. \tag{16}$$

The constant value of η expresses the fact that the choice (16) corresponds to the p-version.

4 The Two-Dimensional Model Case

In $\Omega = (0,1) \times (0,1)$ we consider the regular grid

$$I = \{(i,j) : 1 \le i, j \le N\}, \quad N = 2^p. \tag{17}$$

Each index $(i,j) \in I$ is associated with the (collocation) point $\xi_{ij} = ((i-\frac{1}{2})h, (j-\frac{1}{2})h) \in \mathbb{R}^2$, where $h := 1/N$. The positions ξ_{ij} are used in (4) and (5).

4.1 The Cluster Tree $T_1 = T(I)$

The natural partitioning of I uses a division of the underlying squares into four quarters. The clusters

$$t_{\alpha,\beta}^\ell := \{(i,j) : \ 2^{p-\ell}\alpha + 1 \le i \le 2^{p-\ell}(\alpha+1), \ 2^{p-\ell}\beta + 1 \le j \le 2^{p-\ell}(\beta+1)\} \tag{18}$$

with $\alpha, \beta \in \{0, \ldots, 2^\ell - 1\}$ belong to level ℓ. Hence, the tree T consisting of all clusters of level $\ell \in \{0, \ldots, p\}$ is a quadtree. The number of clusters on level ℓ equals $O(2^{2\ell})$.

Each index $(i,j) \in I$ is associated with the square[8]

$$X_{ij} := \{(x,y) : (i-1)h \le x \le ih, (j-1)h \le y \le jh\}, \tag{19}$$

which may be regarded as the support of the piecewise constant function for the index (i,j). Note that on level $\ell = 0$ ($t_{00}^0 = I$) we have one big square $X(t_{00}^0)$ (cf. (6)), while for $\ell = p$ we have 4^p tiny squares $X(t_{\alpha,\beta}^p)$. Using the definitions (7) and (8), we obtain the diameter

$$\mathrm{diam}(t) = \sqrt{2}\, 2^{p-\ell} h = \sqrt{2}/2^\ell \tag{20}$$

for clusters of level ℓ. Let t, t' be two clusters of level ℓ characterised by (α, β) and (α', β') (cf. (18)). Then

$$\mathrm{dist}(t, t') = 2^{-\ell}\sqrt{\delta(\alpha - \alpha')^2 + \delta(\beta - \beta')^2} \qquad \text{with } \delta(k) := \max\{0, |k| - 1\}. \tag{21}$$

4.2 The Block-Cluster Tree $T_2 = T(I \times I)$

Let $T_2 = T(I \times I)$ be defined according to Construction 3.1. An obvious result is stated in

Remark 4.1 *Let $b = (t_1, t_2) \in T(I \times I)$. Then $t_1, t_2 \in T$ belong to the same level $\ell \in \{0, \ldots, p\}$.*

Using $\min\{\mathrm{diam}(t_1), \mathrm{diam}(t_2)\} = \sqrt{d}/2^\ell$ and $\mathrm{dist}(t_1, t_2)$ from (21), we observe that $b \in T(I \times I)$ is admissible for the choice $2\eta = \sqrt{2}$, if the squares $t_1, t_2 \in T(I)$ have a relative position as indicated in Fig. 1a: The square X_1 corresponding to t_1 is the crossed square, while X_2 must be outside the bold area. In the case of $d = 2$ and $\eta = 1/\sqrt{2}$, the admissible T_2-partitioning P_2 is described in the following subsection.

---

[8] The grid can also be associated with a regular *triangulation* and, e.g., the supports X_{ij} of piecewise linear functions. This would lead to another cluster tree. The asymptotic complexity bounds turn out to be the same as for the present choice.
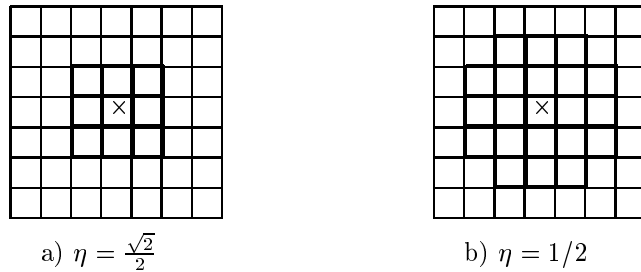
Figure 1: Unacceptable clusters for a given cluster "×" depending on the threshold constant η

4.3 $\mathcal{H}$-Matrices

The admissible T_2-partitioning P_2 is to be defined. Because of the regular structure of the grid, the block partitioning $P_2 = \mathcal{L}(T_2')$ corresponds to a well-structured subtree $T_2' \subset T_2$. This allows the direct constructive definition of the T_2-partitioning P_2 and the corresponding $\mathcal{H}$-matrices based on the following recursive procedure (cf. [6, Subsection 2.3]).

In [6, Section 5], we have introduced $\mathcal{H}$-matrices which could be explained by the three formats $\mathcal{H}$ (diagonal format), $\mathcal{N}$ (right-neighbour format) and $\mathcal{N}^*$ (left-neighbour format). Now the diagonal format is denoted by the symbol $\square$ and instead of two 'neighbour formats' we have eight types denoted by the directions $\rightarrow, \leftarrow, \uparrow, \downarrow, \searrow, \swarrow, \nwarrow, \nearrow$.

For any $t \in T_1$, $\mathcal{H}$-matrices over the index set $t \times t$ have the $\square$-format defined as follows. If t is a leaf (level $\ell = p$, $\#t = 1$), the matrix is 1×1. Otherwise, t has four sons s_i ($1 \le i \le 4$). The related square $X(t)$ splits into the four smaller squares $X_i := X(s_i)$. In the following visualisations, the 4 sons of a square are numbered as follows: $\begin{array}{|c|c|} \hline X_1 & X_2 \\ \hline X_4 & X_3 \\ \hline \end{array}$. Correspondingly, the matrix $A_\square$ over $t \times t$ has a 4×4-block structure:

$$A_\square = \begin{array}{|c|c|c|c|} \hline b_{11} & b_{12} & b_{13} & b_{14} \\ \hline b_{21} & b_{22} & b_{23} & b_{24} \\ \hline b_{31} & b_{32} & b_{33} & b_{34} \\ \hline b_{41} & b_{42} & b_{43} & b_{44} \\ \hline \end{array} = \begin{array}{|c|c|c|c|} \hline \square & \rightarrow & \searrow & \downarrow \\ \hline \leftarrow & \square & \downarrow & \swarrow \\ \hline \nwarrow & \uparrow & \square & \leftarrow \\ \hline \uparrow & \nearrow & \rightarrow & \square \\ \hline \end{array} \tag{22}$$

If t is of level $\ell = p - 1$, all blocks b_{ij} are of level p and of trivial size 1×1. In the following, we assume that the b_{ij} are nontrivial, i.e., $\ell < p - 1$.

All squares $X(s_i)$ and $X(s_j)$ touch by at least one corner point, hence $\text{dist}(X(s_i), X(s_j)) = 0$. Therefore, the vertices $(s_i, s_j) \in T_2$ are not admissible and deserve a further decomposition. The type of block decomposition depends on the relative position of s_i, s_j.

The diagonal blocks b_{ii} ($1 \le i \le 4$) belong to the index pairs (s_i, s_i) and have again format $\square$.

The block b_{12} has a block format denoted by the arrow $\rightarrow$ directing from s_1 to the right neighbour s_2.

The squares $X(s_1)$ and $X(s_3)$ are diagonally neighboured. The corresponding symbol of block b_{13} is $\searrow$.

The block b_{14} corresponds to the squares $X(s_1)$ and $X(s_4)$ (the latter one is situated below the former one). This leads to the $\downarrow$-format.

Similarly, the formats of b_{ij} ($i \ge 2$) are determined (see (22)).

Next, we have to describe the formats different from $\square$. We start with the $\rightarrow$-format:

$$A_\rightarrow = \begin{array}{|c|c|c|c|} \hline b_{a\alpha} & b_{a\beta} & b_{a\gamma} & b_{a\delta} \\ \hline b_{b\alpha} & b_{b\beta} & b_{b\gamma} & b_{b\delta} \\ \hline b_{c\alpha} & b_{c\beta} & b_{c\gamma} & b_{c\delta} \\ \hline b_{d\alpha} & b_{d\beta} & b_{d\gamma} & b_{d\delta} \\ \hline \end{array} = \begin{array}{|c|c|c|c|} \hline R & R & R & R \\ \hline \rightarrow & R & R & \searrow \\ \hline \nearrow & R & R & \rightarrow \\ \hline R & R & R & R \\ \hline \end{array}. \tag{23}$$

The matrix $A_\rightarrow$ corresponds to the index pair $(s, s') \in T_2$, where $X(s')$ is the right neighbouring square of $X(s)$. The sons $\{a, b, c, d\}$ of s and the sons $\{\alpha, \beta, \gamma, \delta\}$ of s' correspond to the squares situated as follows: $\begin{array}{|c|c|c|c|} \hline a & b & \alpha & \beta \\ \hline d & c & \delta & \gamma \\ \hline \end{array}$. The squares $X(a)$ and $X(\alpha)$ satisfy $\text{diam}(a) = \text{diam}(\alpha) = \sqrt{2}h_{\ell-1}$ ($h_{\ell-1} = 2^{1-\ell}$) and $\text{dist}(a, \alpha) = h_{\ell-1}$. Hence, (9) holds with $\eta = \sqrt{2}/2$ and the pair $(a, \alpha) \in T_2$ is admissible. By definition, the block $b_{a\alpha}$ can be represented by an Rk-matrix (this format is denoted by 'R'). A different situation arises

for $(b, \alpha) \in T_2$, where $X(\alpha)$ is the direct right neighbouring square of b. Therefore, the block $b_{b\alpha}$ has the $\rightarrow$-format. The complete result is described in (23).

The format

$$A_{\leftarrow} = \begin{array}{|c|c|c|c|} \hline R & \leftarrow & \swarrow & R \\ \hline R & R & R & R \\ \hline R & R & R & R \\ \hline R & \nwarrow & \leftarrow & R \\ \hline \end{array} \tag{24}$$

is transposed to (23). Similarly,

$$A_{\uparrow} = \begin{array}{|c|c|c|c|} \hline R & R & \nearrow & \uparrow \\ \hline R & R & \uparrow & \nwarrow \\ \hline R & R & R & R \\ \hline R & R & R & R \\ \hline \end{array}, \quad A_{\downarrow} = \begin{array}{|c|c|c|c|} \hline R & R & R & R \\ \hline R & R & R & R \\ \hline \swarrow & \downarrow & R & R \\ \hline \downarrow & \searrow & R & R \\ \hline \end{array}. \tag{25}$$

The format of $A_{\nearrow}$ is even simpler. The blocks $b_{a\alpha}, \ldots$ correspond to pairs of squares situated as follows:

$$\begin{array}{cc} \begin{array}{|c|c|} \hline \alpha & \beta \\ \hline \delta & \gamma \\ \hline \end{array} & \\ \begin{array}{|c|c|} \hline a & b \\ \hline d & c \\ \hline \end{array} & \end{array}$$

. Only the (b, δ)-block leads to $\mathrm{dist}(X(b), X(\delta)) = 0$ and requires a further decomposition. All other pairs of squares have a sufficiently large distance; therefore, those blocks $b_{a\alpha}, \ldots$ are defined to be Rk-matrices:

$$A_{\nearrow} = \begin{array}{|c|c|c|c|} \hline R & R & R & R \\ \hline R & R & R & \nearrow \\ \hline R & R & R & R \\ \hline R & R & R & R \\ \hline \end{array}, \quad A_{\swarrow} = \begin{array}{|c|c|c|c|} \hline R & R & R & R \\ \hline R & R & R & R \\ \hline R & R & R & R \\ \hline R & \swarrow & R & R \\ \hline \end{array}. \tag{26}$$

Similarly,

$$A_{\nwarrow} = \begin{array}{|c|c|c|c|} \hline R & R & \nwarrow & R \\ \hline R & R & R & R \\ \hline R & R & R & R \\ \hline R & R & R & R \\ \hline \end{array}, \quad A_{\searrow} = \begin{array}{|c|c|c|c|} \hline R & R & R & R \\ \hline R & R & R & R \\ \hline \searrow & R & R & R \\ \hline R & R & R & R \\ \hline \end{array}. \tag{27}$$

The recursions (22)-(27) define a subtree T_2' of T_2. The root $I \times I$ is of type $\square$ (level $\ell = 0$) and has 16 sons (the 16 blocks of $A_{\square}$). According to (22), 4 sons are of type $\square$ (level $\ell - 1$), 2 sons of each of the types $\rightarrow, \leftarrow, \uparrow, \downarrow$ and 1 son of each of the types $\searrow, \swarrow, \nwarrow, \nearrow$ (see Fig. 2a). A vertex of type $\rightarrow$ has 12 sons of rank-k-type (R), 2 sons of type $\rightarrow$ and 1 son of each of the types $\searrow, \nearrow$ (see Fig. 2b). Figs. 2c-d show the tree structure for the remaining types. The leaves of the subtree T_2' are reached if the vertex has type R (i.e., condition (9) satisfied) or if level p is reached (blocks of size 1×1). In this particular situation, the T_2-partitioning $P_2 = \mathcal{L}(T_2')$ is the minimal admissible partitioning which also results from Algorithm 3.8 for the choice $\eta = \sqrt{2}/2$.

Figure 2 gives rise to the graph of Fig. 3, whose vertices are the formats. This graph is a tree except the cycles induced by the edges of all formats $\neq R$ to itself. The edges are weighted by the multiplicity already shown in Fig. 2. The discussion in the next Subsection will demonstrate the following remark.

Remark 4.2 *a) For the following complexity considerations it is essential that only the format $\square$ has a self-reference with weight 4, whereas all other weights are ≤ 3.*
b) The complexity order does not depend on the number of different formats. For instance, choosing η smaller than $1/\sqrt{2}$ (as in Fig. 1b) one would need more formats, but again only format $\square$ has a self-reference with weight 4.

Formally, the recursions (22)-(27) must be used to define the T_2-partitioning P_2, while in a second step the $\mathcal{H}$-matrix set $\mathcal{M}_{\mathcal{H},k}(I \times I, P_2)$ is defined by Definition 3.12. Instead, we can give a direct definition of the
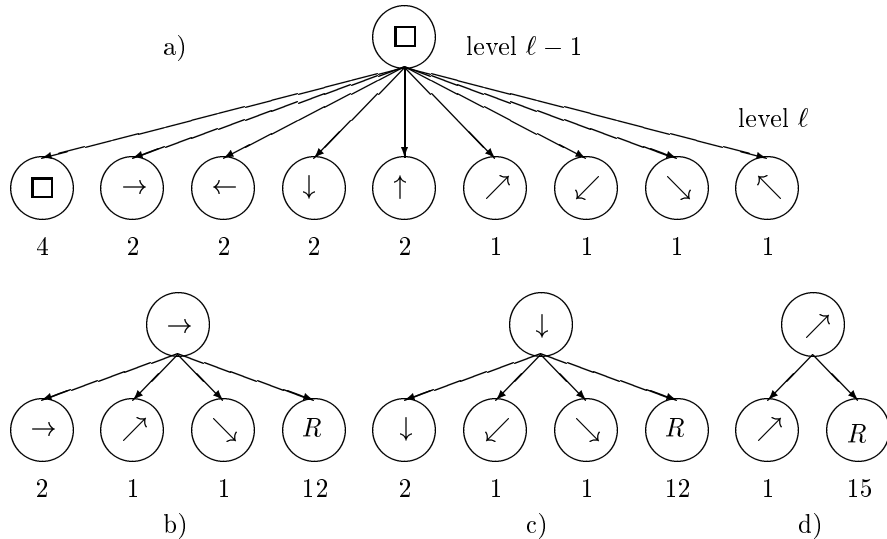
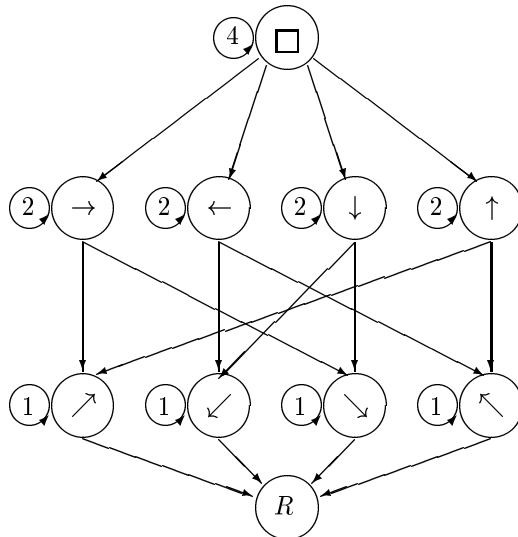Figure 2: The subtrees of the diagonal and typical auxiliary formats

Figure 3: The graph of the involved formats

matrix sets $\mathcal{M}_{\ell,p}^*$ with upper index $* \in \{R, \square, \rightarrow, \leftarrow, \uparrow, \downarrow, \searrow, \swarrow, \nwarrow, \nearrow\}$ and level number $0 \le \ell \le p \in \mathbb{N}_0$ (note that by (17), the index set I depends on p). First we define

$$\mathcal{M}_{\ell,p}(t_1, t_2) := \mathbb{K}^{t_1 \times t_2}, \qquad \text{where } t_1, t_2 \in T_1 \text{ belong to level } \ell,$$

i.e., $i \in t_1$ are the row indices and $j \in t_2$ the column indices of $A \in \mathcal{M}_\ell(t_1, t_2)$ (note that by Remark 4.1, the block matrices are of this kind). For $\ell = 0$, $t_1 = t_2 = I$ is the only vertex of that level, but in general $t_1 \ne t_2$ is possible. The level-ℓ-matrices and the corresponding Rk-matrices are denoted by

$$\mathcal{M}_{\ell,p} := \{A \in \mathcal{M}_{\ell,p}(t_1, t_2) : t_1, t_2 \in T_1 \text{ belong to level } \ell\},$$
$$\mathcal{M}_{\ell,p}^R := \{A \in \mathcal{M}_{\ell,p} : \text{rank}(A) = k\}.$$

Here $k \ge 1$ is fixed. The following recursive definition starts from $\ell = p$ and ends with $\ell = 0$. Since $\#t_1 = \#t_2 = 1$ for level $\ell = p$, we have that

$$\mathcal{M}_{p,p}^* \text{ is the set of } 1 \times 1\text{-matrices for all } * \in \{R, \square, \rightarrow, \leftarrow, \uparrow, \downarrow, \searrow, \swarrow, \nwarrow, \nearrow\}.$$

10

For $\ell < p$, the sons $S(t_1) = \{a, b, c, d\}$, $S(t_2) = \{\alpha, \beta, \gamma, \delta\}$ of the vertices t_1, t_2 are assumed to have the geometric constellation as described in the beginning of this Subsection (i.e., b $[\beta]$ is the right neighbour of a $[\alpha]$, etc.).

- Definition of $\mathcal{M}_{\ell,p}^{\nearrow}$: For $\ell < p$, a matrix $A \in \mathcal{M}_{\ell,p}(t_1, t_2)$ belongs to $\mathcal{M}_{\ell,p}^{\nearrow}$, if its block matrices in $A = \{A_{ij}\}_{i \in \{a,b,c,d\}, j \in \{\alpha,\beta,\gamma,\delta\}}$ satisfy $A_{b,\delta} \in \mathcal{M}_{\ell+1,p}^{\nearrow}$ and $A_{ij} \in \mathcal{M}_{\ell+1,p}^R$, otherwise (cf. (26)).

- Similarly, $\mathcal{M}_{\ell,p}^{\searrow}, \mathcal{M}_{\ell,p}^{\nwarrow}, \mathcal{M}_{\ell,p}^{\swarrow}$ are defined (cf. (26), (27)).

- Definition of $\mathcal{M}_{\ell,p}^{\rightarrow}$: For $\ell < p$, a matrix $A \in \mathcal{M}_{\ell,p}(t_1, t_2)$ belongs to $\mathcal{M}_{\ell,p}^{\rightarrow}$, if its block matrices in $A = \{A_{ij}\}_{i \in \{a,b,c,d\}, j \in \{\alpha,\beta,\gamma,\delta\}}$ satisfy $A_{b\alpha}, A_{c\delta} \in \mathcal{M}_{\ell+1,p}^{\rightarrow}$, $A_{c\alpha} \in \mathcal{M}_{\ell+1,p}^{\nearrow}$, $A_{b\delta} \in \mathcal{M}_{\ell+1,p}^{\searrow}$, and $A_{ij} \in \mathcal{M}_{\ell+1,p}^R$ otherwise (cf. (23)).

- Similarly, $\mathcal{M}_{\ell,p}^{\leftarrow}$, $\mathcal{M}_{\ell,p}^{\downarrow}$, $\mathcal{M}_{\ell,p}^{\uparrow}$ are defined.

- Definition of $\mathcal{M}_{\ell,p}^{\square}$: For $\ell < p$, a matrix $A \in \mathcal{M}_{\ell,p}(t_1, t_1)$ belongs to $\mathcal{M}_{\ell,p}^{\square}$, if its block matrices in $A = \{A_{ij}\}_{i,j \in \{a,b,c,d\}}$ satisfy $A_{ii} \in \mathcal{M}_{\ell+1,p}^{\square}$, $A_{ab}, A_{dc} \in \mathcal{M}_{\ell+1,p}^{\rightarrow}$, $A_{ba}, A_{cd} \in \mathcal{M}_{\ell+1,p}^{\leftarrow}$, $A_{ad}, A_{bc} \in \mathcal{M}_{\ell+1,p}^{\downarrow}$, $A_{ac} \in \mathcal{M}_{\ell+1,p}^{\searrow}$, $A_{ca} \in \mathcal{M}_{\ell+1,p}^{\nwarrow}$, $A_{bd} \in \mathcal{M}_{\ell+1,p}^{\swarrow}$, $A_{db} \in \mathcal{M}_{\ell+1,p}^{\nearrow}$.

Then $\mathcal{M}_{\mathcal{H},k}(I \times I, P_2) = \mathcal{M}_{0,p}^{\square}$ holds.

When using the grid (17) for difference or finite element discretisations of differential equations, we obtain a five-, seven-, or nine-point formula as discretisation matrix. The next lemma implies that such a matrix can be *exactly* represented by an $\mathcal{H}$-matrix (see also the later Lemma 5.7).

Lemma 4.3 *If the matrix A has a nine-point or an even sparser pattern, it is in the set $\mathcal{M}_{\mathcal{H},k}(I \times I, P_2)$ for any $k \geq 1$.*

Proof. By definition, a nine-point matrix has non-zero entries only for index pairs $(p, q) \in I \times I$, where $X_p \cap X_q \neq \emptyset$ holds for the squares introduced in (19). Let $b = (t_1, t_2) \in P_2 \subset I \times I$ be the block of the partitioning with $(p, q) \in b$. The previous characterisation yields $\mathrm{dist}(t_1, t_2) \leq \mathrm{dist}(X_p, X_q) = 0$. Hence, condition (9) cannot be satisfied. Since b belongs to an admissible partitioning, it must be leaf, i.e., it is a 1×1 block. Obviously, a 1×1 block represents the matrix entry A_{pq} exactly. ∎

4.4 Complexity

In the following, we discuss the storage requirements $\mathcal{N}_{st}^{\square}$ and the cost $\mathcal{N}_{MV}^{\square}$ of the matrix-vector multiplication. The complexity discussion for the format $\square$ first requires the study of the expenses for the other formats. Since matrices from $\mathcal{M}_{\ell,p}^{\nearrow}, \mathcal{M}_{\ell,p}^{\searrow}, \mathcal{M}_{\ell,p}^{\nwarrow}, \mathcal{M}_{\ell,p}^{\swarrow}$ behave similarly, we denote their format by the collective symbol "$\times$" (diagonal neighbourhood), while "$+$" refers to $\mathcal{M}_{\ell,p}^{\rightarrow}, \mathcal{M}_{\ell,p}^{\leftarrow}, \mathcal{M}_{\ell,p}^{\downarrow}, \mathcal{M}_{\ell,p}^{\uparrow}$.

Note that the maximal level number p does not exceed $O(|\log h|)$. The rank number k is chosen to be $k = 1$, in order to present concrete constants in the leading terms.

4.4.1 Storage

Below, the number $\mathcal{N}_{st}^*(p)$ describes the storage requirements of an matrix $A \in \mathcal{M}_{0,p}^*$ of the format $* \in \{\square, +, \times\}$.

Lemma 4.4 *Let $k = 1$ and $n = \#I = 4^p$. The storage size of matrices of the different formats amounts to*

$$
\begin{aligned}
\mathcal{N}_{st}^{\square}(p) &= (1 + 54p)n + O(p), \\
\mathcal{N}_{st}^+(p) &= 22n + O(1), \\
\mathcal{N}_{st}^{\times}(p) &= 10n + O(1).
\end{aligned}
$$

Proof. Note that $\mathcal{N}_{st}^{R1}(p) = 2n = 2 \cdot 4^p$. Due to Definition 3.12, we obtain the recurrence formulae

$$
\begin{aligned}
\mathcal{N}_{st}^{\times}(p) &= \mathcal{N}_{st}^{\times}(p-1) + 15\mathcal{N}_{st}^{R1}(p-1), \\
\mathcal{N}_{st}^+(p) &= 2\mathcal{N}_{st}^+(p-1) + 2\mathcal{N}_{st}^{\times}(p-1) + 12\mathcal{N}_{st}^{R1}(p-1), \\
\mathcal{N}_{st}^{\square}(p) &= 4\mathcal{N}_{st}^{\square}(p-1) + 8\mathcal{N}_{st}^+(p-1) + 4\mathcal{N}_{st}^{\times}(p-1),
\end{aligned}
\tag{28}
$$

with starting value $\mathcal{N}_{st}(0) = 1$ for all formats. The first equation in (28) implies $\mathcal{N}_{st}^{\times}(p) = 10n - 9$. Inserting this result into the second recurrence yields $\mathcal{N}_{st}^{+}(p) = 22n + O(1)$. Therefore, the last recurrence becomes $\mathcal{N}_{st}^{\square}(p) = 4\mathcal{N}_{st}^{\square}(p - 1) + 54n + O(1)$. Its solution is $\mathcal{N}_{st}^{\square}(p) = (1 + 54p)n + O(p)$. ∎

4.4.2 Matrix-Vector Multiplication

Lemma 4.5 *The cost for the matrix-vector multiplication is*

$$
\begin{aligned}
\mathcal{N}_{MV}^{\square}(p) &= (1 + 82p)n + O(p), \\
\mathcal{N}_{MV}^{+}(p) &= 30n + O(1), \\
\mathcal{N}_{MV}^{\times}(p) &= 19n + O(1).
\end{aligned}
\tag{29}
$$

Proof. We recall $\mathcal{N}_{MV}^{R1}(p) = 3n$. Consider type '×'. The multiplication of the 16 blocks at level $p - 1$ with the (partial) vector costs $\mathcal{N}_{MV}^{\times}(p - 1) + 15\mathcal{N}_{MV}^{R1}(p - 1)$. The summation of the results costs $3n$ additions. This leads to $\mathcal{N}_{MV}^{\times}(p) = \mathcal{N}_{MV}^{\times}(p - 1) + 15\mathcal{N}_{MV}^{R1}(p - 1) + 3n = \mathcal{N}_{MV}^{\times}(p - 1) + \frac{57}{4}n_p$ and $\mathcal{N}_{MV}^{\times}(0) = 1$. Its solution is $\mathcal{N}_{MV}^{\times}(p) = 19\,n - 18$.

Similarly, $\mathcal{N}_{MV}^{+}(p) = 2\mathcal{N}_{MV}^{+}(p - 1) + 2\mathcal{N}_{MV}^{\times}(p - 1) + 12\mathcal{N}_{MV}^{R1}(p - 1) + 3n$ yields $\mathcal{N}_{MV}^{+}(p) = 30n + O(1)$. Finally, $\mathcal{N}_{MV}^{\square}(p) = 4\mathcal{N}_{MV}^{\square}(p - 1) + 8\mathcal{N}_{MV}^{+}(p - 1) + 4\mathcal{N}_{MV}^{\times}(p - 1) + 3n = 4\mathcal{N}_{MV}^{\square}(p - 1) + 82n + O(1)$ implies the result of the Lemma. ∎

The estimate (29) is similar to the bound $\mathcal{N}_{MV}(p) = 11pn + O(n)$ obtained in [6] for the 1D index set I with $\eta = 1/2$. Clearly, the corresponding constant in (29) depends on the spatial dimension (compare also Theorem 6.2 for the 3D case).

4.4.3 Matrix Addition, Multiplication and Inversion

As in [6], one can introduce the approximate *addition* $+_{\square}$, *multiplication* $*_{\square}$, and *inversion* of matrices from $\mathcal{M}_{p,p}^{\square}$ retaining the corresponding hierarchical matrix structure. The formatted operations $+_{\square}$ and $*_{\square}$ are defined similarly to the case of 1D-$\mathcal{H}$-matrices considered in [6]. In fact, the complexity analysis of $+_{\square}$ is rather simple and yields $\mathcal{N}_{\square + \square}(p) = O(pn)$.

The proof of $\mathcal{N}_{\square * \square}(p) = O(p^2 n)$ is more lengthy, since various combinations of factors occur.

The inversion is based on blockwise transformations involving the addition and multiplication addressed above. While in the case of [6] the $\mathcal{H}$-matrix was treated as a 2×2 block matrix, the matrix (22) has now a 4×4 block pattern. This does not change the complexity order $\mathcal{N}_{\text{Inversion}}(p) = O(p^2 n)$ obtained in [6].

5 Construction for General 2D-Meshes

We consider an (unstructured) quasi-uniform triangulation $\mathcal{T}_{\overline{h}}$ of $\Omega \subset \mathbb{R}^2$ characterised by the maximal mesh size $\overline{h} := \max\{\overline{d}_\tau : \tau \in \mathcal{T}_{\overline{h}}\}$, where $\overline{d}_\tau$ is the diameter of the Chebyshev sphere of the triangle τ (cf. Footnote 7). Assuming also shape regularity, there are generic constants $c_1, c_2 > 0$ such that

$$
c_1 \overline{d}_\tau \leq \overline{h} \leq c_2 \underline{d}_\tau \qquad \text{for all } \tau \in \mathcal{T}_{\overline{h}},
\tag{30}
$$

where $\underline{d}_\tau$ denotes the diameter of the inscribed circle for an (closed) element τ of $\mathcal{T}_{\overline{h}}$. In fact, we are not restricted to triangles τ. Any elements satisfying (30) are allowed (isoparametric triangles, quadrangles, etc.).

For simplicity, we consider piecewise constant functions on $\tau \in \mathcal{T}_{\overline{h}}$. Then each index $\alpha \in I$ corresponds to a basis function with support $X_\alpha = \tau_\alpha \in \mathcal{T}_{\overline{h}}$. The Chebyshev centre of τ is denoted by ξ_τ (or ξ_α if $\tau = \tau_\alpha$).

In order to construct $\mathcal{H}$-matrix structures, we have to define a suitable cluster tree $T(I)$ (cf. Subsection 4.1). Proposals can be found in [9]. Here, we give a construction based on the uniform tensor-product grid discussed in the previous section. Since the regular grid is needed only for reference, we call it the *fictitious grid*. We do not claim that the presented construction of $T(I)$ is optimal, but it leads to a straightforward proof of the complexity bounds.

5.1 How to Map the Fictitious Hierarchy onto the Unstructured Grid

Without loss of generality we may assume $\Omega \subset \Omega^f := (0, 1) \times (0, 1)$ and

$$
\mu(\Omega) \geq \underline{c}\,\mu(\Omega^f) = \underline{c} > 0,
\tag{31}
$$

where μ denotes the two-dimensional measure. In Ω^f we consider the uniform tensor-product grid $\mathcal{T}_h$ from Section 4. Its index set is denoted by $I^f := \{(i,j) : 1 \leq i, j \leq N\}$, $N = 2^p$ (the superscript 'f' stands for 'fictitious'), while I is the index set of the unknowns of the unstructured grid.

The grid size of $\mathcal{T}_h$ is assumed to be the largest $h = 2^{-p}$ satisfying

$$h < \frac{1}{2\sqrt{2}} \min_{\alpha,\beta \in I,\ \tau_\alpha \cap \tau_\beta \neq \emptyset} (\underline{d}_\alpha + \underline{d}_\beta) \tag{32}$$

with $\underline{d}_\alpha := \underline{d}_{\tau_\alpha}$ from (30). For each index $\alpha \in I$, the Chebyshev centre ξ_α belongs to at least one of the squares X_{ij} of $\mathcal{T}_h$ ($(i,j) \in I^f$; cf. (19)). Selecting one of the possible indices in the multiple case, we are able to define a mapping $F : I \to I^f$ ($\alpha \mapsto F(\alpha) = (i,j)$) via $\xi_\alpha \in X_{F(\alpha)}$. The following remark allows us to define F^{-1} on $F(I) \subset I^f$.

Remark 5.1 *Under condition (32), the mapping F is injective.*

Proof. Let $\alpha \neq \beta$. Then $|\xi_\alpha - \xi_\beta| \geq \frac{1}{2}(\underline{d}_\alpha + \underline{d}_\beta) > \sqrt{2}h = \text{diam}\, X_{ij}$ contradicts $F(\alpha) = F(\beta) = (i,j) \in I^f$. ∎

For any subset $t^f \subset I^f$ (not only for $t^f \subset F(I)$), we define

$$F^{-1}(t^f) := \{\alpha \in I : F(\alpha) \in t^f\} \subset I.$$

Since I^f is the regular grid from Section 4, the cluster tree $T(I^f)$ is already described. F gives rise to the cluster tree for the index set I :

$$T(I) = \{F^{-1}(t^f) : t^f \in T(I^f)\}.$$

The arising tree $T(I)$ meets the conditions of Definition 2.1, but is unusual since some of the vertices $t \in T(I)$ may represent the empty set ($F^{-1}(t^f) = \emptyset$ if $t^f \cap F(I) = \emptyset$). Moreover, if only one of the sons $s \in S(t)$ is non-empty, this son s represents the same subset as the father t. Although, in practice, this tree $T(I)$ could be simplified, we use the tree in the given form since then $T(I)$ and $T(I^f)$ are isomorphic.

As seen in Section 3.1, the cluster tree $T(I)$ determines the block-cluster tree $T_2 = T(I \times I)$, which defines the $\mathcal{H}$-matrix structure. The elements of T_2 are pairs (t_1, t_2) with $t_1, t_2 \in T(I)$. In the case of Construction 3.1, $T_2 = F^{-1}(T_2^f)$ holds, where $T_2^f = T(I^f \times I^f)$ and $F^{-1}((t_1, t_2)) := (F^{-1}(t_1), F^{-1}(t_2))$ for $t_1, t_2 \subset I^f$. Otherwise, we use $T_2 := F^{-1}(T_2^f)$ as definition for T_2.

Let $P_2^f \in T(I^f \times I^f)$ be any admissible $\mathcal{H}$-partitioning for the fictitious grid satisfying (9) with the constant $\eta_f < 1$. Below we will characterise the tolerance constant $\eta \geq \eta_f$ needed for the definition of admissible clusters from the induced partitioning $P_2 \in T(I \times I)$.

Lemma 5.2 *All $t_1, t_2, t \in T_1 = T(I)$ satisfy*

$$\text{diam}\, t \leq \text{diam}\, F(t) + \overline{h}, \tag{33}$$

$$\text{dist}(t_1, t_2) \geq \text{dist}(F(t_1), F(t_2)) - \overline{h}. \tag{34}$$

Proof. Let $x \in \tau_x \in \mathcal{T}_{\overline{h}}$ and $y \in \tau_y \in \mathcal{T}_{\overline{h}}$ for triangles $\tau_x, \tau_y \subset t$ with the Chebyshev centres $\xi_x \in \tau_x, \xi_y \in \tau_y$. Then $|\xi_x - \xi_y| \leq \text{diam}\, F(t)$, while $|\xi_x - x| \leq \text{diam}(\tau_x)/2 \leq \overline{h}/2$ and $|\xi_y - y| \leq \text{diam}(\tau_y)/2 \leq \overline{h}/2$. Hence, $|x - y| \leq \text{diam}\, F(t) + \overline{h}$ yields (33). Similarly, (34) is proved. ∎

Since the image $F(t) = \{F(\alpha) : \alpha \in t\}$ of a cluster $t \in T(I)$ is in general different from the clusters in $T(I^f)$, we introduce mappings F_ℓ for all levels $0 \leq \ell \leq p$. The sets

$$T^\ell(I^f) := \{t^f \in T(I^f) : t^f \text{ is a cluster of level } \ell\} \qquad (0 \leq \ell \leq p)$$

can also be defined by $T^0(I^f) := \{I^f\}$, $T^{\ell+1}(I^f) := \bigcup_{t^f \in T^\ell(I^f)} S(t^f)$ for $0 \leq \ell < p$ and yield a level-wise decomposition of the tree $T(I^f) = \bigcup_{0 \leq \ell \leq p} T^\ell(I^f)$. For $t \subset I$, we define $level(t) := \min\{0 \leq \ell \leq p : F(t) \subset t^f$ for some $t^f \in T^\ell(I^f)\}$.

The mapping F_ℓ is defined on all subsets $t \subset I$ with $level(t) \leq \ell$ and its value is $F_\ell(t) = t^f$ if $t^f \in T^\ell(I^f)$ satisfies $F(t) \subset t^f$. Hence, $F_\ell(t)$ denotes the 'rounding up' of $F(t)$ to an I^f-cluster of level ℓ. Note that $F^{-1}(F_\ell(t)) = t$.

We recall that $(F_\ell(t_1), F_\ell(t_2)) \in P_2^f$ is either a leaf of $T_2^f = T(I^f \times I^f)$ or satisfies the admissibility condition (9): $\min(\text{diam}\, F_\ell(t_1), \text{diam}\, F_\ell(t_2)) \leq 2\eta_f\, \text{dist}(F_\ell(t_1), F_\ell(t_2))$ for the corresponding level ℓ. Because of Remark 4.1, the clusters t_i ($i = 1, 2$) belong to the same level (say ℓ). Assuming the latter inequality, we are interested in the question whether (t_1, t_2) also satisfies condition (9) for a suitable parameter η.

Lemma 5.3 *Assume $t_1, t_2 \in T_1$, $\mathrm{level}(t_1) = \mathrm{level}(t_2) = \ell$ and $\sqrt{2}/2^\ell \geq (1 + 4\eta_f)\overline{h}$. Then the admissibility condition*

$$\min(\mathrm{diam}\, F_\ell(t_1), \mathrm{diam}\, F_\ell(t_2)) \leq 2\eta_f \, \mathrm{dist}(F_\ell(t_1), F_\ell(t_2))$$

implies $\min(\mathrm{diam}(t_1), \mathrm{diam}(t_2)) \leq 2\eta \, \mathrm{dist}(t_1, t_2)$ *for* $\eta := 2\eta_f$.

Proof. Set $A := \mathrm{diam}\, F_\ell(t_1) = \mathrm{diam}\, F_\ell(t_2) = \sqrt{2}/2^\ell$ (cf. (20)) and $B := \mathrm{dist}(F_\ell(t_1), F_\ell(t_2))$. The inequalities (33) and (34) together with $\mathrm{diam}\, F(t_i) \leq A$ and $\mathrm{dist}(F(t_1), F(t_2)) \geq B$ show

$$\frac{\min(\mathrm{diam}(t_1), \mathrm{diam}(t_2))}{\mathrm{dist}(t_1, t_2)} \leq \frac{A + \overline{h}}{B - \overline{h}}.$$

Note that $\frac{A}{B} \leq 2\eta_f$. The assumption $A \geq (1 + 4\eta_f)\overline{h}$ allows us to bound

$$\frac{A + \overline{h}}{B - \overline{h}} - \frac{A}{B} = \overline{h}\,\frac{A + B}{B(B - \overline{h})} = \frac{A/B + 1}{B/\overline{h} - 1} \leq \frac{1 + 2\eta_f}{A/(2\eta_f \overline{h}) - 1}$$

by $2\eta_f$. Hence, $\min(\mathrm{diam}(t_1), \mathrm{diam}(t_2)) \leq 4\eta_f \, \mathrm{dist}(t_1, t_2) = 2\eta \, \mathrm{dist}(t_1, t_2)$ for the choice $\eta = 2\eta_f$. ∎

Corollary 5.4 *The modified assumption* $\mathrm{diam}\, F(t_1) = \mathrm{diam}\, F(t_2) \geq (\frac{1}{\varepsilon} + (2 + \frac{2}{\varepsilon}\eta_f))\overline{h}$ *for some* $\varepsilon > 0$ *leads to* $\min(\mathrm{diam}(t_1), \mathrm{diam}(t_2)) \leq 2\eta \, \mathrm{dist}(t_1, t_2)$ *with* $\eta := (1 + \varepsilon)\eta_f$. *Hence, any* $\eta_f < 1$ *allows a choice* $\eta < 1$.

The condition $\sqrt{2}/2^\ell \geq (1 + 4\eta_f)\overline{h}$ from Lemma 5.3 is not satisfied in general, e.g., for $\ell = p$ we have $\sqrt{2}/2^p = \sqrt{2}h < \overline{h}$ (cf. (32)). However, there is a constant δp such that all clusters t of level $\ell \leq p - \delta p$ fulfil $\sqrt{2}/2^\ell \geq (1 + 4\eta_f)\overline{h}$ as stated in the next lemma, where we may insert $c := 5 \geq 1 + 4\eta_f$.

Lemma 5.5 *Given a constant c, there is a constant $\delta p \in \mathbb{N}$ independent of p so that $\sqrt{2}/2^\ell \geq c\overline{h}$ for all clusters $t \in T$ of level $\ell \leq p - \delta p$.*

Proof. Let t be of level $\ell \leq p - \delta p$. The definition of $h = 2^{-p}$ by (32) together with (30) yields $2^{1-p} = 2h \geq \frac{1}{2\sqrt{2}}\min(\underline{d}_\alpha + \underline{d}_\beta) \geq \frac{1}{c_2\sqrt{2}}\overline{h}$. Hence, $\sqrt{2}/2^{p-\delta p} \geq (2^{\delta p - 1}/c_2)\overline{h}$. Choose δp such that $2^{\delta p - 1}/c_2 \geq c$. ∎

We have to describe an admissible T_2-partitioning P_2, where the parameter η from (9) is defined by $\eta = 2\eta_f$ (see Lemma 5.3). The first trial is to use $P_2^{**} := F^{-1}(P_2^f)$, where P_2^f is the admissible T_2'-partitioning corresponding to η_f. Due to the preceding lemmata, this leads to admissible blocks (t_1, t_2), provided they belong to a level $\ell \leq p - \delta p$. It remains to modify P_2^{**} at the levels ℓ with $p - \delta p < \ell \leq p$. By Remark 2.2, there is a subtree T_2^{**} of $T(I \times I)$ with $P_2^{**} = \mathcal{L}(T_2^{**})$. Construct the smaller tree T_2^* in the following way:

1) Delete all vertices belonging to levels $\ell > p - \delta p$ and

2) insert the sons (i, j), $i \in t_1$, $j \in t_2$ for all *non*-admissible blocks $(t_1, t_2) \in T_2^{**}$ at level $\ell = p - \delta p$.

Then the final T_2-partitioning P_2^* is $P_2^* = \mathcal{L}(T_2^*)$. The matrix interpretation is that all non-admissible blocks of level $p - \delta p$ are full submatrices.

Finally, $\mathcal{M}^* := \mathcal{M}_{\mathcal{H},k}(I \times I, P_2^*)$ defines the $\mathcal{H}$-matrix set corresponding to the unstructured mesh (cf. (11)).

Lemma 5.6 *There holds $n_f \leq \mathrm{const}\, n$, where $n_f := \#I^f$ and $n = \#I$. Moreover,*

$$\mathcal{N}_{st}^*(n) = O(n \log n) \quad \text{and} \quad \mathcal{N}_{MV}^*(n) = O(n \log n)$$

are the respective costs of the storage and the matrix-vector multiplication for matrices from $\mathcal{M}^*$.

Proof. First we consider the auxiliary partitioning $P_2^{**} = F^{-1}(P_2^f) = \mathcal{L}(T_2^{**})$ from above. Since T_2^{**} is isomorphic to a subtree of $T_2^f := T(I^f \times I^f)$, the expenses $\mathcal{N}_{st}^{**}$, $\mathcal{N}_{MV}^{**}$ corresponding to the format $\mathcal{M}^{**} := \mathcal{M}_{\mathcal{H},k}(I \times I, P_2^{**})$ are less or equal to the bounds $O(n_f \log n_f)$ in Lemmata 4.4-4.5. Since δp is a constant, the costs $\mathcal{N}_{st}^*$, $\mathcal{N}_{MV}^*$ are also bounded by $O(n_f \log n_f)$ (note that the same recurrence formulae hold, but the starting value may be increased).

It remains to replace the fictitious dimension n_f in the latter bound by the true dimension n. By (31) we have $n_f = h^{-2} = \mu(\Omega^f) \cdot h^{-2} \leq \frac{\mu(\Omega)}{c}h^{-2}$. The proof of Lemma 5.5 has shown $2h \geq \overline{h}/(c_2\sqrt{2})$, so that $n_f \leq \mu(\Omega)\frac{8c_2^2}{c}\overline{h}^{-2}$. The left inequality in (30) yields $\mu(\Omega) = \sum_{\tau \in \mathcal{T}_h}\mu(\tau) \leq \sum_{\tau \in \mathcal{T}_h}\frac{\pi}{4}\overline{d}_\tau^2 \leq n\frac{\pi}{4c_1^2}\overline{h}^2$. The last two estimates prove $n_f \leq \mathrm{const} \cdot n$ with $\mathrm{const} = \frac{2\pi}{c}(\frac{c_2}{c_1})^2$. ∎

Lemma 4.3 and its proof generalise to all FE stiffness matrices.

Lemma 5.7 *A finite element stiffness matrix belongs to the set $\mathcal{M}^* = \mathcal{M}_{\mathcal{H},k}(I \times I, P_2^*)$ for any $k \geq 1$.*

5.2 Two-Dimensional Manifolds

The above defined matrix formats $\mathcal{M}_{0,p}^{\square}$ and $\mathcal{M}^*$ enable data-sparse $\mathcal{H}$–approximations for a wide class of finite element stiffness matrices corresponding to boundary value problems in $\Omega \subset \mathbb{R}^2$. Applications in boundary element methods (BEM) are based on manifolds (surfaces) instead of flat domains.

In a first step we study the surface of a polyhedron (§5.2.1). Curvilinear surfaces are considered in §5.2.2.

5.2.1 $\mathcal{H}$-Formats for Polyhedrons

Consider a polyhedron $\Gamma \subset \mathbb{R}^3$ composed of M plane faces Γ_i $(1 \le i \le M)$. On each Γ_i a quasi-uniform mesh is given which meets the conditions required in §5. We assume that all pairs of adjacent faces form an angle $\omega \in [\omega_0, 2\pi - \omega_0]$, where $0 < \omega_0 \le \pi$.

To begin with, we assume that for a given admissibility parameter $\eta < 1$ the inequality

$$\min(\operatorname{diam}\Gamma_i, \operatorname{diam}\Gamma_j) \le 2\eta \operatorname{dist}(\Gamma_i, \Gamma_j) \tag{35}$$

holds for all *disjoint and non-adjacent* faces Γ_i, Γ_j. Note that the distance is measured by the *Euclidean distance* in $\mathbb{R}^3$.

In this situation we construct the cluster tree as follows. Let T^i be the $\mathcal{H}$-tree for the face Γ_i, i.e., the root of T^i is the index set I^i corresponding to the unknowns[9] associated with Γ_i $(1 \le i \le M)$. The global set of indices is $I = \cup_{i=1}^M I^i$. The cluster tree $T_1 = T(I)$ is defined as the union of the disjoint trees T^i together with the new root I possessing the M sons I^i $(1 \le i \le M)$. The block-cluster tree is again denoted by T_2.

In the following, we propose a matrix format corresponding to the index set I. Given a block $b = (t_1, t_2) \in T_2$, three different cases can occur:

(i) t_1, t_2 belong to the same face, i.e., $t_1, t_2 \subset I^i$ for some $i \le M$.

(ii) t_1, t_2 belong to adjacent faces.

(iii) t_1, t_2 belong to disjoint and non-adjacent faces.

Case (i) corresponds to the plane case of §5.

In Case (ii), two faces Γ_1, Γ_2 with a common edge e are involved. Turning Γ_2 into the plane of Γ_1, we obtain Γ_1 and $\tilde{\Gamma}_2$ contained in $\Omega := \Gamma_1 \cup \tilde{\Gamma}_2 \subset \mathbb{R}^2$. Choose the matrix structure as in §5 with admissibility parameter $\eta \sin\frac{\omega}{2}$, where ω is the angle between Γ_1 and Γ_2. Let $t_1 \subset \Gamma_1$, $t_2 \subset \Gamma_2$ and denote the corresponding cluster in the rotated copy $\tilde{\Gamma}_2$ by $\tilde{t}_2$. One checks that $\min(\operatorname{diam} t_1, \operatorname{diam} \tilde{t}_2) \le 2\eta \sin\frac{\omega}{2} \operatorname{dist}(t_1, \tilde{t}_2)$ implies $\min(\operatorname{diam} t_1, \operatorname{diam} t_2) \le 2\eta \operatorname{dist}(t_1, t_2)$. Therefore, the chosen partitioning is admissible.

Case (iii) is trivial, since assumption (35) ensures $\min(\operatorname{diam} t_1, \operatorname{diam} t_2) \le 2\eta \operatorname{dist}(t_1, t_2)$.

Altogether, we have obtained an admissible partitioning P_2 comparable to the plane case of §5 with η (partially) replaced by the smaller parameter $\eta \sin\frac{\omega}{2}$.

It remains to discuss the case, where the chosen constant η does not satisfy (35). In this case divide each face into several smaller ones Γ_i' $(1 \le i \le M')$ with $M' > M$. If the subdivision is fine enough, we have

$$\min(\operatorname{diam}\Gamma_i', \operatorname{diam}\Gamma_j') \le 2\eta \operatorname{dist}(\Gamma_i', \Gamma_j') \quad \text{for } \Gamma_i' \subset \Gamma_i, \ \Gamma_j' \subset \Gamma_j,$$

where Γ_i, Γ_j are disjoint and non-adjacent (unrefined) faces. Since η is a fixed constant, also M' is fixed. The same arguments as above can be used to construct an admissible partitioning with similar structure as in the plain case. As in Lemmata 4.4 and 4.5 we derive

Corollary 5.8 *Under the assumption from above, the costs for storage, matrix-vector multiplication, and the further operations have the same complexity as in the plane case of §5.*

The assumption $\omega \in [\omega_0, 2\pi - \omega_0]$ may lead to the impression that small angles cause difficulties. This is not the case. An important example is a slender aerofoil. Here it is well-known that the cluster tree must be constructed differently: The clusters should contain the neighbouring parts from the upper *and* lower side of the wing.

Finally, we mention a special case, where the complexity is even better than mentioned before.

---

[9] According to the example of piecewise constant functions, we assume the index sets I^i to be disjoint. If $I^i \cap I^j \ne \emptyset$ $(i \ne j)$ due to unknowns belonging to the edges, obvious modifications are required.

Remark 5.9 *Consider the double-layer potential for the second order PDEs with constant coefficients in the case of piecewise flat surfaces. Define the structure of the approximating $\mathcal{H}$-matrix as before. Then $\mathcal{N}_{st}$ and $\mathcal{N}_{MV}$ are of the order $O(n)$ instead of $O(n \log n)$.*

The reason is the fact that the kernel function satisfies $\partial k(x,y)/\partial n(x) = 0$ for $x, y \in \Gamma_i$ on any plane face Γ_i of the surface Γ.

5.2.2 Curved Manifold

A general manifold is described by an atlas of mappings. The usual practice is to start from a (reference) polyhedron Γ_{ref} and to define a bi-Lipschitz mapping $\varphi : \Gamma_{ref} \to \Gamma$ (cf. [3]) with Lipschitz constants c_1, c_2:

$$c_1 |x - y| \leq |\varphi(x) - \varphi(y)| \leq c_2 |x - y| \qquad \text{for all } x, y \in \Gamma_{ref}.$$

Let T^{ref} be the cluster tree from §5.2.1 for the reference boundary Γ_{ref}. The corresponding tree for Γ is then defined by $T = \varphi(T^{ref})$. Choose an admissible T^{ref}-partitioning P_2^{ref} with admissibility parameter η_{ref}. Then the resulting T-partitioning $P_2 = \varphi(P_2^{ref})$ is admissible with the parameter $\eta = \frac{c_2}{c_1} \eta_{ref}$, as
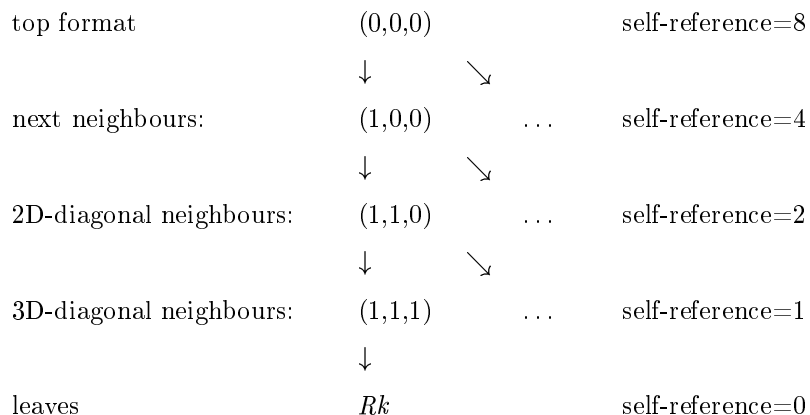
$$\min(\operatorname{diam} \varphi(t_1), \operatorname{diam} \varphi(t_2)) \leq c_2 \min(\operatorname{diam}(t_1), \operatorname{diam}(t_2)) \leq 2c_2 \eta_{ref} \operatorname{dist}(t_1, t_2) \leq \frac{2c_2}{c_1} \eta_{ref} \operatorname{dist}(\varphi(t_1), \varphi(t_2)).$$

Since the matrix-formats corresponding to P_2^{ref} and P_2 are identical, we obtain the same complexity bounds of the computational cost as in §5.2.1.

6 The Three-Dimensional Case

In this section, we introduce the formats for matrices operating in the vector space associated with an index set I for the cell-centred tensor product grid $I_h^3 = I_h \times I_h \times I_h$ in $\Omega = (0,1)^3$ with the mesh size $h = 2^{-p}$ and $\#I = 8^p$. Similar to the 2D-case, the cluster tree $T = T_1$ is defined by the regular refinement (subdivision into eight equal parts) of the initial index set I. $T_1(I)$ gives rise to the block-cluster tree $T(I \times I)$, in which we determine the admissible partitioning according to the admissibility condition (9). In Definition 6.1 below, we choose the constant $\eta = \sqrt{3}/2$ which corresponds to the 3D counterpart of Fig. 1a.

The natural notation of indices from I_h^3 uses triples $(i, j, k) \in \mathbb{N}^3$ with $1 \leq i, j, k \leq 2^p$. As in the 2D case, we can describe the partitioning by a number of formats $\mathcal{M}_{\ell,p}^{\alpha,\beta,\gamma}$, where (α, β, γ) with $\alpha, \beta, \gamma \in \{-1, 0, 1\}$ indicates the shift in the following sense. Let $b = (t, t')$ be a block, where $t, t' \subset I$ are clusters. If $t = t'$, we have a diagonal block and the shift is given by $(\alpha, \beta, \gamma) = (0, 0, 0)$. For these blocks we introduce the 'top format' $\mathcal{M}_{\ell,p}^{0,0,0}$. If $t = (i_0, j_0, k_0) + \{(i, j, k) : 1 \leq i, j, k \leq 2^{p-\ell}\}$ and $t' = (i_0 + 2^{p-\ell}, j_0, k_0) + \{(i, j, k) : 1 \leq i, j, k \leq 2^{p-\ell}\}$ are two clusters (cubes of length $2^{p-\ell}$ in $\mathbb{Z}^3$), their relation is given by the shift $(1, 0, 0)$ indicating the direct neighbourhood in x-direction. Then, for $b = (t, t')$ we use the format $\mathcal{M}_{\ell,p}^{1,0,0}$. Similarly, the other formats $\mathcal{M}_{\ell,p}^{-1,0,0}$, $\mathcal{M}_{\ell,p}^{0,\pm 1,0}$, $\mathcal{M}_{\ell,p}^{0,0,\pm 1}$ ("next neighbours"), $\mathcal{M}_{\ell,p}^{1,1,0}$,... ("2D-diagonal neighbours") and $\mathcal{M}_{\ell,p}^{\pm 1,\pm 1,\pm 1}$ ("3D-diagonal neighbours") are involved. In Definition 6.1 these formats contain the same format at the next level ("self-reference") and other formats as depicted in the graph corresponding to Fig. 3:

top format	(0,0,0)		self-reference=8
	$\downarrow$ $\searrow$		
next neighbours:	(1,0,0)	...	self-reference=4
	$\downarrow$ $\searrow$		
2D-diagonal neighbours:	(1,1,0)	...	self-reference=2
	$\downarrow$ $\searrow$		
3D-diagonal neighbours:	(1,1,1)	...	self-reference=1
	$\downarrow$		
leaves	Rk		self-reference=0

(see also Fig. 4b). Let $\sigma = \{a,b,c,d,e,f,g,h\}$ be the set of the eight sons of a cluster situated as shown in Fig. 4a. For example the block-matrix with columns from a and rows from b is denoted by A_{ab}.
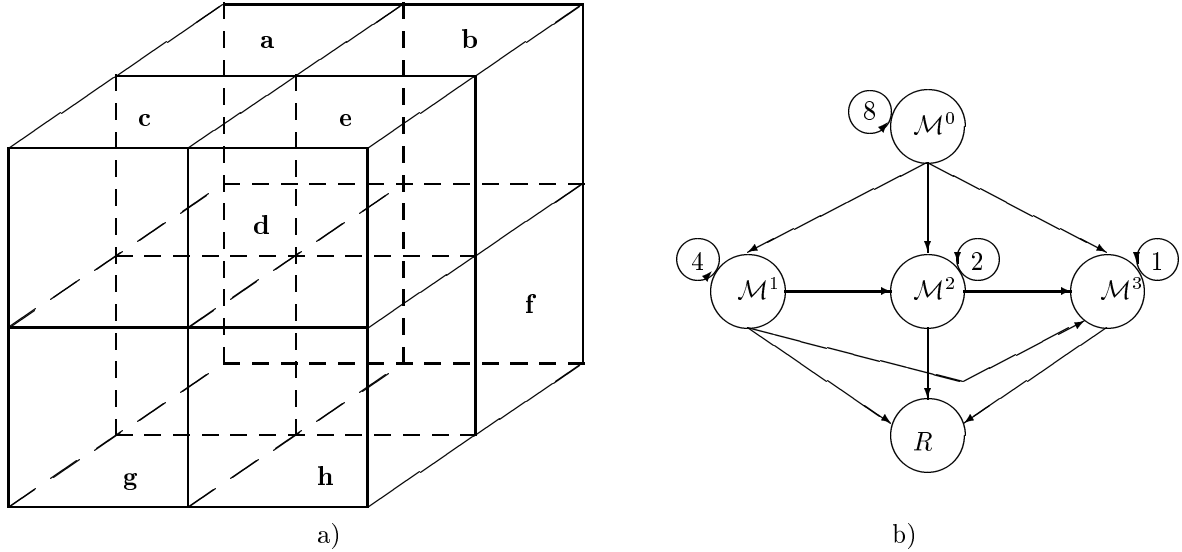

a)   b)

Figure 4: (a) Indexing for the $3D$ clusters, where $a = (0,0,0)$, $b = (1,0,0)$, $c = (0,-1,0)$, $d = (0,0,-1)$, $e = (-1,-1,0)$, $f = (1,0,-1)$, $g = (0,-1,-1)$, $h = (1,-1,-1)$. (b) Graph of the $3D$ formats $\mathcal{M}^m$ ($0 \le m \le 3$) corresponding to $|\alpha| + |\beta| + |\gamma| = m$.

In the following, we define the matrix formats $\mathcal{M}^{\alpha,\beta,\gamma}_{\ell,p}$ recursively with respect to the degree $|\alpha| + |\beta| + |\gamma|$. The notations $\mathcal{M}_{\ell,p}(t_1,t_2)$, $\mathcal{M}_{\ell,p}$, and $\mathcal{M}^R_{\ell,p}$ have the same meaning as in Subsection 4.3.

Definition 6.1 *a) For $\ell = p$, $\mathcal{M}^{\alpha,\beta,\gamma}_{p,p}$ is the set of 1×1-matrices for all $\alpha,\beta,\gamma \in \{-1,0,1\}$. For $\ell < p$, the formats are described in b-e).*

b) Let $\sigma = \{a,b,c,d,e,f,g,h\}$ be the 8 cubes as indicated in Fig. 4a and $\sigma' = \{a',b',c',d',e',f',g',h'\}$ the similar set of clusters shifted in the $(1,1,1)$-direction so that b and g' have one corner point in common. A matrix $A \in \mathcal{M}_{\ell,p}(\sigma,\sigma')$ belongs to $\mathcal{M}^{1,1,1}_{\ell,p}$, if its block matrices in $A = \{A_{ij}\}_{i\in\sigma,j\in\sigma'}$ satisfy $A_{bg'} \in \mathcal{M}^{1,1,1}_{\ell+1,p}$ and $A_{ij} \in \mathcal{M}^R_{\ell+1,p}$, otherwise. Similarly, one defines $\mathcal{M}^{\alpha,\beta,\gamma}_{\ell,p}$ for other combinations subject to $|\alpha| = |\beta| = |\gamma| = 1$.

c) Let $\sigma' = \{a',b',c',d',e',f',g',h'\}$ result from a shift of $\sigma = \{a,b,c,d,e,f,g,h\}$ in the direction $(1,1,0)$ so that the pairs (b,c'), (f,g') of cubes have a common edge. Then $A = \{A_{ij}\}_{i\in\sigma,j\in\sigma'} \in \mathcal{M}^{1,1,0}_{\ell,p}$ holds if the submatrices have the formats $A_{b,c'}$, $A_{f,g'} \in \mathcal{M}^{1,1,0}_{\ell+1,p}$, $A_{f,c'} \in \mathcal{M}^{1,1,1}_{\ell+1,p}$, $A_{b,g'} \in \mathcal{M}^{1,1,-1}_{\ell+1,p}$, and $A_{ij} \in \mathcal{M}^R_{\ell+1,p}$, otherwise. Similarly, one defines $\mathcal{M}^{\alpha,\beta,\gamma}_{\ell,p}$ for other combinations with $|\alpha| + |\beta| + |\gamma| = 2$.

d) Let $\sigma' = \{a',b',c',d',e',f',g',h'\}$ be resulting from a shift of $\sigma = \{a,b,c,d,e,f,g,h\}$ in the direction $(1,0,0)$ so that, e.g., b and a' have a common face. Then $A = \{A_{ij}\}_{i\in\sigma,j\in\sigma'} \in \mathcal{M}^{1,0,0}_{\ell,p}$ holds if $A_{b,a'}$, $A_{f,d'}$, $A_{e,c'}$, $A_{h,g'} \in \mathcal{M}^{1,0,0}_{\ell+1,p}$, $A_{e,a'}$, $A_{h,d'} \in \mathcal{M}^{1,1,0}_{\ell+1,p}$, $A_{b,c'}$, $A_{f,g'} \in \mathcal{M}^{1,-1,0}_{\ell+1,p}$, $A_{b,d'}$, $A_{e,g'} \in \mathcal{M}^{1,0,-1}_{\ell+1,p}$, $A_{f,a'}$, $A_{h,c'} \in \mathcal{M}^{1,0,1}_{\ell+1,p}$, $A_{h,a'} \in \mathcal{M}^{1,1,1}_{\ell+1,p}$, $A_{f,c'} \in \mathcal{M}^{1,-1,1}_{\ell+1,p}$, $A_{b,g'} \in \mathcal{M}^{1,-1,-1}_{\ell+1,p}$, $A_{e,d'} \in \mathcal{M}^{1,1,-1}_{\ell+1,p}$, and $A_{ij} \in \mathcal{M}^R_{\ell+1,p}$, otherwise. Similarly, for other combinations with $|\alpha| + |\beta| + |\gamma| = 1$.

e) Finally, let $\sigma' = \sigma$. Then $A = \{A_{ij}\}_{i,j\in\sigma} \in \mathcal{M}^{0,0,0}_{\ell,p}$ holds if $A_{ii} \in \mathcal{M}^{0,0,0}_{\ell+1,p}$, $A_{ab}, A_{ce}, A_{df}, A_{gh} \in \mathcal{M}^{1,0,0}_{\ell+1,p}$, $A_{ca}, A_{eb}, A_{gd}, A_{hf} \in \mathcal{M}^{0,1,0}_{\ell+1,p}$, $A_{da}, A_{he}, A_{fb}, A_{gc} \in \mathcal{M}^{0,0,1}_{\ell+1,p}$, $A_{cb}, A_{gf} \in \mathcal{M}^{1,1,0}_{\ell+1,p}$, $A_{ae}, A_{dh} \in \mathcal{M}^{1,-1,0}_{\ell+1,p}$, $A_{ga}, A_{hb} \in \mathcal{M}^{0,1,1}_{\ell+1,p}$, $A_{cd}, A_{ef} \in \mathcal{M}^{0,1,-1}_{\ell+1,p}$, $A_{af}, A_{ch} \in \mathcal{M}^{1,0,-1}_{\ell+1,p}$, $A_{db}, A_{ge} \in \mathcal{M}^{1,0,1}_{\ell+1,p}$, $A_{gb} \in \mathcal{M}^{1,1,1}_{\ell+1,p}$, $A_{cf} \in \mathcal{M}^{1,1,-1}_{\ell+1,p}$, $A_{ah} \in \mathcal{M}^{1,-1,-1}_{\ell+1,p}$, $A_{de} \in \mathcal{M}^{1,-1,1}_{\ell+1,p}$, and $A_{ji} \in \mathcal{M}^{-\alpha,-\beta,-\gamma}_{\ell+1,p}$ if $A_{ij} \in \mathcal{M}^{\alpha,\beta,\gamma}_{\ell+1,p}$.

The calculation of the storage and matrix-vector multiplication complexity for the described formats is a result of four staggered recurrence formulae. Below, we give the results (with exact constants) for the corresponding matrix-vector multiplication[10]. Here, we use the notation $\mathcal{M}^m_p := \cup\{\mathcal{M}^{\alpha,\beta,\gamma}_{q-p,q} : q \ge \ell, |\alpha| + |\beta| + |\gamma| = m\}$. Note that all $A \in \mathcal{M}^m_p$ are $8^p \times 8^p$ matrices.

---

[10]More details will be in a forthcoming report.

17

Theorem 6.2 *Let $n = 8^p$ and let $k = 1$ be the rank of the Rk-blocks. Then for $\mathcal{M}_p^m$-matrices the matrix-vector multiplication costs $\mathcal{N}_{MV}^m(p)$ equal*

$$
\begin{aligned}
\mathcal{N}_{MV}^3(p) &= 35 \cdot n + O(1); & \mathcal{N}_{MV}^2(p) &= 51 \cdot n + O(1), \\
\mathcal{N}_{MV}^1(p) &= 187 \cdot n + O(1); & \mathcal{N}_{MV}^0(p) &= 756 \cdot pn + O(p).
\end{aligned}
$$

Proof. The desired estimate for $\mathcal{N}_{MV}^0(p)$ follows from the recurrence

$$
\mathcal{N}_{MV}^0(p) = 8\mathcal{N}_{MV}^0(p-1) + 24 \ \mathcal{N}_{MV}^1(p-1) + 24\mathcal{N}_{MV}^2(p-1) + 8\mathcal{N}_{MV}^3(p-1) + 7n
$$

taking into account $\mathcal{N}_{MV}^0(0) = 1$ and substituting the results for the auxiliary formats. ∎

In the case of a general finite element mesh in a 3D-domain $\Omega \subset \mathbb{R}^3$, we can extend the considerations of §5 to three dimensions as well. We conclude that the $\mathcal{H}$-matrix format is also applicable to general 3D finite element problems as well as for volume integral formulations.

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] W. Dahmen, S. Prössdorf and R. Schneider, *Wavelet approximation methods for pseudodifferential equations I: Stability and convergence*, Math. Z. **215** (1994) 583-620.

[3] I. G. Graham, W. Hackbusch and S. A. Sauter. *Discrete boundary element methods on general meshes in 3D.* Bath Mathematics Preprint 97/19, University of Bath, November 1997. To appear in Numer. Math.

[4] W. Hackbusch: *The panel clustering algorithm.* MAFELAP 1990 (J. R. Whiteman, ed.). Academic Press, London, 1990. pp. 339–348.

[5] W. Hackbusch: *Integral Equations. Theory and Numerical Treatment.* ISNM 128. Birkhäuser, Basel, 1995.

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

[7] 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.

[8] W. Hackbusch and S. A. Sauter: *On the efficient use of the Galerkin method to solve Fredholm integral equations.* Applications of Mathematics **38** (1993) 301–322.

[9] C. Lage: *Softwareentwicklung zur Randelementmethode: Analyse und Entwurf effizienter Techniken.* Doctoral thesis, Universität Kiel 1996.

[10] S. A. Sauter: *Über die effiziente Verwendung des Galerkin-Verfahrens zur Lösung Fredholmscher Integralgleichungen.* Doctoral thesis, Universität Kiel 1992.

[11] T. von Petersdorff and C. Schwab, *Wavelet approximations for first kind boundary integral equations,* Numer. Math. **74** (1996) 479-516.