Raphael Watschinger, BSc

A Directional Approximation of the Helmholtz Kernel and its Application to Fast Matrix-Vector Multiplications

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Mathematics

submitted to

Graz University of Technology

Supervisor

Assoc. Prof. Dr. G. Of

Institute of Applied Mathematics

Graz, February 2019

Abstract

In this thesis a directional approximation of the Helmholtz kernel and its application to fast matrix-vector multiplications is considered. An application of this technique is the solution of boundary value problems for the Helmholtz equation by means of fast boundary element methods. The oscillating behavior of the Helmholtz kernel for large wave numbers does not allow for an application of standard methods. Instead directional methods can be used, which are covered in this thesis.

Directional methods rely on appropriate directional approximations of the Helmholtz kernel. Such an approximation based on tensor interpolation is defined and analyzed. This allows for the definition of a fast directional matrix-vector multiplication based on clustering methods. It is shown, that the storage and time complexity of this fast matrix-vector multiplication is of quasilinear order in the number of columns and rows of the matrix under suitable assumptions on the wave number and the geometry. The algorithm corresponding to the fast directioanl matrix-vector multiplication depends in particular on two parameters, whose selection has to be investigated. For this purpose the algorithm is implemented and a parameter study is conducted, whose results are summarized in a parameter selection strategy.

Contents

Introduction

Many processes in physical and industrial applications are modeled by linear partial differential equations on suitable domains Ω with appropriate boundary conditions. To understand these processes one needs to solve the corresponding boundary value problems. An example is the Dirichlet boundary value problem for the Laplace operator, which is given by

$$\Delta u(x) = f(x),\ x \in \Omega, \qquad u(x) = h(x),\ x \in \partial\Omega,$$

where Δ denotes the Laplace operator, $\partial\Omega$ the boundary of Ω, f a source term, and h a Dirichlet datum. In general the solution of such a boundary value problem cannot be computed exactly, but has to be approximated using numerical discretization methods. Finite element methods (FEM) are widely used for this purpose. They base on a decomposition of the domain Ω into small elements – often triangles in 2D or tetrahedrons in 3D – and the solution of a variational formulation of the boundary value problem in suitable discrete ansatz spaces. Typically the functions in the test and trial spaces have local supports, i.e. they are zero on most of the elements of Ω. The variational formulation can then be written as a matrix equation with a sparsely populated matrix A which is solved with appropriate direct or iterative solvers.

Boundary element methods (BEM) are another class of methods to solve certain boundary value problems. If a fundamental solution f of the partial differential equation is known and there is no source term, then one can represent the solution of the boundary value problem by means of boundary integral equations. The idea of BEM is to discretize these boundary integral equations. For the discretization a decomposition of the boundary $\partial\Omega$ of Ω into small elements suffices. In particular, the degrees of freedom are typically much lower for BEM than for FEM. In 3D for example, if h is the characteristic size of an element of the decomposition of Ω or $\partial\Omega$ and $n = 1/h$, then the degrees of freedom for FEM are $\mathcal{O}(n^3)$ while only $\mathcal{O}(n^2)$ for BEM. However, the BEM system matrices are in general densely populated. In particular, the required storage for such a matrix is $\mathcal{O}(n^4)$ in contrast to $\mathcal{O}(n^3)$ for sparsely populated FEM matrices. Similarly, the computation times for the solution of the matrix equations are considerably higher for BEM than for FEM. This makes the use of classical BEM prohibitive for large-scale problems.

In the 1980s Rokhlin and Greengard proposed the fast multipole method (FMM) [20, 22, 34], which allowed to resolve this deficiency of BEM for a large class of boundary value problems. In fact, the FMM allows to reduce the complexity of the application of BEM system matrices to an almost linear order, which leads to a significant

speed up of iterative solvers based on matrix-vector multiplications. The FMM relies on a hierarchical partitioning of the considered geometry and the splitting of the matrix application into a nearfield part, which has to be evaluated directly, and a farfield part, which can be evaluated efficiently using suitable expansions of the fundamental solution. Since then, many other methods for the fast solution of BEM have been developed, e.g. Panel Clustering [27], Adapative Cross Approximation [1, 3], Wavelets [12], hierarchical matrices [23, 25], and $\mathcal{H}^2$–matrices [26].

A special treatment is required for boundary value problems for the Helmholtz equation

$$\Delta u + \kappa^2 u = 0,$$

where $\kappa > 0$ is the so-called wave number. This equation can be used to model the propagation of waves and is therefore of particular interest in acoustics and electrodynamics among other fields. The fundamental solution of the Helmholtz equation in $\mathbb{R}^3$ is the Helmholtz kernel f defined by

$$f(x, y) := \frac{\exp(i\kappa|x - y|)}{4\pi|x - y|}.$$

For large κ this kernel is highly oscillatory due to the term $\exp(i\kappa|x - y|)$. This oscillatory behavior makes standard FMM schemes [11, 18], based for example on the expansion of the kernel into spherical harmonics, inefficient for large κ. In [21, 35] a different approach based on diagonal forms of translation operators is considered, which works good in high frequency regimes, i.e. for large κ, but yields numerically unstable schemes if applied in low-frequency regimes. In the context of the FMM there have been several approaches to overcome the problems related to the tranistion between low and high frequency regime, see e.g. [10, 13, 29]. Cheng et al. [10], for example, combine the aforementioned strategies by using spherical harmonic expansions in the low frequency setting and a transition to diagonal forms of translation operators, also denoted as farfield signatures in their work, in the high frequency regime. The corresponding algorithm works accurately in both frequency regimes and has a runtime complexity of $\mathcal{O}(N)$ if low-frequency computations dominate, or $\mathcal{O}(N \log(N))$ in the contrary case, where N denotes the number of nodes in the discretization of the domain. This method is widely used in the context of FMM [9].

Directional methods are another class of methods which overcome the problems of standard FMM methods in the high frequency regime. These methods are based on a directional approximation of the Helmholtz kernel. In the setting of fast matrix operations such approximations were first considered in [8], picked up in [15], and considered more recently in [2, 5, 6, 7, 30, 31]. The basic idea is that the Helmholtz kernel f can be smoothened in a cone centered at a direction c by a multiplication with a plane wave term of the form $\exp(-i\kappa\langle c, x - y\rangle)$, which allows for a low rank approximation of the kernel locally, for example by polynomial interpolation. With such an approximation one can proceed similarly as in the case of typical clustering

methods by constructing a hierarchical partitioning of the geometry and splitting the computations in nearfield and appropriate farfield parts. The main difference to standard FMM schemes is that suitable sets of directions have to be considered in the farfield evaluation in the high frequency regime and that the related partitioning of the matrix is finer.

In this thesis a directional approximation of the Helmholtz kernel and its application to fast matrix-vector multiplications is discussed in detail. In Chapter 1 we follow mainly the lines of [7]. First we recall basic definitions and results concerning univariate and tensor interpolation. Then we construct a directional single-level approximation of the kernel f on pairs of axis-parallel boxes X and Y, by multiplicating f with a plane wave term as described above and interpolating the resulting function on the boxes X and Y. Directional reinterpolation schemes are used to transform these single-level approximations into multi-level approximations, which are needed to define an efficient algorithm in Chapter 2. To ensure a sufficient approximation quality we introduce three admissibility criteria in Section 1.2.2. One of these criteria is used to control the distance of boxes in the low frequency regime, another one to control the distance of boxes in the high frequency regime, and the last to assign suitable directions to each pair of boxes. These criteria and suitable assumptions on the directions used for the directional reinterpolation allow us to show exponential convergence of the approximation errors of the directional single- and multi-level approximations of the kernel f on appropriate pairs of boxes X and Y. The results which we proof here are slightly sharper than those in [7].

Chapter 2 is dedicated to the derivation of a fast directional matrix-vector multiplication for a matrix A generated by the Helmholtz kernel (cf. (1.19) for the definition of such a matrix). We start by constructing box cluster trees, which are used for the hierarchical partitioning of the considered geometry into axis-parallel boxes similarly as in classical FMM. Then we construct appropriate sets of directions for the directional approximation, similarly as in [15, 30]. For the sake of efficiency the same directions are used for all boxes on a level of a box cluster tree. We describe how to assign two boxes a suitable direction from the constructed set of directions and discuss which directions should be used for the directional reinterpolation. In particular, we show that the chosen directions are such that the assumptions of the theorems on the approximation errors in Chapter 1 hold. Such a detailed discussion of the properties of the constructed sets of directions has not been conducted before in literature, to our best knowledge. After this section on directions we define block trees, which are used to find and organize all pairs of boxes in two box cluster trees for which the admissibility criteria in Section 1.2.2 hold. This will enable us to describe Algorithm 2.42 for the fast directional matrix-vector multiplication, similarly as in [5]. In the remainder of Chapter 2 we first talk about some details which should be considered in an implementation of this algorithm to improve its performance. Then we analyze its complexity. Inspired by the complexity discussion in [5] we show that the

complexity of our fast directional matrix-vector multiplication is of order $\mathcal{O}(N \log(N))$ under suitable assumptions on the wave number κ, the geometry, and the degrees of freedom N.

In Chapter 3 we present the results of some numerical experiments, for which we have implemented Algorithm 2.42 in C++. We show the effect of the parameters η_2 and ℓ_{hf} on the runtime and accuracy of this algorithm in a detailed parameter study for points distributed on the surface of a cube. Here, the parameter η_2 is a separation parameter originating from the admissibility criteria in Section 1.2.2 and controls ultimately which pairs of boxes are admissible. The parameter ℓ_{hf} is introduced in the construction of the directions in Section 2.1.2 and controls which boxes are in the high frequency regime. The observations from this parameter study lead to a parameter selection strategy, which is presented in Section 3.1.4 and applied to the fast directional matrix-vector multiplication for points distributed on a sphere subsequently. We conclude Chapter 3 by presenting computation results which show why one aspect of the implementation details presented in Chapter 2, namely suitable compression schemes, are crucial for a good performance of the presented directional method.

1 A Directional Approximation of the Helmholtz Kernel

Fast methods for the solution of boundary value problems by means of boundary element methods rely typically on suitable separable approximations of the underlying fundamental solution. The fundamental solution of the Helmholtz equation

$$\Delta u + \kappa^2 u = 0,$$

with the wave number $\kappa > 0$, is given by the Helmholtz kernel

$$f(x,y) := \frac{\exp(i\kappa|x-y|)}{4\pi|x-y|}. \tag{1.1}$$

As discussed in the introduction, the oscillations caused by the term $\exp(i\kappa|x-y|)$ for large κ require special approximation methods. One such method which is suitable for the approximation of the Helmholtz kernel in all frequency regimes is directional approximation. We dedicate this chapter to the description and analysis of a directional approximation of the Helmholtz kernel based on interpolation.

We start by recalling some basic definitions and results of univariate and tensor polynomial interpolation in Section 1.1. In particular, we discuss a 1D interpolation error estimate, which can be found in [7] and plays an important rule in the later error estimates of the directional approximation. We conclude this section with a demonstration on how polynomial interpolation can be used for an efficient approximation of suitable matrices.

Section 1.2 is dedicated to the introduction of a directional approximation of the Helmholtz kernel f. Here we follow the lines of [7]. We start by defining a directional single-level approximation of f, which is based on the multiplication of f with a plane wave term $\exp(-i\kappa\langle c, x-y\rangle)$ for some appropriate direction c satisfying $|c| = 1$ and subsequent tensor polynomial interpolation on suitable axis-parallel boxes X and Y. Then we discuss directional reinterpolation schemes leading to a multi-level approximation of the Helmholtz kernel. The fast directional matrix-vector multiplication derived in Chapter 2 is based on this multi-level approximation. In the last part of Section 1.2 we discuss admissibility criteria which the boxes X and Y and the direction c need to fulfill such that the quality of the directional approximation with direction c on these boxes is sufficiently good. These criteria are again taken from [7].

In Section 1.3 we present error estimates of the directional single- and multi-level approximations of the Helmholtz kernel f. We show in Theorem 1.15 that the directional multi-level approximation error converges exponentially to zero with increasing interpolation degree m. Its proof relies on the exponential decay of the directional single-level approximation errors, which is proved in Theorem 1.26 in Section 1.3.1, and an estimate of the norms of the reinterpolation operators, which are used to transform a single-level approximation into a multi-level approximation. The latter estimate can be found in Theorem 1.35 in Section 1.3.2. Througout Section 1.3 we follow the lines of [7]. However, the error estimates of the directional multi-level approximation of f in Theorem 1.15 and of the directional single-level approximation of f in Theorem 1.26 are slightly sharper than the corresponding results in [7].

1.1 Polynomial interpolation

The approximation of a function by means of polynomial interpolation plays an important role in the present work. Hence we devote this section to the discussion of some of its aspects. First we recall some definitions and motivate why we use the Chebyshev nodes as interpolation points. Here we follow the lines of [7], but adapt the notation slightly. Then we discuss a result, which gives us exponential convergence of the interpolation error in the supremum norm. Finally we demonstrate how polynomial interpolation can be used to approximate matrices generated by smooth functions.

1.1.1 Basic definitions and results concerning interpolation

Definition 1.1 (Interpolation Operators)**.** Let Π_m be the space of all polynomials of degree $m \in \mathbb{N}_0$ and let $\{\xi_{[a,b],j}\}_{j=0}^m$ be $m+1$ distinct interpolation points in an interval $[a,b]$. Define the interpolation operator

$$\mathcal{I}_{[a,b]}^{(m)} : C([a,b]) \to \Pi_m$$

by identifying $\mathcal{I}_{[a,b]}^{(m)}[g]$ with the unique polynomial of degree m that satisfies

$$\mathcal{I}_{[a,b]}^{(m)}[g](\xi_{[a,b],j}) = g(\xi_{[a,b],j}), \qquad \text{for all } j \in \{0,...,m\}.$$

For an n dimensional, axis-parallel box $B = [a_1,b_1] \times ... \times [a_n,b_n]$ consider the set of interpolation points $\{\xi_{B,\nu}\}_{\nu \in M}$, where $\xi_{B,\nu} = (\xi_{[a_1,b_1],\nu_1}, ..., \xi_{[a_n,b_n],\nu_n})$ for a multi-index $\nu = (\nu_1,...,\nu_n)$ in the set $M = \{0,...,m\}^n$. Here $\{\xi_{[a_k,b_k],j}\}_{j=0}^m$ are again $m+1$ distinct interpolation points in $[a_k,b_k]$ for all $k \in \{1,...,n\}$. We consider the tensor product space $\bigotimes_{k=1}^n \Pi_m$, which consists of all multivariate polynomials π in x_1, $...x_n$ of the form

$$\pi(x_1,...,x_n) = \sum_{k_1=1}^{m} ... \sum_{k_n=1}^{m} c_{k_1,...,k_n} x_1^{k_1} ... x_n^{k_n}.$$

The interpolation operator

$$\mathcal{I}_B^{(m)} : C(B) \to \bigotimes_{k=1}^{n} \Pi_m$$

is defined by identifying $\mathcal{I}_B^{(m)}[g]$ with the unique polynomial in $\bigotimes_{k=1}^{n} \Pi_m$ that satisfies

$$\mathcal{I}_B^{(m)}[g](\xi_{B,\nu}) = g(\xi_{B,\nu}), \qquad \text{for all } \nu \in M.$$

Definition 1.2 (Lagrange polynomials)**.** Let $\{\xi_{[a,b],j}\}_{j=0}^{m} \subset [a,b]$ be $m+1$ distinct interpolation points. We define the one dimensional Lagrange polynomials

$$L_{[a,b],k}^{(m)}(x) := \prod_{\substack{j=0 \\ j \neq k}}^{m} \frac{x - \xi_{[a,b],j}}{\xi_{[a,b],k} - \xi_{[a,b],j}}, \qquad \text{for all } k \in \{0, ..., m\}.$$

Similarly, for a set of distinct interpolation points $\{\xi_{B,\nu}\}_{\nu \in M}$ in an axis-parallel box $B = [a_1, b_1] \times ... \times [a_n, b_n]$ we define the n-dimensional Lagrange polynomials

$$L_{B,\nu}^{(m)}(x) := \prod_{j=1}^{n} L_{[a_j,b_j],\nu_j}^{(m)}(x_j), \qquad \text{for all } \nu \in M = \{0, ..., m\}^n. \tag{1.2}$$

Remark 1.3. In the notation of the interpolation operators $\mathcal{I}_{[a,b]}^{(m)}$ and $\mathcal{I}_B^{(m)}$ as well as the Lagrange polynomials $L_{[a,b],k}^{(m)}$ and $L_{B,\nu}^{(m)}$ we neglect their dependency on the interpolation points $\{\xi_{[a,b],j}\}_{j=0}^{m}$ and $\{\xi_{B,\nu}\}_{\nu \in M}$, respectively. However, in this work we will never use different interpolation points on the same intervals or boxes simultaneously, and hence this notation is used to support readability.

With the Lagrange polynomials we can represent $\mathcal{I}_{[a,b]}^{(m)}[g]$ and $\mathcal{I}_B^{(m)}[h]$ for all functions $g \in C([a,b])$ and $h \in C(B)$, respectively, by

$$\mathcal{I}_{[a,b]}^{(m)}[g] = \sum_{j=0}^{m} g(\xi_{[a,b],j}) L_{[a,b],j}^{(m)},$$
$$\mathcal{I}_B^{(m)}[h] = \sum_{\nu \in M} h(\xi_{B,\nu}) L_{B,\nu}^{(m)}.$$

Throughout this work we will use the Chebyshev nodes as interpolation points, which are defined next.

Definition 1.4 (Chebyshev polynomials and Chebyshev nodes)**.** For all $m \in \mathbb{N}_0$ the Chebyshev polynomial T_m on the interval $[-1, 1]$ is defined by

$$T_m(x) = \cos(m \arccos(x)). \tag{1.3}$$

For every $m \in \mathbb{N}_0$ the $m+1$ Chebyshev nodes $\{\xi_{[-1,1],j}\}_{j=0}^m$ in the interval $[-1,1]$ are the roots of T_{m+1}, which are given by

$$\xi_{[-1,1],j} = \cos\left(\frac{2j+1}{2(m+1)}\pi\right), \qquad \text{for all } j \in \{0,...,m\}.$$

For an arbitrary, non-empty interval $[a,b]$ we define the transformed Chebyshev nodes $\{\xi_{[a,b],j}\}_{j=0}^m$ by

$$\xi_{[a,b],j} = \Phi_{[a,b]}(\xi_{[-1,1],j}), \qquad \text{for all } j \in \{0,...,m\}, \tag{1.4}$$

where the affine linear mapping $\Phi_{[a,b]}$ is defined by

$$\Phi_{[a,b]} : [-1,1] \to [a,b], \qquad x \mapsto \frac{1}{2}(a+b+x(b-a)). \tag{1.5}$$

The Chebyshev nodes are a particularly good choice for the interpolation points, which is motivated by the following theorem.

Theorem 1.5 (cf. [16, Theorem 1] and [32, Theorem 1.2])**.** *Let the Lebesgue constant Λ_m be defined by*

$$\Lambda_m := \|\mathcal{I}_{[a,b]}^{(m)}\| = \sup_{0 \neq g \in C([a,b])} \frac{\|\mathcal{I}_{[a,b]}^{(m)}[g]\|_{\infty,[a,b]}}{\|g\|_{\infty,[a,b]}}. \tag{1.6}$$

There exists a constant $c > 0$ such that

$$\Lambda_m \geq \frac{2}{\pi}\log(m+1) - c, \tag{1.7}$$

independently of the choice of the interpolation points. For the special choice of the transformed Chebyshev interpolation points, there holds

$$\Lambda_m \leq \frac{2}{\pi}\log(m+1) + 1. \tag{1.8}$$

Sketch of proof. The case of a general interval $[a,b]$ can be easily deduced from the case of the interval $[-1,1]$. On the interval $[-1,1]$ one can show that

$$\Lambda_m = \|\lambda_{m+1}\|_{\infty,[-1,1]}, \qquad \text{where } \lambda_{m+1}(x) := \sum_{j=0}^m |L_{[-1,1],j}^m(x)|,$$

holds for an arbitrary choice of interpolation points. Then assertion (1.7) is given in [16, Theorem 1] and assertion (1.8) in [32, Theorem 1.2]. □

1.1.2 A 1D interpolation error estimate

For the error estimate of the directional approximation of the Helmholtz kernel in Section 1.3 we need to estimate the error made when replacing a function $g \in C([-1,1])$ by its interpolant $\mathcal{I}^{(m)}_{[-1,1]}[g] \in \Pi_m$. In [7] a useful result is given, which grants exponential convergence of the error in the supremum norm for suitable functions g. We discuss this result here.

Let us start with some definitions. For the sake of completeness we recall the definition of a holomorphic function.

Definition 1.6 (Holomorhpic functions)**.** Let $U \subset \mathbb{C}$ be open. A function $g : U \to \mathbb{C}$ is called complex differentiable in a point $z_0 \in U$ if the limit

$$g'(z_0) = \lim_{z \to z_0} \frac{g(z) - g(z_0)}{z - z_0}$$

exists. If g is complex differentiable for all points $z \in U$ then g is called holomorphic on U.

For the approximation result we have to deal with holomorphic functions on Bernstein elliptic discs, which we define next.

Definition 1.7 ([7, cf. equation (3.22)])**.** Let $\rho \in \mathbb{R}$ with $\rho > 1$ be given. The Bernstein elliptic disc D_ρ is defined by

$$D_\rho := \left\{ z = x + iy : x, y \in \mathbb{R},\ \left(\frac{2x}{\rho + \rho^{-1}}\right)^2 + \left(\frac{2y}{\rho - \rho^{-1}}\right)^2 < 1 \right\}. \tag{1.9}$$

These definitions allow us to formulate the following proposition, which is the key result of this section.

Proposition 1.8 ([7, cf. Lemma 3.11])**.** *Let $\rho \in \mathbb{R}$ satisfy $\rho > 1$. Let $f : D_\rho \to \mathbb{C}$ be holomorphic and $C_f := \max\{|f(z)| : z \in \overline{D_\rho}\}$. Then there exists a polynomial $\pi_m \in \Pi_m$ of degree m such that*

$$\|f - \pi_m\|_\infty \leq \frac{2C_f}{\rho - 1} \rho^{-(m+1)}. \tag{1.10}$$

Proof. Here we follow [14, cf. Chapter 7; proof of Theorem 8.1]. We define the function

$$g(t) := f(\cos(t)).$$

This function is obviously even and 2π-periodic. Furthermore it is continuously differentiable for all $t \in \mathbb{R}$. This follows immediately from the fact that f is holomorphic on

D_ρ and hence continuously differentiable in $(-1-\epsilon, 1+\epsilon)$ for an $\epsilon > 0 \in \mathbb{R}$ sufficiently small. Therefore we can express g for all t as a Fourier series, i.e.

$$g(t) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} a_k \cos(kt), \qquad a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\cos(t))\cos(kt)\,\mathrm{d}t.$$

By the definition of g and the Chebyshev polynomials in (1.3) it follows that

$$f(x) = g(\arccos(x)) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} a_k T_k(x),$$

i.e. we can expand f on $[-1,1]$ into a series of Chebyshev polynomials. Let us define

$$\pi_m(x) = \frac{1}{2}a_0 + \sum_{k=1}^{m} a_k T_k(x).$$

We want to show that (1.10) holds for this choice of π_m. Since $|T_k(x)| \leq 1$ for all $k \in \mathbb{N}$ and all $x \in [-1,1]$ we get

$$\|f - \pi_m\|_\infty \leq \sum_{k=m+1}^{\infty} |a_k|.$$

Let us therefore estimate $|a_k|$. By using the substitution $z = e^{it}$, we rewrite the defining integral of a_k into the line integral

$$a_k = \frac{1}{\pi i}\int_{\partial B_1(0)} f\left(\frac{z+z^{-1}}{2}\right)\frac{z^k + z^{-k}}{2}\frac{1}{z}\,\mathrm{d}z,$$

where $\partial B_1(0)$ is the circle defined by $\partial B_1(0) := \{z \in \mathbb{C} : |z| = 1\}$. We split this integral into two parts

$$a_k = \underbrace{\frac{1}{2\pi i}\int_{\partial B_1(0)} f\left(\frac{z+z^{-1}}{2}\right) z^{k-1}\,\mathrm{d}z}_{:=I_1} + \underbrace{\frac{1}{2\pi i}\int_{\partial B_1(0)} f\left(\frac{z+z^{-1}}{2}\right) z^{-(k+1)}\,\mathrm{d}z}_{:=I_2},$$

and estimate both parts separately.

Let $\partial B_\mu(0) := \{z \in \mathbb{C} : |z| = \mu\}$ be the circle with radius μ. An easy computation shows that the mapping $\Psi(z) = (z+z^{-1})/2$ maps the circles $\partial B_\mu(0)$ and $\partial B_{1/\mu}(0)$ for all $\mu > 1$ injectively to the boundary ∂D_μ of the Bernstein elliptic discs D_μ and the circle $\partial B_1(0)$ to the interval $[-1,1]$. Hence, it maps the annulus

$$A := \{z \in \mathbb{C} : 1/\rho < |z| < \rho\}$$

to the Bernstein elliptic disc D_μ. As a consequence the function $f \circ \Psi$ and also the integrands above are holomorphic on A. This allows us to apply the homotopic

version of Cauchy's integral theorem [17, cf. Chapter 4, Theorem 1.3]. For all $\hat{\rho}$ such that $1 < \hat{\rho} < \rho$ we get

$$\begin{aligned} |I_1| &= \frac{1}{2\pi} \left| \int_{\partial B_1(0)} f\left(\frac{z+z^{-1}}{2}\right) z^{k-1} \,\mathrm{d}z \right| = \frac{1}{2\pi} \left| \int_{\partial B_{1/\hat{\rho}}(0)} f\left(\frac{z+z^{-1}}{2}\right) z^{k-1} \,\mathrm{d}z \right| \\ &\leq \frac{1}{2\pi} \int_{\partial B_{1/\hat{\rho}}(0)} \left| f\left(\frac{z+z^{-1}}{2}\right) \right| |z^{k-1}| \,\mathrm{d}z \leq \frac{1}{2\pi} C_f \hat{\rho}^{-k+1} 2\pi \hat{\rho}^{-1} = C_f \hat{\rho}^{-k}, \end{aligned}$$

and similarly

$$|I_2| = \frac{1}{2\pi} \left| \int_{\partial B_{\hat{\rho}}(0)} f\left(\frac{z+z^{-1}}{2}\right) z^{-(k+1)} \,\mathrm{d}z \right| \leq C_f \hat{\rho}^{-k}.$$

This yields

$$\|f - \pi_m\|_\infty \leq \sum_{k=m+1}^{\infty} |a_k| \leq 2C_f \sum_{k=m+1}^{\infty} \hat{\rho}^{-k} = \frac{2C_f}{\hat{\rho}-1} \hat{\rho}^{-(m+1)}.$$

By taking the limit $\hat{\rho} \to \rho$ we get assertion (1.10). □

A direct consequence of this proposition is the following theorem, which is the basis for the error estimate of the directional approximation of the Helmholtz kernel in Section 1.3.

Theorem 1.9. *Let $\rho > 1 \in \mathbb{R}$ and $f : D_\rho \to \mathbb{C}$ be holomorphic. Let Λ_m be given by (1.6) and let $C_f := \max\{|f(z)| : z \in \overline{D_\rho}\}$. Then there holds*

$$\|f - \mathcal{I}^{(m)}_{[-1,1]}[f]\|_{\infty,[-1,1]} \leq \frac{2C_f}{\rho-1}(\Lambda_m + 1)\rho^{-(m+1)}. \tag{1.11}$$

Proof. The proof is a direct consequence of Proposition 1.8 and the projection property of the interpolation operator $\mathcal{I}^{(m)}_{[-1,1]}$. In fact we get for π_m as in Proposition 1.8

$$\begin{aligned} \|f - \mathcal{I}^{(m)}_{[-1,1]}[f]\|_{\infty,[-1,1]} &= \|f - \pi_m + \mathcal{I}^{(m)}_{[-1,1]}[\pi_m] - \mathcal{I}^{(m)}_{[-1,1]}[f]\|_{\infty,[-1,1]} \\ &\leq \|f - \pi_m\|_{\infty,[-1,1]} + \|\mathcal{I}^{(m)}_{[-1,1]}[\pi_m] - \mathcal{I}^{(m)}_{[-1,1]}[f]\|_{\infty,[-1,1]} \\ &\leq (1 + \Lambda_m)\|f - \pi_m\|_{\infty,[-1,1]} \leq \frac{2C_f}{\rho-1}(\Lambda_m + 1)\rho^{-(m+1)}. \end{aligned}$$

□

Remark 1.10. By Theorem 1.5 the Lebesgue constant increases only logarithmically in m if the Chebyshev nodes are chosen as interpolation points. Hence Theorem 1.9 gives us exponential convergence of the interpolation error.

1.1.3 Matrix approximation via tensor interpolation

Suppose we want to approximate a matrix $A \in \mathbb{C}^{N_X \times N_Y}$ generated by a complex valued function $f_A \in C(X \times Y)$, i.e.

$$A[j,k] = f_A(x_j, y_k),$$

where $\{x_j\}_{j=1}^{N_X}$ and $\{y_k\}_{k=1}^{N_Y}$ are points in $X \subset \mathbb{R}^3$ and $Y \subset \mathbb{R}^3$, respectively. Assuming that X and Y are axis parallel boxes we can apply tensor interpolation in X and Y and get

$$\mathcal{I}_X^{(m)} \otimes \mathcal{I}_Y^{(m)}[f_A](x,y) = \sum_{\alpha \in M} \sum_{\beta \in M} f_A(\xi_{X,\alpha}, \xi_{Y,\beta}) L_{X,\alpha}(x) L_{Y,\beta}(y),$$

where as in the previous section M is the set $\{0,...,m\}^3$, and $\{\xi_{X,\alpha}\}_{\alpha \in M}$ and $\{\xi_{y,\beta}\}_{\beta \in M}$ are some interpolation points in the respective boxes. With this approximation we get

$$A[j,k] \approx \sum_{\alpha \in M} \sum_{\beta \in M} f_A(\xi_{X,\alpha}, \xi_{Y,\beta}) L_{X,\alpha}(x_j) L_{Y,\beta}(y_k).$$

This can be rewritten in matrix notation as follows. Let the interpolation matrices $L_X \in \mathbb{C}^{N_X \times (m+1)^3}$, $L_Y \in \mathbb{C}^{N_Y \times (m+1)^3}$ and the coupling matrix $A_{X,Y} \in \mathbb{C}^{(m+1)^3 \times (m+1)^3}$ be defined by

$$\begin{aligned}
L_X[j,k] &:= L_{X,\alpha_k}(x_j), && j \in \{1,...,N_X\}, k \in \{1,...,(m+1)^3\},\\
L_Y[j,k] &:= L_{Y,\beta_k}(y_j), && j \in \{1,...,N_Y\}, k \in \{1,...,(m+1)^3\},\\
A_{X,Y}[j,k] &:= f_A(\xi_{X,\alpha_j}, \xi_{Y,\beta_k}), && j,k \in \{1,...,(m+1)^3\}.
\end{aligned}$$

Then the matrix A is approximated by

$$A \approx L_X A_{X,Y} L_Y^*, \tag{1.12}$$

where we used that all entries of L_Y are real. This leads to an efficient approximation of A if $(m+1)^3$ is sufficiently smaller than N_X and N_Y. Indeed, instead of $N_X N_Y$ only

$$(N_X + N_Y + (m+1)^3)(m+1)^3$$

entries have to be computed and stored.

The quality of the above approximation depends on the interpolation degree m and the behavior of the function f_A in $X \times Y$. Generally, higher interpolation degrees lead to lower approximation errors but at the same time to higher computational and storage costs. However, if the function f_A is sufficiently smooth in $X \times Y$, a low interpolation degree m might suffice for a good approximation.

1.2 A directional multi-level approximation

In this section we want to discuss how to approximate the Helmholtz kernel f in (1.1) efficiently on appropriate sets $X \times Y$ using tensor interpolation. For general sets this is not possible. In fact, the singularities of f for $x = y$ and its oscillatory behavior for large wave numbers κ or large sets X or Y have a negative impact on the approximation quality. The singular behavior can be overcome with a classical separation criterion of clustering methods for the sets X and Y. We will introduce such a criterion in Section 1.2.2. The oscillatory behavior, however, requires a special treatment. In fact, applying a classical approach that uses only a separation criterion and does not take the oscillatory behavior into account is limited in a high frequency regime, which is demonstrated for example in [30, Chapter 3, Example 6]. For this purpose we introduce a directional multi-level approximation strategy, following the lines of [7]. First we present a single level approximation of f, which forms the basis of the multi-level approximation. Then we will motivate the necessity of a multi-level scheme and introduce it. Finally we will introduce necessary admissibility conditions for the sets X and Y which guarantee a good approximation.

1.2.1 Derivation of the directional single- and multi-level approximation

We follow an idea which first appeared in [8] and [15] and was later used in [7] and [30] among other. It states that the oscillatory part $\exp(i\kappa|x-y|)$ of f can be smoothened by a plane wave term $\exp(-i\kappa\langle x-y, c\rangle)$ in a cone around c. We discuss this in detail here.

Let $c \in \mathbb{R}^3$ be a vector with $|c| = 1$. We can rewrite the Helmholtz kernel f by

$$f(x,y) = f_c(x,y)\exp(i\kappa\langle x-y, c\rangle), \tag{1.13}$$

where f_c is defined by

$$f_c(x,y) := f(x,y)\exp(-i\kappa\langle x-y, c\rangle) = \frac{\exp(i\kappa(|x-y| - \langle x-y, c\rangle))}{4\pi|x-y|}. \tag{1.14}$$

The modified kernel function f_c is somewhat smoother than f on suitable boxes X and Y due to the additional plane wave term. In fact, if two vectors x and y satisfy $(x-y)/|x-y| \approx c$, then this plane wave term behaves like the reciprocal of the oscillating factor of f. Therefore it dampens the oscillations locally.

Due to its higher smoothness we interpolate f_c instead of f on suitable axis-parallel boxes X and Y and get

$$\tilde{f}^{(m)}_{c,X,Y}(x,y) := \mathcal{I}^{(m)}_X \otimes \mathcal{I}^{(m)}_Y[f_c](x,y) = \sum_{\nu\in M}\sum_{\mu\in M} f_c(\xi_{X,\nu}, \xi_{Y,\mu}) L^{(m)}_{X,\nu}(x) L^{(m)}_{Y,\mu}(y).$$

Here we used the interpolation operators and their representation given in Section 1.1. Finally we get an approximation of f on the set $X \times Y$ by substituting f_c in (1.13) with $\tilde{f}^{(m)}_{c,X,Y}$ yielding

$$\begin{aligned}\tilde{f}^{(m)}_{X,Y}(x,y) &:= \sum_{\nu\in M}\sum_{\mu\in M} f_c(\xi_{X,\nu},\xi_{Y,\mu}) L^{(m)}_{X,\nu}(x) L^{(m)}_{Y,\mu}(y)\exp(i\kappa\langle x-y,c\rangle)\\ &= \sum_{\nu\in M}\sum_{\mu\in M} f_c(\xi_{X,\nu},\xi_{Y,\mu})\left[L^{(m)}_{X,\nu}(x)\exp(i\kappa\langle x,c\rangle)\right]\left[L^{(m)}_{Y,\mu}\exp(-i\kappa\langle y,c\rangle)\right].\end{aligned}$$

This leads to the following definiton.

Definition 1.11. Let $Z \subset \mathbb{R}^3$ be an axis-parallel box. Let $c \in \mathbb{R}^3$ be a direction satisfying $|c| = 1$ and let $\{L^{(m)}_{Z,\nu}\}_{\nu\in M}$ be the Lagrange polynomials corresponding to a given set of interpolation points. Then the directionally modified Lagrange polynomials $L^{(m)}_{Z,c,\nu}$ are defined by

$$L^{(m)}_{Z,c,\nu} := L^{(m)}_{Z,\nu}\exp\left(i\kappa\langle\cdot,c\rangle\right). \tag{1.15}$$

Moreover, define the directional interpolation operator by

$$\mathcal{I}^{(m)}_{Z,c} : C(Z)\to C(Z), \qquad \mathcal{I}^{(m)}_{Z,c}[g] := \exp(i\kappa\langle\cdot,c\rangle)\mathcal{I}^{(m)}_{Z}[\exp(-i\kappa\langle\cdot,c\rangle)g], \tag{1.16}$$

and the directional tensor interpolation operator $\mathcal{I}^{(m)}_{X\times Y,c}$ for two boxes X and Y by

$$\mathcal{I}^{(m)}_{X\times Y,c} := \mathcal{I}^{(m)}_{X,c}\otimes\mathcal{I}^{(m)}_{Y,-c}. \tag{1.17}$$

Then the directional single-level approximation $\tilde{f}^{(m)}_{X,Y}$ of the Helmholtz kernel f on the set $X \times Y$ is defined by

$$\tilde{f}^{(m)}_{X,Y} := \mathcal{I}^{(m)}_{X\times Y,c}[f] = \sum_{\nu\in M}\sum_{\mu\in M} f_c(\xi_{X,\nu},\xi_{Y,\mu}) L^{(m)}_{X,c,\nu}(x)\overline{L^{(m)}_{Y,c,\mu}}. \tag{1.18}$$

Similarly as in Section 1.1.3 we can use the single-level approximation $\tilde{f}^{(m)}_{X,Y}$ of f to approximate a matrix $A \in \mathbb{C}^{N_X\times N_Y}$ with entries

$$A[j,k] = f(x_j,y_k) = \frac{\exp(i\kappa|x_j-y_k|)}{4\pi|x_j-y_k|} \tag{1.19}$$

by the matrix $\widetilde{A}$ with entries

$$\widetilde{A}[j,k] = \tilde{f}^{(m)}_{X,Y}(x_j,y_k) = \sum_{\nu\in M}\sum_{\mu\in M} f_c(\xi_{X,\nu},\xi_{Y,\mu}) L^{(m)}_{X,c,\nu}(x_j)\overline{L^{(m)}_{Y,c,\mu}}(y_k),$$

for some points $\{x_j\}_{j=1}^{N_X} \subset X$ and $\{y_k\}_{k=1}^{N_Y} \subset Y$ where X and Y are suitable boxes. Analogously to (1.12) we can write this in matrix notation, i.e.

$$\widetilde{A} = L_{X,c} A_{c,X,Y} L^*_{Y,c}, \tag{1.20}$$

where we define the coupling matrix $A_{c,X,Y} \in \mathbb{C}^{(m+1)^3 \times (m+1)^3}$ by

$$A_{c,X,Y}[j,k] := f_c(\xi_{X,\alpha_j}, \xi_{Y,\beta_k}), \quad j,k \in \{1,...,(m+1)^3\}, \tag{1.21}$$

and the directional interpolation matrices $L_{X,c} \in \mathbb{C}^{N_X \times (m+1)^3}$ and $L_{Y,c} \in \mathbb{C}^{N_Y \times (m+1)^3}$ by

$$\begin{aligned} L_{X,c}[j,k] &:= L^{(m)}_{X,c,\alpha_k}(x_j), \quad j \in \{1,...,N_X\}, k \in \{1,...,(m+1)^3\}, \\ L_{Y,c}[j,k] &:= L^{(m)}_{Y,c,\beta_k}(y_j), \quad j \in \{1,...,N_Y\}, k \in \{1,...,(m+1)^3\}. \end{aligned} \tag{1.22}$$

In Chapter 2 we describe how to approximate a matrix generated by the Helmholtz kernel by partitioning it appropriately. For this purpose we will need directional interpolation matrices for a sequence of boxes $\{X_j\}_{j=0}^L$ in $\mathbb{R}^3$ satisfying $X_0 \supset ... \supset X_L$ and a sequence of directions $\{c_j\}_{j=0}^L$, where c_{j+1} is a direction close to c_j for all $j \in \{0,...,L-1\}$ in some appropriate sense. Assembling or storing all these matrices directly would not be efficient. Hence we want to find a way to express L_{X_j,c_j} by $L_{X_{j+1},c_{j+1}}$ for points in $X_j \cap X_{j+1}$. In other words, for all $\nu \in M$ we want to express the function $L^{(m)}_{X_j,c_j,\nu}$ on the box X_{j+1} by the functions $\{L^{(m)}_{X_{j+1},c_{j+1},\tilde{\nu}}\}_{\tilde{\nu} \in M}$. Again we follow [7, Section 2.2.2].

Let us rewrite $L^{(m)}_{X_j,c_j,\nu}(x)$ by

$$L^{(m)}_{X_j,c_j,\nu}(x) = \exp(i\kappa\langle x, c_{j+1}\rangle) \left[\exp(i\kappa\langle x, c_j - c_{j+1}\rangle) L^{(m)}_{X_j,\nu}(x)\right].$$

If c_{j+1} is sufficiently close to c_j the term in square brackets is smooth and we can interpolate it in the box X_{j+1}, yielding

$$\exp(i\kappa\langle x, c_j - c_{j+1}\rangle) L^{(m)}_{X_j,\nu}(x) \approx \sum_{\tilde{\nu} \in M} \exp(i\kappa\langle \xi_{X_{j+1},\tilde{\nu}}, c_j - c_{j+1}\rangle) L^{(m)}_{X_j,\nu}(\xi_{X_{j+1},\tilde{\nu}}) L^{(m)}_{X_{j+1},\tilde{\nu}}(x).$$

Using the definition of $L^{(m)}_{X_{j+1},c_{j+1},\tilde{\nu}}$ in (1.15) we get the desired result

$$L^{(m)}_{X_j,c_j,\nu}(x) \approx \sum_{\tilde{\nu} \in M} \left(\left[\exp(i\kappa\langle \xi_{X_{j+1},\tilde{\nu}}, c_j - c_{j+1}\rangle) L^{(m)}_{X_j,\nu}(\xi_{X_{j+1},\tilde{\nu}})\right] L^{(m)}_{X_{j+1},c_{j+1},\tilde{\nu}}(x) \right),$$

or in short form

$$L^{(m)}_{X_j,c_j,\nu}|_{X_{j+1}} \approx \mathcal{I}^{(m)}_{X_{j+1},c_{j+1}} \left[L^{(m)}_{X_j,c_j,\nu}\right], \tag{1.23}$$

where we make use of the interpolation operator defined in (1.16).

By means of equation (1.23) we can approximate the directional interpolation matrices in the following way. Suppose $L_{X_j,c_j}|_{\hat{X}_{j+1} \times (m+1)^3}$ is the submatrix of L_{X_j,c_j} containing all rows corresponding to the points in X_{j+1}. Then there holds

$$L_{X_j,c_j}|_{\hat{X}_{j+1} \times (m+1)^3} \approx L_{X_{j+1},c_{j+1}} E_{X_{j+1},c_j}, \tag{1.24}$$

where the entries of the transfer matrix $E_{X_{j+1},c_j} \in \mathbb{C}^{(m+1)^3\times(m+1)^3}$ are defined by

$$E_{X_{j+1},c_j}[k,\ell] := \exp(i\kappa\langle\xi_{X_{j+1},\nu_k}, c_j - c_{j+1}\rangle)L^{(m)}_{X_j,\nu_\ell}(\xi_{X_{j+1},\nu_k}), \tag{1.25}$$

for all $k,\ell \in \{1,...,(m+1)^3\}$.

Finally we can define the directional multi-level approximation of f on a suitable pair of boxes $X \times Y$ by combining (1.18) and (1.23).

Definition 1.12 (Directional multi-level approximation)**.** Let two sequences of boxes $\{X_j\}_{j=0}^L$ and $\{Y_j\}_{j=0}^L$ be given such that $X = X_0 \supset ... \supset X_L$ and $Y = Y_0 \supset ... \supset Y_L$. Let furthermore $\{c_j\}_{j=0}^L$ be a sequence of directions in $\mathbb{R}^3$. Then the directional multi-level approximation $\hat{f}^{(m)}_{X,Y}|_{X_L\times Y_L}$ of f on the set $X_L \times Y_L$ is defined by

$$\hat{f}^{(m)}_{X,Y}|_{X_L\times Y_L} = \mathcal{I}^{(m)}_{X_L\times Y_L,c_L} \circ ... \circ \mathcal{I}^{(m)}_{X_0\times Y_0,c_0}[f]. \tag{1.26}$$

1.2.2 Directional admissibility conditions

The question arises under which conditions on the boxes X and Y and the direction c the approximation (1.26) yields good results. Following [7] we introduce three useful criteria, which will allow us to prove exponential convergence of the approximation error in Theorem 1.26.

Definition 1.13 (Directional admissibility [7, cf. Section 3.3])**.** Let $X, Y \subset \mathbb{R}^3$ be two axis-parallel boxes and let a direction $c \in \mathbb{R}^3$ with $|c| = 1$ or $c = 0$ be given. Denote the midpoints of X and Y by m_X and m_Y, respectively. Let two constants $\eta_1 > 0$ and $\eta_2 > 0$ be chosen. Define the diameter $\operatorname{diam}(X)$ and the distance $\operatorname{dist}(X,Y)$ by

$$\begin{aligned}\operatorname{diam}(X) &= \sup_{x_1,x_2\in X} |x_1 - x_2|,\\ \operatorname{dist}(X,Y) &= \inf_{x\in X, y\in Y} |x - y|.\end{aligned}$$

We say that X and Y are directionally admissible with respect to c if the separation criterion

$$\max\{\operatorname{diam}(X), \operatorname{diam}(Y)\} \le \eta_2 \operatorname{dist}(X,Y), \tag{A1}$$

and the two cone admissibility criteria

$$\kappa\left|\frac{m_X - m_Y}{|m_X - m_Y|} - c\right| \le \frac{\eta_1}{\max\{\operatorname{diam}(X), \operatorname{diam}(Y)\}}, \tag{A2}$$

$$\kappa \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\}^2 \le \eta_2 \operatorname{dist}(X,Y), \tag{A3}$$

are satisfied. Criteria (A1) – (A3) are denoted as the directional admissibility criteria.

In Theorems 1.15 and 1.26 we will show that the criteria (A1) – (A3) are sufficient to guarantee exponential convergence of the directional multi-level and single-level approximation error in the interpolation degree m. However (A1) – (A3) do not hold for arbitrary boxes X and Y. If we approximate a matrix A as in (1.19) for boxes which do not satisfy these criteria, we will not get satisfactory results in general. However, we can partition the matrix and use the directional approximation for suitable submatrices. We discuss this in Chapter 2. Here we continue with the discussion of the admissibility criteria.

Let us start with (A1). This is a classical separation criterion, which can for example be found in [24, Section 4.1.3]. It guarantees that the boxes X and Y do not overlap, which is desirable because for $x = y$ the term $1/|x - y|$ in the Helmholtz kernel is singular. This separation of boxes allows for low rank approximations of many kernels corresponding to fundamental solutions of elliptic partial differential equations, e.g. the Laplace kernel $(4\pi|x-y|)^{-1}$. The exponential term $\exp(i\kappa|x-y|)$, however, prohibits low rank approximations of the Helmholtz kernel for large boxes and large κ, even if the boxes satisfy (A1).

Next we discuss (A3). We observe that (A1) and (A3) are closely connected. Both control the distance of two boxes X and Y. The bound on the distance in (A3), however, does not only depend on the diameter of the boxes X and Y but also on the wave number κ. Furthermore we note that (A3) is more restrictive for large box diameters while (A1) is more restrictive for small diameters. Indeed, (A1) follows immediately from (A3) if $\kappa \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\} > 1$. On the other hand (A3) follows from (A1) if $\kappa \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\} \leq 1$.

The criterion (A3) can also be interpreted differently. It can be written in the form

$$\frac{\max\{\operatorname{diam}(X), \operatorname{diam}(Y)\}}{\operatorname{dist}(X,Y)} \leq \frac{\eta_2}{\kappa \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\}}.$$

As discussed in [5, Section 3] this can be understood as a bound on the angle between all vectors $x - y$ for $x \in X$ and $y \in Y$ that shrinks if κ or $\max\{\operatorname{diam}(X), \operatorname{diam}(Y)\}$ increases. Hence, (A3) guarantees that the angle between $x - y$ and a direction c is small if the angle between the difference of the midpoints $m_X - m_Y$ and c is small, which we will enforce by (A2).

The effect of (A3) is schematically depicted in 2D in Figure 1.1. For a fixed box X, a fixed wave number $\kappa > 0$ and two parameters η_2 we marked all boxes Y such that X and Y satisfy (A3). We see that for larger η_2 the admissibility condition is less restrictive, i.e. more boxes Y are admissible with respect to X. Visualizing (A1) would result in a similar figure with slight changes of the allowed distances.

Finally we discuss the criterion (A2). This criterion will be used to assign two boxes X and Y a suitable direction c from a given set of directions for the directional approximation. If we only need to compute the directional approximation $L_{X,c} A_{c,X,Y} L^*_{Y_k,c}$ of a matrix A as in (1.19) for points in two boxes X and Y we can

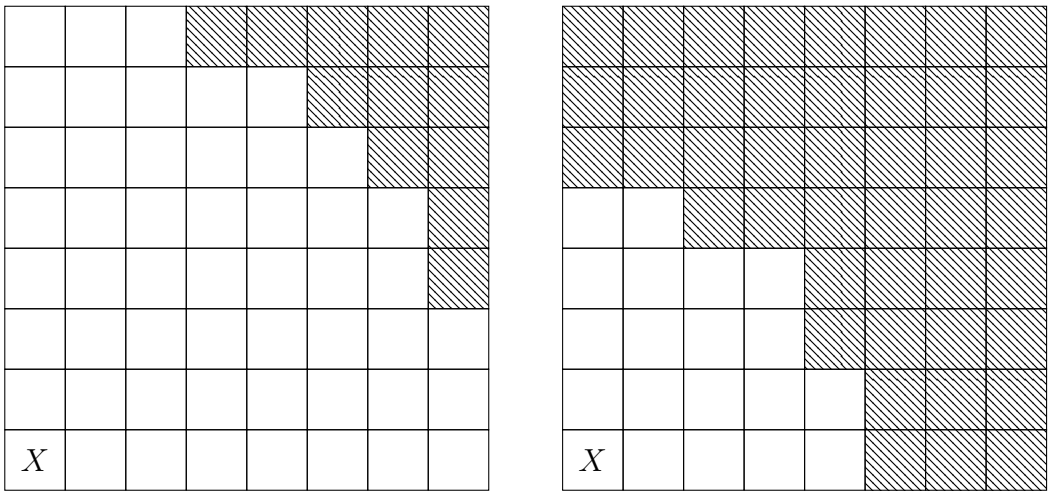

Figure 1.1: Schematic visualization of the admissibility condition (A3): The hatched boxes are such that (A3) is fulfilled for the given box X and some values of κ and η_2. In the right image η_2 is chosen larger.

simply choose the direction $c = (m_X - m_Y)/|m_X - m_Y|$ and do not need the criterion (A2). However, in the setting of Chapter 2 we will have to approximate such matrices for points in a given box X and several boxes Y_k at the same time using the directional multi-level approximation. Choosing the same direction c for multiple boxes Y_k is then useful, because it enables us to reuse the transfer matrices defined in (1.25), i.e. we do not need individual transfer matrices for each box Y_k. The criterion (A2) tells us if a given direction c is sufficiently close to the normalized difference of the midpoints $(m_X - m_Y)/|m_X - m_Y|$ such that the directional approximation of a matrix corresponding to points in the boxes X and Y with the direction c is reasonable. In particular we have to provide a set of directions, such that for each pair of boxes X and Y there exists a direction c in this set for which (A2) holds. Less directions are needed for this purpose if the boxes X and Y are small, since the right hand side of (A2) increases with decreasing diameters of the boxes X and Y.

In Section 2.1.2 we describe a possible strategy to construct suitable sets of directions. Here we continue with the discussion of (A2). The case

$$\frac{\eta_1}{\kappa \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\}} > 1, \tag{1.27}$$

is of special interest. If (1.27) holds, then (A2) is satisfied for $c = 0$. This means that no specific direction is needed and we can use standard interpolation instead of the directional interpolation. The condition (1.27) obviously depends on the boxes X and Y. This leads to the following definition.

Definition 1.14. Let X and Y be two axis-parallel boxes in $\mathbb{R}^3$ and $\eta_1 > 0$. We say that the pair (X, Y) is in the low frequency regime (with respect to η_1) if

$$\frac{\eta_1}{\kappa \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\}} > 1. \tag{1.28}$$

Else we say that (X, Y) is in the high frequency regime (with respect to η_1).

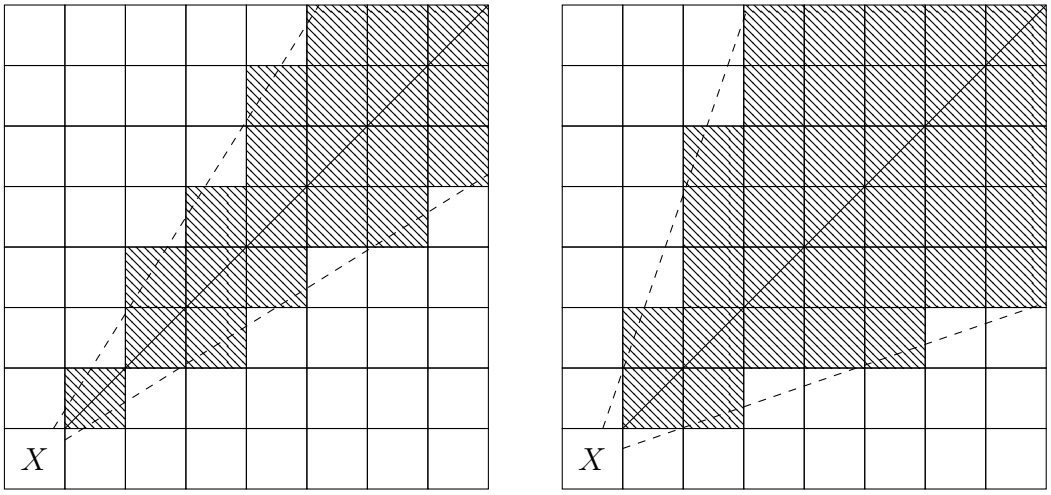

Figure 1.2: Schematic visualization of the admissibility condition (A2): Both pictures show a box X, a direction c which is depicted by the diagonal solid line, and some hatched boxes. These hatched boxes are such that (A2) is fulfilled for given X and c, and some values of κ and η_1. The variable η_1 is chosen larger in the right image. The dashed lines bound the area in which the midpoints of a box must lie so that (A2) holds.

The effect of criterion (A2) is schematically depicted in Figure 1.2 for the 2D case and two different choices of η_1. We observe that for a fixed box X and a fixed direction c only those boxes are admissible whose midpoints are contained in a circular arc around c. In 3D the same holds for cones instead of arcs. In Figure 1.2 it becomes clear that a larger parameter η_1 leads to a larger cone, i.e. in total less directions are necessary to cover the whole space. However, the approximation quality suffers away from the central direction c of the cone, in particular for large η_1. Indeed we will show in the analysis of the directional single-level approximation error in Section 1.3.1 that larger η_1 lead to worse constants in the error estimate.

1.3 Error analysis of the directional multi-level approximation

Having introduced the directional multi-level approximation in (1.26) in the previous section, it is now our goal to analyze it. In particular our goal is to estimate the error

$$\|f - \hat{f}^{(m)}_{X,Y}\|_{\infty, X_L \times Y_L}. \tag{1.29}$$

For this purpose we follow again the lines of [7, Sections 3 and 5]. In contrast to this work, we do not deal with general interpolation points, but choose the Chebyshev nodes as stated in Section 1.1. Furthermore we simplify some results, for example the estimate in Theorem 1.35, which corresponds to [7, Theorem 5.7]. For small interpolation degrees m however, this has no effect on the main result presented in Theorem 1.15. On the other hand we refine some results, for example the estimate in Theorem 1.26 concerning the approximation error of the directional single-level approximation, which is similar to [7, Corollary 3.11]. This leads to a better rate of convergence of the error in (1.29). Further differences are emphasized throughout this chapter.

We start the discussion by presenting the main result of this section, which is the estimate of (1.29). Its proof motivates the necessity of the discussion of the directional single-level approximation (1.18) in Section 1.3.1 and the directional reinterpolation in Section 1.3.2. We note that throughout the discussion we assume the wave number κ to be a positive real number.

Theorem 1.15. *Let $\{X_j\}_{j=0}^L$ and $\{Y_j\}_{j=0}^L$ be two sequences of axis-parallel boxes in $\mathbb{R}^3$ such that $X = X_0 \supset ... \supset X_L$, $Y = Y_0 \supset ... \supset Y_L$ and*

$$X_k := [a^{(X)}_{k,1}, b^{(X)}_{k,1}] \times ... \times [a^{(X)}_{k,3}, b^{(X)}_{k,3}], \qquad Y_k := [a^{(Y)}_{k,1}, b^{(Y)}_{k,1}] \times ... \times [a^{(Y)}_{k,3}, b^{(Y)}_{k,3}],$$

for all $k \in \{0, ..., L\}$. Assume that there exists a $\bar{q} < 1$ such that

$$\frac{b^{(X)}_{k,j} - a^{(X)}_{k,j}}{b^{(X)}_{k-1,j} - a^{(X)}_{k-1,j}} \leq \bar{q}, \qquad \frac{b^{(Y)}_{k,j} - a^{(Y)}_{k,j}}{b^{(Y)}_{k-1,j} - a^{(Y)}_{k-1,j}} \leq \bar{q}, \tag{1.30}$$

for all $k \in \{1, ..., L\}$ and $j \in \{1, ..., 3\}$. Let $\sigma < 1$ be such that

$$\max\{s(X_k), s(Y_k)\} \leq \sigma \max\{\operatorname{diam}(X_k), \operatorname{diam}(Y_k)\}, \tag{1.31}$$

for all $k \in \{0, ..., L\}$, where for an axis parallel box $B = [a_1, b_1] \times ... \times [a_3, b_3]$ we define by $s(B)$ its longest edge, i.e.

$$s(B) := \max_{j \in \{1,...,3\}} \{|b_j - a_j|\}.$$

Let $\{c_k\}_{k=0}^L$ *be a sequence of directions in* $\mathbb{R}^3$ *such that either* $|c_k| = 1$ *or* $c_k = 0$ *for all* $k \in \{0, ..., L\}$. *Assume that there exists a* $\gamma \in \mathbb{R}$ *such that*

$$\kappa \max\{\operatorname{diam}(X_{k-1}), \operatorname{diam}(Y_{k-1})\}|c_k - c_{k-1}| \leq \gamma. \tag{1.32}$$

Assume further, that X, Y, *and* c_0 *satisfy the admissibility criteria* (A1)–(A3) *and that* X_k, Y_k *and* c_k *satisfy* (A2) *for all* $k \in \{1, ..., L\}$ *for some* $\eta_1 > 0$, $\eta_2 > 0$ *independent of* k. *Let the directional multi-level approximation* $\hat{f}_{X,Y}^{(m)}$ *of the Helmholtz kernel* f *on the set* $X_L \times Y_L$ *be defined by* (1.26) *and let*

$$\hat{\rho} := 1 + \frac{2}{\eta_2}. \tag{1.33}$$

Then there exists a constant C *depending only on* η_1, η_2, $\bar{q}$, σ, *and* L, *and an* $m_0 \in \mathbb{N}$ *depending only on* γ *and* $\bar{q}$, *such that for all* $m \geq m_0$

$$\|f - \hat{f}_{X,Y}^{(m)}\|_{\infty, X_L \times Y_L} \leq \frac{C}{\operatorname{dist}(X, Y)} \hat{\rho}^{-(m+1)}. \tag{1.34}$$

Remark 1.16. Let us discuss the assumptions of Theorem 1.15.

- In equation (1.30) we require the edges $[a_{k,j}^{(X)}, b_{k,j}^{(X)}]$ and $[a_{k,j}^{(Y)}, b_{k,j}^{(Y)}]$ of the boxes X_k and Y_k, respectively, to decrease for increasing k in each step at least by a factor $\bar{q}$. In our later application we will focus on boxes in a uniform box cluster tree (cf. Algorithm 2.3). In Theorem 2.29 we will see that in this setting (1.30) is always satisfied with $\bar{q} = 1/2$.
- Equation (1.31) is satisfied in particular if $s(X_k) \leq \sigma \operatorname{diam}(X_k)$ as well as $s(Y_k) \leq \sigma \operatorname{diam}(Y_k)$ for all $k \in \{0, ..., L\}$. Hence one can think of σ as a measure of the isotropy of the boxes X_k and Y_k. It is easy to see that for cubes we can choose $\sigma = 1/\sqrt{3}$, while σ has to be chosen close to 1 for boxes where one edge is considerably larger than the others. This is undesirable as for $\sigma \to 1$ the constant C in (1.34) becomes unbounded.
- The directions c_k in Theorem 1.15 are allowed to be zero. In fact if the pair of boxes (X_k, Y_k) is in the low frequency regime according to Definition 1.14, then (A2) holds for $c_k = 0$ as we discussed in Section 1.2.2. In applications there is typically a $\hat{k} \in \mathbb{N}$ such that $c_k \neq 0$ for all $k \leq \hat{k}$ and $c_k = 0$ for all $k > \hat{k}$, since we require the diameters of X_k and Y_k to shrink for increasing k,.
- In equation (1.32) we assure that the directions c_{k-1} and c_k with $k \in \{1, ..., L\}$ are close to each other, i.e. their difference is bounded. The bound in (1.32) depends on the wave number $\kappa > 0$ and the diameters of the boxes X_{k-1} and Y_{k-1}. As these diameters decrease for increasing k due to (1.30), we allow larger differences of the directions for larger k. In Section 2.1.2 we will introduce some sets of directions and discuss this assumption further.

Proof of Theorem 1.15. Let us start by rewriting the difference $f - \hat{f}^{(m)}_{X,Y}$ on the set $X_L \times Y_L$ as a telescopic sum. Similarly as in [7, Section 2.2.3] we have

$$\begin{aligned} f - \hat{f}^{(m)}_{X,Y} = f - \mathcal{I}^{(m)}_{X_L \times Y_L, c_\ell}[f] \\ + \sum_{\ell=0}^{L-1} \left(\mathcal{I}^{(m)}_{X_L \times Y_L, c_L} \circ ... \circ \mathcal{I}^{(m)}_{X_{\ell+1} \times Y_{\ell+1}, c_{\ell+1}} \right) \left[f - \mathcal{I}^{(m)}_{X_\ell \times Y_\ell, c_\ell}[f] \right]. \end{aligned}$$

We proceed by applying the triangle inequality and the operator norm inequality to get

$$\begin{aligned} \|f - \hat{f}^{(m)}_{X,Y}\|_{\infty, X_L \times Y_L}\| \leq \|f - \mathcal{I}^{(m)}_{X_L \times Y_L, c_L}[f]\|_{\infty, X_L \times Y_L} + \\ \sum_{\ell=0}^{L-1} \|\mathcal{I}^{(m)}_{X_L \times Y_L, c_L} \circ ... \circ \mathcal{I}^{(m)}_{X_{\ell+1} \times Y_{\ell+1}, c_{\ell+1}}\|_{\infty, L \leftarrow \ell} \|f - \mathcal{I}^{(m)}_{X_\ell \times Y_\ell, c_\ell}[f]\|_{\infty, X_\ell \times Y_\ell}, \end{aligned} \tag{1.35}$$

where we write $\|\cdot\|_{\infty, L \leftarrow \ell}$ for the operator norm $\|\cdot\|_{\infty, X_L \times Y_L \leftarrow X_\ell \times Y_\ell}$. Therefore, we need to control the norms of the reinterpolation operators $\mathcal{I}^{(m)}_{X_L \times Y_L, c_L} \circ ... \circ \mathcal{I}^{(m)}_{X_{\ell+1} \times Y_{\ell+1}, c_{\ell+1}}$ as well as the approximation error of the single-level approximations $\mathcal{I}^{(m)}_{X_\ell \times Y_\ell, c_\ell}[f]$ for all $\ell \in \{0, ..., L-1\}$. We start with the former.

Due to the assumptions on the boxes $\{X_k\}_{k=0}^L$, $\{Y_k\}_{k=0}^L$ and the directions $\{c_k\}_{k=0}^L$ we can apply Theorem 1.35 and (1.80) to estimate the norm of the reinterpolation operators for all $\ell \in \{0, ..., L-1\}$ by

$$\|\mathcal{I}^{(m)}_{X_L \times Y_L, c_L} \circ ... \circ \mathcal{I}^{(m)}_{X_{\ell+1} \times Y_{\ell+1}, c_{\ell+1}}\|_{\infty, L \leftarrow \ell} \leq \Lambda_m^6 C(\bar{q}, L)^6, \tag{1.36}$$

for a constant $C(\bar{q}, L)$ depending only on $\bar{q}$ and L, if $m \geq m_0$, where m_0 depends solely on γ and $\bar{q}$.

Next we want to estimate the single-level approximation error. For this purpose we show that all three admissibility criteria (A1) – (A3) hold for X_ℓ, Y_ℓ and c_ℓ, for all $\ell \in \{0, ..., L\}$, which allows us to apply Theorem 1.26. By assumption (A2) already holds for all such ℓ, whereas (A1) and (A3) only hold for $\ell = 0$. However, we observe that

$$\operatorname{diam}(X_\ell) \leq \operatorname{diam}(X_0), \qquad \operatorname{diam}(Y_\ell) \leq \operatorname{diam}(Y_0),$$

which holds true due to the inclusions $X_\ell \subset X_{\ell-1}$ and $Y_\ell \subset Y_{\ell-1}$ for all $\ell \in \{1, ..., L\}$. Furthermore these inclusions imply

$$\operatorname{dist}(X_0, Y_0) \leq \operatorname{dist}(X_\ell, Y_\ell). \tag{1.37}$$

As (A1) holds for $X = X_0$ and $Y = Y_0$ we get

$$\begin{aligned} \max\{\operatorname{diam}(X_\ell), \operatorname{diam}(Y_\ell)\} &\leq \max\{\operatorname{diam}(X_0), \operatorname{diam}(Y_0)\} \\ &\leq \eta_2 \operatorname{dist}(X_0, Y_0) \leq \eta_2 \operatorname{dist}(X_\ell, Y_\ell), \end{aligned}$$

which is (A1) for X_ℓ and Y_ℓ. Analogously (A3) follows for all $\ell > 0$. Therefore, we can apply Theorem 1.26 to estimate the single-level approximations

$$\|f - \mathcal{I}^{(m)}_{X_\ell \times Y_\ell, c_\ell}[f]\|_{\infty, X_\ell \times Y_\ell} \leq \frac{C_{\mathrm{sl}}(m, \sigma, \eta_1, \eta_2)}{\operatorname{dist}(X_\ell, Y_\ell)} \hat{\rho}^{-(m+1)}, \tag{1.38}$$

where $C_{\mathrm{sl}}(m, \sigma, \eta_1, \eta_2)$ is a function which decays exponentially in m.

Inserting the two estimates (1.36) and (1.38) in (1.35) and using the estimate (1.37) yields

$$\|f - \hat{f}^{(m)}_{X,Y}\|_{\infty, X_L \times Y_L} \leq \frac{C_{\mathrm{sl}}(m, \sigma, \eta_1, \eta_2)}{\operatorname{dist}(X_0, Y_0)} \left(1 + \sum_{\ell=0}^{L-1} \Lambda_m^6 C(\bar{q}, L)^6\right) \hat{\rho}^{-(m+1)},$$

for all $m \geq m_0$. Due to the exponential decay of C_{sl} in m and the logarithmic behaviour of Λ_m seen in (1.8) the supremum

$$C := \sup_{m \in \mathbb{N}} \left\{ C_{\mathrm{sl}}(m, \sigma, \eta_1, \eta_2) \left(1 + L\ \Lambda_m^6\ C(\bar{q}, L)^6\right) \right\}$$

is bounded. With this constant C, which depends only on η_1, η_2, $\bar{q}$, σ and L, the assertion follows. □

In the following two sections we derive the results concerning the directional single-level approximation and the directional reinterpolation, which we used in the previous proof.

1.3.1 Error analysis of the directional single-level approximation

The estimate of the error (1.34) of the directional multi-level approximation relies on the estimate of the error of the directional single-level approximation $\tilde{f}^{(m)}_{X,Y}$ given in Definition 1.11. In [7, Section 3] an error analysis of this single-level approximation based on interpolation is given, which we repeat here. The procedure is as follows: First the error estimate is reduced to a one dimensional interpolation error estimate of certain functions $f_{dp}(t)$ on $[-1, 1]$. These functions have a holomorphic extension on suitable domains, where they can be bounded uniformly, if the boxes X and Y satisfy the admissibility criteria (A1) – (A3). This allows us to use the results from Section 1.1.2 to proof Theorem 1.26, which is the main result in this section.

We start our considerations with the following lemma, which relates the tensor interpolation error and the univariate interpolation error.

Lemma 1.17 ([7, Lemma 3.3])**.** *Let $B = \prod_{i=1}^6 [a_i, b_i]$ be a non-empty box in $\mathbb{R}^6$ and let $g \in C(B)$. Define for all $j \in \{1, ..., 6\}$ and $x \in B$ the functions*

$$g_{x,j} : [a_j, b_j] \to \mathbb{C}, \qquad t \mapsto g(x_1, ..., x_{j-1}, t, x_{j+1}, ..., x_6) \tag{1.39}$$

Then

$$\|g - \mathcal{I}_B^{(m)}[g]\|_{\infty,B} \leq \sum_{j=1}^{6} \Lambda_m^{j-1} \sup_{x\in B} \|g_{x,j} - \mathcal{I}_{[a_j,b_j]}^{(m)}[g_{x,j}]\|_{\infty,[a_j,b_j]}, \tag{1.40}$$

where Λ_m is the Lebesgue constant given in (1.6)*.*

Proof. We start by rewriting the difference $g - \mathcal{I}_B^{(m)}[g]$ into the telescopic sum

$$\begin{aligned} g - \mathcal{I}_B^{(m)}[g] = &\left(\left(I - \mathcal{I}_{[a_1,b_1]}^{(m)}\right) \otimes I \otimes ... \otimes I\right)[g] \\ &+ \sum_{j=2}^{6} \left(\mathcal{I}_{[a_1,b_1]}^{(m)} \otimes ... \otimes \mathcal{I}_{[a_{j-1},b_{j-1}]}^{(m)} \otimes \left(I - \mathcal{I}_{[a_j,b_j]}^{(m)}\right) \otimes I \otimes ... \otimes I\right)[g], \end{aligned}$$

where I is the identity on the corresponding sets and we use the multilinearity of the tensor product.

Next we note, that we can write the norm

$$\begin{aligned} \|g\|_{\infty,B} &= \sup_{x\in B} |g(x)| = \sup_{\hat{x}^{(j)}\in \hat{B}_j} \left(\|g(x_1, ..., x_{j-1}, \cdot, x_{j+1}, ..., x_6)\|_{\infty,[a_j,b_j]}\right) \\ &= \sup_{x\in B} \|g_{x,j}\|_{\infty,[a_j,b_j]}, \end{aligned} \tag{1.41}$$

for all $j \in \{1, ..., 6\}$, where $\hat{x}^{(j)} := (x_1, ..., x_{j-1}, x_{j+1}, ..., x_6)$ for all points $x \in B$ and $\hat{B}_j := \prod_{i=1, i\neq j}^{6} [a_i, b_i]$. An immediate consequence is

$$\begin{aligned} &\left\| \left(\mathcal{I}_{[a_1,b_1]}^{(m)} \otimes ... \otimes \mathcal{I}_{[a_{j-1},b_{j-1}]}^{(m)} \otimes \left(I - \mathcal{I}_{[a_j,b_j]}^{(m)}\right) \otimes I \otimes ... \otimes I\right)[g] \right\|_{\infty,B} \\ &\quad \leq \Lambda_m^{j-1} \left\| \left(I \otimes ... \otimes I \otimes \left(I - \mathcal{I}_{[a_j,b_j]}^{(m)}\right) \otimes I \otimes ... \otimes I\right)[g] \right\|_{\infty,B} \\ &\quad = \Lambda_m^{j-1} \sup_{x\in B} \|g_{x,j} - \mathcal{I}_{[a_j,b_j]}^{(m)}[g_{x,j}]\|_{\infty,[a_j,b_j]}, \end{aligned} \tag{1.42}$$

which follows by using (1.41) and the operator norm inequality (1.6) for the interpolation operators $\mathcal{I}_{[a_k,b_k]}^{(m)}$ repeatedly for all $k \in \{1, ..., j-1\}$. By estimating $\|g - \mathcal{I}_B^{(m)}[g]\|_{\infty,B}$ using the telescopic sum representation, the triangle inequality, and (1.42) we get (1.40). □

With the previous lemma we can proof the following proposition, which reduces the error estimate of the directional single-level approximation to a class of one dimensional interpolation error estimates.

Proposition 1.18 ([7, cf. Lemma 3.4])**.** *Let $X, Y \subset \mathbb{R}^3$ be two axis-parallel boxes such that*

$$X = \prod_{i=1}^{3} [a_i, b_i], \qquad Y = \prod_{i=4}^{6} [a_i, b_i].$$

Let $s_{\max}$ be the length of the longest edge of X and Y, i.e.

$$s_{\max} := \max\{|b_i - a_i| : i \in \{1, ..., 6\}\}. \tag{1.43}$$

Define the functions f_{dp} for two vectors $d, p \in \mathbb{R}^3$ by

$$f_{dp} : [-1, 1] \to \mathbb{C}, \qquad t \mapsto \frac{\exp(i\kappa(|d - tp| - \langle d - tp, c\rangle))}{4\pi|d - tp|}. \tag{1.44}$$

Suppose there exists an $\varepsilon \geq 0$ in $\mathbb{R}$ such that

$$\|f_{dp} - \mathcal{I}^{(m)}_{[-1,1]}[f_{dp}]\|_{\infty,[-1,1]} \leq \varepsilon, \tag{1.45}$$

for all vectors $d, p \in \mathbb{R}^3$ with $p \neq 0$ satisfying

$$2|p| \leq s_{\max}, \tag{1.46a}$$

$$d - tp \in X - Y = \{x - y : x \in X, y \in Y\}, \quad \text{for all } t \in [-1, 1]. \tag{1.46b}$$

Then the directional single-level approximation error is bounded by

$$\|f - \tilde{f}^{(m)}_{XY}\|_{\infty,X\times Y} \leq 6\Lambda_m^5 \varepsilon. \tag{1.47}$$

Proof. Note that by definition there holds

$$f(x, y) = f_c(x, y) \exp(i\kappa\langle x - y, c\rangle),$$
$$\tilde{f}^{(m)}_{XY}(x, y) = \mathcal{I}^{(m)}_{X\times Y}[f_c](x, y) \exp(i\kappa\langle x - y, c\rangle).$$

Due to $|\exp(i\kappa\langle x - y, c\rangle)| = 1$ for all $x \in X$ and $y \in Y$ we get

$$\|f - \tilde{f}^{(m)}_{XY}\|_{\infty,X\times Y} = \|f_c - \mathcal{I}^{(m)}_{X\times Y}[f_c]\|_{\infty,X\times Y}. \tag{1.48}$$

We apply Lemma 1.17 to estimate this last norm. For this purpose we consider the six-dimensional box $B := X \times Y = \prod_{j=1}^{6}[a_j, b_j]$ and think of f_c as a function in $C(B)$. In particular we identify $f_c(z) = f_c(x, y)$ for all $z \in B$ given by $z = (x, y)$. Then by (1.40) there holds

$$\|f_c - \mathcal{I}^{(m)}_{X\times Y}[f_c]\|_{\infty,X\times Y} \leq \Lambda_m^5 \sum_{j=1}^{6} \sup_{z\in B} \|[f_c]_{z,j} - \mathcal{I}^{(m)}_{[a_j,b_j]}[[f_c]_{z,j}]\|_{\infty,[a_j,b_j]}. \tag{1.49}$$

Here we used that $\Lambda_m^j < \Lambda_m^5$ for all $j \in \{0, ..., 5\}$.

Let $\Phi_{[a_j,b_j]}$ be the affine linear transformation from $[-1, 1]$ to $[a_j, b_j]$, defined in (1.5). Suppose we can find a pair $d, p \in \mathbb{R}^3$ for all $z \in B$ and $j \in \{1, ..., 6\}$ such that (1.46a) and (1.46b) hold and furthermore

$$[f_c]_{z,j} \circ \Phi_{[a_j,b_j]} - \mathcal{I}^{(m)}_{[a_j,b_j]}[[f_c]_{z,j}] \circ \Phi_{[a_j,b_j]} = f_{dp} - \mathcal{I}^{(m)}_{[-1,1]}[f_{dp}] \tag{1.50}$$

in $[-1,1]$. Then assumption (1.45) yields

$$\sup_{z\in B} \|[f_c]_{z,j} - \mathcal{I}^{(m)}_{[a_j,b_j]}[[f_c]_{z,j}]\|_{\infty,[a_j,b_j]} \leq \varepsilon,$$

for all $j \in \{1,...,6\}$. Combining this with (1.49) concludes the proof.

Hence, let us construct $d, p \in \mathbb{R}^3$ for given $z = (x,y) \in B$ and $j \in \{1,...,6\}$ such that (1.50) as well as (1.46a) and (1.46b) hold. For $j \in \{1,...,3\}$ equation (1.50) implies

$$\begin{aligned} (d-tp)[k] &= x_k - y_k, && k \neq j, \\ (d-tp)[k] &= \frac{a_j+b_j}{2} - y_j + t\,\frac{b_j-a_j}{2} && k = j, \end{aligned}$$

which immediately gives us d and p. Similarly, for $j \in \{4,...,6\}$ we get d and p from

$$\begin{aligned} (d-tp)[k] &= x_k - y_k, && k \neq j-3, \\ (d-tp)[k] &= x_k - \frac{a_j+b_j}{2} - t\,\frac{b_j-a_j}{2}, && k = j-3. \end{aligned}$$

In both cases (1.46a) and (1.46b) hold for the resulting vectors d and p. □

Our next goal is to show (1.45) for all $d, p \in \mathbb{R}^3$ such that (1.46a) and (1.46b) hold. As in [7, Sections 3.2 – 3.4] we will show that under these conditions f_{dp} has a holomorphic extension and use Theorem 1.9 to show the assertion.

In order to find a holomorphic extension of f_{dp} we have to find a suitable extension of the mapping $t \mapsto |d-tp| = \sqrt{\langle d-tp, d-tp\rangle_{\mathbb{R}}}$, where $\langle x,y\rangle_{\mathbb{R}} = \sum_{k=1}^3 x_i y_i$ for two vectors x and y in $\mathbb{C}^3$. For all $z = r\exp(i\varphi) \in \mathbb{C}$ we can define the principal branch of the square root of z by

$$\sqrt{z} = \sqrt{r}\exp(i\varphi/2).$$

This square root is holomorphic in $\mathbb{C} \setminus \mathbb{R}_{\leq 0}$, which can for example be seen by checking that it satisfies the Cauchy-Riemann equations in polar coordinates. Hence we can extend $t \mapsto |d-tp|$ to the function $n_{dp} : z \mapsto \sqrt{\langle d-zp, d-zp\rangle_{\mathbb{R}}}$, which is holomorphic for all $z \in U_{dp} := \{z \in \mathbb{C} : \langle d-zp, d-zp\rangle_{\mathbb{R}} \notin \mathbb{R}_{\leq 0}\}$. The following lemma characterizes n_{dp} and the set U_{dp}.

Lemma 1.19 ([7, Lemma 3.5])**.** *Let $d, p \in \mathbb{R}^3$ with $p \neq 0$. Define*

$$\begin{aligned} n_{dp}(z) : U_{dp} = \{z \in \mathbb{C} : \langle d-zp, d-zp\rangle_{\mathbb{R}} \notin \mathbb{R}_{\leq 0}\} &\to \mathbb{C} \\ z &\mapsto \sqrt{\langle d-zp, d-zp\rangle_{\mathbb{R}}}, \end{aligned} \tag{1.51}$$

and the complex number w by

$$w := w_r + iw_i, \qquad w_r := \langle d,p\rangle/|p|^2, \qquad w_i := \sqrt{|d|^2/|p|^2 - w_r^2}.$$

Then n_{dp} is well-defined and holomorphic on U_{dp} and there holds

$$U_{dp} = \mathbb{C} \setminus \{z = w_r + iy : y \in \mathbb{R}, |y| \geq w_i\}, \tag{1.52}$$

$$n_{dp}(z) = |p|\sqrt{(w - z)(\overline{w} - z)}. \tag{1.53}$$

Sketch of proof. The function n_{dp} is the decomposition of the principal branch of the square root and the holomorphic mapping

$$g : \mathbb{C} \to \mathbb{C}, \qquad z \mapsto \langle d - zp, d - zp\rangle_{\mathbb{R}},$$

which maps U_{dp} to the set $\mathbb{C} \setminus \mathbb{R}_{\leq 0}$, where the square root is well-defined and holomorphic. As a consequence n_{dp} is well-defined and holomorphic.

It remains to show the identities (1.52) and (1.53). A direct computation yields

$$\langle d - zp, d - zp\rangle_{\mathbb{R}} = |p|^2(w - z)(\overline{w} - z), \tag{1.54}$$

which immediately gives us (1.53). To show the remaining identity we set

$$S := \{z = w_r + iy : y \in \mathbb{R}, |y| \geq w_i\}.$$

It is an easy exercise to show that $g(z) \in \mathbb{R}_{\leq 0}$ holds if and only if $z \in S$. Hence we get $U_{dp} = \mathbb{C} \setminus S$, which is (1.52). □

Lemma 1.19 allows us to extend the function $f_{dp} : [-1, 1] \to \mathbb{C}$ defined in (1.44) holomorphically to the set U_{dp} by substituting $|d - tp|$ by $n_{dp}(t)$ and setting

$$\hat{f}_{dp}(z) = \frac{\exp(i\kappa(n_{dp}(z) - \langle d - zp, c\rangle_{\mathbb{R}}))}{4\pi n_{dp}(z)}. \tag{1.55}$$

For the sake of readability we will identify f_{dp} with its holomorphic extension $\hat{f}_{dp}$ in the rest of the section.

For the application of Theorem 1.9 we need that f_{dp} is holomorphic on a Bernstein elliptic disc D_ρ as defined in (1.9) for some suitable ρ. For this purpose we prove the following lemma.

Lemma 1.20 ([7, Lemmata 3.6 and 3.12])**.** *Let $d, p \in \mathbb{R}^3$ with $p \neq 0$ and U_{dp} be defined as in Lemma 1.19. Let $\zeta \in \mathbb{R}$ be defined by*

$$\zeta := \min\left\{\frac{|d - tp|}{|p|} : t \in [-1, 1]\right\}. \tag{1.56}$$

Let $r \in [0, \zeta)$, $\rho := \sqrt{1 + r^2} + r$ and define

$$U_r := \left\{z \in \mathbb{C} : \min_{t\in[-1,1]} |z - t| \leq r\right\}. \tag{1.57}$$

Then there holds

$$\overline{D_\rho} \subset U_r \subset U_{dp}. \tag{1.58}$$

Proof. We start with the inclusion $U_r \subset U_{dp}$. For this purpose we recall that by defintion $\mathbb{C} \setminus U_{dp} = \{z = w_r + iy : y \in \mathbb{R}, |y| \geq w_i\}$, with w defined in Lemma 1.19. The closest points of this set to the real line interval $[-1, 1]$ are obviously the two points w and $\overline{w}$. For a $t \in [-1, 1]$ the distance $|w - t| = |\overline{w} - t|$ is given by

$$|w - t|^2 = (w - t)(\overline{w} - t) = \frac{\langle d - tp, d - tp\rangle_{\mathbb{R}}}{|p|^2} = \frac{|d - tp|^2}{|p|^2}, \tag{1.59}$$

where we used (1.54). Hence the distance of the real line interval $[-1, 1]$ to the set $\mathbb{C} \setminus U_{dp}$ is given by ζ defined in (1.56). This means that for all $r \in [0, \zeta)$ and all $t \in [-1, 1]$ the closed balls $\overline{B_r(t)} := \{z \in \mathbb{C} : |z - t| \leq r\}$ are contained in U_{dp} and therefore also their union $U_r \subset U_{dp}$.

For the second inclusion $\overline{D_\rho} \subset U_r$ we follow the lines of [4, Proof of Lemma 4.77]. We want to show that for all $z \in \overline{D_\rho}$ there exists a $t \in [-1, 1]$ such that $|z - t| \leq r$. For this purpose we first recall, that for $z = x + iy \in \overline{D_\rho}, x, y \in \mathbb{R}$ there holds

$$\left(\frac{2x}{\rho + 1/\rho}\right)^2 + \left(\frac{2y}{\rho - 1/\rho}\right)^2 \leq 1.$$

By definition of ρ there holds

$$\frac{\rho - 1/\rho}{2} = \frac{\rho^2 - 1}{2\rho} = \frac{1 + r^2 + 2r\sqrt{1 + r^2} + r^2 - 1}{2\rho} = \frac{2r(r + \sqrt{1 + r^2})}{2\rho} = r$$

and therefore

$$\frac{\rho + 1/\rho}{2} = r + \frac{1}{\rho}.$$

Hence for all z as above we get

$$|x| \leq \frac{\rho + 1/\rho}{2} = r + 1/\rho, \qquad |y| \leq \frac{\rho - 1/\rho}{2} = r. \tag{1.60}$$

Next we distinguish three cases: For $x \in [-1, 1]$ we choose $t = x$. Then there holds $|z - t| = |y| \leq r$ by (1.60) and hence $z \in U_r$. For $x > 1$ we choose $t = 1$ and get $|z - t| \leq r$ due to

$$1 \geq \left(\frac{2x}{\rho + 1/\rho}\right)^2 + \left(\frac{2y}{\rho - 1/\rho}\right)^2 = \frac{x^2}{(r + 1/\rho)^2} + \frac{y^2}{r^2} > \frac{(x - 1)^2}{r^2} + \frac{y^2}{r^2} = \frac{|z - t|}{r^2},$$

if we can show that

$$\frac{x^2}{(r + 1/\rho)^2} > \frac{(x - 1)^2}{r^2}.$$

As $x > 1$ this is equivalent to

$$xr > (x - 1)(r + 1/\rho),$$

and

$$r + 1/\rho > x/\rho.$$

This last equation, however, holds true because $\rho > 1$ and $x \leq r + 1/\rho$ due to (1.60). The case $x < -1$ is treated analogously, and hence we have $z \in U_r$ for all $z \in \overline{D_\rho}$, which is the desired result. □

By Lemma 1.20 f_{dp} is holomorphic in a Bernstein elliptic disc $\overline{D_\rho} \subset U_r$ for $r \in [0, \zeta)$. This allows us to apply Theorem 1.9 to estimate $\|f_{dp} - \mathcal{I}^{(m)}_{[-1,1]}[f_{dp}]\|_{\infty,[-1,1]}$. The estimate depends on the maximum of $|f_{dp}|$ in $\overline{D_\rho}$. The following two lemmata allow us to estimate this maximum on the superset U_r of $\overline{D_\rho}$.

Lemma 1.21 (Bound for n_{dp}, [7, cf. Lemma 3.6])**.** *Let $d, p \in \mathbb{R}^3$ with $p \neq 0$ and let n_{dp} be defined as in Lemma 1.19. Furthermore let ζ and U_r for some $r \in [0, \zeta)$ be defined as in Lemma 1.20. Then for all $z \in U_r$ there holds*

$$|n_{dp}(z)| \geq |p|(\zeta - r). \tag{1.61}$$

Proof. By Lemma 1.21 we have

$$n_{dp}(z) = |p|\sqrt{(w - z)(\overline{w} - z)}.$$

Hence it suffices to bound $|w - z|$ and $|\overline{w} - z|$ for all $z \in U_r$. Let $z \in U_r$ be given. By definition of U_r we can find a $t \in [-1, 1]$ satisfying $|z - t| \leq r$. Due to (1.59) and the definition (1.56) of ζ we get

$$\begin{aligned} |w - z| &= |w - t + t - z| \geq |w - t| - |z - t| \geq \zeta - r, \\ |\overline{w} - z| &\geq |\overline{w} - t| - |z - t| = |w - t| - |z - t| \geq \zeta - r, \end{aligned}$$

and finally

$$|n_{dp}(z)| = |p|\sqrt{|w - z||\overline{w} - z|} \geq |p|(\zeta - r).$$

□

Lemma 1.22 (Exponent bound, [7, cf. Lemma 3.8])**.** *Let $d, p \in \mathbb{R}^3$ with $p \neq 0$ and let $c \in \mathbb{R}^3$ with $|c| = 1$. Let ζ and U_r for some $r \in [0, \zeta)$ be defined as in Lemma 1.20. Then for every $z \in U_r$ there exists a $t \in [-1, 1]$ such that*

$$|\exp(i\kappa(n_{dp}(z) - \langle d - zp, c\rangle_{\mathbb{R}}))| \leq \exp\left(\kappa|p|\left(\left|\frac{d - tp}{|d - tp|} - c\right| r + \frac{r^2}{2(\zeta - r)}\right)\right). \tag{1.62}$$

Proof. We observe first that

$$|\exp(i\kappa(n_{dp}(z) - \langle d - zp, c\rangle_{\mathbb{R}}))| \leq \exp(\kappa|\Im(n_{dp}(z) - \langle d - zp, c\rangle_{\mathbb{R}})|),$$

where we used that $\kappa > 0$. Hence it remains to show that

$$|\Im(n_{dp}(z) - \langle d - zp, c\rangle_{\mathbb{R}})| \leq |p| \left(\left| \frac{d-tp}{|d-tp|} - c \right| r + \frac{r^2}{2(\zeta - r)} \right). \tag{1.63}$$

For this purpose fix $z \in U_r$. Due to the definition of U_r in (1.57) there exists a $t \in [-1, 1]$ such that $|t - z| \leq r$. Let $\gamma(s) = t + (z - t)s$ be the straight line with endpoints $\gamma(0) = t$ and $\gamma(1) = z$. We will use a Taylor expansion of $n_{dp} \circ \gamma$ around 0 to estimate $n_{dp}(z) - \langle d - zp, c\rangle_{\mathbb{R}}$. For this purpose we compute the derivatives of n_{dp} defined in (1.51):

$$\begin{aligned} n'_{dp}(z) &= -\frac{\langle p, d - zp\rangle_{\mathbb{R}}}{\langle d - zp, d - zp\rangle_{\mathbb{R}}^{1/2}} = -\frac{\langle p, d - zp\rangle_{\mathbb{R}}}{n_{dp}(z)}, \\ n''_{dp}(z) &= \frac{|p|^2 n_{dp}(z) + n'_{dp}(z)\langle p, d - zp\rangle_{\mathbb{R}}}{n_{dp}(z)^2} = \frac{|p|^2 n_{dp}(z)^2 - \langle p, d - zp\rangle_{\mathbb{R}}^2}{n_{dp}(z)^3}. \end{aligned}$$

Then there holds

$$\begin{aligned} n_{dp}(z) &= n_{dp}(\gamma(1)) = n_{dp}(\gamma(0)) + n'_{dp}(\gamma(0))\gamma'(0) + \int_0^1 n''_{dp}(\gamma(s))(\gamma'(s))^2(1-s)\,\mathrm{d}s \\ &= n_{dp}(t) + n'_{dp}(t)(z-t) + (z-t)^2 \int_0^1 n''_{dp}(\gamma(s))(1-s)\,\mathrm{d}s, \end{aligned}$$

where we used that $\gamma'(s) = (z - t)$ and $\gamma''(s) = 0$. This yields

$$\begin{aligned} &n_{dp}(z) - \langle d - zp, c\rangle_{\mathbb{R}} \\ &\quad = \underbrace{n_{dp}(t) + n'_{dp}(t)(z-t) - \langle d - zp, c\rangle_{\mathbb{R}}}_{=:S_1} + \underbrace{(z-t)^2 \int_0^1 n''_{dp}(\gamma(s))(1-s)\,\mathrm{d}s}_{=:S_2}. \end{aligned} \tag{1.64}$$

Let us first consider the summand S_1. We see for all $t \in [-1, 1]$ that

$$\begin{aligned} S_1 &= n_{dp}(t) + n'_{dp}(t)(z-t) - \langle d - zp, c\rangle_{\mathbb{R}} \\ &= |d - tp| - (z-t)\frac{\langle d - tp, p\rangle_{\mathbb{R}}}{|d-tp|} - \langle d - zp, c\rangle_{\mathbb{R}} \\ &= \left\langle \frac{d-tp}{|d-tp|}, d - tp \right\rangle_{\mathbb{R}} - \left\langle \frac{d-tp}{|d-tp|}, zp - tp \right\rangle_{\mathbb{R}} - \langle d - zp, c\rangle_{\mathbb{R}} \\ &= \left\langle \frac{d-tp}{|d-tp|}, d - zp \right\rangle_{\mathbb{R}} - \langle d - zp, c\rangle_{\mathbb{R}} = \left\langle \frac{d-tp}{|d-tp|} - c, d - zp \right\rangle_{\mathbb{R}}. \end{aligned}$$

By decomposing $z = x + iy$, $x, y \in \mathbb{R}$, we get

$$\Im\left(\left\langle \frac{d-tp}{|d-tp|} - c, d - zp \right\rangle_{\mathbb{R}} \right) = -\left\langle \frac{d-tp}{|d-tp|} - c, yp \right\rangle_{\mathbb{R}},$$

and furthermore the inequality $|y| \leq |z-t| \leq r$. Hence, with the Cauchy–Schwarz inequality we get

$$|\Im(S_1)| = \left| \left\langle \frac{d-tp}{|d-tp|} - c, yp \right\rangle_{\mathbb{R}} \right| \leq |y||p| \left| \frac{d-tp}{|d-tp|} - c \right| \leq r|p| \left| \frac{d-tp}{|d-tp|} - c \right|. \quad (1.65)$$

Next we want to estimate the summand S_2 in (1.64). For this purpose we take a closer look at the integrand. Due to

$$\begin{aligned} &|p|^2\langle d-zp, d-zp\rangle_{\mathbb{R}} - \langle d-zp, p\rangle_{\mathbb{R}}^2 \\ &\quad = |d|^2|p|^2 - 2|p|^2\langle d, zp\rangle_{\mathbb{R}} + z^2|p|^4 - (\langle d, p\rangle_{\mathbb{R}} - z|p|^2)^2 = |d|^2|p|^2 - \langle d, p\rangle_{\mathbb{R}}^2, \end{aligned}$$

which holds for all $z \in \mathbb{C}$, we can bound the numerator of $n''_{dp}(\gamma(s))$ by

$$\begin{aligned} &|\langle d-\gamma(s)p, d-\gamma(s)p\rangle_{\mathbb{R}}|p|^2 - \langle d-\gamma(s)p, p\rangle_{\mathbb{R}}^2| = ||d|^2|p|^2 - \langle d, p\rangle_{\mathbb{R}}^2| = |d|^2|p|^2 - \langle d, p\rangle_{\mathbb{R}}^2 \\ &\quad = |p|^2\langle d-tp, d-tp\rangle_{\mathbb{R}} - \langle d-tp, p\rangle_{\mathbb{R}}^2 \leq |d-tp|^2|p|^2 = |p|^4|w-t|^2, \end{aligned}$$

where we used (1.53) for the last equality. On the other hand by using (1.53) we can estimate $|n_{dp}(\gamma(s))|$ in the denominator of $|n''_{dp}(\gamma(s))|$ for all $s \in [0,1]$ by

$$\begin{aligned} |n_{dp}(\gamma(s))| &= |p|\ |w-\gamma(s)|^{1/2}|\overline{w}-\gamma(s)|^{1/2} \\ &= |p|\ |w-t-(z-t)s|^{1/2}|\overline{w}-t-(z-t)s|^{1/2} \\ &\geq |p|(|w-t|-|z-t|s)^{1/2}(|\overline{w}-t|-|z-t|s))^{1/2} = |p|(|w-t|-|z-t|s). \end{aligned}$$

Hence, we have with $|z-t| \leq r$

$$|S_2| \leq |z-t|^2 \int_0^1 |n''_{dp}(\gamma(s))|(1-s)\,\mathrm{d}s \leq r^2 \int_0^1 \frac{|p||w-t|^2(1-s)}{(|w-t|-|z-t|s)^3}\,\mathrm{d}s. \quad (1.66)$$

By using a partial fraction decomposition one can show for $c_1, c_2 \in \mathbb{R}$ with $c_2 \neq 0$ and $c_1 \neq c_2$ that

$$\int_0^1 \frac{(1-s)}{(c_1-c_2s)^3}\,\mathrm{d}s = \left(\frac{1}{c_2^2(c_1-c_2s)} + \frac{c_2-c_1}{2c_2^2(c_1-c_2s)^2} \right) \Bigg|_0^1 = \frac{1}{2c_1^2(c_1-c_2)}.$$

We can combine this with our estimate in (1.66) to get

$$\begin{aligned} |S_2| &\leq r^2|p||w-t|^2 \int_0^1 \frac{(1-s)}{(|w-t|-|z-t|s)^3}\,\mathrm{d}s \\ &= \frac{r^2|p||w-t|^2}{2|w-t|^2(|w-t|-r)} \leq \frac{|p|r^2}{2(\zeta-r)}, \end{aligned} \quad (1.67)$$

where we used $c_1 = |w-t| \geq \zeta > r \geq |z-t| = c_2$, which holds by definition (1.56) of ζ. By combining (1.64), (1.65) and (1.67) we finally have

$$|\Im(n_{dp}(z) - \langle d-zp, c\rangle_{\mathbb{R}})| \leq |\Im(S_1)| + |S_2| \leq |p| \left(\left| \frac{d-tp}{|d-tp|} - c \right| r + \frac{r^2}{2(\zeta-r)} \right).$$

□

The bounds in Lemmata 1.21 and 1.22 depend on the given vectors d and p and furthermore on the wave number κ. However, we need a bound of f_{dp}, which is independent of d and p, to get a uniform interpolation error by Theorem 1.9, which we need for the application of Proposition 1.18. Furthermore the interpolation error should not depend on κ. By use of the admissibility criteria in Definition 1.13 we can overcome these dependencies.

Proposition 1.23 (Uniform bound of f_{dp} [7, cf. Lemma 3.9 and Theorem 3.13])**.** *Let $c \in \mathbb{R}^3$ with $|c| = 1$ or $c = 0$ and two axis-parallel boxes $X, Y \subset \mathbb{R}^3$ be given by*

$$X = \prod_{i=1}^{3} [a_i, b_i], \qquad Y = \prod_{i=4}^{6} [a_i, b_i].$$

Let $s_{\max}$ be defined by (1.43) *and let $\sigma < 1$ be such that*

$$s_{\max} \leq \sigma \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\}. \tag{1.68}$$

Assume that X, Y and c satisfy the admissibility criteria (A1) – (A3) *for some $\eta_1 > 0$ and $\eta_2 > 0$. Let $d, p \in \mathbb{R}^3$ with $p \neq 0$ satisfy* (1.46a) *and* (1.46b)*. Define*

$$R := \frac{2}{\eta_2}, \tag{1.69}$$

and let U_R be defined as in Lemma 1.20. Then there exists a constant $C(\sigma, \eta_1, \eta_2)$ depending only on σ, η_1 and η_2 such that

$$|f_{dp}(z)| \leq \frac{C(\sigma, \eta_1, \eta_2)}{\operatorname{dist}(X, Y)}, \qquad \text{for all } z \in U_R. \tag{1.70}$$

Proof. We start the proof by showing

$$\left| \frac{d - tp}{|d - tp|} - c \right| \leq \frac{\eta_1 + \eta_2}{2\kappa |p|}, \qquad \text{for all } t \in [-1, 1], \tag{1.71}$$

which is [7, Lemma 3.9]. For this purpose, let $q := \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\}$, let $t \in [-1, 1]$, and let $x \in X$ and $y \in Y$ be such that $(d - tp) = (x - y)$. Such points x and y exist due to (1.46b). Therefore we get $|d - tp| = |x - y| \geq \operatorname{dist}(X, Y)$. Obviously also the midpoints m_X and m_Y of X and Y, respectively, satisfy $|m_X - m_Y| \geq \operatorname{dist}(X, Y)$. Hence, by (A3) we get

$$|x - y| \geq \operatorname{dist}(X, Y) \geq \frac{\kappa}{\eta_2} q^2, \qquad |m_X - m_Y| \geq \frac{\kappa}{\eta_2} q^2,$$

i.e. $(x - y)$ and $(m_X - m_Y)$ lie outside of the ball with radius $q^2 \kappa / \eta_2$ centered at zero. By means of Lemma 2.12 we see that if we project these two points onto the surface of this ball we do not increase their distance, which means that

$$\frac{\kappa q^2}{\eta_2} \left| \frac{x - y}{|x - y|} - \frac{m_X - m_Y}{|m_X - m_Y|} \right| \leq |(x - y) - (m_X - m_Y)|.$$

Furthermore there holds

$$|(x-y)-(m_X-m_Y)| \le |x-m_X|+|y-m_Y| \le \frac{\operatorname{diam}(X)}{2}+\frac{\operatorname{diam}(Y)}{2} \le q.$$

By using these two estimates together with (A2) and the estimate $2|p| \le s_{\max} < q$, which is a consequence of (1.46a), we have

$$\begin{aligned}\left|\frac{d-tp}{|d-tp|}-c\right| = \left|\frac{x-y}{|x-y|}-c\right| &\le \left|\frac{x-y}{|x-y|}-\frac{m_X-m_Y}{|m_X-m_Y|}\right| + \left|\frac{m_X-m_Y}{|m_X-m_Y|}-c\right| \\ &\le |(x-y)-(m_X-m_Y)|\frac{\eta_2}{\kappa q^2}+\frac{\eta_1}{\kappa q} \le \frac{\eta_1+\eta_2}{\kappa q} = \frac{\eta_1+\eta_2}{2\kappa|p|}.\end{aligned}$$

With this result and Lemmata 1.21 and 1.22 we are able to prove (1.70). We start by estimating the denominator n_{dp} of f_{dp}. For this purpose, we recall the definition of ζ in Lemma 1.20. We observe that

$$\zeta = \min\left\{\frac{|d-tp|}{|p|} : t \in [-1,1]\right\} \ge \frac{\operatorname{dist}(X,Y)}{|p|}.$$

Due to (1.46a), (1.68) and the admissibility criterion (A1) we can further see that

$$\zeta \ge \frac{2\operatorname{dist}(X,Y)}{s_{\max}} \ge \frac{2\operatorname{dist}(X,Y)}{\sigma \max\{\operatorname{diam}(X),\operatorname{diam}(Y)\}} \ge \frac{2}{\sigma\eta_2} = \frac{R}{\sigma}. \tag{1.72}$$

Combining these estimates with (1.61) yields

$$|n_{dp}(z)| \ge |p|(\zeta - R) \ge \operatorname{dist}(X,Y)\,(1-R/\zeta) > \operatorname{dist}(X,Y)\,(1-\sigma), \tag{1.73}$$

for all $z \in U_R$.

For the estimate of the exponential term of f_{dp} we note that due to (1.46a) and the admissibility criterion (A3) there holds

$$\zeta \ge \frac{2\operatorname{dist}(X,Y)}{s_{\max}} \ge \frac{2\kappa \max\{\operatorname{diam}(X),\operatorname{diam}(Y)\}^2}{\eta_2 s_{\max}} \ge \frac{2\kappa}{\eta_2} s_{\max} > \frac{4\kappa|p|}{\eta_2}.$$

By applying Lemma 1.22 and using the definition of R in (1.69) as well as (1.71) and (1.72), we get for all $z \in U_R$ the estimate

$$\begin{aligned}|\exp(i\kappa(n_{dp}(z)-\langle d-zp,c\rangle_{\mathbb{R}}))| &\le \exp\left(\kappa|p|\left(\left|\frac{d-tp}{|d-tp|}-c\right|R+\frac{R^2}{2(\zeta-R)}\right)\right) \\ &\le \exp\left(\kappa|p|\left(\frac{\eta_1+\eta_2}{2\kappa|p|}R+\frac{R^2}{2\zeta(1-R/\zeta)}\right)\right) \\ &\le \exp\left(\kappa|p|\left(\frac{\eta_1+\eta_2}{2\kappa|p|}R+\frac{\eta_2 R^2}{8\kappa|p|(1-R/\zeta)}\right)\right) \\ &\le \exp\left(\frac{\eta_1+\eta_2}{\eta_2}+\frac{1}{2\eta_2(1-\sigma)}\right).\end{aligned}$$

We combine this last estimate and (1.73) for all $z \in U_R$ to end up with

$$|f_{dp}(z)| = \frac{|\exp(i\kappa(n_{dp}(z) - \langle d - zp, c\rangle_{\mathbb{R}}))|}{4\pi|n_{dp}(z)|} \leq \frac{\exp\left((\eta_1+\eta_2)/\eta_2 + 1/(2\eta_2(1-\sigma))\right)}{4\pi(1-\sigma)\operatorname{dist}(X,Y)} = \frac{C(\sigma,\eta_1,\eta_2)}{\operatorname{dist}(X,Y)}.$$

□

Remark 1.24. In [7, Proof of Theorem 3.13] the authors use that

$$|f_{dp}(z)| \leq \frac{\exp(\eta_1+\eta_2)}{\pi \operatorname{dist}(X,Y)},$$

for all points z in $U_{\hat{r}}$, where

$$\hat{r} := \min\left\{1, \frac{3}{2\eta_2}\right\},$$

to estimate the interpolation error of f_{dp} by means of Theorem 1.9. In comparison the bound which we have shown in the last proof has the advantage that it is strictly decreasing for increasing η_2. Hence, it seems acceptable to choose larger values of η_2 in the admissibility criteria (A1) and (A3). In addition our bound holds on the larger set U_R, which leads to a better convergence rate of the interpolation error. The only drawback we have is the additional parameter σ which describes the shape of the boxes X and Y. In fact, the term $C(\sigma,\eta_1,\eta_2)$ becomes unbounded for σ converging to 1. However, for rather isotropic boxes X and Y this is not the case.

The estimate of the interpolation error of f_{dp} is a direct consequence of the previous results. We formulate it in the following proposition.

Proposition 1.25 (Approximation of f_{dp} [7, cf. Theorem 3.13])**.** *Let c, X, Y, d, p and R be given as in Proposition 1.23. In particular assume that X, Y and c satisfy the admissibility criteria* (A1) – (A3) *for some $\eta_1 > 0$ and $\eta_2 > 0$. Furthermore let $\sigma < 1$ be such that* (1.68) *holds. Define*

$$\rho := \sqrt{R^2+1} + R. \tag{1.74}$$

Let Λ_m be the Lebesgue constant defined by (1.6) *and let $C(\sigma,\eta_1,\eta_2)$ be the constant from Proposition 1.23. Then there holds*

$$\|f_{dp} - \mathcal{I}^{(m)}_{[-1,1]}[f_{dp}]\|_{\infty,[-1,1]} \leq \frac{2C(\sigma,\eta_1,\eta_2)}{(\rho-1)\operatorname{dist}(X,Y)}(\Lambda_m+1)\rho^{-(m+1)}. \tag{1.75}$$

Proof. It suffices to collect the above results. The function f_{dp} has a holomorphic extension (1.55) on the set U_{dp} defined in Lemma 1.19. Due to Proposition 1.23 we know that this extension, which we also denote with f_{dp}, is bounded by $C(\eta_1,\eta_2)/\operatorname{dist}(X,Y)$ on U_R, which is a superset of the Bernstein elliptic disc D_ρ and a subset of U_{dp} due to Lemma 1.20. Hence we can apply Theorem 1.9, which yields the assertion. □

Finally we come to the main result of this section, the error estimate of the directional single-level approximation.

Theorem 1.26 (Directional single-level approximation error [7, cf. Corollary 3.14])**.** *Let $c \in \mathbb{R}^3$ with $|c| = 1$ or $c = 0$ and let X, $Y \subset \mathbb{R}^3$ be two axis-parallel boxes. Assume that X, Y and c satisfy the admissibility criteria* (A1) – (A3) *for some $\eta_1 > 0$ and $\eta_2 > 0$. Let $s_{\max}$ be defined as in* (1.43) *and assume that* (1.68) *holds for $\sigma < 1$. Let furthermore*

$$\hat{\rho} := 1 + \frac{2}{\eta_2}. \tag{1.76}$$

Then there exists a function $C_{\mathrm{sl}}(m, \sigma, \eta_1, \eta_2)$, which decays exponentially in m such that

$$\|f - \tilde{f}^{(m)}_{X,Y}\|_{\infty, X\times Y} \leq \frac{C_{\mathrm{sl}}(m, \sigma, \eta_1, \eta_2)}{\operatorname{dist}(X, Y)} \hat{\rho}^{-(m+1)}. \tag{1.77}$$

Proof. By Propositions 1.18 and 1.25 we have

$$\|f - \tilde{f}^{(m)}_{X,Y}\|_{\infty, X\times Y} \leq 6\Lambda_m^5(\Lambda_m + 1)\frac{2C(\sigma, \eta_1, \eta_2)}{(\rho - 1)\operatorname{dist}(X, Y)}\rho^{-(m+1)}, \tag{1.78}$$

with $\rho = \sqrt{R^2 + 1} + R$ and $R = 2/\eta_2$. We set $\rho = \alpha\hat{\rho}$, where

$$\alpha := \frac{\rho}{\hat{\rho}} = \frac{\sqrt{1 + 4/\eta_2^2} + 2/\eta_2}{1 + 2/\eta_2} > 1.$$

With this definition we can write (1.78) in the form

$$\|f - \tilde{f}^{(m)}_{X,Y}\|_{\infty, X\times Y} \leq \frac{C_{\mathrm{sl}}(m, \sigma, \eta_1, \eta_2)}{\operatorname{dist}(X, Y)} \hat{\rho}^{-(m+1)},$$

by defining

$$C_{\mathrm{sl}}(m, \sigma, \eta_1, \eta_2) := 12\Lambda_m^5(\Lambda_m + 1)\frac{C(\sigma, \eta_1, \eta_2)}{(\rho - 1)}\alpha^{-(m+1)}.$$

Due to the estimate (1.8) the Lebesgue constant for the Chebyshev nodes increases only logarithmically. Consequently $C_{\mathrm{sl}}(m, \sigma, \eta_1, \eta_2)$ converges to 0 exponentially if m goes to infinity due to the exponential decay of $\alpha^{-(m+1)}$. □

Remark 1.27. We note that due to the exponential decay of $C_{\mathrm{sl}}(m, \sigma, \eta_1, \eta_2)$ there exists a constant C depending only on σ, η_1, and η_2 such that

$$\|f - \tilde{f}^{(m)}_{X,Y}\|_{\infty, X\times Y} \leq \frac{C}{\operatorname{dist}(X, Y)} \hat{\rho}^{-(m+1)}. \tag{1.79}$$

However, we need the decaying behavior of C_{sl} in the proof of Theorem 1.15.

1.3.2 Estimate of the directional reinterpolation

Another key ingredient for the directional multi-level error analysis is the estimate of the nested operator norms

$$\|\mathcal{I}^{(m)}_{X_L\times Y_L,c_L}\circ ...\circ \mathcal{I}^{(m)}_{X_{\ell+1}\times Y_{\ell+1},c_{\ell+1}}\|_{\infty,X_L\times Y_L\leftarrow X_\ell\times Y_\ell},$$

where $\ell\in\{0,...,L-1\}$, $\{X_k\}_{k=0}^L$ and $\{Y_k\}_{k=0}^L$ are two sequences of nested axis-parallel boxes in $\mathbb{R}^3$ and the operators $\mathcal{I}^{(m)}_{X_k\times Y_k,c_k}$ are given in (1.17) for all $k\in\{0,...,L\}$. As before we follow the lines of [7]. We will first show, that this operator norm can be written as the product of norms of nested operators in one dimension. Then we will show an appropriate estimate for these nested operators by focusing first on a single-level and then transfering the obtained result to the multi-level case.

We start by recalling that

$$\mathcal{I}^{(m)}_{X_L\times Y_L,c_L}\circ ...\circ \mathcal{I}^{(m)}_{X_{\ell+1}\times Y_{\ell+1},c_{\ell+1}}=\left(\mathcal{I}^{(m)}_{X_L,c_L}\circ ...\circ \mathcal{I}^{(m)}_{X_{\ell+1},c_{\ell+1}}\right)\otimes\left(\mathcal{I}^{(m)}_{Y_L,-c_L}\circ ...\circ \mathcal{I}^{(m)}_{Y_{\ell+1},-c_{\ell+1}}\right).$$

From this definition we directly see that

$$\begin{aligned}&\|\mathcal{I}^{(m)}_{X_L\times Y_L,c_L}\circ ...\circ \mathcal{I}^{(m)}_{X_{\ell+1}\times Y_{\ell+1},c_{\ell+1}}\|_{\infty,X_L\times Y_L\leftarrow X_\ell\times Y_\ell}\\ &\quad\le \|\mathcal{I}^{(m)}_{X_L,c_L}\circ ...\circ \mathcal{I}^{(m)}_{X_{\ell+1},c_{\ell+1}}\|_{\infty,X_L\leftarrow X_\ell}\|\mathcal{I}^{(m)}_{Y_L,-c_L}\circ ...\circ \mathcal{I}^{(m)}_{Y_{\ell+1},-c_{\ell+1}}\|_{\infty,Y_L\leftarrow Y_\ell}.\end{aligned}\tag{1.80}$$

Similarly the following assertion follows.

Lemma 1.28. *Let $\{X_k\}_{k=0}^L$ be a sequence of axis-parallel boxes with $X_0\supset ...\supset X_L$. In particular let*

$$X_k=[a_{k,1},b_{k,1}]\times ...\times [a_{k,3},b_{k,3}]=:J_{k,1}\times ...\times J_{k,3}.\tag{1.81}$$

Let $\{c_k\}_{k=0}^L\subset\mathbb{R}^3$ be a sequence of directions satisfying $|c_k|=1$ or $c_k=0$ for all $k\in\{0,...,L\}$. Define the one-dimensional directional interpolation operators

$$\mathcal{I}^{(m)}_{k,j}:C(J_{k,j})\to C(J_{k,j}),\qquad u\mapsto \exp(i\kappa c_{k,j}\,\cdot)\mathcal{I}^{(m)}_{J_{k,j}}[\exp(-i\kappa c_{k,j}\,\cdot)u],\tag{1.82}$$

for all $k\in\{0,...,L\}$ and $j\in\{1,...,3\}$. Then there holds

$$\|\mathcal{I}^{(m)}_{X_L,c_L}\circ ...\circ \mathcal{I}^{(m)}_{X_{\ell+1},c_{\ell+1}}\|_{\infty,X_L\leftarrow X_\ell}\le\prod_{j=1}^3\|\mathcal{I}^{(m)}_{L,j}\circ ...\circ \mathcal{I}^{(m)}_{\ell+1,j}\|_{\infty,J_{L,j}\leftarrow J_{\ell,j}}.\tag{1.83}$$

Proof. By construction we have

$$\mathcal{I}^{(m)}_{X_k,c_k}=\mathcal{I}^{(m)}_{k,1}\otimes ...\otimes \mathcal{I}^{(m)}_{k,3},$$

for all $k\in\{0,...,L\}$. This property transfers to the composed operators

$$\mathcal{I}^{(m)}_{X_L,c_L}\circ ...\circ \mathcal{I}^{(m)}_{X_{\ell+1},c_{\ell+1}}=\left(\mathcal{I}^{(m)}_{L,1}\circ ...\circ \mathcal{I}^{(m)}_{\ell+1,1}\right)\otimes ...\otimes\left(\mathcal{I}^{(m)}_{L,3}\circ ...\circ \mathcal{I}^{(m)}_{\ell+1,3}\right).$$

As in the proof of Lemma 1.17 the assertion follows by using (1.41). □

The previous lemma motivates the necessity to estimate

$$\|\mathcal{I}_{L,j}^{(m)} \circ ... \circ \mathcal{I}_{\ell+1,j}^{(m)}\|_{J_{L,j} \leftarrow J_{\ell,j}},$$

for all $j \in \{1, ..., 3\}$. We start by considering a single level ℓ and estimating the directional interpolation error

$$\| \exp(i\kappa c_{l-1} \cdot)\pi - \mathcal{I}_{\ell}^{(m)}[\exp(i\kappa c_{\ell-1} \cdot)\pi]\|_{\infty, J_\ell}, \tag{1.84}$$

for all polynomials $\pi \in \Pi_m$, where one can think of $\mathcal{I}_{\ell}^{(m)}$ as one of the operators $\mathcal{I}_{\ell,j}^{(m)}$. For this estimate we introduce two technical lemmata.

Lemma 1.29 ([7, Lemma 5.3])**.** *Let $[a, b] \subset [-1, 1]$ be a non-empty interval and $h := (b - a)/2$. Define for $\alpha > 1$ the transformed Bernstein elliptic disc*

$$D_{\alpha}^{a,b} := \left\{ z = x + iy : x, y \in \mathbb{R},\ \left(\frac{2(x - (a + b)/2)}{h(\alpha + \alpha^{-1})} \right)^2 + \left(\frac{2y}{h(\alpha - \alpha^{-1})} \right)^2 < 1 \right\}. \tag{1.85}$$

Let $\varepsilon \in (0, 1)$ be fixed. Then there exists a $\rho_0 > 1$ depending only on ε such that

$$D_{(1-\varepsilon)\rho/h}^{a,b} \subset D_{\rho}, \qquad \text{for all } \rho > \rho_0. \tag{1.86}$$

Proof. Let us denote by $B_r(\xi)$ the open ball around ξ with radius $r > 0$, i.e.

$$B_r(\xi) = \{ z \in \mathbb{C} : |z - \xi| < r \}.$$

By definition of D_ρ in (1.9) there holds

$$B_{(\rho - 1/\rho)/2}(0) \subset D_{\rho},$$

and similarly by (1.85)

$$D_{\alpha}^{a,b} \subset B_{h(\alpha + 1/\alpha)/2}((a + b)/2).$$

Here we consider $\alpha = (1 - \varepsilon)\rho/h$. The point $(a + b)/2$ satisfies $|(a + b)/2| \leq 1$. Hence, for all $z \in B_{h(\alpha+1/\alpha)/2}((a + b)/2)$ we have

$$|z| \leq \left| z - \frac{a + b}{2} \right| + \left| \frac{a + b}{2} \right| \leq \frac{h(\alpha + 1/\alpha)}{2} + 1,$$

which means that $z \in B_{h(\alpha+1/\alpha)/2+1}(0)$. Therefore, it suffices to show that

$$B_{h(\alpha+1/\alpha)/2+1}(0) \subset B_{(\rho-1/\rho)/2}(0),$$

to proof the assertion of the lemma. This inclusion obviously holds true if and only if

$$1 + \frac{h(\alpha + 1/\alpha)}{2} \leq \frac{\rho - 1/\rho}{2}. \tag{1.87}$$

By inserting the definition of α the left hand side can be written as

$$1 + \frac{h(\alpha + 1/\alpha)}{2} = 1 + \frac{h}{2}\left(\frac{(1-\varepsilon)\rho}{h} + \frac{h}{(1-\varepsilon)\rho}\right) = 1 + \frac{(1-\varepsilon)^2\rho^2 + h^2}{2(1-\varepsilon)\rho}.$$

Therefore (1.87) is equivalent to

$$g_\varepsilon(\rho) := \rho^2 - 1 - 2\rho - (1-\varepsilon)\rho^2 - \frac{h^2}{1-\varepsilon} = \varepsilon\rho^2 - 2\rho - 1 - \frac{h^2}{1-\varepsilon} \geq 0.$$

This condition obviously holds if $\rho > \tilde{\rho}_0$, where $\tilde{\rho}_0$ is the larger of the two roots of g_ε, i.e.

$$\tilde{\rho}_0 = \frac{1 + \sqrt{1 + \varepsilon(1 + h^2/(1-\varepsilon))}}{\varepsilon}.$$

The root $\tilde{\rho}_0$ can be bounded by

$$\tilde{\rho}_0 < \rho_0 := \frac{1 + \sqrt{1 + \varepsilon(1 + 1/(1-\varepsilon))}}{\varepsilon},$$

because $h \in (0, 1]$. This ρ_0 depends only on ε, and by construction for all $\rho > \rho_0$ the desired inclusion holds. □

Lemma 1.30 ([14, Chapter 4, Theorem 2.2])**.** *Let $p_m \in \Pi_m$ be given by*

$$p_m(z) = \sum_{j=0}^{m} a_j z^j, \qquad a_j \in \mathbb{C}.$$

Assume that

$$|p_m(x)| \leq M, \qquad \text{for all } x \in [-1, 1]. \tag{1.88}$$

Let D_ρ for $\rho > 1$ denote the Bernstein elliptic disc defined in (1.9)*. Then there holds*

$$|p_m(z)| \leq M\rho^m, \qquad \text{for all } z \in D_\rho \text{ with } \rho > 1. \tag{1.89}$$

Proof. The proof is based on the properties of the mapping

$$\Psi : \mathbb{C} \setminus \{0\} \to \mathbb{C}, \qquad z \mapsto \Psi(z) = \frac{1}{2}\left(z + \frac{1}{z}\right),$$

and the maximum modulus principle for holomorphic functions [17, cf. Theorem 5.4].

As mentioned in the proof of Proposition 1.8 the function Ψ maps circles $\partial B_{1/\rho}(0)$ for $\rho > 1$ to the boundary ∂D_ρ of the Bernstein elliptic discs D_ρ, and maps $\partial B_1(0)$ to the interval $[-1, 1]$. Let us consider the function

$$F(w) := w^m (p_m \circ \Psi)(w) = w^m \sum_{j=0}^{m} a_j \left(\frac{w + 1/w}{2}\right)^j = \sum_{j=0}^{m} a_j w^{m-j} \left(\frac{w^2 + 1}{2}\right)^j,$$

which is a polynomial in w. In particular it is holomorphic on $B_1(0)$ and continuous on $\overline{B_1(0)}$. For w with $|w| = 1$ there holds $(w + 1/w)/2) \in [-1, 1]$ and hence by (1.88)

$$|F(w)| = |p_m\left(\frac{w + 1/w}{2}\right) w^m| \leq M. \tag{1.90}$$

By the maximum modulus principle we get that $|F|$ attains its maximum for some w with $|w| = 1$, and hence (1.90) holds for all $w \in \overline{B_1(0)}$.

We use this estimate to show the assertion. For an arbitrary $z \in D_\rho$, $z \notin [-1, 1]$ there exists a $\mu < \rho$ such that $z \in \partial D_\mu$. Therefore we can find a w with $|w| = 1/\mu$ such that $\Psi(w) = z$. In particular we get

$$|p_m(z)| = |(p_m \circ \Psi)(w)w^m w^{-m}| = |F(w)|\ |w|^{-m} \leq M\mu^m \leq M\rho^m.$$

□

With these two lemmata we can show the following result, which plays a central role in the estimate of (1.84).

Lemma 1.31 ([7, cf. Lemma 5.4])**.** *Let $J_1 \subset J_0$ be two closed intervals. Let $\bar{q} \in \mathbb{R}$ be such that*

$$|J_1|/|J_0| \leq \bar{q} < 1, \tag{1.91}$$

and denote $h_0 := |J_0|/2$, $h_1 := |J_1|/2$. Let $\gamma \in \mathbb{R}$ and c_0, $c_1 \in \mathbb{R}$ satisfy

$$|\kappa h_0(c_0 - c_1)| \leq \gamma. \tag{1.92}$$

Define $\hat{q} := (1 + \bar{q})/2$. Then there exists an $m_0 > 0$ depending only on $\bar{q}$ and γ, such that for all $m \in \mathbb{N}$ with $m \geq m_0$ and all polynomials $\pi \in \Pi_m$ there holds

$$\inf_{v \in \Pi_m} \|\exp(i\kappa c_0 \cdot)\pi - \exp(i\kappa c_1 \cdot)v\|_{\infty, J_1} \leq \hat{q}^m \|\pi\|_{\infty, J_0}. \tag{1.93}$$

Proof. Let Φ be the affine linear mapping from $[-1, 1]$ to J_0. Define the interval $\hat{J}_1 := [a, b] := \Phi^{(-1)}(J_1)$ and $h := (b - a)/2$. Then by construction there holds $h = h_1/h_0 \leq \bar{q}$. With $\hat{\pi} := \pi \circ \Phi$ and due to $|\exp(-i\kappa c_1 \cdot)| = 1$ we get

$$\inf_{v \in \Pi_m} \|\exp(i\kappa c_0 \cdot)\pi - \exp(i\kappa c_1 \cdot)v\|_{\infty, J_1} = \inf_{\hat{v} \in \Pi_m} \|\exp(i\kappa h_0(c_0 - c_1)\cdot)\hat{\pi} - \hat{v}\|_{\infty, \hat{J}_1}. \tag{1.94}$$

The function $\exp(i\kappa h_0(c_0 - c_1)\cdot)\hat{\pi}$ can be extended holomorphically to the whole complex plain. Hence, the right hand side in (1.94) can be estimated using an analogous result as in Proposition 1.8 for $D_\alpha^{a,b}$ instead of D_ρ and arbitrary $\alpha > 1$. This yields

$$\begin{aligned} \inf_{\hat{v} \in \Pi_m} \|\exp(i\kappa h_0(c_0 - c_1)\cdot)\hat{\pi} - \hat{v}\|_{\infty, \hat{J}_1} &\leq \frac{2\alpha^{-m}}{\alpha - 1} \|\exp((i\kappa h_0(c_0 - c_1)\cdot)\hat{\pi}\|_{\infty, D_\alpha^{a,b}} \\ &\leq \frac{2\alpha^{-m}}{\alpha - 1} \exp\left(|\kappa h_0(c_0 - c_1)| \left| h\frac{\alpha - 1/\alpha}{2} \right|\right) \|\hat{\pi}\|_{\infty, D_\alpha^{a,b}}. \end{aligned} \tag{1.95}$$

For the estimate of the exponential term in the last inequality we used that for all $z \in D_\alpha^{a,b}$ there holds $|\Im(z)| \le h(\alpha - 1/\alpha)/2$. For $\alpha := (1-\varepsilon)\rho/\bar{q} \le (1-\varepsilon)\rho/h$ we get by Lemma 1.29 that $D_\alpha^{a,b} \subset D_\rho$ for all $\rho > \rho_0$, and with Lemma 1.30 it follows that

$$\|\hat{\pi}\|_{\infty, D_\alpha^{a,b}} \le \|\hat{\pi}\|_{\infty, D_\rho} \le \rho^m \|\hat{\pi}\|_{\infty,[-1,1]}. \tag{1.96}$$

The parameter ε in the definition of α is chosen such that $\varepsilon \in (0, 1 - \bar{q}/\hat{q})$. For the subsequent estimates we further introduce a parameter $\beta > 0$ such that

$$\hat{q} = \frac{\bar{q}}{1-\varepsilon} \exp\left(\frac{\gamma(1-\varepsilon)\beta}{2}\right).$$

Such a β exists due to the choice of ε. Finally we set $\rho = \beta m$ with $m \ge \lceil \rho_0/\beta \rceil$ for ρ in the definition of α and proceed with the estimate in (1.95).

Due to (1.92) and (1.96) we have

$$\begin{aligned}\frac{2\alpha^{-m}}{\alpha - 1} \exp\left(|\kappa h_0 (c_0 - c_1)| \frac{h(\alpha - 1/\alpha)}{2}\right) \|\hat{\pi}\|_{\infty, D_\alpha^{a,b}} \\ \le \frac{2\alpha^{-m}}{\alpha - 1} \exp\left(\gamma \frac{h(\alpha - 1/\alpha)}{2}\right) \rho^m \|\hat{\pi}\|_{\infty,[-1,1]}.\end{aligned}$$

By using $h(\alpha - 1/\alpha)/2 = h\alpha(1 - 1/\alpha^2)/2 \le \alpha\bar{q}/2$, which holds true as $h \le \bar{q}$ if $\alpha > 1$, and the definition of α, ρ and β we have

$$\begin{aligned}\left(\frac{\rho}{\alpha}\right)^m \exp\left(\gamma h \frac{\alpha - 1/\alpha}{2}\right) &\le \left(\frac{\bar{q}}{1-\varepsilon}\right)^m \exp\left(\frac{\gamma \bar{q} \alpha}{2}\right) = \left(\frac{\bar{q}}{1-\varepsilon}\right)^m \exp\left(\frac{\gamma(1-\varepsilon)\beta m}{2}\right) \\ &= \left(\frac{\bar{q}}{1-\varepsilon} \exp\left(\frac{\gamma(1-\varepsilon)\beta}{2}\right)\right)^m = \hat{q}^m.\end{aligned}$$

Hence, we can estimate (1.95) further by

$$\inf_{\hat{v} \in \Pi_m} \| \exp(i\kappa h_0 (c_0 - c_1)\cdot)\hat{\pi} - \hat{v}\|_{\infty, \hat{J}_1} \le \frac{2}{\alpha - 1} \hat{q}^m \|\hat{\pi}\|_{\infty,[-1,1]} \le \hat{q}^m \|\pi\|_{\infty, J_0}, \tag{1.97}$$

where we used the identity $\|\hat{\pi}\|_{\infty,[-1,1]} = \|\pi\|_{\infty, J_0}$ and the condition $\alpha \ge 3$ for the last inequality. This condition and all the above steps hold true if

$$m \ge m_0 := \max\left\{ \left\lceil \frac{\rho_0}{\beta} \right\rceil, \left\lceil \frac{3\bar{q}}{(1-\varepsilon)\beta} \right\rceil \right\}.$$

Therefore, combining (1.94) and (1.97) yields the assertion for such m. □

Lemma 1.31 allows us to estimate (1.84), which we do in the following lemma.

Lemma 1.32 ([7, cf. Lemma 5.5])**.** *Let J_0, J_1, $\bar{q}$ and $\hat{q}$ be as in Lemma 1.31. In particular let $|J_1|/|J_0| \leq \bar{q} < 1$. Let furthermore c_0, c_1 and $\gamma \in \mathbb{R}$ be such that* (1.92) *holds. Define*

$$\mathcal{I}_1^{(m)} : C(J_1) \to C(J_1), \qquad u \mapsto \exp(i\kappa c_1 \cdot)\mathcal{I}_{J_1}^{(m)}[\exp(-i\kappa c_1 \cdot)u], \tag{1.98}$$

and let Λ_m be the Lebesgue constant defined in (1.6)*. Then for m_0 as in Lemma 1.31, for all $m \in \mathbb{N}$ with $m \geq m_0$, and for all polynomials $\pi \in \Pi_m$ there holds*

$$\| \exp(i\kappa c_0 \cdot)\pi - \mathcal{I}_1^{(m)}[\exp(i\kappa c_0 \cdot)\pi]\|_{\infty,J_1} \leq (1 + \Lambda_m)\hat{q}^m \|\pi\|_{\infty,J_0} \tag{1.99}$$

Proof. We see, that for arbitrary $v \in \Pi_m$ there holds

$$\begin{aligned}\exp(i\kappa c_0 \cdot)\pi - \mathcal{I}_1^{(m)}[\exp(i\kappa c_0 \cdot)\pi] &= \exp(i\kappa c_0 \cdot)\pi - \exp(i\kappa c_1 \cdot)v \\ &\quad - \exp(i\kappa c_1 \cdot)\mathcal{I}_{J_1}^{(m)}[\exp(-i\kappa c_1 \cdot)(\exp(i\kappa c_0 \cdot)\pi - \exp(i\kappa c_1 \cdot)v)],\end{aligned}$$

where we used, that $\mathcal{I}_{J_1}^{(m)} v = v$. As a consequence

$$\| \exp(i\kappa c_0 \cdot)\pi - \mathcal{I}_1^{(m)}[\exp(i\kappa c_0 \cdot)\pi]\|_{\infty,J_1} \leq (1 + \Lambda_m)\| \exp(i\kappa c_0 \cdot)\pi - \exp(i\kappa c_1 \cdot)v\|_{\infty,J_1}.$$

As this holds for arbitrary v we can take the infimum on the right hand side and by Lemma 1.31 we get

$$\begin{aligned}&\| \exp(i\kappa c_0 \cdot)\pi - \mathcal{I}_1^{(m)}[\exp(i\kappa c_0 \cdot)\pi]\|_{\infty,J_1} \\ &\quad \leq (1 + \Lambda_m) \inf_{v \in \Pi_m} \| \exp(i\kappa c_0 \cdot)\pi - \exp(i\kappa c_1 \cdot)v\|_{\infty,J_1} \leq (1 + \Lambda_m)\hat{q}^m \|\pi\|_{\infty,J_0}.\end{aligned}$$

□

Finally we are ready to estimate the norms of the nested operators $\mathcal{I}_L^{(m)} \circ ... \circ \mathcal{I}_{l+1}^{(m)}$.

Theorem 1.33 ([7, cf. Theorem 5.6])**.** *Let $\{J_k\}_{k=0}^L$ be a sequence of non-empty intervals, such that $J_0 \supset J_1 \supset ... \supset J_L$. Let $\bar{q} \in \mathbb{R}$ be such that*

$$|J_k|/|J_{k-1}| \leq \bar{q} < 1, \qquad \text{for all } k \in \{1, ..., L\}, \tag{1.100}$$

and let $\hat{q} = (1 + \bar{q})/2$. Define $h_k := |J_k|/2$ for all $k \in \{0, ..., L\}$. Let c_0, ..., $c_L \in \mathbb{R}$ and $\gamma \in \mathbb{R}$ satisfy

$$|\kappa h_{k-1}(c_{k-1} - c_k)| \leq \gamma \qquad \text{for all } k \in \{1, ..., L\}. \tag{1.101}$$

Furthermore, let the operators $\mathcal{I}_k^{(m)}$, $k \in \{0, ..., L\}$ be defined similarly as in (1.98) *for the intervals J_k instead of J_1 and c_k instead of c_1. Then there exists an $m_0 \in \mathbb{N}$ depending on γ and $\bar{q}$ only, such that*

$$\|\mathcal{I}_L^{(m)} \circ ... \circ \mathcal{I}_\ell^{(m)}\|_{\infty,J_L \leftarrow J_{\ell-1}} \leq \Lambda_m(1 + \varepsilon_{m,L-\ell}), \tag{1.102}$$

$$\varepsilon_{m,L-\ell} := (1 + (1 + \Lambda_m)\hat{q}^m)^{L-\ell} - 1, \tag{1.103}$$

for all $m \geq m_0$ and $\ell \in \{1, ..., L\}$. In particular there exists a constant $C(\bar{q}, L)$ depending solely on $\bar{q}$ and L such that

$$\varepsilon_{m,k} + 1 \leq C(\bar{q}, L) \quad \text{for all } m \geq m_0,\ k \in \{1, ..., L-1\}. \tag{1.104}$$

Proof. We show (1.102) for $\ell = 1$ and general L. For general ℓ the assertion then follows by a simple index shift.

Define the operators $E_k := I - \mathcal{I}_k^{(m)} \circ ... \circ \mathcal{I}_2^{(m)}$ for $k > 2$ and $E_2 := (I - \mathcal{I}_2^{(m)})$. We start by showing that for arbitrary $\pi \in \Pi_m$ and all $k \in \{2, ..., L\}$ there holds

$$\|E_k[\exp(i\kappa c_1 \cdot)\pi]\|_{\infty,J_k} \leq \varepsilon_{m,k-1}\|\pi\|_{\infty,J_1}, \tag{1.105a}$$

$$\|\mathcal{I}_k^{(m)} \circ ... \circ \mathcal{I}_2^{(m)}[\exp(i\kappa c_1 \cdot)\pi]\|_{\infty,J_k} \leq (1 + \varepsilon_{m,k-1})\|\pi\|_{\infty,J_1}. \tag{1.105b}$$

This is done by induction over k. Due to Lemma 1.32 we have

$$\|(I - \mathcal{I}_k^{(m)})[\exp(i\kappa c_{k-1} \cdot)\pi]\|_{\infty,J_k} \leq (1 + \Lambda_m)\hat{q}^m\|\pi\|_{\infty,J_{k-1}} \tag{1.106}$$

for all $k \in \{1, ..., L\}$. For $k = 2$ this corresponds to the base case of (1.105a). The base case of (1.105b) follows immediately from

$$\begin{aligned}\|\mathcal{I}_2^{(m)}[\exp(i\kappa c_1 \cdot)\pi]\|_{\infty,J_2} &\leq \|\exp(i\kappa c_1 \cdot)\pi\|_{\infty,J_2} + \|E_2[\exp(i\kappa c_1 \cdot)\pi]\|_{\infty,J_2} \\ &\leq (1 + \varepsilon_{m,1})\|\pi\|_{\infty,J_1}.\end{aligned} \tag{1.107}$$

Hence, we can assume that (1.105a) and (1.105b) hold for all k up to $k = n - 1 < L$. To show these two inequalities for $k = n$, we observe that we can write E_n for $n > 2$ as the telescopic sum

$$E_n = (I - \mathcal{I}_2^{(m)}) + (I - \mathcal{I}_3^{(m)}) \circ \mathcal{I}_2^{(m)} + ... + (I - \mathcal{I}_n^{(m)}) \circ \mathcal{I}_{n-1}^{(m)} \circ ... \circ \mathcal{I}_2^{(m)}, \tag{1.108}$$

and estimate the norms of all the summands. For this purpose we use that for all $k \in \{2, ..., L\}$ there holds

$$\mathcal{I}_k^{(m)} \circ ... \circ \mathcal{I}_2^{(m)}[\exp(i\kappa c_1 \cdot)\pi] = \exp(i\kappa c_k \cdot)\hat{\pi},$$

for some appropriate $\hat{\pi} \in \Pi_m$. Using this equation together with (1.106) yields

$$\begin{aligned}&\|(I - \mathcal{I}_{k+1}^{(m)}) \circ \mathcal{I}_k^{(m)} \circ ... \circ \mathcal{I}_2^{(m)}[\exp(i\kappa c_1 \cdot)\pi]\|_{\infty,J_{k+1}} = \|(I - \mathcal{I}_{k+1}^{(m)})[\exp(i\kappa c_k \cdot)\hat{\pi}]\|_{\infty,J_{k+1}} \\ &\quad \leq (1 + \Lambda_m)\hat{q}^m\|\exp(i\kappa c_k \cdot)\hat{\pi}\|_{\infty,J_k} = (1 + \Lambda_m)\hat{q}^m\|\mathcal{I}_k^{(m)} \circ ... \circ \mathcal{I}_2^{(m)}[\exp(i\kappa c_1 \cdot)\pi]\|_{\infty,J_k}.\end{aligned}$$

This estimate, the representation (1.108) and the induction hypothesis (1.105b) for $k \leq n - 1$ are used to show

$$\begin{aligned}&\|E_n[\exp(i\kappa c_1 \cdot)\pi]\|_{\infty,J_\ell} \leq \sum_{k=2}^{n-1} \|(I - \mathcal{I}_{k+1}^{(m)}) \circ \mathcal{I}_k^{(m)} \circ ... \circ \mathcal{I}_2^{(m)}[\exp(i\kappa c_1 \cdot)\pi]\|_{\infty,J_{k+1}} \\ &+ \|(I - \mathcal{I}_2^{(m)})[\exp(i\kappa c_1 \cdot)\pi]\|_{\infty,J_2} \\ &\leq (1 + \Lambda_m)\hat{q}^m \left(\|\exp(i\kappa c_1 \cdot)\pi\|_{\infty,J_1} + \sum_{k=2}^{n-1} \|\mathcal{I}_k^{(m)} \circ ... \circ \mathcal{I}_2^{(m)}[\exp(i\kappa c_1 \cdot)\pi]\|_{\infty,J_k} \right) \\ &\leq (1 + \Lambda_m)\hat{q}^m \sum_{k=1}^{n-1} (1 + \varepsilon_{m,k-1})\|\pi\|_{\infty,J_1} \\ &= (1 + \Lambda_m)\hat{q}^m \frac{(1 + (1 + \Lambda_m)\hat{q}^m)^{n-1} - 1}{(1 + (1 + \Lambda_m)\hat{q}^m) - 1}\|\pi\|_{\infty,J_1} = \varepsilon_{m,n-1}\|\pi\|_{\infty,J_1}.\end{aligned}$$

The estimate (1.105b) for $k = n$ follows as in the base case (1.107) from (1.105a) by the triangle inequality. Therefore we have shown, that (1.105a) and (1.105b) hold for all $k \in \{1, ..., L\}$.

Let now u be an arbitrary function in $C(J_1)$. We define the polynomial π_1 by

$$\pi_1 := \mathcal{I}_{J_1}^{(m)}[\exp(-i\kappa c_1 \cdot)u].$$

Then by Definition (1.6) of the Lebesgue constant Λ_m and due to $|\exp(-i\kappa c_1 x)| = 1$ for all $x \in \mathbb{R}$, there holds

$$\|\pi_1\|_{\infty,J_1} \leq \Lambda_m \|u\|_{\infty,J_1}.$$

Using (1.105b) yields

$$\begin{aligned}\|\mathcal{I}_L^{(m)} \circ ... \circ \mathcal{I}_1^{(m)}[u]\|_{\infty,J_L} &= \|\mathcal{I}_L^{(m)} \circ ... \circ \mathcal{I}_2^{(m)}[\exp(i\kappa c_1 \cdot)\pi_1]\|_{\infty,J_L} \\ &\leq (1 + \varepsilon_{m,L-1})\|\pi_1\|_{\infty,J_1} \leq (1 + \varepsilon_{m,L-1})\Lambda_m \|u\|_{\infty,J_1}.\end{aligned}$$

This is (1.102) for $\ell = 1$, which we wanted to show.

What remains to show is the boundedness of $\varepsilon_{m,k}$ for all $k \in \{1, ..., L-1\}$. Obviously there holds $\varepsilon_{m,k} + 1 \leq \varepsilon_{m,L-1} + 1$ for all $k \leq L - 1$. Due to (1.8) and the definition of $\hat{q}$ there holds

$$(1 + \Lambda_m)\hat{q}^m \leq \left(2 + \frac{\pi}{2}\log(m+1)\right)\left(\frac{1+\bar{q}}{2}\right)^m \to 0 \quad \text{for } m \to \infty.$$

Hence, there exists a constant $C(\bar{q})$ depending only on $\bar{q}$ such that

$$(1 + \Lambda_m)\hat{q}^m \leq C(\bar{q}), \qquad \text{for all } m \in \mathbb{N}.$$

As a consequence, for all $m \in \mathbb{N}$ we have

$$\varepsilon_{m,L-1} + 1 \leq (1 + C(\bar{q}))^{L-1} = C(\bar{q}, L).$$

□

Remark 1.34. In [7] it is shown that $\varepsilon_{m,L}$ decays exponentially in m. As this behavior is only observable for sufficiently large m and is not crucial for the further estimate we have chosen to show only the boundedness of $\varepsilon_{m,\ell}$ for $\ell \in \{1, ..., L-1\}$.

We come to the main result of this section.

Theorem 1.35 ([7, cf. Theorem 5.7])**.** *Let $\{X_k\}_{k=0}^L$ be a sequence of axis-parallel boxes in $\mathbb{R}^3$, such that $X_0 \supset ... \supset X_L$ and $X_k = [a_{k,1}, b_{k,1}] \times ... \times [a_{k,3}, b_{k,3}]$ for all $k \in \{0, ..., L\}$. Assume that for some $\bar{q} < 1$ there holds*

$$\frac{b_{k,j} - a_{k,j}}{b_{k-1,j} - a_{k-1,j}} \leq \bar{q}, \qquad \textit{for all } k \in \{1, ..., L\} \textit{ and } j \in \{1, ..., 3\}. \tag{1.109}$$

Let $\{c_k\}_{k=0}^L \subset \mathbb{R}^3$ *be a sequence of directions with* $|c_k| = 1$ *or* $c_k = 0$ *and let* $\gamma \in \mathbb{R}$ *be such that*

$$\kappa \operatorname{diam}(X_{k-1})\,|c_{k-1} - c_k| \leq \gamma \text{ for all } k \in \{1, ..., L\}. \tag{1.110}$$

Let $\varepsilon_{m,k}$, $C(\bar{q}, l)$, *and* m_0 *be as in Theorem 1.33. Then there holds*

$$\|\mathcal{I}^{(m)}_{X_L,c_L} \circ ... \circ \mathcal{I}^{(m)}_{X_{\ell+1},c_{\ell+1}}\|_{\infty, X_L \leftarrow X_\ell} \leq \Lambda_m^3 (1 + \varepsilon_{m,L-\ell+1})^3 \leq \Lambda_m^3 C(\bar{q}, L)^3, \tag{1.111}$$

for all $m \geq m_0$

Proof. By Lemma 1.28 it suffices to estimate the norms

$$\|\mathcal{I}^{(m)}_{L,j} \circ ... \circ \mathcal{I}^{(m)}_{\ell+1,j}\|_{\infty, J_{L,j} \leftarrow J_{\ell,j}}$$

for $j \in \{1, ..., 3\}$ where $J_{k,j} = [a_{k,j}, b_{k,j}]$ and $\mathcal{I}^{(m)}_{k,j}$ is defined as in (1.82). For this purpose we fix j and define $h_k := b_{k,j} - a_{k,j}$ for all $k \in \{0, ..., L\}$. Due to (1.110) there holds

$$\kappa h_{k-1} |c_{k-1,j} - c_{k,j}| \leq \kappa \operatorname{diam}(X_{k-1})\,|c_{k-1} - c_k| \leq \gamma.$$

Hence by Theorem 1.33 we have

$$\|\mathcal{I}^{(m)}_{L,j} \circ ... \circ \mathcal{I}^{(m)}_{\ell+1,j}\|_{\infty, J_{L,j} \leftarrow J_{\ell,j}} \leq \Lambda_m (1 + \varepsilon_{m,L-\ell}) \leq \Lambda_m C(\bar{q}, L).$$

As this holds true for all $j \in \{1, ..., 3\}$ the assertion follows from (1.83). □

2 Fast Directional Matrix-Vector Multiplication

In this chapter we present an efficient method for the computation of the matrix-vector product for a fully populated matrix $A \in \mathbb{C}^{N_X \times N_Y}$ with entries

$$A[j,k] = f(x_j, y_k), \tag{2.1}$$

where f is the Helmholtz kernel defined in (1.1) and $P_X = \{x_j\}_{j=1}^{N_X}$ and $P_Y = \{y_k\}_{k=1}^{N_Y}$ are two sets of points in $\mathbb{R}^3$. Similar matrices have to be considered in the solution of boundary value problems for the Helmholtz equation via boundary element methods. The corresponding matrix equations can be solved with iterative solvers, for which only matrix-vector products have to be evaluated. However, using standard matrix-vector multiplication is prohibitive for large N_X and N_Y due to the asymptotic runtime of order $\mathcal{O}(N_X N_Y)$. This motivates the discussion of the fast directional matrix-vector multiplication in this chapter. This method is based on a hierarchical partitioning of the sets of points into boxes and the directional multi-level approximation of the Helmholtz kernel f on suitable pairs of such boxes. Basically, this allows us to approximate suitable subblocks of the matrix A similarly as in (1.20), but in a multi-level setting.

In Section 2.1 we give a complete description of the fast directional matrix-vector multiplication. We start with the construction of uniform box cluster trees in Section 2.1.1 which provide us with a hierarchical partitioning of the geometry. The uniformity of the boxes brings benefits which are described later in Section 2.2.1.

In Section 2.1.2 we construct suitable directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ for the directional approximation for all boxes at a given level ℓ of a box cluster tree. This construction is similar to the construction of the directions in [15, 30]: We choose a level ℓ_{hf} as the first level for which directions are used and increase the number of directions for decreasing level ℓ by a hierarchical refinement strategy. After that we define functions dir_ℓ which are used to assign two boxes at level ℓ in two box cluster trees a suitable direction in $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ which should be used for the directional approximation. Furthermore we discuss suitable choices of directions for the directional multi-level approximation. We conclude the section by proving in Corollary 2.17 and Theorem 2.19 that the assumptions of Theorem 1.15 concerning the directions, i.e. (1.32) and the admissibility criterion (A2), hold for constants γ and η_1 depending only on the product κq_{hf}, where q_{hf} is the maximum of the diameters of boxes at level ℓ_{hf} in the corresponding

uniform box cluster trees. To our best knowledge, such a detailed discussion of the suitability of the constructed directions is missing in comparable works.

Section 2.1.3 is dedicated to the partitioning of the matrix A in (2.1). The basic idea is to find all pairs of boxes in two box cluster trees which satisfy the admissibility criteria (A1) and (A3) but whose parents do not. The matrix entries corresponding to the points in such pairs of boxes will then form an admissible block in the partition of the matrix A. All other matrix entries are grouped in inadmissible blocks. Introducing block trees similarly as in [5] allows us to construct and organize such a partition.

After all these preparatory sections we discuss the fast directional matrix-vector multiplication itself in Section 2.1.4. Again we follow [5] for this purpose. We split the computation of the matrix-vector product in a nearfield and a farfield part. The nearfield part is computed directly in Algorithm 2.33. The farfield part is computed using an approximation of the admissible blocks of the matrix A based on the directional approximation of the Helmholtz kernel, similarly as in (1.20). Algorithm 2.38 shows how to compute this farfield part efficiently. All necessary steps for the application of the fast directional matrix-vector multiplication, including the construction of the underlying tree structures and directions, are finally summarized in Algorithm 2.42.

In Section 2.2 we discuss some implementation details which should be considered for an efficient implementation of Algorithm 2.42. In particular we discuss how the uniformity of the box cluster trees can be exploited and talk shortly about possible compression strategies. Then we analyze the complexity of Algorithm 2.42 in Section 2.2.2. Inspired by the complexity analysis in [5] we show that the complexity of the introduced fast directional matrix-vector multiplication is of order $\mathcal{O}(N \log(N))$ under suitable assumptions on the wave number κ, the geometry, and the maximal number of points $N = \max\{N_X, N_Y\}$.

2.1 Matrix partitioning and fast matrix-vector multiplication

In general we cannot approximate the matrix A given in (2.1) directly. Instead we have to partition it and approximate appropriate subblocks. One strategy to create such a matrix partition is clustering, which is based on the geometry or rather the position of the points $\{x_j\}_j^{N_X}$ and $\{y_k\}_k^{N_Y}$. The idea is to assign suitable subsets of these points to axis-parallel boxes, the so-called clusters, and to find pairs of these boxes which satisfy the admissibility criteria (A1) – (A3). Then we can approximate the related subblocks of A corresponding to the points in such pairs of admissible boxes. To do this efficiently we introduce hierarchical structures to organize the boxes and pairs of boxes appropriately. This idea can be found for example in [24, Section 5.2]. In our context, however, the directions which have to be introduced due to (A2) are new. So far we have not talked about a suitable strategy to choose these directions. We

will devote Section 2.1.2 to this topic, after the discussion of clustering by means of box cluster trees in Section 2.1.1. The two remaining subsections will deal with the partitioning of the matrix A and the fast matrix-vector multiplication.

2.1.1 Box cluster trees

The basis of the partitioning of a matrix via clustering are cluster trees. These hierarchical structures are used to organize a set of points in suitable sets on several levels. We define so-called box cluster trees in which these sets correspond to axis-parallel boxes. Additionally we discuss a possibility to construct a special type of such trees.

Definition 2.1 (Box cluster trees)**.** Let $P_X = \{x_j\}_{j=1}^{N_X}$ be a set of pairwise disjoint points in an axis-parallel box $X \subset \mathbb{R}^3$. For a box $t \subset X$ define the set of points of P_X contained in t by

$$P_t := \{x \in P_X : x \in t\}. \tag{2.2}$$

A box cluster tree $\mathcal{T}_X(P_X)$, or shortly $\mathcal{T}_X$, corresponding to P_X and X is a tree satisfying the following conditions:

- Each vertex t corresponds to an axis-parallel box in $\mathbb{R}^3$, which contains at least one point $x \in P_X$.
- The root of the tree is the box X itself.
- For two children t_1 and t_2 of a vertex t with $t_1 \neq t_2$ there holds $P_{t_1} \cap P_{t_2} = \emptyset$.
- There holds
$$P_t = \bigcup_{t' \in \mathrm{child}(t)} P_{t'}.$$

For a box $t \in \mathcal{T}_X$ we define the index set

$$\hat{t} := \{j \in \{1, ..., N_X\} : x_j \in P_t\}. \tag{2.3}$$

We denote the leaves of $\mathcal{T}_X$ by

$$\mathcal{L}_X := \{t \in \mathcal{T}_X : \mathrm{child}(t) = \emptyset\}. \tag{2.4}$$

Furthermore we divide the tree $\mathcal{T}_X$ into levels. For this purpose we define

$$\begin{aligned} \mathcal{T}_X^{(0)} &= \{X\}, \\ \mathcal{T}_X^{(\ell)} &= \{t \in \mathcal{T}_X : \mathrm{parent}(t) \in \mathcal{T}_X^{(\ell-1)}\}, \quad \text{for all } \ell \in \mathbb{N}. \end{aligned} \tag{2.5}$$

For a box $t \in \mathcal{T}_X$ we can then define

$$\mathrm{level}(t) := \{\ell \in \mathbb{N} : t \in \mathcal{T}_X^{(\ell)}\}. \tag{2.6}$$

Finally we denote the depth of a box cluster tree by

$$p(\mathcal{T}_X) = \max\{\mathrm{level}(t) : t \in \mathcal{T}_X\}. \tag{2.7}$$

Remark 2.2. In regard to Definition 2.1 we note:

- In the literature the nodes of cluster trees are sometimes not geometrical objects like boxes, but subsets of an index set I, e.g. $\{1, ..., N_X\}$ in our case. We refer to [24, Definition 5.3.1] for a suitable definition of such cluster trees. The concept of these cluster trees is more general. However, we prefer to keep the geometric background in the focus, which motivates our definition of box cluster trees.
- A priori the index set $\hat{t}$ of the points in a box t in a box cluster tree does not have to consist of consecutive numbers. However, we prefer such sets of consecutive indices when we use box cluster trees for the partitioning of corresponding matrices, which we do in Section 2.1.3. Hence, we reorder the points in a box cluster tree suitably in implementations.

There are various possibilities to construct a box cluster tree corresponding to a set of points P_X. We present here the construction of uniform box cluster trees similarly as in [20, 22], which is also considered in [30, Section 3.1.1].

Algorithm 2.3 (Uniform box cluster tree)**.** Let $n_{\max} \in \mathbb{N}$ and $L_{\max} \in \mathbb{N}_0$. Let P_X be a set of points in $\mathbb{R}^3$. Then we construct a uniform box cluster tree $\mathcal{T}_X(P_X)$ in the following way:

1. Construct a box $X = (a_1, b_1] \times ... \times (a_3, b_3]$ such that all $x \in P_X$ are contained in X and $b_j - a_j = 2h_0$ for an $h_0 \in \mathbb{R}$ and all $j \in \{1, ..., 3\}$. This means that X is a cube which contains P_X and whose edges have the length $2h_0$.
2. If $N_X > n_{\max}$ and $L_{\max} > 0$ we divide X into 8 uniform boxes X_1^1, ..., X_8^1 by subdividing the intervals $(a_j, b_j]$ for all $j \in \{1, ..., 3\}$ into $(a_j, (a_j + b_j)/2]$ and $((a_j+b_j)/2, b_j]$ and forming the boxes X_k^1 for $k \in \{1, ..., 8\}$ as all possible products of such intervals. All formed boxes, which contain at least one point in P_X, are added to the tree $\mathcal{T}_X$ as children of X.
3. Recursively we subdivide all boxes $t \in \mathcal{T}_X$ analogously as in step 2 until each leave $t \in \mathcal{T}_X$ contains either at most $n_{\max}$ points or satisfies $\text{level}(t) = L_{\max}$.

Remark 2.4. Let us comment on the uniform box cluster tree $\mathcal{T}_X(P_X)$ constructed in Algorithm 2.3:

- $\mathcal{T}_X(P_X)$ is obviously a box cluster tree as defined in Definition 2.1. In particular for each point $x \in P_X$ there is a unique leave $t \in \mathcal{L}_X$ such that $x \in t$.
- We choose the initial box X to be a non-closed box only to avoid overlapping boxes in $\mathcal{T}_X^{(l)}$ for all $\ell \in \mathbb{N}$. In an implementation one can choose X and all its descendants as closed boxes. Then, however, one needs to clarify to which box a point $x \in t_1 \cap t_2$ with $t_1, t_2 \in \mathcal{T}_X^{(\ell)}$ for some $\ell > 0$ belongs.
- In step 1 of the construction of $\mathcal{T}_X$ in Algorithm 2.3 we do not explicitly state how to construct the initial box X. One possibility is to proceed similarly as in

[30, Section 3.1.1]. For this purpose choose $\delta > 0$ and set

$$h_0 = \frac{1}{2} \max_{j=1,\dots,3} \left\{ \max_{x \in P_X} x_j - \min_{x \in P_X} x_j \right\} + \delta.$$

Then a suitable box X is given by

$$X = \prod_{j=1}^{3} (a_j, a_j + 2h_0], \qquad a_j := \min_{x \in P_X} x_j - \delta.$$

If one allows closed boxes one can also choose $\delta = 0$ in the definitions above.

- The boxes in $\mathcal{T}_X(P_X)$ depend mostly on the initial box X and not essentially on the actual position of the points in P_X. Therefore it is possible that for a level ℓ in $\mathcal{T}_X^{(\ell)}$ there exist boxes with very few points and at the same time boxes with a lot of points. In particular the constructed tree can be unbalanced. Other construction principles for box cluster trees such as bisection [30, cf. balanced cluster trees in Section 3.1.1] can tailor the tree to the point sets. The resulting box cluster trees are typically balanced. However, they lack the uniformity of the constructed boxes.
- The big advantage of a uniform box cluster tree $\mathcal{T}_X$ is the uniformity of its boxes. For each level ℓ all boxes in $\mathcal{T}_X^{(l)}$ are identical up to translation. In Section 2.2.1 we describe how to exploit this property to avoid recomputations and reduce storage costs in the fast matrix-vector multiplication algorithm. Furthermore, due to this uniformity Definition 2.5 is well-defined.

Definition 2.5. Let $\mathcal{T}_X$ be a uniform box cluster tree as constructed in Algorithm 2.3. Let the depth $p(\mathcal{T}_X)$ be defined as in (2.7). Let $\ell \in \{0, ..., p(\mathcal{T}_X)\}$ and $a_\ell, b_\ell \in \mathbb{R}$ such that all boxes $t \in \mathcal{T}_X^{(\ell)}$ are identical to the cube $[a_\ell, b_\ell]^3$ up to translation. Then we define the half edge length $h_\ell(\mathcal{T}_X)$ of a box at level ℓ by

$$h_\ell(\mathcal{T}_X) := \frac{b_\ell - a_\ell}{2} \tag{2.8}$$

and its diameter $q_\ell(\mathcal{T}_X)$ by

$$q_\ell(\mathcal{T}_X) := \operatorname{diam}\left([a_\ell, b_\ell]^3\right) = \sqrt{3}(b_\ell - a_\ell). \tag{2.9}$$

For a matrix A as in (2.1) and two corresponding sets of points P_X and P_Y we construct two uniform box cluster trees $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$. In Section 2.1.3 we describe how to find suitable pairs of boxes (t, s) with $t \in \mathcal{T}_X$ and $s \in \mathcal{T}_Y$ such that the admissibility criteria (A1) – (A3) in Section 1.2.2 hold for t, s, and an appropriate direction c. This will allow us to partition the matrix A into suitable subblocks which can be approximated efficiently using the directional single-level approximation in (1.20) or its multi-level counterpart. First, however, we have to discuss how to choose suitable directions.

2.1.2 The choice of directions

In the discussion of the admissibility criterion (A2) in Section 1.2.2 we have motivated why we should only consider a fixed number of directions for the directional approximation. In regard of the box cluster trees considered in Section 2.1.1 it makes sense to choose the same directions for all boxes t at level ℓ of a tree. We have already noted in Section 1.2.2 that less directions are needed for smaller box diameters as the right-hand side in (A2) gets larger for decreasing diameters. This means that less directions are needed for higher levels in a box cluster tree.

Below we combine the ideas in [15, Section 4.1], [30, Section 3.3.3], and [5, Section 3] to construct a set of directions for each level. In the former two works the directions are generated by subdividing $\mathbb{R}^3$ into a set of cones. Starting with six initial cones for some level ℓ_{hf} these cones are gradually subdivided for lower tree levels. In [5] the author constructs the directions by splitting the surface of the cube $[-1,1]^3$ in suitable squares and choosing their midpoints as directions. In particular it is shown that this construction method yields suitable directions to satisfy the admissibility criterion (A2). We elaborate this in Proposition 2.13 similarly .

We will further discuss a suitable choice of directions in the multi-level setting. In particular we will show that the error estimate in Theorem 1.15 is applicable. To our best knowledge these properties have not been discussed in detail yet. The estimate (1.32) will follow rather easily in Corollary 2.17 for a suitable choice of the constant γ. Finally we will show that the admissibility criterion (A2) holds for all occuring pairs of boxes in Theorem 2.19 for our choice of directions.

Algorithm 2.6. Let the highest level in the high frequency regime $\ell_{\mathrm{hf}} \in \mathbb{N}_0$ be chosen suitably. For each level $\ell \in \mathbb{N}_0$ we construct a set of N_ℓ directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ as follows:

- For low frequency levels $\ell > \ell_{\mathrm{hf}}$ we set $N_\ell = 1$ and the direction $c_1^{(\ell)} = 0$.
- For $\ell = \ell_{\mathrm{hf}}$ we consider the six faces of the cube $[-1,1]^3$ and denote them by $\{E_j^{(\ell_{\mathrm{hf}})}\}_{j=1}^6$. We set $N_{\ell_{\mathrm{hf}}} = 6$ and choose the directions $\{c_j^{(\ell_{\mathrm{hf}})}\}_{j=1}^6$ such that $c_j^{(\ell_{\mathrm{hf}})}$ is the midpoint of $E_j^{(\ell_{\mathrm{hf}})}$.
- For all other high frequency levels $\ell < \ell_{\mathrm{hf}}$ we set $N_\ell = 6\ 4^{\ell_{\mathrm{hf}}-\ell}$. For each $j \in \{1,...,6\}$ we divide the face $E_j^{(\ell_{\mathrm{hf}})}$ uniformly into $4^{\ell_{\mathrm{hf}}-\ell}$ squares $E_k^{(\ell)}$ with

$$k \in \{(j-1)4^{\ell_{\mathrm{hf}}-\ell} + p : p \in \{1,...,4^{\ell_{\mathrm{hf}}-\ell}\}\}.$$

  We construct $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ by choosing $c_k^{(\ell)}$ as the normalized midpoint of $E_k^{(\ell)}$ for each $k \in \{1,...,N_\ell\}$.

Remark 2.7. In Algorithm 2.6 we do not explain how to order the directions $\{c_j^{(\ell)}\}$ for $\ell \in \{0,...,\ell_{\mathrm{hf}}\}$. However, a consistent ordering of these directions will be needed for Definition 2.9, where we define a mapping that will allow us to find for each vector in $\mathbb{R}^3$ a close direction. Hence, let us sketch a consistent ordering strategy.

First we have to order the six faces $E_j^{(\ell_{\mathrm{hf}})}$ of the cube $[-1,1]^3$. Next we need to find for each of these faces and all levels $\ell \in \{0,...,\ell_{\mathrm{hf}}-1\}$ a suitable ordering of the squares $E_k^{(\ell)}$, in which we subdivide them. This can be done by choosing a fixed orientation for each of the faces $E_j^{(\ell_{\mathrm{hf}})}$ and then enumerating the squares $E_k^{(\ell)}$ row by row. By construction the ordering of these faces $E_j^{(\ell_{\mathrm{hf}})}$ leads to an ordering of the directions.

Example 2.8. To clarify the construction of the directions in Algorithm 2.6 we exemplarily compute some of the directions for a given $\ell_{\mathrm{hf}} > 0$.

For $\ell = \ell_{\mathrm{hf}}$ one of the six faces $E_j^{(\ell_{\mathrm{hf}})}$ is given by

$$E = \{1\} \times [-1,1] \times [-1,1]. \tag{2.10}$$

Its midpoint $c = (1,0,0)$ is one of the six directions $\{c_j^{(\ell_{\mathrm{hf}})}\}_{j=1}^6$ at level ℓ_{hf}.

If $\ell = \ell_{\mathrm{hf}} - 1$ we have to subdivide each of the six faces $E_j^{(\ell_{\mathrm{hf}})}$ uniformly into four squares. For example we divide the face E in (2.10) into the faces

$$\begin{aligned} &\{1\} \times [-1,0] \times [-1,0], && \{1\} \times [-1,0] \times [0,1], \\ &\{1\} \times [0,1] \times [-1,0], && \{1\} \times [0,1] \times [0,1]. \end{aligned}$$

The vector $\tilde{c} = (1,1/2,1/2)$ is the midpoint of the face $\{1\} \times [0,1] \times [0,1]$. By normalizing it we end up with $c = \sqrt{2/3}\ (1,1/2,1/2)$ which is one of the 24 directions $\{c_j^{(\ell_{\mathrm{hf}}-1)}\}_{j=1}^{24}$.

We proceed by defining for each level $\ell \in \mathbb{N}_0$ a mapping $\mathrm{dir}_{(\ell)}$, which maps a vector in $\mathbb{R}^3 \backslash \{0\}$ to one of the directions $c_j^{(\ell)}$. The idea is to select a direction $c_j^{(\ell)}$ for each $v \neq 0$ and each high frequency level $\ell \leq \ell_{\mathrm{hf}}$ such that the intersection point of the ray $\{\lambda v : \lambda > 0\}$ and the surface of the cube $[-1,1]^3$ lies in the face $E_j^{(\ell)}$, whose normalized middlepoint is $c_j^{(\ell)}$ by construction in Algorithm 2.6. We will use these mappings later to choose suitable directions for the directional multi-level approximation. There we will not only need to find a suitable direction for two boxes t and s such that (A2) holds, but will also have to switch from a direction on a certain level to a close direction on the next level.

Definition 2.9. Let $\ell_{\mathrm{hf}} \in \mathbb{N}_0$ and let the directions $\{c_j^{(\ell)}\}$ and the faces $\{E_j^{(\ell)}\}$ be constructed as in Algorithm 2.6. We define the mapping

$$\psi_Q : \mathbb{R}^3 \backslash \{0\} \to \partial([-1,1]^3), \qquad v \mapsto \frac{1}{\max\limits_{j \in \{1,...,3\}} |v_j|}\, v. \tag{2.11}$$

Furthermore we define the mapping $\mathrm{dir}_{(\ell)} : \mathbb{R}^3 \to \{c_j^{(\ell)}\}_{j=1}^{N_\ell} \cup \{0\}$ for each $\ell \in \mathbb{N}_0$ as follows:

- If $\ell > \ell_{\mathrm{hf}}$ we set $\mathrm{dir}_{(\ell)}(v) = 0$ for all $v \in \mathbb{R}^3$.

- If $\ell \leq \ell_{\text{hf}}$ then we set $\text{dir}_{(\ell)}(0) = 0$. For all $v \in \mathbb{R}^3 \backslash \{0\}$ we define

$$j(v) := \min\{j : \psi_Q(v) \in E_j^{(\ell)}\},$$

  and set $\text{dir}_{(\ell)}(v) = c_{j(v)}^{(\ell)}$.

Remark 2.10. Let $v \in \mathbb{R}^3 \backslash \{0\}$. If $\psi_Q(v) \in E_{j_0}^{(0)}$ for some $j_0 \in \{1, ..., N_0\}$ then there exists a sequence $\{j_\ell\}_{\ell=0}^{\ell_{\text{hf}}}$ such that $E_{j_0}^{(0)} \subset E_{j_1}^{(1)} \subset ... \subset E_{j_{\ell_{\text{hf}}}}^{(\ell_{\text{hf}})}$ and $\psi_Q(v) \in E_{j_\ell}^{(\ell)}$ for all $\ell \in \{0, ..., \ell_{\text{hf}}\}$ due to the nested construction of the squares $E_k^{(\ell)}$ in Algorithm 2.6. Suppose that $\psi_Q(v)$ does not lie on a corner or edge of $E_{j_0}^{(0)}$. Then the sequence $\{j_{(\ell)}\}_{\ell=0}^{\ell_{\text{hf}}}$ is unique and $\text{dir}_{(\ell)}(v) = c_{j_\ell}^{(\ell)}$, i.e. $\text{dir}_{(\ell)}$ maps v to the normalized midpoint of $E_{j_\ell}^{(\ell)}$. In particular it follows that

$$\text{dir}_{(\ell+1)}(v) = \text{dir}_{(\ell+1)}(\text{dir}_{(\ell)}(v)) \tag{2.12}$$

for all $\ell \in \mathbb{N}_0$. Indeed, if $\ell < \ell_{\text{hf}}$ then $\psi_Q(c_{j_\ell}^{(\ell)}) \in E_{j_\ell}^{(\ell)} \subset E_{j_{\ell+1}}^{(\ell+1)}$ which yields

$$\text{dir}_{(\ell+1)}(\text{dir}_{(\ell)}(v)) = \text{dir}_{(\ell+1)}(c_{j_\ell}^{(\ell)}) = c_{j_{\ell+1}}^{(\ell+1)} = \text{dir}_{(\ell+1)}(v).$$

If $\ell \geq \ell_{\text{hf}}$ then $\text{dir}_{(\ell+1)}$ maps all vectors in $\mathbb{R}^3$ to zero and hence (2.12) holds trivially.

On the other hand, if $\psi_Q(v)$ lies on a corner or edge of $E_{j_0}^{(0)}$, then the sequence $\{j_\ell\}_{\ell=0}^{\ell_{\text{hf}}}$ is not unique anymore. However, we assume that the ordering of the squares $\{E_k^{(\ell)}\}_{k=1}^{N_\ell}$ is such that (2.12) holds. This can be shown for the ordering sketched in Remark 2.7. Hence, we can assume from here on that (2.12) holds for all $v \in \mathbb{R}^3 \backslash \{0\}$.

With the help of the mappings $\text{dir}_{(\ell)}$ we can define for each pair of boxes (t, s) a direction $c_\ell(t, s)$ in the set of directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ constructed in Algorithm 2.6, which is close to the normalized difference $(m_t - m_s)/|m_t - m_s|$ of the midpoints of t and s.

Definition 2.11. Let $t, s \subset \mathbb{R}^3$ be two boxes in $\mathbb{R}^3$ and m_t and m_s their midpoints. Let $m_t \neq m_s$. Then we define for all $\ell \in \mathbb{N}_0$ the direction $c_{(\ell)}(t, s)$ by

$$c_{(\ell)}(t, s) := \text{dir}_{(\ell)}\left(\frac{m_t - m_s}{|m_t - m_s|}\right). \tag{2.13}$$

The question arises if (A2) is satisfied for two boxes t, s, and the direction $c_{(\ell)}(t, s)$ for some $\ell \in \mathbb{N}$ or more precisely how large η_1 has to be chosen such that it is satisfied. We discuss this in Proposition 2.13. First we show that we can estimate the difference of two directions c_1, $c_2 \in \mathbb{R}^3$ with $|c_1| = |c_2| = 1$ by the distance of their images on $\partial([-1, 1]^3)$ under the mapping ψ_Q defined in (2.11). This result, which is an immediate consequence of the following lemma, plays a central role in the proofs of all following results in this section.

Lemma 2.12 (Projection [5, Lemma 7])**.** *Let $x,\ y \in \mathbb{R}^3$ with $|x|, |y| \geq 1$. Then*

$$\left| \frac{x}{|x|} - \frac{y}{|y|} \right| \leq |x - y|. \tag{2.14}$$

Proof. Without loss of generality we can assume that

$$1 \leq |x| \leq |y|. \tag{2.15}$$

Let $\lambda = \langle x, y \rangle / (|x||y|^2)$. Then there holds

$$\left\langle \frac{x}{|x|} - \lambda y, y \right\rangle = \frac{\langle x, y \rangle}{|x|} - \frac{\langle x, y \rangle}{|x||y|^2} |y|^2 = 0.$$

Using this property and (2.15) yields

$$\begin{aligned}
\left| \frac{x}{|x|} - \frac{y}{|y|} \right|^2 &= \left| \frac{x}{|x|} - \lambda y + \left(\lambda - \frac{1}{|y|} \right) y \right|^2 \\
&= \left| \frac{x}{|x|} - \lambda y \right|^2 + \left(\lambda - \frac{1}{|y|} \right)^2 |y|^2 + 2 \left(\lambda - \frac{1}{|y|} \right) \left\langle \frac{x}{|x|} - \lambda y, y \right\rangle \\
&= \left| \frac{x}{|x|} - \lambda y \right|^2 + \left(\frac{|x||y| - \langle x, y \rangle}{|x||y|^2} \right)^2 |y|^2 \\
&\leq \left| \frac{x}{|x|} - \lambda y \right|^2 + \left(\frac{|y|^2 - \langle x, y \rangle}{|x||y|^2} \right)^2 |y|^2 \\
&= \left| \frac{x}{|x|} - \lambda y \right|^2 + \left(\lambda - \frac{1}{|x|} \right)^2 |y|^2 + 2 \left(\lambda - \frac{1}{|x|} \right) \left\langle \frac{x}{|x|} - \lambda y, y \right\rangle \\
&= \left| \frac{x}{|x|} - \lambda y + \lambda y - \frac{y}{|x|} \right|^2 = \frac{|x - y|^2}{|x|^2} \leq |x - y|^2.
\end{aligned}$$

□

Proposition 2.13. *Let $\ell_{\mathrm{hf}} \in \mathbb{N}_0$. Let the directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ be constructed as in Algorithm 2.6 and the mappings $\mathrm{dir}_{(\ell)}$ be defined as in Definition 2.9. Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees constructed by Algorithm 2.3. For a pair of boxes (t, s) with $t \cap s = \emptyset$ and $\ell \in \mathbb{N}_0$ let the direction $c_{(\ell)}(t, s)$ be defined by* (2.13)*. Let $\kappa > 0$ and let the characteristic high frequency diameter q_{hf} be defined by*

$$q_{\mathrm{hf}} := \max\{\mathrm{diam}\,(X), \mathrm{diam}\,(Y)\} 2^{-\ell_{\mathrm{hf}}}. \tag{2.16}$$

Finally let

$$\eta_1 = \sqrt{2} \kappa q_{\mathrm{hf}}. \tag{2.17}$$

Then for all levels $\ell \in \mathbb{N}_0$ and all boxes $t \in \mathcal{T}_X^{(\ell)}$, $s \in \mathcal{T}_Y^{(\ell)}$ with midpoints m_t and m_s satisfying $m_t \neq m_s$ there holds the admissibility condition (A2) *for the constructed direction $c_{(\ell)}(t,s)$, i.e.*

$$\kappa \left| \frac{m_t - m_s}{|m_t - m_s|} - c_{(\ell)}(t,s) \right| \leq \frac{\eta_1}{\max\{\operatorname{diam}(t), \operatorname{diam}(s)\}}. \tag{2.18}$$

Proof. We start by showing the estimate

$$|c - \operatorname{dir}_{(\ell)}(c)| \leq \sqrt{2}\, 2^{\ell - \ell_{\mathrm{hf}}} \tag{2.19}$$

for all $c \in S^2 := \{c \in \mathbb{R}^3 : |c| = 1\}$ and all $\ell \in \mathbb{N}_0$. For $\ell > \ell_{\mathrm{hf}}$ this is trivially satisfied as in this case $|c| = 1$, $\operatorname{dir}_{(\ell)}(c) = 0$ and the right hand side is greater than 1.

Hence, let $\ell \in \{0, ..., \ell_{\mathrm{hf}}\}$. Consider the faces $\{E_k^{(\ell)}\}_{k=1}^{N_\ell}$ from the construction of the directions $\{c_k^{(\ell)}\}_{k=1}^{N_\ell}$ in Algorithm 2.6. All these faces are squares with edge length $2^{\ell+1-\ell_{\mathrm{hf}}}$. By definition of the function ψ_Q in (2.11) and by construction of the directions $\{c_k^{(\ell)}\}_{k=1}^{N_\ell}$, their midpoints are given by $\{\psi_Q(c_k^{(\ell)})\}_{k=1}^{N_\ell}$. The distance of an arbitrary point $y \in E_k^{(\ell)}$ and its midpoint $\psi_Q(c_k^{(\ell)})$ can therefore be estimated by

$$|y - \psi_Q(c_k^{(\ell)})| \leq \sqrt{2}\, 2^{\ell - \ell_{\mathrm{hf}}}. \tag{2.20}$$

Let $c \in S^2$ be fixed. By Definition 2.9 the direction $\operatorname{dir}_{(\ell)}(c)$ is such that $\psi_Q(c)$ and $\psi_Q(\operatorname{dir}_{(\ell)}(c))$ are contained in the same square $E_k^{(\ell)}$ for a $k \in \{1, ..., N_\ell\}$. Hence, by using Lemma 2.12 and (2.20) we get

$$|c - \operatorname{dir}_{(\ell)}(c)| \leq |\psi_Q(c) - \psi_Q(\operatorname{dir}_{(\ell)}(c))| = |\psi_Q(c) - c_k^{(\ell)}| \leq \sqrt{2}\, 2^{\ell - \ell_{\mathrm{hf}}}.$$

For the proof of the main assertion (2.18) let η_1 be given by (2.17). Due to the uniformity of the cluster trees we have

$$\max\{\operatorname{diam}(t), \operatorname{diam}(s)\} = \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\} 2^{-\ell} = q_{\mathrm{hf}} 2^{\ell_{\mathrm{hf}} - \ell}$$

for all $\ell \in \mathbb{N}_0$ and arbitrary boxes $t \in \mathcal{T}_X^{(\ell)}$ and $s \in \mathcal{T}_Y^{(\ell)}$. Together with (2.19) and the definition of $c_{(\ell)}(t,s)$ in (2.13) this yields

$$\begin{aligned} \kappa \left| \frac{m_t - m_s}{|m_t - m_s|} - c_{(\ell)}(t,s) \right| &= \kappa \left| \frac{m_t - m_s}{|m_t - m_s|} - \operatorname{dir}_{(\ell)}\left(\frac{m_t - m_s}{|m_t - m_s|} \right) \right| \\ &\leq \kappa \sqrt{2}\, 2^{\ell - \ell_{\mathrm{hf}}} = \frac{\eta_1}{q_{\mathrm{hf}} 2^{\ell_{\mathrm{hf}} - \ell}} = \frac{\eta_1}{\max\{\operatorname{diam}(t), \operatorname{diam}(s)\}}. \end{aligned}$$

□

Remark 2.14. Let the assumptions of Proposition 2.13 hold and let $\ell > \ell_{\mathrm{hf}}$. Then one can easily check that for two boxes $t \in \mathcal{T}_X^{(\ell)}$ and $s \in \mathcal{T}_Y^{(\ell)}$ the pair (t,s) is in the low frequency regime with respect to η_1 (cf. Definition 1.14). Hence it is reasonable that the direction $c_{(\ell)}(t,s)$ is chosen to be zero for such boxes.

Remark 2.15. If the depths $p(\mathcal{T}_X)$ and $p(\mathcal{T}_Y)$ are at least ℓ_{hf}, then due to the uniformity of the cluster trees the characteristic high frequency diameter q_{hf} in (2.16) corresponds to the maximum of the diameters of boxes at level ℓ_{hf} in $\mathcal{T}_X$ and $\mathcal{T}_Y$, i.e.

$$q_{\text{hf}} = \max\{q_{\ell_{\text{hf}}}(\mathcal{T}_X), q_{\ell_{\text{hf}}}(\mathcal{T}_Y)\}. \tag{2.21}$$

We have yet to choose suitable directions for the directional reinterpolation. Let us recall the discussion in Section 1.3. For two sequences of boxes $\{t_j\}_{j=0}^L$ and $\{s_j\}_{j=0}^L$ with $t_0 \supset ... \supset t_L$ and $s_0 \supset ... \supset s_L$, and a sequence of directions $\{c_j\}_{j=0}^L$ we define the directional multi-level approximation $\hat{f}_{t_0,s_0}^{(m)}$ on $t_L \times s_L$ as in (1.26). By Theorem 1.15 we get exponential convergence of the approximation error

$$\|f - \hat{f}_{t_0,s_0}^{(m)}\|_{\infty,t_L\times s_L}$$

in m under suitable assumptions. In particular there should exist $\gamma \in \mathbb{R}$ and $\eta_1 \in \mathbb{R}$ such that for all $\ell \in \{1, ..., L\}$ there holds (1.32), i.e.

$$\kappa \max\{\operatorname{diam}(t_{\ell-1}), \operatorname{diam}(s_{\ell-1})\}|c_\ell - c_{\ell-1}| \leq \gamma,$$

and the admissibility criterion (A2) should be satisfied for the boxes t_ℓ, s_ℓ, and the direction c_ℓ. We will show in Corollary 2.17 and Theorem 2.19 that these conditions hold true for sequences of boxes $\{t_j\}_{j=0}^L$ and $\{s_j\}_{j=0}^L$ in uniform box cluster trees $\mathcal{T}_X$ and $\mathcal{T}_Y$, where $t_0 \in \mathcal{T}_X^{(\ell_0)}$ and $s_0 \in \mathcal{T}_Y^{(\ell_0)}$ for some $\ell_0 \in \mathbb{N}_0$, and the directions $\{c_j\}_{j=0}^L$ such that $c_j = c_{(j+\ell_0)}(t_0, s_0)$. The choice of $c_0 = c_{(\ell_0)}(t_0, s_0)$ was already motivated in Proposition 2.13. We can think of c_j as the naturally inherited directions from c_0 by means of the mappings $\operatorname{dir}_{(\ell_0+j)}$ given in Definition 2.9. Indeed there holds $c_j = \operatorname{dir}_{(\ell_0)+j)}(c_{j-1})$ by definition of the directions $c_{(\ell_0+j)}(t_0, s_0)$ in (2.13) and (2.12). This will allow us to show (1.32) in Corollary 2.17 by proving the following theorem.

Theorem 2.16. *Let $\kappa > 0$ and $\ell_{\text{hf}} \in \mathbb{N}_0$. Let the directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ be constructed as in Algorithm 2.6 and the mappings $\operatorname{dir}_{(\ell)}$ be defined as in Definition 2.9. Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees constructed by Algorithm 2.3. Let the diameters $q_\ell(\mathcal{T}_X)$ and $q_\ell(\mathcal{T}_Y)$ for all $\ell \leq p(\mathcal{T}_X)$ and $\ell \leq p(\mathcal{T}_Y)$, respectively, be defined by* (2.9) *and let q_{hf} be defined by* (2.16)*. Then for all $\ell \leq \min\{p(\mathcal{T}_X), p(\mathcal{T}_Y)\}$ and all $j \in \{1, ..., N_\ell\}$ we have*

$$\kappa \max\{q_\ell(\mathcal{T}_X), q_\ell(\mathcal{T}_Y)\}|c_j^{(\ell)} - \operatorname{dir}_{(\ell+1)}(c_j^{(\ell)})| \leq \gamma, \tag{2.22}$$

where the constant γ is given by

$$\gamma = \sqrt{2}\kappa q_{\text{hf}}. \tag{2.23}$$

Proof. Without loss of generality we assume that $\min\{p(\mathcal{T}_X), p(\mathcal{T}_Y)\} > \ell_{\text{hf}}$. Then we show (2.22) for all $\ell \leq \min\{p(\mathcal{T}_X), p(\mathcal{T}_Y)\}$ by distinguishing three cases.

If $\ell > \ell_{\mathrm{hf}}$, i.e. the level ℓ is in the low frequency regime, then the directions $c_j^{(\ell)}$ and $\mathrm{dir}_{(\ell+1)}(c_j^{(\ell)})$ are zero by construction. Hence (2.22) holds trivially.

If $\ell = \ell_{\mathrm{hf}}$ then $\mathrm{dir}_{(\ell+1)}(c_j^{(\ell)}) = 0$ for all j. Due to (2.21) we get

$$\kappa \max\{q_\ell(\mathcal{T}_X), q_\ell(\mathcal{T}_Y)\}|c_j^{(\ell)} - \mathrm{dir}_{(\ell+1)}(c_j^{(\ell)})| = \kappa q_{\mathrm{hf}}|c_j^{\ell}| = \kappa q_{\mathrm{hf}} \leq \gamma,$$

for γ in (2.23).

Finally let $\ell < \ell_{\mathrm{hf}}$. We consider a fixed direction $c_j^{(\ell)}$ and the related direction $c_k^{(\ell+1)} = \mathrm{dir}_{(\ell+1)}(c_j^{(\ell)})$. Let the mapping $\psi_Q : \mathbb{R}^3\backslash\{0\} \to \partial([-1,1]^3)$ be given by (2.11). By construction of the directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ in Algorithm 2.6 we know that $\psi_Q(c_j^{(\ell)})$ is the midpoint of the square $E_j^{(\ell)}$ with edge length $2^{\ell+1-\ell_{\mathrm{hf}}}$ and $\psi_Q(c_k^{(\ell+1)})$ is the midpoint of the square $E_k^{(\ell+1)}$ with edge length $2^{\ell+2-\ell_{\mathrm{hf}}}$, where the squares $E_j^{(\ell)}$ and $E_k^{(\ell+1)}$ are defined as in Algorithm 2.6. Furthermore $\psi_Q(c_j^{(\ell)})$ lies in the square $E_k^{(\ell+1)}$ by Definition 2.9 of the mapping $\mathrm{dir}_{(\ell+1)}$, since $\mathrm{dir}_{(\ell+1)}(c_j^{(\ell)}) = c_k^{(\ell+1)}$. Hence, by construction $E_j^{(\ell)}$ is one of the squares which we get when we subdivide $E_k^{(\ell+1)}$ uniformly into four squares. The distance of their midpoints is therefore given by

$$|\psi_Q(c_k^{(\ell+1)}) - \psi_Q(c_j^{(\ell)})| = \frac{1}{4}\,\mathrm{diam}\left(E_k^{(\ell+1)}\right) = \sqrt{2}\;2^{\ell-\ell_{\mathrm{hf}}}.$$

Together with Lemma 2.12 this yields

$$|c_j^{(\ell)} - \mathrm{dir}_{(\ell+1)}(c_j^{(\ell)})| \leq |\psi_Q(c_k^{(\ell+1)}) - \psi_Q(c_j^{(\ell)})| = \sqrt{2}\;2^{\ell-\ell_{\mathrm{hf}}},$$

and due to the uniformity of the cluster trees we finally get

$$\begin{aligned}\kappa \max\{q_\ell(\mathcal{T}_X), q_\ell(\mathcal{T}_Y)\}|c_j^{(\ell)} - \mathrm{dir}_{(\ell+1)}(c_j^{(\ell)})| &= \kappa q_{\mathrm{hf}} 2^{\ell_{\mathrm{hf}}-\ell}|c_j^{(\ell)} - \mathrm{dir}_{(\ell+1)}(c_j^{(\ell)})| \\ &\leq \kappa q_{\mathrm{hf}} 2^{\ell_{\mathrm{hf}}-\ell}\sqrt{2}\;2^{\ell-\ell_{\mathrm{hf}}} = \gamma. \qquad \square\end{aligned}$$

Corollary 2.17. *Let the assumptions of Theorem 2.16 hold. Let $\{t_j\}_{j=0}^L$ and $\{s_j\}_{j=0}^L$ be two sequences of boxes in $\mathcal{T}_X$ and $\mathcal{T}_Y$, respectively, and let $\ell_0 \in \mathbb{N}_0$ be such that*

$$t_0 \supset \ldots \supset t_L, \quad s_0 \supset \ldots \supset s_L, \quad t_j \in \mathcal{T}_X^{(j+\ell_0)}, s_j \in \mathcal{T}_Y^{(j+\ell_0)} \text{ for all } j \in \{0,\ldots,L\}. \tag{2.24}$$

Assume further that $s_0 \cap t_0 = \emptyset$ and define the directions

$$c_j = c_{(\ell_0+j)}(t_0, s_0) \qquad \text{for all } j \in \{0,\ldots,L\}, \tag{2.25}$$

with $c_{(\ell_0+j)}(t_0,s_0)$ defined as in (2.13). Then for all $\ell \in \{1,\ldots,L\}$ there holds (1.32), i.e. we have

$$\kappa \max\{\mathrm{diam}\,(t_{\ell-1}), \mathrm{diam}\,(s_{\ell-1})\}|c_\ell - c_{\ell-1}| \leq \gamma,$$

for the constant γ in (2.23).

Proof. Due to (2.12) and the definition of the directions $c_{(\ell_0+j)}(t_0, s_0)$ there holds

$$c_j = c_{(\ell_0+j)}(t_0, s_0) = \mathrm{dir}_{(\ell_0+j)}\left(\frac{m_{t_0} - m_{s_0}}{|m_{t_0} - m_{s_0}|}\right) = \mathrm{dir}_{(\ell_0+j)}\left(\mathrm{dir}_{(\ell_0+j-1)}\left(\frac{m_{t_0} - m_{s_0}}{|m_{t_0} - m_{s_0}|}\right)\right)$$
$$= \mathrm{dir}_{(\ell_0+j)}(c_{(\ell_0+j-1)}(t_0, s_0)) = \mathrm{dir}_{(\ell_0+j)}(c_{j-1}),$$

for all $j \in \{1, ..., L\}$. Hence, the assertion is a direct consequence of Theorem 2.16. □

Let $\{t_j\}_{j=0}^L$ and $\{s_j\}_{j=0}^L$ be two sequences of boxes in two uniform box cluster trees $\mathcal{T}_X$ and $\mathcal{T}_Y$ such that (2.24) holds. Let the directions $\{c_j\}_{j=0}^L$ be given by (2.25). We want to show that there exists an $\eta_1 > 0$ such that (A2) holds for the boxes t_j, s_j, and the direction c_j for all $j \in \{0, ..., L\}$. Unfortunately we cannot apply Proposition 2.13 for this purpose. The reason is that in general

$$c_j = c_{(\ell_0+j)}(t_0, s_0) \neq c_{(\ell_0+j)}(t_j, s_j), \tag{2.26}$$

i.e. in general c_j is not the best direction to approximate $(m_{t_j} - m_{s_j})/|m_{t_j} - m_{s_j}|$. However, by assuming that t_0 and s_0 satisfy the admissibility condition (A3) we can show that (A2) holds for the boxes t_j, s_j, and the direction c_j, for all $j \in \{1, ..., L\}$. For this purpose, however, η_1 has to be chosen greater than in Proposition 2.13. To show this we need the following lemma.

Lemma 2.18. *Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees and let q_{hf} be defined by* (2.16)*. Let two sequences of boxes $\{t_j\}_{j=0}^L$ and $\{s_j\}_{j=0}^L$ in $\mathcal{T}_X$ and $\mathcal{T}_Y$, respectively, satisfy* (2.24)*. Let $\kappa > 0$ and $\ell_{\mathrm{hf}} \in \mathbb{N}_0$. Let the directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ be constructed as in Algorithm 2.6 and the directions $c_{(\ell)}(t, s)$ for two boxes t and s with $t \cap s = \emptyset$ be defined by* (2.13)*. Assume that t_0 and s_0 satisfy the admissibility criterion* (A3) *for an η_2 such that*

$$\eta_2 \leq 4\kappa q_{\mathrm{hf}}. \tag{2.27}$$

Then for all $j \in \{1, ..., L\}$ there holds

$$|c_{(j+\ell_0)}(t_0, s_0) - c_{(j+\ell_0)}(t_j, s_j)| \leq \sqrt{2}\ 2^{\ell_0+j+1-\ell_{\mathrm{hf}}}. \tag{2.28}$$

Proof. For $j > \ell_{\mathrm{hf}} - \ell_0$ the assertion is trivially satisfied as $c_{(j+\ell_0)}(t_0, s_0) = 0$ and $c_{(j+\ell_0)}(t_j, s_j) = 0$. Hence let j be such that $1 \leq j \leq \ell_{\mathrm{hf}} - \ell_0$. Let k_1 and k_2 be such that $c_{(j+\ell_0)}(t_0, s_0) = c_{k_1}^{(\ell_0+j)}$ and $c_{(j+\ell_0)}(t_j, s_j) = c_{k_2}^{(\ell_0+j)}$, and let $E_{k_1}^{(\ell_0+j)}$ and $E_{k_2}^{(\ell_0+j)}$ be the corresponding squares used for the construction of these directions in Algorithm 2.6. The idea of the proof is to show that

$$E_{k_1}^{(\ell_0+j)} \cap E_{k_2}^{(\ell_0+j)} \neq \emptyset, \tag{2.29}$$

i.e. that these squares share at least a common corner. Then (2.28) will follow immediately. Indeed, one can easily check that the maximal distance of the midpoints of

two squares in $\{E_j^{(\ell_0+j)}\}_{j=1}^{N_{\ell_0+j}}$ satisfying (2.29) is given by $\sqrt{2}\ 2^{\ell_0+j+1-\ell_{\rm hf}}$. Hence, by Lemma 2.12 we get

$$\begin{aligned}|c_{(j+\ell_0)}(t_0,s_0)-c_{(j+\ell_0)}(t_j,s_j)| &= |c_{k_1}^{(\ell_0+j)}-c_{k_2}^{(\ell_0+j)}| \le |\psi_Q(c_{k_1}^{(\ell_0+j)})-\psi_Q(c_{k_1}^{(\ell_0+j)})| \\ &\le \sqrt{2}\ 2^{\ell_0+j+1-\ell_{\rm hf}},\end{aligned}$$

which is (2.28).

To show (2.29) we consider the vectors $\tilde{c}_{k_1}=\psi_Q(m_{t_0}-m_{s_0})$ and $\tilde{c}_{k_2}=\psi_Q(m_{t_j}-m_{s_j})$, where the mapping ψ_Q is defined in (2.11) and m_{t_j} and m_{s_j} are the midpoints of t_j and s_j, respectively. From $c_{(j+\ell_0)}(t_0,s_0)=c_{k_1}^{(\ell_0+j)}$ we get that $\tilde{c}_{k_1}\in E_{k_1}^{(\ell_0+j)}$ by (2.13) and the Defintion 2.9 of $\mathrm{dir}_{(\ell_0+j)}$. Likewise we have $\tilde{c}_{k_2}\in E_{k_2}^{(\ell_0+j)}$. Our goal is to show the componentwise estimate

$$|\tilde{c}_{k_1,n}-\tilde{c}_{k_2,n}|<2^{\ell_0+j+1-\ell_{\rm hf}},\quad \text{for all } n\in\{1,...,3\} \text{ and } j\in\{1,...,\ell_{\rm hf}-\ell_0\}. \tag{2.30}$$

From these estimates we immediately get (2.29), since the edges of the squares $E_{k_1}^{(\ell_0+j)}$ and $E_{k_2}^{(\ell_0+j)}$ have length $2^{\ell_0+j+1-\ell_{hf}}$ and are arranged on a regular grid on the surface of $\partial([-1,1]^3)$. Instead of showing (2.30) directly we show for arbitrary $x\in t_0$ and $y\in s_0$ the componentwise estimate

$$|(\psi_Q(x-y))_n-\tilde{c}_{k_1,n}|<2^{3+\ell_0-\ell_{\rm hf}},\qquad \text{for all } n\in\{1,...,3\}. \tag{2.31}$$

For $j\ge 2$ this yields (2.30) if we choose x and y as the midpoints of the boxes t_j and s_j, i.e. $x=m_{t_j}$, and $y=m_{s_j}$. The case $j=1$ is treated separately afterwards.

For the proof of (2.31) we first define the half edge lengths $h_{t_0}=h_{\ell_0}(\mathcal{T}_X)$ and $h_{s_0}=h_{\ell_0}(\mathcal{T}_Y)$ as in (2.8) and set $h_{\max}:=\max\{h_{t_0},h_{s_0}\}$. Let us write the midpoint $m_{t_0}=(d_1,d_2,d_3)$ and assume without loss of generality that $m_{s_0}=0$. Due to symmetry arguments we can assume that $d_1\ge d_2\ge 0$ and $d_1\ge d_3\ge 0$. This yields $d_1>h_{t_0}+h_{s_0}$. Indeed, in the contrary case we would get $0\le d_n\le h_{t_0}+h_{s_0}$ for all $n\in\{1,...,3\}$ and hence $\mathrm{dist}\,(t_0,s_0)=0$ in contradiction to the assumption that t_0 and s_0 satisfy (A3). For arbitrary $x\in t_0$ and $y\in s_0$ we know by definition of h_{t_0} and h_{s_0} that

$$|x_n-m_{t_0,n}|\le h_{t_0},\qquad |y_n-m_{s_0,n}|=|y_n|\le h_{s_0},\qquad \text{for all } n\in\{1,...,3\}.$$

In particular we can write $x-y=m_{t_0}+z=(d_1+z_1,d_2+z_2,d_3+z_3)$ with

$$|z_n|=|x_n-y_n-d_n|\le|x_n-d_n|+|y_n|\le h_{t_0}+h_{s_0}\quad \text{for all } n\in\{1,...,3\}. \tag{2.32}$$

For the computation of $\psi_Q(x-y)$ we need the maximum $\mu=\max_n\{|x_n-y_n|\}$. From

$$\begin{aligned}|x_n-y_n|=|d_n+z_n|\le d_n+|z_n|\le d_1+h_{t_0}+h_{s_0}\quad &\text{and}\\ \mu\ge|x_1-y_1|=|d_1+z_1|\ge d_1-h_{t_0}-h_{s_0}&\end{aligned}$$

we get the estimate

$$|\mu - d_1| \leq h_{t_0} + h_{s_0}.$$

Together with (2.32) this yields the componentwise estimate

$$\begin{aligned} |\tilde{c}_{k_1,n} - (\psi_Q(x-y))_n| &= \left| \frac{d_n}{d_1} - \frac{d_n + z_n}{\mu} \right| = \left| \frac{d_n(\mu - d_1) - z_n d_1}{d_1 \mu} \right| \\ &\leq \frac{d_n |\mu - d_1| + |z_n| d_1}{d_1 \mu} \leq \frac{4 d_1 h_{\max}}{d_1 \mu} \leq \frac{4 h_{\max}}{d_1 - h_{t_0} - h_{s_0}} \end{aligned} \tag{2.33}$$

for all $n \in \{1, ..., 3\}$. To continue the estimate in (2.33) we have to bound $d_1 - h_{t_0} - h_{s_0}$ from below. For this purpose we note that due to our assumption on the midpoints of the boxes t_0 and s_0 there holds

$$\begin{aligned} \operatorname{dist}(t_0, s_0)^2 &= \sum_{n=1}^{3} \operatorname{dist}([-h_{s_0}, h_{s_0}], [d_n - h_{t_0}, d_n + h_{t_0}])^2 \\ &= \sum_{n=1}^{3} \max\{0, (d_n - h_{t_0} - h_{s_0})^2\} \leq 3(d_1 - h_{t_0} - h_{s_0})^2, \end{aligned}$$

where for the last inequality we used $d_1 \geq \max\{d_2, d_3, h_{t_0} + h_{s_0}\}$. Together with the relation $\max\{\operatorname{diam}(t_0), \operatorname{diam}(s_0)\} = 2\sqrt{3} h_{\max}$ and the assumption that (A3) holds for t_0 and s_0 this yields

$$(d_1 - h_{t_0} - h_{s_0}) \geq \frac{1}{\sqrt{3}} \operatorname{dist}(t_0, s_0) \geq \frac{\kappa}{\sqrt{3}\eta_2} \max\{\operatorname{diam}(t_0), \operatorname{diam}(s_0)\}^2 = \frac{4\sqrt{3}\kappa h_{\max}^2}{\eta_2}.$$

In particular with (2.27) and the equality

$$\begin{aligned} q_{\mathrm{hf}} &= 2^{-\ell_{\mathrm{hf}}} \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\} = \sqrt{3}\, 2^{1-\ell_{\mathrm{hf}}} \max\{h_0(\mathcal{T}_X), h_0(\mathcal{T}_Y)\} \\ &= \sqrt{3}\, 2^{1+\ell_0-\ell_{\mathrm{hf}}} h_{\max}, \end{aligned}$$

which holds by definition of q_{hf} in (2.16) and due to the uniformity of the cluster trees, we get

$$\frac{4 h_{\max}}{d_1 - h_{t_0} - h_{s_0}} \leq \frac{\eta_2}{\sqrt{3}\kappa h_{\max}} \leq \frac{4\kappa q_{\mathrm{hf}}}{\sqrt{3}\kappa h_{\max}} = 2^{3+\ell_0-\ell_{\mathrm{hf}}}.$$

Hence, (2.31) follows from (2.33). In particular we have shown (2.30) for $j \geq 2$.

For $j = 1$ we repeat the same computations but restrict ourselves to pairs (x, y) in the set

$$\left\{ (x,y) : x - y = (d_1 + z_1, d_2 + z_2, d_3 + z_3),\ |z_n| \leq \frac{h_{t_0} + h_{s_0}}{2} \text{ for all } n \in \{1, ..., 3\} \right\},$$

in which the middlepoints (m_{t_1}, m_{s_1}) lie by construction of the uniform cluster trees. In this way we get an analogous estimate like (2.33) but with a factor 2 instead of a factor 4 in the numerator of the right hand side. Similarly as before we can conclude that (2.30) holds for $j = 1$. This completes the proof. □

Theorem 2.19. *Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees and let q_{hf} be defined by* (2.16)*. Let two sequences of boxes $\{t_j\}_{j=0}^L$ and $\{s_j\}_{j=0}^L$ in $\mathcal{T}_X$ and $\mathcal{T}_Y$, respectively, satisfy* (2.24)*. Let for all $j \in \{1, ..., L\}$ the midpoints of t_j and s_j be given by m_{t_j} and m_{s_j}. Let $\kappa > 0$ and $\ell_{\mathrm{hf}} \in \mathbb{N}_0$. Let the directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ be constructed as in Algorithm 2.6 and the sequence of directions $\{c_j\}_{j=0}^L$ be defined by* (2.25)*. Assume that t_0 and s_0 satisfy the admissibility criterion* (A3) *for an η_2 satisfying* (2.27)*. Then for all $j \in \{0, ..., L\}$ the admissibility criterion* (A2) *is satisfied for the boxes t_j, s_j, and the direction c_j, i.e.*

$$\kappa \left| \frac{m_{t_j} - m_{s_j}}{|m_{t_j} - m_{s_j}|} - c_j \right| \leq \frac{\eta_1}{\max\{\operatorname{diam}(t_j), \operatorname{diam}(s_j)\}},$$

for the admissibility constant

$$\eta_1 = 3\sqrt{2}\kappa q_{\mathrm{hf}}. \tag{2.34}$$

Proof. For $j = 0$ the assertion follows directly from Proposition 2.13. Hence we fix $j > 0$. By means of the triangle inequality we get

$$\kappa \left| \frac{m_{t_j} - m_{s_j}}{|m_{t_j} - m_{s_j}|} - c_j \right| \leq \underbrace{\kappa \left| \frac{m_{t_j} - m_{s_j}}{|m_{t_j} - m_{s_j}|} - c_{(\ell_0+j)}(t_j, s_j) \right|}_{=:S_1} + \underbrace{\kappa \left| c_{(\ell_0+j)}(t_j, s_j) - c_j \right|}_{=:S_2}.$$

From Proposition 2.13 we get

$$S_1 \leq \frac{\sqrt{2}\kappa q_{\mathrm{hf}}}{\max\{\operatorname{diam}(t_j), \operatorname{diam}(s_j)\}}.$$

With Lemma 2.18 and the relation

$$\max\{\operatorname{diam}(t_j), \operatorname{diam}(s_j)\} = q_{\mathrm{hf}} 2^{\ell_{\mathrm{hf}} - \ell_0 - j}$$

we can estimate S_2 by

$$S_2 \leq \kappa\sqrt{2}\ 2^{\ell_0+j+1-\ell_{\mathrm{hf}}} = \frac{2\sqrt{2}\kappa q_{\mathrm{hf}}}{q_{\mathrm{hf}} 2^{\ell_{\mathrm{hf}} - \ell_0 - j}} = \frac{2\sqrt{2}\kappa q_{\mathrm{hf}}}{\max\{\operatorname{diam}(t_j), \operatorname{diam}(s_j)\}}.$$

Hence, we end up with

$$\kappa \left| \frac{m_{t_j} - m_{s_j}}{|m_{t_j} - m_{s_j}|} - c_j \right| \leq \frac{3\sqrt{2}\kappa q_{\mathrm{hf}}}{\max\{\operatorname{diam}(t_j), \operatorname{diam}(s_j)\}} = \frac{\eta_1}{\max\{\operatorname{diam}(t_j), \operatorname{diam}(s_j)\}}. \qquad \square$$

Remark 2.20. Let us recall the error estimate of the directional multi-level approximation in Theorem 1.15. From the proof of Proposition 1.23 we know that the constant C in (1.34) gets smaller for decreasing η_1 in the admissibility criterion (A2). Furthermore the proof of Lemma 1.31 yields that (1.34) is applicable for smaller degrees m for

decreasing γ in (1.32) because m_0 decreases with γ. Hence by Theorems 2.16 and 2.19 we should choose κq_{hf} as small as possible, because then γ and η_1 are small by (2.23) and (2.34), respectively. For this purpose we need to choose the level ℓ_{hf}, which is used for the construction of the directions, sufficiently large. On the other hand we need that $\eta_2 \leq 4\kappa q_{\mathrm{hf}}$ for the application of Theorem 2.19. Furthermore the number of directions per level ℓ, which is given by $N_\ell = 6\ 4^{\ell_{\mathrm{hf}}-\ell}$, increases exponentially with increasing ℓ_{hf}. Hence, the storage and computational costs increase with increasing ℓ_{hf}. This means that we should choose ℓ_{hf} such that κq_{hf} is neither to large nor to small. In Chapter 3 we will run some numerical experiments and further investigate how the choice of ℓ_{hf} influences the quality and the computational costs of the directional multi-level approximation.

2.1.3 Matrix partitioning

Let the matrix A be defined by (2.1) for two sets of points P_X and P_Y. In Section 2.1.1 we described how to build box cluster trees $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$ corresponding to these sets of points. Here we want to discuss how these trees can be used to partition the matrix A into appropriate subblocks, which can then be approximated using a directional approximation as in (1.20). The principle idea is to find pairs (t, s) of boxes $t \in \mathcal{T}_X(P_X)$ and $s \in \mathcal{T}_Y(P_Y)$ that satisfy the admissibility criteria (A1) and (A3) in Section 1.2.2 and assign them a suitable direction c such that (A2) holds. For this purpose we will introduce block trees as in [5, Definition 2].

We start by defining the near- and farfield of a box in the setting of box cluster trees.

Definition 2.21 (Near- and farfield)**.** Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two box cluster trees as in Definition 2.1. Let $\eta_2 > 0$. For a box $t \in \mathcal{T}_X^{(\ell)}$ at level ℓ we define the farfield $\mathcal{F}(t)$ (with respect to $\mathcal{T}_X$, $\mathcal{T}_Y$, and η_2) by

$$\mathcal{F}(t) = \{s \in \mathcal{T}_Y^{(\ell)} : t \text{ and } s \text{ satisfy (A1) and (A3)}\}. \tag{2.35}$$

Furthermore we define the nearfield $\mathcal{N}(t)$ by

$$\mathcal{N}(t) = \mathcal{T}_Y^{(\ell)} \setminus \mathcal{F}(t). \tag{2.36}$$

Analogously we define the nearfield $\mathcal{N}(s)$ and farfield $\mathcal{F}(s)$ for a box s in $\mathcal{T}_Y^{(\ell)}$.

Remark 2.22. In regard to Definition 2.21 we note:

- For the definition of the near- and farfield we only consider boxes at the same level in their respective cluster trees. This is not necessary in general. However, it fits to the Definition 2.24 of the block trees, which we will use to partition the matrix in (1.19).

- Let $t \in \mathcal{T}_X^{(\ell)}$ and $s \in \mathcal{T}_Y^{(\ell)}$. We see that $s \in \mathcal{N}(t)$ if and only if $t \in \mathcal{N}(s)$. Likewise there holds $s \in \mathcal{F}(t)$ if and only if $t \in \mathcal{F}(s)$, i.e. the property of being in the near- or farfield of a box is symmetric.
- The admissibility criterion (A2) is not relevant for the definition of the near- or farfield of a box. The reason is that for two boxes $t \in \mathcal{T}_X^{(\ell)}$ and $s \in \mathcal{T}_Y^{(\ell)}$ with $s \in \mathcal{F}(t)$ and the set of directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ constructed in Algorithm 2.6, we can find a direction $c \in \{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ such that (A2) holds for suitable η_1 (cf. Proposition 2.13 and Theorem 2.19).

For a pair of boxes (t, s) with $t \in \mathcal{T}_X^{(\ell)}$ and $s \in \mathcal{T}_Y^{(\ell)}$ such that $s \in \mathcal{F}(t)$ we can use the directional single-level approximation $\tilde{f}_{t,s}^{(m)}$ in (1.18) to approximate f on $t \times s$. In particular we can approximate the subblock of A corresponding to the points in $t \times s$ as in (1.20). However, if $\text{parent}(s) \in \mathcal{F}(\text{parent}(t))$ then it is more efficient to approximate the larger subblock corresponding to the points in $\text{parent}(t) \times \text{parent}(s)$. This means that for all levels ℓ we are interested in finding the pairs of boxes (t, s) in $\mathcal{T}_X^{(\ell)} \times \mathcal{T}_Y^{(\ell)}$ such that $s \in \mathcal{F}(t)$ and $\text{parent}(s) \in \mathcal{N}(\text{parent}(t))$. This motivates the following definition.

Definition 2.23 (Hierarchical admissibility)**.** Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two box cluster trees and let $\ell > 0$. Let $t \in \mathcal{T}_X^{(\ell)}$ and $s \in \mathcal{T}_Y^{(\ell)}$. We say that t and s are hierarchically admissible if $s \in \mathcal{F}(t)$ and $\text{parent}(s) \in \mathcal{N}(\text{parent}(t))$.

It is our goal to find all pairs of hierarchically admissible boxes in two uniform box cluster trees $\mathcal{T}_X$ and $\mathcal{T}_Y$. A suitable structure for this purpose are block trees, which we define in the following.

Definition 2.24 (Block Trees [5, cf. Definition 2])**.** Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two box cluster trees as in Definition 2.1. A tree $\mathcal{T}_{X\times Y}$ is a block tree corresponding to $\mathcal{T}_X$ and $\mathcal{T}_Y$ if the following properties are satisfied:

- For each node $b \in \mathcal{T}_{X\times Y}$ there exist boxes $t \in \mathcal{T}_X$ and $s \in \mathcal{T}_Y$ such that $b = (t, s)$.
- The root $r \in \mathcal{T}_{X\times Y}$ is given by the pair $r = (X, Y)$.
- For each $b = (t, s) \in \mathcal{T}_{X\times Y}$ there holds

$$\text{child}(b) \neq \emptyset \implies \text{child}(b) = \text{child}(t) \times \text{child}(s). \tag{2.37}$$

The tree $\mathcal{T}_{X\times Y}$ is divided into levels by defining

$$\begin{aligned} \mathcal{T}_{X\times Y}^{(0)} &= \{(X, Y)\}, \\ \mathcal{T}_{X\times Y}^{(\ell)} &= \{(t, s) \in \mathcal{T}_{X\times Y} : \text{parent}(t, s) \in \mathcal{T}_{X\times Y}^{(\ell-1)}\}, \quad \text{for all } \ell \in \mathbb{N}. \end{aligned} \tag{2.38}$$

We denote the nodes in a block tree $\mathcal{T}_{X\times Y}$ as blocks and the set of its leaves by

$$\mathcal{L}_{X\times Y} := \{b \in \mathcal{T}_{X\times Y} : \text{child}(b) = \emptyset\}. \tag{2.39}$$

Remark 2.25. In [5, Definition 2] the nodes of block trees are not pairs of boxes, but pairs of indices. We changed the definition to be consistent, as in our setting also the nodes of the corresponding trees are not indices but boxes.
We further note that for a block $b = (t, s) \in \mathcal{T}_{X\times Y}^{(\ell)}$ there holds $t \in \mathcal{T}_X^{(\ell)}$ and $s \in \mathcal{T}_Y^{(\ell)}$ by definition, i.e. for a block at level ℓ in $\mathcal{T}_{X\times Y}$ its boxes are also at level ℓ in the corresponding box cluster trees.
Finally we observe that the blocks at level ℓ in a block tree $\mathcal{T}_{X\times Y}$ together with the leaves at levels less than ℓ form a partition of a suitable matrix A. In particular this holds true for the set of leaves $\mathcal{L}_{X\times Y}$, which is relevant for us. Let us describe a partition induced by such leaves of a block tree more in detail. For this purpose, let A be a matrix as in (2.1) for two sets of points $P_X = \{x_j\}_{j=1}^{N_X}$ and $P_Y = \{y_k\}_{k=1}^{N_Y}$. Let $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$ be two box cluster trees and $\mathcal{T}_{X\times Y}$ the corresponding block tree. Every pair (x_j, y_k) of points $x_j \in P_X$ and $y_k \in P_Y$ is included in exactly one block $(t, s) \in \mathcal{L}_{X\times Y}$. This follows immediately from (2.37) and the fact that every point $x \in P_X$ or $y \in P_Y$ is contained in exactly one leaf of $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$. Hence, a block $b = (t, s) \in \mathcal{L}_{X\times Y}$ corresponds to the subblock of A formed by the row indices in $\hat{t}$ and column indices in $\hat{s}$, where $\hat{t}$ and $\hat{s}$ are the sets of indices of points in t and s as defined in (2.3).

A block tree whose leaves yield a suitable partition for the directional approximation of the subblocks of A is constructed in Algorithm 2.26.

Algorithm 2.26. Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees and let $\eta_2 > 0$. Let the farfield $\mathcal{F}(t)$ for boxes t in $\mathcal{T}_X$ be defined by (2.35). Then we construct $\mathcal{T}_{X\times Y}$ as follows:

1. Insert $r = (X, Y)$ as root in $\mathcal{T}_{X\times Y}$.
2. If either $Y \in \mathcal{F}(X)$ or one of the trees $\mathcal{T}_X$ or $\mathcal{T}_Y$ has depth zero then we stop the construction. Else we add all pairs (t, s) such that $t \in \text{child}(X)$ and $s \in \text{child}(Y)$ as children of (X, Y) to $\mathcal{T}_{X\times Y}$.
3. Recursively we add blocks to $\mathcal{T}_{X\times Y}$ as in step 2 until for each leaf (t, s) in $\mathcal{T}_{X\times Y}$ there holds either $s \in \mathcal{F}(t)$ or $\text{child}(t) = \emptyset$ or $\text{child}(s) = \emptyset$.

We can divide the leaves of a block tree $\mathcal{T}_{X\times Y}$ constructed by means of Algorithm 2.26 into admissible and inadmissible blocks.

Definition 2.27 ([5, cf. page 8])**.** Let $\mathcal{T}_{X\times Y}$ be a block tree constructed as in Algorithm 2.26. We define the set of all admissible leaves of $\mathcal{T}_{X\times Y}$ by

$$\mathcal{L}_{X\times Y}^{+} := \{b = (t, s) \in \mathcal{L}_{X\times Y} : s \in \mathcal{F}(t)\} \tag{2.40}$$

and the set of all inadmissible leaves by

$$\mathcal{L}_{X\times Y}^{-} = \mathcal{L}_{X\times Y} \setminus \mathcal{L}_{X\times Y}^{+}. \tag{2.41}$$

Remark 2.28. Let us comment on the construction of a block tree $\mathcal{T}_{X\times Y}$ described in Algorithm 2.26:

- By construction all admissible leaves $b = (t, s) \in \mathcal{L}^+_{X\times Y}$ are such that t and s are hierarchically admissible. Indeed by construction we know that $s' \in \mathcal{N}(t')$ for $(t', s') = \text{parent}(b)$ and since $b \in \mathcal{L}^+_{X\times Y}$ there holds $s \in \mathcal{F}(t)$.
- In Algorithm 2.26 we do not specify the choice of directions for the admissible leaves $(t, s) \in \mathcal{L}^+_{X\times Y}$. However, we have seen in Proposition 2.13 that by choosing the direction $c_{(\ell)}(t, s)$ for $(t, s) \in \mathcal{T}^{(\ell)}_{X\times Y}$ we satisfy the admissibility criterion (A2) for a suitable $\eta_1 > 0$.
- The construction of the tree $\mathcal{T}_{X\times Y}$ in Algorithm 2.26 is possibly unsuitable if the roots X and Y of the uniform box cluster trees $\mathcal{T}_X$ and $\mathcal{T}_Y$ are very different in size. The reason is that the admissibility criteria (A1) and (A3) take into account only the larger of the two diameters of two boxes. To overcome this problem we can consider only pairs (t, s) of boxes t and s with comparable diameters instead of identical levels. The required adjustments of the block trees in this setting are straightforward.

Let us consider the partition of the matrix A in (2.1) induced by the block tree $\mathcal{T}_{X\times Y}$ constructed by Algorithm 2.26. The question arises if all subblocks $A|_{\hat{t}\times\hat{s}}$ of A corresponding to admissible blocks $(t, s) \in \mathcal{L}^+_{X\times Y}$ can be approximated sufficiently well by means of the directional single- and multi-level approximation. For this purpose we show in the following theorem that the assumptions of Theorem 1.15 hold for the corresponding boxes and directions.

Theorem 2.29. *Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees with depths $p(\mathcal{T}_X)$ and $p(\mathcal{T}_Y)$, respectively. Let $\ell_{\text{hf}} \in \mathbb{N}_0$ and let the directions $\{c^{(\ell)}_j\}^{N_\ell}_{j=1}$ be constructed as in Algorithm 2.6. Let $\kappa > 0$ and q_{hf} be defined by* (2.16)*. Let $\mathcal{T}_{X\times Y}$ be a block tree corresponding to $\mathcal{T}_X$ and $\mathcal{T}_Y$ constructed as in Algorithm 2.26 for an $\eta_2 \le 4\kappa q_{\text{hf}}$. Let $b = (t_0, s_0)$ be a block in $\mathcal{L}^+_{X\times Y}$ at level ℓ_0. Let $\{t_j\}^L_{j=0}$ and $\{s_j\}^L_{j=0}$ be two sequences in $\mathcal{T}_X$ and $\mathcal{T}_Y$, respectively, such that*

$$t_j \in \text{child}(t_{j-1}), \quad s_j \in \text{child}(s_{j-1}), \quad \text{for all } j \in 1, ..., L$$

and the directions $\{c_j\}^L_{j=0}$ be given by $c_j = c_{(\ell_0+j)}(t_0, s_0)$ as defined in (2.13)*. Let $\eta_1 = 3\sqrt{2}\kappa q_{\text{hf}}$ and let $\hat{\rho} = 1+2/\eta_2$ as in* (1.33)*. Then the assumptions of Theorem 1.15 on $\{t_j\}^L_{j=0}$, $\{s_j\}^L_{j=0}$ and $\{c_j\}^L_{j=0}$ are satisfied. In particular there exists a constant C depending only on η_2, $\max\{p(\mathcal{T}_X, p(\mathcal{T}_Y)\}$, and the product κq_{hf}, and an $m_0 \in \mathbb{N}$ depending only on κq_{hf}, such that for all $m \ge m_0$ the estimate* (1.34) *on the directional multi-level approximation error holds, i.e.*

$$\|f - \hat{f}^{(m)}_{t_0,s_0}\|_{\infty,t_L\times s_L} \le \frac{C}{\text{dist}\,(t_0, s_0)}\hat{\rho}^{-(m+1)}.$$

Proof. We start by showing (1.30) and (1.31). By construction all boxes at a given level ℓ in a general uniform box cluster tree $\mathcal{T}$ are cubes which are identical up to translation. Furthermore the half edge length $h_\ell(\mathcal{T})$ of these boxes satisfies

$$h_{\ell-1}(\mathcal{T}) = \frac{1}{2} h_\ell(\mathcal{T}).$$

Hence we immediately get (1.30) with $\bar{q} = 1/2$ for the edge lengths of the boxes in $\{t_j\}_{j=0}^L$ and $\{s_j\}_{j=0}^L$. Furthermore the diameters $q_\ell(\mathcal{T})$ of boxes at level ℓ of a uniform box cluster tree satisfy

$$q_\ell(\mathcal{T}) = 2\sqrt{3} h_\ell.$$

Therefore we have

$$s(t_j) = \frac{1}{\sqrt{3}} \operatorname{diam}(t_j), \qquad s(s_j) = \frac{1}{\sqrt{3}} \operatorname{diam}(s_j),$$

where $s(t_j)$ and $s(s_j)$ denote the length of the longest edge of t_j and s_j, respectively. In particular there holds (1.31) with $\sigma = 1/\sqrt{3}$.

Next we observe that $s_0 \in \mathcal{F}(t_0)$ due to $(t_0, s_0) \in \mathcal{L}^+_{X\times Y}$, which means that t_0 and s_0 satisfy (A1) and (A3). Furthermore the criterion (A2) holds for $\eta_1 = 3\sqrt{2}\kappa q_{\mathrm{hf}}$, the boxes t_j, s_j, and the direction c_j for all $j \in \{0, ..., L\}$ due to Theorem 2.19. Moreover we have (1.32) for $\gamma = \sqrt{2}\kappa q_{\mathrm{hf}}$ due to Theorem 2.16. Hence all assumptions of Theorem 1.15 are satisfied and (1.34) follows.

Let us finally check the dependencies of C and m_0. By Theorem 1.15 C depends on η_1, η_2, $\bar{q}$, σ and L. In our setting η_1 depends only on the product κq_{hf}. Furthermore $\bar{q}$ and σ are constants independent of the particular trees, and L is bounded by the maximal depth $\max\{p(T_X), p(T_Y)\}$. Hence C depends only on η_2, $\max\{p(T_X), p(T_Y)\}$, and the product κq_{hf}. Similarly m_0 depends only on the product κq_{hf} as γ does so too. □

2.1.4 An efficient matrix-vector multiplication

With the block trees constructed in Algorithm 2.26 we can define an efficient matrix-vector multiplication for the matrix A in (2.1). We focus first on the directional single-level approximation of the admissible subblocks of A in Algorithm 2.35. Then we discuss the necessary changes if a multi-level approximation is used instead in Algorithm 2.38. Both ideas can be found in [5, Section 4] and [30, Section 3.3.5].

We start by clarifying our notation.

Definition 2.30. Let the matrix A be defined as in (2.1) for two sets of points P_X and P_Y. Let $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$ be the corresponding uniform box cluster trees and the block tree $\mathcal{T}_{X\times Y}$ be constructed as in Algorithm 2.26. For a block $(t, s) \in \mathcal{L}^+_{X\times Y}$ denote the set of indices $\hat{t}$ and $\hat{s}$ by (2.3). Then we denote the subblock of A corresponding to the row indices in $\hat{t}$ and the column indices in $\hat{s}$ by $A|_{\hat{t}\times\hat{s}}$.

For a vector $v \in \mathbb{C}^N$ and a set of indices $\hat{s} \subset \{1, ..., N\}$ we further denote the subvector of v corresponding to the indices in $\hat{s}$ by $v|_{\hat{s}}$.

Remark 2.31. If the points in the respective box cluster trees $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$ are reordered, then the subblocks $A|_{\hat{t}\times\hat{s}}$ of a matrix A are subblocks in the classical sense, i.e. they consist of consecutive rows and columns. While in theory this reordering has no effect, in implementations it can cause speedups due to better memory alignment.

For a block $(t, s) \in \mathcal{L}^+_{X\times Y}$ at level ℓ and the direction $c = c_{(\ell)}(t, s)$ in (2.13) we approximate the subblock $A|_{\hat{t}\times\hat{s}}$ as in (1.20) by

$$A|_{\hat{t}\times\hat{s}} = L_{t,c}A_{c,t,s}L^*_{s,c}.$$

For blocks $(t, s) \in \mathcal{L}^-_{X\times Y}$ on the other hand we cannot approximate $A|_{\hat{t}\times\hat{s}}$. Hence it is reasonable to split the multiplication $g = Av$ into two parts.

Definition 2.32. In the setting of Definition 2.30 let $v \in \mathbb{C}^{N_Y}$ and $g = Av$. Define the matrix of admissible subblocks A^+ by

$$A^+|_{\hat{t}\times\hat{s}} := \begin{cases} A|_{\hat{t}\times\hat{s}}, & \text{for } (t, s) \in \mathcal{L}^+_{X\times Y}, \\ 0, & \text{for } (t, s) \in \mathcal{L}^-_{X\times Y}, \end{cases} \tag{2.42}$$

and the matrix of inadmissible subblocks A^- by

$$A^-|_{\hat{t}\times\hat{s}} := \begin{cases} 0, & \text{for } (t, s) \in \mathcal{L}^+_{X\times Y}, \\ A|_{\hat{t}\times\hat{s}}, & \text{for } (t, s) \in \mathcal{L}^-_{X\times Y}, \end{cases} \tag{2.43}$$

Then we define the nearfield part g^- and the farfield part g^+ of g by

$$g^+ := A^+g, \qquad g^- := A^-g. \tag{2.44}$$

The computation of the nearfield part g^- is straightforward.

Algorithm 2.33 (Computation of g^-)**.** Let A be a matrix as in (1.19) and $\mathcal{T}_{X\times Y}$ a corresponding block tree as in Definition 2.30. Let $\mathcal{L}^-_{X\times Y}$ be defined as in (2.41). Let $v \in \mathbb{C}^{N_Y}$ and the nearfield part g^- of a vector $g = Av$ be defined by (2.44). Then for all $j \in \{1, ..., N_X\}$ there holds componentwise

$$g^-_j = \sum_{\substack{(t,s)\in\mathcal{L}^-_{X\times Y} \\ j\in\hat{t}}} \left(A|_{\hat{t}\times\hat{s}}v|_{\hat{s}}\right)_j. \tag{2.45}$$

The farfield part g^+ is not computed exactly but approximated via the directional approximation (1.20) of the admissible subblocks or its multi-level counterpart. In Algorithm 2.35 we describe the single-level approximation of g^+ as presented in [5, Section 4]. For this purpose we use the following definition.

Definition 2.34. Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees and $\mathcal{T}_{X\times Y}$ the corresponding block tree constructed by Algorithm 2.26. Let $\ell_{\mathrm{hf}} \in \mathbb{N}_0$ and the directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ for all $\ell \in \mathbb{N}_0$ be constructed by Algorithm 2.6. Furthermore let the direction $c_{(\ell)}(t,s)$ for two boxes t and s be defined by (2.13). Then we define for all $\ell \in \mathbb{N}_0$ and all $t \in \mathcal{T}_X^{(\ell)}$ the set of active directions

$$D_t := \{c \in \{c_j^{(\ell)}\}_{j=1}^{N_\ell} : \exists\, s \in \mathcal{T}_Y^{(\ell)} \text{ such that } (t,s) \in \mathcal{L}_{X\times Y}^{+} \text{ and } c = c_{(\ell)}(t,s)\}. \quad (2.46)$$

Analogously, we define the set of directions D_s for $s \in \mathcal{T}_Y^{(\ell)}$.

Algorithm 2.35 (Directional single-level approximation of g^+ [5, cf. Section 4])**.** Let A be a matrix as in (1.19) and $\mathcal{T}_{X\times Y}$ a corresponding block tree as in Definition 2.30. Let $\mathcal{L}_{X\times Y}$ be defined as in (2.41). Let $v \in \mathbb{C}^{N_Y}$ and the farfield part g^+ of a vector $g = Av$ be defined by (2.44). Choose the interpolation degree $m \in \mathbb{N}$. Let the coupling matrices $A_{c,t,s}$ for two boxes t and s and a direction c be defined by (1.21). Let the directional interpolation matrices $L_{t,c}$ and $L_{s,c}$ be defined by (1.22). Let the sets of directions D_t and D_s be defined by (2.46). Then we compute the directional single-level approximation $\tilde{g}^+$ of g^+ in three steps:

1. Forward transformation (S2M)
   For all $s \in \mathcal{T}_Y$ and all active directions $c \in D_t$ we compute the coefficient-vectors
   $$\tilde{v}_{sc} := L_{s,c}^*(v|_{\hat{s}}).$$
2. Coupling step (M2L)
   For all $t \in \mathcal{T}_X^{(\ell)}$ and directions $c \in D_t$ we compute the coefficient-vectors
   $$\tilde{g}_{tc} := \sum_{\substack{(t,s)\in\mathcal{L}_{X\times Y}^{+} \\ c=c_{(\ell)}(t,s)}} A_{c,t,s}\tilde{v}_{sc}.$$
3. Backward transformation (L2T)
   For all $j \in \{1,...,N_X\}$ we compute componentwise
   $$\tilde{g}_j^+ = \sum_{\substack{t\in\mathcal{T}_X, c\in D_t \\ j\in\hat{t}}} \left(L_{t,c}\tilde{g}_{tc}\right)_j.$$

Remark 2.36. In Algorithm 2.35 we added the abbreviations S2M, M2L and L2T in parentheses to the three steps. These abbreviations originate from the fast multipole method originally presented in [20, 22, 34]. The basic idea of this fast summation scheme is the same that we exploit: Suitable expansions are used to express the effect of the kernel function of a certain region s on another distant region t. To emphasize the similarities of the computation steps in our algorithm we use the same abbreviations

for these steps, even though we are using interpolation schemes instead of multipole and local expansions, which are used in the classical fast multipole method and from which the names are taken from. Similarly we will use the abbreviations M2M and L2L for steps two and four of Algorithm 2.38.

As discussed in Section 1.2.1 the approximation of a subblock $A|_{\hat{t}\times\hat{s}}$ and, hence, the approximation of the farfield part g^+ becomes really efficient only if we do not compute the directional interpolation matrices $L_{t,c}$ and $L_{s,c}$ directly. Instead, we have to approximate $L_{t,c}$ with the transfer matrices $E_{t',c}$ and the matrices $L_{t',c'}$ for $t' \in \text{child}(t)$ as in (1.24). If the box t has eight sons t'_1, ..., t'_8 and the points in t are ordered according to the numeration of its sons this approximation has the form

$$L_{t,c} = \begin{bmatrix} L_{t'_1,c'} E_{t'_1,c} \\ \dots \\ L_{t'_8,c'} E_{t'_8,c} \end{bmatrix}.$$

In Section 2.1.2 we discussed a suitable choice of c'. Let us recall this here.

We have seen in Theorems 2.16 and 2.19 that for two sequences of boxes $\{t_j\}_{j=0}^L$ and $\{s_j\}_{j=0}^L$ satisfying (2.24) the directions $\{c_j\}_{j=0}^L$ with $c_j = c_{(\ell_0+j)}(t_0, s_0)$ as defined in (2.13) are a reasonable choice. In the proof of Corollary 2.17 we have shown that $c_j = \text{dir}_{(\ell_0+j)}(c_{j-1})$ holds for all $j \in \{1, ..., L\}$, where the mappings $\text{dir}_{(\ell)}$ are given in Definition 2.9. Hence, we see that for a box t at level ℓ we should choose the direction $c' = \text{dir}_{(\ell+1)}(c)$ for the approximation of $L_{t,c}$ as in (1.24).

To keep track of all directions which we need to consider for given boxes t and s we introduce the sets $\hat{D}_t$ and $\hat{D}_s$ of inherited directions.

Definition 2.37. Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees and $\mathcal{T}_{X\times Y}$ the corresponding block tree. Let $\ell_{\text{hf}} \in \mathbb{N}_0$ and the directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ for all $\ell \in \mathbb{N}_0$ be constructed by Algorithm 2.6. Furthermore let the mappings $\text{dir}_{(\ell)}$ for all $\ell \in \mathbb{N}_0$ be given by Definition 2.9. Let the set D_t for a box $t \in \mathcal{T}_X$ be defined by (2.46) and denote by $\text{anc}(t)$ the set of all the ancestors of t in $\mathcal{T}_X$. Then we define for all $\ell \in \mathbb{N}$ and all $t \in \mathcal{T}_X^{(\ell)}$ the set of inherited directions $\hat{D}_t$ by

$$\hat{D}_t = \{\hat{c} \in \{c_j^{(\ell)}\}_{j=1}^{N_\ell} : \exists\, t' \in \text{anc}(t), \exists\, c \in D_{t'} \text{ such that } \hat{c} = \text{dir}_{(\ell)}(c)\}. \tag{2.47}$$

Analogously, we define the set of inherited directions $\hat{D}_s$ for $s \in \mathcal{T}_Y^{(\ell)}$.

Finally we are ready to describe the computation of the directional multi-level approximation of the farfield part g^+.

Algorithm 2.38 (Directional multi-level approximation of g^+ [5, cf. Section 4])**.** Let A be a matrix as in (1.19) and $\mathcal{T}_{X\times Y}$ a corresponding block tree as in Definition 2.30. Let $\mathcal{L}_{X\times Y}$ be defined as in (2.41). Let $v \in \mathbb{C}^{N_Y}$ and the farfield part g^+ of a vector $g = Av$ be defined by (2.44). Choose the interpolation degree $m \in \mathbb{N}$. Let the

coupling matrices $A_{c,t,s}$ for two boxes t and s and a direction c be defined by (1.21), the directional interpolation matrices $L_{t,c}$ and $L_{s,c}$ by (1.22), and the transfer matrices $E_{t,c}$ and $E_{s,c}$ by (1.25). Let the sets of directions D_t and D_s be defined by (2.46) and the sets of inherited directions $\hat{D}_t$ and $\hat{D}_s$ by (2.47). Then we compute the directional multi-level approximation $\hat{g}^+$ of g^+ in five steps:

1. Direct forward transformation (S2M)
   For all leaves $s \in \mathcal{L}_Y$ and all active or inherited directions $c \in \hat{D}_s \cup D_s$ we compute the coefficient-vectors
   $$\hat{v}_{sc} := L_{s,c}^*(v|_{\hat{s}}).$$

2. Forward transfer (M2M)
   Starting at level $\ell = p(\mathcal{T}_Y) - 1$, where $p(\mathcal{T}_Y)$ is the depth of $\mathcal{T}_Y$, we compute for decreasing ℓ for all boxes $s \in \mathcal{T}_Y^{(\ell)} \backslash \mathcal{L}_Y$ and directions $c \in \hat{D}_s \cup D_s$ the coefficient-vectors
   $$\hat{v}_{sc} := \sum_{s' \in \mathrm{child}(s)} E_{s',c}^* \hat{v}_{s'c'}, \quad \text{where } c' = \mathrm{dir}_{(\ell+1)}(c).$$

3. Coupling step (M2L)
   For all $t \in \mathcal{T}_X^{(\ell)}$ and directions $c \in D_t$ we compute the coefficient-vectors
   $$\tilde{g}_{tc} := \sum_{\substack{(t,s) \in \mathcal{L}_{X \times Y}^+ \\ c = c_{(\ell)}(t,s)}} A_{c,t,s} \hat{v}_{sc}.$$
   Further we set $\tilde{g}_{tc} := 0$ for all $t \in \mathcal{T}_X^{(\ell)}$ and directions $c \in \hat{D}_t \backslash D_t$.

4. Backward transfer (L2L)
   Starting at level $\ell = 0$ we compute for incremental ℓ for all boxes $t' \in \mathcal{T}_X^{(\ell)}$ and directions $c' \in \hat{D}_{t'} \cup D_{t'}$ the coefficent-vectors
   $$\hat{g}_{t'c'} = \tilde{g}_{t'c'} + \sum_{\substack{c \in \hat{D}_t \cup D_t \\ \mathrm{dir}_{(\ell)}(c) = c'}} E_{t',c} \hat{g}_{tc}, \quad \text{where } t = \mathrm{parent}(t').$$

5. Direct backward transformation (L2T)
   For all leaves $t \in \mathcal{L}_X$ such that $D_t \cup \hat{D}_t \neq \emptyset$ we compute
   $$\hat{g}^+|_{\hat{t}} = \sum_{c \in \hat{D}(t) \cup D_t} L_{t,c} \hat{g}_{tc}.$$
   For the remaining leaves we set
   $$\hat{g}^+|_{\hat{t}} = 0.$$

Remark 2.39. We observe that in Algorithm 2.38 there is no explicit transition from the high frequency regime, i.e. levels $\ell \leq \ell_{\mathrm{hf}}$, to the low frequency regime, i.e. levels $\ell > \ell_{\mathrm{hf}}$. Indeed we use the same matrix operators for all levels. For levels $\ell > \ell_{\mathrm{hf}}$, however, the directions used for the computation of the directional interpolation matrices $L_{t,c}$ and the transfer matrices $E_{t',c}$ are zero. This means that we use standard interpolation and standard reinterpolation for such levels. This is the implicit transition between the two frequency regimes in Algorithm 2.38.

Example 2.40. To clarify the steps in Algorithm 2.38 we examplarily compute the matrix-vector product $g = A|_{\hat{t}\times\hat{s}} v_s$ of an admissible subblock $A|_{\hat{t}\times\hat{s}} \in \mathbb{C}^{N_t \times N_s}$ and a corresponding vector $v \in \mathbb{C}^{N_s}$. For this purpose we use a directional two-level approximation of $A|_{\hat{t}\times\hat{s}} \in \mathbb{C}^{N_t \times N_s}$. For the children s'_1, ..., s'_8, and t'_1, ..., t'_8 of s and t, respectively, the directions c, and the inherited direction c' this approximation can be written in the form

$$A|_{\hat{t}\times\hat{s}} \approx L_{t,c} A_{c,t,s} L^*_{s,c} \approx \begin{bmatrix} L_{t'_1,c'} E_{t'_1,c} \\ \dots \\ L_{t'_8,c'} E_{t'_8,c} \end{bmatrix} A_{c,t,s} \begin{bmatrix} E^*_{s'_1,c} L^*_{s'_1,c'} & \dots & E^*_{s'_8,c} L^*_{s'_8,c'} \end{bmatrix} \tag{2.48}$$

The matrix vector product of this two-level approximation (2.48) and the vector v is computed in the following five steps:

1. For all boxes s'_1, ..., s'_8 and the direction c' we compute the coefficient-vectors

$$\hat{v}_{s'_j c'} = L^*_{s'_j,c'}(v|_{\hat{s}'_j}) \quad j \in \{1, ..., 8\}.$$

2. We compute the coefficient-vector $\hat{v}_{sc}$ by

$$\hat{v}_{sc} = \begin{bmatrix} E^*_{s'_1,c} L^*_{s'_1,c'} & \dots & E^*_{s'_8,c} L^*_{s'_8,c'} \end{bmatrix} v = \sum_{j=1}^{8} E_{s'_j,c} L^*_{s'_j,c'}(v|_{\hat{s}'_j}) = \sum_{j=1}^{8} E_{s'_j,c} \hat{v}_{s'_j c'}.$$

3. Next we compute the coefficient-vector

$$\tilde{g}_{tc} = A_{c,t,s} \hat{v}_{sc}$$

and set $\tilde{g}_{t'_j c'} = 0$ for all $j \in \{1, ..., 8\}$.

4. We set $\hat{g}_{tc} = \tilde{g}_{tc}$ and compute for all boxes t'_1, ..., t'_8 the coefficient-vectors

$$\hat{g}_{t'_j c'} = \underbrace{\tilde{g}_{t'_j c'}}_{=0} + E_{t'_j,c} \hat{g}_{tc}, \quad j \in \{1, ..., 8\}.$$

5. Finally we compute the approximation $\hat{g}$ of g blockwise by

$$\hat{g}|_{\hat{t}'_j} = L_{t'_j,c'} \hat{g}_{t'_j c'}.$$

The five steps of this computation correspond to the five steps in Algorithm 2.38. The few differences are easily explained:

- In the M2L step of Algorithm 2.38 we collect in the coefficient vector $\tilde{g}_{tc}$ all contributions from boxes s, such that t, s and c satisfy the three admissibility criteria (A1) – (A3) from Section 1.2.2. In our example we only consider one such box, hence we need no sum.
- In the L2L step the coefficient-vector $g_{t'c'}$ is computed using two parts: the coefficient-vector $\tilde{g}_{t'c'}$, which is formed by direct contributions in the M2L step, and the inherited part, which is the sum in step four of Algorithm 2.38. This inherited part is what we compute in step four of our example. However, we only have one active direction c for the box t in our example and, therefore, need no summation over all directions.

Remark 2.41. The directional multi-level approximation of g^+ in Algorithm 2.38 relies on the directional approximation of the Helmholtz kernel f in (1.1) on sets $t_0 \times s_0$ such that $(t_0, s_0) \in \mathcal{L}^+_{X\times Y}$. Due to the forward transfer and the backward transfer we end up with approximations of the form

$$\left(\mathcal{I}^{(m)}_{t_{L_1},c_{L_1}} \circ \ldots \circ \mathcal{I}^{(m)}_{t_0,c_0}\right) \otimes \left(\mathcal{I}^{(m)}_{s_{L_2},-c_{L_2}} \circ \ldots \circ \mathcal{I}^{(m)}_{s_0,-c_0}\right)[f], \tag{2.49}$$

where the interpolation operators of the form $\mathcal{I}^{(m)}_{t,c}$ are defined in (1.16) and L_1 and L_2 are the lengths of the sequences of boxes $\{t_j\}_{j=0}^{L_1}$ and $\{s_j\}_{j=0}^{L_2}$ such that t_{L_1} and s_{L_2} are leaves in the respective box cluster trees. In case that $L_1 = L_2$ we can estimate the approximation error by Theorem 1.15. However, L_1 and L_2 in (2.49) are not necessarily equal. Let without loss of generality $L_1 \leq L_2$. Then there holds

$$\begin{aligned}\left(\mathcal{I}^{(m)}_{t_{L_1},c_{L_1}} \circ \ldots \circ \mathcal{I}^{(m)}_{t_0,c_0}\right) \otimes \left(\mathcal{I}^{(m)}_{s_{L_2},c_{L_2}} \circ \ldots \circ \mathcal{I}^{(m)}_{s_0,c_0}\right)[f] \\ = \left(I \otimes \mathcal{I}^{(m)}_{t_{L_2},c_{L_2}} \circ \ldots \circ \mathcal{I}^{(m)}_{t_{L_1+1},c_{L_1+1}}\right)\left[\mathcal{I}^{(m)}_{t_{L_1}\times s_{L_1},c_{L_1}} \circ \ldots \circ \mathcal{I}^{(m)}_{t_0\times s_0,c_0}[f]\right],\end{aligned}$$

where the operators of the form $\mathcal{I}^{(m)}_{t\times s,c}$ are defined in (1.17). If $c_j = 0$ for all $j \geq L_1$ then the operator $I \otimes \mathcal{I}^{(m)}_{t_{L_2},c_{L_2}} \circ \ldots \circ \mathcal{I}^{(m)}_{t_{L_1+1},c_{L_1+1}}$ induces no further error as the interpolation operators are non-directional, i.e. standard interpolation operators, and the function $\mathcal{I}^{(m)}_{t_{L_1}\times s_{L_1},c_{L_1}} \circ \ldots \circ \mathcal{I}^{(m)}_{t_0\times s_0,c_0}[f]$ is a polynomial in both its arguments. For general directions this is no longer true. In this case one has to estimate the error when re-interpolating a directional Lagrange polynomial. In [7, Corollary 5.8] it is shown that this error decays exponentially for sufficiently large m, i.e. exponential convergence of the approximation error is maintained.

By using Algorithm 2.33 for the computation of g^- and Algorithm 2.38 for the computation of $\hat{g}^+$ we can now efficiently approximate the matrix-vector product $g = Av$ for a matrix A as in (2.1) and a vector $v \in \mathbb{C}^{N_Y}$. In Section 2.2.2 we will discuss

the complexity of this approach. Here we recapitulate all the necessary steps for the approximation.

Algorithm 2.42 (Fast directional matrix-vector multiplication)**.** Let A be a matrix as defined in (2.1) for two sets of points $P_X = \{x_j\}_{j=1}^{N_X}$ and $P_Y = \{y_k\}_{k=1}^{N_Y}$. Let $v \in \mathbb{C}^{N_Y}$. Choose a separation constant η_2, the highest level in the high frequency regime ℓ_{hf}, and the interpolation degree $m \in \mathbb{N}$. Then an approximation $\hat{g} = g^- + \hat{g}^+$ of the matrix-vector product $g = Av$ is computed as follows:

1. Construct uniform box cluster trees $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$ by Algorithm 2.3.
2. Construct the sets of directions $\{c_j\}_{j=1}^{N_\ell}$ for the chosen ℓ_{hf} by Algorithm 2.6.
3. Construct the block tree $\mathcal{T}_{X\times Y}$ corresponding to $\mathcal{T}_X(P_X)$, $\mathcal{T}_Y(P_Y)$ and the separation paramter η_2 by Algorithm 2.26.
4. Compute the nearfield part g^- by Algorithm 2.33.
5. Compute the directional multi-level approximation $\hat{g}^+$ of the farfield part g^+ by Algorithm 2.38 and set $\hat{g} = g^- + \hat{g}^+$.

Remark 2.43. In Algorithm 2.42 we compute the approximation $\hat{g} = g^- + \hat{g}^+$ of the matrix vector product $g = g^- + g^+ = Av$ for the matrix $A \in \mathbb{C}^{N_X \times N_Y}$ in (2.1) and any vector $v \in \mathbb{C}^{N_Y}$, where g^+ and g^- are defined in (2.44). In applications we would like to control the error

$$\|g - \hat{g}\|_2 = \|\hat{g}^+ - g^+\|_2 = \|\hat{A}^+ v - A^+ v\|_2 \leq \|\hat{A}^+ - A^+\| \|v\|_2,$$

where A^+ is the matrix of admissible subblocks in (2.42) and $\hat{A}^+$ the approximation of A^+ which we get when we approximate each of its subblocks using the directional multi-level approximation. Hence we need to estimate the difference $\hat{A}^+ - A^+$ in a suitable norm, for example, the Frobenius norm $\|\cdot\|_F$ or the spectral norm $\|\cdot\|_2$. For this purpose, we observe, that the non-zero blocks of $A+$ are generated by the Helmholtz kernel f defined in (1.1) while in $\hat{A}^+$ they are generated by a (possibly asymmetric) multi-level approximation of f as in (2.49). We have discussed in Remark 2.41 that the approximation error of the directional multi-level approximation of f is exponentially converging to zero with increasing interpolation degree m, which follows from (1.34) in Theorem 1.15 in case of the standard directional multi-level approximation. These estimates on the generating functions suffice to show that the approximation errors $\|A^+ - \hat{A}^+\|_F$ and $\|A^+ - \hat{A}^+\|_2$ converge exponentially to zero with increasing m. Details can be found in [4, Section 4.6].

2.2 Implementation details and complexity discussion

The fast directional matrix-vector multiplication presented in Algorithm 2.42 is based on the clustering of the points corresponding to the matrix A in (2.1). In Section 2.1 we have focused on a uniform clustering strategy. We have already pointed out in Remark 2.4 that we can exploit this uniformity in the implementation. We will discuss this in the following section. Furthermore we talk about a possible recompression of the occurring coupling matrices. Finally we will analyze the complexity of the fast matrix-vector multiplication in Section 2.2.2.

2.2.1 Implementation details

In this section we do not give a complete description of an implementation of the fast directional matrix-vector multiplication described in Algorithm 2.42. Instead we focus on some details which allow us to reduce the computational costs. In particular we describe how to exploit the uniformity of the box cluster trees for the storage and application of the coupling matrices $A_{c,t,s}$ defined in (1.21) and the transfer matrices $E_{t',c}$ and $E_{s',c}$ defined in (1.25). Throughout this section we let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees as constructed in Algorithm 2.3 and $\mathcal{T}_{X\times Y}$ the corresponding block tree as constructed in Algorithm 2.26. We let $\ell_{\mathrm{hf}} \in \mathbb{N}_0$ be the highest level in the high frequency regime. Furthermore we let the directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ be constructed as in Algorithm 2.6 and the mappings $\mathrm{dir}_{(\ell)}$ be given as in Definition 2.9.

We start by considering the transfer matrices defined in (1.25). It suffices to focus on the cluster tree $\mathcal{T}_X$ as the properties which we show hold for general uniform box cluster trees. Let $t \in \mathcal{T}_X^{(\ell)}$ be a non-leaf box at level ℓ and $t' \in \mathrm{child}(t)$. The transfer matrix $E_{t',c}$ has the entries

$$E_{t',c}[j,k] = \exp(i\kappa\langle \xi_{t',\nu_j}, c - c'\rangle) L_{t,\nu_k}^{(m)}(\xi_{t',\nu_j}), \quad j,k \in \{1,...,(m+1)^3\},$$

where $c' = \mathrm{dir}_{(\ell+1)}(c)$, $\{\xi_{t',\nu_j}\}_{\nu_j \in M}$ are the interpolation points in t' and $\{L_{t,\nu_k}^{(m)}\}_{\nu_k \in M}$ are the Lagrange polynomials corresponding to the interpolation points in t for the set of multi-indices $M = \{0,...,m\}^3$. We can split this matrix into a directional and a non-directional part by setting

$$E_{t',c} = E_{t',c}^{\mathrm{d}} E_{t'}, \tag{2.50}$$

where we define the directional part by

$$E_{t',c}^{\mathrm{d}} := \mathrm{diag}\left(\{\exp(i\kappa\langle \xi_{t',\tilde{\nu}}, c - c'\rangle)\}_{\tilde{\nu}\in M}\right) \tag{2.51}$$

and the non-directional part by

$$E_{t'}[j,k] := L_{t,\nu_k}^{(m)}(\xi_{t',\nu_j}), \quad j,k \in \{1,...,(m+1)^3\}. \tag{2.52}$$

In particular we can split the multiplication $g = E_{t',c}v$ for a vector $v \in \mathbb{C}^{(m+1)^3}$ into

$$g = E^{\mathrm{d}}_{t',c} g_{\mathrm{nd}}, \qquad g_{\mathrm{nd}} = E_{t'} v.$$

Let us consider the directional part $E^{\mathrm{d}}_{t',c}$ of the transfer matrix $E_{t',c}$. We distinguish two cases. If $\ell \in \{0, ..., \ell_{\mathrm{hf}}\}$, i.e. ℓ is a level in the high frequency regime, then for each box $t \in \mathcal{T}_X^{(\ell)}$ we have to compute a matrix $E^{\mathrm{d}}_{t',c}$ for all children t' of t and all directions c in the set of active and inherited directions D_t and $\hat{D}_t$ given in Definitions 2.34 and 2.37. As these matrices are diagonal, we only have to compute $(m+1)^3$ entries instead of $(m+1)^6$, which are needed for the transfer matrices $E_{t',c}$. If $\ell > \ell_{\mathrm{hf}}$ we see that $E^{\mathrm{d}}_{t',c}$ corresponds to the identity matrix because in this case $c = c' = 0$. Hence, for such levels no directional part $E^{\mathrm{d}}_{t',c}$ of a transfer matrix $E_{t',c}$ has to be computed.

Let us consider the non-directional part $E_{t'}$ defined in (2.52). We observe that the value of the Lagrange polynomial $L^{(m)}_{t,\nu}$ depends only on the position of the evaluation point x relative to the box t. If t and x are translated by the same vector $w \in \mathbb{R}^3$ then the value does not change. Hence, we have

$$L^{(m)}_{t_1,\nu}(\xi_{t'_1,\tilde{\nu}}) = L^{(m)}_{t_2,\nu}(\xi_{t'_2,\tilde{\nu}}), \quad \text{for all } \nu, \tilde{\nu} \in M \tag{2.53}$$

for two arbitrary boxes $t_1, t_2 \in \mathcal{T}_X^{(\ell)}$ and their children $t'_1 \in \mathrm{child}(t_1)$, $t'_2 \in \mathrm{child}(t_2)$ if t'_1 and t'_2 have the same position relative to t_1 and t_2, respectively. In particular, for such boxes there holds $E_{t'_1} = E_{t'_2}$, i.e. the non-directional parts of the transfer matrices coincide. The boxes in $\mathcal{T}_X$ have this translation invariance property due to the uniformity of $\mathcal{T}_X$. In fact, each box $t \in \mathcal{T}_X$ has at most 8 children whose position relative to t is fixed. Hence, we only have to compute and store at most eight non-directional transfer matrices E'_t at each level $\ell \in \{0, ..., p(\mathcal{T}_X) - 1\}$ of the box cluster tree $\mathcal{T}_X$, where $p(\mathcal{T}_X)$ defines the depth of $\mathcal{T}_X$.

Next we consider the coupling matrices $A_{c,t,s}$ defined in (1.21). Their entries are given by

$$A_{c,t,s}[j,k] = f_c(\xi_{t,\nu_j}, \xi_{s,\nu_k}) = \frac{\exp(i\kappa(|\xi_{t,\nu_j} - \xi_{s,\nu_k}| - \langle \xi_{t,\nu_j} - \xi_{s,\nu_k}, c\rangle))}{4\pi|\xi_{t,\nu_j} - \xi_{s,\nu_k}|},$$

where $j, k \in \{1, ..., (m+1)^3\}$. The generating function f_c of these matrices depends only on the difference of its arguments. Hence, f_c is invariant under a translation of both its arguments by the same vector. In particular, if two pairs of boxes (t_1, s_1) and (t_2, s_2) are identical up to translation we get

$$A_{c,t_1,s_1} = A_{c,t_2,s_2}. \tag{2.54}$$

We observe that two admissible leaf-blocks $(t_1, s_1), (t_2, s_2) \in \mathcal{L}^+_{X\times Y}$ at the same level ℓ of the block tree are identical up to translation if the difference of their midpoints $m_{t_1} - m_{s_1}$ and $m_{t_2} - m_{s_2}$ is the same. Furthermore the direction used to compute

the coupling matrix is $c = c_{(\ell)}(t_1, s_1) = c_{(\ell)}(t_2, s_2)$ for such blocks by definition of the directions $c_{(\ell)}(t, s)$ in (2.13). Hence, the coupling matrices corresponding to (t_1, s_1) and (t_2, s_2) satisfy (2.54). In particular, it suffices to compute and store all required coupling matrices for all levels ℓ once for a reference configuration and assign them to the appropriate blocks $(t, s) \in \mathcal{L}^+_{X\times Y}$.

Another aspect which improves the performance of an implementation further is a compression of the coupling matrices (cf. [6, 30]). The problem is that the number of rows and columns of these matrices, which is $(m + 1)^3$, increases cubically in the interpolation degree m. Hence, we would like to approximate them further. For this purpose we would like to find a low rank approximation of $A_{c,t,s}$, i.e. we would like to find a $k \in \mathbb{N}$ and two matrices $U_{c,t,s}$, $V_{c,t,s} \in \mathbb{C}^{(m+1)^3\times k}$ such that

$$A_{c,t,s} \approx U_{c,t,s} V^*_{c,t,s}. \tag{2.55}$$

We assume that such an approximation exists, because the coupling matrices are generated by functions which are reasonably smooth by construction. This allows to reduce the effort for the storage and application of the matrix $A_{c,t,s}$ from $\mathcal{O}((m+1)^6)$ to $\mathcal{O}(k(m+1)^3)$.

There are several possibilities to construct a low rank approximation of a given matrix $A \in \mathbb{C}^{n\times n}$. The best low rank approximation for a given rank k can be found by computing a singular value decomposition (SVD) of A and truncating the resulting matrices appropriately, i.e by taking only the k largest singular values. However, the computation of a SVD for a general matrix is quite expensive. Indeed, the number of floating point operations needed to compute the SVD of a matrix $A \in \mathbb{C}^{n\times n}$ is of order $\mathcal{O}(n^3)$ [19, Section 5.4.5]. Hence a truncated singular value decomposition is unsuitable for our implementation.

Instead, we use the adaptive cross approximation (ACA), which was originally presented in [1, 3], for the compression of the coupling matrices. This method is based on a gradual approximation of a matrix A by rank one matrices. A detailed description of the general method and the partially pivoted ACA which we used in our implementation can be found in [33, Section 3.2.2, in particular Algorithm 3.9]. An advantage of this method is that it can be used without the knowledge of the full matrix A. Furthermore for the computation of a rank r approximation of a quadratic matrix $A \in \mathbb{C}^{n\times n}$ we only need $\mathcal{O}(r^2 n)$ floating point operations in general, see [33, Remarks on Algorithm 3.9].

2.2.2 Complexity analysis

The standard matrix-vector multiplication $g = Av$ for a dense matrix $A \in \mathbb{C}^{N_X\times N_Y}$ and a vector $v \in \mathbb{C}^{N_Y}$ has a complexity of $\mathcal{O}(N_X N_Y)$ in terms of storage and number of floating point operations needed for the computation. Hence, for large N_X and N_Y its use is prohibitive. This was the motivation for the derivation of the directional approximation of the matrix A in (2.1) and the corresponding fast directional matrix-vector

multiplication presented in Algorithm 2.42. Here we want to analyze the complexity of this algorithm. However, we skip the discussion of the algorithms to construct the box cluster trees and the block tree, as these are standard.

The results which we show in Theorems 2.45, 2.47 and 2.54 hold for rather general assumptions of the underlying box cluster trees. As a consequence, they do not imply an optimal or near optimal complexity in many application cases. Indeed, only for points distributed more or less uniformly in a 3D domain we achieve a complexity of $\mathcal{O}(N \log(N))$ under suitable assumptions on the domain and the wave number κ (cf. (2.68c)). In [30, Section 3.3.6] a complexity of $\mathcal{O}(N \log(N))$ is heuristically motivated for various distributions of points. A more sophisticated discussion for the 2D case can be found in [5, Section 5, cf. Theorem 12] that leads to a complexity of $\mathcal{O}(N + \kappa^2 \log(N))$ under suitable assumptions on the distribution of the points. Since κ^2 is typically of order N in applications, the complexity is of order $\mathcal{O}(N \log(N))$ also in this case.

We start our complexity estimate with an immediate but important observation.

Remark 2.44. Let us consider Algorithm 2.38, which is used to approximate the farfield part g^+ of the matrix-vector product $g = Av$ (cf. Definition 2.32). In this algorithm we multiply every directional interpolation matrix $L_{t,c}$ and $L_{s,c}$, every transfer matrix $E_{t',c}$ and $E_{s',c}$ and every coupling matrix $A_{c,t,s}$ with a suitable vector exactly once. The same holds for the matrices $A|_{\hat{t}\times\hat{s}}$ in Algorithm 2.33, which is used for the computation of the nearfield part g^- of $g = Av$. Furthermore every entry of these matrices can be computed with $\mathcal{O}(1)$ operations. Hence, it suffices to discuss only the storage requirements of Algorithms 2.33 and 2.38 as these are directly proportional to the number of floating point operations required for their execution.

Let us first count the directional interpolation matrices and transfer matrices.

Theorem 2.45. *Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees and $\mathcal{T}_{X\times Y}$ the corresponding block tree. Let $p_{\max} := \max\{p(\mathcal{T}_X), p(\mathcal{T}_Y)\}$ be the maximum of the depths of the trees $\mathcal{T}_X$ and $\mathcal{T}_Y$. Let the highest level in the high frequency regime $\ell_{\mathrm{hf}} \in \mathbb{N}_0$ satisfy*

$$\ell_{\mathrm{hf}} \leq p_{\max} - 1, \tag{2.56}$$

where $p(\mathcal{T}_X)$ and $p(\mathcal{T}_Y)$ are the depths of the corresponding trees. Let the directions $\{c_j^{(\ell)}\}_{j=1}^{N_\ell}$ be constructed as in Algorithm 2.6. Then there exists a constant C_{LE} such that the number N_{E} of all transfer matrices $E_{t',c}$ and $E_{s',c}$ and the number N_{L} of all directional interpolation matrices $L_{t,c}$ and $L_{s,c}$ in Algorithm 2.38 is bounded by

$$N_{\mathrm{L}} + N_{\mathrm{E}} \leq C_{\mathrm{LE}}\, 8^{p_{\max}}. \tag{2.57}$$

Proof. We start by estimating the number of transfer matrices for boxes in the tree $\mathcal{T}_X$, which are needed in step four (L2L) of Algorithm 2.38. For all non-leaf boxes t in $\mathcal{T}_X^{(\ell)} \backslash \mathcal{L}_X$ we need a transfer matrix $E_{t',c}$ for each child t' of t and each direction

$c \in D_t \cup \hat{D}_t$, where the set of active directions D_t is defined in (2.46) and the set of inherited directions $\hat{D}_t$ in (2.47). At each level ℓ of a uniform box cluster tree there exist at most 8^ℓ boxes. Furthermore the number N_ℓ of directions at level ℓ is $6 \cdot 4^{\ell_{\mathrm{hf}}-\ell}$ if $\ell \leq \ell_{\mathrm{hf}}$ and 1 else. As a consequence we can estimate the number $N_{\mathrm{E},X}^{(\ell)}$ of transfer matrices at level $\ell \leq \ell_{\mathrm{hf}}$ by

$$\begin{aligned} N_{\mathrm{E},X}^{(\ell)} &= \sum_{t \in \mathcal{T}_X^{(\ell)}} \#\operatorname{child}(t) \#(D_t \cup \hat{D}_t) \leq \sum_{t \in \mathcal{T}_X^{(\ell)}} 8 \cdot 6 \cdot 4^{\ell_{\mathrm{hf}}-\ell} \\ &\leq 8^{\ell+1} \cdot 6 \cdot 4^{\ell_{\mathrm{hf}}-\ell} \leq 48 \cdot 4^{\ell_{\mathrm{hf}}} \cdot 2^\ell \end{aligned} \tag{2.58}$$

and at level $\ell > \ell_{\mathrm{hf}}$ by

$$N_{\mathrm{E},X}^{(\ell)} = \sum_{t \in \mathcal{T}_X^{(\ell)}} \#\operatorname{child}(t) \leq 8^{\ell+1}.$$

Let us assume without loss of generality that $p(\mathcal{T}_X) = p_{\max}$. The number $N_{\mathrm{E},X}$ of all transfer matrices can then be estimated by

$$\begin{aligned} N_{\mathrm{E},X} &= \sum_{\ell=0}^{p(\mathcal{T}_X)-1} N_{\mathrm{E},X}^{(\ell)} \leq \sum_{\ell=0}^{\ell_{\mathrm{hf}}} 48 \cdot 4^{\ell_{\mathrm{hf}}} \cdot 2^\ell + \sum_{\ell=\ell_{\mathrm{hf}}+1}^{p(\mathcal{T}_X)-1} 8^{\ell+1} \leq 96 \cdot 8^{\ell_{\mathrm{hf}}} + (8^{p(\mathcal{T}_X)+1} - 8^{\ell_{\mathrm{hf}}+2}) \\ &= 4 \cdot 8^{\ell_{\mathrm{hf}}+1} + 8 \cdot 8^{p(\mathcal{T}_X)} \leq 12 \cdot 8^{p(\mathcal{T}_X)} = 12 \cdot 8^{p_{\max}}, \end{aligned} \tag{2.59}$$

if we assume that $\ell_{\mathrm{hf}} \leq p(\mathcal{T}_X) - 1 = p_{\max} - 1$.

Next we estimate the number $N_{\mathrm{L},X}$ of all directional interpolation matrices $L_{t,c}$ for boxes in $\mathcal{T}_X$, which are required in step five (L2T) of Algorithm 2.38. Let us first assume that all leaf boxes $t \in \mathcal{L}_X$ are at a level $\ell > \ell_{\mathrm{hf}}$ in $\mathcal{T}_X$. In this case we need for each leaf box exactly one interpolation matrix, which does not depend on any direction. Since a uniform box cluster tree with depth $p(\mathcal{T}_X)$ has at most $8^{p(\mathcal{T}_X)}$ leaves, we can estimate the number $N_{\mathrm{L},X}$ of interpolation matrices in $\mathcal{T}_X$ by

$$N_{\mathrm{L},X} \leq 8^{p(\mathcal{T}_X)} \leq 8^{p_{\max}}.$$

Together with (2.59) this yields

$$N_{\mathrm{E},X} + N_{\mathrm{L},X} \leq 13 \cdot 8^{p_{\max}}. \tag{2.60}$$

Let us now consider the general case, where we allow leaves at levels $\ell \leq \ell_{\mathrm{hf}}$. For all such leaves $t \in \mathcal{T}_X^{(\ell)}$ we need a directional interpolation matrix for all directions $c \in D_t \cup \hat{D}_t$, i.e. at most $6 \cdot 4^{\ell_{\mathrm{hf}}-\ell}$ per box. However, for such leaf boxes we do not need any transfer-matrices, but have counted nonetheless $8 \cdot 6 \cdot 4^{\ell_{\mathrm{hf}}-\ell}$ such matrices in the estimate of $N_{\mathrm{E},X}^{(\ell)}$ in (2.58). Hence, the estimate (2.60) holds also in the general case.

Analogously we denote by $N_{\mathrm{E},Y}$ the number of transfer matrices and by $N_{\mathrm{L},Y}$ the number of directional interpolation matrices for boxes in $\mathcal{T}_Y$ which are needed in steps two (M2M) and one (S2M) of Algorithm 2.38, respectively. As above we get

$$N_{\mathrm{E},Y} + N_{\mathrm{L},Y} \leq 13 \cdot 8^{p_{\max}}.$$

Together with (2.60) this yields

$$N_{\mathrm{E}} + N_{\mathrm{L}} = N_{\mathrm{E},X} + N_{\mathrm{E},Y} + N_{\mathrm{L},X} + N_{\mathrm{L},Y} \leq 26 \cdot 8^{p_{\max}} = C_{\mathrm{LE}}\, 8^{p_{\max}}. \qquad \square$$

Remark 2.46. Let all assumptions of Theorem 2.45 hold, but $\ell_{\mathrm{hf}} \geq p_{\max} - 1$. By repeating the proof of Theorem 2.45 we end up with the estimate

$$N_{\mathrm{E}} \leq 96 \cdot 4^{\ell_{\mathrm{hf}} - p_{\max}} \cdot 8^{p_{\max}}, \tag{2.61}$$

where N_{E} still denotes the number of transfer matrices. In the proof we ignored that the number of directions per box t at level ℓ is not only bounded by the total number of directions at level ℓ but also by the number of admissible boxes s such that $(t, s) \in \mathcal{L}^+_{X \times Y}$. Hence the estimate (2.61) is in general far from optimal. However, we observe for sufficiently large ℓ_{hf} that a different direction will be used for each admissible box. In this case the number of transfer matrices and directional matrices exceeds the number of coupling matrices, which is undesirable. Hence, we should not choose ℓ_{hf} too large in general.

In the following theorem we estimate the number N_{C} of all coupling matrices in Algorithm 2.38. The proof of this theorem is inspired by [5, proof of Lemma 8]. In particular, the strategy which we use to count hierarchically admissible clusters is identical.

Theorem 2.47. *Let $\kappa > 0$. Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be two uniform box cluster trees and $\mathcal{T}_{X \times Y}$ the corresponding block tree constructed as in Algorithm 2.26 for a given $\eta_2 > 0$. Let the diameters $q_\ell(\mathcal{T}_X)$ and $q_\ell(\mathcal{T}_Y)$ be defined as in* (2.9) *and $p(\mathcal{T}_X)$ and $p(\mathcal{T}_Y)$ be the depths of the corresponding box cluster trees. Define the depth $p_{\min} := \min\{p(\mathcal{T}_X), p(\mathcal{T}_Y)\}$ and the diameters $q_\ell := \max\{q_\ell(\mathcal{T}_X), q_\ell(\mathcal{T}_Y)\}$ for all $\ell \in \{0, ..., p_{\min}\}$. Let $C_{\mathrm{un}} \geq 1$ be such that*

$$\frac{q_\ell}{q_\ell(\mathcal{T}_X)} \leq C_{\mathrm{un}}. \tag{2.62}$$

Then there exists a constant $C(C_{\mathrm{un}}, \eta_2)$ depending only on C_{un} and η_2, such that the number N_{C} of all coupling matrices $A_{c,t,s}$ in Algorithm 2.38 is bounded by

$$N_{\mathrm{C}} \leq C(C_{\mathrm{un}}, \eta_2) \left(p_{\min} (q_0 \kappa)^3 + 8^{p_{\min}} \right). \tag{2.63}$$

Proof. [5, cf. proof of Lemma 8] If there exists a leaf in $\mathcal{T}_{X \times Y}$ at level zero then $N_{\mathrm{C}} \leq 1$ and the assertion is trivial. Hence we can assume that no such leaf exists. We want to

estimate the number of coupling matrices levelwise. Therefore we denote the number of all coupling matrices at level $\ell \in \{1, ..., p_{\min}\}$ by $N_{C,\ell}$.

For the estimate of $N_{C,\ell}$ we observe, that the number of coupling matrices needed for a box $s' \in \mathcal{T}_Y^{(\ell)}$ in step three of Algorithm 2.38 corresponds to the number of boxes $t' \in \mathcal{T}_X^{(\ell)}$ such that t' and s' satisfy $(t', s') \in \mathcal{L}_{X\times Y}^+$. The interaction list $\mathcal{I}^+(s')$ defined by

$$\mathcal{I}^+(s') := \{t' \in \mathcal{T}_Y^{(\ell)} : (t', s') \in \mathcal{L}_{X\times Y}^+\} \tag{2.64}$$

collects all such boxes. The parent $t \in \mathcal{T}_X^{(\ell-1)}$ of a box t' in $\mathcal{I}^+(s')$ is contained in the nearfield $\mathcal{N}(s)$ of $s = \text{parent}(s') \in \mathcal{T}_Y^{(\ell-1)}$ by construction of $\mathcal{T}_{X\times Y}$ in Algoritm 2.26. Therefore, we can estimate the number of boxes $t' \in \mathcal{I}^+(s')$ by estimating the number of boxes $t \in \mathcal{N}(s)$. In fact there holds

$$\#\mathcal{I}^+(s') \leq 8\#\mathcal{N}(s),$$

because every box $t \in \mathcal{N}(s)$ has at most eight children. To estimate $\#\mathcal{N}(s)$ we construct a ball $B_{r_{\ell-1}}(m_s)$ with radius $r_{\ell-1}$ centered at the midpoint m_s of s which covers $\mathcal{N}(s)$, i.e.

$$\bigcup_{t\in\mathcal{N}_s} t \subset B_{r_{\ell-1}}(m_s).$$

Then there holds

$$\#\mathcal{N}(s) \leq \frac{|B_{r_{\ell-1}}(m_s)|}{v_{\ell-1}(\mathcal{T}_X)} = \frac{(4\pi/3)r_{\ell-1}^3}{3^{-3/2}q_{\ell-1}(\mathcal{T}_X)^3} = 4\pi\sqrt{3}\left(\frac{r_{\ell-1}}{q_{\ell-1}(\mathcal{T}_X)}\right)^3,$$

where $v_{\ell-1}(\mathcal{T}_X) = 3^{-3/2}q_{\ell-1}(\mathcal{T}_X)^3$ denotes the volume of a box $t \in \mathcal{T}_X^{(\ell-1)}$. In particular we get

$$\begin{aligned} N_{C,\ell} &= \sum_{s'\in\mathcal{T}_Y^{(\ell)}} \#\mathcal{I}^+(s') \leq \sum_{s'x\in\mathcal{T}_Y^{(\ell)}} 8\#\mathcal{N}(\text{parent}(s')) \\ &\leq 8^\ell \cdot 32\pi\sqrt{3}\left(\frac{r_{\ell-1}}{q_{\ell-1}(\mathcal{T}_X)}\right)^3 = 8^\ell\, C\left(\frac{r_{\ell-1}}{q_{\ell-1}(\mathcal{T}_X)}\right)^3, \end{aligned} \tag{2.65}$$

with the constant $C = 32\pi\sqrt{3}$.

In the following we construct a ball $B_{r_{\ell-1}}(m_s)$ which covers the nearfield $\mathcal{N}(s)$ of a box $s \in \mathcal{T}_Y^{(\ell-1)}$. For this purpose we have to consider the admissibility criteria (A1) and (A3). Let $\tilde{\ell}$ be such that $\kappa q_{\tilde{\ell}-1} > 1$ and $\kappa q_{\tilde{\ell}} \leq 1$. Then the admissibility criterion (A3) implies (A1) for boxes $t \in \mathcal{T}_X^{(\ell-1)}$ and $s \in \mathcal{T}_Y^{(\ell-1)}$ if $\ell \leq \tilde{\ell}$. In particular there holds $t \in \mathcal{N}(s)$ if and only if (A3) is violated. Similarly, if $\ell > \tilde{\ell}$ then $t \in \mathcal{N}(s)$ holds if and only if (A1) is violated. We distinguish these two cases to determine a suitable ball $B_{r_{\ell-1}}(m_s)$.

Let first $\ell \in \{1, ...\tilde{\ell}\}$. We consider a fixed box $t \in \mathcal{T}_X^{(\ell-1)}$ in the nearfield $\mathcal{N}(s)$ of $s \in \mathcal{T}_Y^{(\ell-1)}$ and let $z \in \bar{t}$ be such that

$$|z - m_s| = \max_{x \in t} |x - m_s|.$$

We know that $\eta_2 \operatorname{dist}(t, s) < \kappa q_{\ell-1}^2$, because (A3) is violated since $t \in \mathcal{N}(s)$. Hence, there exist $x \in \bar{t}$ and $y \in \bar{s}$ such that

$$|x - y| < \frac{q_{\ell-1}}{\eta_2}.$$

Therefore we can estimate

$$\begin{aligned} |z - m_s| &\leq |z - x| + |x - y| + |y - m_s| < q_{\ell-1}(\mathcal{T}_X) + \frac{\kappa q_{\ell-1}^2}{\eta_2} + \frac{q_{\ell-1}(\mathcal{T}_Y)}{2} \\ &\leq \frac{3q_{\ell-1}}{2} + \frac{\kappa q_{\ell-1}^2}{\eta_2} < \left(\frac{3}{2} + \frac{1}{\eta_2}\right) \kappa q_{\ell-1}^2, \end{aligned}$$

where we used $q_{\ell-1} < \kappa q_{\ell-1}^2$ for the last estimate. In particular every box $t \in \mathcal{N}(s)$ is contained in the ball $B_{r_{\ell-1}}(m_s)$ with radius

$$r_{\ell-1} = \left(\frac{3}{2} + \frac{1}{\eta_2}\right) \kappa q_{\ell-1}^2.$$

Together with (2.62) and (2.65) we get

$$\begin{aligned} N_{C,\ell} &\leq 8^\ell \, C \left(\frac{r_{\ell-1}}{q_{\ell-1}(\mathcal{T}_X)}\right)^3 \leq 8^\ell \, C \left(\frac{3}{2} + \frac{1}{\eta_2}\right)^3 \left(\frac{q_{\ell-1}}{q_{\ell-1}(\mathcal{T}_X)}\right)^3 (\kappa q_{\ell-1})^3 \\ &\leq 8^\ell \, C \left(\frac{3}{2} + \frac{1}{\eta_2}\right)^3 C_{\text{un}}^3 (\kappa q_{\ell-1})^3 \end{aligned} \tag{2.66}$$

for all $\ell \in \{1, ...\tilde{\ell}\}$.

Let now $\ell \in \{\tilde{\ell} + 1, ..., p_{\min}\}$. Then a box $t \in \mathcal{T}_X^{(\ell-1)}$ is in the nearfield $\mathcal{N}(s)$ of $s \in \mathcal{T}_Y^{(\ell-1)}$ if (A1) is violated, i.e. $\eta_2 \operatorname{dist}(t, s) < q_{\ell-1}$. Analogously as above this yields that the nearfield $\mathcal{N}(s)$ is covered by the ball $B_{r_{\ell-1}}(m_s)$ with radius

$$r_{\ell-1} = \left(\frac{3}{2} + \frac{1}{\eta_2}\right) q_{\ell-1}.$$

In particular for all $\ell \in \{\tilde{\ell} + 1, ..., p_{\min}\}$ there holds the estimate

$$N_{C,\ell} \leq 8^\ell \, C \left(\frac{3}{2} + \frac{1}{\eta_2}\right)^3 C_{\text{un}}^3. \tag{2.67}$$

To estimate the number N_{C} of all needed coupling matrices it suffices to sum up the estimates of $N_{\mathrm{C},\ell}$ over all levels $\ell \in \{1, ..., p_{\min}\}$. With $q_0 = 2^\ell q_\ell$ and the estimates (2.66) and (2.67) we get

$$\begin{aligned} N_{\mathrm{C}} = \sum_{\ell=1}^{p_{\min}} N_{\mathrm{C},\ell} &\leq \sum_{\ell=1}^{\tilde{\ell}} 8\ C \left(\frac{3}{2} + \frac{1}{\eta_2}\right)^3 C_{\mathrm{un}}^3 (\kappa q_0)^3 + \sum_{\ell=\tilde{\ell}+1}^{p_{\min}} 8^\ell\ C \left(\frac{3}{2} + \frac{1}{\eta_2}\right)^3 C_{\mathrm{un}}^3 \\ &\leq 8\ \tilde{\ell}\ (q_0\kappa)^3\ C\ C_{\mathrm{un}}^3 \left(\frac{3}{2} + \frac{1}{\eta_2}\right)^3 + 8^{p_{\min}+1} C\ C_{\mathrm{un}}^3 \left(\frac{3}{2} + \frac{1}{\eta_2}\right)^3 \\ &< C(C_{\mathrm{un}}, \eta_2) \left(p_{\min}(q_0\kappa)^3 + 8^{p_{\min}}\right), \end{aligned}$$

in case that $\tilde{\ell} < p_{\min}$, with the constant

$$C(C_{\mathrm{un}}, \eta_2) := 8\ C\ C_{\mathrm{un}}^3 \left(\frac{3}{2} + \frac{1}{\eta_2}\right)^3.$$

In case that $\tilde{\ell} \geq p_{\min}$ we similarly get the estimate

$$N_{\mathrm{C}} \leq C(C_{\mathrm{un}}, \eta_2)\ p_{\min}\ (q_0\kappa)^3 \leq C(C_{\mathrm{un}}, \eta_2) \left(p_{\min}(q_0\kappa)^3 + 8^{p_{\min}}\right). \qquad \square$$

In Theorems 2.45 and 2.47 we have estimated the number of matrices which are needed for the application of Algorithm 2.38. Hence we can now estimate its complexity.

Theorem 2.48. *Let A be a matrix as in (2.1) corresponding to two sets of points P_X and P_Y with cardinalities N_X and N_Y. Let $N := \max\{N_X, N_Y\}$. Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be the uniform box cluster trees corresponding to these sets of points as constructed in Algorithm 2.3. In particular let each leaf box contain at most $n_{\max}$ points. Let $\mathcal{T}_{X\times Y}$ be the block tree corresponding to $\mathcal{T}_X$ and $\mathcal{T}_Y$. Let all the assumptions of Theorems 2.45 and 2.47 hold. Furthermore, let m be the interpolation degree for the directional approximation. Assume that there exist constants C_{ad}, $C_{\max}$, and C_{geo} such that*

$$n_{\max} \leq C_{\max}(m+1)^3, \tag{2.68a}$$

$$p_{\max} \leq \log_8(N) + C_{\mathrm{ad}}, \tag{2.68b}$$

$$\kappa q_0 \leq C_{\mathrm{geo}} \sqrt[3]{N}. \tag{2.68c}$$

Then there exists a constant $C(C_{\max}, C_{\mathrm{LE}}, C_{\mathrm{ad}}, C_{\mathrm{un}}, C_{\mathrm{geo}}, \eta_2)$ such that Algorithm 2.38 requires not more than

$$C(C_{\max}, C_{\mathrm{LE}}, C_{\mathrm{ad}}, C_{\mathrm{un}}, C_{\mathrm{geo}}, \eta_2)(m+1)^6 N \log(N) \tag{2.69}$$

units of storage.

Proof. We observe that due to (2.68a) a directional interpolation matrix $A_{c,t,s}$ has at most $C_{\max}(m+1)^6$ entries, where m is the interpolation degree. Since all transfer and coupling matrices have exactly $(m+1)^6$ entries, we see that each matrix used in Algorithm 2.38 requires at most $C_{\max}(m+1)^6$ units of storage. By Theorems 2.45 and 2.47 we know that the total number of all such matrices is bounded by

$$\begin{aligned} N_{\mathrm{L}} + N_{\mathrm{E}} + N_{\mathrm{C}} &\leq C_{\mathrm{LE}}\ 8^{p_{\max}} + C(C_{\mathrm{un}}, \eta_2)\left(p_{\min}(q_0\kappa)^3 + 8^{p_{\min}+1}\right) \\ &\leq C_{\mathrm{LE}}\ 8^{C_{\mathrm{ad}}} N + C(C_{\mathrm{un}}, \eta_2)\left((\log_8(N) + C_{\mathrm{ad}})C_{\mathrm{geo}}^3 N + 8^{C_{\mathrm{ad}}+1} N\right) \\ &\leq C(C_{\mathrm{LE}}, C_{\mathrm{ad}}, C_{\mathrm{un}}, C_{\mathrm{geo}}, \eta_2) N \log(N), \end{aligned}$$

for a suitable constant $C(C_{\mathrm{LE}}, C_{\mathrm{ad}}, C_{\mathrm{un}}, C_{\mathrm{geo}}, \eta_2)$. Hence, we need at most

$$C(C_{\mathrm{LE}}, C_{\mathrm{ad}}, C_{\mathrm{un}}, C_{\mathrm{geo}}, \eta_2) N \log(N) C_{\max}(m+1)^6,$$

units of storage. By setting

$$C(C_{\max}, C_{\mathrm{LE}}, C_{\mathrm{ad}}, C_{\mathrm{un}}, C_{\mathrm{geo}}, \eta_2) = C_{\max}\ C(C_{\mathrm{LE}}, C_{\mathrm{ad}}, C_{\mathrm{un}}, C_{\mathrm{geo}}, \eta_2)$$

we end up with (2.69). □

Remark 2.49. The constant $C_{\max}$ in assumption (2.68a) can be controlled by choosing $n_{\max}$ in the construction of the uniform box cluster trees sufficiently low. The assumptions (2.68b) and (2.68c) on the other hand depend substantially on the position of the points in P_X and P_Y. For points which are distributed more or less uniformly in a 3D domain they are reasonable. However, in the setting of boundary element methods the points are typically distributed on 2D manifolds. In particular we would expect in such applications that

$$\begin{aligned} p_{\max} &\approx \log_4(N), \\ \kappa q_0 &\leq C_{\mathrm{geo}}\sqrt{N}. \end{aligned}$$

With these assumptions we would end up with a quadratic complexity in N by repeating the previous proof, which is unsatisfactory. Hence we need to refine our estimates in Theorems 2.45 and 2.47 for points on 2D manifolds to end up with a complexity of $\mathcal{O}((m+1)^6 N \log(N))$. However, we do not further discuss this here but refer to [5, Section 5] instead.

In Section 2.2.1 we have seen that we only have to compute coupling matrices once for a reference configuration due to the uniformity of the box cluster trees $\mathcal{T}_X$ and $\mathcal{T}_Y$. In the following proposition we discuss the effect of this approach for the required storage.

Proposition 2.50. *Let $\kappa > 0$. Let two uniform box cluster trees $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$ and the corresponding block tree $\mathcal{T}_{X\times Y}$ be given as in Theorem 2.47. In particular let*

$p_{\min} = \min\{p(\mathcal{T}_X), p(\mathcal{T}_Y)\}$ be the minimal depth of the two box cluster trees and the constant $C(C_{\mathrm{un}}, \eta_2)$ be the same as in (2.63). Then we can estimate the number N_{SC} of coupling matrices $A_{c,t,s}$ which we need to store by

$$N_{\mathrm{SC}} \leq p_{\min} \max\{C(C_{\mathrm{un}}, \eta_2), \sqrt{C(C_{\mathrm{un}}, \eta_2)}(\kappa q_0)^{3/2}\}. \tag{2.70}$$

Let furthermore N_X and N_Y be the cardinalities of the sets P_X and P_Y and their maximum be given by $N = \max\{N_X, N_Y\}$. Let m be the interpolation degree chosen for the directional approximation. Assume that there exist constants C_{p} and $C_{\mathrm{geo},2}$ such that

$$p_{\min} \leq C_{\mathrm{p}} \log(N), \tag{2.71a}$$

$$\kappa q_0 \leq C_{\mathrm{geo},2}\sqrt{N}. \tag{2.71b}$$

Then there exists a constant $C(C_{\mathrm{un}}, \eta_2, C_{\mathrm{p}}, C_{\mathrm{geo},2})$ depending only on C_{un}, η_2, C_{p}, and $C_{\mathrm{geo},2}$ such that the number M_{SC} of entries of coupling matrices $A_{c,t,s}$ which we need to store is bounded by

$$M_{\mathrm{SC}} \leq C(C_{\mathrm{un}}, \eta_2, C_{\mathrm{p}}, C_{\mathrm{geo},2})(m+1)^6 N^{3/4} \log(N). \tag{2.72}$$

Proof. The number $N_{\mathrm{SC},\ell}$ of coupling matrices which we need to store at level ℓ is the same as the maximal possible number of boxes $t' \in \mathcal{T}_X^{(\ell)}$ in the interaction list $\mathcal{I}^+(s')$ of a box $s' \in \mathcal{T}_Y^{(\ell)}$, which we defined in (2.64). Analogously as in the proof of Theorem 2.47 (cf. equations (2.66) and (2.67)) it follows that

$$N_{\mathrm{SC},\ell} \leq C(C_{\mathrm{un}}, \eta_2) \max\{1, (\kappa q_{\ell-1})^3\}.$$

We further know that there can be at most 8^ℓ boxes in $\mathcal{T}_X^{(\ell)}$ and that $q_\ell = 2^{-\ell} q_0$ due to the uniformity of the box cluster trees $\mathcal{T}_X$ and $\mathcal{T}_Y$. Hence, there holds the even stronger estimate

$$N_{\mathrm{SC},\ell} \leq \min\{8^\ell, C(C_{\mathrm{un}}, \eta_2) \max\{1, (\kappa q_0)^3 8^{1-\ell}\}\} = \overline{N}_{\mathrm{SC},\ell}$$

Let us compute the maximum of $\overline{N}_{\mathrm{SC},\ell}$. Such a maximum exists, since 8^ℓ is strictly increasing and $\max\{1, (\kappa q_0)^3 8^{1-\ell}\}$ is decreasing. In particular there holds

$$\max_{\ell \in \{0,\dots,p_{\min}\}} \overline{N}_{\mathrm{SC},\ell} \geq \min_{\ell \in \mathbb{N}_0} C(C_{\mathrm{un}}, \eta_2) \max\{1, (\kappa q_0)^3 8^{1-\ell}\} = C(C_{\mathrm{un}}, \eta_2). \tag{2.73}$$

If equality holds in (2.73) we have found the maximum. Else we compute the intersection point of the expressions 8^ℓ and $C(C_{\mathrm{un}}, \eta_2)(\kappa q_0)^3 8^{1-\ell}$, i.e. we want to find ℓ^* such that

$$8^{\ell^*} = C(C_{\mathrm{un}}, \eta_2)(\kappa q_0)^3 8^{1-\ell^*}.$$

This is equivalent to

$$8^{\ell^*} = \sqrt{8\ C(C_{\mathrm{un}}, \eta_2)}(\kappa q_0)^{3/2}.$$

In particular we get the estimate

$$\begin{aligned}\max_{\ell\in\{0,\dots,p_{\min}\}} N_{\mathrm{SC},\ell} \le 8^{\ell^*} &= \sqrt{8\ C(C_{\mathrm{un}},\eta_2)}(\kappa q_0)^{3/2} \\ &\le \max\left\{C(C_{\mathrm{un}},\eta_2), \sqrt{8\ C(C_{\mathrm{un}},\eta_2)}(\kappa q_0)^{3/2}\right\}.\end{aligned}$$

Summation over all numbers $N_{\mathrm{SC},\ell}$ finally yields

$$N_{\mathrm{SC}} = \sum_{\ell=1}^{p_{\min}} N_{\mathrm{SC},\ell} \le p_{\min} \max\{C(C_{\mathrm{un}},\eta_2), \sqrt{8\ C(C_{\mathrm{un}},\eta_2)}(\kappa q_0)^{3/2}\}.$$

Here we assumed that at level $\ell = 0$ no coupling matrix is stored, because else we need exactly one coupling matrix and the assertion is trivial.

Let now (2.71a) and (2.71b) hold. We know that each coupling matrix has exactly $(m+1)^6$ entries, where m is the interpolation degree. Hence, we get

$$\begin{aligned}M_{\mathrm{SC}} = (m+1)^6 N_{\mathrm{SC}} &\le \sqrt{8}\ C(C_{\mathrm{un}},\eta_2)(m+1)^6 p_{\min}\max\{1,(\kappa q_0)^{3/2}\} \\ &\le C(C_{\mathrm{un}},\eta_2,C_{\mathrm{p}},C_{\mathrm{geo},2})(m+1)^6 N^{3/4}\log(N),\end{aligned}$$

with the constant $C(C_{\mathrm{un}},\eta_2,C_{\mathrm{p}},C_{\mathrm{geo},2}) = \sqrt{8}\ C(C_{\mathrm{un}},\eta_2)C_{\mathrm{p}}C_{\mathrm{geo},2}$. □

Remark 2.51. Assumptions (2.71a) and (2.71b) are comparable to assumptions (2.68b) and (2.68c) in Theorem 2.48, but weaker. In particular, assumption (2.71b) is generally fulfilled in applications if the points in P_X and P_Y are scattered more or less uniformly in a 3D domain or on a 2D manifold. This is also assumed in the complexity discussion in [30, Section 3.3.6]. Under these assumptions the costs to store the coupling matrices are of order $\mathcal{O}(N^{3/4}\log(N))$. However, we still need to assign the stored coupling matrices to the corresponding admissible blocks $(t,s) \in \mathcal{L}^+_{X\times Y}$. For this purpose we need N_{C} assignments, where N_{C} denotes the number of all needed coupling matrices. Each of these assignments needs only $\mathcal{O}(1)$ storage. Hence, for general applications they play a subordinate role. However, asymptotically they require $\mathcal{O}(N_{\mathrm{C}})$ storage and hence the asymptotic storage requirements do not change, if we store every coupling matrix only once.

Next we consider the nearfield matrices $A|_{\hat{t}\times\hat{s}}$ corresponding to inadmissible blocks in the block tree $\mathcal{T}_{X\times Y}$. These matrices are not approximated but used directly for the computation in Algorithm 2.33. Hence, we estimate the number of entries of these matrices to estimate the required storage in Theorem 2.54. In the formulation and the proof of this theorem we make use of the following definition.

Definition 2.52. Let $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$ be two uniform box cluster trees constructed as in Algorithm 2.3. In particular let $n_{\max} \in \mathbb{N}$ be such that all leaves $t \in \mathcal{L}_X$ and $s \in \mathcal{L}_Y$ contain at most $n_{\max}$ points. Let $\mathcal{T}_{X\times Y}$ be the block tree corresponding

to $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$. Let $C_{\mathrm{m}} \geq 1$ and let the sets of indices $\hat{t}$ and $\hat{s}$ for boxes $t \in \mathcal{T}_X$ and $s \in \mathcal{T}_Y$ be defined by (2.3). Then we define the set of controllable leaves $\mathcal{L}^+_{Y,X\times Y}$ in $\mathcal{T}_Y$ and the remaining leaves $\mathcal{L}^-_{Y,X\times Y}$ by

$$\begin{aligned} \mathcal{L}^+_{Y,X\times Y} &:= \bigcup_{\substack{\ell\in\{0,\dots,p_{\min}\}\\ \kappa q_\ell<1}} \{s \in \mathcal{T}_Y^{(\ell)} \cap \mathcal{L}_Y : \#\hat{t} \leq C_{\mathrm{m}} n_{\max} \text{ if } (t,s) \in \mathcal{L}^+_{X\times Y}\}, \\ \mathcal{L}^-_{Y,X\times Y} &:= \mathcal{L}_Y \backslash \mathcal{L}^+_{Y,X\times Y}. \end{aligned} \tag{2.74}$$

Analogously we define the set of controllable leaves $\mathcal{L}^+_{X,X\times Y}$ in $\mathcal{T}_X$ and the remaining leaves $\mathcal{L}^-_{X,X\times Y}$.

Remark 2.53. In [5, Section 5, Equations (15) and (16)] the author assumes for all leaves in the cluster trees similar conditions as we do for the leaves in the sets $\mathcal{L}^+_{Y,X\times Y}$ and $\mathcal{L}^+_{Y,X\times Y}$ to control the contribution of the inadmissible blocks in the complexity estimates. In particular, it is assumed that all leaves of the respective cluster trees satisfy $\kappa q_\ell < 1$, i.e. they have sufficiently large levels. In our case we cannot assume such a property in general, because there can be leaves at any level of a tree due to the construction of the uniform box cluster trees in Algorithm 2.3. Only for suitable geometries or box cluster trees constructed with a bisection method, for example, this assumption is reasonable. In the estimate which we present in Theorem 2.54 we assume that the number of leaves which do not have such nice properties, i.e. the leaves in $\mathcal{L}^-_{Y,X\times Y}$ and $\mathcal{L}^-_{Y,X\times Y}$ are only few.

Theorem 2.54. *Let the matrix A be defined as in (2.1) and P_X and P_Y be the corresponding sets of points with cardinalities N_X and N_Y. Let $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$ be two uniform box cluster trees constructed as in Algorithm 2.3. In particular let $n_{\max} \in \mathbb{N}$ be such that all leaves $t \in \mathcal{L}_X$ and $s \in \mathcal{L}_Y$ contain at most $n_{\max}$ points. Let $p_{\min} = \min\{p(\mathcal{T}_X), p(\mathcal{T}_Y)\}$ be the minimal depth of the two box cluster trees. Let $q_\ell := \max\{q_\ell(\mathcal{T}_X), q_\ell(\mathcal{T}_Y)\}$ be the maximum of the diameters of boxes at level ℓ in the two trees and let $C_{\mathrm{un}} \geq 1$ be such that*

$$\frac{q_\ell}{\min\{q_\ell(\mathcal{T}_X), q_\ell(\mathcal{T}_Y)\}} \leq C_{\mathrm{un}}. \tag{2.75}$$

Let $\mathcal{T}_{X\times Y}$ be the block tree corresponding to $\mathcal{T}_X(P_X)$ and $\mathcal{T}_Y(P_Y)$. Let $C_{\mathrm{m}} \geq 1$ and let the controllable leaves $\mathcal{L}^+_{Y,X\times Y}$ and $\mathcal{L}^+_{X,X\times Y}$, as well as the remaining leaves $\mathcal{L}^-_{Y,X\times Y}$ and $\mathcal{L}^-_{X,X\times Y}$ be given by Definition 2.52. Finally let $C_{\mathrm{in}} \in \mathbb{N}_0$ be such that

$$\#\mathcal{L}^-_{X,X\times Y} \leq C_{\mathrm{in}}, \quad \#\mathcal{L}^-_{Y,X\times Y} \leq C_{\mathrm{in}}. \tag{2.76}$$

Then there exists a constant $\hat{C}(C_{\mathrm{un}}, \eta_2)$ depending only on C_{un} and η_2 such that the number M_{D} of entries of all nearfield matrices $A|_{\hat{t}\times\hat{s}}$ corresponding to inadmissible blocks $(t,s) \in \mathcal{L}^-_{X\times Y}$ is bounded by

$$M_{\mathrm{D}} \leq C_{\mathrm{in}} n_{\max}(N_X + N_Y) + \hat{C}(C_{\mathrm{un}}, \eta_2) C_{\mathrm{m}} n_{\max}^2 8^{p_{\min}+1}. \tag{2.77}$$

In particular, if $p_{\min} \leq \log_8(N) + C_{\mathrm{ad}}$ *for* $N = \max\{N_X, N_Y\}$ *and a constant* $C_{\mathrm{ad}} \in \mathbb{N}_0$*, then there exists a constant* C_{nf} *depending only on* C_{ad}*,* C_{in}*,* C_{m}*,* C_{un}*,* $n_{\max}$*, and* η_2 *such that Algorithm 2.33 does not need more than* $C_{\mathrm{nf}}N$ *units of storage.*

Proof. For each inadmissible block $(t, s) \in \mathcal{L}^-_{X\times Y}$ we know that either $t \in \mathcal{L}_X$ or $s \in \mathcal{L}_Y$. Hence, we start by considering the leaves in $\mathcal{L}_Y$. Our goal is to estimate the number $M_{\mathrm{D},Y}$ of matrix entries corresponding to all blocks $(t, s) \in \mathcal{L}^-_{X\times Y}$ with $s \in \mathcal{L}_Y$. We split this estimate into two parts. First we consider the number $M^-_{\mathrm{D},Y}$ of matrix entries corresponding to all blocks $(t, s) \in \mathcal{L}^-_{X\times Y}$ such that $s \in \mathcal{L}^-_{Y,X\times Y}$. For $s \in \mathcal{L}^-_{Y,X\times Y}$ the nearfield can contain all N_X points in P_X. However, the number of boxes in $\mathcal{L}^-_{Y,X\times Y}$ is bounded by C_{in} by assumption. Furthermore we know that the number of points in s is bounded by $n_{\max}$ by construction of the tree $\mathcal{T}_X$ in Algorithm 2.3. Hence, we get

$$M^-_{\mathrm{D},Y} = \sum_{s\in\mathcal{L}^-_{Y,X\times Y}} \#\hat{s} \sum_{\{t:(s,t)\in\mathcal{L}^-_{X\times Y}\}} \#\hat{t} \leq C_{\mathrm{in}} n_{\max} N_X. \tag{2.78}$$

Next we consider the number $M^+_{\mathrm{D},Y}$ of matrix entries corresponding to all blocks $(t, s) \in \mathcal{L}^-_{X\times Y}$ such that $s \in \mathcal{L}^+_{Y,X\times Y}$. Each matrix corresponding to such a block has at most $C_{\mathrm{m}} n^2_{\max}$ entries, as $\#\hat{s} \leq n_{\max}$ by construction of $\mathcal{T}_Y$ and $\#\hat{t} \leq C_{\mathrm{m}} n_{\max}$ by definition of $\mathcal{L}^+_{Y,X\times Y}$ in (2.74). Hence, it suffices to count the number of all these blocks. Let $s \in \mathcal{L}^+_{Y,X\times Y}$. We know that $\ell = \mathrm{level}(s)$ is such that $\kappa q_\ell \leq 1$ by definition of $\mathcal{L}^+_{Y,X\times Y}$. Similarly as in the proof of Theorem 2.47 we see that the nearfield $\mathcal{N}(s)$ is included in a ball with radius $(3/2 + 1/\eta_2)q_\ell$ centered at the midpoint m_s of s. This allows us to estimate the number of boxes t in the nearfield $\mathcal{N}(s)$ of s by

$$\#\mathcal{N}(s) \leq 4\pi\sqrt{3}\left(\frac{3}{2} + \frac{1}{\eta_2}\right)^3 C^3_{\mathrm{un}} =: \tilde{C}(C_{\mathrm{un}}, \eta_2),$$

similarly as in (2.67). With this estimate we get

$$M^+_{\mathrm{D},Y,\ell} := \sum_{\substack{s\in\mathcal{L}^+_{Y,X\times Y}\\ \mathrm{level}(s)=\ell}} \#\hat{s} \sum_{\left\{t:(s,t)\in\mathcal{L}^-_{X\times Y}\right\}} \#\hat{t} \leq \tilde{C}(C_{\mathrm{un}}, \eta_2) C_{\mathrm{m}} n^2_{\max} 8^\ell,$$

where we further used that the number of boxes $s \in \mathcal{L}_Y$ at level ℓ is bounded by 8^ℓ. Therefore we can estimate $M^+_{\mathrm{D},Y}$ by

$$M^+_{\mathrm{D},Y} = \sum_{\substack{\ell\in\{0,\dots,p_{\min}\}\\ \kappa q_\ell\leq 1}} M^+_{\mathrm{D},Y,\ell} \leq \sum_{\ell=0}^{p_{\min}} \tilde{C}(C_{\mathrm{un}}, \eta_2) C_{\mathrm{m}} n^2_{\max} 8^\ell \leq \tilde{C}(C_{\mathrm{un}}, \eta_2) C_{\mathrm{m}} n^2_{\max} 8^{p_{\min}+1}.$$

Finally we get

$$M_{\mathrm{D},Y} = M^-_{\mathrm{D},Y} + M^+_{\mathrm{D},Y} \leq C_{\mathrm{in}} n_{\max} N_X + \tilde{C}(C_{\mathrm{un}}, \eta_2) C_{\mathrm{m}} n^2_{\max} 8^{p_{\min}+1}.$$

In the same manner we can show that

$$M_{\mathrm{D},X} \leq C_{\mathrm{in}} n_{\max} N_Y + \tilde{C}(C_{\mathrm{un}}, \eta_2) C_{\mathrm{m}} n_{\max}^2 8^{p_{\min}+1},$$

where $M_{\mathrm{D},X}$ is the number of matrix entries corresponding to all blocks $(t, s) \in \mathcal{L}_{X\times Y}$ with $t \in \mathcal{L}_X$. The assertion (2.77) follows by summation of the estimates on $M_{\mathrm{D},Y}$ and $M_{\mathrm{D},X}$ with the constant $\hat{C}(C_{\mathrm{un}}, \eta_2) = 2\tilde{C}(C_{\mathrm{un}}, \eta_2)$.

To show the linear complexity of Algorithm 2.33 it suffices to estimate

$$8^{p_{\min}} \leq 8^{C_{\mathrm{ad}}+1} N$$

in (2.77), replace N_X and N_Y by N, and collect the constants. □

Remark 2.55. We have already pointed out in Remark 2.49 that the assumption $p_{\min} \leq \log_8(N) + C_{\mathrm{ad}}$ is generally unsuitable for points distributed on a 2D manifold. Again, we refer to [5, Section 5] for a detailed discussion of the 2D setting.

The following theorem concerning the complexity of Algorithm 2.42 is a summary of the main results of this section.

Theorem 2.56. *Let the matrix A be defined in* (2.1) *and P_X and P_Y be the corresponding sets of points with cardinalities N_X and N_Y. Let $N := \max\{N_X, N_Y\}$. Let $\mathcal{T}_X$ and $\mathcal{T}_Y$ be the uniform box cluster trees constructed by Algorithm 2.3 for the sets P_X and P_Y and $\mathcal{T}_{X\times Y}$ the corresponding block tree constructed by Algorithm 2.26 for some $\eta_2 > 0$. Assume that there exists a constant $C_{\mathrm{tree}} > 0$ such that the construction of the trees $\mathcal{T}_X$, $\mathcal{T}_Y$, and $\mathcal{T}_{X\times Y}$ does not require more than $C_{\mathrm{tree}} N \log(N)$ units of storage and floating point operations. Assume further that the assumptions of Theorems 2.48 and 2.54 hold. Then the run-time and storage complexity of Algorithm 2.42 is of order $\mathcal{O}(N \log(N))$.*

Proof. Step one and three of Algorithm 2.42, i.e. the construction of the uniform box cluster trees $\mathcal{T}_X$ and $\mathcal{T}_Y$ and the block tree $\mathcal{T}_{X\times Y}$, have the desired complexity by assumption. The same holds for steps four and five due to Theorems 2.48 and 2.54. Hence, it remains to show that step two in Algorithm 2.42 has a complexity of $\mathcal{O}(N \log(N))$. As every direction can be constructed with $\mathcal{O}(1)$ floating point operations it suffices to estimate the number of all constructed directions.

For $\ell > \ell_{\mathrm{hf}}$ there exists only the trivial direction 0 which has to be stored only once. For all other levels $\ell \leq \ell_{\mathrm{hf}}$ the number of directions is given by $6 \cdot 4^{\ell_{\mathrm{hf}}-\ell}$. Hence, in total we need to compute and store at most

$$1 + \sum_{\ell=0}^{\ell_{\mathrm{hf}}} 6 \cdot 4^{\ell_{\mathrm{hf}}-\ell} = 1 + \frac{6 \cdot (4^{\ell_{\mathrm{hf}}+1} - 1)}{4 - 1} \leq 2 \cdot 4^{\ell_{\mathrm{hf}}+1}$$

directions. By assumptions (2.56) and (2.68b) we have $\ell_{\mathrm{hf}} + 1 \leq p_{\max} \leq \log_8(N) + C_{\mathrm{ad}}$. Therefore the number of directions is of order $\mathcal{O}(N)$, since

$$2 \cdot 4^{\ell_{\mathrm{hf}}+1} \leq 2 \cdot 4^{C_{\mathrm{ad}}+\log_8(N)} \leq 2 \cdot 4^{C_{\mathrm{ad}}+\log_4(N)} = 2 \cdot 4^{C_{\mathrm{ad}}} \cdot N.$$ □

3 Numerical Experiments

The fast directional matrix vector multiplication in Algorithm 2.42 allows us to approximate the matrix-vector product of a matrix $A \in \mathbb{C}^{N_X \times N_Y}$ as in (2.1) and a vector $v \in \mathbb{C}^{N_Y}$. The required number of floating point operations for this algorithm is of order $\mathcal{O}(N \log(N))$ with $N := \max\{N_X, N_Y\}$ as seen in Theorem 2.56. Hence it is significantly less than the complexity $\mathcal{O}(N_X N_Y)$ of the direct evaluation, if N_X and N_Y are large. Furthermore we have seen that the approximation error converges exponentially to zero with increasing interpolation degree m, which is used for the directional approximation. This interpolation degree m as well as the separation constant η_2 and the highest level in the high-frequency regime ℓ_{hf} have to be chosen for the application of Algorithm 2.42. The choice of the latter two parameters is not immediate. If η_2 is small, then we expect lower approximation errors by Theorem 2.29, but at the same time a higher number of floating point operations for the evaluation due to Theorems 2.48 and 2.54. Similarly large ℓ_{hf} reduce the approximation error by Theorem 2.29 but increase the computational effort, which is discussed in Remark 2.46. This motivates the necessity of numerical experiments and a detailed parameter study, which we present below. To our best knowledge, such a parameter study is missing in literature.

For the numerical experiments we implemented Algorithm 2.42 in C++. The basis of our code is an implementation of a fast multipole method for the Laplace equation presented in [28]. We used the construction of the uniform box cluster trees (cf. Algorithm 2.3) of this implementation and adapted the other existing structures to our needs. In particular we had to adapt the admissibility criteria for the construction of the block tree (cf. Algorithm 2.26) and to construct the sets of directions (cf. Algorithm 2.6). In our implementation we take into account also the details of Section 2.2.1. In particular we store each coupling matrix $A_{c,t,s}$ exactly once and compress it using a partially pivoted ACA. The only other matrices we store are the non-directional transfer matrices $E_{t'}$ introduced in Section 2.2.1. All other matrices appearing in Algorithm 2.38 are used exactly once in a matrix vector multiplication. Hence, we do not store them, but assemble them when needed. We note that all computations in this chapter were executed sequentially on a PC with two Intel Xeon E5620 CPUs running at 2.4 GHz and a total of 24 GiB of RAM.

For our numerical tests we focused on matrices $A \in \mathbb{C}^{N \times N}$ with entries

$$A[j,k] = \begin{cases} \frac{\exp(i\kappa|x_j - x_k|)}{4\pi|x_j - x_k|} & \text{if } j \neq k, \\ 0 & \text{if } j = k, \end{cases} \tag{3.1}$$

where $P_X = \{x_j\}_{j=1}^N$ is a set of points in $\mathbb{R}^3$. These matrices correspond to those in (2.1), if the sets of points P_X and P_Y are chosen to be equal and the diagonal of the matrix is set to zero to get rid of the singularities. We note that Algorithm 2.42 is also applicable for a matrix A as in (3.1), because the change of the diagonal affects only the computation of the nearfield part g^- of the matrix vector product $g = Av$, which is evaluated directly in Algorithm 2.42. Matrices of similar shape are of particular interest for the solution of boundary value problems by means of boundary element methods. In this setting the points in P_X are scattered on a 2D manifold in $\mathbb{R}^3$, which is why we focus on such sets of points.

In the following sections we consider matrices $A \in \mathbb{C}^{N\times N}$ defined by (3.1) for various sets of points P_X and compute the matrix-vector product $g = Av$ directly and approximately using the fast directional matrix vector multiplication in Algorithm 2.42 for some vectors $v \in \mathbb{C}^N$. In Section 3.1 we choose the points on the surface of the cube $[-1, 1]^3$. We investigate how the choice of parameters influences the approximation quality, the memory requirements, and the runtimes of the fast directional matrix-vector mlieultiplication. For this purpose we execute Algorithm 2.42 for various choices of the parameters m, η_2, and ℓ_{hf} and compare the resulting vector $\hat{g}$ with the exact result $g = Av$ by computing the relative error $\|g - \hat{g}\|_2/\|g\|_2$. The comparison of the runtimes and approximation qualities in Sections 3.1.2 and 3.1.3 motivates a parameter selection strategy for η_2 and ℓ_{hf} which we discuss in Section 3.1.4. In Section 3.2 we check if the parameters chosen according to these parameter choice strategies are reasonable also in the case of points distributed on the sphere $S^2 = \{x \in \mathbb{R}^3 : |x| = 1\}$.

We conclude the chapter by presenting computation results of two implementations of Algorithm 2.42: one in which all stored coupling matrices $A_{c,t,s}$ are compressed via ACA and one in which they are not. By comparing the relevant runtimes, approximation qualities, and the required storage for the coupling matrices we show that compression strategies are crucial for a good performance of Algorithm 2.42.

Before we start with the parameter study we note that in other works the parameter η_1 from the criterion (A2) is chosen instead of ℓ_{hf} as a parameter for the algorithm (cf. [5, 7]). However, these two parameters are closely connected, as seen in Theorem 2.19.

3.1 Parameter study on the surface of a cube

To gain a better understanding of the effect of the parameters η_2 and ℓ_{hf} on Algorithm 2.42 it makes sense to consider a simple geometry for a basic parameter study. Hence, we consider only points on the surface of the cube $[-1, 1]^3$ throughout this section. In Section 3.1.1 we describe a way to construct sets of points $P_X(L)$ whose corresponding box cluster trees $\mathcal{T}_X$ are balanced. Then we start our parameter study by first varying the separation constant η_2 in some numerical experiments in Section 3.1.2. This gives us an insight into suitable choices of η_2. We repeat the same for

the parameter ℓ_{hf} in Section 3.1.3. In Section 3.1.4 we finally summarize our observations and discuss a possible parameter selection strategy. We note that the results in Sections 3.1.2 and 3.1.3 are only a suitable selection of a large number of numerical experiments which we conducted.

3.1.1 Construction of the point sets

To make the interpretation of the results easier we construct a set of points P_X on the surface of $[-1,1]^3$ such that the uniform box cluster tree $\mathcal{T}_X(P_X)$ constructed by Algorithm 2.3 is balanced. In particular we want that all leaves of the tree are on the same level and contain a similar number of points. For this purpose we fix the number $n_{\max} = 150$ of maximally allowed points in a leaf. We choose this number since it guarantees that inadmissible blocks in the later matrix partition are reasonably small, and at the same time admissible blocks corresponding to boxes at leaf level are still large enough such that it is reasonable to approximate them. Then we choose a level $L \in \mathbb{N}$ for the leaves and divide each face of the surface $\partial([-1,1]^3)$ uniformly into 4^L squares. On any square which does not touch an edge or corner of $\partial([-1,1]^3)$ we distribute 12×12 points on an equidistant grid on the inside of this square, to ensure that the resulting points are pairwise disjoint. If a square touches an edge but no corner then we distribute only 10×7 points on this square and in case it touches a corner only 7×7. We denote the resulting set of points by $P_X(L)$. If we use Algorithm 2.3 to construct a uniform cluster tree $\mathcal{T}_X$ for the set of points $P_X(L)$ and choose the box $X = [-1,1]^3$ as initial box, then all of the leaves of $\mathcal{T}_X$ are at level L and contain at least 140 and at most 147 points. The set $P_X(2)$ is depicted in Figure 3.1.

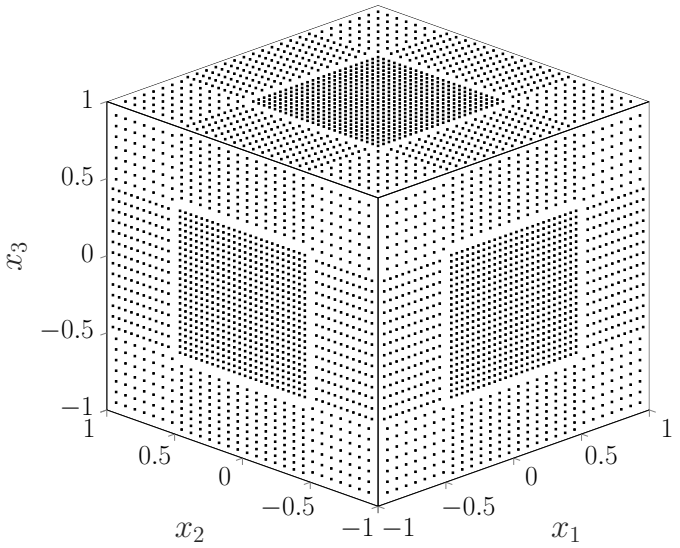

Figure 3.1: The set of points $P_X(2)$ on $\partial([-1,1]^3)$

One could argue that this choice of points is inappropriate, because the regularity of the points could lead to unrealistically low approximation errors due to the interpolation schemes used for the directional approximation (cf. Section 1.2.1). However, we use Chebyshev nodes as interpolation points whereas the points in $P_X(L)$ lie on equidistant grids. Hence, the points in $P_X(L)$ do not coincide with the interpolation points in general and we can assume that this choice of points does not have a strong impact on the approximation errors.

In the following sections we choose the wave number κ depending on the sets of points $P_X(L)$. The choice $\kappa = 3.14 \cdot 2^{L-2}$ seems to be reasonable for the set $P_X(L)$ for a fixed L. For this wave number the wavelength $\lambda = 2\pi/\kappa$ is approximately 2^{3-L}. By construction of $P_X(L)$ there are about 12 points per length 2^{1-L} and hence 48 points per wavelength. This suffices to resolve the waves of the oscillatory factor $\exp(i\kappa|x-y|)$ sufficiently good, which is typically required in applications.

3.1.2 Variation of the separation parameter

We want to investigate the influence of the separation parameter η_2 on the fast directional matrix vector multiplication in Algorithm 2.42. For this purpose we consider the points $P_X(4)$ constructed in Section 3.1.1, whose cardinality is $N = 194040$, and the corresponding box cluster tree $\mathcal{T}_X$. We choose the wave number $\kappa = 12.56$. Additionally, we fix the highest level $\ell_{\text{hf}} = 3$ in the high frequency regime, i.e. we let only the leaf level of $\mathcal{T}_X$ be in the low frequency regime. Then we construct the block tree $\mathcal{T}_{X\times X}(\eta_2)$ by Algorithm 2.26 for various choices of η_2. The notation $\mathcal{T}_{X\times X}(\eta_2)$ is used to emphasize the dependence on the parameter η_2. Finally we compute the approximation $\hat{g}$ of the matrix-vector product $g = Av$ for the matrix $A \in \mathbb{C}^{N\times N}$ in (3.1) and a vector $v \in \mathbb{C}^N$, which we choose randomly but fixed for all the computations. This approximation consists of the approximate farfield part $\hat{g}^+$ which is computed by Algorithm 2.38 and the nearfield part g^- computed by Algorithm 2.33. The choice of η_2 influences both parts.

Table 3.1 contains the numbers of admissible leaf blocks $\mathcal{L}^+_{X\times X}(\eta_2)$ at the relevant levels $\ell \in \{2, ..., 4\}$ in $\mathcal{T}_{X\times X}(\eta_2)$, their sum, and the percentage of entries of the matrix A, which are used for the compuation of the nearfield part g^- in Algorithm 2.33, for some values of η_2. These values were chosen suitably to show the effect of the change of η_2.

The nearfield percentage is given in the last column of the table. We see that for increasing η_2 it decreases until it reaches the fixed value 0.7202 for $\eta_2 \geq 5$. The reason is that for increasing η_2 the admissibility criteria (A1) and (A3) become less restrictive, i.e. more and more pairs of boxes (t, s) with t and s in $\mathcal{T}_X$ become admissible and as a consequence the nearfield percentage decreases. If we choose η_2 large enough only pairs (t, s) satisfying $\operatorname{dist}(t, s) = 0$ remain inadmissible. This is the case for $\eta_2 = 5$ in Table 3.1. Increasing η_2 even further has no more effect on the nearfield percentage, as these remaining pairs of boxes can never be admissible.

Table 3.1: Number of admissible leaf blocks in $\mathcal{T}^+_{X\times X}(\eta_2)$ (levelwise and total) and the percentage (near. %) of matrix entries used to compute the nearfield part g^- in Algorithm 2.33; (set of points $P_X(4)$, $\kappa = 12.56$)

η_2	level 2	level 3	level 4	total	near. %
1	0	56	1649400	1649456	9.6439
2	0	47360	724200	771560	3.7384
3	0	65912	393072	458984	2.0736
4	0	71648	286848	358496	1.6923
5	0	78392	170592	248984	0.7202
6	8	78240	166368	244616	0.7202
7	176	73080	132192	205448	0.7202
10	1352	41472	49536	92360	0.7202
19	2504	12960	49536	65000	0.7202

Next we consider columns two to five of Table 3.1. The first three of these columns contain the number of admissible leaf blocks in $\mathcal{L}^+_{X\times X}$ counted at each of the levels $\ell \in \{2, ..., 4\}$ separately, and the last their sum. We note that this sum corresponds to the total number of M2L operations used for the computation of the farfield part $\hat{g}^+$ in Algorithm 2.38. For increasing η_2 the total number of admissible leaf blocks decreases strictly. Again the reason is that for increasing η_2 the admissibility criteria (A1) and (A3) are less restrictive. As a consequence non-leaf blocks in $\mathcal{T}_{X\times X}(\eta_2)$ can become admissible and are transformed into leaves by truncating the underlying subtrees. This reduces the total number of leaves in $\mathcal{T}_{X\times X}$ drastically, as every block (t, s) in $\mathcal{T}_{X\times X}$ has at least 16 children. Indeed every non-leaf box t in $\mathcal{T}_X(4)$ has at least four children due to the uniformity of the construction of the set $P_X(4)$ and the corresponding box cluster tree $\mathcal{T}_X$ in Section 3.1.1. Hence, the children of a block (t, s), which are formed by all pairs of children of t and s, are at least 16. Another effect of the increase of η_2 is that previously inadmissible leaves in $\mathcal{T}_{X\times X}(\eta_2)$ can become admissible, which leads to an increase of the number of admissible blocks. However, the reduction of the total number of leaves outweighs this transformation of inadmissible blocks in general. This explains the decreasing behavior of the total number of admissible blocks. Similarly one can explain the steady decrease of the number of admissible blocks at level four, the initial increase and later decrease at level three and the increase at level two.

The numbers in Table 3.1 suggest to choose η_2 not too small. Indeed, we see that for $\eta_2 = 1$ the nearfield percentage is almost 10, and the total number of blocks in the block tree, which correspond to the subblocks in which we partition the matrix A, is 1649456, which is more than $8N$. Optimal in terms of the total number of admissible blocks is the choice $\eta_2 = 19$. However, we have not yet considered the approximation quality. For this purpose we compute the approximations $\hat{g}$ of the matrix vector product $g = Av$ for the matrix A in (3.1) and a given vector v for several choices of the separation parameter η_2 and interpolation degrees m. The resulting relative errors

are presented in Table 3.2. We note that these errors are in general also influenced by the choice of the stopping criterion used for the compression of the coupling matrices by means of the ACA. We used the stopping criterion in [33, Algorithm 3.9, Step 3.5] with $\varepsilon_{\mathrm{ACA}} = 10^{-5}$. A comparison between the errors in the compressed case and the uncompressed case showes that the influence of the compression on the errors in Table 3.2 is neglectible for this stopping criterion and our choices of parameters. We note that this choice of $\varepsilon_{\mathrm{ACA}}$ is not optimal for all the following computations. In particular for some choices of η_2 and m we could choose larger $\varepsilon_{\mathrm{ACA}}$ and achieve a better compression without disturbing the approximation quality. Only for the sake of simplicity we consider the fixed parameter $\varepsilon_{\mathrm{ACA}} = 10^{-5}$ as long as it does not influence the approximation quality.

Table 3.2: Relative errors $\|g - \hat{g}\|_2/\|g\|_2$ for various separation parameters η_2 and interpolation degrees m and the estimated and predicted rate of convergence (erc and prc); (set of points $P_X(4)$, $\kappa = 12.56$, $\ell_{\mathrm{hf}} = 3$, $\varepsilon_{\mathrm{ACA}} = 10^{-5}$)

η_2	$m = 2$	$m = 3$	$m = 4$	$m = 5$	erc	prc
1	2.19_{-2}	2.13_{-3}	1.63_{-4}	1.71_{-5}	0.10	0.33
2	2.30_{-2}	2.37_{-3}	2.21_{-4}	2.40_{-5}	0.11	0.50
3	2.47_{-2}	2.53_{-3}	2.41_{-4}	2.51_{-5}	0.10	0.60
4	2.70_{-2}	2.81_{-3}	2.95_{-4}	3.23_{-5}	0.11	0.67
5	2.81_{-2}	3.31_{-3}	3.17_{-4}	3.22_{-5}	0.10	0.71
6	2.85_{-2}	3.35_{-3}	2.92_{-4}	3.17_{-5}	0.11	0.75
7	3.83_{-2}	6.43_{-3}	8.40_{-4}	1.11_{-4}	0.13	0.78
10	6.41_{-2}	1.12_{-2}	2.17_{-3}	3.71_{-4}	0.17	0.83
19	8.46_{-2}	1.90_{-2}	6.41_{-3}	1.60_{-3}	0.25	0.90

In Table 3.2 we observe exponential convergence of the relative errors with increasing interpolation degree m for all choices of the separation parameter η_2. This is what we expect from Theorem 1.15 and Remark 2.43. The last two columns in Table 3.2 contain an estimate of the rate of convergence of the relative errors, which we compute by taking the quotient of the relative errors for $m = 5$ and $m = 4$, and a prediction of this rate by Theorem 1.15. Although the estimate is very rough, it suggests that the rate of convergence is much better than the predicted rate $1/\hat{\rho} = 1/(1 + 2/\eta_2)$. The qualitative statements of Theorem 1.15 seem to apply, however. In particular, the errors for a fixed interpolation degree m increase with increasing η_2. Only the errors for $\eta_2 = 6$ and $m = 4$ and $m = 5$, respectively, step out of row, as they are lower than the respective errors for $\eta_2 = 5$. Similarly the rate of convergence seems to increase with increasing η_2, in particular for larger values of η_2.

All in all the results in Table 3.2 suggest to choose η_2 rather small. While the approximation errors do not differ to much for $\eta_2 \leq 6$ they get significantly worse for higher η_2. On the other hand we recall that the numbers in Table 3.1 suggested to

choose η_2 as high as possible. Indeed, we observed that the percentage of nearfield operations reaches its minimum for $\eta_2 \geq 5$ and that the total number of admissible blocks is strictly decreasing for increasing η_2. Hence by Tables 3.1 and 3.2 the choices $\eta_2 = 5$ and $\eta_2 = 6$ seem to be reasonable.

Finally we want to check if these choices are also reasonable in regard to the computation times. In Table 3.3 we show the computation times needed for the evaluation of Algorithm 2.42 for several choices of the separation parameter η_2. In regard to Table 3.2 we consider the interpolation degree $m = 3$ if $\eta_2 \leq 7$, and $m = 4$ if $\eta_2 > 7$. These choices lead to relative errors which are of similar magnitude and can therefore be compared well.

Table 3.3: Time in seconds required for the computation of the fast directional matrix-vector product $\hat{g}$ for several choices of η_2 and m and the set of points $P_X(4)$. The columns two to four contain the times needed for the setup, the computation of the nearfield part g^-, and the computation of the approximate farfield part $\hat{g}^+$. Column five contains the time required for the computation of $\hat{g}$ after setup, i.e. the sum of columns three and four. The last line contains the time required for the exact computation of $g = Av$. ($\kappa = 12.56$, $\ell_{\text{hf}} = 3$, $\varepsilon_{\text{ACA}} = 10^{-5}$)

(η_2, m)	setup	g^-	$\hat{g}^+$	$\hat{g}$
$(1, 3)$	13.35	361.20	3.54	364.75
$(2, 3)$	7.39	135.93	2.07	138.01
$(3, 3)$	5.25	74.23	1.51	75.75
$(4, 3)$	4.71	59.65	1.28	60.94
$(5, 3)$	4.25	23.40	1.08	24.49
$(6, 3)$	5.93	23.38	1.08	24.46
$(7, 3)$	5.37	23.39	0.98	24.38
$(10, 4)$	6.77	23.38	1.32	24.71
$(19, 4)$	5.09	23.39	1.17	24.57
exact				3810.26

The setup times in Table 3.3 include the times needed for the construction of the uniform box cluster tree $\mathcal{T}_X$ and the block tree $\mathcal{T}_{X\times X}(\eta_2)$ as well as the times to assemble the non-directional transfer matrices $E_{t'}$ introduced in Section 2.2.1 and to assemble the compressed coupling matrices $A_{c,t,s}$. We present them to show that they are not extraordinarily high, but in the range of the computation times in column three and four. If the matrix A in (3.1) is used several times for matrix-vector multiplications this setup is only required once. Therefore, we neglect the setup times in the sums in column five.

The nearfield computation times, i.e. the times required for the computation of g^-, are given in column three of Table 3.3. Let us compare them with the time needed for the computation of the exact matrix-vector product $g = Av$ which is given in the

last line of the same table. We see that they behave as we expect when considering the nearfield percentages in Table 3.1. In particular, the nearfield computation times are constant for $\eta_2 \geq 5$, and attain their minimum for such η_2. In comparison to the farfield computation times the nearfield computation times are significantly higher. Hence, it seems to be reasonable to choose η_2 such that the nearfield computation time is minimal.

Let us finally consider the farfield computation times in column five of Table 3.3. As expected we observe a decrease of these times for increasing η_2 and fixed $m = 3$ up to $\eta_2 = 7$. The slight increase of the farfield computation time for $\eta_2 = 10$ and $m = 4$ is what we expect due to the increase of the interpolation degree. Overall, the farfield computation times are in a similar range for $\eta_2 \geq 4$.

In summary, we observe that Tables 3.1 and 3.3 give similar insights into the choice of η_2. Both suggest to choose η_2 sufficiently large. In particular the choice $\eta_2 \geq 5$ seems to be reasonable as for such η_2 the nearfield percentage is minimized. Together with the results in Table 3.2 on the approximation quality we see that the choices $\eta_2 = 5$ and $\eta_2 = 6$ seem to be preferable. For the sake of completeness we list the relative errors and farfield times for these parameters and interpolation degrees $m \in \{2, ..., 5\}$ in Table 3.4.

Table 3.4: Relative errors $\|g - \hat{g}\|_2 / \|g\|_2$ and computation times in seconds for the farfield part $\hat{g}$ for $\ell_{\text{hf}} = 3$, $\eta_2 \in \{5, 6\}$, and $m \in \{2, ..., 5\}$; (set of points $P_X(4)$, $\kappa = 12.56$, $\varepsilon_{\text{ACA}} = 10^{-5}$)

$\ell_{\text{hf}} = 3$	$\eta_2 = 5$		$\eta_2 = 6$	
m	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$
2	2.81_{-2}	0.47	2.85_{-2}	0.47
3	3.31_{-3}	1.08	3.35_{-3}	1.08
4	3.17_{-4}	2.12	2.92_{-4}	2.12
5	3.22_{-5}	3.86	3.17_{-5}	3.87

Up to this point we have only talked about the set of points $P_X(4)$ and the corresponding matrix-vector multiplication. Let us now consider the set $P_X(5)$ with cardinality $N = 829176$. As mentioned in Section 3.1.1 we choose the wave number $\kappa = 25.12$ for this set of points. Furthermore we set the highest level in the high frequency regime $\ell_{\text{hf}} = 4$ to end up again with only one level in the low frequency regime. Analogously as for the set $P_X(4)$ we consider the matrix A in (3.1) and fix a randomly constructed vector $v \in \mathbb{C}^N$ for the matrix-vector multiplication. As before our goal is to find a suitable separation constant η_2 for the fast directional matrix-vector multiplication. Therefore, we used Algorithm 2.42 for the computation of this fast directional matrix-vector product $\hat{g}$ for several choices of η_2 and interpolation degrees m and compared them with the exact result $g = Av$. The results are given in Table 3.5.

Table 3.5: Relative errors $\|g - \hat{g}\|_2/\|g\|_2$ for various separation parameters η_2 and interpolation degrees m and the estimated and predicted rate of convergence (erc and prc); (set of points $P_X(5)$, $\kappa = 25.12$, $\ell_{\text{hf}} = 4$, $\varepsilon_{\text{ACA}} = 10^{-5}$)

η_2	$m = 2$	$m = 3$	$m = 4$	$m = 5$	erc	prc
4	4.81_{-2}	7.04_{-3}	9.17_{-4}	1.36_{-4}	0.15	0.67
5	4.80_{-2}	8.17_{-3}	1.12_{-3}	1.64_{-4}	0.15	0.71
6	4.84_{-2}	8.38_{-3}	1.14_{-3}	1.72_{-4}	0.15	0.75
7	5.06_{-2}	9.31_{-3}	1.45_{-3}	2.19_{-4}	0.15	0.78
10	6.07_{-2}	1.10_{-2}	2.54_{-3}	3.86_{-4}	0.15	0.83

This time we have only considered a smaller number of separation parameters η_2 due to our findings for the set of points $P_X(4)$. Indeed, we have seen that choosing η_2 too small leads to large computation times, while choosing η_2 too large leads to large errors. In Table 3.5 we consider the approximation quality. We observe again an increase of the approximation errors and a decrease of the estimated rates of convergence for increasing η_2. The differences are only small, however. Only for $\eta_2 = 10$ and $m \in \{4, 5\}$ the relative errors are about a factor two to three worse than the corresponding errors for the other choices of η_2.

In Table 3.6 we show the computation times which are needed for the setup and computation of the fast directional matrix-vector product $\hat{g}$ as well as the time for the exact computation of $g = Av$. Similarly as in Table 3.3 we choose the interpolation degree m such that a similar approximation quality is achieved for all choices of η_2. Choosing $m = 4$ for all η_2 seems to be reasonable for this purpose.

Table 3.6: Time in seconds required for the computation of the fast directional matrix-vector product $\hat{g}$ for several choices of η_2 and m and the set of points $P_X(5)$. The columns two to four contain the times needed for the setup, the computation of the nearfield part g^-, and the computation of the approximate farfield part $\hat{g}^+$. Column five contains the time required for the computation of $\hat{g}$ after setup, i.e. the sum of columns three and four. The last line contains the time required for the exact computation of $g = Av$. ($\kappa = 25.12$, $\ell_{\text{hf}} = 4$, $\varepsilon_{\text{ACA}} = 10^{-5}$)

η_2	setup	g^-	$\hat{g}^+$	$\hat{g}$
$(4, 4)$	20.45	244.27	13.34	257.63
$(5, 4)$	16.49	95.62	10.89	106.54
$(6, 4)$	14.56	95.92	10.21	106.15
$(7, 4)$	13.07	95.38	9.12	104.52
$(10, 4)$	10.81	95.81	6.50	102.34
exact				70575.70

First we want to emphasize the huge difference between the times needed for the computation of the approximate matrix-vector products, which lie in the range of a few minutes, and the computation of the exact product, which requires more than 19 hours on our PC. This underlines the benefits of the fast directional matrix-vector multiplication. Furthermore the large time required for the exact computation suggests to choose η_2 not to small. If we choose $\eta_2 = 1$, for example, the nearfield percentage, i.e. the percentage of entries of A needed for the computation of the nearfield part g^-, is about 2.1, and we expect the nearfield computation time for $\eta_2 = 1$ to be more than 20 minutes.

Let us further comment on the nearfield computation time. As in the case of the set of points $P_X(4)$ it reaches its minimum for $\eta_2 \geq 5$. Also the farfield computation times behave as in the case before, i.e. they decrease for increasing η_2. The decrease is greater this time. In particular for $\eta_2 = 10$ the farfield computation time is clearly lower compared to the rest.

When considering the results in Tables 3.5 and 3.6 it is hard to speak of an optimal choice of η_2. While choosing $\eta_2 \in \{5, 6\}$ leads to slightly better relative errors choosing $\eta_2 = 10$ is preferable in terms of the computation time for the farfield part $\hat{g}^+$. However, we note that in some applications an error twice as large might enforce a higher interpolation degree. In such a case the choices $\eta_2 = 5$ and $\eta_2 = 6$ are preferable. Nonetheless we collect the relative errors and farfield computation times for all three choices of η_2 in the separate Table 3.7.

Table 3.7: Relative errors $\|g - \hat{g}\|_2/\|g\|_2$ and computation times in seconds for the farfield part $\hat{g}$ for $\ell_{\text{hf}} = 3$, $\eta_2 \in \{5, 6\}$, and $m \in \{2, ..., 5\}$; (set of points $P_X(5)$, $\kappa = 25.12$, $\varepsilon_{\text{ACA}} = 10^{-5}$)

$\ell_{\text{hf}} = 4$	$\eta_2 = 5$		$\eta_2 = 6$		$\eta_2 = 10$	
m	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$
2	4.80_{-2}	2.35	4.84_{-2}	2.16	6.07_{-2}	1.22
3	8.17_{-3}	5.49	8.38_{-3}	5.12	1.10_{-2}	3.14
4	1.12_{-3}	10.89	1.14_{-3}	10.21	2.54_{-3}	6.50
5	1.64_{-4}	20.43	1.72_{-4}	19.62	3.86_{-4}	13.07

By the computations and discussions above we have now a basic idea on how to choose the separation parameter η_2. What we have not considered yet is how the parameter ℓ_{hf} influences the fast directional matrix-vector multiplication in applications. This is the topic of the following section.

3.1.3 Variation of the number of high frequency levels

In Section 3.1.2 we have chosen the highest level in the high frequency regime $\ell_{\text{hf}} = 3$ for the set of points $P_X(4)$ and $\ell_{\text{hf}} = 4$ for $P_X(5)$. Here we want to investigate if these choices are reasonable. For this purpose we repeat some of the computations from Section 3.1.2 for other choices of ℓ_{hf}. Again, we consider first the set of points $P_X(4)$ and the wave number $\kappa = 12.56$.

Table 3.8: Relative errors $\|g - \hat{g}\|_2/\|g\|_2$ and computation times in seconds for the farfield part $\hat{g}$ for $\ell_{\text{hf}} = 2$ and several choices of m and η_2; (set of points $P_X(4)$, $\kappa = 12.56$, $\varepsilon_{\text{ACA}} = 10^{-5}$)

$\ell_{\text{hf}} = 2$	$\eta_2 = 2$		$\eta_2 = 5$		$\eta_2 = 6$	
m	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$
2	1.46_{-1}	1.04	1.67_{-1}	0.49	1.69_{-1}	0.44
3	2.67_{-2}	2.04	3.09_{-2}	1.11	3.15_{-2}	1.02
4	4.19_{-3}	3.86	4.83_{-3}	1.97	5.00_{-3}	1.95
5	5.60_{-4}	7.10	6.43_{-4}	3.43	6.80_{-4}	3.41

In Table 3.8 we see some computational results for $\ell_{\text{hf}} = 2$. We only list the relative errors and farfield computation times, since the choice of ℓ_{hf} does not influence the nearfield computation times and the setup time only slightly due to the different sets of directions which have to be constructed. The separation parameters $\eta_2 = 5$ and $\eta_2 = 6$ are considered, since these are the reasonable choices for $\ell_{\text{hf}} = 3$ as seen in Section 3.1.2. The results for $\eta_2 = 2$ are listed for comparison, as we expect the accuracy to be better for this choice of η_2. The results for η_1 would be misleading. The reason is that for $\eta_2 = 1$ there are almost no admissible blocks at level three in the box cluster tree, which we see in Table 3.1. Changing ℓ_{hf} from three to two, however, affects only these blocks.

If we compare the relative errors in Table 3.4 for $\ell_{\text{hf}} = 3$ and $\eta_2 \in \{5, 6\}$ with their counterparts in Table 3.8 for $\ell_{\text{hf}} = 2$, then we note that the approximation errors are significantly worse for $\ell_{\text{hf}} = 2$. At the same time the computation times decrease only slightly if we choose $\ell_{\text{hf}} = 2$ instead of $\ell_{\text{hf}} = 3$. The same holds for $\eta_2 = 2$, which can be seen by comparing the results in Table 3.8 with the corresponding results in Tables 3.2 and 3.3. Hence, the choice $\ell_{\text{hf}} = 2$ appears to be inappropriate for the considered data.

In Table 3.9 we see the relative errors and farfield computation times for the levels $\ell_{\text{hf}} = 4$ and $\ell_{\text{hf}} = 5$ and the separation parameters $\eta_2 \in \{1, 5, 6\}$. For $m = 4$ and $m = 5$ we adapted the parameter ε_{ACA} for the stopping criterion of the ACA, to end up with a similar accuracy as in the uncompressed case. In particular we have chosen $\varepsilon_{\text{ACA}} = 10^{-6}$ for $m = 4$, and $\varepsilon_{\text{ACA}} = 10^{-7}$ for $m = 5$ instead of $\varepsilon_{\text{ACA}} = 10^{-5}$. The results for $\eta_2 = 1$ are used for comparison purposes. The relative errors for $\eta_2 = 1$ and $\ell_{\text{hf}} \in \{4, 5\}$ are considerably lower than the corresponding errors for $\eta_2 = 1$ and

Table 3.9: Relative errors $\|g - \hat{g}\|_2/\|g\|_2$ and computation times in seconds for the farfield part $\hat{g}$ for $\ell_{\text{hf}} = 4$ and $\ell_{\text{hf}} = 5$, several choices of η_2, and the interpolation degree $m \in \{2, ..., 5\}$; (set of points $P_X(4)$, $\kappa = 12.56$)

$\ell_{\text{hf}} = 4$	$\eta_2 = 1$		$\eta_2 = 5$		$\eta_2 = 6$	
$(m, \varepsilon_{\text{ACA}})$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$
$(2, 10^{-5})$	1.64_{-3}	2.06	1.31_{-2}	0.98	1.31_{-2}	0.98
$(3, 10^{-5})$	1.19_{-4}	4.17	2.09_{-3}	1.81	2.10_{-3}	1.80
$(4, 10^{-6})$	7.72_{-6}	9.56	3.25_{-4}	3.99	3.26_{-4}	4.02
$(5, 10^{-7})$	3.97_{-7}	20.90	3.63_{-5}	8.52	3.65_{-5}	8.50
$\ell_{\text{hf}} = 5$	$\eta_2 = 1$		$\eta_2 = 5$		$\eta_2 = 6$	
$(m, \varepsilon_{\text{ACA}})$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$
$(2, 10^{-5})$	1.14_{-3}	2.95	1.07_{-2}	2.04	1.07_{-2}	2.23
$(3, 10^{-5})$	1.25_{-4}	5.55	1.60_{-3}	3.78	1.62_{-3}	3.77
$(4, 10^{-6})$	8.08_{-6}	11.79	2.31_{-4}	7.57	2.32_{-4}	7.57
$(5, 10^{-7})$	5.44_{-7}	24.01	2.03_{-5}	15.65	2.05_{-5}	15.71

$\ell_{\text{hf}} = 3$ in Table 3.2. Unfortunately, this does not hold for $\eta_2 \in \{5, 6\}$. For $\ell_{\text{hf}} = 4$ and $m \geq 4$ the relative errors are even slightly worse, while for $\ell_{\text{hf}} = 5$ they are only slightly better. The farfield computation times on the other hand increase clearly. Hence, the choice $\ell_{\text{hf}} = 3$ which we considered in Section 3.1.2 seems to be reasonable for the set of points $P_X(4)$.

Next we consider the set $P_X(5)$ for varying ℓ_{hf} to see if also for this set of points the choice $\ell_{\text{hf}} = 4$ of Section 3.1.2 is preferable. We start by considering the results for $\ell_{\text{hf}} = 3$ in Table 3.10.

Table 3.10: Relative errors $\|g - \hat{g}\|_2/\|g\|_2$ and computation times in seconds for the farfield part $\hat{g}$ for $\ell_{\text{hf}} = 3$, $\eta_2 \in \{5, 6, 10\}$, and $m \in \{2, ..., 5\}$; (set of points $P_X(5)$, $\kappa = 25.12$, $\varepsilon_{\text{ACA}} = 10^{-5}$)

$\ell_{\text{hf}} = 3$	$\eta_2 = 5$		$\eta_2 = 6$		$\eta_2 = 10$	
m	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$
2	1.85_{-1}	2.16	1.83_{-1}	2.14	1.87_{-1}	1.05
3	3.70_{-2}	5.03	3.69_{-2}	4.67	3.69_{-2}	2.66
4	6.43_{-3}	9.63	6.43_{-3}	8.92	6.71_{-3}	5.16
5	1.01_{-3}	17.02	1.01_{-3}	15.82	1.06_{-3}	9.30

Let us compare this results with those in Table 3.7. As in the case of the set $P_X(4)$ we see that reducing ℓ_{hf} leads to significantly larger errors while the computation times decrease in comparison only slightly. We draw the conclusion that for $P_X(5)$ and $\kappa = 25.12$ we should choose the level $\ell_{\text{hf}} \geq 4$.

In Table 3.11 we show the computational results for the choices $\ell_{\text{hf}} = 5$ and $\ell_{\text{hf}} = 6$.

Table 3.11: Relative errors $\|g - \hat{g}\|_2/\|g\|_2$ and computation times in seconds for the farfield part $\hat{g}$ for $\ell_{\text{hf}} = 5$ and $\ell_{\text{hf}} = 6$, several choices of η_2, and the interpolation degree $m \in \{2, ..., 5\}$; (set of points $P_X(5)$, $\kappa = 25.12$)

$\ell_{\text{hf}} = 5$	$\eta_2 = 5$		$\eta_2 = 6$		$\eta_2 = 10$	
$(m, \varepsilon_{\text{ACA}})$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$
$(2, 10^{-5})$	1.84_{-2}	4.47	1.87_{-2}	4.31	2.42_{-2}	3.32
$(3, 10^{-5})$	2.74_{-3}	9.38	2.84_{-3}	9.07	4.72_{-3}	6.94
$(4, 10^{-5})$	4.99_{-4}	18.19	5.03_{-4}	17.66	1.13_{-3}	14.03
$(5, 10^{-6})$	4.51_{-5}	39.76	4.71_{-5}	39.07	1.77_{-4}	31.88
$\ell_{\text{hf}} = 6$	$\eta_2 = 5$		$\eta_2 = 6$		$\eta_2 = 10$	
$(m, \varepsilon_{\text{ACA}})$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$
$(2, 10^{-5})$	1.30_{-2}	10.33	1.31_{-2}	10.23	1.75_{-2}	9.32
$(3, 10^{-5})$	1.87_{-3}	19.84	1.91_{-3}	19.63	2.42_{-3}	17.88
$(4, 10^{-5})$	3.50_{-4}	39.04	3.53_{-4}	39.08	6.61_{-4}	37.34
$(5, 10^{-6})$	2.36_{-5}	85.94	2.44_{-5}	86.22	8.24_{-5}	82.64

We have adapted the parameter ε_{ACA} for the stopping criterion of the ACA here again whenever needed. Let us first consider the results for $\ell_{\text{hf}} = 5$ in comparison with the results for $\ell_{\text{hf}} = 4$ in Table 3.7. We observe that the errors for given η_2 and m decrease by a factor between two and four, while the computation times almost double when we choose $\ell_{\text{hf}} = 5$ instead of ℓ_{hf}. This makes a direct comparison difficult. Comparing the results for given η_2 and m for the choice $\ell_{\text{hf}} = 5$ with the results for $\ell_{\text{hf}} = 4$ for the same η_2 but the interpolation degree $m + 1$ simplifies the discussion. Indeed, we see that then the approximation errors for $\ell_{\text{hf}} = 4$ are more than fifty percent lower while the computation times are only slightly higher. Hence, it seems to be preferable to choose $\ell_{\text{hf}} = 4$ and increase the interpolation degree if higher accuracy is needed.

The results for $\ell_{\text{hf}} = 6$ can be compared well with the results for $\ell_{\text{hf}} = 5$. We observe that the relative errors are more or less halved, while the computation times are about doubled for fixed η_2 and m when we increase $\ell_{\text{hf}} = 5$ by one. Increasing the interpolation degree m instead leads also to similar computation times, but far better approximation errors. Hence the choice $\ell_{\text{hf}} = 6$ appears to be worse then the choice $\ell_{\text{hf}} = 5$, and in particular worse then the choice $\ell_{\text{hf}} = 4$. Again, we come to the conclusion that the choice $\ell_{\text{hf}} = 4$ considered in Section 3.1.2 is reasonable for the set of points $P_X(5)$.

3.1.4 A parameter selection strategy

Based on the insights which we gained in the previous two sections we present a strategy to choose the parameters η_2 and ℓ_{hf} for the execution of Algorithm 2.42. This strategy must be understood as a rule of thumb and not a mathematically sound parameter choice rule, as it bases only on the few observations and conclusions we drew from the simple setting of points on the surface of a cube.

We start with the choice of the level ℓ_{hf}. This level is used to distinguish the low and high frequency levels in the box cluster trees in Algorithm 2.42. Boxes at levels $\ell \leq \ell_{\mathrm{hf}}$ are considered to be in a high frequency setting, while boxes at level $\ell > \ell_{\mathrm{hf}}$ are in a low frequency setting. Hence, it seems reasonable to choose ℓ_{hf} not arbitrarily but depending on the size of the boxes, i.e. the geometry, and furthermore on the wave number κ. An implicit distinction of the frequency regimes is already incorporated in the admissibility criteria (A1) and (A3). Indeed, we have seen in Section 1.2.2 that for two boxes t and s the value of $\kappa \max\{\operatorname{diam}(t), \operatorname{diam}(s)\}$ decides which of these criteria is stronger and has to be considered therefore. If

$$\kappa \max\{\operatorname{diam}(t), \operatorname{diam}(s)\} \leq 1,$$

then the standard separation criterion (A1) is stronger, else the newly introduced (high frequency) criterion (A3). This motivates to choose the level ℓ_{hf} depending on the products κq_ℓ, where q_ℓ denotes the diameters of the boxes at level ℓ in a uniform box cluster tree. In particular, it seems reasonable to choose ℓ_{hf} such that

$$\kappa q_{\ell_{\mathrm{hf}}} > C \quad \text{and} \quad \kappa q_{\ell_{\mathrm{hf}}+1} \leq C \tag{3.2}$$

for a suitable constant C. We determine such a constant by investigating the choices of ℓ_{hf} in the above parameter study.

In Section 3.1.2 ℓ_{hf} we chose $\ell_{\mathrm{hf}} = 3$ for the wave number $\kappa = 12.56$ and the points $P_X(4)$ on $\partial([-1,1]^3)$, constructed in Section 3.1.1. For the set $P_X(5)$ and $\kappa = 25.12$ we chose $\ell_{\mathrm{hf}} = 4$. In Section 3.1.3 we observed that these choices are reasonable compared to other choices of ℓ_{hf}. In case of the set $P_X(4)$, the wave number $\kappa = 12.56$, and the level $\ell_{\mathrm{hf}} = 3$ there holds

$$\kappa q_{\ell_{\mathrm{hf}}+1} = 12.56\ q_4 = 12.56 \cdot \frac{\sqrt{3}}{8} \approx 2.72.$$

The same can be shown for the set $P_X(5)$, the wave number $\kappa = 25.12$ and the level $\ell_{\mathrm{hf}} = 4$, i.e. in both cases (3.2) holds for $C \approx 2.72$. In Section 3.1.3 we saw that the choices $\ell_{\mathrm{hf}} = 2$ and $\ell_{\mathrm{hf}} = 4$ for $P_X(4)$ and $\kappa = 12.56$ lead to worse relative errors and worse computation times, respectively. The corresponding constants C for these levels would be 5.44 and 1.36, respectively. Similar effects were observable for $P_X(5)$, $\kappa = 25.12$, and the choices $\ell_{\mathrm{hf}} = 3$ and $\ell_{\mathrm{hf}} = 5$. However, the choice $\ell_{\mathrm{hf}} = 5$ appeared to be only slightly worse than $\ell_{\mathrm{hf}} = 4$ in this example. Hence we suggest to choose ℓ_{hf} such that (3.2) holds for a C close to 2.72, preferring the case $1.36 < C \leq 2.72$. We summarize this in the following remark

Remark 3.1 (Parameter choice rule for ℓ_{hf}). For given κ and a box cluster tree $\mathcal{T}_X$ choose the highest level in the high frequency regime ℓ_{hf} such that (3.2) holds for a constant C close to 2.72, preferably $1.36 < C \leq 2.72$.

A drawback of the above parameter choice rule is the inconsistency of the high frequency regimes introduced by this choice of ℓ_{hf} and the natural distinction of regimes introduced by the admissibility criteria (A1) and (A3). In the case of the set of points $P_X(4)$, the wave number κ, and the level $\ell_{\mathrm{hf}} = 3$, for example, the level $\ell = 4$ satisfies $\ell > \ell_{\mathrm{hf}}$, which means that it is in the low frequency regime corresponding to ℓ_{hf}. However, we use the (high frequency) admissibility criterion (A3) to control the distance of boxes at level four, because $q_4\kappa = 1.36 > 1$ and, therefore, the criterion (A3) is stronger than (A1). Hence, it would be reasonable to substitute (A3) with a criterion like

$$\kappa \max\{\operatorname{diam}(X), \operatorname{diam}(Y)\}^2 \leq \eta_2 q_{\ell_{\mathrm{hf}}+1} \operatorname{dist}(X, Y),$$

for which the frequency regimes would coincide. With this criterion the theoretical results on the approximation quality and the complexity would still hold up to a change of constants. However, this new criterion would influence the construction of the block trees in Algorithm 2.26 and in particular also the choice of the separation constant η_2, which is why we keep the criterion (A3).

Finally we discuss the choice of the separation parameter η_2. We recall that this parameter appears in the admissibility criteria (A1) and (A3) and therefore controls which blocks (t, s) in the block tree $\mathcal{T}_{X\times X}$ are admissible. In Section 3.1.2 we observed that it is crucial for low computation times to choose η_2 such that the number of inadmissible blocks in $\mathcal{T}_{X\times X}(\eta_2)$ is minimal. This means that for the chosen η_2 the only remaining inadmissible blocks (t, s) in $\mathcal{T}_{X\times X}(\eta_2)$ should satisfy $\operatorname{dist}(t, s) = 0$. For general geometries this is probably a too strong requirement. Indeed, if there are leaves at high levels ℓ in the box cluster tree $\mathcal{T}_X$, then there are corresponding leaves at level ℓ in the block tree. These leaves become admissible only if η_2 is chosen considerably high, which can lead to a bad approximation quality. Hence minimizing the number of all inadmissible blocks seems to be inappropriate. However, we can choose η_2 such that the number of all inadmissible blocks at low frequency levels $\ell > \ell_{\mathrm{hf}}$ is minimized. Let us describe how to choose η_2 for this purpose.

In a uniform box cluster tree $\mathcal{T}_X$ the boxes are arranged on a regular grid. Two boxes t and s do not touch if there is at least a box between them in this grid. This means that the minimal distance for two boxes t and s at level ℓ in $\mathcal{T}_X$ with $\operatorname{dist}(t, s) \neq 0$ is one edge length of a box at level ℓ, which is $q_\ell/\sqrt{3}$, where q_ℓ denotes again the diameter of the boxes at level ℓ. Our goal is to choose η_2 such that these boxes are admissible, if $\ell > \ell_{\mathrm{hf}}$. For this purpose we distinguish two cases.

Let first $\ell \geq \ell_{\mathrm{hf}} + 1$ such that $\kappa q_\ell > 1$. Then for all boxes t and s at level ℓ the admissibility criterion (A3) implies (A1). In particular t and s with $\operatorname{dist}(t, s) = q_\ell/\sqrt{3}$

are admissible if (A3) holds, i.e.

$$\kappa q_\ell^2 \leq \eta_2 \operatorname{dist}(t,s) = \frac{\eta_2 q_\ell}{\sqrt{3}},$$

which is equivalent to

$$\eta_2 \geq \sqrt{3}\kappa q_\ell. \tag{3.3}$$

Let now $\ell \geq \ell_{\mathrm{hf}}+1$ but $\kappa q_\ell \leq 1$. Then t and s with $\operatorname{dist}(t,s) = q_\ell/\sqrt{3}$ are admissible if (A1) holds, which means

$$q_\ell \leq \eta_2 \operatorname{dist}(t,s) = \frac{\eta_2 q_\ell}{\sqrt{3}},$$

or equivalently

$$\eta_2 \geq \sqrt{3}. \tag{3.4}$$

By combining (3.3) and (3.4) we see that the number of inadmissible blocks at all levels $\ell > \ell_{\mathrm{hf}}$ is minimized, if we choose

$$\eta_2 \geq \sqrt{3}\max\{1, \kappa q_{\ell_{\mathrm{hf}}+1}\}.$$

In particular a suitable choice for η_2 is given by

$$\left\lceil \sqrt{3}\max\{1, \kappa q_{\ell_{\mathrm{hf}}+1}\} \right\rceil, \tag{3.5}$$

where $\lceil \alpha \rceil$ denotes the least integer greater than or equal to α as usual. Again, we summarize this parameter choice rule in a separate remark.

Remark 3.2 (Parameter choice rule for η_2). For a given κ, a given uniform box cluster tree $\mathcal{T}_X$ and a fixed level ℓ_{hf} choose the separation parameter η_2 such that (3.5) holds.

In the discussion of the parameter choice of ℓ_{hf} we observed, that in the computations in Section 3.1.2 the product $\kappa q_{\ell_{\mathrm{hf}}+1}$ was approximately 2.72 for the set of points $P_X(4)$, $\kappa = 12.56$, and $\ell_{\mathrm{hf}} = 3$, as well as for $P_X(5)$, $\kappa = 25.12$ and $\ell_{\mathrm{hf}} = 4$. Hence in these examples there holds $\sqrt{3}\kappa q_{\ell_{\mathrm{hf}}+1} \approx 4.71$ and by the parameter choice rule in Remark 3.2 we should choose $\eta_2 = 5$ in both settings. This choice appeared to be reasonable in the computations, in particular with respect to the approximation quality.

3.2 Further numerical experiments

In all previous computations we focused on sets of points on the surface of the cube $[-1,1]^3$. Furthermore these points were chosen such that corresponding uniform box cluster trees were balanced. In this section we want to consider a more general setting. For this purpose we consider points on the surface of the sphere $S^2 = \{x \in \mathbb{R}^3 : |x| = 1\}$ and investigate the fast directional matrix vector multiplication for the corresponding matrix A in (3.1), similarly as in Sections 3.1.2 and 3.1.3. It is our goal to see if the choice of the parameters according to Remarks 3.1 and 3.2 is reasonable in such a more general setting.

For the computations in this section we consider the sets of points $P_X^S(L)$. These sets are constructed by projecting the points in the sets $P_X(L)$ constructed in Section 3.1.1 onto the sphere S^2. In Figure 3.2 we present the set $P_X^S(2)$.

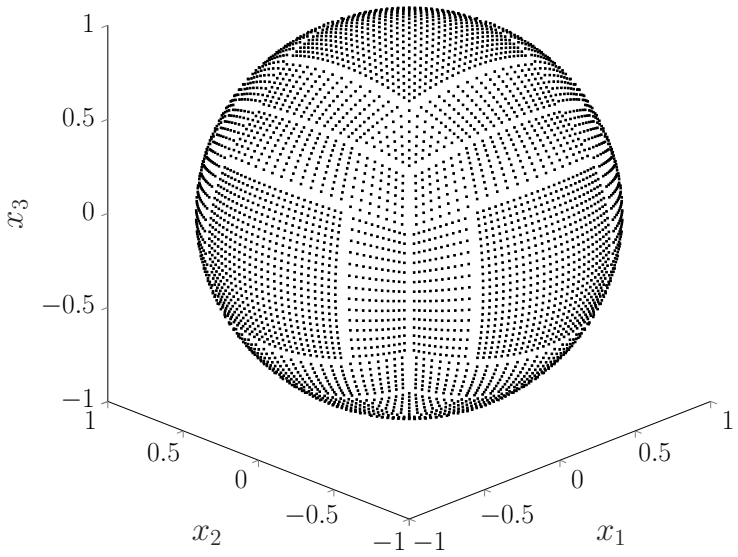

Figure 3.2: The set of points $P_X^S(2)$ on the sphere S^2

The uniform box cluster trees for these geometries are constructed by Algorithm 2.3 using the initial box $X = [-1,1]^3$ and allowing at most $n_{\max} = 150$ points per leaf box. The resulting box cluster trees are not balanced anymore, which we observe in Table 3.12. In this table we compare the number of leaves in the box cluster trees $\mathcal{T}_X(P_X(5))$ and $\mathcal{T}_X(P_X^S(5))$. The entries for levels less than four and greater than seven are missing, since for these levels there are no leaves in either of the two trees. In $\mathcal{T}_X(P_X(5))$ all leaves are at level five, which is the depth of this tree. In $\mathcal{T}_X(P_X^S(5))$ the majority of the leaves is at level six, but also an other levels we find some leaves. This is the typical situation for general sets of points.

Table 3.12: Comparison of the number of leaves in the trees $\mathcal{T}_X(P_X(5))$ and $\mathcal{T}_X(P_X^S(5))$ corresponding to the points $P_X(5)$ on the surface of $[-1,1]^3$ and the points $P_X^S(5)$ on the sphere S^2

box cluster tree	level 4	level 5	level 6	level 7
$\mathcal{T}_X(P_X(5))$	0	5768	0	0
$\mathcal{T}_X(P_X^S(5))$	240	1760	13824	168

Let us check if the fast directional matrix vector multiplication given in Algorithm 2.42 performs well for such general sets of points. For the following computations we consider the set of points $P_X^S(5)$ with cardinality $N = 829176$ and choose the wave number $\kappa = 25.12$, i.e. the same as for $P_X(5)$ in the calculations in Sections 3.1.2 and 3.1.3. We consider the matrix A in (3.1) for the points in $P_X^S(5)$, use Algorithm 2.42 to compute the approximation $\hat{g} = \hat{g}^+ + g^-$ of the matrix-vector product $g = Av$ for a random but fixed vector v, and compare the computational results for several choices of the parameters ℓ_{hf}, η_2, and interpolation degrees $m \in \{2, ..., 5\}$. In all computations the parameter ε_{ACA} for the stopping criterion of the ACA is a priori chosen sufficiently high to avoid an influence on the relative errors. In particular, we choose a lower parameter for $m = 5$.

Table 3.13: Relative errors $\|g - \hat{g}\|_2/\|g\|_2$ and computation times in seconds for the farfield part $\hat{g}^+$ for $\eta_2 \in \{4, 5, 10\}$, $\ell_{\text{hf}} \in \{3, ..., 5\}$, and interpolation degrees $m \in \{2, ..., 5\}$, together with the corresponding computation times in seconds for the nearfield part g^-. (set of points $P_X^S(5)$, $\kappa = 25.12$)

		$\eta_2 = 4$		$\eta_2 = 5$		$\eta_2 = 10$	
	$(m, \varepsilon_{\text{ACA}})$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$	rel. error	time $\hat{g}^+$
$\ell_{\text{hf}} = 3$	$(2, 10^{-5})$	2.65_{-1}	5.00	2.51_{-1}	3.17	2.81_{-1}	2.01
	$(3, 10^{-5})$	4.75_{-2}	10.92	4.75_{-2}	7.58	5.09_{-2}	5.16
	$(4, 10^{-5})$	9.09_{-3}	20.67	9.45_{-3}	14.56	9.92_{-3}	10.04
	$(5, 10^{-6})$	1.57_{-3}	36.16	1.62_{-3}	25.78	1.68_{-3}	18.10
$\ell_{\text{hf}} = 4$	$(2, 10^{-5})$	6.02_{-2}	5.05	6.80_{-2}	3.38	8.15_{-2}	2.22
	$(3, 10^{-5})$	8.74_{-3}	11.40	1.06_{-2}	8.04	1.50_{-2}	5.65
	$(4, 10^{-5})$	1.21_{-3}	21.85	1.50_{-3}	15.73	2.66_{-3}	11.58
	$(5, 10^{-6})$	1.59_{-4}	39.20	1.96_{-4}	28.90	4.50_{-4}	21.72
$\ell_{\text{hf}} = 5$	$(2, 10^{-5})$	1.70_{-2}	6.23	1.98_{-2}	4.55	3.20_{-2}	3.39
	$(3, 10^{-5})$	1.95_{-3}	14.01	2.49_{-3}	10.65	5.36_{-3}	8.33
	$(4, 10^{-5})$	2.26_{-4}	28.13	3.19_{-4}	22.36	9.74_{-4}	18.28
	$(5, 10^{-6})$	2.46_{-5}	66.70	3.58_{-5}	52.65	1.58_{-4}	45.09
	time g^-	375.80		224.70		125.28	

Let us first consider the relative errors $\|g - \hat{g}\|_2/\|g\|_2$ and the computation times of the farfield part $\hat{g}^+$ for different η_2 and ℓ_{hf} in Table 3.13. The parameter rules in Remarks 3.1 and 3.2 suggest to choose $\ell_{\text{hf}} = 4$ and $\eta_2 = 5$, as in the case of the set of points $P_X(5)$. Hence, we take a closer look at this choice.

Compared to the relative errors for $\ell_{\text{hf}} = 3$ the relative errors for $\ell_{\text{hf}} = 4$ are clearly lower for fixed η_2 and interpolation degree m, while the farfield computation times do not differ much. Therefore, ℓ_{hf} should not be chosen smaller than four. The comparison of the errors for $\ell_{\text{hf}} = 4$ and $\ell_{\text{hf}} = 5$ is more complicated. Both choices have their advantages. While choosing $\ell_{\text{hf}} = 4$ leads to better farfield computation times and a lower increase of these times for increasing m, choosing $\ell_{\text{hf}} = 5$ leads to better relative errors and a faster decrease of these errors. However, none of the two choices appears to be significantly better than the other.

Next we want to compare the results for different choices of η_2 for fixed $\ell_{\text{hf}} \in \{4, 5\}$. We observe that the choice $\eta_2 = 5$ is a good compromise when considering the relative errors and farfield computation times. The farfield computation times for $\eta_2 = 5$ are lower than the times for $\eta_2 = 4$ but higher than those for $\eta_2 = 10$. At the same time the errors for $\eta_2 = 5$ are worse than the errors for $\eta_2 = 4$ but better than those for $\eta_2 = 10$. We note further, that the errors for $\eta_2 = 4$ and $\eta_2 = 5$ appear to converge faster to zero with increasing m than the errors for $\eta_2 = 10$, while the computation times seem to behave similarly for all three choices with increasing m.

What we have ignored so far are the nearfield computation times, which are given in the last line of Table 3.13. Comparing them shows that the only acceptable choice in this example is $\eta_2 = 10$. Indeed, for the other choices of η_2 the nearfield times are significantly higher, in particular in regard to the farfield computation times. Only if the desired accuracy is high enough, i.e. when the farfield computation times dominate the time required to compute $\hat{g} = \hat{g}^+ + g^-$, we expect a better performance for the choices $\eta_2 = 4$ and $\eta_2 = 5$ due to the better convergence of the errors.

The conclusion which we draw from the test on the surface of the sphere is that the parameter selection strategy presented in Section 3.1.4 is only partially applicable to general geometries. While the choice of ℓ_{hf} according to the parameter choice rule 3.1 seems to be reasonable, the choice of η_2 according to the parameter choice rule 3.2 is too restrictive for general examples. In particular, if only a low accuracy is required minimizing the number of all inadmissible blocks of the matrix A seems to be crucial for a good performance.

3.3 Effect of the compression via ACA

For all the computations in the previous sections we used an implementation of Algorithm 2.38 in which we compressed the occuring coupling matrices $A_{c,t,s}$ in step three with the adaptive cross approximation (ACA). Here we present an example which shows the benefits of this compression.

Let us consider the set of points $P_X(5)$ with cardinality $N = 829176$ on the surface of the cube $[-1, 1]^3$, as constructed in Section 3.1.1. Let the matrix A be given by (3.1) for this set of points and the wave number $\kappa = 25.12$. We repeat the same computations as in Section 3.1.2 with the same vector v and the parameters $\eta_2 = 5$ and $\ell_{\mathrm{hf}} = 4$, once with compression and once without. Then we compare the computation times of the farfield part $\hat{g}^+$, the required storage for the coupling matrices, and the relative errors $\|\hat{g} - g\|_2/\|g\|_2$, where $g = Av$ is the exact matrix vector product and $\hat{g}$ its approximation computed by Algorithm 2.42. The results are given in Table 3.14.

Table 3.14: Relative errors $\|g-\hat{g}\|_2/\|g\|_2$, computation times in seconds for the farfield part $\hat{g}^+$ and required storage for the coupling matrices $A_{c,t,s}$ for implementations of Algorithm 2.42 with compression (ACA) and without compression (direct). (set of points $P_X(5)$, $\kappa = 25.12$)

	rel. error		time $\hat{g}^+$		storage	
$(m, \varepsilon_{\mathrm{ACA}})$	ACA	direct	ACA	direct	ACA	direct
$(3, 10^{-3})$	8.15_{-3}	8.16_{-3}	3.29	46.85	107.6 MiB	614 MiB
$(4, 10^{-5})$	1.12_{-3}	1.10_{-3}	10.89	134.02	436.4 MiB	2.3 GiB
$(5, 10^{-5})$	1.64_{-4}	1.64_{-4}	20.43	446.69	754.4 MiB	6.8 GiB

In contrast to Section 3.1.2 we reduce the parameter $\varepsilon_{\mathrm{ACA}}$ for $m = 3$, choosing $\varepsilon_{\mathrm{ACA}} = 10^{-3}$ instead of $\varepsilon_{\mathrm{ACA}} = 10^{-5}$. In Table 3.14 we observe that this has no negative effect on the approximation quality. However, it further reduces the required storage for the coupling matrices and the computation time of the farfield part $\hat{g}^+$. Indeed, we see in Table 3.7 that for $\eta_2 = 5$, $m = 3$, and $\varepsilon_{\mathrm{ACA}} = 10^{-5}$ the computation time is 5.49 while it is 3.29 for $\varepsilon_{\mathrm{ACA}} = 10^{-3}$. For $m = 4$ choosing $\varepsilon_{\mathrm{ACA}} = 10^{-4}$ leads to an error of $1.39 \cdot 10^{-3}$, which is why we retain $\varepsilon_{\mathrm{ACA}} = 10^{-5}$.

A comparison of the results in Table 3.14 shows why a compression of the coupling matrices $A_{c,t,s}$ is crucial. While we observe that the relative errors for the implementation of Algorithm 2.42 with and without compression are nearly identical, the computation times and storage requirements for the coupling matrices differ essentially. For $m = 3$ and $m = 4$ the required storage is reduced by a factor greater than five, and the farfield computation times by a factor greater than twelve. Even better, for $m = 5$ the storage is reduced by a factor greater than nine and the farfield computation time by a factor greater than 21.

In addition we see in Table 3.15 that also the setup time is significantly lower for the implementation including the compression of the coupling matrices. We recall that

these setup times include the required time for steps one to three in Algorithm 2.42, i.e. the construction of the tree structures and directions, as well as the times to compute and store the non-directional transfer matrices E_t' introduced in Section 2.2.1 and the coupling matrices $A_{c,t,s}$.

Table 3.15: Setup times in seconds for the implementations of Algorithm 2.42 with compression (ACA) and without compression (direct). ($\kappa = 25.12$, set of points $P_X(5)$)

	time setup	
$(m, \varepsilon_{\text{ACA}})$	ACA	direct
$(3, 10^{-3})$	4.64	51.60
$(4, 10^{-5})$	16.49	150.96
$(5, 10^{-5})$	28.51	371.22

The difference in the setup times is caused by the different computation of the coupling matrices $A_{c,t,s}$, since the two considered implementations are identical up to this computation. In Section 2.2.1 we have mentioned that the approximation of a coupling matrix $A_{c,t,s}$ by means of ACA requires only $\mathcal{O}(r^2(m+1)^3)$ floating point operations, where r is the rank of the resulting approximation, while the direct computation of such a matrix requires $\mathcal{O}((m+1)^6)$ floating point operations, which is significantly more if $r^2 \ll (m+1)^3$. This explains the significantly lower setup times for the implementation of Algorithm 2.42 with compression compared with the implementation without compression.

In conclusion, we see that by compressing the coupling matrices $A_{c,t,s}$ by means of ACA we do not only reduce the farfield computation time but also the setup time of Algorithm 2.42. The only drawback is the additional parameter ε_{ACA} which has to be introduced for the stopping criterion of the ACA. However, it seems sufficient to choose this parameter close to the expected relative error to achieve reasonable approximations.

Conclusion and Outlook

In this thesis we discussed directional single- and multi-level-approximations of the Helmholtz kernel f based on polynomial tensor interpolation. We introduced admissibility criteria that guarantee a suitable approximation quality. In particular we proved that the directional single- and multi-level approximation errors converge exponentially to zero with increasing polynomial degree m. Subsequently we derived an algorithm for fast directional matrix-vector multiplications which is applicable for matrices generated by the Helmholtz kernel. For this purpose we used a clustering strategy based on uniform box cluster trees. We constructed hierarchical sets of directions and proved that they are suitable for the directional approximation. An appropriate partitioning strategy allowed us then to split the matrix-vector product in a nearfield and farfield part, and to approximate the latter using the directional multi-level approximation of the kernel f. We analyzed the complexity of the resulting Algorithm 2.42. In particular, we proved that the storage and runtime complexity of this algorithm is of order $\mathcal{O}(N \log(N))$ under suitable assumptions on the wave number κ, the geometry, and the degrees of freedom N. The accuaracy and performance of the fast directional matrix-vector multiplication was further investigated in numerical experiments for which we implemented it in C++. To understand the effects of the separation parameter η_2 and the parameter ℓ_{hf} on the accuracy and runtimes we conducted a parameter study. Its results lead to a heuristic parameter selection strategy.

There are still a few details in this thesis which are missing or could be improved:

- There is room for improvement in our complexity estimate of Algorithm 2.42. We have already pointed out that the estimates which we proved are valid in general only for points distributed more or less uniformly in a 3D domain. The complexity esimate in [5], which holds for suitable 2D manifolds is not directly applicable in our setting, since it requires rather balanced cluster trees. However, we are optimistic that it can be adapted to our setting.
- The numerical tests on the surface of the sphere presented in Section 3.2 showed that further investigation is required for the development of reliable parameter selection strategies for the parameters η_2 and ℓ_{hf} in Algorithm 2.42. In particular a broader range of wave numbers and degrees of freedom should be considered in the numerical examples in future works.
- One of the motivations for the derivation of the fast directional matrix-vector multiplication was its relevance for the solution of boundary value problems for

the Helmholtz equation by means of boundary element methods. In this thesis, however, we have not considered such an application, because it would require some further adaptions of the theory and the implementation. Let us shortly discuss some of the differences and obstacles. In the setting of boundary element methods, we typically do not deal with sets of loose points on some 2D manifold, but sets of elements of the discretization of the boundary. In case of Galerkin discretizations the matrix entries are furthermore given by some surface integrals over these elements, and not simple function values. The matrix entries in inadmissible blocks need to be evaluated directly. Note that this requires to evaluate possibly singular integrals. On the other hand, the matrix entries in admissible blocks can be approximated using some quadrature formulas, resulting in a similar setting which we considered in this thesis, i.e. sets of points and function values as matrix indices. For the computations it makes sense to keep the points corresponding to an element together. However, if we build a uniform box cluster tree by Algorithm 2.3 then the points corresponding to a common element can be contained in different boxes. To overcome this problem we could pad all boxes at a certain level in the uniform cluster tree equally such that each element would be completely contained in a box. While the uniformity of the boxes would be preserved like this, other basic properties would be lost with this approach. All these changes would in particular require some new parameter studies for η_2 and ℓ_{hf}.

Bibliography

[1] M. Bebendorf. Approximation of boundary element matrices. *Numer. Math.*, 86(4):565–589, 2000.

[2] M. Bebendorf, C. Kuske, and R. Venn. Wideband nested cross approximation for Helmholtz problems. *Numer. Math.*, 130(1):1–34, 2015.

[3] M. Bebendorf and S. Rjasanow. Adaptive low-rank approximation of collocation matrices. *Computing*, 70(1):1–24, 2003.

[4] S. Börm. *Efficient Numerical Methods for Non-local Operators*, volume 14 of *Tracts in Mathematics*. European Mathematical Society, Zürich, 2010.

[5] S. Börm. Directional $\mathcal{H}^2$-matrix compression for high-frequency problems. *Numer. Linear. Algebra. Appl.*, 24(6):e2112, 2017.

[6] S. Börm and C. Börst. Hybrid matrix compression for high-frequency problems. Preprint, arXiv:1809.04384, 2018.

[7] S. Börm and J. M. Melenk. Approximation of the high-frequency Helmholtz kernel by nested directional interpolation: error analysis. *Numer. Math.*, 137(1):1–34, 2017.

[8] A. Brandt. Multilevel computations of integral transforms and particle interactions with oscillatory kernels. *Comput. Phys. Commun.*, 65(1):24 – 38, 1991.

[9] D. Brunner, M. Junge, P. Rapp, M. Bebendorf, and L. Gaul. Comparison of the fast multipole method with hierarchical matrices for the Helmholtz-BEM. *CMES Comput. Model. Eng. Sci.*, 58(2):131–159, 2010.

[10] H. Cheng, W. Y. Crutchfield, Z. Gimbutas, L. F. Greengard, J. F. Ethridge, J. Huang, V. Rokhlin, N. Yarvin, and J. Zhao. A wideband fast multipole method for the Helmholtz equation in three dimensions. *J. Comput. Phys.*, 216:300–325, 2006.

[11] R. Coifman, V. Rokhlin, and S. Wandzura. The Fast Multipole Method for the Wave Equation: A Pedestrian Prescription. *IEEE Trans. Antennas Propag.*, 35(3):7–12, 1993.

[12] W. Dahmen, S. Prössdorf, and R. Schneider. Wavelet approximation methods for pseudodifferential equations. II. Matrix compression and fast solution. *Adv. Comput. Math.*, 1(3-4):259–335, 1993.

[13] E. Darve and P. Havé. Efficient fast multipole method for low-frequency scattering. *J. Comput. Phys.*, 197(1):341–363, 2004.

[14] R. A. Devore and G. G. Lorentz. *Constructive Approximation*, volume 303 of *A Series of Comprehensive Studies in Mathematics*. Springer, Berlin-Heidelberg-New York, 1993.

[15] B. Engquist and L. Ying. Fast directional multilevel algorithms for oscillatory kernels. *SIAM J. Sci. Comput*, 29(4):1710–1737, 2007.

[16] P. Erdös. Problems and results on the theory of interpolation, II. *Acta. Math. Acad. Sci. H.*, 12:235–244, 1961.

[17] W. Fischer and I. Lieb. *Einführung in die komplexe Analysis*. Vieweg + Teubner, Wiesbaden, 2010.

[18] K. Giebermann. A new version of panel clustering for the boundary element method. Preprint, Institut für Angewandte Mathematik Universität Bonn, 1999.

[19] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins University Press, Baltimore, 1996.

[20] L. Greengard. *The Rapid Evaluation of Potential Fields in Particle Systems*. PhD thesis, Yale University, New Haven, CT, USA, 1987.

[21] L. Greengard, J. Huang, R. V., and S. Wandzura. Accelerating Fast Multipole Methods for the Helmholtz Equation at Low Frequencies. *IEEE Comput. Sci. Eng*, pages 32–38, 1998.

[22] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comput. Phys.*, 73(2):325–348, 1987.

[23] W. Hackbusch. A sparse matrix arithmetic based on $\mathcal{H}$-matrices. part I: Introduction to $\mathcal{H}$-matrices. *Computing*, 62:89–108, 1999.

[24] W. Hackbusch. *Hierarchische Matrizen*. Springer, Berlin-Heidelberg, 2009.

[25] W. Hackbusch and B. Khoromskij. A sparse $\mathcal{H}$-matrix arithmetic, part II: Application to multi-dimensional problems. *Computing*, 64, 2000.

[26] W. Hackbusch, B. Khoromskij, and S. A. Sauter. On $\mathcal{H}^2$-matrices. In *Lectures on applied mathematics (Munich, 1999)*, pages 9–29. Springer, Berlin, 2000.

[27] W. Hackbusch and Z. Nowak. On the fast matrix multiplication in the boundary element method by panel clustering. *Numer. Math.*, 54(4):463–492, 1989.

[28] F. Himmelbauer. Analyse und Implementierung einer adaptiven schnellen Multipolmethode für die Potentialrechnung. Masterarbeit, Institut für Numerische Mathematik, Technische Universität Graz, 2017.

[29] B. Hu and W. C. Chew. Fast inhomogeneous plane wave algorithm for scattering from objects above the multilayered medium. *IEEE Trans. Geosci. Remote Sens.*, 39(5):1028–1038, May 2001.

[30] M. Messner. *Fast Boundary Element Methods in Acoustics*. Number 13 in Monographic Series TU Graz / Computation in Engineering and Science. Verlag der Technischen Universität Graz, 2012.

[31] M. Messner, M. Schanz, and E. Darve. Fast directional multilevel summation for oscillatory kernels based on Chebyshev interpolation. *J. Comput. Phys.*, 231(4):1175–1196, 2012.

[32] T. J. Rivlin. *Chebyshev polynomials*. Wiley, New York, 1990.

[33] S. Rjasanow and O. Steinbach. *The Fast Solution of Boundary Integral Equations*. Springer-Verlag, Berlin, Heidelberg, 2007.

[34] V. Rokhlin. Rapid solution of integral equations of classical potential theory. *J. Comput. Phys.*, 60:187–207, 1985.

[35] V. Rokhlin. Diagonal forms of translation operators for Helmholtz equation in three dimensions. *Appl. Comput. Harmon. A.*, 1:82–93, 1993.

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitely indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Date	Signature