ASC Report No. 33/2015

Approximation of the high-frequency Helmholtz kernel by nested directional interpolation

S. Börm and J.M. Melenk

Institute for Analysis and Scientific Computing —
Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-05-6

Most recent ASC Reports

32/2015 *F. Achleitner, A. Arnold, E.A. Carlen*
On hypocoercive BGK models

31/2015 *M. Feischl, T. Führer, M. Niederer, S. Strommer, A. Steinboeck, and D. Praetorius*
Efficient numerical computation
of direct exchange areas
in thermal radiation analysis

30/2015 *B. Düring, P. Fuchs, and A. Jüngel*
A higher-order gradient flow scheme
for a singular one-dimensional
diffusion equation

29/2015 *C. Erath and D. Praetorius*
Adaptive finite volume methods with convergence rates

28/2015 *W. Auzinger, O. Koch, M. Schöbinger, E. Weinmüller*
A new version of the code bvpsuite for singular BVPs in ODEs: Nonlinear solver and its application to m-Laplacians.

27/2015 *C. Lehrenfeld, J. Schöberl*
High order exactly divergence-free Hybrid Discontinuous Galerkin Methods for unsteady incompressible flows

26/2015 *A. Jüngel, and S. Schuchnigg*
Entropy-dissipating semi-discrete Runge-Kutta schemes for nonlinear diffusion equations

25/2015 *W. Auzinger, H. Hofstätter, D. Ketcheson, O. Koch*
Practical splitting methods for the adaptive integration of nonlinear evolution equations.
Part I: Construction of optimized schemes and pairs of schemes

24/2015 *J. Burkotová, M. Hubner, I. Rachunková, and E.B. Weinmüller*
Asymptotic properties of Kneser solutions to nonlinear second order ODEs with regularly varying coefficients

23/2015 *C. Abert, G. Hrkac, D. Praetorius, M. Ruggeri, D. Suess*
Coupling of dynamical micromagnetism and a stationary spin drift-diffusion equation: A step towards a fully implicit spintronics frame

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: admin@asc.tuwien.ac.at
WWW: http://www.asc.tuwien.ac.at
FAX: +43-1-58801-10196

ISBN 978-3-902627-05-6

Approximation of the high-frequency Helmholtz kernel by nested directional interpolation

Steffen Börm and Jens M. Melenk

November 2, 2015

We present a data-sparse approximation scheme for integral operators associated with the Helmholtz equation in the high-frequency regime. The technique combines the directional approximation [8, 10] with nested tensor interpolation to achieve polylogarithmic-linear complexity.

We rigorously prove that the directional interpolation converges exponentially with the asymptotically optimal rate and that the nested interpolation, which is required to obtain an efficient hierarchical algorithm, preserves the exponential convergence.

1 Introduction

We consider the integral operator

$$\mathcal{G}[u](x) := \int_\Gamma g(x,y)u(y)\,dy,$$

where $\Gamma \subseteq \mathbb{R}^3$ is a two-dimensional surface and

$$g(x,y) = \frac{\exp(\mathbf{i}\kappa\|x-y\|)}{4\pi\|x-y\|} \tag{1}$$

denotes the *Helmholtz kernel function* with the *wave number* $\kappa \in \mathbb{R}_{\geq 0}$.

Applying a standard Galerkin discretization scheme with a finite element basis $(\varphi_i)_{i\in\mathcal{I}}$ leads to the *stiffness matrix* $G \in \mathbb{C}^{\mathcal{I}\times\mathcal{I}}$ given by

$$g_{ij} = \int_\Gamma \varphi_i(x) \int_\Gamma g(x,y)\varphi_j(y)\,dy\,dx \qquad \text{for all } i,j \in \mathcal{I}, \tag{2}$$

where we assume that the basis functions are sufficiently smooth to ensure that the integrals are well-defined even for $x = y$.

Since $g(x,y) \neq 0$ for all $x \neq y$, the matrix G is fully populated, and special matrix compression techniques have to be devised and employed in order to achieve logarithmic-linear complexity. Most successful modern techniques rely on a multilevel decomposition of $\Gamma \times \Gamma$ into subdomains where the kernel function g can be suitably approximated.

Fast multipole methods [20, 13] were originally developed for the case $\kappa = 0$, i.e., the Laplace kernel. They can be generalized to the high frequency Helmholtz case by using appropriate expansions and taking advantage of the fact that the operators appearing in the expansion can be diagonalized [21]. The expansion becomes unstable for low frequencies, but this problem can be addressed by using an alternate expansion that is suitable for this case [12]. Implementing the fast multipole method is a comparatively challenging task, and since the error analysis relies heavily on the special form of the kernel function, it is not obvious how to generalize the analysis to other kernels arising in wave propagation problems.

Butterfly schemes (sometimes also known as multi-level matrix decomposition algorithms, MLMDA) [18], on the other hand, work directly with the matrix entries to approximate submatrices by products of permutation and block-diagonal matrices. These algorithms lead to comparatively simple algebraic structures that can be handled efficiently, but they will typically not reach the theoretical asymptotic efficiency that is available to fast multipole methods. The butterfly representation is closely connected to general $\mathcal{H}$-matrix representations [15, 11] in that it replaces general low-rank factorizations by the more efficient butterfly factorizations. Numerical experiments [14] indicate that, similar to $\mathcal{H}$-matrix methods, the butterfly representation can also be used to approximate the LR factorization of G, and this would give rise to efficient preconditioners.

Directional methods [8, 10, 17, 1] take advantage of the fact that, once a direction for a plane wave is chosen, the Helmholtz kernel (1) can be written as a product of this plane wave and a function that is smooth inside a conical domain. The smooth part can then be approximated, for example, by polynomials in this conical domain so that degenerate approximations of g become available there. In turn these degenerate approximation lead to low-rank blocks for the stiffness matrix G as in $\mathcal{H}$-matrices and even $\mathcal{H}^2$-matrices (see [16, 6, 4] and in particular [1] in the context of Helmholtz problems). In fact, the use of matrix formats such as $\mathcal{H}^2$-matrices is an essential ingredient for the logarithmic-linear complexity of schemes that are based on directional approximations. Such matrix formats were implicitly used in [17] and then explicitly in [1] with directional cluster bases.

Directional methods are usually introduced using adaptive compression techniques intended to improve efficiency. These techniques yield very convincing compression rates in practice, but rely on a stability assumption for incomplete LU factorizations that — to the best of our knowledge — has not been fully proven.

The purpose of our paper is to close this gap by investigating a non-adaptive directional approximation scheme based on polynomial interpolation. We present a complete and rigorous mathematical proof of the exponential convergence of our method. For the reader's convenience, the scheme as well as the exponential convergence result are collected at the end of Section 4.

We point out that our approximation scheme can be combined with stable recom-

pression algorithms [3, 4] to obtain quasi-optimal compression rates while preserving the rigorous error bounds.

The present work has two main mathematical results: first we prove that the product of the Helmholtz kernel function g and a plane wave

$$g_c(x,y) = g(x,y)\exp(-\mathbf{i}\kappa\langle x-y,c\rangle) = \frac{\exp(\mathbf{i}\kappa(\|x-y\| - \langle x-y,c\rangle))}{4\pi\|x-y\|}$$

can be approximated by tensor interpolation if x and y are in domains satisfying suitable admissibility conditions (see Section 4 and Theorem 27). Due to $\exp(\mathbf{i}\kappa\langle x-y,c\rangle) = \exp(\mathbf{i}\kappa\langle x,c\rangle)\overline{\exp(\mathbf{i}\kappa\langle y,c\rangle)}$, this result immediately gives rise to a low-rank approximation of g. Given that analytic functions are to be approximated, the *via regia* is to bound the holomorphic extensions of the reduced kernel function similar to [4, Section 4.4].

In order to obtain the polylogarithmic-linear complexity, we require an approximation that uses nested expansion systems. Our second result (see Theorem 30) shows that these systems can also be constructed by interpolation and that the resulting approximation also converges exponentially as long as the degree m of the underlying polynomial interpolation scheme satisfies the condition $m \in \mathcal{O}(\log(\log(\kappa)))$. This is a fairly weak condition, since we have to require $m \in \mathcal{O}(\log(\kappa))$ to ensure that the error introduced by the matrix compression is of the order of the discretization error. The analysis leading to Theorem 30 is non-standard since it involves iterated polynomial interpolation. The basic issues also had to be adressed in [1]; however, since the present analysis relies on polynomial interpolation instead of cross approximation, we have significantly stronger tools for the analysis at our disposal and are therefore able to obtain stronger results.

It is worth stressing that the directional methods are not particularly tailored to the specific kernel g considered here. Therefore it should be possible to generalize our approach very easily to other kernels associated with accoustic, electromagnetic or elastic wave equations.

2 Plane wave interpolation

2.1 Tensor interpolation

Polynomial interpolation on tensor product domains feature importantly in the present paper. We therefore fix some notation and assumptions. For $m \in \mathbb{N}_0$, we denote by Π_m the space of univariate polynomials of degree m. Let $\xi_0,\dots,\xi_m \in [-1,1]$ be distinct interpolation points and define the associated Lagrange polynomials by

$$L_\nu(z) := \prod_{\substack{\mu=0\\ \mu\neq\nu}}^{m} \frac{z-\xi_\mu}{\xi_\nu-\xi_\mu} \qquad \text{for all } \nu\in\{0,\dots,m\},\ z\in\mathbb{C}. \tag{3}$$

The corresponding interpolation operator is given by

$$\mathfrak{I} : C[-1,1] \to \Pi_m, \qquad f \mapsto \sum_{\nu=0}^{m} f(\xi_\nu)L_\nu.$$

An interpolation operator $\mathfrak{I}$ on $[-1,1]$ induces in a natural way an interpolation operator $\mathfrak{I}_{[a,b]} : C([a,b]) \to \Pi_m$ on an interval $[a,b]$ by

$$\mathfrak{I}_{[a,b]}[f] := (\mathbf{I}[f \circ \Phi_{[a,b]}]) \circ \Phi_{[a,b]}^{-1},$$

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

$$\Phi_{[a,b]} : [-1,1] \to [a,b], \qquad \xi \mapsto \frac{b+a}{2} + \frac{b-a}{2}\xi. \tag{4}$$

This operator can be written explicitly as

$$\mathfrak{I}_{[a,b]} : C([a,b]) \to \Pi_m, \qquad f \mapsto \sum_{\nu=0}^{m} f(\xi_{[a,b],\nu}) L_{[a,b],\nu}$$

with interpolation nodes $\xi_{[a,b],\nu}$ and Lagrange polynomials $L_{[a,b],\nu}$ given by

$$\xi_{[a,b],\nu} = \Phi_{[a,b]}(\xi_\nu), \quad L_{[a,b],\nu}(z) = \prod_{\substack{\mu=0 \\ \mu \neq \nu}}^{m} \frac{z - \xi_{[a,b],\mu}}{\xi_{[a,b],\nu} - \xi_{[a,b],\mu}} \quad \text{for all } \nu \in \{0,\dots,m\},\ z \in \mathbb{C}.$$

Tensor product interpolation generalizes this procedure to multiple dimensions. For an axis-parallel n-dimensional box

$$B := [a_1,b_1] \times \cdots \times [a_n,b_n]$$

we introduce the interpolation points and Langrange interpolation polynomials by

$$\begin{aligned} \xi_{B,\nu} &:= (\xi_{[a_1,b_1],\nu_1}, \dots, \xi_{[a_n,b_n],\nu_n}) \in B, \\ L_{B,\nu}(x) &:= L_{[a_1,b_1],\nu_1}(x_1) \cdots L_{[a_n,b_n],\nu_n}(x_n) \quad \text{for all } x \in \mathbb{C}^n,\ \nu \in M := \{0,\dots,m\}^n; \end{aligned}$$

the tensor interpolation operator is then written in the familiar form

$$\mathfrak{I}_B[f] = \sum_{\nu \in M} f(\xi_{B,\nu}) L_{B,\nu} \qquad \text{for all } f \in C(B). \tag{5}$$

2.2 Plane wave approximation

The function

$$x \mapsto \exp(\mathbf{i}\kappa\|x-y\|)$$

appearing in the Helmholtz kernel (1) represents a spherical wave originating at y. If the wave number κ is large, the function oscillates rapidly and standard approximation techniques such as polynomial interpolation require a fairly large number of terms to reach a required accuracy. In order to overcome this obstacle, we follow the idea of Brandt [8] and Engquist and Ying [10]. That is, we introduce a vector $c \in \mathbb{R}^3$ with $\|c\| = 1$ and use $\langle x-y, c\rangle$ as an approximation of $\|x-y\|$:

$$\exp(\mathbf{i}\kappa\|x-y\|) = \exp(\mathbf{i}\kappa\langle x-y,c\rangle) \exp(\mathbf{i}\kappa(\|x-y\| - \langle x-y,c\rangle)$$

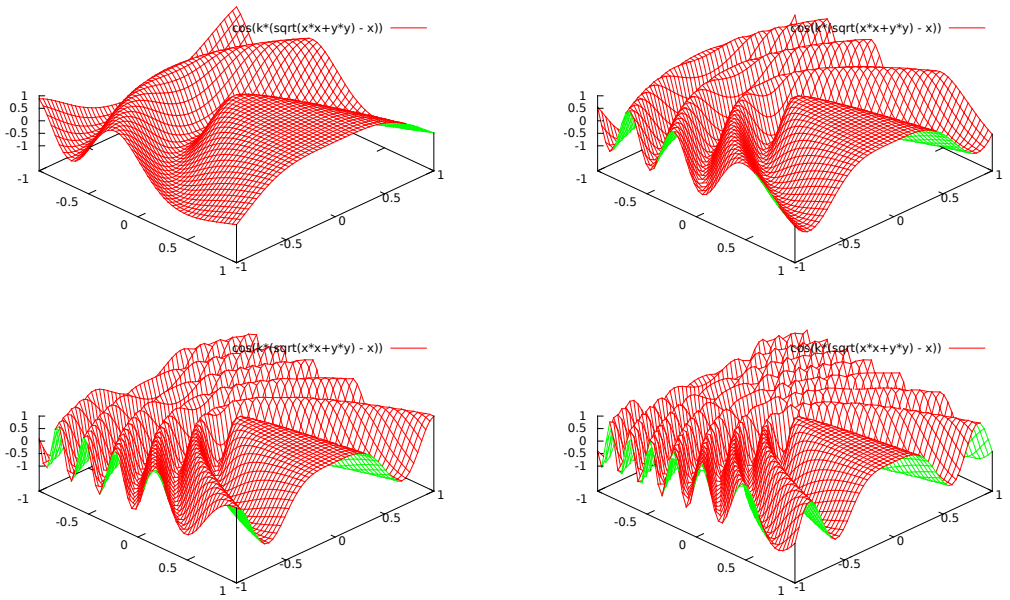

Figure 1: $x \mapsto \cos(\kappa(\|x\| - x_1))$ in $[-1,1] \times [-1,1]$ for $\kappa \in \{5, 10, 15, 20\}$

$$= \exp(\mathbf{i}\kappa\langle x, c\rangle) \exp(-\mathbf{i}\kappa\langle y, c\rangle) \exp(\mathbf{i}\kappa(\|x-y\| - \langle x-y, c\rangle)). \tag{6}$$

The first factor depends only on x, the second only on y. Hence, in order to obtain a separable form, i.e., a (short) sum of products of functions of x and y only, we have to require this of the third term in (6). We have

$$\begin{aligned}\|x-y\| - \langle x-y, c\rangle &= \|x-y\| - \|x-y\|\langle \frac{x-y}{\|x-y\|}, c\rangle = \|x-y\| \left(1 - \langle \frac{x-y}{\|x-y\|}, c\rangle\right) \\ &= \|x-y\|(1 - \cos\angle(x-y, c)) \approx \frac{1}{2}\|x-y\| \sin^2\angle(x-y, c).\end{aligned}$$

In this case, the argument of the exponential function satisfies

$$\mathbf{i}\kappa(\|x-y\| - \langle x-y, c\rangle) \approx \frac{1}{2}\mathbf{i}\kappa\|x-y\| \sin^2\angle(x-y, c).$$

We observe that we can compensate a large wave number κ if we ensure that the angle between $x-y$ and c is sufficiently small. In this case, the third term in (6) is a smooth function (cf. Figure 1) that can be approximated using standard techniques.

In this paper, we use standard tensor product interpolation as described in Section 2.1: given axis-parallel target and source boxes $B_t, B_s \subseteq \mathbb{R}^3$, we approximate the modified Helmholtz kernel function

$$g_c(x,y) := \frac{\exp(\mathbf{i}\kappa(\|x-y\| - \langle x-y, c\rangle))}{4\pi\|x-y\|} \tag{7}$$

by its interpolating polynomial

$$\begin{aligned}\tilde{g}_{c,ts}(x,y) &:= \mathfrak{I}_{B_t\times B_s}[g_c](x,y) \\ &= \sum_{\nu\in M}\sum_{\mu\in M} g_c(\xi_{t,\nu},\xi_{s,\mu})L_{t,\nu}(x)L_{s,\mu}(y) \qquad \text{for all } x\in B_t,\ y\in B_s.\end{aligned} \tag{8}$$

Compared to the notation of Section 2.1 we have made a slight change: we denote the interpolation points by $\xi_{t,\nu}$ and $\xi_{s,\mu}$ instead of $\xi_{B_t,\nu}$ and $\xi_{B_s,\mu}$ and the Lagrange polynomials by $L_{t,\nu}$ and $L_{s,\mu}$ instead of $L_{B_t,\nu}$ and $L_{B_s,\mu}$.

Feeding the approximation (8) of g_c back into (6) leads to an approximation of the kernel function g by

$$\begin{aligned}\tilde{g}_{ts}(x,y) &:= \exp(\mathbf{i}\kappa\langle x-y,c\rangle)\tilde{g}_{c,ts}(x,y) \\ &= \sum_{\nu\in M}\sum_{\mu\in M} g_c(\xi_{t,\nu},\xi_{s,\mu})\exp(\mathbf{i}\kappa\langle x,c\rangle)L_{t,\nu}(x)\overline{\exp(\mathbf{i}\kappa\langle y,c\rangle)L_{s,\mu}(y)}.\end{aligned}$$

By introducing the modified Lagrange polynomials

$$\begin{aligned}L_{tc,\nu}(x) &:= \exp(\mathbf{i}\kappa\langle x,c\rangle)L_{t,\nu}(x), \\ L_{sc,\mu}(y) &:= \exp(\mathbf{i}\kappa\langle y,c\rangle)L_{s,\mu}(y) \qquad \text{for all } x\in B_t,\ y\in B_s,\end{aligned}$$

our approximation to the kernel g takes the form

$$\tilde{g}_{ts}(x,y) = \sum_{\nu\in M}\sum_{\mu\in M} g_c(\xi_{t,\nu},\xi_{s,\mu})L_{tc,\nu}(x)\overline{L_{sc,\mu}(y)} \qquad \text{for all } x\in B_t,\ y\in B_s. \tag{9}$$

If M has only few elements, then (9) is a short sum of *separated* functions. Such a structure is at the heart of many multilevel schemes for non-local operators. The analysis of this scheme has to address the following questions:

- Identify appropriate *admissibility* conditions for boxes B_t, B_s so that (9) is a good approximation with small M; this is discussed in Section 4.
- Incorporate the fact that the direction c depends on the pair (B_t, B_s) into a multilevel structure. Since for a given box B_t possibly a large number of directions have to be taken into account, we have to be careful in order to preserve the efficiency of our scheme. This is addressed through the new technique of *directional $\mathcal{H}^2$-matrices* introduced in Section 3.

3 Directional $\mathcal{H}^2$-matrices

An approximation of the form (9) is only admissible if the boxes B_t and B_s satisfy certain conditions. In particular, it is not admissible for the entire domain $\Gamma\times\Gamma$, so we have to split the domain into subdomains that either admit an approximation or are small enough to be handled directly.

This approach immediately gives rise to the $\mathcal{H}^2$-*matrix representation* [16, 6, 4], which we have to generalize to include the directions $c \in \mathbb{R}^3$ required by our approximation scheme. We call the result *directional* $\mathcal{H}^2$-*matrix representation* ($\mathcal{DH}^2$-matrix representation for short). Our definition of $\mathcal{DH}^2$-matrices is not identical to the one used in [1], since we do not switch to an $\mathcal{H}$-matrix representation for the low-frequency case, but use $\mathcal{H}^2$-matrix representations for all blocks.

Definition 1 (Cluster tree) *Let $\mathcal{T}$ be a labeled tree such that the label $\hat{t}$ of each node $t \in \mathcal{T}$ is a subset of the index set $\mathcal{I}$. We call $\mathcal{T}$ a* cluster tree *for $\mathcal{I}$ if*

- *the root $r \in \mathcal{T}$ is assigned $\hat{r} = \mathcal{I}$,*
- *the index sets of siblings are disjoint, i.e.,*

$$t_1 \neq t_2 \Longrightarrow \hat{t}_1 \cap \hat{t}_2 = \emptyset \qquad \text{for all } t \in \mathcal{T},\ t_1, t_2 \in \operatorname{sons}(t),\ \text{and}$$

- *the index sets of a cluster's sons are a partition of their father's index set, i.e.,*

$$\hat{t} = \bigcup_{t' \in \operatorname{sons}(t)} \hat{t}' \qquad \text{for all } t \in \mathcal{T} \text{ with } \operatorname{sons}(t) \neq \emptyset.$$

A cluster tree for $\mathcal{I}$ is usually denoted by $\mathcal{T}_\mathcal{I}$. Its nodes are called clusters.

A cluster tree $\mathcal{T}_\mathcal{I}$ can be split into levels: we let $\mathcal{T}_\mathcal{I}^{(0)}$ be the set containing only the root of $\mathcal{T}_\mathcal{I}$ and define

$$\mathcal{T}_\mathcal{I}^{(\ell)} := \{t' \in \mathcal{T}_\mathcal{I} \ :\ t' \in \operatorname{sons}(t) \text{ for a } t \in \mathcal{T}_\mathcal{I}^{(\ell-1)}\} \qquad \text{for all } \ell \in \mathbb{N}.$$

For each cluster $t \in \mathcal{T}_\mathcal{I}$, there is exactly one $\ell \in \mathbb{N}_0$ such that $t \in \mathcal{T}_\mathcal{I}^{(\ell)}$. We call this the *level number* of t and denote it by $\operatorname{level}(t) = \ell$. The maximal level

$$p_\mathcal{I} := \max\{\operatorname{level}(t) \ :\ t \in \mathcal{T}_\mathcal{I}\}$$

is called the *depth* of the cluster tree.

Pairs of clusters (t, s) correspond to subsets $\hat{t} \times \hat{s}$ of $\mathcal{I} \times \mathcal{I}$, and by extension to submatrices of $G \in \mathbb{C}^{\mathcal{I} \times \mathcal{I}}$. These pairs inherit the hierarchical structure provided by the cluster tree.

Definition 2 (Block tree) *Let $\mathcal{T}$ be a labeled tree, and let $\mathcal{T}_\mathcal{I}$ be a cluster tree for the index set $\mathcal{I}$ with root $r_\mathcal{I}$. We call $\mathcal{T}$ a* block tree *for $\mathcal{T}_\mathcal{I}$ if*

- *for each $b \in \mathcal{T}$ there are $t, s \in \mathcal{T}_\mathcal{I}$ such that $b = (t, s)$,*
- *the root $r \in \mathcal{T}$ satisfies $r = (r_\mathcal{I}, r_\mathcal{I})$,*
- *the label of $b = (t, s) \in \mathcal{T}$ is given by $\hat{b} = \hat{t} \times \hat{s}$, and*

- *for each* $b = (t, s) \in \mathcal{T}$ *we have*

$$\operatorname{sons}(b) \neq \emptyset \Longrightarrow \operatorname{sons}(b) = \operatorname{sons}(t) \times \operatorname{sons}(s).$$

A block tree for $\mathcal{T}_\mathcal{I}$ *is usually denoted by* $\mathcal{T}_{\mathcal{I}\times\mathcal{I}}$*. Its nodes are called* blocks*.*

In the following, we assume that a cluster tree $\mathcal{T}_\mathcal{I}$ for the index set $\mathcal{I}$ and a block tree $\mathcal{T}_{\mathcal{I}\times\mathcal{I}}$ for $\mathcal{T}_\mathcal{I}$ are given.

We have to identify submatrices, corresponding to blocks, that can be approximated efficiently. Considering the form (2) of the matrix entries, we require the approximation $\tilde{g}_{ts}$ of the kernel function g to be valid in the entire support of the basis functions φ_i and φ_j for $i \in \hat{t}$ and $j \in \hat{s}$.

Definition 3 (Bounding box) *Let* $t \in \mathcal{T}_\mathcal{I}$ *be a cluster. An axis-parallel box* $B_t \subseteq \mathbb{R}^3$ *is called a* bounding box *for* t *if*

$$\operatorname{supp}(\varphi_i) \subseteq B_t \qquad \text{for all } i \in \hat{t}.$$

In practice we can construct bounding boxes of minimal size by a simple and fast recursive algorithm [5, Example 2.2].

Our approximation scheme (9) requires a direction for the plane wave. In order to obtain the optimal order of complexity, we fix a set of directions for each level of the cluster tree and introduce a connection between the directions for a cluster t and the directions for its sons $t' \in \operatorname{sons}(t)$.

Definition 4 (Hierarchical directions) *A family* $(\mathcal{D}_\ell)_{\ell=0}^\infty$ *of finite subsets of* $\mathbb{R}^3$ *is called a* family of hierarchical directions *if*

$$\|c\| = 1 \vee c = 0 \qquad \text{for all } c \in \mathcal{D}_\ell,\ \ell \in \mathbb{N}_0.$$

A family $(\operatorname{sd}_\ell)_{\ell=0}^\infty$ *of mappings* $\operatorname{sd}_\ell : \mathcal{D}_\ell \to \mathcal{D}_{\ell+1}$ *is called a* family of compatible son mappings *if*

$$\|c - \operatorname{sd}_\ell(c)\| \leq \|c - \tilde{c}\| \qquad \text{for all } c \in \mathcal{D}_\ell,\ \tilde{c} \in \mathcal{D}_{\ell+1},\ \ell \in \mathbb{N}_0.$$

Given a cluster tree $\mathcal{T}_\mathcal{I}$*, a family of hierarchical directions, and a family of compatible son mappings, we write*

$$\mathcal{D}_t := \mathcal{D}_{\operatorname{level}(t)}, \qquad \operatorname{sd}_t(c) := \operatorname{sd}_{\operatorname{level}(t)}(c) \qquad \text{for all } t \in \mathcal{T}_\mathcal{I},\ c \in \mathcal{D}_{\operatorname{level}(t)}.$$

Remark 5 *The "direction"* $c = 0$ *is included in Definition 4 in order to include the low-frequency case in our scheme in a convenient way.*

Remark 6 (Implementation) *In practice, we only have to define* $\mathcal{D}_\ell$ *for* $\ell \leq p_\mathcal{I}$ *and* sd_ℓ *for* $\ell < p_\mathcal{I}$*. Our definition admits infinite levels only to avoid special cases.*

In the following, we fix a cluster tree $\mathcal{T}_\mathcal{I}$, a family $(\mathcal{D}_\ell)_{\ell=0}^\infty$ of hierarchical directions and a family $(\mathrm{sd}_\ell)_{\ell=0}^\infty$ of compatible son mappings.

Assume that a block $b = (t,s) \in \mathcal{T}_{\mathcal{I}\times\mathcal{I}}$ and a direction $c = c_b \in \mathcal{D}_t = \mathcal{D}_s$ is given. Replacing g in (2) with the directional approximation

$$\tilde{g}_{ts}(x,y) = \sum_{\nu\in M}\sum_{\mu\in M} g_c(\xi_{t,\nu},\xi_{s,\nu}) L_{tc,\nu}(x)\overline{L_{sc,\mu}(y)} \qquad \text{for all } x\in B_t,\ y\in B_s$$

as defined in (9) yields for the entries $(g_{ij})_{i\in\hat{t},j\in\hat{s}}$ of the stiffness matrix G

$$\begin{aligned} g_{ij} &= \int_\Gamma \varphi_i(x) \int_\Gamma g(x,y)\varphi_j(y)\,dy\,dx \approx \int_\Gamma \varphi_i(x)\int_\Gamma \tilde{g}_{ts}(x,y)\varphi_j(y)\,dy\,dx \\ &= \sum_{\nu\in M}\sum_{\mu\in M} \underbrace{g_c(\xi_{t,\nu},\xi_{s,\mu})}_{=:s_{b,\nu\mu}} \underbrace{\int_\Gamma \varphi_i(x) L_{tc,\nu}(x)\,dx}_{=:v_{tc,i\nu}} \underbrace{\int_\Gamma \varphi_j(y)\overline{L_{sc,\mu}(y)}\,dy}_{=\overline{v_{sc,j\mu}}} \qquad (10) \\ &= \sum_{\nu\in M}\sum_{\mu\in M} s_{b,\nu\mu} v_{tc,i\nu}\overline{v_{sc,j\mu}} = (V_{tc}S_bV_{sc}^*)_{ij} \qquad \text{for all } i\in\hat{t},\ j\in\hat{s}. \end{aligned}$$

Due to $S_b \in \mathbb{C}^{k\times k}$ with $k = \#M$, this is a factorized low-rank approximation

$$G|_{\hat{t}\times\hat{s}} \approx V_{tc}S_bV_{sc}^*. \qquad (11)$$

In order to handle the matrices V_{tc} efficiently, we rely on the ideas of fast multipole methods and $\mathcal{H}^2$-matrices and assume that these matrices have a multilevel representation: For t and a direction $c\in\mathcal{D}_t$ consider a son $t'\in\mathrm{sons}(t)$ and corresponding direction $c' = \mathrm{sd}_t(c)$ as introduced in Definition 4 and look for a (small) transfer matrix $E_{t'c}\in\mathbb{C}^{k\times k}$ such that

$$V_{tc}|_{\hat{t}'\times k} \approx V_{t'c'}E_{t'c}. \qquad (12)$$

This approximation brings about a complexity reduction since only the small matrices $E_{t'c}\in\mathbb{C}^{k\times k}$ need to be stored instead of $V_{tc}\in\mathbb{C}^{\hat{t}\times k}$. A motivation for the construction of $E_{t'c}$ is obtained from (10) and consists in replacing the function $L_{tc,\nu}(x) = \exp(\mathbf{i}\kappa\langle x,c\rangle)L_{t,\nu}(x)$ by an approximation. Setting $c' = \mathrm{sd}_t(c)$ we have

$$\begin{aligned} L_{ct,\nu}(x) &= \exp(\mathbf{i}\kappa\langle x,c\rangle)L_{t,\nu}(x) = \exp(\mathbf{i}\kappa\langle x,c'\rangle)\exp(\mathbf{i}\kappa\langle x,c-c'\rangle)L_{t,\nu}(x) \\ &\approx \exp(\mathbf{i}\kappa\langle x,c'\rangle)\mathfrak{I}_{B_{t'}}[\exp(\mathbf{i}\kappa\langle\cdot,c-c'\rangle)L_{t,\nu}](x) \qquad (13) \\ &= \exp(\mathbf{i}\kappa\langle x,c'\rangle)\sum_{\nu'\in M} \underbrace{\exp(\mathbf{i}\kappa\langle\xi_{t',\nu'},c-c'\rangle)L_{t,\nu}(\xi_{t',\nu'})}_{=:e_{t'c,\nu'\nu}} L_{t',\nu'}(x). \qquad (14) \end{aligned}$$

We will take this as the definition of $E_{t'c}$. The approximation step (13) has to be justified, and we will do this in Section 9 under the assumption that $c-c'$ is small (relative to κ and relative to the size of $B_{t'}$); in that case, the function $x\mapsto\exp(\mathbf{i}\kappa\langle x,c-c'\rangle)$ does not vary much, and we will show that the interpolation error is small on $B_{t'}$. This approach immediately yields

$$v_{tc,i\nu} = \int_\Gamma \varphi_i(x)L_{tc,\nu}(x)\,dx \approx \sum_{\nu'\in M} e_{t'c,\nu'\nu}\int_\Gamma \varphi_i(x)L_{t'c',\nu'}(x)\,dx = (V_{t'c'}E_{t'c})_{i\nu}$$

for all $i \in \hat{t}'$ and $\nu \in M$, which is equivalent to (12).

The notation $E_{t'c}$ is well-defined since the father $t \in \mathcal{T}_\mathcal{I}$ is uniquely determined by $t' \in \operatorname{sons}(t)$ due to the tree structure and the direction $c' = \operatorname{sd}_t(c) \in \mathcal{D}_{t'}$ is uniquely determined by $c \in \mathcal{D}_t$ due to our Definition 4.

Our (approximate) equation (12) gives rise to the following, purely algebraic definition.

Definition 7 (Directional cluster basis) *Let M be a finite index set, and let $V = (V_{tc})_{t\in\mathcal{T}_\mathcal{I}, c\in\mathcal{D}_t}$ be a family of matrices. We call it a* directional cluster basis *if*

- *$V_{tc} \in \mathbb{C}^{\hat{t}\times M}$ for all $t \in \mathcal{T}_\mathcal{I}$ and $c \in \mathcal{D}_t$, and*
- *there is a family $E = (E_{t'c})_{t\in\mathcal{T}_\mathcal{I}, t'\in\operatorname{sons}(t), c\in\mathcal{D}_t}$ such that*

$$V_{tc}|_{\hat{t}'\times k} = V_{t'c'}E_{t'c} \qquad \text{for all } t \in \mathcal{T}_\mathcal{I},\ t' \in \operatorname{sons}(t),\ c \in \mathcal{D}_t,\ c' = \operatorname{sd}_t(c). \tag{15}$$

The elements of the family E are called transfer matrices *for the directional cluster basis V, and $k := \#M$ is called its* rank.

We can now define the class of matrices that is the subject of this article: we denote the *leaves* of the block tree $\mathcal{T}_{\mathcal{I}\times\mathcal{I}}$ by

$$\mathcal{L}_{\mathcal{I}\times\mathcal{I}} := \{b \in \mathcal{T}_{\mathcal{I}\times\mathcal{I}} \ :\ \operatorname{sons}(b) = \emptyset\}.$$

The corresponding sets $\hat{b} \subseteq \mathcal{I}\times\mathcal{I}$ form a disjoint partition of $\mathcal{I}\times\mathcal{I}$, so a matrix G is uniquely determined by the submatrices $G|_{\hat{b}}$ for $b \in \mathcal{L}_{\mathcal{I}\times\mathcal{I}}$. For most of these submatrices, we can find an approximation of the form (11). These matrices are called *admissible* and collected in a subset

$$\mathcal{L}^+_{\mathcal{I}\times\mathcal{I}} := \{b \in \mathcal{L}_{\mathcal{I}\times\mathcal{I}} \ :\ b \text{ is admissible}\}.$$

The remaining blocks are called *inadmissible* and collected in the set

$$\mathcal{L}^-_{\mathcal{I}\times\mathcal{I}} := \mathcal{L}_{\mathcal{I}\times\mathcal{I}} \setminus \mathcal{L}^+_{\mathcal{I}\times\mathcal{I}}.$$

How to decide whether a block is admissible or not is the topic of Section 4.

Definition 8 (Directional $\mathcal{H}^2$-matrix) *Let V and W be directional cluster bases for $\mathcal{T}_\mathcal{I}$. Let $G \in \mathbb{C}^{\mathcal{I}\times\mathcal{I}}$ be a matrix. We call it a* directional $\mathcal{H}^2$-matrix *(or simply: an $\mathcal{DH}^2$-matrix) if there are families $S = (S_b)_{b\in\mathcal{L}^+_{\mathcal{I}\times\mathcal{I}}}$ and $(c_b)_{b\in\mathcal{L}^+_{\mathcal{I}\times\mathcal{I}}}$ such that*

- *$S_b \in \mathbb{C}^{k\times k}$ and $c_b \in \mathcal{D}_t = \mathcal{D}_s$ for all $b = (t, s) \in \mathcal{L}^+_{\mathcal{I}\times\mathcal{I}}$, and*
- *$G|_{\hat{t}\times\hat{s}} = V_{tc}S_bW^*_{sc}$ with $c = c_b$ for all $b = (t, s) \in \mathcal{L}^+_{\mathcal{I}\times\mathcal{I}}$.*

The elements of the family S are called coupling matrices*, and c_b is called the* block direction *for $b \in \mathcal{T}_{\mathcal{I}\times\mathcal{I}}$. The cluster bases V and W are called the* row cluster basis *and* column cluster basis*, respectively.*

*A $\mathcal{DH}^2$-*matrix representation *of a $\mathcal{DH}^2$-matrix G consists of V, W, S and the family $(G|_{\hat{b}})_{b\in\mathcal{L}^-_{\mathcal{I}\times\mathcal{I}}}$ of* nearfield matrices *corresponding to the inadmissible leaves of $\mathcal{T}_{\mathcal{I}\times\mathcal{I}}$.*

Remark 9 (Storage) *It is possible to prove that a $\mathcal{DH}^2$-matrix representation requires $\mathcal{O}(nk+\kappa^2k^2\log n)$ units of storage if the cluster tree is constructed by standard algorithms (e.g., [11, 4]) and the following conditions hold:*

$$|\{y\in\Gamma \ : \ \|x-y\|\leq r\}| \lesssim r^2, \qquad \text{for all } x\in\mathbb{R}^3,\ r\in\mathbb{R}_{\geq 0}, \tag{16a}$$

$$\operatorname{diam}^2(B_t) \lesssim |B_t\cap\Gamma| \qquad \text{for all } t\in\mathcal{T}_{\mathcal{I}}, \tag{16b}$$

$$\#\{t\in\mathcal{T}_{\mathcal{I}}^{(\ell)} \ : \ x\in B_t\} \lesssim 1 \qquad \text{for all } x\in\Gamma,\ \ell\in\mathbb{N}_0, \tag{16c}$$

$$\eta_2\operatorname{dist}(B_t,B_s)<\operatorname{diam}(B_t) \Longrightarrow \#\hat{s}\lesssim\#\hat{t} \qquad \text{for all } t\in\mathcal{L}_{\mathcal{I}},\ s\in\mathcal{T}_{\mathcal{I}},\ \operatorname{level}(t)=\operatorname{level}(s). \tag{16d}$$

Here $|X|$ denotes the surface measure of a measurable set $X\subseteq\Gamma$. The conditions (16a) and (16b) require Γ to be "essentially two-dimensional", the condition (16c) limits the overlap of clusters, and the condition (16d) corresponds to fairly weak mesh regularity. The proof can be found in [2, Theorem 11].

4 Admissibility and main result

In order to construct a $\mathcal{DH}^2$-matrix approximation, we have to find a cluster tree $\mathcal{T}_{\mathcal{I}}$, a block tree $\mathcal{T}_{\mathcal{I}\times\mathcal{I}}$, and a family of hierarchical directions $(\mathcal{D}_t)_{t\in\mathcal{T}_{\mathcal{I}}}$ such that

$$G|_{\hat{t}\times\hat{s}}\approx V_{tc}S_bW_{sc}^* \qquad \text{for all } b=(t,s)\in\mathcal{L}^+_{\mathcal{I}\times\mathcal{I}},\ c=c_b.$$

An analysis of the approximation (8) (cf., e.g., [17]) indicates that boxes B_t, B_s and the direction c have to satisfy the following three *admissibility conditions*:

$$\kappa\left\|\frac{m_t-m_s}{\|m_t-m_s\|}-c\right\|\leq\frac{\eta_1}{\max\{\operatorname{diam}(B_t),\operatorname{diam}(B_s)\}}, \tag{17a}$$

$$\kappa\max\{\operatorname{diam}^2(B_t),\operatorname{diam}^2(B_s)\}\leq\eta_2\operatorname{dist}(B_t,B_s),\ \text{and} \tag{17b}$$

$$\max\{\operatorname{diam}(B_t),\operatorname{diam}(B_s)\}\leq\eta_2\operatorname{dist}(B_t,B_s), \tag{17c}$$

where m_t and m_s denote the centers of the bounding boxes B_t and B_s and $\eta_1,\eta_2>0$ are parameters that can be chosen to balance storage requirements and accuracy.

The first condition (17a) ensures that the direction c of the plane-wave approximation is sufficiently close to the direction of the wave traveling from m_t to m_s.

The second condition (17b) is equivalent to

$$\frac{\max\{\operatorname{diam}(B_t),\operatorname{diam}(B_s)\}}{\operatorname{dist}(B_t,B_s)}\leq\frac{\eta_2}{\kappa\max\{\operatorname{diam}(B_t),\operatorname{diam}(B_s)\}};$$

it ensures that the angle between all vectors $x-y$ for $x\in B_t$ and $y\in B_s$ is bounded and that this bound shrinks when the wave number or the cluster diameter grows.

The third condition (17c) provides an upper bound for the same angle that is independent of the wave number and cluster diameter. This is the standard admissibility condition that is also used for the Laplace equation or linear elasticity.

In order to obtain a simple algorithm, we treat the first condition (17a) separately from the others: for each level ℓ of the cluster tree, we compute the maximal diameter

$$\delta_\ell := \max\{\operatorname{diam}(B_t) \ : \ t \in \mathcal{T}_\mathcal{I}^{(\ell)}\}$$

of all bounding boxes and then fix a set of directions $\mathcal{D}_\ell$

$$\|z - c\| \leq \frac{\eta_1}{\kappa\delta_\ell} \qquad \text{for all } z \in \mathbb{R}^3,\ \|z\| = 1. \tag{18}$$

Since all clusters on level ℓ share this set $\mathcal{D}_\ell$ of directions, the condition (17a) is guaranteed since $(m_t - m_s)/\|m_t - m_s\|$ is a unit vector for all $t, s \in \mathcal{T}_\mathcal{I}$. We also note that condition (18) is trivially satisfied with $c = 0$ for small δ_ℓ, viz., $\eta_1/(\kappa\delta_\ell) \geq 1$. We therefore require

$$c = 0 \qquad \text{if } \delta_\ell \leq \eta_1/\kappa. \tag{19}$$

In our numerical experiments, we construct the sets $\mathcal{D}_\ell$ by splitting the surface of the cube $[-1, 1]^3$ into squares with diameter $\leq 2\eta_2/(\kappa\delta_\ell)$, considering these squares' midpoints $\tilde{c}$, and projecting them by $c := \tilde{c}/\|\tilde{c}\|$ to the unit sphere. By construction, each point on the cube's surface has a distance of less than $\eta_2/(\kappa\delta_\ell)$ to one of the midpoints, and we only have to prove that the same holds for the points' projections to the unit sphere. This is a consequence of the following general result:

Lemma 10 (Projection) *Let $x, y \in \mathbb{R}^n$ with $\|x\|, \|y\| \geq 1$. We have*

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

Proof. [2, Lemma 6] □

The remaining admissibility conditions (17b) and (17c) do not depend on the directions and can be used to construct a suitable block tree by standard algorithms [11, 4].

With the Lagrange interpolation polynomials $L_{t,\nu}$ associated with the cluster t and the chosen polynomial interpolation scheme, our approximation scheme is summarized in Figures 2 and 3.

We close this section by assessing the error incurred by our approximation scheme. Our approximation scheme yields a matrix $\widetilde{G} \in \mathbb{C}^{\mathcal{I}\times\mathcal{I}}$ that is close to $G \in \mathbb{C}^{\mathcal{I}\times\mathcal{I}}$:

Theorem 11 *Let $t, s \in \mathcal{T}_\mathcal{I}$ and $c \in \mathcal{D}_t$ with* $\operatorname{level}(t) = \operatorname{level}(s)$ *satisfy the admissibility conditions (17) and (19).*

Let the 1D interpolation operator satisfy (23). Assume additionally that the bounding boxes B_t, $t \in \mathcal{T}_\mathcal{I}$, have the contraction property (46) for some $\overline{q} \in (0, 1)$.

Then there exist constants C, b, $K > 0$ depending only on the parameters η_1, η_2, $\overline{q}$ as well as C_Λ, λ of (23) such that under the side condition $m \geq K \log(\log \kappa)$ the following error estimates hold for the difference between the Galerkin matrix $G \in \mathbb{C}^{\mathcal{I}\times\mathcal{I}}$ and its approximation $\widetilde{G} \in \mathbb{C}^{\mathcal{I}\times\mathcal{I}}$:

$$\sup_{(i,j)\in\hat{t}\times\hat{s}} \frac{|g_{ij} - \widetilde{g}_{ij}|}{\|\varphi_i\|_{L^1(\Gamma)}\|\varphi_j\|_{L^1(\Gamma)}} \leq C \begin{cases} \exp(-bm)\frac{1}{\operatorname{dist}(B_t,B_s)} & \text{if } (t,s) \text{ is an admissible block,} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Follows from combining Theorems 27 and 30. □

procedure build_basis(t);
if sons(t) $= \emptyset$ **then**
  for $c \in \mathcal{D}_t$ **do**
    for $i \in \hat{t}$, $\nu \in M$ **do**
      $v_{tc,i\nu} \leftarrow \int_\Gamma \varphi_i(x) \exp(\mathbf{i}\kappa\langle x, c\rangle) L_{t,\nu}(x)\, dx$
else begin
  for $t' \in \mathrm{sons}(t)$ **do** build_basis(t');
  for $c \in \mathcal{D}_t$ **do begin**
    $c' \leftarrow \mathrm{sd}_t(c)$;
    for $\nu, \nu' \in M$ **do**
      $e_{t'c,\nu'\nu} \leftarrow \exp(\mathbf{i}\kappa\langle \xi_{t',\nu'}, c - c'\rangle) L_{t,\nu}(\xi_{t',\nu'})$
  end
end

Figure 2: Construction of the directional cluster basis

procedure build_coupling($b = (t, s)$);
if sons(b) $= \emptyset$ **then**
  if b satisfies (17b) and (17c) **then begin**
    Choose $c_b = c \in \mathcal{D}_t = \mathcal{D}_s$ with (17a);
    for $\nu, \mu \in M$ **do begin**
      $r \leftarrow \xi_{t,\nu} - \xi_{s,\mu}$;
      $s_{b,\nu\mu} \leftarrow \frac{\exp(\mathbf{i}\kappa(\|r\| - \langle r, c\rangle))}{4\pi\|r\|}$
    end
  end
  else
    for $i \in \hat{t}$, $j \in \hat{s}$ **do** $g_{ij} \leftarrow \int_\Gamma \varphi_i(x) \int_\Gamma g(x, y)\varphi_j(y)\, dy\, dx$
end
else
  for $b' = (t', s') \in \mathrm{sons}(b)$ **do** build_coupling($b' = (t', s')$)
end

Figure 3: Construction of the directional cluster basis

5 Matrix-vector multiplication

Let G be a $\mathcal{DH}^2$-matrix for the directional cluster bases V and W, and let $x \in \mathbb{C}^{\mathcal{I}}$. For an efficient evaluation of the matrix-vector product $y = Gx$. we follow the familiar approach of fast multipole and $\mathcal{H}^2$-matrix techniques: since the submatrices are factorized into three terms

$$G|_{\hat{t}\times\hat{s}} = V_{tc} S_b W_{sc}^* \qquad \text{for all } b = (t,s) \in \mathcal{L}^+_{\mathcal{I}\times\mathcal{I}},$$

the algorithm is split into three phases: in the first phase, called the *forward transformation*, we multiply by W_{sc}^* and compute

$$\widehat{x}_{sc} = W_{sc}^* x|_{\hat{s}} \qquad \text{for all } s \in \mathcal{T}_{\mathcal{I}},\ c \in \mathcal{D}_s; \tag{20a}$$

in the second phase, the *coupling step*, we multiply these coefficient vectors by the coupling matrices S_b and obtain

$$\widehat{y}_{tc} := \sum_{\substack{b=(t,s)\in\mathcal{L}^+_{\mathcal{I}\times\mathcal{I}} \\ c=c_b}} S_b \widehat{x}_{sc} \qquad \text{for all } t \in \mathcal{T}_{\mathcal{I}},\ c \in \mathcal{D}_t; \tag{20b}$$

and in the final phase, the *backward transformation*, we multiply by V_{tc} to get the result

$$y_i = \sum_{\substack{t\in\mathcal{T}_{\mathcal{I}},\ c\in\mathcal{D}_t \\ i\in\hat{t}}} (V_{tc}\widehat{y}_{tc})_i \qquad \text{for all } i \in \mathcal{I}. \tag{20c}$$

The first and third phase can be handled efficiently by using the transfer matrices $E_{t'c}$: let $s \in \mathcal{T}_{\mathcal{I}}$ with $\operatorname{sons}(s) \neq \emptyset$, and let $c \in \mathcal{D}_s$. Due to Definition 1, the set $\{\hat{s}' \ :\ s' \in \operatorname{sons}(s)\}$ is a disjoint partition of the index set $\hat{s}$. Combined with (15), this implies

$$W_{sc}^* x|_{\hat{s}} = \sum_{s'\in\operatorname{sons}(s)} (W_{sc}|_{\hat{s}'\times k})^* x|_{\hat{s}'} = \sum_{s'\in\operatorname{sons}(s)} E_{s'c}^* V_{s'c'}^* x|_{\hat{s}'} = \sum_{s'\in\operatorname{sons}(s)} E_{s'c}^* \widehat{x}_{s'c'},$$

and we can prepare *all* coefficient vectors $\widehat{x}_{sc}$ by the simple recursion given on the left of Figure 4. By similar arguments we find that the third phase can also be handled by the recursion given on the right of Figure 4.

The submatrices corresponding to inadmissible leaves $b = (t,s) \in \mathcal{L}^-_{\mathcal{I}\times\mathcal{I}}$ are stored as standard arrays and can be evaluated accordingly.

We see that the algorithms use each of the matrices of the $\mathcal{DH}^2$-matrix representation exactly once, so the bound provided by Remark 9 for the storage requirements yields an $\mathcal{O}(nk + \kappa^2 k^2 \log n)$ complexity of a matrix-vector multiplication.

6 Interpolation error analysis

Our goal is to prove that our approximation scheme converges exponentially and that the rate of convergence does not depend on the mesh resolution or the wave number κ.

procedure forward(s, x, **var** $\widehat{x}$);
if sons(s) $= \emptyset$ **then**
  for $c \in \mathcal{D}_s$ **do** $\widehat{x}_{sc} \leftarrow W_{sc}^* x|_{\hat{s}}$
else begin
  for $s' \in$ sons(s) **do** forward(s', x, $\widehat{x}$);
  for $c \in \mathcal{D}_s$ **do begin**
    $\widehat{x}_{sc} \leftarrow 0$;
    for $s' \in$ sons(s) **do**
      $\widehat{x}_{sc} \leftarrow \widehat{x}_{sc} + E_{s'c}^* \widehat{x}_{s'c'}$
  end
end

procedure backward(t, **var** $\widehat{y}$, y);
if sons(t) $= \emptyset$ **then**
  for $c \in \mathcal{D}_s$ **do** $y|_{\hat{t}} \leftarrow y|_{\hat{t}} + V_{tc} \widehat{y}_{tc}$
else begin
  for $c \in \mathcal{D}_t$ **do**
    for $t' \in$ sons(t) **do**
      $\widehat{y}_{t'c'} \leftarrow \widehat{y}_{t'c'} + E_{t'c} \widehat{y}_{tc}$;
  for $t' \in$ sons(t) **do** backward(t', $\widehat{y}$, y)
end

Figure 4: Fast forward and backward transformation

We use the maximum norm

$$\|f\|_{\infty,B} := \max\{|f(x)| \ : \ x \in B\} \qquad \text{for all } f \in C(B)$$

for compact sets B to measure the approximation error.

For bounding boxes $B_t, B_s \subseteq \mathbb{R}^3$ and a direction $c \in \mathbb{R}^3$, we immediately find

$$|\exp(\mathbf{i}\kappa\langle x - y, c\rangle)| = 1 \qquad \text{for all } x \in B_t,\ y \in B_s,$$

and we can conclude that multiplication with a plane wave does not change the maximum norm. This implies

$$\|g - \tilde{g}_{ts}\|_{\infty,B_t \times B_s} = \|g_c - \tilde{g}_{c,ts}\|_{\infty,B_t \times B_s} \tag{21}$$

for the approximations $\tilde{g}_{ts}$ and $\tilde{g}_{c,ts}$ defined in (9) and (8). This equation allows us to focus on interpolation error estimates for the modified function g_c.

6.1 Tensor interpolation

The error analysis of our scheme has to gauge two sources of error: the interpolation error associated with (8) and the interpolation error arising from (13). Both cases require error estimates for tensor interpolation.

Definition 12 (Lebesgue constant) *The* Lebesgue constant *of the interpolation operator* $\mathfrak{I} : C([-1,1]) \to \Pi_m$ *is given by*

$$\Lambda_m := \max\left\{ \sum_{\nu=0}^{m} |L_\nu(x)| \ : \ x \in [-1,1] \right\}$$

and is the optimal constant for the stability estimate

$$\|\mathfrak{I}[f]\|_{\infty,[-1,1]} \leq \Lambda_m \|f\|_{\infty,[-1,1]} \qquad \textit{for all } f \in C([-1,1]). \tag{22}$$

Remark 13 (Chebyshev interpolation) *A good choice for $\mathfrak{I}$ is interpolation in the Chebyshev points*

$$\xi_\nu = \cos\left(\frac{2\nu+1}{2m+2}\right) \qquad \text{for all } \nu \in \{0,\dots,m\}.$$

According to [19], we have $\Lambda_m \leq \frac{2}{\pi}\ln(m+1)+1$, and this is very close to the best stability constant a Lagrange interpolation scheme can reach.

The convergence theory of the present paper is formulated more generally for interpolation operators $\mathfrak{I}$ that satisfy the condition

$$\Lambda_m \leq C_\Lambda (m+1)^\lambda \qquad \text{for all } m \in \mathbb{N}_0, \tag{23}$$

where $C_\Lambda \in \mathbb{R}_{>0}$ and $\lambda \in \mathbb{R}_{\geq 1}$ are independent constants. In the case of Chebyshev interpolation, we can use $C_\Lambda = \lambda = 1$.

Recall the tensor product interpolation operator $\mathfrak{I}_B$ from (5) that is obtained from the 1D interpolation operator $\mathfrak{I}$. We will analyze the approximation properties of $\mathfrak{I}_B$ in terms of one-dimensional interpolation operators. To that end, we consider for an axis-parallel n-dimensional box

$$B = [a_1, b_1] \times \cdots \times [a_n, b_n] \tag{24}$$

the *partial interpolation operators* for each coordinate direction $\iota \in \{1,\dots,n\}$ by applying $\mathfrak{I}_{[a_\iota,b_\iota]}$ only in this direction. The corresponding operators are given by

$$\begin{aligned}\mathfrak{I}_{B,\iota} &: C(B) \to C(B),\\ &f \mapsto \left(x \mapsto \sum_{\nu=0}^{m} f(x_1,\dots,x_{\iota-1},\xi_{[a_\iota,b_\iota],\nu},x_{\iota+1},\dots,x_n) L_{[a_\iota,b_\iota],\nu}(x_\iota)\right).\end{aligned}$$

A simple induction shows that the operator $\mathfrak{I}_B$ of (5) can be written as

$$\mathfrak{I}_B := \mathfrak{I}_{B,1} \circ \cdots \circ \mathfrak{I}_{B,n} = \mathfrak{I}_{B,n} \circ \cdots \circ \mathfrak{I}_{B,1}. \tag{25}$$

It is easy to see that, if the underlying interpolation operator $\mathfrak{I} : C([-1,1]) \to \Pi_m$ satisfies (22), then

$$\begin{aligned}\|\mathfrak{I}_{[a,b]}[f]\|_{\infty,[a,b]} &\leq \Lambda_m \|f\|_{\infty,[a,b]} && \text{for all } f \in C([a,b]) \text{ and}\\ \|\mathfrak{I}_{B,\iota}[f]\|_{\infty,B} &\leq \Lambda_m \|f\|_{\infty,B} && \text{for all } f \in C(B),\ \iota \in \{1,\dots,n\}.\end{aligned} \tag{26}$$

This leads to error estimates for the tensor interpolation operator:

Lemma 14 (Tensor interpolation) *Let B be given by (24). For $f \in C(B)$ define*

$$f_{x,\iota} : [-1,1] \to \mathbb{R}, \qquad t \mapsto f\left(x_1,\dots,x_{\iota-1},\Phi_{[a_\iota,b_\iota]}(t),x_{\iota+1},\dots,x_n\right) \tag{27}$$

for all $x \in B$ and $\iota \in \{1,\dots,n\}$. If we have

$$\|f_{x,\iota} - \mathfrak{I}[f_{x,\iota}]\|_{\infty,[-1,1]} \leq \epsilon \qquad \text{for all } x \in B,\ \iota \in \{1,\dots,n\}, \tag{28}$$

then

$$\|f - \mathfrak{I}_B[f]\|_{\infty,B} \leq n\Lambda_m^{n-1}\epsilon.$$

Proof. Let $x \in B$, $\iota \in \{1, \dots, n\}$, and $t := \Phi^{-1}_{[a_\iota, b_\iota]}(x_\iota)$. By definition (27) we have

$$f_{x,\iota}(t) = f(x_1, \dots, x_{\iota-1}, \Phi_{[a_\iota,b_\iota]}(t), x_{\iota+1}, \dots, x_n) = f(x),$$

and a straightforward computation gives us

$$\mathfrak{I}[f_{x,\iota}](t) = \mathfrak{I}_{B,\iota}[f](x),$$

so the assumption (28) implies

$$\|f - \mathfrak{I}_{B,\iota}[f]\|_{\infty,B} \leq \epsilon \qquad \text{for all } \iota \in \{1, \dots, d\}.$$

Using a telescoping sum, we obtain with the stability property (26)

$$\begin{aligned}\|f - \mathfrak{I}_B[f]\|_{\infty,B} &\leq \sum_{\iota=1}^{n} \|\mathfrak{I}_{B,1} \circ \cdots \circ \mathfrak{I}_{B,\iota-1}[f - \mathfrak{I}_{B,\iota}[f]]\|_{\infty,B} \\ &\leq \sum_{\iota=1}^{n} \Lambda_m^{\iota-1} \|f - \mathfrak{I}_{B,\iota}[f]\|_{\infty,B} \leq n\Lambda_m^{n-1}\epsilon,\end{aligned}$$

which is the required error estimate. □

6.2 Reduction to univariate approximation

We have seen in (21) that an interpolation error estimate for g_c implies an estimate for the directional approximation $\tilde{g}_{ts}$ of g.

In order to apply Lemma 14, we let

$$f : B_t \times B_s \to \mathbb{C}, \qquad (x, y) \mapsto g_c(x, y).$$

We have to investigate the functions $f_{(x,y),\iota}$ introduced in (27) for $\iota \in \{1, \dots, 6\}$. If $\iota = 1$ holds, the first component of x varies in $[a_{t,1}, b_{t,s}]$, and we have

$$f_{(x,y),\iota}(t) = \frac{\exp(\mathbf{i}\kappa(\|d - tp\| - \langle d - tp, c\rangle))}{\|d - tp\|} \qquad \text{for all } t \in [-1, 1], \tag{29}$$

with the vectors

$$d := \begin{pmatrix} (a_{t,1} + b_{t,1})/2 - y_1 \\ x_2 - y_2 \\ x_3 - y_3 \end{pmatrix}, \qquad p := \begin{pmatrix} (a_{t,1} - b_{t,1})/2 \\ 0 \\ 0 \end{pmatrix}.$$

For $\iota \in \{2, 3\}$, we obtain similar results by considering the second and third component of the vectors instead of the first.

For $\iota = 4$, the first component of y varies in $[a_{s,1}, b_{s,1}]$, and we can again use (29) by choosing the vectors

$$d := \begin{pmatrix} x_1 - (a_{s,1} + b_{s,1})/2 \\ x_2 - y_2 \\ x_3 - y_3 \end{pmatrix}, \qquad p := \begin{pmatrix} (a_{s,1} - b_{s,1})/2 \\ 0 \\ 0 \end{pmatrix}.$$

We can handle $\iota \in \{5,6\}$ in a similar fashion.

In all six cases, we have

$$\|p\| \leq \frac{\max\{\operatorname{diam}(B_t), \operatorname{diam}(B_s)\}}{2}, \tag{30a}$$

and since $\Phi_{[a,b]}$ maps into $[a,b]$, we also have

$$d - \tau p \in B_t - B_s = \{x - y \ : \ x \in B_t, \ y \in B_s\} \qquad \text{for all } \tau \in [-1,1]. \tag{30b}$$

Combining these inequalities with the admissibility conditions (17) yields bounds that can be used to estimate the interpolation error of $f_{(x,y),\iota}$.

In order to keep the notation simple and the results general, we consider general vectors $p, d \in \mathbb{R}^3$ satisfying the conditions (30a) and (30b).

Lemma 15 (Univariate formulation) *Let $t, s \in \mathcal{T}_{\mathcal{I}}$ satisfy the admissibility conditions (17). Let $\epsilon \in \mathbb{R}_{\geq 0}$.*

For $p, d \in \mathbb{R}^3$ satisfying (30a) and (30b), we define

$$g_1 : [-1,1] \to \mathbb{C}, \qquad t \mapsto \frac{\exp(\mathbf{i}\kappa(\|d - tp\| - \langle d - tp, c\rangle))}{4\pi\|d - tp\|}. \tag{31}$$

If we have

$$\|g_1 - \mathfrak{I}[g_1]\|_{\infty,[-1,1]} \leq \epsilon \tag{32}$$

for all such $p, d \in \mathbb{R}^3$, we find for $\tilde{g}_{ts}$ given by (9)

$$\|g - \tilde{g}_{ts}\|_{\infty, B_t \times B_s} \leq 6\Lambda_m^5 \epsilon.$$

Proof. We have already seen that $f_{(x,y),\iota}$, $\iota \in \{1,\ldots,6\}$, corresponds to specific vectors $d, p \in \mathbb{R}^3$ satisfying (30a) and (30b). Since we have (32) at our disposal for *all* pairs of vectors satisfying these conditions, we can apply Lemma 14 to obtain the required estimate. □

7 Holomorphic extension of g_1

In order to obtain bounds for the interpolation error of the functions g_1 defined in (31), we consider its holomorphic extension into a neighbourhood of the interval $[-1,1]$.

This extension can be constructed by combining holomorphic extensions of the inner product, the Euclidean norm, and the exponential function.

For the inner product, we use the simple approach

$$\langle x, y\rangle = \sum_{j=1}^{n} x_i y_i \qquad \text{for all } x, y \in \mathbb{C}^n.$$

This is *not* a sequilinear form, but the restriction to the real subspace $\mathbb{R}^n$ is the standard Euclidean inner product.

The exponential function can be extended to the entire complex plane by using its globally convergent power series representation.

Finally, the holomorphic extension of the Euclidean norm

$$\|x\| = \sqrt{\langle x, x\rangle} \qquad \text{for all } x \in \mathbb{R}^n$$

requires a suitable extension of the square root, which cannot be defined in all of $\mathbb{C}$. We choose the principal branch given by

$$\sqrt{z} = \sqrt{|z|}\frac{z + |z|}{|z + |z||} \qquad \text{for all } z \in \mathbb{C} \setminus \mathbb{R}_{\leq 0}, \tag{33}$$

which is holomorphic in $\mathbb{C} \setminus \mathbb{R}_{\leq 0}$. In order to identify a subset of $\mathbb{C}$ in which $z \mapsto \sqrt{\langle d - zp, d - zp\rangle}$ is holomorphic, we have to determine the values $z \in \mathbb{C}$ satisfying

$$\langle d - zp, d - zp\rangle \notin \mathbb{R}_{\leq 0}.$$

Lemma 16 (Holomorphic extension) *Let $n \in \mathbb{N}$ and $d, p \in \mathbb{R}^n$ with $p \neq 0$ and define*

$$w_r := \langle d, p\rangle / \|p\|^2, \qquad w_i := \sqrt{\|d\|^2/\|p\|^2 - w_r^2}, \qquad w := w_r + iw_i,$$
$$U_d := \mathbb{C} \setminus \{w_r + iy \ : \ y \in \mathbb{R},\ |y| \geq w_i\}.$$

We have

$$\langle d - zp, d - zp\rangle = \|p\|^2(w - z)(\bar{w} - z) \qquad \text{for all } z \in \mathbb{C}. \tag{34}$$

The function

$$f : U_d \to \mathbb{C} \setminus \{0\}, \qquad z \mapsto \sqrt{\langle d - zp, d - zp\rangle} = \|p\|\sqrt{(w - z)(\bar{w} - z)},$$

is well-defined and holomorphic.

Proof. The equality (34) follows from a direct computation using $d, p \in \mathbb{R}^n$ and $|w| = \|d\|/\|p\|$: We have

$$\begin{aligned}\langle d - zp, d - zp\rangle &= \|d\|^2 - 2\langle d, p\rangle z + \|p\|^2 z^2 = \|p\|^2|w|^2 - \|p\|^2(w + \bar{w})z + \|p\|^2 z^2 \\ &= \|p\|^2(w - z)(\bar{w} - z).\end{aligned}$$

In order to show that f is well-defined, it suffices to demonstrate

$$z \in U_d \Longrightarrow \langle d - zp, d - zp\rangle \in \mathbb{C} \setminus \mathbb{R}_{\leq 0} \qquad \text{for all } z \in \mathbb{C}.$$

We use contraposition: let $z \in \mathbb{C}$ be such that $\langle d - zp, d - zp\rangle \in \mathbb{R}_{\leq 0}$. We will show that this implies $z \notin U_d$. Let $x, y \in \mathbb{R}$ be such that $z = x + iy$.

$$\begin{aligned}\langle d - zp, d - zp\rangle &= \|p\|^2(w - z)(\bar{w} - z) \\ &= \|p\|^2((w_r - x) + i(w_i - y))\,((w_r - x) + i(-w_i - y))\end{aligned}$$

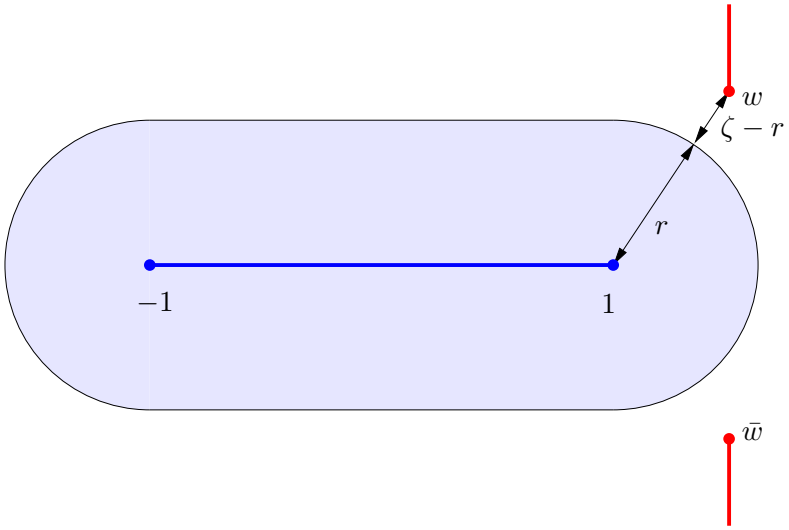

Figure 5: Domain U_r in relation to the interval $[-1, 1]$ and $\{w, \bar{w}\}$

$$= \|p\|^2((w_r - x)^2 - 2i(w_r - x)y + w_i^2 - y^2).$$

Due to $\langle d - zp, d - zp\rangle \in \mathbb{R}_{\leq 0}$, the imaginary part vanishes and the real part is non-positive, i.e., we have

$$0 = 2(w_r - x)y, \qquad 0 \geq (w_r - x)^2 + w_i^2 - y^2 \tag{35}$$

due to $p \neq 0$. If $y = 0$ holds, the inequality implies

$$0 \geq (w_r - x)^2 + w_i^2 \geq 0,$$

and we get $0 = w_r - x$ and $0 = w_i \leq |y|$, i.e., $z \notin U_d$.

Otherwise, i.e., if $y \neq 0$ holds, the equation in (35) yields $w_r - x = 0$, and the inequality gives us

$$y^2 \geq (w_r - x)^2 + w_i^2 = w_i^2,$$

and therefore $|y| \geq w_i$ and again $z \notin U_d$.

Contraposition yields that $z \in U_d$ implies $\langle d - zp, d - zp\rangle \in \mathbb{C} \setminus \mathbb{R}_{\leq 0}$.

Since $z \mapsto \langle d - zp, d - zp\rangle$ is holomorphic in U_d and maps into the domain of the holomorphic principal square root, the composed function f is also holomorphic. □

In order to obtain bounds for the derivatives of g_1 and the corresponding interpolation error, we require bounds for its holomorphic extension. We start by investigating the extension f of the norm $t \mapsto \|d - tp\|$.

Lemma 17 (Norm estimates) *Let $n \in \mathbb{N}$ and $d, p \in \mathbb{R}^n$ with $p \neq 0$, and let f, w and U_d be defined as in Lemma 16.*

We define

$$\zeta_t := \frac{\|d - tp\|}{\|p\|}, \qquad U_{t,r} := \{z \in \mathbb{C} \ : \ |z - t| \le r\} \quad \text{for all } t, r \in \mathbb{R}, \tag{36a}$$

$$\zeta := \min\{\zeta_t \ : \ t \in [-1, 1]\}, \quad U_r := \bigcup_{t \in [-1,1]} U_{t,r} \qquad \text{for all } r \in \mathbb{R}. \tag{36b}$$

Then we have

$$\begin{aligned} U_{t,r} &\subseteq U_d, & |f(z)| &\ge \|p\|(\zeta_t - r) & &\text{for all } t \in \mathbb{R},\ r \in [0, \zeta_t),\ z \in U_{t,r}, \\ U_r &\subseteq U_d, & |f(z)| &\ge \|p\|(\zeta - r) & &\text{for all } r \in [0, \zeta),\ z \in U_r. \end{aligned}$$

Proof. Due to (34), we have

$$\|d - tp\|^2 = \|p\|^2 (w - t)(\bar{w} - t) = \|p\|^2 (w - t)\overline{(w - t)} = \|p\|^2 |w - t|^2 \quad \text{for all } t \in \mathbb{R}.$$

Let $t \in \mathbb{R}$. We have

$$\zeta_t^2 = \frac{\|d - tp\|^2}{\|p\|^2} = \frac{\|p\|^2 |w - t|^2}{\|p\|^2} = |w - t|^2 \tag{37}$$

and therefore $\zeta_t = |w - t|$.

Let $\zeta_t > 0$. In order to prove $U_{t,r} \subseteq U_d$, we only have to show that $w_r + iy \notin U_{t,r}$ holds for all $y \in \mathbb{R}$ with $|y| \ge w_i$. Let $z := w_r + iy$ with $y \in \mathbb{R}$ and $|y| \ge w_i$. Due to (37), we have

$$|z - t|^2 = (w_r - t)^2 + y^2 \ge (w_r - t)^2 + w_i^2 = |w - t|^2 = \zeta_t^2,$$

and this implies $z \notin U_{t,r}$ for all $r \in [0, \zeta_t)$.

Now that we have proven $U_{t,r} \subseteq U_d$, we can consider the lower bound. Let $r \in [0, \zeta)$ and $z \in U_{t,r}$. Using again (37), we find

$$\begin{aligned} |w - z| &= |w - t + t - z| \ge |w - t| - |t - z| \ge \zeta_t - r > 0, \\ |\bar{w} - z| &= |\bar{w} - t + t - z| \ge |\bar{w} - t| - |t - z| = |w - t| - |t - z| \ge \zeta_t - r > 0, \end{aligned}$$

and this implies

$$|f(z)| = \|p\| \sqrt{|w - z|\,|\bar{w} - z|} \ge \|p\|(\zeta_t - r).$$

Assume now $\zeta > 0$, and let $r \in [0, \zeta)$ and $z \in U_r$. By definition, we can find $t \in [-1, 1]$ such that $z \in U_{t,r}$. Since $\zeta = \min_{\tau \in [-1,1]} \zeta_\tau \le \zeta_t$, we get

$$z \in U_d, \qquad |f(z)| \ge \|p\|(\zeta_t - r) \ge \|p\|(\zeta - r),$$

and the proof is complete. □

We also require a bound for the exponential function appearing in (1). Due to

$$|\exp(x + \mathbf{i}y)| = |\exp(x)|\,|\exp(\mathbf{i}y)| \le \exp(|x|) \qquad \text{for all } x, y \in \mathbb{R},$$

we only have to find an estimate for the real part of the argument of the exponential function. Introducing

$$f_e : U_d \to \mathbb{C}, \qquad z \mapsto f(z) - \langle d - zp, c\rangle, \tag{38}$$

we have

$$\begin{aligned}\exp(\mathbf{i}\kappa(f(z) - \langle d - zp, c\rangle)) &= \exp(\mathbf{i}\kappa f_e(z)),\\ |\exp(\mathbf{i}\kappa(f(z) - \langle d - zp, c\rangle))| &\leq \exp(\kappa|\Im(f_e(z))|),\end{aligned}$$

and our next goal is to find an upper bound for $|\Im(f_e(z))|$. Following the approach of [17], we apply a Taylor expansion of f around $t \in [-1, 1]$ to obtain the required estimate.

Lemma 18 *Let $\zeta \in \mathbb{R}_{>0}$ and $r \in [0, \zeta)$. We have*

$$\int_0^1 \frac{1-s}{(\zeta - rs)^3}\, ds = \frac{1}{2\zeta^2(\zeta - r)}.$$

Proof. The proof is straightforward for $r = 0$. For $r > 0$, the function

$$h : [0, 1] \to \mathbb{R}, \qquad s \mapsto \frac{(1 + \zeta/r) - 2s}{2r(\zeta - rs)^2},$$

is the antiderivative of the integrand, and we can use the fundamental theorem of calculus to obtain the equation. □

Lemma 19 (Taylor expansion) *Let $d, p \in \mathbb{R}^3$ with $p \neq 0$ and $c \in \mathbb{R}^3$ with $\|c\| = 1$. Let $t \in \mathbb{R}$ and*

$$\zeta_t := \frac{\|d - tp\|}{\|p\|} > 0$$

as before. Let $r \in [0, \zeta_t)$, and let $U_{t,r}$, U_d and f be defined as in Lemmas 16, 17. Let f_e be given by (38). Then, we have for $z \in U_{t,r}$ the representation

$$f_e(z) = \left\langle d - zp, \frac{d - tp}{\|d - tp\|} - c \right\rangle + \int_0^1 \frac{\|d\|^2\|p\|^2 \sin^2 \angle(d, p)}{f(t + (z - t)s)^3}(1 - s)\, ds(z - t)^2. \tag{39}$$

The imaginary part of $f_e(z)$ can be bounded by

$$|\Im(f_e(z))| \leq \|p\| \left(\left\| \frac{d - tp}{\|d - tp\|} - c \right\| r + \frac{1}{2(\zeta_t - r)} r^2 \right). \tag{40}$$

Proof. Let $z \in U_{t,r}$. Due to Lemma 17, this implies $z \in U_d$. We have

$$\begin{aligned} f(z) &= \sqrt{\langle d - zp, d - zp\rangle} = \langle d - zp, d - zp\rangle^{1/2},\\ f'(z) &= -\frac{\langle p, d - zp\rangle}{\langle d - zp, d - zp\rangle^{1/2}} = \frac{-\langle p, d - zp\rangle}{f(z)}, \end{aligned}$$

$$f''(z) = \frac{\langle p,p\rangle f(z) + \langle p, d-zp\rangle f'(z)}{f(z)^2} = \frac{\langle p,p\rangle f(z) - \langle p, d-zp\rangle^2/f(z)}{f(z)^2}$$
$$= \frac{\langle p,p\rangle f(z)^2 - \langle p, d-zp\rangle^2}{f(z)^3} = \frac{\langle d-zp, d-zp\rangle\langle p,p\rangle - \langle p, d-zp\rangle^2}{f(z)^3}.$$

We use a Taylor expansion of $f(z)$ around t. More precisely, with the parametrization

$$\hat{z} : [0,1] \to U_d, \qquad\qquad s \mapsto \hat{z}_s := t + (z-t)s,$$

we have

$$f(\hat{z}_0) = f(t), \qquad f(\hat{z}_1) = f(z), \qquad \hat{z}'(s) = z - t \qquad \text{for all } s \in [0,1],$$

and the Taylor expansion of $f \circ \hat{z}$ around $s = 0$ yields

$$f(z) = (f \circ \hat{z})(1) = f(t) + f'(t)(z-t) + \int_0^1 f''(\hat{z}_s)(1-s)\, ds (z-t)^2.$$

Hence, we obtain for the function f_e

$$\begin{aligned}
f_e(z) &= f(z) - \langle d - zp, c\rangle \\
&= f(t) + f'(t)(z-t) + \int_0^1 f''(\hat{z}_s)(1-s)\, ds (z-t)^2 - \langle d-zp, c\rangle \\
&= \|d-tp\| - \langle d-zp, c\rangle - \frac{\langle d-tp, p\rangle}{\|d-tp\|}(z-t) \\
&\quad + \int_0^1 \frac{\langle d-\hat{z}_s p, d-\hat{z}_s p\rangle\langle p,p\rangle - \langle p, d-\hat{z}_s p\rangle^2}{f(\hat{z}_s)^3}(1-s)\, ds (z-t)^2 \\
&= \left\langle d-tp, \frac{d-tp}{\|d-tp\|}\right\rangle - \langle d-zp, c\rangle - \left\langle (z-t)p, \frac{d-tp}{\|d-tp\|}\right\rangle \\
&\quad + \int_0^1 \frac{\langle d-\hat{z}_s p, d-\hat{z}_s p\rangle\langle p,p\rangle - \langle p, d-\hat{z}_s p\rangle^2}{f(\hat{z}_s)^3}(1-s)\, ds (z-t)^2 \\
&= \left\langle d-zp, \frac{d-tp}{\|d-tp\|} - c\right\rangle \\
&\quad + \int_0^1 \frac{\langle d-\hat{z}_s p, d-\hat{z}_s p\rangle\langle p,p\rangle - \langle p, d-\hat{z}_s p\rangle^2}{f(\hat{z}_s)^3}(1-s)\, ds (z-t)^2.
\end{aligned}$$

We take a closer look at the integrand. For any $z \in \mathbb{C}$, we find

$$\begin{aligned}
&\langle d-zp, d-zp\rangle\langle p,p\rangle - \langle d-zp, p\rangle^2 \\
&\quad = (\langle d,d\rangle - 2z\langle d,p\rangle + z^2\langle p,p\rangle)\langle p,p\rangle - (\langle d,p\rangle^2 - 2z\langle d,p\rangle\langle p,p\rangle + z^2\langle p,p\rangle^2) \\
&\quad = \|d\|^2\|p\|^2 - 2z\langle d,p\rangle\|p\|^2 + z^2\|p\|^4 - \langle d,p\rangle^2 + 2z\langle d,p\rangle\|p\|^2 - z^2\|p\|^4 \\
&\quad = \|d\|^2\|p\|^2 - \langle d,p\rangle^2 = \|d\|^2\|p\|^2 \sin^2 \angle(d,p). \qquad (41)
\end{aligned}$$

With this equation, we obtain (39).

For any $s \in [0,1]$, applying (41) twice yields

$$\begin{aligned}|\langle d-\hat{z}_s p, d-\hat{z}_s p\rangle\langle p,p\rangle - \langle d-\hat{z}_s p, p\rangle^2| &= \|d\|^2\|p\|^2 - \langle d,p\rangle^2 \\ &= \langle d-tp, d-tp\rangle\langle p,p\rangle - \langle d-tp,p\rangle^2 = \|d-tp\|^2\|p\|^2 - \langle d-tp,p\rangle^2 \\ &= \|d-tp\|^2\|p\|^2 \sin^2\angle(d-tp,p) \leq \|d-tp\|^2\|p\|^2,\end{aligned}$$

and we obtain

$$\left|\frac{\langle d-\hat{z}_s p, d-\hat{z}_s p\rangle\langle p,p\rangle - \langle p, d-\hat{z}_s p\rangle^2}{f(\hat{z}_s)^3}\right| \leq \frac{\|d-tp\|^2\|p\|^2}{|f(\hat{z}_s)^3|}.$$

By definition, we have $|\hat{z}_s - t| = |(z-t)s| \leq rs$ and therefore $\hat{z}_s \in U_{t,rs}$. Hence, Lemma 17 yields $|f(\hat{z}_s)| \geq \|p\|(\zeta_t - rs)$, and we can use Lemma 18 to find

$$\begin{aligned}&\left|\int_0^1 \frac{\langle d-\hat{z}_s p, d-\hat{z}_s p\rangle\langle p,p\rangle - \langle p, d-\hat{z}_s p\rangle^2}{f(\hat{z}_s)^3}(1-s)\,ds(z-t)^2\right| \\ &\quad\leq \frac{\|d-tp\|^2\|p\|^2}{\|p\|^3}\int_0^1 \frac{1-s}{(\zeta_t - rs)^3}\,ds|z-t|^2 \leq \|p\|\frac{\|d-tp\|^2}{2\|p\|^2\zeta_t^2(\zeta_t - r)}r^2 \\ &\quad= \|p\|\frac{\zeta_t^2}{2\zeta_t^2(\zeta_t - r)}r^2 = \|p\|\frac{1}{2(\zeta_t - r)}r^2.\end{aligned}$$

Write $z = x + \mathbf{i}y$ with x, $y \in \mathbb{R}$. This implies $z - t = (x-t) + \mathbf{i}y$ and $|z-t| \geq |y|$, so we can apply the Cauchy-Schwarz inequality to get

$$\begin{aligned}\Im\left(\left\langle d - zp, \frac{d-tp}{\|d-tp\|} - c\right\rangle\right) = \left\langle yp, \frac{d-tp}{\|d-tp\|} - c\right\rangle &\leq |y|\,\|p\|\left\|\frac{d-tp}{\|d-tp\|} - c\right\| \\ &\leq \|p\|\left\|\frac{d-tp}{\|d-tp\|} - c\right\||z-t| \leq \|p\|\left\|\frac{d-tp}{\|d-tp\|} - c\right\| r.\end{aligned}$$

Combining both estimates with (39) gives us (40). □

Theorem 20 (Derivatives) *Let d, p, $c \in \mathbb{R}^3$ with $\|c\| = 1$. Define*

$$\alpha := \left\|\frac{d}{\|d\|} - c\right\|, \qquad \beta := \frac{\|p\|}{\|d\|}. \tag{42}$$

If $\beta \leq 1/2$, then

$$|g_1^{(m)}(0)| \leq m!\,\frac{\exp(\kappa\|p\|(\alpha+\beta))}{(1-\beta)\|d\|} \qquad \text{for all } m \in \mathbb{N}.$$

Proof. If $p = 0$, then g_1 is a constant, and the estimate is trivial.

Let $p \neq 0$ and $\zeta_0 := \|d\|/\|p\| = 1/\beta$. Assume $\beta \leq 1/2$. Then, we have $\zeta_0 > 1$.

Let $z \in \mathbb{C}$ with $|z| \leq 1$. Applying Lemma 17 with $t = 0$ and $r = 1 < \zeta_0$ yields

$$|f(z)| \geq \|p\|(\zeta_0 - 1),$$

and with Lemma 19 this leads to

$$|\Im(f_e(z))| \leq \|p\| \left(\alpha + \frac{1}{2(\zeta_0 - 1)}\right) = \|p\| \left(\alpha + \frac{\beta}{2(1-\beta)}\right) \leq \|p\|(\alpha + \beta).$$

By the Cauchy formula for derivatives, we find

$$\begin{aligned}
|g_1^{(m)}(0)| &\leq m! \max\{|g_1(z)| \ : \ z \in \mathbb{C}, \ |z| = 1\} \\
&\leq m! \max\left\{\frac{\exp(\kappa|\Im(f_e(z))|)}{|f(z)|} \ : \ z \in \mathbb{C}, \ |z| = 1\right\} \\
&\leq m! \frac{\exp(\kappa\|p\|(\alpha+\beta))}{\|p\|(\zeta_0 - 1)} = m! \frac{\beta}{1-\beta} \frac{\exp(\kappa\|p\|(\alpha+\beta))}{\|p\|} \\
&= m! \frac{\exp(\kappa\|p\|(\alpha+\beta))}{(1-\beta)\|d\|}.
\end{aligned}$$

□

Corollary 21 (Directional derivatives) *Let x, y, p_x, p_y, $c \in \mathbb{R}^3$ with $x \neq y$ and $\|c\| = 1$. Let*

$$\alpha := \left\| \frac{x-y}{\|x-y\|} - c \right\|, \qquad \beta := \frac{\|p_x - p_y\|}{\|x-y\|}.$$

If $\beta \leq 1/2$, then we have

$$|\partial^m_{(p_x,p_y)} g_c(x,y)| \leq m! \frac{\exp(\kappa\|p_x - p_y\|(\alpha+\beta))}{(1-\beta)\|x-y\|} \qquad \text{for all } m \in \mathbb{N}.$$

Proof. We let $p := p_y - p_x$ and observe

$$g_1(t) = g_c(x - tp, y) = g_c(x - tp_y + tp_x, y) = g_c(x + tp_x, y + tp_y) \quad \text{for all } t \in [-1,1].$$

This implies

$$g_1^{(m)}(0) = \partial^m_{(p_x,p_y)} g_c(x,y) \qquad \text{for all } m \in \mathbb{N}.$$

An application of Theorem 20 completes the proof. □

8 Polynomial best approximation

Corollary 21 is already sufficient to obtain convergence results for the Taylor expansion of g_c, but the requirement $\|p\| < \|d\|$ forces us to use rather strong admissibility conditions.

We can obtain better convergence rate estimates by considering an approximation by interpolation. If a stable interpolation scheme, e.g., Chebyshev interpolation, is used, we only have to prove that the function can be approximated by a polynomial, since standard best-approximation arguments immediately result in an almost optimal bound for the interpolation error.

In order to prove that g_1 can be approximated, we rely on the theory presented in [9, Chapter 7].

Definition 22 (Bernstein ellipse) *We define*

$$\mathcal{E}_\varrho := \left\{ z = x + iy \ : \ x, y \in \mathbb{R}, \left(\frac{2x}{\varrho + 1/\varrho}\right)^2 + \left(\frac{2y}{\varrho - 1/\varrho}\right)^2 = 1 \right\},$$

$$\mathcal{D}_\varrho := \left\{ z = x + iy \ : \ x, y \in \mathbb{R}, \left(\frac{2x}{\varrho + 1/\varrho}\right)^2 + \left(\frac{2y}{\varrho - 1/\varrho}\right)^2 < 1 \right\} \quad \text{for all } \varrho \in \mathbb{R}_{>1}.$$

The set $\mathcal{E}_\varrho$ is called a Bernstein ellipse*, the set $\mathcal{D}_\varrho$ is the interior of $\mathcal{E}_\varrho$.*

Lemma 23 (Existence) *Let $\hat{\varrho} \in \mathbb{R}_{>1}$, and let $f : \mathcal{D}_{\hat{\varrho}} \to \mathbb{C}$ be holomorphic. Given $\varrho \in (1, \hat{\varrho})$ and $m \in \mathbb{N}$, there is a polynomial $\pi \in \Pi_m$ of degree m such that*

$$\|f - \pi\|_{\infty,[-1,1]} \leq \frac{2}{\varrho - 1} \varrho^{-m} \max\{|f(z)| \ : \ z \in \overline{\mathcal{D}}_\varrho\}.$$

Proof. This is [9, eqn. (8.7), Chap. 7]. □

Our Lemmas 17 and 19 can be used to obtain bounds for the holomorphic extension of g_1 in the domains U_r. In order to apply Lemma 23, we simply have to find $\varrho > 1$ such that $\overline{\mathcal{D}}_\varrho \subseteq U_r$.

Lemma 24 (Inclusion) *Let $r \in \mathbb{R}_{>0}$, and let $\varrho := \sqrt{r^2 + 1} + r$.*

For each $z \in \overline{\mathcal{D}}_\varrho$, there is a $t \in [-1, 1]$ such that $z \in U_{t,r}$. In particular, we have

$$\overline{\mathcal{D}}_\varrho \subseteq U_r.$$

Proof. We start by observing

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

$$\frac{\varrho + 1/\varrho}{2} = \frac{2\sqrt{r^2+1}}{2} = \sqrt{r^2+1}, \qquad \frac{\varrho - 1/\varrho}{2} = \frac{2r}{2} = r.$$

Let $z \in \overline{\mathcal{D}}_\varrho$. We fix $x, y \in \mathbb{R}$ such that $z = x + iy$. Definition 22 implies

$$1 \geq \left(\frac{2x}{\varrho + 1/\varrho}\right)^2 + \left(\frac{2y}{\varrho - 1/\varrho}\right)^2 = \frac{x^2}{r^2+1} + \frac{y^2}{r^2}. \tag{43}$$

If $x \in [-1, 1]$, we let $t = x$ and find $|z - t| = |y|$. Since (43) implies $|y| \leq r$, we have $z \in U_{t,r}$.

If $x > 1$, we let $t = 1$. The inequality (43) yields

$$x^2 = (r^2+1)\left(1 - \frac{y^2}{r^2}\right) = r^2 + 1 - \frac{r^2+1}{r^2} y^2 \leq r^2 + 1 - y^2, \tag{44}$$

$$(x-1)^2 = x^2 - 2x + 1 \leq r^2 + 2 - y^2 - 2x < r^2 - y^2,$$

and we have $|z - 1|^2 = (x-1)^2 + y^2 < r^2$, i.e., $z \in U_{t,r}$.

If $x < -1$, we let $t = -1$ and use (44) again to find

$$(x+1)^2 = x^2 + 2x + 1 \leq r^2 + 2 - y^2 + 2x < r^2 - y^2$$

Since this implies $|z + 1|^2 = (x+1)^2 + y^2 < r^2$, we have $z \in U_{t,r}$. □

Theorem 25 (Approximation) *Let d, p, $c \in \mathbb{R}^3$ with $\|c\| = 1$. Define*

$$\alpha := \max\left\{\left\|\frac{d-tp}{\|d-tp\|} - c\right\| \ : \ t \in [-1,1]\right\}, \qquad \delta := \inf\{\|d-tp\| \ : \ t \in [-1,1]\}$$

and assume $\delta > 0$. Let $\zeta := \delta/\|p\|$, $r \in (0,\zeta)$, and $\varrho = \sqrt{r^2+1} + r$. Then for each $m \in \mathbb{N}$ there exists a polynomial $\pi \in \Pi_m$ of degree m such that for the function g_1 of (31)

$$\|g_1 - \pi\| \le \frac{2}{r}\frac{\exp(\gamma(r))}{4\pi\delta(1-r/\zeta)}\varrho^{-m},$$

where

$$\gamma(r) := \kappa\|p\| r\left(\alpha + \frac{r}{2(\zeta - r)}\right).$$

Proof. We observe

$$\zeta_t := \frac{\|d-tp\|}{\|p\|} \ge \zeta \qquad \text{for all } t \in [-1,1].$$

Let $z \in U_r$, and let $t \in [-1,1]$ with $|z-t| \le r$. Lemma 19 yields

$$\begin{aligned}\kappa|\Im(f_e(z))| &\le \kappa\|p\|\left(\left\|\frac{d-tp}{\|d-tp\|} - c\right\| r + \frac{1}{2(\zeta_t - r)}r^2\right)\\ &\le \kappa\|p\|\left(\alpha r + \frac{1}{2(\zeta-r)}r^2\right) = \gamma(r).\end{aligned}$$

Lemma 17 yields

$$|f(z)| \ge \|p\|(\zeta - r) = \|p\|\zeta(1-r/\zeta) = \delta(1-r/\zeta),$$

and we conclude

$$|g_1(z)| \le \frac{\exp(\kappa|\Im(f_e(z))|)}{4\pi|f(z)|} \le \frac{\exp(\gamma(r))}{4\pi\delta(1-r/\zeta)}.$$

We have established this bound for all $z \in U_r$. Due to Lemma 24, this implies that the bound holds for all $z \in \overline{\mathcal{D}}_\varrho$, so we can apply Lemma 23 and use $\varrho - 1 = \sqrt{r^2+1} + r - 1 \ge r$ to complete the proof. □

Lemma 26 (Approximate directions) *Let $t, s \in \mathcal{T}_\mathcal{I}$ be given with* $\mathrm{level}(t) = \mathrm{level}(s)$ *and let $c \in \mathcal{D}_t = \mathcal{D}_s$ be chosen such that the admissibility conditions (17) hold.*

Let d, $p \in \mathbb{R}^3$ be vectors satisfying (30a) and (30b). Then we have

$$\left\|\frac{d-\gamma p}{\|d-\gamma p\|} - c\right\| \le \frac{\eta_1 + \eta_2}{\kappa \max\{\mathrm{diam}(B_t), \mathrm{diam}(B_s)\}} \qquad \textit{for all } \gamma \in [-1,1].$$

Proof. Let $q := \max\{\operatorname{diam}(B_t), \operatorname{diam}(B_s)\}$.

Let $\gamma \in [-1,1]$. Due to (30b) and (17b), we have

$$\|d - \gamma p\| \geq \operatorname{dist}(B_t, B_s) \geq \frac{\kappa}{\eta_2} q^2.$$

Due to $m_t \in B_t$ and $m_s \in B_s$, we can apply (17b) to find

$$\|m_t - m_s\| \geq \operatorname{dist}(B_t, B_s) \geq \frac{\kappa}{\eta_2} q^2.$$

The last two estimates imply

$$\left\| (d - \gamma p) \frac{\eta_2}{\kappa q^2} \right\| \geq 1, \qquad \left\| (m_t - m_s) \frac{\eta_2}{\kappa q^2} \right\| \geq 1,$$

so we can apply Lemma 10 to obtain

$$\left\| \frac{d - \gamma p}{\|d - \gamma p\|} - \frac{m_t - m_s}{\|m_t - m_s\|} \right\| \leq \|(d - \gamma p) - (m_t - m_s)\| \frac{\eta_2}{\kappa q^2}.$$

Due to (30b), we can find $x \in B_t$ and $y \in B_s$ such that $d - \gamma p = x - y$ and obtain

$$\begin{aligned} \|(d - \gamma p) - (m_t - m_s)\| &= \|(x - m_t) - (y - m_s)\| \leq \|x - m_t\| + \|y - m_s\| \\ &\leq q/2 + q/2 = q. \end{aligned}$$

Combining the estimates yields

$$\begin{aligned} \left\| \frac{d - \gamma p}{\|d - \gamma p\|} - c \right\| &\leq \left\| \frac{d - \gamma p}{\|d - \gamma p\|} - \frac{m_t - m_s}{\|m_t - m_s\|} \right\| + \left\| \frac{m_t - m_s}{\|m_t - m_s\|} - c \right\| \\ &\leq \|(d - \gamma p) - (m_t - m_s)\| \frac{\eta_2}{\kappa q^2} + \frac{\eta_1}{\kappa q} \leq \frac{\eta_2 + \eta_1}{\kappa q}. \end{aligned}$$

This is the required estimate. □

Theorem 27 (Interpolation error) *Let $t, s \in \mathcal{T}_\mathcal{I}$ be clusters with* $\operatorname{level}(t) = \operatorname{level}(s)$ *and let $c \in \mathcal{D}_t = \mathcal{D}_s$ be chosen such that the admissibility conditions (17) hold.*

Let $r \in (0, 1/\eta_2)$ and $\varrho = \sqrt{r^2 + 1} + r$. Then, we have for the approximation $\tilde{g}_{ts} := \exp(\mathbf{i}\kappa\langle x - y, c\rangle)\mathfrak{I}_{B_t \times B_s}[\exp(-\mathbf{i}\kappa\langle x - y, c\rangle)g]$ the bound

$$\|g - \tilde{g}_{ts}\|_{\infty, B_t \times B_s} \leq C_{\mathrm{in}}(\eta_1, \eta_2, r) \frac{\Lambda_m^5 (1 + \Lambda_m)}{\operatorname{dist}(B_t, B_s)} \varrho^{-m} \qquad \text{for all } m \in \mathbb{N}_0,$$

where

$$C_{\mathrm{in}}(\eta_1, \eta_2, r) = \frac{12}{r} \frac{\exp\left(r(\eta_1 + \eta_2) + r^2 \frac{\eta_2}{2(1 - r\eta_2)}\right)}{4\pi(1 - r\eta_2)}.$$

Proof. Lemma 26 yields

$$\alpha := \max\left\{\left\|\frac{t-\gamma p}{\|t-\gamma p\|}\right\| \ : \ \gamma\in[-1,1]\right\} \le \frac{\eta_1+\eta_2}{\kappa\max\{\operatorname{diam}(B_t),\operatorname{diam}(B_s)\}},$$

and the condition (30b) implies

$$\delta := \inf\{\|d-\gamma p\| \ : \ \gamma\in[-1,1]\} \ge \operatorname{dist}(B_t,B_s).$$

With (30a) and the standard admissibility condition (17c) we get

$$\zeta := \delta/\|p\| \ge \frac{\operatorname{dist}(B_t,B_s)}{\max\{\operatorname{diam}(B_t),\operatorname{diam}(B_s)\}} \ge 1/\eta_2 > 0, \qquad \text{i.e., } \ 1/\zeta \le \eta_2.$$

Since $r<1/\eta_2$ implies $r<\zeta$, we can apply Theorem 25 to find a $\pi\in\Pi_m$ with

$$\|g_1-\pi\|_{\infty,[-1,1]} \le \frac{2}{r}\frac{\exp(\gamma(r))}{4\pi\delta(1-r/\zeta)}\varrho^{-m}. \tag{45}$$

Using (30a) in combination with the directional admissibility condition (17a) and the parabolic admissibility condition (17b), we find

$$\begin{aligned}
\gamma(r) &= \kappa\|p\| r\left(\alpha+\frac{r}{2(\zeta-r)}\right)\\
&\le \kappa\|p\| r\frac{\eta_1+\eta_2}{\kappa\max\{\operatorname{diam}(B_t),\operatorname{diam}(B_s)\}} + \kappa\frac{\|p\|^2r^2}{2\delta(1-r/\zeta)}\\
&\le r(\eta_1+\eta_2) + r^2\frac{\kappa\max\{\operatorname{diam}(B_t)^2,\operatorname{diam}(B_s)^2\}}{2\operatorname{dist}(B_t,B_s)(1-r/\zeta)}\\
&\le r(\eta_1+\eta_2) + r^2\frac{\eta_2}{2(1-r/\zeta)} \le r(\eta_1+\eta_2) + r^2\frac{\eta_2}{2(1-r\eta_2)}.
\end{aligned}$$

Combining this estimate with (45) yields

$$\|g_1-\pi\|_{\infty,[-1,1]} \le \frac{C_{\mathrm{in}}(\eta_1,\eta_2,r)}{6\operatorname{dist}(B_t,B_s)}\varrho^{-m}.$$

With (22) we get the estimate

$$\begin{aligned}
\|g_1-\mathfrak{I}[g_1]\|_{\infty,[-1,1]} &= \|(g_1-\pi)-\mathfrak{I}[g_1-\pi]\|_{\infty,[-1,1]} \le (1+\Lambda_m)\|g_1-\pi\|_{\infty,[-1,1]}\\
&\le (1+\Lambda_m)\frac{C_{\mathrm{in}}(\eta_1,\eta_2,r)}{6\operatorname{dist}(B_t,B_s)}\varrho^{-m},
\end{aligned}$$

and Lemma 15 finally gives us

$$\begin{aligned}
\|g_c-\mathfrak{I}_{B_t\times B_s}[g_c]\|_{\infty,B_t\times B_s} &\le 6\Lambda_m^5(1+\Lambda_m)\frac{C_{\mathrm{in}}(\eta_1,\eta_2,r)}{6\operatorname{dist}(B_t,B_s)}\varrho^{-m}\\
&= C_{\mathrm{in}}(\eta_1,\eta_2,r)\frac{\Lambda_m^5(1+\Lambda_m)}{\operatorname{dist}(B_t,B_s)}\varrho^{-m}.
\end{aligned}$$

□

9 Nested approximation

As discussed in Section 3, an $\mathcal{H}^2$-structure is essential to reduce the complexity of our matrix representation. The crucial step that permitted this structure is the approximation step (13). In this section, we analyze the impact of this step. The analysis is non-standard since the functions $L_{tc,\nu}$ are recursively approximated. Structurally similar analyses can be found in [7, 22].

Since we will frequently have to interpolate functions of the form

$$z \mapsto \exp(\mathbf{i}\kappa c z) u(z)$$

for functions u, we introduce the short notation

$$\exp(\mathbf{i}\kappa c \cdot) u$$

for these functions, where "$\cdot$" marks the position of the variable z and the multiplication of the functions $\exp(\mathbf{i}\kappa c\cdot)$ and u is defined pointwise.

In order to perform the analysis, we introduce the notion of descendants of a cluster $t \in \mathcal{T}_\mathcal{I}$:

Definition 28 (descendant) *Let $\mathcal{T}_\mathcal{I}$ be a cluster tree. For $t \in \mathcal{T}_\mathcal{I}$ define*

$$\begin{aligned}\operatorname{sons}^*(t) := \{\tau \in \mathcal{T}_\mathcal{I} \,|\, &\tau \text{ is leaf of } \mathcal{T}_\mathcal{I} \text{ and } \exists\, (\tau_i)_{i=0}^L \text{ s.t.} \\ &t = \tau_0,\ \tau = \tau_L,\ \tau_{i+1} \in \operatorname{sons}(\tau_i),\ i = 0, \ldots, L-1\}.\end{aligned}$$

We call $(\tau_i)_{i=0}^L$ the cluster chain *connecting t with τ. For given $t \in \mathcal{T}_\mathcal{I}$, $\tau \in \operatorname{sons}^*(t)$, the cluster chain $(\tau_i)_{i=0}^L$ is unique. We call L the* length *of the cluster chain connecting t with $\tau \in \operatorname{sons}^*(t)$.*

Let $\mathcal{T}_\mathcal{I}$ be a cluster tree with hierarchical directions $(c_t)_{t\in\mathcal{T}_\mathcal{I}}$. For given $t \in \mathcal{T}_\mathcal{I}$, $\tau \in \operatorname{sons}^*(t)$ with cluster chain $(\tau_i)_{i=0}^L$ the effective *length of the cluster chain is defined by*

$$L_0 := \begin{cases} L & \text{if } c_{\tau_i} \neq 0 \text{ for all } i \in \{0, \ldots, L\} \\ \min\{i \,|\, c_{\tau_i} = 0\} & \text{otherwise.} \end{cases}$$

Remark 29 *The cluster chain length L is bounded by the depth of the tree $\mathcal{T}_\mathcal{I}$. The effective cluster chain length L_0 is bounded by $C \log \kappa$ if the following two conditions are satisfied:*

- *the directions $c \in \mathcal{D}_t$ are selected such that $c = 0$ for $\kappa \operatorname{diam}(B_t) \leq \eta_1$, and*
- $\operatorname{diam}(B_{t'})/\operatorname{diam}(B_t) \leq q < 1$ *for all $t \in \mathcal{T}_\mathcal{I}$, $t' \in \operatorname{sons}(t)$.*

In this situation, the condition in (47) below is satisfied for $m \geq K \log(\log \kappa)$.

The step (13) shows that approximation operators $\mathfrak{I}_{t'}^c$ of the following form appear for $t \in \mathcal{T}_\mathcal{I}$, $t' \in \operatorname{sons}(t)$, $c \in \mathcal{D}_t$, $c' = \operatorname{sd}_t(c)$:

$$u \mapsto \mathfrak{I}_{t'}^c[u] := \exp(\mathbf{i}\kappa\langle \cdot, c'\rangle)\mathfrak{I}_{B_{t'}}[\exp(-\mathbf{i}\kappa\langle \cdot, c'\rangle)u] \qquad \text{for all } u \in C(B_{t'}).$$

Note that these operator reduce to the usual polynomial interpolation if $c' = 0$. With the aid of the operators $\mathfrak{I}^c_{t'}$, we define the following iterated operators: let $t \in \mathcal{T}_{\mathcal{I}}$ and $\tau \in \text{sons}^*(t)$. Denote $(\tau_i)_{i=0}^L$ the corresponding cluster chain. Define with the effective length L_0 of the cluster chain

$$\mathfrak{I}^c_{t,\tau} := \mathfrak{I}^c_{\tau_{L_0}} \circ \mathfrak{I}^c_{\tau_{L_0-1}} \circ \cdots \circ \mathfrak{I}^c_{\tau_1}$$

(If $L_0 = 0$, then $\mathfrak{I}^c_{t,\tau} = \text{Id}$.) These operators arise in our algorithm. Indeed, for admissible blocks $b = (t, s)$, the interpolant (9) is given by

$$\begin{aligned}\tilde{g}_{ts}(x,y) &= \mathfrak{I}^c_t \otimes \mathfrak{I}^c_s[g](x,y) \\ &= \exp(\mathbf{i}\kappa\langle x-y, c\rangle) \sum_{\nu,\mu\in M} g_c(\xi_{t,\nu}, \xi_{s,\mu}) L_{t,\nu}(x)\overline{L_{s,\mu}(y)} \quad \text{for all } x \in B_t,\ y \in B_s,\end{aligned}$$

and is approximated by our algorithm by

$$\begin{aligned}\tilde{g}_{ts,\tau\sigma}(x,y) &:= \mathfrak{I}^c_{t,\tau} \otimes \mathfrak{I}^c_{s,\sigma}[g](x,y) \\ &= \sum_{\nu,\mu\in M} g_c(\xi_{t,\nu}, \xi_{s,\mu}) \widetilde{L}_{t\tau c,\nu}(x)\overline{\widetilde{L}_{s\sigma c,\mu}(y)} \qquad \text{for all } x \in B_\tau,\ y \in B_\sigma\end{aligned}$$

for $\tau \in \text{sons}^*(t)$, $\sigma \in \text{sons}^*(s)$, where the functions $\widetilde{L}_{t\tau c,\nu}$ are defined by

$$\widetilde{L}_{t\tau c,\nu}(x) = \mathfrak{I}^c_{t,\tau}[L_{tc,\nu}](x), \qquad \text{for all } x \in B_\tau,\ \tau \in \text{sons}^*(t).$$

The analysis is therefore reduced to gauging the difference $L_{tc,\nu} - \widetilde{L}_{t\tau c,\nu}$. This is achieved in the following theorem.

Theorem 30 *Assume the 1D interpolation operator $\mathfrak{I}$ satisfies (23). Assume the existence of $\overline{q} \in (0,1)$ such that for any $t \in \mathcal{T}_{\mathcal{I}}$, $t' \in \text{sons}(t)$, the corresponding bounding boxes $B_t = [a_{t,1}, b_{t,1}] \times [a_{t,2}, b_{t,2}] \times [a_{t,3}, b_{t,3}]$ and $B_{t'} = [a'_{t,1}, b'_{t,1}] \times [a'_{t,2}, b'_{t,2}] \times [a'_{t,3}, b'_{t,3}]$ satisfy the contraction property*

$$\frac{b'_{t,i} - a'_{t,i}}{b_{t,i} - a_{t,i}} \le \overline{q} \qquad i = 1,2,3. \tag{46}$$

Assume (18). Choose $q \in (\overline{q}, 1)$. Then there exists K depending only on $\overline{q}$, the chosen q, the constants C_Λ, λ of (23), and η_1 such that for any $t \in \mathcal{T}_{\mathcal{I}}$ and $\tau \in \text{sons}^*(t)$ with effective cluster chain length L_0 the following holds:

$$m \ge K \log(L_0 + 2) \qquad \Longrightarrow \qquad \|L_{tc,\nu} - \widetilde{L}_{t\tau c,\nu}\|_{\infty,B_\tau} \le q^m. \tag{47}$$

Proof. The proof is a consequence of Theorem 36 (with $n = 3$) and the following observations: first, the parameter L in the statement of Theorem 36 is replaced by the effective cluster chain length L_0, and second

$$\|L_{tc,\nu}\|_{\infty,B_t} = \|\exp(\mathbf{i}\kappa\langle c, \cdot\rangle) L_{t,\nu}\|_{\infty,B_t} = \|L_{t,\nu}\|_{\infty,B_t} \le \Lambda_m^3,$$

where, in the last step, we used that $L_{t,\nu}$ is a Lagrange interpolation polynomial for the box B_t. □

9.1 Recursive reinterpolation in 1D

The following definition captures the essence of the reinterpolation process analyzed in Theorem 30 in a 1D setting:

Definition 31 *(i) We call the sequence of closed intervals* $\mathfrak{C} := (J_i)_{i=0}^L$ *an* interval chain *if* $J_0 \supset J_1 \supset \cdots \supset J_L$*. If*

$$\frac{|J_{i+1}|}{|J_i|} \le \overline{q} < 1, \qquad i = 0, \dots, L-1,$$

for a $\overline{q} \in \mathbb{R}$*, we call* $\overline{q}$ *a* contraction factor *of the interval chain.*

(ii) Let $\mathfrak{I} : C([-1,1]) \to \Pi_m$ *be a polynomial interpolation operator with Lebesgue constant* Λ_m*. For a sequence* $(c_i)_{i=0}^L$ *in* $\mathbb{R}$*, a parameter* $\kappa \in \mathbb{R}$*, and an interval chain* $\mathfrak{C} = (J_i)_{i=0}^L$ *we define the operators* $\mathfrak{I}_i^c$ *by*

$$\mathfrak{I}_i^c u := \exp(\mathbf{i}\kappa c_i \cdot)\mathfrak{I}_{J_i}[\exp(-\mathbf{i}\kappa c_i \cdot)u] \qquad \text{for all } i \in \{0, \dots, L\}, \tag{48}$$

where $\mathfrak{I}_{J_i}$ *is the interpolation operator* $\mathfrak{I}$ *scaled to the interval* J_i*.*

(iii) The iterated *interpolation operator is defined by*

$$\mathfrak{I}_{\mathfrak{C}}^c u := \begin{cases} \mathfrak{I}_L^c \circ \mathfrak{I}_{L-1}^c \circ \cdots \circ \mathfrak{I}_1^c, & L \ge 1 \\ \mathrm{Id} & L = 0. \end{cases} \tag{49}$$

We say that $\mathfrak{I}_{\mathfrak{C}}^c$ *is based on the interval chain* $\mathfrak{C}$*, the coefficients* $(c_i)_{i=0}^L$*, and the interpolation operator* $\mathfrak{I}$ *with underlying Lebesgue constant* Λ_m*.*

We need the following result concerning the inclusion of a scaled Bernstein ellipse in another one:

Lemma 32 *Let* $-1 \le a < b \le 1$ *and* $h := (b-a)/2$*. For* $\alpha > 1$ *denote by* $\mathcal{D}_\alpha^{a,b}$ *the interior of the "scaled" Bernstein ellipse for the interval* $[a, b]$*:*

$$\mathcal{D}_\alpha^{a,b} := \frac{a+b}{2} + h\mathcal{D}_\alpha,$$

where $\mathcal{D}_\varrho$ *is given in Definition 22. Fix* $\varepsilon \in (0,1)$*. Then there is a* $\varrho_0 > 1$ *(depending solely on* ε*) such that*

$$\mathcal{D}_{(1-\varepsilon)\varrho/h}^{a,b} \subset \mathcal{D}_\varrho \qquad \forall \varrho \ge \varrho_0. \tag{50}$$

Proof. We exploit that for large ϱ the Bernstein ellipse $\mathcal{E}_\varrho$ is essentially a circle of radius $\varrho/2$. We start from the following inclusion of balls in Bernstein ellipses and *vice versa*:

$$B_{(\varrho - 1/\varrho)/2}(0) \subset \mathcal{D}_\varrho \subset B_{(\varrho + 1/\varrho)/2}(0),$$

where $B_r(x) = \{|z-x| < r \ : \ z \in \mathbb{C}\}$ denotes the open disc around x of radius r. Hence, we have to show (for ϱ sufficiently large) that for $\alpha = (1-\varepsilon)\varrho/h$ we have

$$\begin{aligned}\mathcal{D}_\alpha^{a,b} &= \frac{a+b}{2} + h\mathcal{D}_\alpha \subset \frac{a+b}{2} + hB_{(\alpha+1/\alpha)/2}(0) \\ &= B_{h(\alpha+1/\alpha)/2}((a+b)/2) \overset{!}{\subset} B_{(\varrho-1/\varrho)/2}(0);\end{aligned}$$

all inclusions are geometrically clear with the exception of the last one. To ensure that one, we require

$$1 + h\frac{\alpha+1/\alpha}{2} \leq \frac{\varrho-1/\varrho}{2}. \tag{51}$$

Inserting the condition $\alpha = (1-\varepsilon)\varrho/h$ and rearranging terms, we see that (51) is true if we ensure

$$\varepsilon\varrho^2 \geq 2\varrho + 1 + \frac{h^2}{1-\varepsilon}. \tag{52}$$

In view of $h \in [0,2]$, this last condition is certainly met if

$$\varrho \geq \varrho_0 := \frac{1+\sqrt{1+\varepsilon(1+4/(1-\varepsilon))}}{\varepsilon},$$

which finishes the argument. □

Lemma 33 *Fix $0 < \overline{q} < 1$ and $\gamma > 0$. Choose $q \in (\overline{q}, 1)$. Then there is $m_0 > 0$ depending only on $\overline{q}$, γ, and the chosen q such that the following is true:*

Let $J_1 \subset J_0$ be two closed intervals with $|J_1|/|J_0| \leq \overline{q} < 1$. Denote $h_0 := |J_0|/2$, $h_1 := |J_1|/2$. Let κ, c_0, $c_1 \in \mathbb{R}$ and assume that

$$|\kappa h_0(c_0 - c_1)| \leq \gamma.$$

Then for all $m \geq m_0$ and all $\pi \in \Pi_m$

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

Proof. Let $\Phi : [-1,1] \to J_0$ be the orientation preserving affine bijection as in Section 2.1. Set $\widehat{\pi} := \pi \circ \Phi$, $[a,b] := \widehat{J_1} := \Phi^{-1}(J_1)$. Set $\widehat{h} := h_1/h_0 = (b-a)/2 \leq \overline{q}$. We have

$$\inf_{v\in\Pi_m} \|\exp(\mathbf{i}\kappa c_0\cdot)\pi - \exp(\mathbf{i}\kappa c_1\cdot)v\|_{\infty,J_1} = \inf_{v\in\Pi_m} \|\exp(\mathbf{i}\kappa h_0(c_0-c_1)\cdot)\widehat{\pi} - v\|_{\infty,\widehat{J_1}}.$$

By the polynomial approximation results of Lemma 23, we estimate for arbitrary $\alpha > 1$ and $m \in \mathbb{N}_0$

$$\begin{aligned}\inf_{v\in\Pi_m} \|\exp(\mathbf{i}\kappa h_0(c_0-c_1)\cdot)\widehat{\pi} - v\|_{\infty,\widehat{J_1}} &\leq \frac{2\alpha^{-m}}{\alpha-1} \|\exp(\mathbf{i}\kappa h_0(c_0-c_1)\cdot)\widehat{\pi}\|_{\infty,\mathcal{D}_\alpha^{a,b}} \\ &\leq \frac{2\alpha^{-m}}{\alpha-1} \exp\left(|\kappa h_0(c_0-c_1)|\widehat{h}\frac{\alpha-1/\alpha}{2}\right) \|\widehat{\pi}\|_{\infty,\mathcal{D}_\alpha^{a,b}}.\end{aligned}$$

We now choose α in dependence on m. Fix $\varepsilon \in (0, 1-\overline{q})$ (so that $\widehat{h}/(1-\varepsilon) < 1$) and choose $\beta > 0$ such that (for the q of the statement of the lemma)

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

We set $\varrho = \beta m$ and $\alpha = (1-\varepsilon)\varrho/\widehat{h} = (1-\varepsilon)\beta m/\widehat{h}$. Lemma 32 implies $\mathcal{D}_\alpha^{a,b} \subset \mathcal{D}_\varrho$ if $\beta m = \varrho \geq \varrho_0$. We note that this condition imposes the condition $m \geq m_0' := \varrho_0/\beta$. Furthermore, the Bernstein estimate [9, Thm. 2.2, Chap. 4], gives

$$\|\widehat{\pi}\|_{\infty,\mathcal{D}_\alpha^{a,b}} \leq \|\widehat{\pi}\|_{\infty,\mathcal{D}_\varrho} \leq \varrho^m \|\widehat{\pi}\|_{\infty,(-1,1)}.$$

Hence we obtain

$$\begin{aligned}
\inf_{v\in\Pi_m} \|\exp(\mathbf{i}\kappa h_0(c_0-c_1)\cdot)\widehat{\pi} - v\|_{\infty,\widehat{J}_1} &\leq \frac{2}{\alpha-1}\left(\frac{\varrho}{\alpha}\right)^m \exp\left(\gamma\widehat{h}\frac{\alpha-1/\alpha}{2}\right)\|\widehat{\pi}\|_{\infty,(-1,1)} \\
&\leq \frac{2}{\alpha-1}\left(\frac{\widehat{h}}{1-\varepsilon}\right)^m \exp\left(\frac{\gamma\widehat{h}\alpha}{2}\right)\|\widehat{\pi}\|_{\infty,(-1,1)} \\
&= \frac{2}{\alpha-1}\left(\frac{\widehat{h}}{1-\varepsilon}\right)^m \exp\left(m\frac{\gamma(1-\varepsilon)\beta}{2}\right)\|\widehat{\pi}\|_{\infty,(-1,1)} \\
&= \frac{2}{\alpha-1}\left(\frac{\widehat{h}}{1-\varepsilon}\exp\left(\frac{\gamma(1-\varepsilon)\beta}{2}\right)\right)^m \|\widehat{\pi}\|_{\infty,(-1,1)} = \frac{2}{\alpha-1}q^m\|\widehat{\pi}\|_{\infty,(-1,1)} \\
&\leq q^m\|\widehat{\pi}\|_{\infty,(-1,1)},
\end{aligned}$$

where, in the last step we used that $\alpha \to \infty$ as $m \to \infty$; more precisely, we assumed

$$m \geq m_0 := \max\left\{m_0', \left\lceil 3\frac{\widehat{h}}{(1-\varepsilon)\beta}\right\rceil)\right\},$$

thus ensuring $\alpha \geq 3$. □

Lemma 34 *Let $\mathfrak{I} : C([-1,1]) \to \Pi_m$ be an interpolation operator with Lebesgue constant Λ_m. Then, for $\mathfrak{I}_i^c$ as defined in (48), there holds for arbitrary $\pi \in \Pi_m$*

$$\begin{aligned}
&\|\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi - \mathfrak{I}_i^c[\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi]\|_{\infty,J_i} \\
&\qquad\qquad \leq (1+\Lambda_m)\inf_{v\in\Pi_m}\|\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi - \exp(\mathbf{i}\kappa c_i\cdot)v\|_{\infty,J_i}.
\end{aligned} \tag{53}$$

Proof. Let π, $v \in \Pi_m$ be arbitrary. Write

$$\begin{aligned}
&\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi - \mathfrak{I}_i^c[\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi] \\
&\qquad = \exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi - \exp(\mathbf{i}\kappa c_i\cdot)v \\
&\qquad\quad - \exp(\mathbf{i}\kappa c_i\cdot)\mathfrak{I}_{J_i}[\exp(-\mathbf{i}\kappa c_i\cdot)\{\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi - \exp(\mathbf{i}\kappa c_i\cdot)v\}].
\end{aligned}$$

Hence, by the stability of $\mathfrak{I}$ we get

$$\begin{aligned}\|\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi - \mathfrak{I}_i^c[\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi]\|_{\infty,J_i} \\ \leq (1+\Lambda_m)\|\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi - \exp(\mathbf{i}\kappa c_i\cdot)v\|_{\infty,J_i},\end{aligned}$$

which concludes the proof. □

Lemma 35 (Stability of reinterpolation) *Let $\kappa \in \mathbb{R}$. Let $\mathfrak{C} := (J_i)_{i=0}^L$ be an* interval chain *with contraction factor $\overline{q} < 1$. Write $h_i = |J_i|/2$. Let $c_0, \dots, c_L \in \mathbb{R}$ be such that*

$$|\kappa h_i(c_i - c_{i+1})| \leq \gamma, \qquad i = 0, \dots, L-1, \tag{54}$$

for some $\gamma > 0$. Let $\mathfrak{I}_{\mathfrak{C}}^c$ be an iterated interpolation operator based on $\mathfrak{C}$, $(c_i)_{i=0}^L$ and a polynomial interpolation operator $\mathfrak{I}$ with Lebesgue constant Λ_m. Fix $q \in (\overline{q}, 1)$. Then there is $m_0 > 0$, which depends solely on γ, $\overline{q}$, and the chosen q, such that for all $m \geq m_0$

$$\|(\mathrm{Id} - \mathfrak{I}_L^c \circ \cdots \circ \mathfrak{I}_1^c)[\exp(\mathbf{i}\kappa c_0\cdot)\pi]\|_{\infty,J_L} \leq \varepsilon_{m,L}\|\exp(\mathbf{i}\kappa c_0\cdot)\pi\|_{\infty,J_0}, \qquad \forall \pi \in \Pi_m, \tag{55a}$$

$$\|\mathfrak{I}_{\mathfrak{C}}^c\|_{C(J_L)\leftarrow C(J_1)} \leq \Lambda_m\,(1+\varepsilon_{m,L-1})\,, \qquad L \geq 1, \tag{55b}$$

$$\varepsilon_{m,L} := (1 + (1+\Lambda_m)q^m)^L - 1. \tag{55c}$$

Assume additionally that the Lebesgue constant Λ_m of the underlying interpolation operator $\mathfrak{I}$ satisfies (23) for some C_Λ, $\lambda > 0$. Fix $\widehat{q} \in (q, 1)$. Then there is $K > 0$, which depends solely on γ, $\overline{q}$, the chosen $\widehat{q}$, and the constants C_Λ, λ, such that the following implication holds:

$$m \geq K\log(L+2) \qquad \Longrightarrow \varepsilon_{m,L} \leq \widehat{q}^m. \tag{56}$$

Proof. Step 1. (stability of $\mathfrak{I}_i^c$). Let m_0 be given by Lemma 33. Then, combining Lemmas 33 and 34, the following stability and approximation assertions hold for arbitrary $\pi \in \Pi_m$ and $i \in \{1, \dots, L\}$:

$$\|\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi - \mathfrak{I}_i^c[\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi]\|_{\infty,J_i} \leq (1+\Lambda_m)q^m\|\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi\|_{\infty,J_{i-1}}, \tag{57a}$$

$$\|\mathfrak{I}_i^c[\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi]\|_{\infty,J_i} \leq (1+(1+\Lambda_m)q^m)\|\exp(\mathbf{i}\kappa c_{i-1}\cdot)\pi\|_{\infty,J_{i-1}}. \tag{57b}$$

Step 2. We note the following telescoping sum:

$$\begin{aligned}E_1 &:= \mathrm{Id} - \mathfrak{I}_L^c \circ \cdots \circ \mathfrak{I}_1^c \\ &= \mathrm{Id} - \mathfrak{I}_1^c + (\mathrm{Id} - \mathfrak{I}_2^c) \circ \mathfrak{I}_1^c + (\mathrm{Id} - \mathfrak{I}_3^c) \circ \mathfrak{I}_2^c \circ \mathfrak{I}_1^c + \cdots \\ &\quad + (\mathrm{Id} - \mathfrak{I}_L^c) \circ \mathfrak{I}_{L-1}^c \circ \mathfrak{I}_{L-2}^c \circ \cdots \circ \mathfrak{I}_1^c.\end{aligned}$$

Let $\pi \in \Pi_m$ be arbitrary and note that, for each i, we have

$$\mathfrak{I}_i^c \circ \mathfrak{I}_{i-1}^c \circ \mathfrak{I}_{i-2}^c \circ \cdots \circ \mathfrak{I}_1^c[\exp(\mathbf{i}\kappa c_0\cdot)\pi] = \exp(\mathbf{i}\kappa c_i\cdot)\pi_i$$

for some $\pi_i \in \Pi_m$. Hence, the stability and approximation assertions (57a), (57b) are applicable, and we arrive at

$$\begin{aligned}\|E_1[\exp(\mathbf{i}\kappa c_0\cdot)\pi]\|_{\infty,J_L} &\le \sum_{j=0}^{L-1}(1+\Lambda_m)q^m\,(1+(1+\Lambda_m)q^m)^j\,\|\exp(\mathbf{i}\kappa c_0\cdot)\pi\|_{\infty,J_0}\\ &= (1+\Lambda_m)q^m\frac{(1+(1+\Lambda_m)q^m)^L-1}{(1+\Lambda_m)q^m}\|\exp(\mathbf{i}\kappa c_0\cdot)\pi\|_{\infty,J_0}\\ &= \big\{(1+(1+\Lambda_m)q^m)^L-1\big\}\,\|\exp(\mathbf{i}\kappa c_0\cdot)\pi\|_{\infty,J_0},\end{aligned}$$

which show (55a).

Step 3. Inspection shows that (55b) is valid for $L=1$. For $L\ge 2$ and $u\in C(J_1)$, define $\pi_1\in\Pi_m$ by $\pi_1 := \mathfrak{I}_{J_1}[\exp(-\mathbf{i}\kappa c_1\cdot)u]$. By definition of $\mathfrak{I}_1^c$ we have

$$\begin{aligned}&\mathfrak{I}_1^c u = \exp(\mathbf{i}\kappa c_1\cdot)\pi_1, \qquad \|\pi_1\|_{\infty,J_1}\le\Lambda_m\|u\|_{\infty,J_1},\\ &\mathfrak{I}_{\mathfrak{C}}^c u = \mathfrak{I}_L^c\circ\cdots\circ\mathfrak{I}_2^c[\exp(\mathbf{i}\kappa c_1\cdot)\pi_1].\end{aligned}$$

The approximation result (55a) is applicable to the interval chain $J_L\subset J_{J-1}\subset\cdots\subset J_2$ instead of $\mathfrak{C}$, and we get

$$\begin{aligned}\|\mathfrak{I}_{\mathfrak{C}}^c u\|_{\infty,J_L} &= \|\mathfrak{I}_L^c\circ\cdots\circ\mathfrak{I}_2^c[\exp(\mathbf{i}\kappa c_1\cdot)\pi_1]\|_{\infty,J_L}\\ &\le \|\exp(\mathbf{i}\kappa c_1\cdot)\pi_1\|_{\infty,J_L} + \|(\mathrm{Id}-\mathfrak{I}_L^c\circ\cdots\circ\mathfrak{I}_2^c)[\exp(\mathbf{i}\kappa c_1\cdot)\pi_1]\|_{\infty,J_L}\\ &\overset{(55a)}{\le} (1+\varepsilon_{m,L-1})\,\Lambda_m\|u\|_{\infty,J_1},\end{aligned}$$

which shows (55b).

Step 4. Fix $\widetilde{q}\in(q,\widehat{q})$. The assumption (23) on Λ_m implies that for sufficiently large m, we have

$$(1+\Lambda_m)q^m \le (1+C_\Lambda(m+1)^\lambda)q^m \le \widetilde{q}^m,$$

so we obtain

$$\varepsilon_{m,L}\le(1+\widetilde{q}^m)^L-1 = (1+\widetilde{q}^m)^{\widetilde{q}^{-m}\widetilde{q}^m L}-1 \le \exp(L\widetilde{q}^m)-1,$$

where we used $\sup_{x>0}(1+x)^{1/x}\le e$. Using the estimate $\exp(x)-1\le ex$, which is valid for $x\in[0,1]$, and assuming that $\widetilde{q}^m L\le 1$ (note that this holds for $m\ge K\log(L+2)$ for sufficiently large K), we obtain

$$\widehat{q}^{-m}\varepsilon_{m,L}\le e\widehat{q}^{-m}\widetilde{q}^m L = e(\widetilde{q}/\widehat{q})^m L = \exp\big(\log L+\log e - m|\log(\widehat{q}/\widetilde{q})|\big),$$

from which the existence of K that is asserted in (56) can be inferred. □

9.2 Recursive reinterpolation in multi-D

Tensor product arguments allow us to generalize Lemma 35 to the multi-dimensional setting. We formulate this as a separate result in Theorem 36. For $n\in\mathbb{N}$ we consider a

sequence $(B_i)_{i=0}^L$ of nondegenerate boxes of the form $B_i = \prod_{j=1}^n [a_{i,j}, b_{i,j}]$. Let $(c_i)_{i=0}^L \subset \mathbb{R}^n$ with $\|c_i\| \le 1$, $i = 0, \dots, L$, and define the interpolation operators

$$\mathfrak{I}_i^c u := \exp(\mathbf{i}\kappa\langle c_i, \cdot\rangle)\mathfrak{I}_{B_i}(\exp(-\mathbf{i}\kappa\langle c_i, \cdot\rangle)u), \qquad i = 0, \dots, L, \tag{58}$$

where the tensor product operator $\mathfrak{I}_{B_i}$ is based on the 1D operator $\mathfrak{I}$ with Lebesgue constant Λ_m. We write $\Pi_m^n := \operatorname{span}\{\prod_{i=1}^n x_i^{\alpha_i}, \,|\, \alpha_i \in \{0, \dots, m\}\}$ for the usual space of polynomials of degree m in each variable. In this setting, we have the following analog of Lemma 35:

Theorem 36 *Let $n \in \mathbb{N}$, $\kappa \in \mathbb{R}$, $L \in \mathbb{N}$. Assume that the sequence $(B_i)_{i=0}^L$ of boxes is nested, i.e., $B_{i+1} \subset B_i$, $i = 0, \dots, L-1$ and that there is $\overline{q} \in (0,1)$ such that*

$$\max_{i=0,\dots,L-1, j=1,\dots,n} \frac{b_{i+1,j} - a_{i+1,j}}{b_{i,j} - a_{i,j}} \le \overline{q}.$$

Assume $(c_i)_{i=0}^L$ satisfies, for some $\gamma > 0$,

$$\kappa \operatorname{diam} B_i \|c_{i+1} - c_i\| \le \gamma, \qquad i = 0, \dots, L-1.$$

Fix $q \in (\overline{q}, 1)$. Then there is $m_0 > 0$ depending only on $\overline{q}$, q, and γ such that for all $m \ge m_0$ and every $\pi \in \Pi_m^n$

$$\|(\mathrm{Id} - \mathfrak{I}_L^c \circ \cdots \circ \mathfrak{I}_1^c)[\exp(\mathbf{i}\kappa\langle c_0, \cdot\rangle)\pi]\|_{\infty, B_L} \le \varepsilon_{m,L} n \Lambda_m^{n-1} \|\exp(\mathbf{i}\kappa\langle c_0, \cdot\rangle)\pi\|_{\infty, B_0}, \tag{59}$$

$$\|\mathfrak{I}_{\mathfrak{C}}^c\|_{C(B_L) \leftarrow C(B_1)} \le \Lambda_m^n \left(1 + \varepsilon_{m,L-1} n \Lambda_m^{n-1}\right); \tag{60}$$

here, $\varepsilon_{m,L}$ is defined in (55c). Assume additionally that the Lebesgue constant Λ_m of the underlying interpolation operator $\mathfrak{I}$ satisfies (23) for some C_Λ, $\lambda > 0$. Fix $\widehat{q} \in (q, 1)$. Then there is $K > 0$, which depends solely on γ, $\overline{q}$, the chosen $\widehat{q}$, the constants C_Λ, λ, as well as n such that the following implication holds:

$$m \ge K \log(L+2) \qquad \Longrightarrow \varepsilon_{m,L} \le \widehat{q}^m. \tag{61}$$

Proof. The proof follows from tensor product arguments as in Lemma 14. First, we observe that the product structure $\exp(\mathbf{i}\kappa\langle x, c_i\rangle) = \prod_{j=1}^n \exp(\mathbf{i}\kappa x_j c_{i,j})$ implies that the operators $\mathfrak{I}_i^c$ of (58) are tensor product operator with univariate operator of exactly the form analyzed in Lemma 35. The tensor product arguments worked out in Lemma 14 therefore give (59). (Note that Lemma 14 does not rely in an essential way on the fact that polynomial interpolation is considered; essential is the combination of a univariate approximation result with a stability assertion.) Inspection of the proof of (55b) shows that (60) is also true. Also the proof of (61) follows in exactly the same way as that of (56). □

10 Numerical experiments

In order to investigate how accurately our theoretical results predict the convergence of an actual implementation of our nested interpolation scheme, we have implemented a

		$\eta_1 = 10,\ \eta_2 = 1$						
		$m = 2$	$m = 3$	$m = 4$	$m = 5$	$m = 6$	$m = 7$	
$n = 3 \times 2^{10}$	$\kappa = 6$	1.4_{-7}	7.3_{-9}	3.2_{-10}	1.4_{-11}	4.4_{-13}	2.0_{-14}	0.04
$n = 2 \times 2^{12}$	$\kappa = 8$	1.9_{-6}	2.3_{-7}	2.5_{-8}	2.3_{-9}	1.8_{-10}	1.3_{-11}	0.09
$n = 3 \times 2^{12}$	$\kappa = 12$	3.0_{-6}	8.3_{-7}	1.8_{-7}	3.3_{-8}	5.1_{-9}	7.2_{-10}	0.19
$n = 2 \times 2^{14}$	$\kappa = 16$	4.2_{-6}	1.4_{-6}	3.9_{-7}	1.1_{-7}	2.5_{-8}	5.2_{-9}	0.26
$n = 3 \times 2^{14}$	$\kappa = 24$	3.1_{-6}	1.2_{-6}	3.7_{-7}	9.7_{-8}			0.32
$n = 2 \times 2^{16}$	$\kappa = 32$	2.0_{-6}	9.5_{-7}	3.4_{-7}				(0.41)
		$\eta_1 = 10,\ \eta_2 = 2$						
		$m = 2$	$m = 3$	$m = 4$	$m = 5$	$m = 6$	$m = 7$	
$n = 3 \times 2^{10}$	$\kappa = 6$	4.2_{-6}	5.8_{-7}	$5.5e_{-8}$	5.5_{-9}	4.7_{-10}	3.6_{-11}	0.10
$n = 2 \times 2^{12}$	$\kappa = 8$	1.8_{-5}	4.5_{-6}	1.3_{-6}	3.3_{-7}	7.4_{-8}	1.5_{-8}	0.24
$n = 3 \times 2^{12}$	$\kappa = 12$	1.9_{-5}	6.9_{-6}	1.8_{-6}	4.3_{-7}	9.5_{-8}	1.8_{-8}	0.25
$n = 2 \times 2^{14}$	$\kappa = 16$	1.2_{-5}	5.0_{-6}	1.7_{-6}	5.1_{-7}	1.3_{-7}	3.1_{-8}	0.30
$n = 3 \times 2^{14}$	$\kappa = 24$	5.3_{-6}	2.8_{-6}	1.4_{-6}	5.5_{-7}	1.9_{-7}		0.43
$n = 2 \times 2^{16}$	$\kappa = 32$	3.3_{-6}	1.8_{-6}	8.4_{-7}				(0.51)

Table 1: Absolute approximation errors $\|G - \widetilde{G}\|_2$, estimated by the power iteration

"pure" version of the algorithm, i.e., a version that does not use adaptive techniques to improve the compression rate. While we acknowledge that for practical applications an algebraic recompression scheme [17, 1, 3, 4] is crucial, we have chosen this approach to avoid pitfalls like unrealistically low errors due to full rank "approximations".

We use the unit sphere as the surface Γ for our test, approximated by a triangular mesh constructed by regularly subdividing the faces of a double pyramid and scaling the resulting vertices to move them to the unit sphere. We use meshes with $n \in \{4608, 8192, 18432, 32768, 73728, 131072\}$ triangles.

The cluster tree is set up by standard geometrical bisection. The algorithm stops subdividing clusters t as soon as the corresponding index set $\hat{t}$ contains not more than 64 indices.

The wave number κ is chosen to provide us with a high-frequency problem: we have $\kappa h \approx 0.6$, where h denotes the maximal mesh width, i.e., we have approximately ten mesh elements per wavelength.

The approximation $\widetilde{G}$ constructed by our algorithm is compared to the original matrix G, and the spectral norm $\|G - \widetilde{G}\|_2$ of the error is approximated by 20 steps of the power iteration applied to the matrix $(G - \widetilde{G})^*(G - \widetilde{G})$.

Table 1 summarizes our results: the rows correspond to the different meshes, while the columns give the spectral error estimates for different interpolation orders $m \in \{2, \ldots, 7\}$. Missing numbers correspond to experiments that did not fit into our machine's main memory.

The last column of Table 1 gives the geometric mean of the error reduction factors,

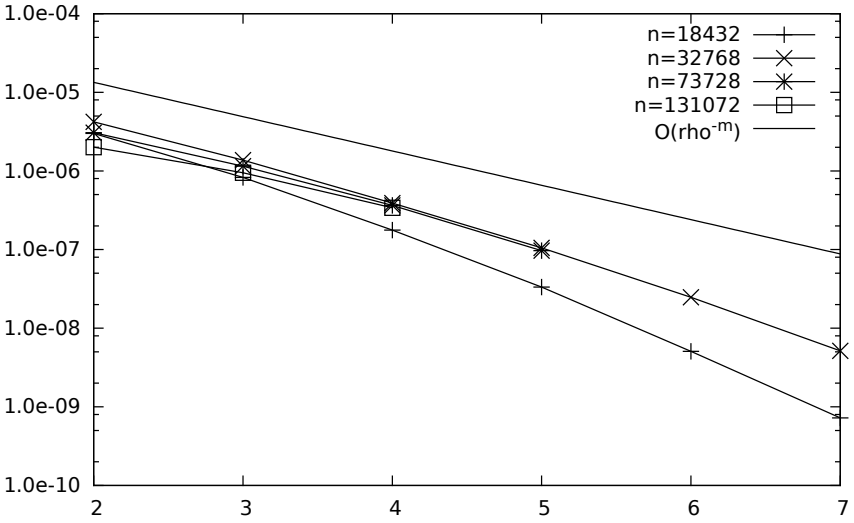

Figure 6: Approximation errors for different meshes and different interpolation orders

and we expect it to be a good approximation of the asymptotic convergence rate.

We investigate two choices for the admissibility parameters η_1 and η_2: for $\eta_1 = 10$, $\eta_2 = 1$, Theorem 27 predicts an asymptotic convergence rate of

$$\varrho = \frac{1}{\sqrt{2}+1} \approx 0.41.$$

We can see that the convergence rates in Table 1 appear to converge to this theoretical bound. This is also illustrated in Figure 6 showing the measured errors for the four finest meshes and the curve predicted by our analysis.

For the second choice $\eta_1 = 10$, $\eta_2 = 2$, we only expect a convergence rate of

$$\varrho = \frac{1}{\sqrt{5/4}+1/2} \approx 0.62.$$

Once again, the measured rates are bounded by the predicted rate.

We conclude that our error estimate appears to be quite accurate, particularly for smaller values of η_2, i.e., for more restrictive admissibility conditions.

References

[1] M. Bebendorf, C. Kuske, and R. Venn. Wideband nested cross approximation for Helmholtz problems. *Num. Math.*, 130(1):1–34, 2015.

[2] S. Börm. Directional $\mathcal{H}^2$-matrix compression for high-frequency problems. Available at `http://arxiv.org`.

[3] S. Börm. Approximation of integral operators by $\mathcal{H}^2$-matrices with adaptive bases. *Computing*, 74(3):249–271, 2005.

[4] S. Börm. *Efficient Numerical Methods for Non-local Operators: $\mathcal{H}^2$-Matrix Compression, Algorithms and Analysis*, volume 14 of *EMS Tracts in Mathematics*. EMS, 2010.

[5] S. Börm, L. Grasedyck, and W. Hackbusch. Introduction to hierarchical matrices with applications. *Eng. Anal. Bound. Elem.*, 27:405–422, 2003.

[6] S. Börm and W. Hackbusch. Data-sparse approximation by adaptive $\mathcal{H}^2$-matrices. *Computing*, 69:1–35, 2002.

[7] S. Börm, M. Löhndorf, and J. M. Melenk. Approximation of integral operators by variable-order interpolation. *Numer. Math.*, 99(4):605–643, 2005.

[8] A. Brandt. Multilevel computations of integral transforms and particle interactions with oscillatory kernels. *Comp. Phys. Comm.*, 65:24–38, 1991.

[9] R. A. DeVore and G. G. Lorentz. *Constructive Approximation*. Springer-Verlag, 1993.

[10] B. Engquist and L. Ying. Fast directional multilevel algorithms for oscillatory kernels. *SIAM J. Sci. Comput.*, 29(4):1710–1737, 2007.

[11] L. Grasedyck and W. Hackbusch. Construction and arithmetics of $\mathcal{H}$-matrices. *Computing*, 70:295–334, 2003.

[12] L. Greengard, J. Huang, V. Rokhlin, and S. Wandzura. Accelerating fast multipole methods for the Helmholtz equation at low frequencies. *IEEE Comp. Sci. Eng.*, 5(3):32–38, 1998.

[13] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comp. Phys.*, 73:325–348, 1987.

[14] H. Guo, J. Hu, and E. Michielssen. On MLMDA/butterfly compressibility of inverse integral operators. *IEEE Ant. Wirel. Prop. Lett.*, 12:31–34, 2013.

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

[16] W. Hackbusch, B. N. Khoromskij, and S. A. Sauter. On $\mathcal{H}^2$-matrices. In H. Bungartz, R. Hoppe, and C. Zenger, editors, *Lectures on Applied Mathematics*, pages 9–29. Springer-Verlag, Berlin, 2000.

[17] M. Messner, M. Schanz, and E. Darve. Fast directional multilevel summation for oscillatory kernels based on Chebyshev interpolation. *J. Comp. Sci.*, 231(4):1175–1196, 2012.

[18] E. Michielssen and A. Boag. A multilevel matrix decomposition algorithm for analyzing scattering from large structures. *IEEE Trans. Antennas and Propagation*, AP-44:1086–1093, 1996.

[19] T. J. Rivlin. *The Chebyshev Polynomials*. Wiley-Interscience, New York, 1990.

[20] V. Rokhlin. Rapid solution of integral equations of classical potential theory. *J. Comp. Phys.*, 60:187–207, 1985.

[21] V. Rokhlin. Diagonal forms of translation operators for the Helmholtz equation in three dimensions. *Appl. Comp. Harm. Anal.*, 1:82–93, 1993.

[22] S. A. Sauter. Variable order panel clustering. *Computing*, 64:223–261, 2000.