Wideband nested cross approximation for Helmholtz problems

M. Bebendorf · C. Kuske · R. Venn

Abstract In this article, the construction of nested bases approximations to discretizations of integral operators with oscillatory kernels is presented. The new method has log-linear complexity and generalizes the adaptive cross approximation method to high-frequency problems. It allows for a continuous and numerically stable transition from low to high frequencies.

Mathematics Subject Classification   41A63 · 65D05 · 65D15 · 35J05

1 Introduction

In this article, the efficient numerical solution of Helmholtz problems

$$-\Delta u - \kappa^2 u = 0 \quad \text{in } \Omega^c, \tag{1a}$$

$$u + \alpha \partial_\nu u = u_0 \quad \text{on } \Gamma := \partial\Omega \tag{1b}$$

used to model acoustics and electromagnetic scattering will be considered. Herein, κ denotes the *wave number* and $\Omega^c := \mathbb{R}^3 \backslash \overline{\Omega}$ the exterior domain of the obstacle $\Omega \subset \mathbb{R}^3$. The parameter α and the right-hand side u_0 appearing in the impedance condition (1b) are given. A convenient way to solve exterior problems is the reformulation as an integral equation [11,13,14] over the boundary Γ of the scatterer Ω. The

M. Bebendorf (✉) · C. Kuske · R. Venn
Institute for Numerical Simulation, Rheinische Friedrich-Wilhelms-Universität Bonn,
Wegelerstr. 6, 53115 Bonn, Germany
e-mail: bebendorf@ins.uni-bonn.de

Galerkin discretization leads to large-scale fully populated matrices $A \in \mathbb{C}^{M \times N}$,

$$a_{ij} = \int_{\Gamma} \int_{\Gamma} K(x, y) \varphi_i(x) \psi_j(y) \, ds_y \, ds_x, \tag{2}$$

with $i \in I := \{1, \ldots, M\}$, $j \in J := \{1, \ldots, N\}$, test and ansatz functions φ_i, ψ_j, having supports $X_i := \operatorname{supp} \varphi_i$ and $Y_j := \operatorname{supp} \psi_j$, respectively. We consider kernel functions K of the form

$$K(x, y) := f(x, y) \exp(i \kappa |x - y|) \tag{3}$$

with an arbitrary *asymptotically smooth* (with respect to x and y) function f, i.e., there are constants $c_{as,1}, c_{as,2} > 0$ such that for all $\alpha, \beta \in \mathbb{N}^3$

$$|\partial_x^\alpha \partial_y^\beta f(x, y)| \le c_{as,1} c_{as,2}^p \alpha! \beta! \frac{|f(x, y)|}{|x - y|^p}, \quad p := |\alpha + \beta|. \tag{4}$$

An example is $K(x, y) = S(x - y)$ used in the single layer ansatz, where $S(x) = \exp(i \kappa |x|)/(4\pi |x|)$ denotes the fundamental solution. Notice that the double layer potential $K(x, y) = \partial_{\nu_y} S(x - y)$ is of the form (3) only if Γ, i.e. the unit outer normal ν, is sufficiently smooth.

Depending on the application, low or high-frequency problems are to be solved. For low-frequency problems, i.e. for $\kappa \operatorname{diam} \Omega \le 1$, the *treecode algorithm* [5] and *fast multipole methods* (FMM) [24–26,35] were introduced to treat A with log-linear complexity. The *panel clustering method* [32] is directed towards more general kernel functions. All previous methods rely on *degenerate approximations*

$$K(x, y) \approx \sum_{i=1}^{k} u_i(x) v_i(y), \quad x \in X, \ y \in Y, \tag{5}$$

using a short sum of products of functions u_i and v_i depending on only one of the two variables x and y chosen from a pair of domains $X \times Y$ which satisfies the *far-field condition*

$$\eta_{\text{low}} \operatorname{dist}(X, Y) \ge \max\{\operatorname{diam} X, \operatorname{diam} Y\} \tag{6}$$

with a given parameter $\eta_{\text{low}} > 0$. Since replacing the kernel function K in the integrals (2) with degenerate approximations (5) leads to matrices of low rank, a more direct approach to the efficient treatment of matrices (2) are algebraic methods such as *mosaic-skeletons* [39] and *hierarchical matrices* [27,29]. These methods also allow to define approximate replacements of usual matrix operations such as addition, multiplication, inversion, and LU factorization; cf. [23]. An efficient and comfortable way to construct low-rank approximations is the *adaptive cross approximation* (ACA) method [6]. The advantage of this approach compared with explicit kernel approximation is that significantly better approximations can be expected due the quasi-optimal

approximation properties; cf. [7]. Furthermore, ACA has the practical advantage that only few of the original entries of A are used for its approximation. A second class are *wavelet compression techniques* [1], which lead to sparse and asymptotically well-conditioned approximations of the coefficient matrix.

It is known that the fundamental solution S (and its derivatives) of any elliptic operator allows for a degenerate approximation (5) on a pair of domains (X, Y) satisfying (6); see [7]. This applies to the *Yukawa operator* $-\Delta + \kappa^2$ for any κ, because the decay of S benefits from the positive shift κ^2. However, the negative shift $-\kappa^2$ in the Helmholtz operator introduces oscillations in S. Hence, for high-frequency Helmholtz problems, i.e. for $\kappa \, \text{diam}\Omega > 1$, the wave number κ enters the degree of degeneracy k in (5) in a way that k grows linearly with κ. In addition to this difficulty, the mesh width h of the discretization has to be chosen such that $\kappa h \sim 1$ for a sufficient accuracy of the solution. We assume that

$$\kappa h \le \frac{1}{2}, \quad h := \max\{\text{diam}X_i, \text{diam}Y_j, \ i \in I, \ j \in J\} \sim 1/\sqrt{N}, \tag{7}$$

which implies that $\kappa \sim \sqrt{N} \sim \sqrt{M}$. Notice that the recent formulation [16] allows to avoid the previous condition and hence leads to significantly smaller N. For high-frequency Helmholtz problems, one- and two-level versions [36,37] were presented with complexity $O(N^{3/2})$ and $O(N^{4/3})$, respectively. Multi-level algorithms [2,18] are able to achieve logarithmic-linear complexity. The previous methods are based on an extensive analytic apparatus that is tailored to the kernel function K. To overcome the instability of some of the employed expansions at low frequencies, a wideband version of FMM was presented in [17]. The $\mathcal{H}^2$-matrix approach presented in [4] is based on the explicit kernel expansions used in [2,36] for two-dimensional problems.

A well-known idea from physical optics (cf. [21]) is to approximate $K(\cdot, y)$ in a given direction $e \in \mathbb{S}^2$ by a plane wave. The desired boundedness of k with respect to κ when approximating

$$\hat{K}(x, y) := K(x, y) \exp(-i\kappa(x - y, e))$$

can be achieved if (6) is replaced by a condition which depends on κ and which is directionally dependent. This is exploited by the fast multipole methods presented in [12,19,20,33] and the so-called butterfly algorithm [15,34]. The aim of this article is to combine this approach with the ease of use of ACA, i.e., our aim is to construct approximations to A with complexity $k^2 N \log N$ using only few of the original entries of A. In this sense, this article generalizes ACA (which achieves log-linear complexity only for low-frequencies) to high-frequency Helmholtz problems. An interesting and important property of the new method is that it will allow for a continuous and numerically stable transition from low to high wave numbers κ by a generalized far-field condition that fades to the usual condition (6) if the wave number becomes small. Since we approximate the operator rather than just its application to a vector, this article is expected to lay ground to future work related to the definition of approximate arithmetic operations and hence to efficient preconditioners for high-frequency problems.

The remaining part of this article is organized as follows. In Sect. 2, we prove estimates like (4) for $\hat{K}$ in a cone around e. The desired asymptotic smoothness of $\hat{K}$ leads to a far-field condition on the pair of domains (X, Y) on which such estimates are valid. In Sect. 3, these conditions will be accounted for by subdividing the matrix indices hierarchically. It will be seen that the number of blocks resulting from this partitioning is too large to allow for hierarchical matrix approximations with log-linear complexity. Therefore, nested bases approximations are required, and in Sect. 4 *directional $\mathcal{H}^2$-matrices* will be introduced as a generalization of usual $\mathcal{H}^2$-matrices [28,31] that incorporate a directional hierarchy. To define approximate arithmetic operations for such matrices, the $\mathcal{H}^2$-matrix operations introduced in [9] have to be generalized to take into account the directional hierarchy.

The results obtained in the first part of this article are amenable to any way of construction of the low-rank approximation, e.g. polynomial interpolation. Sect. 5 is devoted to the construction of directional $\mathcal{H}^2$-matrices using only few of the entries of A. Error estimates for the constructed nested bases are presented and complexity estimates prove the log-linear overall storage and the log-linear number of operations required by the new technique. Finally, Sect. 6 presents numerical experiments that validate our analysis.

2 Directional asymptotic smoothness

In [7] it is proved that the singularity function of any elliptic second-order partial differential operator is asymptotically smooth. The latter property can be used to prove convergence of ACA and hence the existence of degenerate kernel approximations (5). The wave number κ enters the estimates on k in (5) only through the expression $c_\kappa := \max\{1, \kappa \max_{x \in X, y \in Y} |x - y|\}$, which in general becomes unbounded in the limit $\kappa \to \infty$. For parts X of the domain of small size, i.e. $\kappa \operatorname{diam}X \leq 1$, satisfying (6), c_κ and hence k in (5) are bounded independently of κ. This follows from the fact that the recursive construction of domains satisfying (6) ensures that (6) is sharp in the sense that there is a constant $q > 1$ such that

$$\eta_{\text{low}} \operatorname{dist}(X, Y) \leq q \min\{\operatorname{diam}X, \operatorname{diam}Y\}.$$

If the diameters of X and Y are comparable, i.e. $\operatorname{diam}Y \leq c \operatorname{diam}X$ (which is a valid assumption for $\mathcal{H}$-matrix partitions), then

$$\max_{x \in X, y \in Y} |x - y| \leq \operatorname{dist}(X, Y) + \operatorname{diam}X + \operatorname{diam}Y \leq \left(\frac{q}{\eta_{\text{low}}} + 1 + c\right) \operatorname{diam}X$$

and thus

$$c_\kappa \leq \max\{1, \left(\frac{q}{\eta_{\text{low}}} + 1 + c\right) \kappa \operatorname{diam}X\} \leq \frac{q}{\eta_{\text{low}}} + 1 + c \tag{8}$$

is bounded independently of κ. In the other case $\kappa \operatorname{diam}X > 1$, we will not be able to prove asymptotic smoothness with bounded constants. However, a similar property
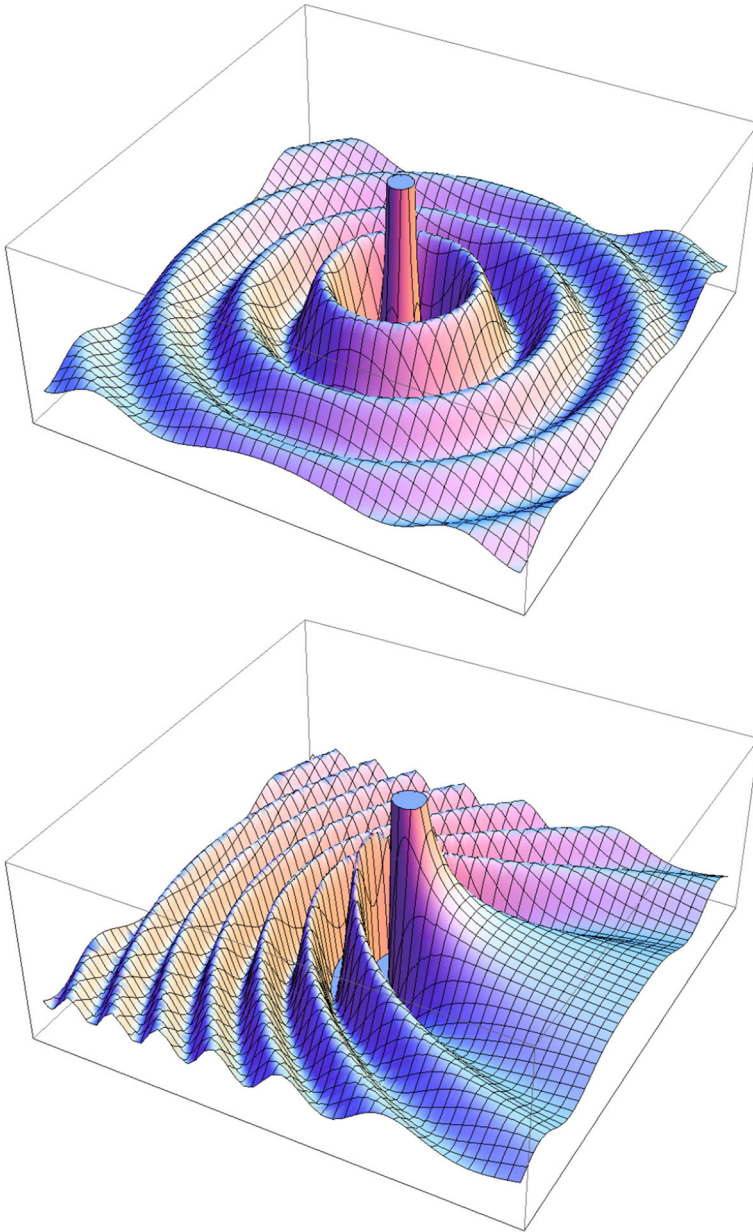
Fig. 1 Re $K(x_1, x_2, 0)$ and Re $\hat{K}(x_1, x_2, 0)$ with $e = (1, 0, 0)^T$

can be proved if the far-field condition (6) is replaced with a frequency dependent condition and if the corresponding far field is subdivided into directions.

For the ease of presentation, K defined in (3) will be investigated as a function of x with fixed y. Hence, after shifting x to $x + y$ we consider

$$\hat{K}(x) = f(x)\exp(i\kappa[|x| - (x, e)]),$$

which can be regarded as K divided by the plane wave $\exp(i\kappa(x, e))$ with some given vector $e \in \mathbb{S}^2$; cf. Fig. 1. It will be shown that $\hat{K}$ is asymptotically smooth with respect to x in a cone around e. To this end, $\hat{K}$ will be investigated in the coordinates $r := |x|$ and

$$\xi := |x| - (x, e).$$

As a first step, we investigate the derivative of ξ in direction e, i.e., we assume that $\cos\varphi > 0$, where $\varphi := \angle(x, e)$. Notice that the desired smoothness cannot be expected for $\cos\varphi \leq 0$; see Fig. 1. Furthermore, it will be seen later that it is required that $\sin\varphi \to 0$ as $\kappa \to \infty$. Hence, the following upper bound on $|\sin\varphi|$ is a valid technical assumption.

Lemma 1 *Let $x \in \mathbb{R}^3$ such that $\cos\varphi > 0$ and $|\sin\varphi| < 1/4$. Then there is a constant $c_{as,3} > 0$ such that*

$$|\partial_e^p \xi(x)| \leq c_{as,3}\, 2^p\, p!\, |x|^{1-p}(\sin\varphi)^2, \quad p \in \mathbb{N}.$$

Proof We may assume that $e = e_1 := (1, 0, 0)^T$. Hence, $x_1 = |x|\cos\angle(x, e_1) > 0$ and $\xi(x) = |x| - x_1$. In order to estimate the p-th order derivative of ξ with respect to x_1, we define

$$b^2 := x_2^2 + x_3^2 = |x|^2 - x_1^2 = (|x|\sin\varphi)^2$$

and extend ξ regarded as a function in x_1 to

$$\hat{\xi}(z) := \sqrt{z^2 + b^2} - z,$$

which is holomorphic in $B_\rho(x_1)$, $\rho := |x|/2$. Consider $z = \alpha + i\beta \in B_\rho(x_1)$. It is easy to see that any ball $B_r(x_1)$ of radius $0 < r < x_1/\sqrt{2}$ is contained in the set $\{(x, y) \in \mathbb{C} : x > |y|\}$. The assumption $|\sin\varphi| < 1/4$ implies that $\cos\varphi > 1/\sqrt{2}$ and hence that $\rho = |x|/2 = x_1/(2\cos\varphi) < x_1/\sqrt{2}$. In particular, we obtain that $\alpha > b > 0$ and $\alpha^2 > \beta^2$. With $A := |z^2 + b^2| = \sqrt{B^2 + 4\alpha^2\beta^2}$, $B := \alpha^2 - \beta^2 + b^2$, we have

$$\sqrt{z^2 + b^2} = \sqrt{\frac{1}{2}(A + B)} + i\,\mathrm{sgn}(\beta)\sqrt{\frac{1}{2}(A - B)}.$$

Due to $|\sqrt{x^2 \pm y^2} - |x|| \leq y^2/\sqrt{x^2 \pm y^2}$ for all $x, y \in \mathbb{R}$, it follows that

$$\left|\frac{1}{2}(A + B) - \alpha^2\right| = \frac{1}{2}|\sqrt{(\alpha^2 + \beta^2 - b^2)^2 + 4\alpha^2 b^2} - (\alpha^2 + \beta^2 - b^2)| \leq 2\frac{b^2\alpha^2}{A}$$

and

$$|\frac{1}{2}(A - B) - \beta^2| = \frac{1}{2}|\sqrt{(\alpha^2 + \beta^2 + b^2)^2 - 4\beta^2 b^2} - (\alpha^2 + \beta^2 + b^2)| \leq 2\frac{b^2\beta^2}{A}.$$

Hence,

$$
\begin{aligned}
|\sqrt{z^2 + b^2} - z|^2 &= |\sqrt{\frac{1}{2}(A + B)} - \alpha|^2 + |\sqrt{\frac{1}{2}(A - B)} - |\beta||^2 \\
&\leq \frac{8b^4}{A^2}\left(\frac{\alpha^4}{A + B} + \frac{\beta^4}{A - B}\right) \\
&\leq \frac{8b^4}{A^2}\left(\frac{\alpha^4}{2\alpha^2 - 4b^2\alpha^2/A} + \frac{\beta^4}{2\beta^2 - 4b^2\beta^2/A}\right) \\
&= \frac{4b^4}{A^2}\frac{|z|^2}{1 - 2b^2/A} \leq \frac{4b^4}{A}\frac{1}{1 - 8\sqrt{2}(\sin\varphi)^2},
\end{aligned}
$$

which follows from

$$
\begin{aligned}
A = |z^2 + b^2| &= \sqrt{|z|^4 + 2b^2(\alpha^2 - \beta^2) + b^4} \geq \sqrt{|z|^4 + b^4} \geq \frac{1}{\sqrt{2}}(|z|^2 + b^2) \\
&\geq \frac{1}{\sqrt{2}}((x_1 - \rho)^2 + b^2) \geq \frac{1}{\sqrt{2}}(|x|^2 - 2\rho|x| + \rho^2) = \frac{|x|^2}{4\sqrt{2}} = \frac{b^2}{4\sqrt{2}(\sin\varphi)^2}
\end{aligned}
$$

and in particular $A \geq |z|^2$. From Cauchy's differentiation formula we obtain

$$
\begin{aligned}
|\partial_{x_1}^p \xi(x)| = |\hat{\xi}^{(p)}(x_1)| &\leq \frac{p!}{2\pi}\int_{\partial B_\rho(x_1)} \frac{|\hat{\xi}(z)|}{|z - x_1|^{p+1}}\, dz \\
&\leq \frac{2^p p!}{|x|^{p-1}}\frac{4\sqrt[4]{2}(\sin\varphi)^2}{\sqrt{1 - 8\sqrt{2}(\sin\varphi)^2}}.
\end{aligned}
$$

$\square$

Using the previous estimate on the derivatives of ξ, we are now able to estimate the derivatives of $\hat{K}$ in directions e and $e_\perp \in \mathbb{S}^2$ perpendicular to e. To this end, we exploit (4) and make use of

$$
\begin{aligned}
|\partial_v^p \hat{K}(x)| &\leq \sum_{i+j=p}\binom{p}{i}|\partial_v^i \exp(i\kappa\xi)|\,|\partial_v^j f(x)| \\
&\leq c_{\mathrm{as},1} p! |\hat{K}(x)| \sum_{i+j=p}\frac{|\partial_v^i \exp(i\kappa\xi)|}{i!}\left(\frac{c_{\mathrm{as},2}}{r}\right)^j,
\end{aligned}
\tag{9}
$$

which holds true for any direction v. Thus, we require an estimate for $|\partial_v^i \exp (i\kappa\xi)|$. This will be done in the following two lemmas for the cases $v = e$ and $v = e_\perp$, respectively.

Lemma 2 *Let x and φ be as in Lemma 1. Let $d > 1/\kappa$ and define $\eta := \kappa d^2/r$, $\gamma := \kappa d|\sin\varphi|$. Then*

$$|\partial_e^p \hat{K}(x)| \le c\, p! \left(\frac{\rho}{d}\right)^p |\hat{K}(x)|, \quad p \in \mathbb{N},$$

where $c := 2c_{as,1} \exp(4c_{as,3}\gamma)$ and $\rho := \max\{c_{as,2}, 4\} \max\{\gamma, \eta\}$.

Proof In order to estimate $|\partial_e^i \exp(i\kappa\xi)|$, we apply Faà di Bruno's formula expressed in terms of Bell polynomials $B_{n,k}$

$$\frac{d^i}{dx^i}(f \circ g)(x) = \sum_{k=0}^{i} f^{(k)}(g(x)) B_{i,k}(g'(x), \ldots, g^{(i-k+1)}(x)).$$

Using Lemma 1 and $\kappa d > 1$, we obtain

$$|\partial_e^i \exp(i\kappa\xi)| = |\sum_{k=0}^{i} \partial_\xi^k \exp(i\kappa\xi) \sum_{\substack{\sum_v j_v = k \\ \sum_v v j_v = i}} \frac{i!}{j_1! j_2! \ldots} \prod_{\ell=1}^{i-k+1} \left(\frac{\partial_e^\ell \xi(x)}{\ell!}\right)^{j_\ell} |$$

$$\le i! \sum_{k=0}^{i} \kappa^k \sum_{\substack{\sum_v j_v = k \\ \sum_v v j_v = i}} \frac{1}{j_1! j_2! \ldots} \prod_{\ell=1}^{i-k+1} \left(\frac{c_{as,3}\, 2^\ell (\sin\varphi)^2}{r^{\ell-1}}\right)^{j_\ell}$$

$$\le 2^i i! \sum_{k=0}^{i} (c_{as,3}\,\kappa (\sin\varphi)^2)^k r^{k-i} \sum_{\substack{\sum_v j_v = k \\ \sum_v v j_v = i}} \frac{1}{j_1! j_2! \ldots}$$

$$= \left(\frac{2}{\kappa d^2}\right)^i i! \sum_{k=0}^{i} (c_{as,3}\gamma^2)^k \eta^{i-k} \sum_{\substack{\sum_v j_v = k \\ \sum_v v j_v = i}} \frac{1}{j_1! j_2! \ldots}$$

$$< \left(\frac{2\tilde{\rho}}{d}\right)^i i! \sum_{k=0}^{i} (c_{as,3}\gamma)^k \sum_{\substack{\sum_v j_v = k \\ \sum_v v j_v = i}} \frac{1}{j_1! j_2! \ldots},$$

where $\tilde{\rho} := \max\{\eta, \gamma\}$ and $j_v = 0$ for all $v > i - k + 1$. From the multinomial theorem for $\boldsymbol{j} \in \mathbb{N}^d$ and $L := i - k + 1$

$$\sum_{|\boldsymbol{j}|=k} \binom{k}{\boldsymbol{j}} x^{\boldsymbol{j}} = (\sum_{i=1}^{L} x_i)^k$$

it follows that

$$\sum_{\substack{\sum_v j_v = k \\ \sum_v v j_v = i}} \frac{1}{j_1! \cdots j_L!} = \frac{2^k}{k!} \sum_{\substack{\sum_v j_v = k \\ \sum_v v j_v = i}} \binom{k}{j} \prod_{\ell=0}^{L} (2^{-\ell})^{j_\ell} \leq \frac{2^k}{k!} \left(\sum_{\ell=0}^{L} 2^{-\ell} \right)^k \leq \frac{4^k}{k!}$$

and hence $|\partial_e^i \exp(i\kappa\xi)| \leq \tilde{c}\, i! (2\tilde{\rho}/d)^i$ with $\tilde{c} := \exp(4 c_{as,3} \gamma)$. Together with (9) this yields

$$\begin{aligned}
|\partial_e^p \hat{K}(x)| &\leq \tilde{c}\, c_{as,1}\, p! \frac{|\hat{K}(x)|}{d^p} \sum_{i+j=p} (2\tilde{\rho})^i \left(\frac{c_{as,2}\, d}{r} \right)^j \\
&\leq \tilde{c}\, c_{as,1}\, p! \left(\frac{2\tilde{\rho}}{d} \right)^p |\hat{K}(x)| \sum_{i+j=p} \left(\frac{c_{as,2}}{2} \right)^j \\
&\leq 2\tilde{c}\, c_{as,1}\, p! \left(\frac{\max\{c_{as,2}, 4\}\tilde{\rho}}{d} \right)^p |\hat{K}(x)|
\end{aligned}$$

due to $d/r = \eta/(\kappa\, d) < \eta \leq \tilde{\rho}$ and $\sum_{j=0}^{p} t^j \leq 2 t^p$ for $t \geq 2$. $\qquad\square$

Lemma 3 *Let $e_\perp \in \mathbb{S}^2$ be any direction perpendicular to $e \in \mathbb{S}^2$ and let d, η, γ as in Lemma 2 such that $\eta, \gamma < 1$. Then*

$$|\partial_{e_\perp}^p \hat{K}(x)| \leq 2 c_{as,1}\, p! \left(\frac{\rho}{d} \right)^p |\hat{K}(x)|,$$

where $\rho := \max\{6/\sqrt{\tilde{\rho}}, 2 c_{as,2}\} \tilde{\rho}$ and $\tilde{\rho} := \max\{\eta, \gamma\}$.

Proof First, we claim that $\partial_{e_\perp}^i \exp(i\kappa\xi)$ consists of at most 3^i summands of the form

$$g(x) := c_g \frac{(i\kappa)^n}{r^{n+2m}} (x, e_\perp)^{2(n+m)-i} \exp(i\kappa\xi),$$

where $n, m \in \mathbb{N}$ with $2(m+n) \geq i \geq m+n$ and $|c_g| \leq 2^i i!$. This can be seen by induction using

$$\begin{aligned}
\partial_{e_\perp} g(x) &= c_g \frac{(i\kappa)^{n+1}}{r^{n+1+2m}} (x, e_\perp)^{2(n+1+m)-(i+1)} \exp(i\kappa\xi) \\
&\quad + (2n + 2m - i) c_g \frac{(i\kappa)^n}{r^{n+2m}} (x, e_\perp)^{2(n+m)-(i+1)} \exp(i\kappa\xi) \\
&\quad - (n + 2m) c_g \frac{(i\kappa)^n}{r^{n+2(m+1)}} (x, e_\perp)^{2(n+m+1)-(i+1)} \exp(i\kappa\xi).
\end{aligned}$$

We estimate

$$\frac{|g(x)|}{i!} \leq \frac{|c_g|}{i!} \frac{\kappa^n}{r^{n+2m}} |(x,e_\perp)|^{2(n+m)-i} \leq 2^i \frac{\kappa^n}{r^{n+2m}} (r|\sin\varphi|)^{2(n+m)-i}$$

$$\leq 2^i \frac{\kappa^{i-(2m+n)}}{r^{i-n}} \left(\frac{\gamma}{d}\right)^{2(n+m)-i}.$$

Here, we used that $|(x,e_\perp)| \leq r|\sin\varphi| = r\gamma/(\kappa d)$. As in the proof of Lemma 2 we have that $d/r < \eta$ and hence

$$\frac{d^i}{i!}|g(x)| = 2^i \frac{\kappa^{i-(2m+n)}}{r^{i-n}} \gamma^{2(n+m)-i} d^{2(i-n-m)} = 2^i \eta^{i-(2m+n)} \gamma^{2(n+m)-i} \left(\frac{d}{r}\right)^{2m}$$

$$\leq 2^i \tilde{\rho}^{2m+n} \leq 2^i \tilde{\rho}^{i/2}.$$

The last estimate follows from $2m + n \geq i/2$. This implies that $|\partial_{e_\perp}^i \exp(i\kappa\xi)| \leq i!(6/d)^i \tilde{\rho}^{i/2}$ and together with (9) we get

$$\frac{d^p}{p!}|\partial_{e_\perp}^p \hat{K}(x)| \leq c_{\mathrm{as},1}|\hat{K}(x)| \sum_{i+j=p} 6^i \tilde{\rho}^{i/2} \left(\frac{c_{\mathrm{as},2}\,d}{r}\right)^j$$

$$\leq c_{\mathrm{as},1}\,(c_{\mathrm{as},2}\,\tilde{\rho})^p|\hat{K}(x)| \sum_{i+j=p} \left(\frac{6}{c_{\mathrm{as},2}\sqrt{\tilde{\rho}}}\right)^i$$

$$\leq c_{\mathrm{as},1}\,(c_{\mathrm{as},2}\,\tilde{\rho})^p|\hat{K}(x)| \sum_{i=0}^p \left(\frac{\hat{c}}{c_{\mathrm{as},2}}\right)^i \leq 2\,c_{\mathrm{as},1}(\hat{c}\,\tilde{\rho})^p|\hat{K}(x)|,$$

where $\hat{c} := \max\{6/\sqrt{\tilde{\rho}}, 2c_{\mathrm{as},2}\}$.   □

We return to the general case of estimating the derivatives of $\hat{K}(x,y)$ for $x \in X$ and $y \in Y$. The last two lemmata show that the derivatives of $\hat{K}$ can be controlled by the parameters η, γ, and d. Let $\chi(X)$ denote the Chebyshev center of X. Using the *angle condition*

$$|\sin\angle(\chi(X) - y, e)| \leq \frac{\gamma_{\mathrm{high}}}{\kappa\,\mathrm{diam}X}, \quad y \in Y, \tag{10}$$

and the *high-frequency far-field condition*

$$\eta_{\mathrm{high}}\mathrm{dist}(X,Y) \geq \kappa(\mathrm{diam}X)^2 \tag{11}$$

with $0 < \gamma_{\mathrm{high}}, \eta_{\mathrm{high}} < 1$, we obtain for the choice $d = \mathrm{diam}X$ and $x \mapsto x - y$ that $d > 1/\kappa$ and

$$\eta = \frac{\kappa\,d^2}{r} = \frac{\kappa(\mathrm{diam}X)^2}{|x-y|} \leq \frac{\kappa(\mathrm{diam}X)^2}{\mathrm{dist}(X,Y)} \leq \eta_{\mathrm{high}}.$$
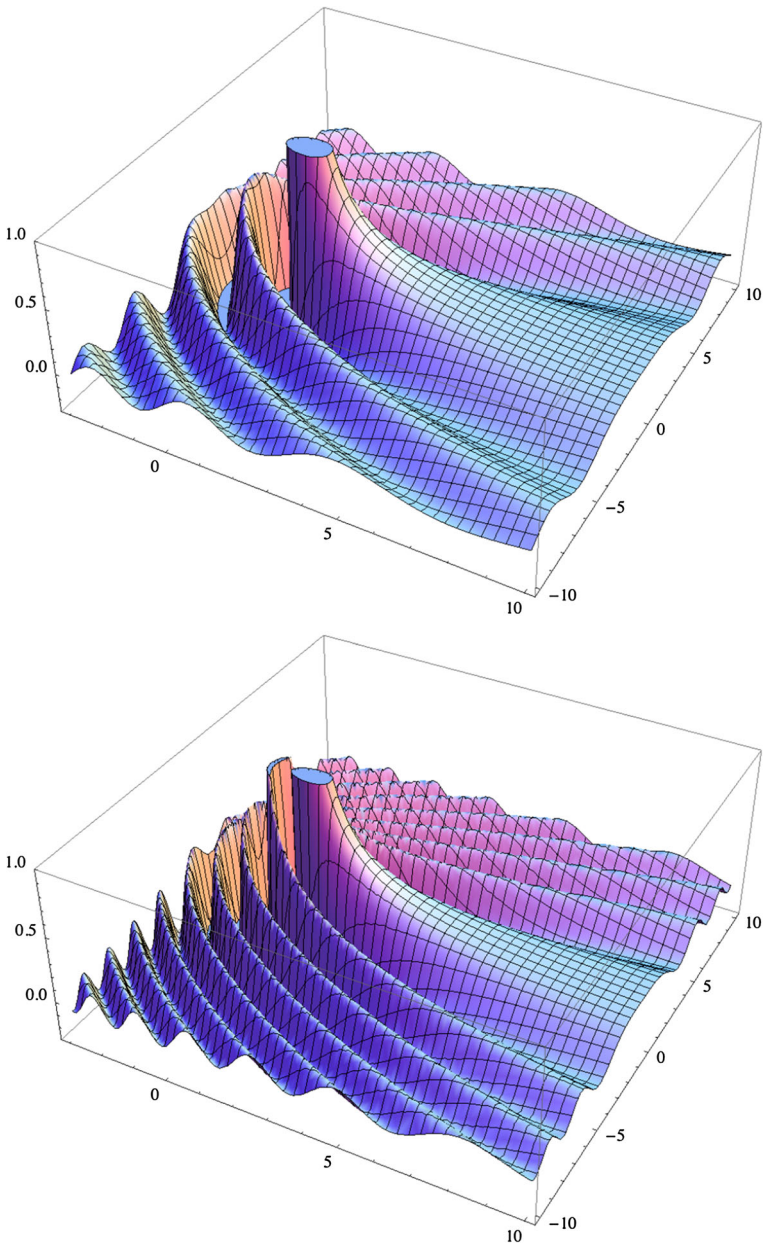
Fig. 2 Re $\hat{K}(x_1, x_2, 0)$ with $e = (1, 0, 0)^T$ for $\kappa = 2$ and $\kappa = 4$

The following lemma shows that $\gamma = \kappa \operatorname{diam} X \sin \sphericalangle(x - y, e)$ is bounded by $\frac{\gamma_{\text{high}} + \eta_{\text{high}}}{1 - \eta_{\text{high}}}$.

Lemma 4 *Let X and Y satisfy* (10) *and* (11). *Then for $x \in X$ and $y \in Y$*

$$| \sin \angle(x - y, e)| \leq \frac{\gamma_{high} + \eta_{high}}{1 - \eta_{high}} \frac{1}{\kappa \, \mathrm{diam} X}.$$

Proof Let $u \times v$ denote the cross product of $u, v \in \mathbb{R}^3$. Then

$$|(x - y) \times e| \leq |x - \chi(X)| + |(\chi(X) - y) \times e| \leq \mathrm{diam} X + |(\chi(X) - y) \times e|.$$

It follows that

$$| \sin \angle(x - y, e)| = \frac{|(x - y) \times e|}{|x - y|} \leq \frac{|(\chi(X) - y) \times e| + \mathrm{diam} X}{|\chi(X) - y| - \mathrm{diam} X}.$$

Due to $\eta_{high}|\chi(X) - y| \geq \kappa(\mathrm{diam} X)^2 \geq \mathrm{diam} X$, we obtain that the denominator of the last expression is bounded from below by $(1 - \eta_{high})|\chi(X) - y|$. Hence,

$$\begin{aligned}
| \sin \angle(x - y, e)| &\leq \frac{1}{1 - \eta_{high}} \frac{|(\chi(X) - y) \times e| + \mathrm{diam} X}{|\chi(X) - y|} \\
&\leq \frac{1}{1 - \eta_{high}} \left(\frac{\gamma_{high}}{\kappa \, \mathrm{diam} X} + \frac{\mathrm{diam} X}{|\chi(X) - y|} \right) \\
&\leq \frac{\gamma_{high} + \eta_{high}}{1 - \eta_{high}} \frac{1}{\kappa \, \mathrm{diam} X}.
\end{aligned}$$

$\square$

As a consequence of the angle condition (10) and the far-field condition (11) we obtain from Lemmas 2 and 3 for $x \in X$ and $y \in Y$

$$\max \left\{ |\partial_{e,x}^p \hat{K}(x, y)|, |\partial_{e_\perp,x}^p \hat{K}(x, y)| \right\} \leq c \, p! \left(\frac{\rho}{\mathrm{diam} X} \right)^p |\hat{K}(x, y)|, \qquad (12)$$

where c is independent of κ, the directions $e, e_\perp \in \mathbb{S}^2$ satisfy $(e, e_\perp) = 0$ and $0 < \rho < 1$ for small enough η_{high} and γ_{high}. Figure 2 shows that the angle condition is not a result of overestimation. The angle of the region in which $\hat{K}$ is smooth decreases with κ.

3 Matrix partitioning

The aim of this section is to partition the set of indices $I \times J$, $I = \{1, \ldots, N\}$ and $J = \{1, \ldots, M\}$, of the matrix defined in (2) into sub-blocks $t \times s$, $t \subset I$ and $s \subset J$, such that the associated supports

$$X_t := \bigcup_{i \in t} X_i \quad \text{and} \quad Y_s := \bigcup_{j \in s} Y_j$$

satisfy (11) in the high-frequency case $\kappa \, \mathrm{diam} X_t > 1$ and (6) if $\kappa \, \mathrm{diam} X_t \leq 1$.

Before we discuss the matrix partition, let us make some assumptions that are in line with the usual finite element discretization. The first assumption is that the overlap of the sets X_i, $i \in I$, is bounded in the sense that there is a constant $\nu > 0$ such that

$$\sum_{i \in t} \mu(X_i) \leq \nu \mu(X_t) \quad \text{for any } t \subset I. \tag{13}$$

Furthermore, the surface measure μ has the property that there is $c_\Gamma > 0$ such that

$$\mu(X) \leq c_\Gamma (\text{diam} X)^2$$

for all $X \subset \Gamma$. The usual way of constructing hierarchical matrix partitions is based on *cluster trees*; see [7,30]. A cluster tree T_I for the index set I is a binary tree with root I, where each $t \in T_I$ and its successors $t', t'' \in T_I$ (if they exist) have the properties $t = t' \cup t''$, $t' \neq \emptyset \neq t''$, and $t' \cap t'' = \emptyset$. We refer to $\mathcal{L}(T_I) = \{t \in T_I : t$ has no successors$\}$ as the leaves of T_I and define

$$T_I^{(\ell)} = \{t \in T_I : \text{dist}(t, I) = \ell\} \subset T_I,$$

where $\text{dist}(t, s)$ is the minimum distance between t and s in T_I. Furthermore,

$$L := L(T_I) := \max\{\text{dist}(t, I), \ t \in T_I\} + 1$$

denotes the depth of T_I. We assume that a cluster tree T_I is constructed such that there are constants $c_g, c_G > 0$ with

$$2^{-\ell}/c_G \leq \mu(X_t) \quad \text{and} \quad (\text{diam} X_t)^2 \leq c_g 2^{-\ell} \tag{14}$$

for all t from the ℓ-th level $T_I^{(\ell)}$ of T_I. We will make use of the notation $S_I(t)$ for the set of sons of a cluster $t \in T_I$. The same properties are also assumed for the sets Y_j, $j \in J$. Under these assumptions it follows that the depth L of the cluster trees is of the order $L \sim \log N \sim \log M$.

A block cluster tree $T_{I \times J}$ is a quad-tree with root $I \times J$ satisfying conditions analogous to a cluster tree. It can be constructed from the cluster trees T_I and T_J in the following way. Starting from the root $I \times J \in T_{I \times J}$, let the sons of a block $t \times s \in T_{I \times J}$ be $S_{I \times J}(t, s) := \emptyset$ if $t \times s$ satisfies the *low-frequency far-field condition*

$$\kappa \min\{\text{diam} X_t, \text{diam} Y_s\} \leq 1, \tag{15a}$$

$$\eta_{\text{low}} \, \text{dist}(X_t, Y_s) \geq \max\{\text{diam} X_t, \text{diam} Y_s\} \tag{15b}$$

or *the high-frequency far-field condition*

$$\kappa \min\{\text{diam} X_t, \text{diam} Y_s\} > 1, \tag{16a}$$

$$\eta_{\text{high}} \, \text{dist}(X_t, Y_s) \geq \kappa \max\{(\text{diam} X_t)^2, (\text{diam} Y_s)^2\} \tag{16b}$$

or $\min\{|t|, |s|\} \le n_{\min}$ with some given constants $\eta_{\text{low}}, \eta_{\text{high}}, n_{\min} > 0$. In all other cases, we set $S_{I \times J}(t, s) := S_I(t) \times S_J(s)$. Notice that for $\kappa = 0$ we obtain the usual far-field condition (6). Furthermore, the transition from the low to the high-frequency regime is continuous in the sense that for $\kappa \max\{\operatorname{diam} X_t, \operatorname{diam} Y_s\} = 1$ the conditions (15b) and (16b) are equivalent with $\eta_{\text{high}} = \eta_{\text{low}}$. We will, however, keep both constants in order to be able to optimize them independently in the implementation of the algorithm.

The set of leaves of $T_{I \times J}$ defines a partition P of $I \times J$. As usual, we partition P into admissible and non-admissible blocks

$$P = P_{\text{adm}} \cup P_{\text{nonadm}},$$

where each $t \times s \in P_{\text{adm}}$ satisfies (15) or (16) and each $t \times s \in P_{\text{nonadm}}$ is small, i.e. satisfies $\min\{|t|, |s|\} \le n_{\min}$. In order to distinguish low and high-frequency blocks, we further subdivide

$$P_{\text{adm}} = P_{\text{low}} \cup P_{\text{high}},$$

where $P_{\text{low}} := \{t \times s \in P : t \times s \text{ satisfies (15)} \}$ and $P_{\text{high}} := P_{\text{adm}} \backslash P_{\text{low}}$.

The following lemma will be the basis for the complexity analysis of the algorithms presented in this article. Notice that this lemma analyzes the so-called *sparsity constant* of hierarchical matrix partitions introduced in [23] for the far-field condition (6). Since the lemma states that this constant is unbounded with respect to κ, usual $\mathcal{H}$-matrices are not able to guarantee logarithmic-linear complexity for the high-frequency far-field condition (16b). Therefore, in the next section a variant of $\mathcal{H}^2$-matrices will be introduced.

Lemma 5 *Let $t \in T_I^{(\ell)}$. The set $\{s \in T_J : t \times s \in P_{high}\}$ has cardinality $\mathcal{O}(1 + 2^{-\ell}\kappa^2)$.*

Proof The assumptions (13) and (14) guarantee that each set

$$N_\rho := \{s \in T_J^{(\ell)} : \max_{y \in Y_s} |\chi(X_t) - y| \le \rho\}, \quad \rho > 0,$$

contains at most $v c_G c_\Gamma 2^\ell (2\rho)^2$ clusters s from the same level ℓ in T_J. This follows from

$$|N_\rho| 2^{-\ell} / c_G \le \sum_{s \in N_\rho} \mu(Y_s) \le v \mu(Y_{N_\rho}) \le v c_\Gamma (2\rho)^2. \tag{17}$$

Let $s \in T_J$ such that $t \times s \in P_{\text{high}}$. Furthermore, let t^* and s^* be the father clusters of t and s, respectively. Suppose that $\max_{y \in Y_s} |\chi(X_t) - y| \ge \rho_0$, where

$$\rho_0 := \kappa / \eta_{\text{high}} \max\{(\operatorname{diam} X_{t^*})^2, (\operatorname{diam} Y_{s^*})^2\} + \operatorname{diam} X_{t^*} + \operatorname{diam} Y_{s^*}.$$

Then

$$\text{dist}(X_{t^*}, Y_{s^*}) \geq \max_{y \in Y_s} |\chi(X_t) - y| - \text{diam}\, X_{t^*} - \text{diam}\, Y_{s^*}$$

$$\geq \kappa/\eta_{\text{high}} \max\{(\text{diam}\, X_{t^*})^2, (\text{diam}\, Y_{s^*})^2\}$$

implies that $t^* \times s^* \in P_{\text{high}}$. Hence, P_{high} cannot contain $t \times s$, which is a contradiction. It follows that

$$\max_{y \in Y_s} |\chi(X_t) - y| < \rho_0 \leq (c_g 2^{-(\ell-1)})^{1/2}(2 + (c_g 2^{-(\ell-1)})^{1/2} \kappa/\eta_{\text{high}}).$$

From (17) we obtain that

$$|\{s \in T_J : t \times s \in P_{\text{high}}\}| \leq 8\nu c_\Gamma c_g c_G (2 + (c_g 2^{-(\ell-1)})^{1/2} \kappa/\eta_{\text{high}})^2$$

and hence the assertion.   □

3.1 Directional subdivision of high-frequency blocks

In the high-frequency regime, i.e. $\kappa \min\{\text{diam}\, X_t, \text{diam}\, Y_s\} > 1$, the matrix block corresponding to $t \times s \in P_{\text{high}}$ cannot be approximated independently of κ unless it is directionally subdivided; see the discussion in Sect. 2. In view of the angle condition (10), we partition the space $\mathbb{R}^3$ recursively into a hierarchy of unbounded pyramids. The first subdivision partitions $\mathbb{R}^3$ into the 6 pyramids defined by the origin and the faces of the unit cube as the pyramids' bases. In each of the next steps, a pyramid is subdivided by dividing its base perpendicular to a largest side of the base into two equally sized halves. A pyramid Z resulting from ν subdivisions satisfies

$$|\sin \angle(x, e)| \leq 2^{(1-\nu)/2} \quad \text{for all } x \in Z, \tag{18}$$

where $e \in \mathbb{S}^2$ denotes the vector pointing from the origin to the center of the base of Z. For a given cluster $t \in T_I^{(\ell)}$ from the ℓ-th level of T_I let ν_t be the smallest non-negative integer such that

$$\nu_t \geq 2(\nu_0 + \log_2 \kappa) - \ell + 1, \tag{19}$$

where $\nu_0 \in \mathbb{N}$ is a fixed value which will be specified later on. Denote by $\mathcal{E}(t)$ the set of directions $e \in \mathbb{S}^2$ associated with all pyramids Z_e after ν_t subdivisions. Notice that depending on t some of the directions will not be used when there are no targets in that particular direction. Such directions can be removed from $\mathcal{E}(t)$. Then

$$|\mathcal{E}(t)| \leq 6 \cdot 2^{\nu_t} \sim \kappa^2 2^{-\ell}. \tag{20}$$

Given $t \in T_I$ and $e \in \mathcal{E}(t)$, for $t' \in S_I(t)$ we define $e' \in \mathcal{E}(t')$ as the axis of the pyramid $Z_{e'}$ from which Z_e results after subdivision. Notice that despite $t' \subset t$ we have $Z_e \subset Z_{e'}$. Furthermore, let

$$\mathcal{F}_e(X_t) := \mathcal{D}_e(X_t) \cap \mathcal{F}(X_t) \quad \text{with}$$

$$\mathcal{D}_e(X) := \left\{ y \in \mathbb{R}^3 : |\sin \angle(\chi(X) - y, e)| \le \frac{\gamma_{high}}{\kappa \operatorname{diam} X} \right\},$$

be the *directional far field* of X_t. Here, $\mathcal{F}(X) := \{y \in \mathbb{R}^3 : \eta_{high} \operatorname{dist}(X, y) \ge \kappa (\operatorname{diam} X)^2\}$ denotes the *far field* of $X \subset \mathbb{R}^3$.

The following lemma will be important for the construction of nested bases used for the approximation. Notice that the directional far fields are nested up to constants.

Lemma 6 *Let* $t' \in S_I(t)$ *and* e' *be defined as above. Then the directional far field satisfies* $\mathcal{F}_e(X_t) \subset \tilde{\mathcal{F}}_{e'}(X_{t'})$ *with*

$$\tilde{\mathcal{F}}_{e'}(X_{t'}) := \left\{ y \in \Gamma : |\sin \angle(\chi(X_{t'}) - y, e')| \le \frac{\tilde{\gamma}}{\kappa \operatorname{diam} X_{t'}} \right\} \cap \mathcal{F}(X_{t'})$$

and the constant $\tilde{\gamma} := 2^{1-\nu_0} \sqrt{c_g} + (\gamma_{high} + \eta_{high})/(1 - \eta_{high})$.

Proof For $y \in \mathcal{F}_e(X_t)$ let $\zeta = \chi(X_{t'}) - y - (\chi(X_{t'}) - y, e)e$, $\zeta' := \chi(X_{t'}) - y - (\chi(X_{t'}) - y, e')e'$, and $\delta := \zeta' - \zeta$. Then

$$|\sin \angle(\chi(X_{t'}) - y, e')| = \frac{|\zeta'|}{|\chi(X_{t'}) - y|} \le \frac{|\zeta| + |\delta|}{|\chi(X_{t'}) - y|}$$

$$= |\sin \angle(\chi(X_{t'}) - y, e)| + \frac{|\delta|}{|\chi(X_{t'}) - y|}$$

and

$$|\delta| = |(\chi(X_{t'}) - y, e)[e - (e, e')e'] - (\chi(X_{t'}) - y, e' - (e, e')e)e'|$$
$$\le |\chi(X_{t'}) - y|[|e - (e, e')e'| + |e' - (e, e')e|] \le 2|\chi(X_{t'}) - y||\sin \angle(e', e)|.$$

Using Lemma 4, we obtain from $\chi(X_{t'}) \in X_{t'} \subset X_t$

$$|\sin \angle(\chi(X_{t'}) - y, e)| \le \frac{\gamma_{high} + \eta_{high}}{1 - \eta_{high}} \frac{1}{\kappa \operatorname{diam} X_t}.$$

From $e' \in \partial Z_e$ it follows

$$|\sin \angle(e', e)| \le 2^{(1-\nu_t)/2} \le \frac{2^{-\nu_0}}{\kappa} 2^{\ell/2} \le \frac{2^{-\nu_0} \sqrt{c_g}}{\kappa \operatorname{diam} X_t}$$

due to (18) and (14). We obtain

$$|\sin \angle(\chi(X_{t'}) - y, e')| \le \left(\frac{\gamma_{high} + \eta_{high}}{1 - \eta_{high}} + 2^{1-\nu_0} \sqrt{c_g} \right) \frac{1}{\kappa \operatorname{diam} X_t}.$$

The inclusion $\mathcal{F}(X_t) \subset \mathcal{F}(X_{t'})$ is obvious due to $X_{t'} \subset X_t$. □

From the previous proof it can be seen that $|\sin \angle(x, e)| \leq \frac{2^{-\nu_0}\sqrt{c_g}}{\kappa \, \mathrm{diam} X_t}$ for all $x \in Z_e$. With Lemma 6 it follows that

$$(\chi(X_t) + Z_e) \cap \mathcal{F}(X_t) \subset \mathcal{F}_{e'}(X_{t'}) \quad \text{for all } e \in \mathcal{E}(t) \tag{21}$$

provided that ν_0 from (19) is chosen sufficiently large and η_{high} is chosen sufficiently small.

Since high-frequency blocks require special attention, we gather column clusters s which are admissible with t in direction $e \in \mathcal{E}(t)$ in the (cluster) far field with direction e

$$F_e(t) := \bigcup \{s \in T_J : \exists \hat{t} \supset t \text{ such that } \hat{t} \times s \in P_{\mathrm{high}} \text{ and } Y_s \subset \mathcal{D}_e(X_t)\}.$$

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

From Lemma 5 we see that $\mathcal{H}$-matrices are not able to achieve logarithmic linear complexity. Therefore we employ a nested representation similar to $\mathcal{H}^2$-matrices introduced in [31]. To account for the required directional subdivision, we generalize the concept of nested cluster bases.

Definition 1 A *directional cluster basis* $\mathcal{U}$ for the rank distribution $(k_t^e)_{t \in T_I, e \in \mathcal{E}(t)}$ is a family $\mathcal{U} = (U_e(t))_{t \in T_I, e \in \mathcal{E}(t)}$ of matrices $U_e(t) \in \mathbb{C}^{t \times k_t^e}$. It is called *nested* if for each $t \in T_I \setminus \mathcal{L}(T_I)$ there are transfer matrices $T_e^{t't} \in \mathbb{C}^{k_{t'}^e \times k_t^e}$ such that for the restriction of the matrix $U_e(t)$ to the rows t' it holds that

$$U_e(t)|_{t'} = U_{e'}(t')T_e^{t't} \quad \text{for all } t' \in S_I(t). \tag{22}$$

For estimating the complexity of storing a nested cluster basis $\mathcal{U}$, we assume that $k_t^e \leq k$ for all $t \in T_I$, $e \in \mathcal{E}(t)$ with some $k \in \mathbb{N}$. It follows from (14) that the depth $L \in \mathbb{N}$ of T_I is of the order $L \sim \log_2(\kappa^2)$. Since the set of leaf clusters $\mathcal{L}(T_I)$ constitutes a partition of I and according to (20) for each cluster $t \in \mathcal{L}(T_I)$ at most $|\mathcal{E}(t)| \sim \kappa^2 2^{-\ell}$ matrices $U_e(t)$ each with at most $k|t|$ entries have to be stored, $k|I|\kappa^2 2^{-L} \sim k N$ units of storage are required for the leaf matrices $U_e(t)$, $t \in \mathcal{L}(T_I)$.

Using (7), we can estimate the storage required for the transfer matrices

$$k^2 \sum_{t \in T_I} |\mathcal{E}(t)| = k^2 \sum_{\ell=0}^{L-1} |T_I^{(\ell)}| \cdot |\mathcal{E}(t)| \sim k^2 \sum_{\ell=0}^{L-1} \kappa^2 = k^2 \kappa^2 L \sim k^2 N \log N.$$

Definition 2 A matrix $M \in \mathbb{C}^{I \times J}$ is a *directional $\mathcal{H}^2$-matrix* if M_b is of low rank for all $b \in P_{\mathrm{low}}$ and there are nested directional cluster bases $\mathcal{U}$ and $\mathcal{V}$ such that for $t \times s \in P_{\mathrm{high}}$

$$M_{ts} = U_e(t)S(t, s)V_{-e}^H(s) \tag{23}$$

with coupling matrices $S(t, s) \in \mathbb{C}^{k_t^e \times k_s^{-e}}$ and $e \in \mathcal{E}(t)$ such that $s \subset F_e(t)$.

The storage cost of the blocks in P_{nonadm} and P_{low} is bounded by the storage cost of a hierarchical matrix, which is known to be $\mathcal{O}(\max\{k, n_{\min}\}N \log N)$. We estimate the storage required for the coupling matrices. According to Lemma 5, the number of blocks $t \times s \in P_{\text{high}}$ is $\mathcal{O}(2^{-\ell}\kappa^2)$ for $t \in T_I^{(\ell)}$. Thus, the coupling matrices require at most

$$k^2 \sum_{t \in T_I} |\{s \in T_J : t \times s \in P_{\text{high}}\}| \sim k^2\kappa^2 \sum_{\ell=0}^{L-1} \sum_{t \in T_I^{(\ell)}} 2^{-\ell} \sim k^2\kappa^2 L \sim k^2 N \log N$$

units of storage. Hence, the overall cost for a directional $\mathcal{H}^2$-matrix is of the order $\mathcal{O}(k^2 N \log N)$.

4.1 Matrix-vector multiplication

Let $M \in \mathbb{C}^{I \times J}$ be a directional $\mathcal{H}^2$-matrix. Since its structure is similar to an $\mathcal{H}^2$-matrix, the matrix-vector multiplication $y := y + Mx$ of M by a vector $x \in \mathbb{C}^J$ can be done via the usual three-phase algorithm (cf. [31]) which we modified to account for the directions e. The following algorithm is a consequence of the decomposition

$$Mx = \sum_{t \times s \in P_{\text{nonadm}}} M_{ts}x_s + \sum_{t \times s \in P_{\text{low}}} W_{ts}Z_{ts}^H x_s + \sum_{t \times s \in P_{\text{high}}} U_e(t)S(t,s)V_{-e}(s)^H x_s$$

with $e \in \mathcal{E}(t)$ such that $s \subset F_e(t)$.

1. *Forward transform*
The auxiliary vectors $\hat{x}_e(s) := V_e(s)^H x_s$, $e \in \mathcal{E}(s)$, are computed for all $s \in T_J$. Exploiting the nestedness of the cluster bases $\mathcal{V}$ (with transfer matrices $\bar{T}_e^{s's}$), one has the following recursive relation

$$\hat{x}_e(s) = V_e(s)^H x_s = \sum_{s' \in S_J(s)} (\bar{T}_e^{s's})^H V_{e'}(s')^H x_{s'} = \sum_{s' \in S_J(s)} (\bar{T}_e^{s's})^H \hat{x}_{e'}(s'),$$

which has to be applied starting from the leaf vectors $\hat{x}_e(s)$, $e \in \mathcal{E}(s)$, $s \in \mathcal{L}(T_J)$.
2. *Far field interaction*
The products $S(t,s)\hat{x}_{-e}(s)$ are computed and summed up over all blocks $t \times s \in P_{\text{high}}$:

$$\hat{y}_e(t) := \sum_{s:t \times s \in P_{\text{high}}} S(t,s)\hat{x}_{-e}(s), \quad e \in \mathcal{E}(t), t \in T_I.$$

3. *Backward transform*
The vectors $\hat{y}_e(t)$ are transformed to the target vector y. The nestedness (22) of $\mathcal{U}$ yields a recursion which descends T_I for the computation of $y :=$ $\sum_{t \in T_I} \sum_{e \in \mathcal{E}(t)} U_e(t)\hat{y}_e(t)$:

(a) Compute $\hat{y}_{e'}(t') := \hat{y}_{e'}(t') + T_e^{t't}\hat{y}_e(t)$ for all $e \in \mathcal{E}(t)$ and all $t' \in S_I(t)$;

(b) Compute $y_t := y_t + \sum_{e \in \mathcal{E}(t)} U_e(t)\hat{y}_e(t)$ for all clusters $t \in \mathcal{L}(T_I)$.

4. *Low-frequency interaction* For all $t \times s \in P_{\text{low}}$ compute $y_t := y_t + W_{ts}Z_{ts}^H x_s$.

5. *Near field interaction*
   For all $t \times s \in P_{\text{nonadm}}$ compute $y_t := y_t + M_{ts}x_s$.

The total number of operations of the latter algorithm is bounded by $\mathcal{O}(k^2 N \log N)$, which can be proven by observing that at most two floating-point operations are performed for each matrix coefficient stored in the directional $\mathcal{H}^2$-matrix representation.

5 Construction of approximations

Our aim is to approximate $A \in \mathbb{C}^{I \times J}$ defined in (2) with a directional $\mathcal{H}^2$-matrix. Blocks in P_{nonadm} are stored entrywise without approximation. From (8) it follows that K is asymptotically smooth with constants independent of κ on domains $X_t \times Y_s$ corresponding to blocks $t \times s \in P_{\text{low}}$. It follows from the convergence analysis in [7] that the adaptive cross approximation

$$A_{ts} \approx A_{t\sigma}(A_{\tau\sigma})^{-1}A_{\tau s} \tag{24}$$

with appropriately chosen $\tau \subset t$ and $\sigma \subset s$, $k_\varepsilon := |\tau| = |\sigma| \sim |\log \varepsilon|^2$ independent of κ, can be used to generate low-rank approximations $W_{ts}Z_{ts}^H \approx A_{ts}$, where $W_{ts} \in \mathbb{C}^{t \times k_\varepsilon}$ and $Z_{ts} \in \mathbb{C}^{s \times k_\varepsilon}$.

In the rest of this section, we will consider the remaining case $t \times s \in P_{\text{high}}$, i.e. we assume that (16) is valid. To be able to apply the results from Sect. 2 and prove existence of low-rank matrix approximations it is required to additionally partition the far field $\mathcal{F}(X_t)$ into subsets $\mathcal{F}_e(X_t)$ corresponding to directions $e \in \mathcal{E}(t)$. Although ACA generates approximations of high quality, the number of blocks is too large (see Lemma 5) to construct and store the approximations as in the low-frequency case. To overcome this difficulty, we consider the approximation (see [8] for the application of this kind of approximation to Laplace-type problems)

$$A_{ts} \approx A_{t\sigma_t^e}(A_{\tau_t\sigma_t^e})^{-1}A_{\tau_t\sigma_s}(A_{\tau_s^{-e}\sigma_s})^{-1}A_{\tau_s^{-e}s}, \quad Y_s \subset \mathcal{F}_e(X_t), \tag{25}$$

instead of (24), which is of type (23) with coupling matrices $S(t,s) = A_{\tau_t\sigma_s}$. The aim of this section is to prove error estimates for the special type of low-rank approximation

$$A_{ts} \approx U_e(t)S(t,s)V_{-e}(s)^H \tag{26}$$

with nested bases $\mathcal{U}$ and $\mathcal{V}$ approximating $A_{t\sigma_t^e}(A_{\tau_t\sigma_t^e})^{-1}$ and $(A_{\tau_s^{-e}\sigma_s})^{-1}A_{\tau_s^{-e}s}$, respectively.

A crucial part of the approximation (25) is the construction of what we call *proper pivots* $\tau_t \subset t$ and $\sigma_t^e \subset F_e(t)$, $|\tau_t| = |\sigma_t^e|$. They have to guarantee that $A_{\tau_t\sigma_t^e}$ is invertible, and for proving error estimates they have to be chosen so that

$$\|A_{ts} - A_{t\sigma_t^e}(A_{\tau_t\sigma_t^e})^{-1}A_{\tau_t s}\| \leq c_R \varepsilon \|A_{tJ}\| \quad \text{for all } s \subset F_e(t)$$

with some $c_R > 0$; cf. Lemma 9. Hence τ_t and σ_t^e represent t and its far field $F_e(t)$, respectively. We refer to [8] for a method for choosing τ_t and σ_t^e. Note that it is sufficient to choose σ_t^e so that

$$Y_{\sigma_t^e} \subset (\chi(X_t) + Z_e) \cap \mathcal{F}(X_t). \tag{27}$$

Here and in what follows, $\varepsilon > 0$ denotes a given accuracy that (up to constants) is to be achieved by the approximations. Let $\{\zeta_1, \ldots, \zeta_{k_\varepsilon}\}$ be a basis of the space

$$\hat{\Pi}_{p_\varepsilon-1}^3 := \{u\, e^{i\kappa(\cdot,e)} : u \in \Pi_{p_\varepsilon-1}^3\},$$

where Π_p^3 denotes the space of polynomials of degree at most p with respect to each of the three variables, $k_\varepsilon := \dim \Pi_{p_\varepsilon-1}^3 = p_\varepsilon^3$, and $p_\varepsilon \in \mathbb{N}$ is the smallest number such that $p_\varepsilon \geq |\log_\rho \varepsilon|$ with ρ from (12).

Assumption 1 Let $t \in T_I$. If $|t| \geq k_\varepsilon$, we assume that there is $\tau_t = \{i_1, \ldots, i_{k_\varepsilon}\} \subset t$ such that the following two conditions are satisfied.

(i) Let $\Lambda \in \mathbb{R}^{|t| \times k_\varepsilon}$ have the entries $\Lambda_{ij} = (\varphi_i/\|\varphi_i\|_{L^2(\Gamma)}, \zeta_j)_{L^2(\Gamma)}$, $i \in t$, $j = 1, \ldots, k_\varepsilon$. There are coefficients $\xi_{i\ell}$ such that

$$\Lambda_{ij} = \sum_{\ell=1}^{k_\varepsilon} \xi_{i\ell} \Lambda_{i_\ell j}, \quad i \in t, \ j = 1, \ldots, k_\varepsilon, \tag{28}$$

and the norm $\|\Xi\|_\infty$ of the matrix $\Xi = (\xi_{i\ell})_{i\ell} \in \mathbb{C}^{t \times k_\varepsilon}$ is bounded by a multiple of 2^{p_ε},

(ii) each matrix $A_{\tau_t F_e(t)}$, $e \in \mathcal{E}(t)$, has full rank.

In the remaining case $|t| < k_\varepsilon$, we set $\tau_t = t$ and assume only that $A_{t F_e(t)}$ has full rank.

To see that these assumptions are reasonable, notice that rank $\Lambda \leq k_\varepsilon$, which implies (28). The coefficients $\xi_{i\ell}$ depend only on the underlying grid. Hence, the boundedness of the norm of Ξ in (i) is an assumption on the regularity of the discretization and on the choice of the pivots τ_t. Notice that the assumed bound grows exponentially with p_ε. In practice, it seems that a polynomial growth $\|\Xi\|_\infty \leq c p_\varepsilon^2$ with respect to p_ε can be achieved. Hence, this should be regarded as a decent assumption.

The set τ_t is not unique. Usually, any sub-set of t having k_ε elements will do. To see this, assume for the time being that $\varphi_i = \delta_{x_i}$. In this case, the sub-matrix $(\Lambda_{i_\ell j})_{\ell j} \in \mathbb{R}^{k_\varepsilon \times k_\varepsilon}$ of Λ is non-singular if and only if $\hat{\Pi}_{p_\varepsilon-1}^3$ is unisolvent with respect to the points x_{i_ℓ}, $\ell = 1, \ldots, k_\varepsilon$, i.e.

$$e^{i\kappa(x_{i_\ell},e)} q(x_{i_\ell}) = 0 \iff q(x_{i_\ell}) = 0, \quad \ell = 1, \ldots, k_\varepsilon,$$

for some $q \in \Pi_{p_\varepsilon-1}^3$ implies that $q = 0$. The set of tupels $(x_{i_1}, \ldots, x_{i_{k_\varepsilon}})$ for which $\Pi_{p_\varepsilon-1}^3$ is not unisolvent is known to be of measure zero (in the parameter domain of

Γ); see [38]. Hence, we may assume that $\tau_t \subset t$ is chosen such that also (ii) is valid. Otherwise, the rank of $A_{tF_e(t)}$ is already bounded by k_ε.

Notice that with the previous assumptions it is possible to guarantee that the rows τ_t used for the approximation of A_{ts} can be chosen independently of $s \subset F_e(t)$. This will be crucial for the construction of nested bases.

In the following lemma, we prove error estimates for the multivariate tensor product Chebyshev interpolant $\mathfrak{I}_p \hat{K}(\cdot, y) \in \Pi_{p-1}^3$ with fixed $y \in Y$ of $\hat{K}(\cdot, y) := K(\cdot, y) \exp(-i\kappa(\cdot - y, e))$.

Lemma 7 *Let $X, Y \subset \mathbb{R}^3$ such that $Y \subset \mathcal{F}_e(X)$. Then*

$$|\hat{K}(x, y) - \mathfrak{I}_{x,p}\hat{K}(x, y)| \leq c_{\mathfrak{I}}(p) \left(\frac{\rho}{\rho + 2}\right)^p \max_{x' \in X} |\hat{K}(x', y)| \quad \text{for all } x \in X, \ y \in Y,$$

where $c_{\mathfrak{I}}(p) := 8ecp(1 + \rho)(1 + \frac{2}{\pi} \log p)^3$ with c and ρ from (12).

Proof Without loss of generality we may assume that X is contained in a cube $Q = \prod_{i=1}^3 Q_i$ which is aligned with e and that $Y \subset \mathcal{F}_e(Q)$. Notice that this can be achieved by slightly modifying the constants $\gamma_{\text{high}}, \eta_{\text{high}}$. Let $\hat{K}_i$ be the function in the i-th argument of $\hat{K}(\cdot, y)$, $i = 1, 2, 3$. From (12) we obtain

$$\|\hat{K}_i^{(k)}\|_{Q_i, \infty} \leq c\, k! \left(\frac{\rho}{\text{diam}\, Q_i}\right)^k \|\hat{K}_i\|_{Q_i, \infty}, \quad k \in \mathbb{N}.$$

Using [10, Lemma 3.13], this implies

$$\min_{q \in \Pi_{p-1}} \|\hat{K}_i - q\|_{Q_i, \infty} \leq \tilde{c}\, p \left(\frac{\rho}{\rho + 2}\right)^p \|\hat{K}_i\|_{Q_i, \infty},$$

where $\tilde{c} := 4ec(1 + \rho)$. With this estimate the proof can be done analogously to Theorem 3.18 in [7]. □

Although the following estimates will hold in any absolute norm, throughout this article the max norm

$$\|A\| := \max_{i=1,\ldots,m} \max_{i=1,\ldots,n} |a_{ij}|$$

of $A \in \mathbb{C}^{m \times n}$ will be used if not otherwise indicated. Notice that the max norm is not a matrix norm but satisfies $\|AB\| \leq \|A\|\|B\|_1$ and $\|AB\| \leq \|A\|_\infty \|B\|$, where $\|\cdot\|_\infty, \|\cdot\|_1$ denote the absolute row and the absolute column sum, respectively.

Lemma 8 *Let assumption (i) be valid and let φ_i, ψ_j, and f in (2) be non-negative. Assume that $c_{as,2}\, \eta_{\text{high}} < 1$. For $t \in T_I$ satisfying $|t| \geq k_\varepsilon$ and $e \in \mathcal{E}(t)$ there is $\Xi \in \mathbb{R}^{t \times k_\varepsilon}$ and an ε-independent constant $c_1 > 0$ such that*

$$\|A_{ts} - \Xi A_{\tau_t s}\| \leq c_1 \varepsilon \|A_{ts}\|$$

for all $s \subset J$ satisfying $Y_s \subset \mathcal{F}_e(X_t)$.

Proof Due to $K(x, y) = \exp(i\kappa(x - y, e))\hat{K}(x, y)$, we can apply Lemma 7 and obtain for $x \in X_t$, $y \in Y_s$ that with $T_{p_\varepsilon}(x, y) := \exp(i\kappa(x, e))\mathfrak{I}_{x, p_\varepsilon}\hat{K}(x, y) \in \hat{\Pi}^3_{p_\varepsilon - 1}$

$$|K(x, y) - T_{p_\varepsilon}(x, y)| = |\hat{K}(x, y) - \mathfrak{I}_{x, p_\varepsilon}\hat{K}(x, y)|$$
$$\leq c_{\mathfrak{I}}(p_\varepsilon)\left(\frac{\rho}{\rho + 2}\right)^{p_\varepsilon} \max_{x' \in X_t} |\hat{K}(x', y)|.$$

According to a remark following (12), we may assume that $\rho < 1$. In addition, without loss of generality we may assume that $\|\varphi_i\|_{L^2(\Gamma)} = 1$. Then assumption (28) is equivalent with

$$\int_\Gamma \left(\varphi_i(x) - \sum_{\ell=1}^{k_\varepsilon} \xi_{i\ell}\varphi_{i_\ell}(x)\right)\zeta(x)\,\mathrm{d}s_x = 0 \quad \text{for all } \zeta \in \hat{\Pi}^3_{p_\varepsilon - 1}.$$

Hence,

$$a_{ij} - \sum_{\ell=1}^{k_\varepsilon} \xi_{i\ell}a_{i_\ell j} = \int_\Gamma \int_\Gamma \left(\varphi_i(x) - \sum_{\ell=1}^{k_\varepsilon} \xi_{i\ell}\varphi_{i_\ell}(x)\right) K(x, y)\psi_j(y)\,\mathrm{d}s_y\,\mathrm{d}s_x$$
$$= \int_\Gamma \left(\varphi_i(x) - \sum_{\ell=1}^{k_\varepsilon} \xi_{i\ell}\varphi_{i_\ell}(x)\right) \int_\Gamma [K(x, y) - T_{p_\varepsilon}(x, y)]\psi_j(y)\,\mathrm{d}s_y\,\mathrm{d}s_x$$

and we see using $\mu(X_i) \leq \pi h^2$ that

$$|a_{ij} - \sum_{\ell=1}^{k_\varepsilon} \xi_{i\ell}a_{i_\ell j}| \leq c_{\mathfrak{I}}(p_\varepsilon)\left(\frac{\rho}{\rho + 2}\right)^{p_\varepsilon}(1 + \|\Xi\|_\infty)\left(\mu(X_i \cup \bigcup_{\ell=1}^{k_\varepsilon} X_{i_\ell})\right)^{1/2}$$
$$\cdot \max_{x' \in X_t} \int_\Gamma |\hat{K}(x', y)||\psi_j(y)|\,\mathrm{d}s_y$$
$$\leq \tilde{c}h\left(\frac{\rho}{\rho + 2}\right)^{p_\varepsilon} \max_{x' \in X_t} \int_\Gamma |\hat{K}(x', y)||\psi_j(y)|\,\mathrm{d}s_y$$

with $\tilde{c} := c_{\mathfrak{I}}(p_\varepsilon)(1 + \|\Xi\|_\infty)\sqrt{(k_\varepsilon + 1)\pi}$. From the Taylor expansion and the asymptotic smoothness of f it can be seen that for $c_{as,2}\,\eta_{high} < 1$

$$|f(x', y)| \leq \hat{c}|f(x, y)| \quad \text{for all } x, x' \in X_t \tag{29}$$

with a constant $\hat{c} > 0$. Estimate (29) and $1 = \|\varphi_i\|_{L^2(\Gamma)} \leq c'h^{-1}\|\varphi_i\|_{L^1(\Gamma)}$ imply for $y \in Y_s$

$$\max_{x' \in X_t} |\hat{K}(x', y)| = \max_{x' \in X_t} |f(x', y)| \leq \hat{c}\min_{x \in X_i} |f(x, y)|$$
$$\leq \hat{c}\,c'h^{-1}\int_\Gamma |\varphi_i(x)||f(x, y)|\,\mathrm{d}s_x.$$

From

$$
\begin{aligned}
|a_{ij}| &= |e^{i\kappa[-|\xi(X_i)-\xi(Y_j)|]}a_{ij}| \\
&\geq \operatorname{Re} \int_\Gamma \int_\Gamma \varphi_i(x)\psi_j(y)f(x,y)e^{i\kappa[|x-y|-|\xi(X_i)-\xi(Y_j)|]}\,ds_y\,ds_x \\
&= \int_\Gamma \int_\Gamma \varphi_i(x)\psi_j(y)f(x,y)\cos(\kappa[|x-y|-|\xi(X_i)-\xi(Y_j)|])\,ds_y\,ds_x \\
&\geq c'' \int_\Gamma \int_\Gamma \varphi_i(x)\psi_j(y)f(x,y)\,ds_y\,ds_x
\end{aligned}
$$

with $c'' > 0$ independent of κ and h due to (cf. (7))

$$
\kappa[|x-y|-|\xi(X_i)-\xi(Y_j)|] \leq \kappa[|x-\xi(X_i)|+|y-\xi(Y_j)|] \leq 2\kappa h < \frac{\pi}{2},
$$

we obtain exploiting the non-negativity of φ_i, ψ_j, and f

$$
|a_{ij} - \sum_{\ell=1}^{k_\varepsilon} \xi_{i\ell}a_{i_\ell j}| \leq \frac{\tilde{c}\hat{c}\,c'}{c''}\left(\frac{\rho}{\rho+2}\right)^{p_\varepsilon}|a_{ij}|.
$$

Hence, the matrix Ξ satisfies

$$
\|A_{ts} - \Xi A_{\tau_t s}\| \leq \frac{\tilde{c}\hat{c}\,c'}{c''(\rho+2)^{p_\varepsilon}}\rho^{p_\varepsilon}\|A_{ts}\| \leq c_1\varepsilon\|A_{ts}\|,
$$

because $\frac{c_{\mathfrak{J}}(p_\varepsilon)\hat{c}c'}{c''(\rho+2)^{p_\varepsilon}}(1+\|\Xi\|_\infty)\sqrt{(p_\varepsilon^3+1)\pi}$ is bounded by an ε-independent constant c_1 from above due to the assumption that $\|\Xi\|_\infty$ is bounded by a multiple of 2^{p_ε} and $c_{\mathfrak{J}}(p_\varepsilon)$ grows at most algebraically with p_ε. $\qquad\square$

The expression $c_S := \max\{c_S^{(r)}, c_S^{(c)}, \|(A_{\tau_s^e\sigma_s})^{-1}A_{\tau_s^e s}\|_1 : s \in T_J\}$, where

$$
c_S^{(c)} := \max\{\|(A_{\tau_t\sigma_t^e})^{-1}A_{\tau_t s}\|_1 : s \subset J,\ Y_s \subset \mathcal{F}_e(X_t),\ e \in \mathcal{E}(t),\ t \in T_I\},
$$

$$
c_S^{(r)} := \max\{\|A_{t\sigma_s}(A_{\tau_s^e\sigma_s})^{-1}\|_\infty : t \subset I,\ X_t \subset \mathcal{F}_e(Y_s),\ e \in \mathcal{E}(s),\ s \in T_J\},
$$

will play a central role in the following error analysis. Notice that c_S does not depend on the cardinality of the clusters. However, it depends on the choice of the pivots $\{\tau_t, \sigma_t^e, \tau_s^e, \sigma_s : t \in T_I,\ s \in T_J\}$ and hence on ε. Numerical experiments show that $c_S \leq cp_\varepsilon^2$.

Lemma 9 *Let Assumption 1 be valid. Then for $t \in T_I$ and $e \in \mathcal{E}(t)$ there are proper pivots (τ_t, σ_t^e), $|\tau_t| = |\sigma_t^e| = \min\{k_\varepsilon, |t|\}$, i.e., for all $s \subset J$ satisfying $Y_s \subset \mathcal{F}_e(X_t)$*

$$
\|A_{ts} - A_{t\sigma_t^e}(A_{\tau_t\sigma_t^e})^{-1}A_{\tau_t s}\| \leq c_2\varepsilon\|A_{tJ}\|, \tag{30}
$$

where $c_2(\varepsilon) := c_1(1 + c_S(\varepsilon))$.

Proof Since $A_{\tau_t F_e(t)}$ is assumed to have full rank, there is $\sigma_t^e \subset F_e(t)$, $|\sigma_t^e| = |\tau_t| = \min\{k_\varepsilon, |t|\}$, such that $A_{\tau_t \sigma_t^e}$ is invertible. If $|t| < k_\varepsilon$, we have $\tau_t = t$ and $A_{t\sigma_t^e}(A_{\tau_t \sigma_t^e})^{-1} = I$. Hence, (30) is trivially satisfied. We may hence assume that $|t| \geq k_\varepsilon$. Let $\Xi \in \mathbb{R}^{t \times k_\varepsilon}$ be as in Lemma 8. We have

$$A_{ts} - A_{t\sigma_t^e}(A_{\tau_t \sigma_t^e})^{-1}A_{\tau_t s} = \{A_{ts} - \Xi A_{\tau_t s}\} - \{A_{t\sigma_t^e} - \Xi A_{\tau_t \sigma_t^e}\}(A_{\tau_t \sigma_t^e})^{-1}A_{\tau_t s}$$

and thus

$$\|A_{ts} - A_{t\sigma_t^e}(A_{\tau_t \sigma_t^e})^{-1}A_{\tau_t s}\| \leq \|A_{ts} - \Xi A_{\tau_t s}\| + \|A_{t\sigma_t^e} - \Xi A_{\tau_t \sigma_t^e}\|c_S$$
$$\leq c_1 \varepsilon(\|A_{ts}\| + c_S\|A_{t\sigma_t^e}\|) \leq c_2 \varepsilon\|A_{tJ}\|.$$

The second last estimate follows from Lemma 8, because $Y_{\sigma_t^e}, Y_s \subset \mathcal{F}_e(X_t)$. □

5.1 Construction of directional cluster bases

Based on the previous estimates, we are going to construct and analyze nested bases approximations. The construction of nested bases is usually done by explicit approximation of the kernel function K; see, for instance, the fast multipole method [37] and the method in [33], which uses interpolation. In this section, we construct the nested bases via a purely algebraic technique which is based on the original matrix entries and thus avoids explicit kernel expansions. In this sense, the presented construction is in the class of kernel independent fast multipole methods; see [3,19,40].

We will define a nested basis $\mathcal{U}$ consisting of matrices $U_e(t) \in \mathbb{C}^{t \times k_\varepsilon}$ for each $t \in T_I$ and $e \in \mathcal{E}(t)$ in a recursive manner starting from the leaves. Due to (16a), it is actually sufficient to consider the sub-tree

$$\hat{T}_I := \{t \in T_I : \kappa \operatorname{diam} X_t > 1\}$$

of T_I. For its leaf clusters $t \in \mathcal{L}(\hat{T}_I)$ and $e \in \mathcal{E}(t)$ we set $U_e(t) = B_e^{tt}$, where for $t' \subset t$

$$B_e^{t't} := A_{t'\sigma_t^e}(A_{\tau_t \sigma_t^e})^{-1}.$$

Assume that matrices $U_e(t')$ have already been constructed for the sons $t' \in S_I(t)$ of $t \in \hat{T}_I \setminus \mathcal{L}(\hat{T}_I)$ and $e \in \mathcal{E}(t')$. Then in view of (22) we define for $e \in \mathcal{E}(t)$

$$U_e(t)|_{t'} := U_{e'}(t')B_e^{\tau_{t'}t}, \quad t' \in S_I(t),$$

where $e' \in \mathcal{E}(t')$ is defined before Lemma 6.

The following lemma estimates the accuracy when expressing $B_e^{t't}$ by the product $B_{e'}^{t't'}B_e^{\tau_{t'}t}$. As stated in [8], $B_e^{t't}$ may be regarded as the algebraic form of an interpolation operator.

Lemma 10 *Let $t' \in \hat{T}_I$ satisfy $t' \subset t \in \hat{T}_I \backslash \mathcal{L}(\hat{T}_I)$ and $e \in \mathcal{E}(t)$. Then for all $s \subset J$ satisfying $Y_s \subset \mathcal{F}_e(X_t)$ it holds that*

$$\| [B_e^{t't} - B_{e'}^{t't'} B_e^{\tau_{t'}t}] A_{\tau_t s} \| \leq c_2 c_S \varepsilon \| A_{t'J} \|$$

with $c_S = c_S(\varepsilon)$ and $c_2 = c_2(\varepsilon)$ from Lemma 9.

Proof It is clear that

$$[B_e^{t't} - B_{e'}^{t't'} B_e^{\tau_{t'}t}] A_{\tau_t s} = [A_{t'\sigma_t^e} - A_{t'\sigma_{t'}^{e'}} (A_{\tau_{t'}\sigma_{t'}^{e'}})^{-1} A_{\tau_{t'}\sigma_t^e}] (A_{\tau_t\sigma_t^e})^{-1} A_{\tau_t s}.$$

Due to (27) and (21), it holds that $Y_{\sigma_t^e} \subset \mathcal{F}_{e'}(X_{t'})$. Hence, Lemma 9 yields

$$\| [B_e^{t't} - B_{e'}^{t't'} B_e^{\tau_{t'}t}] A_{\tau_t s} \| \leq c_2 \varepsilon \| A_{t'J} \| \| (A_{\tau_t\sigma_t^e})^{-1} A_{\tau_t s} \|_1.$$

$\square$

Theorem 1 *Let Assumption 1 be valid. Let $t \in \hat{T}_I$ and let $\ell = L(\hat{T}_t)$ denote the depth of the sub-tree $\hat{T}_t$ of $\hat{T}_I$ rooted at $t \in T_I$. Then for all $e \in \mathcal{E}(t)$*

$$\| [U_e(t) - B_e^{tt}] A_{\tau_t s} \| \leq c_3 \varepsilon \| A_{tJ} \| \quad \text{for all } s \subset F_e(t),$$

where

$$c_3(\varepsilon) := c_1 (1 + c_S(\varepsilon)) \frac{(2c_S(\varepsilon))^\ell}{2c_S(\varepsilon) - 1}.$$

Proof From Lemma 10 we have for $t \in \hat{T}_I \backslash \mathcal{L}(\hat{T}_I)$

$$
\begin{aligned}
\| [U_e(t) - B_e^{tt}] A_{\tau_t s} \| &\leq \sum_{t' \in S_I(t)} \| [U_e(t)|_{t'} - B_e^{t't}] A_{\tau_t s} \| \\
&= \sum_{t' \in S_I(t)} \| [U_{e'}(t') B_e^{\tau_{t'}t} - B_e^{t't}] A_{\tau_t s} \| \\
&\leq \sum_{t' \in S_I(t)} \| [U_{e'}(t') - B_{e'}^{t't'}] B_e^{\tau_{t'}t} A_{\tau_t s} \| + \| [B_e^{t't} - B_{e'}^{t't'} B_e^{\tau_{t'}t}] A_{\tau_t s} \| \\
&\leq \sum_{t' \in S_I(t)} c_S \| [U_{e'}(t') - B_{e'}^{t't'}] A_{\tau_{t'}\sigma_t^e} \| + \varepsilon c_2 c_S \| A_{t'J} \| \\
&\leq 2 c_2 c_S \varepsilon \| A_{tJ} \| + c_S \sum_{t' \in S_I(t)} \| [U_{e'}(t') - B_{e'}^{t't'}] A_{\tau_{t'}\sigma_t^e} \|.
\end{aligned}
$$

We set $\alpha_{t'} := \| [U_{e'}(t') - B_{e'}^{t't'}] A_{\tau_{t'}\sigma_t^e} \|$ for $t' \in S_I(t)$. From (27) and (21) we obtain that $Y_{\sigma_t^e} \subset \mathcal{F}_{e'}(X_{t'})$. Hence, we can apply the latter inequality recursively replacing s by σ_t^e and obtain the recurrence relation

$$\alpha_t \le 2c_2 c_S \varepsilon \|A_{tJ}\| + c_S \sum_{t' \in S_I(t)} \alpha_{t'}, \quad t \in T_I \setminus \mathcal{L}(T_I). \tag{31}$$

We show that

$$\alpha_t \le 2c_2 c_S \varepsilon \frac{(2c_S)^{\ell-1} - 1}{2c_S - 1} \|A_{tJ}\|, \quad t \in T_I, \tag{32}$$

where $\ell = \ell(t)$ denotes the depth of the sub-tree $\hat{T}_t$. This estimate is obviously true for leaf clusters $t \in \mathcal{L}(T_I)$ as $\alpha_t = 0$. Assume that (32) is valid for the sons $S_I(t)$ of $t \in \hat{T}_I \setminus \mathcal{L}(\hat{T}_I)$. Then (31) proves

$$
\begin{aligned}
\alpha_t &\le 2c_2 c_S \varepsilon \|A_{tJ}\| + c_S \sum_{t' \in S_I(t)} \alpha_{t'} \\
&\le 2c_2 c_S \varepsilon \|A_{tJ}\| + 2c_2 c_S^2 \varepsilon \frac{(2c_S)^{\ell-2} - 1}{2c_S - 1} \sum_{t' \in S_I(t)} \|A_{t'J}\| \\
&\le 2c_2 c_S \varepsilon \left(1 + 2c_S \frac{(2c_S)^{\ell-2} - 1}{2c_S - 1} \right) \|A_{tJ}\| = 2c_2 c_S \varepsilon \frac{(2c_S)^{\ell-1} - 1}{2c_S - 1} \|A_{tJ}\|.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\|[U_e(t) - B_e^{tt}] A_{\tau_t s}\| &\le 2c_2 c_S \varepsilon \|A_{tJ}\| + c_S \sum_{t' \in S_I(t)} \alpha_{t'} \\
&\le 2c_2 c_S \varepsilon \frac{(2c_S)^{\ell-1}}{2c_S - 1} \|A_{tJ}\| \le c_2 \varepsilon \frac{(2c_S)^{\ell}}{2c_S - 1} \|A_{tJ}\|
\end{aligned}
$$

together with $c_2 = c_1(1 + c_S)$ proves the assertion. $\qquad\square$

Similar results as for the row clusters t can be obtained for column clusters $s \in T_J$ and $e \in \mathcal{E}(s)$ provided assumptions analogous to Assumption 1 are made. In particular, this defines clusters $\sigma_s \subset s$ and $\tau_s^e \subset F_e'(s)$, $|\tau_s^e| = |\sigma_s| = \min\{k_\varepsilon, |s|\}$, where

$$F_e'(s) := \bigcup \{t \in T_I : \exists \hat{s} \supset s \text{ such that } t \times \hat{s} \in P_{\text{high}} \text{ and } X_t \subset \mathcal{D}_e(Y_s)\}.$$

For $s' \subset s$ we define

$$C_e^{s's} := (A_{\tau_s^e \sigma_s})^{-1} A_{\tau_s^e s'}.$$

For leaf clusters $s \in \mathcal{L}(\hat{T}_J)$ and $e \in \mathcal{E}(s)$ we set $V_e(s) = (C_e^{ss})^H$. Assuming that matrices $V_e(s')$ have already been constructed for the sons $s' \in S_J(s)$ of $s \in \hat{T}_J \setminus \mathcal{L}(\hat{T}_J)$ and $e \in \mathcal{E}(s')$, we define for $e \in \mathcal{E}(s)$

$$V_e(s)|_{s'} := V_{e'}(s')(C_e^{\sigma_{s'} s})^H, \quad s' \in S_J(s),$$

where $e' \in \mathcal{E}(s')$ is defined before Lemma 6. Due to the analogy with Lemma 9 and Theorem 1, we omit the proofs of the following error estimates.

Lemma 11 *There is $c_4 = c_4(\varepsilon) > 0$ such that for all $s \in T_J$ and $e \in \mathcal{E}(s)$ it holds that*

$$\|A_{ts} - A_{t\sigma_s}(A_{\tau_t^e \sigma_s})^{-1} A_{\tau_s^e s}\| \leq c_4 \varepsilon \|A_{Is}\|$$

for all $t \subset I$ satisfying $X_t \subset \mathcal{F}_e(Y_s)$.

Theorem 2 *Let $s \in \hat{T}_J$ and let $\ell = L(\hat{T}_s)$ denote the depth of the cluster tree $\hat{T}_s$. Then there is $c_5 = c_5(\varepsilon) > 0$ such that for all $e \in \mathcal{E}(s)$*

$$\|A_{t\sigma_s}[C_e^{ss} - V_e(s)^H]\| \leq c_5 \varepsilon \|A_{Is}\|, \quad t \subset F_e'(s).$$

Using the previously constructed bases $\mathcal{U}$ and $\mathcal{V}$, we employ $S(t, s) := A_{\tau_t \sigma_s}$ in (26) for each block A_{ts}, $t \times s \in P_{\text{high}}$. In the following theorem, the accuracy of the nested approximation based on the matrix entries of A is estimated.

Theorem 3 *Let all previous assumptions be valid and $t \times s \in P_{\text{high}}$. Then there is $e \in \mathcal{E}(t)$ such that $s \subset F_e(t)$ and the approximation error is bounded by*

$$\|A_{ts} - U_e(t)S(t, s)V_{-e}(s)^H\| \leq c_S(c_2 + c_3)\varepsilon \|A_{tJ}\| + (c_4 + c_5\|U_e(t)\|_\infty)\varepsilon \|A_{Is}\| \tag{33}$$

with c_S, c_2, c_3, c_4, c_5 depending on ε defined above.

Proof From $s \subset F_e(t)$ it follows that $t \subset F_{-e}'(s)$. We have that

$$A_{ts} - B_e^{tt}S(t, s)C_{-e}^{ss} = A_{ts} - A_{t\sigma_s}C_{-e}^{ss} + [A_{t\sigma_s} - B_e^{tt}A_{\tau_t \sigma_s}]C_{-e}^{ss}.$$

From Lemma 9 it follows that $\|A_{t\sigma_s} - B_e^{tt}A_{\tau_t \sigma_s}\| \leq c_2 \varepsilon \|A_{tJ}\|$, and from Lemma 11 we have that $\|A_{t\sigma_s} - A_{t\sigma_s}C_{-e}^{ss}\| \leq c_4 \varepsilon \|A_{Is}\|$. Therefore,

$$\|A_{ts} - B_e^{tt}S(t, s)C_{-e}^{ss}\| \leq \varepsilon(c_4\|A_{Is}\| + c_2\|A_{tJ}\|\|C_{-e}^{ss}\|_1).$$

Furthermore, Theorems 1 and 2 yield

$$\begin{aligned}
&\|U_e(t)S(t, s)V_{-e}(s)^H - B_e^{tt}S(t, s)C_{-e}^{ss}\| \\
&\leq \|U_e(t)\|_\infty\|S(t, s)[V_{-e}(s)^H - C_{-e}^{ss}]\| \\
&\quad + \|[U_e(t) - B_e^{tt}]S(t, s)\|\|C_{-e}^{ss}\|_1 \\
&\leq c_5\varepsilon\|U_e(t)\|_\infty\|A_{Is}\| + c_3 c_S \varepsilon\|A_{tJ}\|,
\end{aligned}$$

which proves the assertion. $\qquad\square$

Remark 1 Furthermore, the $\| \cdot \|_\infty$-norm of $U_e(t)$ can be expected to be close to $\|B_e^{tt}\|$ since $U_e(t)$ approximates B_e^{tt}; cf. Theorem 1. The actual values of $\|U_e(t)\|_\infty$ can easily be evaluated during computation. Hence, it easy to check the quality of the nested approximation via (33). The numerical results in the next section show that these terms are bounded in practice and do not destroy the convergence of the method.

6 Numerical results

We consider the sound soft/hard scattering problem (1) imposing either the Dirichlet datum $u = u_D$ (sound soft) or the Neumann trace $\partial_\nu u = v_N$ (sound hard) on the boundary Γ. We denote $\mathfrak{V}$ the single and $\mathfrak{K}$ the double-layer-operator with the kernels $S(x - y)$ and $\partial_{\nu_y} S(x - y)$, respectively. Using Green's representation formula $u = \mathfrak{K}u - \mathfrak{V}(\partial_\nu u)$ in Ω^c and the jump relations, we end up with the integral equation

$$\mathfrak{V}(\partial_\nu u) = \left(\mathfrak{K} - \frac{1}{2}\mathfrak{I}\right) u \quad \text{on } \Gamma, \tag{34}$$

which can be solved either for the unknown $\partial_\nu u|_\Gamma$ in the case of a Dirichlet problem or for the unknown $u|_\Gamma$ in the Neumann case. The Brakhage–Werner ansatz

$$u = \mathfrak{K}\phi - i\beta\mathfrak{V}\phi \quad \text{in } \Omega^c \tag{35}$$

with an arbitrary coupling parameter $\beta > 0$ uses an unknown density function ϕ that satisfies the integral equation

$$\left(\frac{1}{2}\mathfrak{I} + \mathfrak{K} - i\beta\mathfrak{V}\right)\phi = u_D. \tag{36}$$

In either case, Galerkin discretization of the integral equations leads to a linear system with matrices of the form (2) that can be approximated by directional $\mathcal{H}^2$-matrices. The solution can then be obtained via GMRES using the matrix-vector multiplication from Sect. 4.1, which we proved to have logarithmic-linear complexity.

6.1 Approximation of $\mathfrak{V}$

As a first step, we validate the logarithmic-linear complexity of the directional $\mathcal{H}^2$-matrices (labeled dir$\mathcal{H}^2$-ACA). Moreover, we compare our new approach with the standard $\mathcal{H}$-matrix approximation via ACA (labeled $\mathcal{H}$-ACA) and the $\mathcal{H}^2$-matrix approach from [8] (labeled $\mathcal{H}^2$-ACA). In this section, we focus on the approximation of the single-layer operator $\mathfrak{V}$. Since we assumed κh to be constant, we increase κ with growing number of degrees of freedom N. By "acc." we label the relative error to the full matrix in the Frobenius norm.

Table 1 shows the memory consumption of the approximation of the discretization of $\mathfrak{V}$ with piecewise constant ansatz and test functions on the prolate spheroid, i.e. an ellipsoid with ten times the extension in x-direction. The accuracy for $N = 1,905,242$
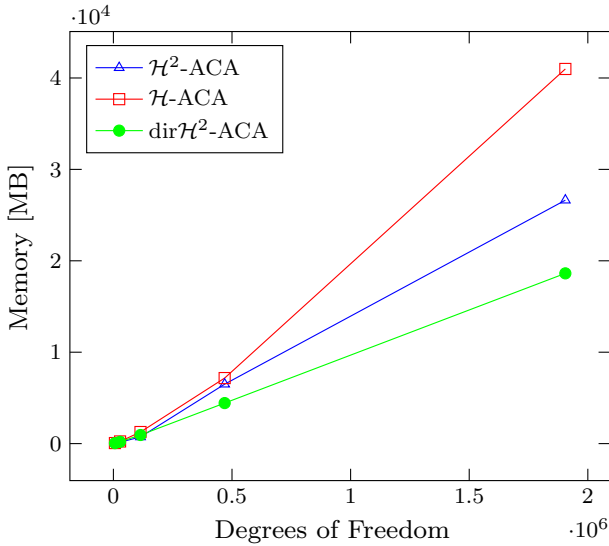
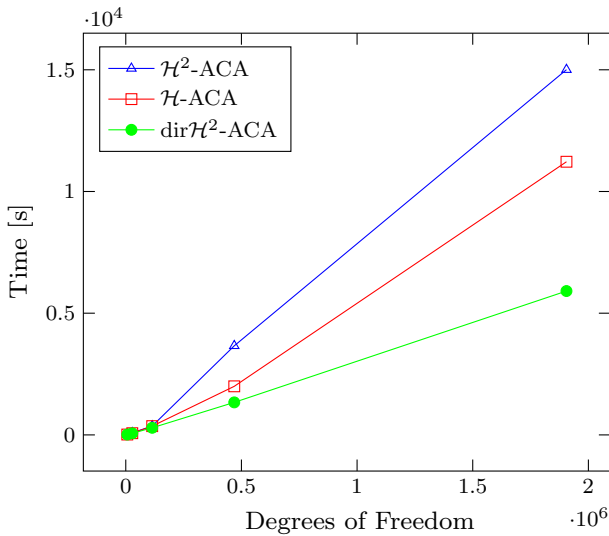**Fig. 3** Memory consumption in MB on prolate spheroid

Fig. 4 Construction time on prolate spheroid

was not computed, because the calculation of the difference to a full matrix in Frobenius norm takes $\mathcal{O}(N^2)$ time.

For small N, the three methods have about the same performance. This is due to fact that directional $\mathcal{H}^2$-matrices adapt to the wave number and fade to usual $\mathcal{H}$-matrices for low frequencies. Nevertheless, $\mathcal{H}^2$-ACA seems to be superior with respect to memory consumption. The latter approach, however, is not convenient for Helmholtz-type

Table 1 Prolate spheroid: $\kappa h \sim 0.15$, $\eta_{high} = 5$

		N	6,496	28,108	114,258	469,010	1,905,242
		κ	16	32	64	128	256
$\mathcal{H}$-ACA	Mem. (MB)		61	236	1,267	7,179	40,987
	Mem./N (KB)		9.8	8.8	11.6	16.1	22.6
	Time (s)		15	70	360	1,996	11,222
	Time/N (ms)		2.3	2.5	3.1	4.3	5.9
	Acc.		7.4_{-5}	6.7_{-5}	7.5_{-5}	1.1_{-4}	–
$\mathcal{H}^2$-ACA	Mem. (MB)		59	154	730	6,501	26,632
	Mem./N (KB)		9.5	5.7	6.7	14.5	14.6
	Time (s)		16	84	364	3,660	15,005
	Time/N (ms)		2.5	2.9	3.2	7.8	7.9
	Acc.		9.8_{-5}	1.3_{-4}	9.9_{-5}	4.5_{-4}	–
dir$\mathcal{H}^2$-ACA	Mem. (MB)		48	197	942	4,434	18,625
	Mem./N (KB)		7.8	7.3	8.6	9.9	10.2
	Time (s)		14	65	292	1,336	5,910
	Time/N (ms)		2.1	2.3	2.6	2.8	3.1
	Acc.		5.6_{-4}	1.0_{-4}	1.5_{-4}	1.4_{-4}	–

Table 2 dir$\mathcal{H}^2$-ACA: L^2-error of $\tilde{u}_D$ with $\kappa h = 0.19$

N	642	2,562	10,242	40,962	163,842	655,362
κ	2	4	8	16	32	64
Time (s) V	2	12	94	755	4,571	27,247
Mem. (MB) V	2	16	105	481	2,145	9,409
KB/N V	3.8	6.6	10.7	12.3	13.7	15.1
Time (s) K	2	13	100	811	4,921	29,539
Mem. (MB) K	2	16	105	471	2,081	9,187
L^2-error	2.6_{-3}	2.1_{-3}	2.0_{-3}	1.9_{-3}	1.9_{-3}	2.0_{-3}

kernel functions as the running time increases for higher wave numbers. The reason for this behaviour is that the blockwise rank cannot be guaranteed to be constant and thus, the number of apriori pivots must be increased accordingly in order to achieve a prescribed accuracy. The gain of dir$\mathcal{H}^2$-ACA in terms of both time and memory becomes visible for larger N (see Figs. 3, 4).

6.2 Neumann problem

We consider the sound hard scattering problem and use piecewise linear ansatz and test functions for the Galerkin discretization of (34). The approximate Dirichlet datum $\tilde{u}_D \approx u|_\Gamma$ is obtained from solving (34) with approximations ($\varepsilon = 10^{-4}$) of the discrete operators V and K of $\mathfrak{V}$ and $\mathfrak{K}$. We use the Neumann datum $v_N := \partial_\nu S(\cdot - x_0)$

Table 3 $\mathcal{H}$-ACA: L^2-error of $\tilde{u}_D$ with $\kappa h = 0.19$

N	642	2,562	10,242	40,962	163,842	655,362
κ	2	4	8	16	32	64
Time (s) V	2	11	64	464	6,890	–
Mem. (MB) V	2	13	80	491	3,150	–
KB/N V	3.1	5.2	8.2	12.6	20.2	–
Time (s) K	3	14	77	520	7,050	–
Mem. (MB) K	2	12	78	478	3,073	–
L^2-error	2.7_{-3}	2.0_{-3}	1.9_{-3}	2.0_{-3}	2.7_{-3}	–

Table 4 dir$\mathcal{H}^2$-ACA: L^2-error of $\tilde{u}_D$ with fixed $\kappa = 6$

N	642	2,562	10,242	40,962	163,842
Time (s) V	6	42	75	286	1,613
Mem. (MB) V	3	15	66	345	1,613
KB/N V	5.1	6.2	6.7	8.8	11.0
Time (s) K	7	47	84	317	1,734
Mem. (MB) K	3	15	64	334	1,669
L^2-error	1.8_{-2}	4.5_{-3}	1.2_{-3}	3.9_{-4}	2.3_{-4}

Table 5 $\mathcal{H}$-ACA: L^2-error of $\tilde{u}_D$ with fixed $\kappa = 6$

N	642	2,562	10,242	40,962	163,842
Time (s) V	2	13	69	379	2,771
Mem. (MB) V	3	17	90	454	2,167
KB/N V	4.5	6.9	9.2	11.6	13.9
Time (s) K	3	17	83	439	3,024
Mem. (MB) K	3	17	89	446	2,122
L^2-error	1.8_{-2}	4.4_{-3}	1.1_{-3}	3.0_{-4}	1.2_{-4}

with an interior point $x_0 \in \Omega$. In this case, we are able to compute the L^2-error $\|\tilde{u}_D - u\|_{L^2(\Gamma)}/\|u\|_{L^2(\Gamma)}$ to the exact Dirichlet trace given by $u|_\Gamma = S(\cdot - x_0)$. Tables 2 and 4 show the behaviour of the error on the sphere with radius 1 for fixed κh and fixed κ, respectively. Furthermore, the CPU time and the memory consumption required by the approximations to V and K is shown and can be seen to behave logarithmic-linear for both fixed and growing wave numbers. As a reference, we made the same computations also using $\mathcal{H}$-matrices. The corresponding results are shown in Tables 3 and 5. As before, the advantage of dir$\mathcal{H}^2$-ACA becomes visible for a growing number of waves. It is remarkable, however, that even in the fixed frequency case the directional approach outperforms $\mathcal{H}$-ACA in terms of memory and computation time for larger degrees of freedom. For smaller degrees of freedom the performance of $\mathcal{H}$-ACA is only slightly better.
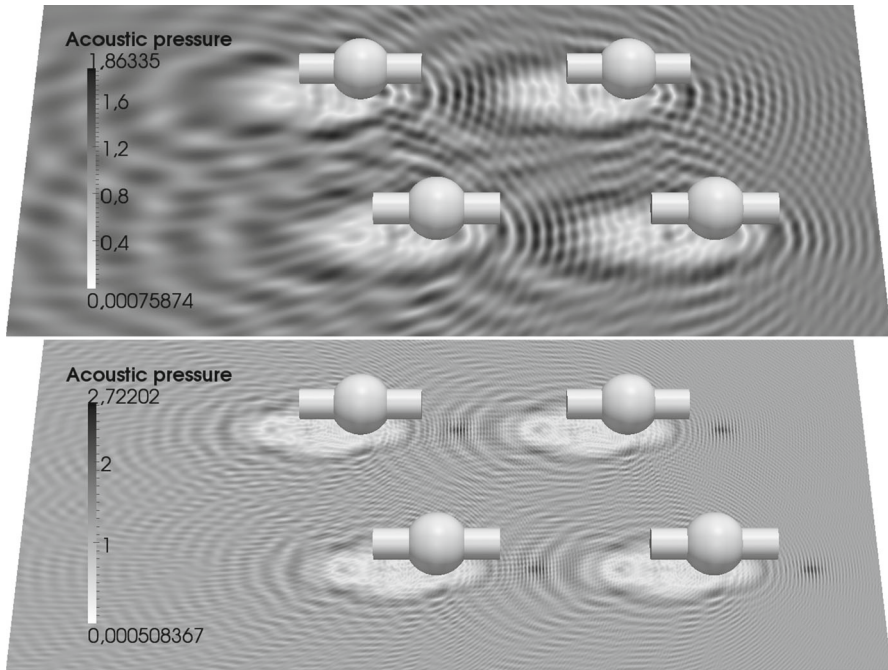
Fig. 5 Pressure field $|u_s + u_i|$ for $\kappa = 10$ and $\kappa = 40$

6.3 Visualization of Dirichlet problem with Brakhage–Werner

We consider the sound soft scattering problem, i.e. we seek a solution $u = u_i + u_s$ of the Helmholtz equation (1), where $u_i(x) := \exp(i\kappa(x, e))$ with $e = (1, 0, 1)^T/\sqrt{2}$ denotes the incident acoustic wave and u_s is the unknown scattered field. The incident wave is reflected on a sound soft obstacle Ω, which is described by the Dirichlet condition $u_s|_\Gamma = -u_i|_\Gamma$.

The obstacle is composed of 4 cylindrical spheres with 507,904 panels and 253,960 vertices. We solve (36) for the unknown density ϕ with piecewise linear ansatz and test functions. Following [22], we use the coupling parameter $\beta = \kappa/2$. In a second step, we evaluate the potential (35) on a uniform discretization of a screen behind the obstacle in order to compute the scattered wave u_s. Figure 5 shows the resulting pressure field of the total wave $|u_i + u_s|$ for $\kappa = 10$ and $\kappa = 40$.

References

1. Alpert, B.K., Beylkin, G., Coifman, R., Rokhlin, V.: Wavelet-like bases for the fast solution of second-kind integral equations. SIAM J. Sci. Comput. **14**, 159–184 (1993)
2. Amini, S., Profit, A.: Multi-level fast multipole solution of the scattering problem. Eng. Anal. Bound. Elements **27**(5), 547–654 (2003)
3. Anderson, C.R.: An implementation of the fast multipole method without multipoles. SIAM J. Sci. Stat. Comput. **13**(4), 923–947 (1992)

4. Banjai, L., Hackbusch, W.: Hierarchical matrix techniques for low- and high-frequency Helmholtz problems. IMA J. Numer. Anal. **28**(1), 46–79 (2008). doi:10.1093/imanum/drm001
5. Barnes, J., Hut, P.: A hierarchical $\mathcal{O}(\mathcal{N} \log \mathcal{N})$ force calculation algorithm. Nature **324**, 446–449 (1986)
6. Bebendorf, M.: Approximation of boundary element matrices. Numer. Math. **86**(4), 565–589 (2000)
7. Bebendorf, M.: Hierarchical matrices: a means to efficiently solve elliptic boundary value problems. In: Lecture Notes in Computational Science and Engineering (LNCSE), vol. 63. Springer, Berlin (2008). ISBN 978-3-540-77146-3
8. Bebendorf, M., Venn, R.: Constructing nested bases approximations from the entries of non-local operators. Numer. Math. **121**(4), 609–635 (2012)
9. Börm, S.: $\mathcal{H}^2$-matrix arithmetics in linear complexity. Computing **77**(1), 1–28 (2006)
10. Börm, S., Löhndorf, M., Melenk, J.M.: Approximation of integral operators by variable-order interpolation. Numer. Math. **99**(4), 605–643 (2005)
11. Brakhage, H., Werner, P.: Über das Dirichletsche Außenraumproblem für die Helmholtzsche Schwingungsgleichung. Arch. Math. **16**, 325–329 (1965)
12. Brandt, A.: Multilevel computations of integral transforms and particle interactions with oscillatory kernels. Comput. Phys. Commun. **65**, 24–38 (1991)
13. Buffa, A., Hiptmair, R.: A coercive combined field integral equation for electromagnetic scattering. SIAM J. Numer. Anal. **42**(2), 621–640 (2003)
14. Burton, A.J., Miller, G.F.: The application of integral equation methods to the numerical solution of boundary value problems. Proc. R. Soc. Lond. **A232**, 201–210 (1971)
15. Candès, E., Demanet, L., Ying, L.: A fast butterfly algorithm for the computation of Fourier integral operators. Multiscale Model. Simul. **7**(4), 1727–1750 (2009). doi:10.1137/080734339
16. Chandler-Wilde, S.N., Graham, I.G., Langdon, S., Spence, E.A.: Numerical-asymptotic boundary integral methods in high-frequency acoustic scattering. Acta Numerica **21**, 89–305 (2012)
17. Cheng, H., Crutchfield, W.Y., Gimbutas, Z., Greengard, L.F., Ethridge, J.F., Huang, J., Rokhlin, V., Yarvin, N., Zhao, J.: A wideband fast multipole method for the Helmholtz equation in three dimensions. J. Comput. Phys. **216**(1), 300–325 (2006). doi:10.1016/j.jcp.2005.12.001
18. Darve, E.: The fast multipole method: numerical implementation. J. Comput. Phys. **160**(1), 195–240 (2000)
19. Engquist, B., Ying, L.: Fast directional multilevel algorithms for oscillatory kernels. SIAM J. Sci. Comput. **29**(4), 1710–1737 (2007, electronic)
20. Engquist, B., Ying, L.: A fast directional algorithm for high frequency acoustic scattering in two dimensions. Commun. Math. Sci. **7**(2), 327–345 (2009)
21. Francia, G.T.D.: Degrees of freedom of an image. J. Opt. Soc. Am. **59**(7), 799–803 (1969)
22. Giebermann, K.: Schnelle Summationsverfahren zur numerischen Lösung von Integralgleichungen für Streuprobleme im $\mathbb{R}^3$. Ph.D. thesis, Universität Karlsruhe (1997)
23. Grasedyck, L., Hackbusch, W.: Construction and arithmetics of $\mathcal{H}$-matrices. Computing **70**, 295–334 (2003)
24. Greengard, L.: The Rapid Evaluation of Potential Fields in Particle Systems. MIT Press, Cambridge (1988)
25. Greengard, L.F., Rokhlin, V.: A fast algorithm for particle simulations. J. Comput. Phys. **73**(2), 325–348 (1987)
26. Greengard, L.F., Rokhlin, V.: A new version of the fast multipole method for the Laplace equation in three dimensions. In: Acta Numerica, vol. 6, pp. 229–269. Cambridge University Press, Cambridge (1997)
27. Hackbusch, W.: A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: introduction to $\mathcal{H}$-matrices. Computing **62**(2), 89–108 (1999)
28. Hackbusch, W., Börm, S.: Data-sparse approximation by adaptive $\mathcal{H}^2$-matrices. Computing **69**(1), 1–35 (2002)
29. Hackbusch, W., Khoromskij, B.N.: A sparse $\mathcal{H}$-matrix arithmetic: general complexity estimates. J. Comput. Appl. Math. **125**(1–2), 479–501 (2000). Numerical analysis 2000, vol. VI, Ordinary differential equations and integral equations
30. Hackbusch, W., Khoromskij, B.N.: A sparse $\mathcal{H}$-matrix arithmetic. Part II: application to multidimensional problems. Computing **64**(1), 21–47 (2000)
31. Hackbusch, W., Khoromskij, B.N., Sauter, S.A.: On $\mathcal{H}^2$-matrices. In: Bungartz, H.J., Hoppe, R.H.W., Zenger, C. (eds.) Lectures on Applied Mathematics, pp. 9–29. Springer, Berlin (2000)

32. Hackbusch, W., Nowak, Z.P.: On the fast matrix multiplication in the boundary element method by panel clustering. Numer. Math. **54**(4), 463–491 (1989)
33. Messner, M., Schanz, M., Darve, E.: Fast directional multilevel summation for oscillatory kernels based on Chebyshev interpolation. J. Comput. Phys. **231**, 1175–1196 (2012)
34. Michielssen, E., Boag, A.: A multilevel matrix decomposition for analyzing scattering from large structures. IEEE Trans. Antennas Propag. **44**, 1086–1093 (1996)
35. Rokhlin, V.: Rapid solution of integral equations of classical potential theory. J. Comput. Phys. **60**(2), 187–207 (1985)
36. Rokhlin, V.: Rapid solution of integral equations of scattering theory in two dimensions. J. Comput. Phys. **86**(2), 414–439 (1990)
37. Rokhlin, V.: Diagonal forms of translation operators for the Helmholtz equation in three dimensions. Appl. Comput. Harmon. Anal. **1**(1), 82–93 (1993)
38. Sauer, T., Xu, Y.: On multivariate Lagrange interpolation. Math. Comput. **64**(211), 1147–1170 (1995)
39. Tyrtyshnikov, E.E.: Mosaic-skeleton approximations. Calcolo **33**(1–2), 47–57 (1998). Toeplitz matrices: structures, algorithms and applications (Cortona, 1996)
40. Ying, L., Biros, G., Zorin, D.: A kernel-independent adaptive fast multipole algorithm in two and three dimensions. J. Comput. Phys. **196**(2), 591–626 (2004)