

**Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig**

**Hierarchical Kronecker tensor-product
approximation to a class of nonlocal
operators in high dimensions**

(revised version: July 2004)

by

Wolfgang Hackbusch and Boris N. Khoromskij

Preprint no.: 16

2004



Hierarchical Kronecker Tensor-Product Approximation to a Class of Nonlocal Operators in High Dimensions

Wolfgang Hackbusch and Boris N. Khoromskij
Max-Planck-Institute for Mathematics in the Sciences,
Inselstr. 22-26, D-04103 Leipzig, Germany.
{wh, bokh}@mis.mpg.de

July 22, 2004

Abstract

The class of \mathcal{H} -matrices allows an approximate matrix arithmetic with almost linear complexity. The combination of the hierarchical and tensor-product format offers the opportunity for efficient data-sparse representations of integral operators and the inverse of elliptic operators in higher dimensions (cf. [19], [7]). In the present paper, we apply the \mathcal{H} -matrix techniques combined with the Kronecker tensor-product approximation to represent integral operators as well as certain functions $\mathcal{F}(A)$ of a discrete elliptic operator A in a hypercube $(0, 1)^d \in \mathbb{R}^d$ in the case of a high spatial dimension d . In particular, we approximate the functions A^{-1} and $\text{sign}(A)$ of a finite difference discretisations $A \in \mathbb{R}^{N \times N}$ with rather general location of the spectrum. The asymptotic complexity of our data-sparse representations can be estimated by $\mathcal{O}(n^p \log^q n)$, $p = 1, 2$, with q independent of d , where $n = N^{1/d}$ is the dimension of the discrete problem in *one* space direction.

AMS Subject Classification: 65F50, 65F30, 46B28, 47A80

Key Words: hierarchical matrices, Kronecker tensor-product, high spatial dimension, Sinc-interpolation, Sinc-quadrature

1 Introduction and Results

In the wide range of applications one requires tractable approximations to certain multi-dimensional nonlocal operators posed in \mathbb{R}^d , $d \geq 2$. In particular, we are interested in the efficient numerical representation to the following classes of nonlocal operators: multi-dimensional integral operators, solution operators of elliptic, parabolic and hyperbolic boundary value problems, Lyapunov and Riccati solution operators in control theory, spectral projection operator associated with matrix *sign* function, density matrix ansatz for solving the Schrödinger and Hartree-Fock equations. As soon as the computational issues are concerned, we come to the challenging problem of accurate representation to the large fully populated matrices (generally given only implicitly) in special data-sparse formats.

The class of hierarchical (\mathcal{H}) matrices allows an approximate matrix arithmetic with almost linear complexity [13]-[17], [10]. An \mathcal{H} -matrix approximation to the class of operator-valued functions of an elliptic operator was developed in [4]-[6], [11]. In multidimensional perspective, even approximations with linear complexity $\mathcal{O}(n^d)$ might not be satisfactory. To relax the “curse of dimensionality” one can try to represent the corresponding data (matrices and vectors) in the tensor-product form with the overall complexity $\mathcal{O}(dn^p \log^q n)$ with $p, q \geq 1$ independent of d . To optimise the exponential p , we can utilise the hierarchical format for each low dimensional component, which further reduces the cost to $\mathcal{O}(dn \log^q n)$.

In the previous paper [7] we apply the \mathcal{H} -matrix techniques combined with the Kronecker tensor-product approximation (cf. [19, 24]) to represent the inverse of a discrete elliptic operator in a hypercube $(0, 1)^d \in \mathbb{R}^d$ in the case of a high spatial dimension d . We recall that the *hierarchical Kronecker tensor-product* (HKT) approximation of a matrix is defined as follows. Given a matrix $A \in \mathbb{C}^{N \times N}$ of dimension $N = n^d$, we try to

approximate A by a matrix $A_{(r)}$ of the form

$$A_{(r)} = \sum_{k=1}^r V_k^1 \otimes \cdots \otimes V_k^d \approx A, \quad (1.1)$$

where the V_k^ℓ are $n \times n$ -matrices and \otimes denotes the Kronecker product operation. Now the crucial parameter is r , called the *Kronecker rank* (cf. [19]). Each elementary Kronecker factor V_k^ℓ is supposed to be represented in the \mathcal{H} -matrix form. In this data-sparse format, we also represent the operator exponential as well as fractional powers of an elliptic operator. The complexity of the HKT approximation can be estimated by $\mathcal{O}(dn \log^q n)$, where q is some fixed constant independent of d . A computational scheme for a low Kronecker-rank approximation to the solution of a tensor system with tensor right-hand side was presented in [9]. Approximations of high order tensors are, e.g., discussed in [27]. Methods based on sparse grid finite elements or wavelets have been successfully applied [3, 21, 12, 23].

The HKT approximation to integral operators posed in \mathbb{R}^d is well understood, since it can be reduced to a separable approximation of the explicitly given kernel function together with an \mathcal{H} -matrix representation of the low-dimensional components (cf. [19] related to the case $d = 2$). In Section 2, we address this topic in the case of rather general shift-invariant kernel functions with $d \geq 2$.

Next we consider the operator-valued functions of an elliptic operator. In the following we use the notation $\mathcal{A}, \mathcal{B}, \dots$ for operators and A, B, \dots for matrices. The basic assumption on the elliptic operator \mathcal{A} given in the form

$$\mathcal{A} = - \sum_{j=1}^d \frac{\partial}{\partial x_j} a_j(x) \frac{\partial}{\partial x_j} + \sum_{j=1}^d b_j(x) \frac{\partial}{\partial x_j} + c(x), \quad x = (x_1, \dots, x_d) \in (0, 1)^d,$$

and the existence of a splitting

$$\mathcal{A} = \sum_{j=1}^d \mathcal{A}_j, \quad \mathcal{A}_k \mathcal{A}_m = \mathcal{A}_m \mathcal{A}_k \quad (1 \leq k, m \leq d), \quad (1.2)$$

with mutually commutative differential operators \mathcal{A}_j , acting on the variable x_j (cf. [7]). In our particular case, to ensure (1.2), we further assume

$$a_j(x) = a_j(x_j), \quad b_j(x) = b_j(x_j) \quad \text{and} \quad c(x) = c_1(x_1) + \dots + c_d(x_d).$$

The main idea to prove an HKT approximation to certain operator-valued functions of \mathcal{A} is based on making use of an integral representation to the operator involving the exponential of \mathcal{A} . For example, a negative fractional power of \mathcal{A} can be represented by (cf. [7])

$$\mathcal{A}^{-\sigma-1} = \frac{1}{\Gamma(\sigma+1)} \int_0^\infty t^\sigma e^{-t\mathcal{A}} dt, \quad \sigma > -1, \quad (1.3)$$

provided the integral is existing. Having at hand such a representation, first, we make use of the fundamental property of the exponential function, namely,

$$\exp(-t\mathcal{A}) = \prod_{j=1}^d \exp(-t\mathcal{A}_j), \quad (1.4)$$

and second, we apply an exponentially convergent quadrature rule to represent the integral (1.3) by a sum involving only factorised expressions

$$\mathcal{A}^{-\sigma-1} \approx \sum_{k=-M}^M c_k t_k^\sigma \prod_{j=1}^d \exp(-t_k \mathcal{A}_j), \quad t_k \in \mathbb{R}.$$

To derive the tensor-product representation, we consider finite difference (FD) discretisations (three-point stencil in each variable) on a uniform tensor-product grid in \mathbb{R}^d with n degrees of freedom in each spatial

direction. The discretisation matrix has the form $A = \sum_{j=1}^d A_j$ with $A, A_j \in \mathbb{R}^{N \times N}$, $N = n^d$, where the matrices A_j are of the tensor-product form

$$A_1 = V^1 \otimes I \otimes \dots \otimes I, \quad A_2 = I \otimes V^2 \otimes \dots \otimes I, \quad \dots, \quad A_d = I \otimes \dots \otimes I \otimes V^d \quad (1.5)$$

with $V^j = \text{tridiag}\{a_j, b_j, c_j\} \in \mathbb{R}^{n \times n}$, $j = 1, \dots, d$, and the identity $I \in \mathbb{R}^{n \times n}$. We require $\Re \text{sp}(A) \subset [a, \infty)$ with $a > 0$, where $\text{sp}(A)$ denotes the spectrum of A . Then we derive the tensor approximation

$$A^{-\sigma-1} \approx \sum_{k=-M}^M c_k t_k^\sigma \bigotimes_{j=1}^d \exp(-t_k V^j) =: A_{(r)} \quad \text{with } r = 2M + 1 \text{ (cf. (1.1))} \quad (1.6)$$

providing the exponential convergence

$$\|A^{-\sigma-1} - A_{(r)}\| \leq C e^{-s\sqrt{M}}$$

with constants C, s not depending on M . Finally, given tolerance $\varepsilon = \mathcal{O}(n^{-\beta})$, $\beta > 0$, to obtain the desired complexity bound $\mathcal{O}(dn \log^q n)$, we apply the \mathcal{H} -matrix approximation to each individual exponent $\exp(-t_k V^j)$ that manifests the linear-logarithmic cost $\mathcal{O}(n \log^q n)$ in n (cf. [4, 7] concerning the existence and construction of the corresponding approximation). Numerical examples on the HKT approximation to the discrete Laplacian inverse are presented in tables at the end of §A.3.

Remark 1.1 *Note that with the choice $\mathcal{A} = -\Delta$, the representation (1.6) would be of particular interest in the cases $\sigma = -1/2$ (preconditioning for the Laplace-Beltrami operator $(-\Delta)^{1/2}$, and for the hypersingular integral operator, e.g., in BEM applications), $\sigma = 0$ (inverse Laplacian), $\sigma = 1$ (inverse biharmonic operator).*

When we study the HKT approximation applied either to a more general class of elliptic operators (say, operators \mathcal{A} with mixed derivatives or indefinite resolvent operators $(zI - \mathcal{A})^{-1}$, $z \in \mathbb{C}$, or to the more general class of operator-valued functions $\mathcal{F}(\mathcal{A})$ (e.g., for $\mathcal{F}(\mathcal{A}) = \text{sign}(\mathcal{A})$) then we find that the strong positiveness and commutativity property required above may fail. The goal of this paper is to present novel more general HKT formats and new approximation techniques which allow to extend the results in [7, 19] to a wider range of practically interesting applications which include, in particular, the functions of operators with mixed derivatives or of indefinite operators. In this paper we focus on the case of matrix-valued functions $\mathcal{F}(A)$, where A represents the discrete elliptic operator though the results can be applied to a more general class of matrices.

The main ideas behind our approach are related to the following issues:

- (a) new representations for the matrices A^{-1} , $(zI - A)^{-1}$, $(zI - A)^{-1} \pm (\bar{z}I - A)^{-1}$ by means of matrix exponentials e^{-tB^2} and e^{-tBB^*} with the corresponding choice of B (cf. (1.7) and (1.8));
- (b) approximation to matrices of the type $\exp(A_k A_m)$ with $k \neq m$ (cf. (1.5));
- (c) new integral representations for $\mathcal{F}(A) = \text{sign}(A)$, which allows exponentially convergent quadratures.

Concerning item (a), we propose the representation of A^{-1} (cf. (1.3) with $\sigma = 0$ and with A substituted by A^2) by

$$A^{-1} = A \int_0^\infty e^{-tA^2} dt, \quad (1.7)$$

which already allows to treat invertible operators with rather general location of the spectrum (say, the Helmholtz operator $\mathcal{A} = -\Delta - \kappa^2$, $\kappa \in \mathbb{R}$). The construction for a sum of resolvents $(zI - A)^{-1} \pm (\bar{z}I - A)^{-1}$ is reduced to the case (1.7). In the general case (e.g., for the resolvent $(zI - A)^{-1}$), we shall use the modification

$$A^{-1} = A^*(AA^*)^{-1} = A^* \int_0^\infty e^{-tAA^*} dt. \quad (1.8)$$

The core of our method is based on efficient quadratures for the integrals in (1.7) and (1.8) combined with a data-sparse representation of the matrices of the form $\exp(A_k A_m)$ with commutative tensor products A_k, A_m .

Remark 1.2 The representation (1.8) is in particular successful for solving least squares problems given by the normal equation of the form $A^\top Au = A^\top f$.

The construction of an HKT approximation to the inverse matrix and, in particular, to the resolvent family $(zI - A)^{-1}$, $z \in \mathbb{C}$, allows to approximate a rather general class of matrix-valued functions, which can be represented by the Dunford-Cauchy integral

$$\mathcal{F}(A) := \frac{1}{2\pi i} \int_{\Gamma} \mathcal{F}(z)(zI - A)^{-1} dz, \quad (1.9)$$

where Γ is a curve containing the spectrum of A . Usually, the curve Γ is chosen symmetrically with respect to the real axis. We assume that the integral in (1.9) is approximated by a proper quadrature formula

$$\mathcal{F}_M(A) := \sum_{k=-M}^M c_k \mathcal{F}(z_k)(z_k I - A)^{-1} \approx \mathcal{F}(A) \quad (1.10)$$

(e.g., Sinc quadrature or Gauss-Lobatto quadrature), where the quadrature points are located symmetrically with respect to the real axis, i.e., $z_k = \bar{z}_{-k}$ and, moreover, $c_k = c_{-k}$, $k = 1, \dots, M$. In general, each term in (1.10) fails to satisfy the assumption $\Re \lambda > 0$ for all $\lambda \in sp(z_k I - A)$. However, if we combine two terms corresponding to the indices k and $-k$, then the result can be represented in the HKT format, and thus, the total sum in (1.10).

In the general case, each term in (1.1) will be amplified by an extra factor S_k in $\mathbb{R}^{N \times N}$. For this purpose, we introduce the generalised tensor-product matrix format

$$A_{(r)} = \sum_{k=1}^r S_k \cdot (V_k^1 \times \dots \times V_k^d) \approx A \quad (1.11)$$

with a matrix S_k having a special data-sparse representation of complexity $\mathcal{O}(n^p \log^d n)$ with $p \leq 2$ (cf. (3.11)). The format (1.11) will be also applied to the matrix-valued function $\mathcal{F}(A) = \text{sign}(A)$, based on an efficient quadrature for the integral representation

$$\text{sign}(A) = \frac{1}{c_f} \int_{\mathbb{R}_+} \frac{f(tA)}{t} dt$$

with certain functions f described in §4.1.

Our study of the previous representations applied to discrete elliptic operators of second order then naturally leads to the understanding of how to treat higher order operators, e.g., the biharmonic operator.

2 HKT Approximation of Integral Operators

Separable approximations to multi-variate functions is a principal ingredient in the HKT representation of integral operators in many spacial dimensions. Given the integral operator in $\Omega := [0, 1]^d \in \mathbb{R}^d$, $d \geq 2$,

$$(\mathcal{A}u)(x) := \int_{\Omega} g(x, y)u(y)dy, \quad x, y \in \Omega$$

with some shift-invariant kernel function $g(x, y) = g(|x - y|)$, which, therefore, can be represented in the form

$$g(x, y) = G(\zeta_1, \dots, \zeta_d) \equiv g(\sqrt{\zeta_1^2 + \dots + \zeta_d^2}),$$

where $\zeta_\ell = |x_\ell - y_\ell| \in [0, 1]$, $\ell = 1, \dots, d$. With some fixed $0 \leq \alpha_0 < 1$, we introduce the auxiliary function

$$F(\zeta_1, \dots, \zeta_d) := (\zeta_1 \dots \zeta_{d-1})^{\alpha_0} G(\zeta_1, \dots, \zeta_d).$$

Assumption 2.1 Suppose that a multi-variate function $F : \mathbb{R}^d \rightarrow \mathbb{R}$ can be approximated by a separable expansion

$$F_r(\zeta_1, \dots, \zeta_d) := \sum_{k=1}^r \Phi_k^1(\zeta_1) \dots \Phi_k^d(\zeta_d) \approx F, \quad (2.1)$$

where the set of functions $\{\Phi_k^\ell(\zeta_\ell)\}$ can be fixed or chosen adaptively.

Consider a Galerkin scheme by tensor-product test functions:

$$\phi^{\mathbf{i}}(x_1, \dots, x_d) = \phi_1^{i_1}(x_1) \cdot \dots \cdot \phi_d^{i_d}(x_d), \quad \mathbf{i} = (i_1, \dots, i_d), \quad i_\ell \in I_n := \{1, \dots, n\}, \quad \ell = 1, \dots, d.$$

Now we approximate the Galerkin stiffness matrix

$$A = \{(\mathcal{A}\phi^{\mathbf{i}}, \phi^{\mathbf{j}})_{L^2}\}_{\mathbf{i}, \mathbf{j} \in I_n^d} \in \mathbb{R}^{N \times N}, \quad N = n^d,$$

by a matrix $A_{(r)}$ of the form (1.1), where the V_k^ℓ are $n \times n$ matrices given by

$$V_k^\ell = \left\{ \int \zeta_\ell^{-\alpha_\ell} \Phi_k^\ell(\zeta_\ell) \phi_\ell^{i_\ell}(x_\ell) \phi_\ell^{j_\ell}(y_\ell) dx_\ell dy_\ell \right\}_{i_\ell, j_\ell=1}^n, \quad \ell = 1, \dots, d$$

with $\alpha_\ell = \alpha_0$, $\ell = 1, \dots, d-1$, and $\alpha_d = 0$. The next lemma shows that the error $\|A - A_{(r)}\|$ in commonly used norms is directly related to the error $\|F - F_r\|_\infty$ of the separable approximation (2.1) of F (see the discussion in [19]).

Lemma 2.2 *Let Assumption 2.1 be valid, then for any $\mathbf{i}, \mathbf{j} \in I_n^d$, we have*

$$|a_{\mathbf{i}, \mathbf{j}} - a_{\mathbf{i}, \mathbf{j}}^r| \leq \|F - F_r\|_\infty \prod_{\ell=1}^d \left\| \zeta_\ell^{-\alpha_\ell} \phi_\ell^{i_\ell} \phi_\ell^{j_\ell} \right\|_{L^1([0,1] \times [0,1])}.$$

Let us further assume that the function $g_{\ell,k}(u, v) := |u - v|^{-\alpha_\ell} \Phi_k^\ell(|u - v|)$, $\ell = 1, \dots, d$, $k = 1, \dots, r$, is asymptotically smooth in $(u, v) \in [0, 1]^2$. Then, for low order piecewise polynomial basis functions, V_k^ℓ can be approximated by a rank- m \mathcal{H} -matrix \tilde{V}_k^ℓ providing an error $\|V_k^\ell - \tilde{V}_k^\ell\| \leq C\eta^m$ with some $\eta < 1$.

Proof. By construction we obtain

$$\begin{aligned} |a_{\mathbf{i}, \mathbf{j}} - a_{\mathbf{i}, \mathbf{j}}^r| &= \left| \int_{\Omega \times \Omega} (F - F_r) \left(\prod_{\ell=1}^d \zeta_\ell^{-\alpha_\ell} \right) \phi^{\mathbf{i}}(x) \phi^{\mathbf{j}}(y) dx dy \right| \\ &\leq \|F - F_r\|_\infty \left\| \left(\prod_{\ell=1}^d \zeta_\ell^{-\alpha_\ell} \right) \phi^{\mathbf{i}}(x) \phi^{\mathbf{j}}(y) \right\|_{L^1(\Omega \times \Omega)}. \end{aligned}$$

Then the first assertion follows by inserting the tensor-product basis. To prove the second statement, we note that V_k^ℓ appears to be the exact Galerkin stiffness matrix for the integral operator with the kernel function $g_{\ell,k}(u, v)$. Then the result follows by the conventional theory of the \mathcal{H} -matrix approximation (cf. [10], [13]-[16]). \blacksquare

Note that due to Lemma 2.2, $\|A - A_{(r)}\|$ can be easily estimated in, say, the Frobenius, l_2 or l_∞ matrix norms.

Several methods of separable approximations to multi-variate functions are presented in Appendix. In the general case, Assumption 2.1 can be validated by using the tensor-product Sinc-interpolation (cf. §A.2.5), since the function $\Phi_k^\ell(|u - v|)$ can be proved to be asymptotically smooth. For the class of kernel functions approximated by the quadrature method the factor $\Phi_k^\ell(|u - v|)$ even appears to be globally smooth (cf. §A).

Lemma 2.3 *For both the tensor-product Sinc-interpolation and quadrature approximation methods, the function $g_{\ell,k}(u, v)$ from Lemma 2.2 is proved to be asymptotically smooth.*

Proof. In the first case we have $g_{\ell,k}(u, v) = |u - v|^{-\alpha_\ell} S(k, \mathfrak{h})(\phi^{-1}(|u - v|))$, where $S(k, \mathfrak{h})$ refers for the k -th Sinc function with step-size \mathfrak{h} , and $\phi^{-1}(x) = \operatorname{arsinh}(\operatorname{arcosh}(\frac{1}{x}))$ (cf. §A.2.4). Since the Sinc function $S(k, \mathfrak{h})(x)$, $x \in \mathbb{R}$, is holomorphic in x , and the factor $|u - v|^{-\alpha_\ell}$ is asymptotically smooth we conclude that $g_{\ell,k}(u, v)$ is also asymptotically smooth (see also discussion in [19]).

In the case of a quadrature method, we have the entire function $\Phi_k^\ell(|u - v|) = \exp(-t_k |u - v|^2)$, $t_k > 0$, which completes the proof. \blacksquare

Application of Lemmata 2.2, 2.3 proves the existence of a low Kronecker rank HKT approximation to the class of multi-dimensional integral operators.

In general, given a tolerance $\varepsilon > 0$, we have the bound $r = \mathcal{O}(|\log h \log \varepsilon \log \log \varepsilon|^{d-1})$, where h is the mesh parameter of the FE discretisation. However, in many practically interesting cases we obtain a dimensionally independent bound $r = \mathcal{O}(|\log h \log \varepsilon \log \log \varepsilon|)$ (see examples in the Appendix).

3 Approximation to A^{-1} for Indefinite Matrices

The HKT representation to the inverse of a discrete elliptic operator plays a central role when treating the case of general matrix-valued functions $\mathcal{F}(A)$, since typically an approximation of the indefinite matrix resolvent $(zI - A)^{-1}$ is needed (cf. (1.10)). In this case a representation like (1.3) is no longer true. To overcome this difficulty, we approximate the matrix resolvent making use of a sum of resolvents with conjugate parameters z and \bar{z} combined with representations (1.7). The alternative is a direct representation of $(zI - A)^{-1}$ by (1.8).

3.1 Using the Representations (1.7) and (1.8)

First, we consider the representation (1.7) provided that the integral exists. For the ease of presentation, we assume that A is diagonalisable, i.e., $A = TDT^{-1}$ with D real diagonal, and $sp(A^2) \subset [1, R]$ with some $R > 1$. Then one can use the efficient quadrature described in the Appendix (cf. Lemma A.8, Remark A.2), which yields the representation of the integral term in (1.7) by the exponential sum $\mathcal{F}_M(A)$:

$$\mathcal{F}(A) := A^{-2} = \int_0^\infty e^{-tA^2} dt \approx \mathcal{F}_M(A) := \sum_{k=-M}^M c_k e^{-t_k A^2}, \quad (3.1)$$

providing exponentially fast convergence $\|\mathcal{F}(A) - \mathcal{F}_M(A)\| \leq C e^{-\pi M / (\log R \log M)}$. Suppose that $\text{cond}(A^2) = n^\beta$, $\beta > 0$, then the approximation error $\varepsilon = n^{-\alpha}$, $\alpha > 0$, can be achieved with $M = \mathcal{O}(\log^q n)$, $q \geq 1$. Figures A.1, A.2 illustrate the efficiency of the quadrature (3.1) depending on the parameter R bounding the condition: $sp(A^2) \subset [1, R]$ (cf. Remark A.2).

We assume that the matrices in (1.5) can be represented in the form $A_j = L_j - \frac{\kappa^2}{d} I$ with $\Re sp(L_j) \subset [a, \infty)$, $a > 0$, and with $\kappa^2 \in \mathbb{R}_+$ (Helmholtz type operators). Due to (1.5), we can write

$$A^2 = \sum_{j=1}^d \left(L_j^2 - 2\kappa^2 L_j + \frac{\kappa^4}{d} I \right) + 2 \sum_{1 \leq i < j \leq d} L_{ij}, \quad L_{ij} := L_i L_j,$$

which implies (using $G_j := L_j^2 - 2\kappa^2 L_j + \frac{\kappa^4}{d} I$)

$$\mathcal{F}_M(A) = \sum_{k=-M}^M S_k \prod_{j=1}^d e^{-t_k G_j} \quad \text{with} \quad S_k = c_k \prod_{1 \leq i < j \leq d} e^{-2t_k L_{ij}}, \quad (3.2)$$

since, by definition, the matrices G_j, L_{ij} ($1 \leq i < j \leq d$) commute pairwise. For the finite difference scheme under consideration we can write

$$L_j = I \otimes \dots \otimes I \otimes B^j \otimes I \otimes \dots \otimes I,$$

with $B^j = \text{tridiag}\{a_j, b_j, c_j\} \in \mathbb{R}^{n \times n}$ situated in the position number $j \in \{1, \dots, d\}$. $I \in \mathbb{R}^{n \times n}$ is the identity. Then denoting

$$G^j := (B^j)^2 - 2\kappa^2 B^j + \frac{\kappa^4}{d} I \in \mathbb{R}^{n \times n}, \quad (3.3)$$

and substituting $B^j - \frac{\kappa^2}{d} I$ (with B^j strongly positive) into the representation (1.5) instead of V^j , implies

$$G_1 = G^1 \otimes I \otimes \dots \otimes I, \quad G_2 = I \otimes G^2 \otimes \dots \otimes I, \quad \dots, \quad G_d = I \otimes \dots \otimes I \otimes G^d \quad (3.4)$$

and

$$L_{ij} = I \otimes \dots \otimes I \otimes B^i \otimes I \otimes \dots \otimes I \otimes B^j \otimes I \otimes \dots \otimes I,$$

where the Kronecker product factors B^i, B^j correspond to the position numbers i and j , respectively. In the following, we assume

$$sp(L_j) \subset [\lambda_{min}, \lambda_{max}], \quad \lambda_{min}, \lambda_{max} \in \mathbb{R}_{>0} \quad (3.5)$$

(an extension to the case of complex eigenvalues is possible).

Lemma 3.1 *Under the assumption (3.5) suppose that*

$$\text{either (a) } (1 - \sqrt{1 - 1/d})\kappa^2 \geq \lambda_{max} \quad \text{or (b) } (1 + \sqrt{1 - 1/d})\kappa^2 \leq \lambda_{min}. \quad (3.6)$$

Then with S_k from (3.2) and G^j from (3.3) the representation

$$\mathcal{F}_M(A) = \sum_{k=-M}^M S_k \bigotimes_{j=1}^d \exp(-t_k G^j) \quad (3.7)$$

provides the error bound $\|\mathcal{F}(A) - \mathcal{F}_M(A)\| \leq C e^{-\pi M / (\log R \log M)}$. Moreover, it can be computed within the tolerance $\varepsilon = n^{-\alpha}$, $\alpha > 0$, with the cost $\mathcal{O}(d^2 n^2 \log R \log^q n)$.

Proof. We start from the representation (3.2). Using (3.4), we immediately obtain

$$\prod_{j=1}^d \exp(-t_k G_j) = \bigotimes_{j=1}^d \exp(-t_k G^j).$$

To analyse the spectrum of G^j in (3.3), we find that under condition (3.6), both roots $\lambda_{1,2} = (1 \pm \sqrt{1 - 1/d})\kappa^2$ of the characteristic polynomial $P(\lambda) = \lambda^2 - 2\kappa^2\lambda + \kappa^4/d$ lie on one side of the interval $[\lambda_{min}, \lambda_{max}]$, so that we have $\min_{\lambda \in [\lambda_{min}, \lambda_{max}]} P(\lambda) = Q > 0$. The property $Q > 0$ means that the corresponding matrix exponentials can be represented by \mathcal{H} -matrices with the cost $\mathcal{O}(n \log^q n)$ (cf. [4, 6]).

In general, the factor S_k in (3.2) cannot be represented exactly in a tensor-product form, hence, the complexity of the product $\prod_{1 \leq i < j \leq d} \exp(-2t_k L_{ij})$ is dominated by the cost for approximating the specific exponential matrices $\exp(-2t_k L_{ij})$ in a data-sparse format. This is possible because, by the definition of L_{ij} and by assumption (3.5), we have $sp(L_{ij}) \subset \mathbb{R}_{>0}$. Since L_{ij} operates on an index set isomorphic to $\mathbb{R}^{n^2 \times n^2}$, the corresponding cost can be estimated by $\mathcal{O}(d^2 n^2 \log^q n)$ provided that we apply an \mathcal{H} -matrix approximation as discussed, e.g., in [7]. ■

Even under the above conditions, the computational complexity is only quadratic with respect to n rather than exponential in d .

Remark 3.2 *Let $\mathcal{A} = -\Delta - \kappa^2$, $\kappa \in \mathbb{R}$, be the Helmholtz operator in $\Omega = (0, b)^d$ and A be the corresponding FD scheme on the uniform grid with mesh-size h . Then we have $B^j = -h^{-2} \text{tridiag}\{1, -2, 1\} \in \mathbb{R}^{n \times n}$, which ensures that $\lambda_{min} = \mathcal{O}(1)$ and $\lambda_{max} = \mathcal{O}(b^2 h^{-2})$, where $h > 0$ is the corresponding mesh size.*

Now, the assumption (a) (cf. (3.6)) on κ means that $\kappa \geq C_0 b h^{-1}$. On the other hand, we have to assume that $\kappa h = q < 1$ (approximation condition). We conclude that the required bound $q \geq C_0 b$ is valid if the size of Ω satisfies $b \leq q/C_0$. Note that the latter condition includes neither the mesh parameter h nor κ . In particular, if $\text{diam}(\Omega) > 1$, we can represent the inverse matrix in the framework of the parallel Schur complement domain decomposition method. In this method we use the Schur complement matrix on the interface, where the approximate inverse is computed only for “small” subdomains forming a geometric decomposition (cf. [18]). Hence, in this way we can reduce the subdomain size to satisfy the condition (3.6).

The assumption (b) in (3.6) simply means that our method works for $\kappa = \mathcal{O}(1)$.

Remark 3.3 *Again, we assume that the matrices in (1.5) have the form $A_j = L_j - \frac{\kappa^2}{d} I$. Rewriting G^j as $G^j = (B^j - \kappa^2 I)^2 + (\frac{1}{d} - 1)\kappa^4 I$, we obtain the modified representation*

$$\mathcal{F}_M(A) = \sum_{k=-M}^M e^{(d-1)t_k \kappa^4} S_k \bigotimes_{j=1}^d \exp(-t_k (B^j - \kappa^2 I)^2) \quad (3.8)$$

(cf. (3.7)), which allows to get rid of the restrictions on κ to ensure the positiveness of G^j (cf. (3.6)). However, if κ is large, the ansatz (3.8) includes “small” matrix exponentials with large coefficients which may cause numerical instabilities.

In the rest of this section, we analyse some cases when the quadratic cost $\mathcal{O}(d^2 n^2 \log^q n)$ in n can be reduced to the linear expense $\mathcal{O}(d^2 n \log^q n)$. Assume that

$$B^j = T \cdot D^j \cdot T^{-1} \in \mathbb{R}^{n \times n}, \quad j = 1, \dots, d, \quad (3.9)$$

with real diagonal matrices $D^j = \text{diag}\{\lambda_1^{(j)}, \dots, \lambda_n^{(j)}\}$ and that T, D^j are numerically available. For example, it is the case if L_j is the finite difference approximation of the one-dimensional Laplacian.

Lemma 3.4 *Let (3.9) hold with real diagonal matrices D^j . Then the matrix $e^{-2t_k L_{ij}}$ can be represented (up to a tolerance $\varepsilon > 0$) in the Kronecker tensor-product form*

$$e^{-2t_k L_{ij}} \approx e^{d-2} \sum_{m=1}^{r_1} I \otimes \dots \otimes I \otimes \mathcal{F}_{m1}(B^i) \otimes I \otimes \dots \otimes I \otimes \mathcal{F}_{m2}(B^j) \otimes I \otimes \dots \otimes I \equiv \mathcal{F}_{ij} \quad (3.10)$$

with $r_1 = \mathcal{O}(\log^2(\varepsilon^{-1}))$ and with some explicitly given functions \mathcal{F}_{m1} and \mathcal{F}_{m2} depending on t_k .

Proof. The assertion is a consequence of (3.9) and the corresponding approximation results for the function $\exp(-xy)$, $x, y \geq 0$ (see the example in §A.6.4). In fact, (3.9) implies

$$e^{-2t_k L_{ij}} = T \exp\{-2t_k I \otimes \dots \otimes I \otimes D^i \otimes I \otimes \dots \otimes I \otimes D^j \otimes I \otimes \dots \otimes I\} T^{-1}.$$

Now the existence of (3.10) is equivalent to the approximability of the exponentials of the Hadamard products

$$\Lambda^{(ij)} = \{\exp(-2t_k \lambda_l^{(i)} \cdot \lambda_m^{(j)})\}_{l,m=1}^n \in \mathbb{R}^{n \times n}, \quad \lambda_l^{(i)} \in sp(B^i), \lambda_m^{(j)} \in sp(B^j),$$

by rank- r_1 -matrices with small r_1 . In turn, the latter task is reduced to the separable approximation of the generating function $\exp(-2t_k xy)$, $x, y \in [\lambda_{min}, \lambda_{max}] \subset \mathbb{R}_+$, $t_k > 0$, which is accomplished by using a *Sinc* interpolation. For this purpose, we consider the modified function $g(x, y) = \frac{x}{1+x} \exp(-2t_k xy)$, $x \in \mathbb{R}_+$. This function $g(x, y)$ can be approximated by a separable ansatz $g_{M_1}(x, y)$ like in §A.6.4. Then we derive

$$|g(x, y) - g_{M_1}(x, y)| \leq (2t_k)^{-1} M_1^{1/2} e^{-cM_1^{1/2}}.$$

Note that we have $t_k \geq CM^{-1}$ with M corresponding to the exterior sum. Then choosing $r_1 = 2M_1 + 1$, we finally obtain the desired separable representation to the function $\frac{1+x}{x} g(x, y)$, $x \in [\lambda_{min}, \lambda_{max}]$. ■

Combining (3.10) with (3.7), we arrive at

$$\mathcal{F}_M(A) = \sum_{k=-M}^M c_k \left\{ \prod_{1 \leq i < j \leq d} \mathcal{F}_{ij} \right\} \bigotimes_{j=1}^d e^{-t_k G^j} \equiv \sum_{k=-M}^M \tilde{S}_k \bigotimes_{j=1}^d e^{-t_k G^j}. \quad (3.11)$$

This representation is of the generalised format (1.11) with the Kronecker rank $r = 2M + 1$. The complexity of (3.11) can be estimated by $\mathcal{O}(Mr_1 d^2 n \log^q n)$ provided that we are able to diagonalise each matrix B^j ($1 \leq j \leq d$) with the cost $\mathcal{O}(n \log n)$ (e.g., if L_j represents a discrete elliptic operator with constant coefficients). As a consequence, the functions $\mathcal{F}_{mp}(B^i)$ can be represented with linear-logarithmic cost in n .

3.2 Sum of Conjugate Resolvents

Using quadrature rules like (1.10) it is often possible to choose $c_k = c_{-k}$ and $z_k = \bar{z}_{-k}$. For such cases we propose the following simplifications to represent a sum of conjugate resolvents.

Let $z = a + ib \in \mathbb{C}$ with $a \geq 0$ and a matrix $L = \sum_{j=1}^d L_j$ be given such that

- (a) $sp(L_j) \subset \mathbb{R}_{>0}$ for all $j = 1, \dots, d$;
- (b) $b^2 + \Re e(a - \mu)^2 > c_L > 0$ for all $\mu \in sp(L)$.

We are interested in the HKT approximation to the matrices

$$\begin{aligned} \mathcal{G}^+(L) &:= (zI - L)^{-1} + (\bar{z}I - L)^{-1} = 2X(b^2 I + X^2)^{-1}, \\ \mathcal{G}^-(L) &:= (zI - L)^{-1} - (\bar{z}I - L)^{-1} = -2ib(b^2 I + X^2)^{-1} \end{aligned} \quad (3.12)$$

with $X = aI - L$. Consider, for example, the matrix $\mathcal{G}^+(L)$. Due to condition **(b)**, one can represent the matrix

$$\mathcal{F}(L) = (2X)^{-1}\mathcal{G}^+(L) = (b^2I + X^2)^{-1}$$

by

$$\mathcal{F}(L) := \int_0^\infty e^{-t(b^2I + X^2)} dt \approx \mathcal{F}_M(L) := \sum_{k=-M}^M c_k e^{-t_k(b^2I + X^2)}$$

(cf. (3.1)). Using the same notation as in §2.1, we substitute $a = \kappa^2$, $A_j = L_j - \frac{a}{d}I$, to obtain

$$b^2I + X^2 = \sum_{j=1}^d (L_j^2 - 2aL_j + (a^2 + b^2)/d) + 2 \sum_{1 \leq i < j \leq d} L_{ij}, \quad L_{ij} := L_i L_j.$$

Taking into account condition **(a)**, we are able to apply Remark 3.3 which, in our particular case, leads to the representation of the form (1.11),

$$\mathcal{F}_M(L) = \sum_{m=-M}^M e^{t_m(a^2 - \frac{a^2 + b^2}{d})} S_m \bigotimes_{j=1}^d \exp(-t_m(B^j - aI)^2) \approx \mathcal{F}(L). \quad (3.13)$$

Opposite to Remark 3.3, the expression (3.13) does not indicate any numerical instabilities because of the following remark.

Remark 3.5 *The Kronecker sum (3.13) includes only coefficients of the size $\mathcal{O}(\varepsilon^{-1})$ which do not cause numerical instabilities. In fact, by construction $\max_m \{t_m\} \leq CM = \mathcal{O}(|\log \varepsilon|)$, where ε is the required accuracy (cf. Appendix). Moreover, by the same reasons, we can show that the quadrature points in (1.10) applied to the Dunford-Cauchy integral also satisfy $\max_{-M \leq k \leq M} \sqrt{a_k^2 + b_k^2} \leq CM$, where $z_k = a_k + ib_k$.*

Representation (3.13) not only prove the existence of an HKT approximation to the considered class of matrix-valued functions but also provides a constructive algorithm for computing such an approximation in real arithmetic. Similarly to (3.11), the corresponding cost is dominated by $\mathcal{O}(Mrd^2n \log^q n)$.

If we are interested to approximate the individual resolvents $(zI - L)^{-1}$, we can apply the representation (1.8). The corresponding construction is completely similar to the previous case. The only difference is in the usage of complex arithmetics to multiply A^* with a real valued integral representing the matrix $\mathcal{F}(L)$ analysed above. The same recipe can be used to approximate A^{-1} in the case of an invertible matrix with a rather general location of the spectrum $sp(A)$.

3.3 Approximation to $\mathcal{F}(A)$

Assume that a given matrix valued function $\mathcal{F}(A)$ allows the Dunford-Cauchy representation. Assume that we are given a quadrature rule (1.10) with symmetric coefficients and quadrature points (i.e., $c_k = c_{-k}$, $z_k = \bar{z}_{-k}$, $k = 1, \dots, M$) that converges exponentially in M , i.e.,

$$\|\mathcal{F}(A) - \mathcal{F}_M(A)\| \leq Ce^{-cM^\alpha} \quad \text{for some } \alpha > 0.$$

For our particular quadratures we have $\alpha \geq 1/2$. Now, we apply the HKT approximation to each couple of conjugate resolvents (cf. §3.2) to obtain a method of complexity $\mathcal{O}(Mrd^2n \log^q n)$ provided that we are able to diagonalise each matrix B^j .

4 HKT Approximation to $\text{sign}(A)$

We recall that the matrix sign function of $A \in \mathbb{R}^{N \times N}$ is defined by

$$\text{sign}(A) := \frac{1}{\pi i} \int_{\Gamma_+} (zI - A)^{-1} dz - I \quad (4.1)$$

with Γ_+ being any simply closed curve in the complex plane whose interior contains all eigenvalues of A with positive real part (cf. §4). The first approach to approximate $\mathcal{F}(A) = \text{sign}(A)$ in the HKT format is based on the efficient quadrature (1.10) (cf. [6] concerning the existence), then the corresponding Kronecker rank $r = 2M + 1$.

The second method to construct the HKT representation to $\text{sign}(A)$ is based on the matrix version of the integral representation

$$\text{sign}(a) = \frac{1}{c_f} \int_0^\infty \frac{f(ta)}{t} dt, \quad c_f > 0, \quad (4.2)$$

where $a \in \mathbb{C}$, $\Re a \neq 0$, with $f : \mathbb{C} \rightarrow \mathbb{C}$ satisfying certain assumptions discussed in §4.1.

To derive efficient representations (1.1) or (1.11), the function f has to be chosen in such a way that the integral (4.2) allows exponentially convergent quadrature and, moreover, f facilitates a good “separability property”. We show that the following examples are satisfactory choices:

$$f_1(t) := t \exp(-t^2), \quad (4.3a)$$

$$f_2(t) := \frac{t}{1 + \alpha t^2}, \quad \alpha > 0, \quad (4.3b)$$

$$f_{3,n}(t) := \frac{j_n(t)}{t^{n-1}}, \quad n = 1, 2, \dots, \quad (4.3c)$$

where $j_n(t)$ are the spherical Bessel functions of the first kind (cf. [8]). In particular, we have

$$j_0(t) = \frac{\sin(t)}{t}, \quad j_1(t) = \frac{\sin(t) - t \cos(t)}{t^2}, \quad j_2(t) = \left(\frac{3}{t^3} - \frac{1}{t}\right) \sin(t) - \frac{3}{t^2} \cos(t).$$

Note that the integral representations (1.7) and (4.4) (the latter with the choice $f(t) = t \exp(-t^2)$) applied to the matrices A^{-1} and $\text{sign}(A)$, respectively, lead to rather similar expressions. Representation (4.2) may have different advantages and limitations depending on the particular choice of the function f and the properties of A (selfadjoint, diagonalisable or rather general).

In §4, we analyse the representations involving the integrands (4.3a-c) in more detail. In particular, we show that the generating functions $f_1(t)$, $f_2(t)$ can be applied to a rather general class of matrices provided that $\Re \text{sp}(A^2) \subset \mathbb{R}_{>0}$. The corresponding complexity is proved to be $\mathcal{O}(d^2 n^2 \log^q n)$. In the case of real-diagonalisable matrices one can use generating functions $f_{3,n}(t)$ expecting the complexity $\mathcal{O}(dn \log^q n)$.

4.1 Using the Representation (4.2)

We consider two classes of matrices:

Case (A) Let A be real-diagonalisable, i.e., $A = T D T^{-1}$, where D is real diagonal. Now the function $f : \mathbb{R} \rightarrow \mathbb{R}$ is supposed to have the properties

$$(A1) \quad f(t) = -f(-t), \quad t \in \mathbb{R},$$

$$(A2) \quad c_f := \int_0^\infty \frac{f(t)}{t} dt \in (0, \infty) \text{ exists as an improper integral.}$$

Case (B) We assume that $\sigma(A) = \sigma_+(A) \cup \sigma_-(A)$, where $\sigma_+(A) := \{\lambda \in \sigma(A) : \Re \lambda > 0\}$, $\sigma_-(A) := \{\lambda \in \sigma(A) : \Re \lambda < 0\}$. The function $f : \mathbb{C} \rightarrow \mathbb{C}$ is supposed to have the properties

$$(B1) \quad f(t) = -f(-t), \quad t \in \mathbb{R}.$$

(B2) The function $f : \mathbb{C} \rightarrow \mathbb{C}$ is analytic in the domain $\Omega = \Omega_+ \cup \Omega_-$ with boundary $\Gamma = \Gamma_+ \cup \Gamma_-$, which is the union of two closed simply connected curves Γ_+ and Γ_- , each of which contains the respective part of spectrum σ_\pm . Moreover,

$$|f(z)| \leq C(1 + |z|)^{-1} \quad \text{for all } z \in \Omega_\theta := \{z : |\arg(z)| \leq \theta < \frac{\pi}{2}\} \text{ with } \Omega \subset \Omega_\theta.$$

(B3) For any $z = r e^{i\theta} \in \Gamma$, we have $c_f := \int_0^{\infty e^{i\theta}} \frac{f(u)}{u} du \in (0, \infty)$, where c_f does not depend on z , with the integration path running along the ray $\{u : u = \rho e^{i\theta}, \rho \in [0, \infty)\}$.

Note that in both cases, f is thought to allow an efficient quadrature for (4.2).

Based on formula (4.2), we derive the integral representation to the matrix $\text{sign}(A)$.

Lemma 4.1 *Let A be a square matrix A such that $0 \notin \Re \sigma(A)$. Let the function f satisfy the Assumptions (A1)-(A2) or (B1)-(B3) in the respective Cases (A) or (B). Then we have*

$$\text{sign}(A) = \frac{1}{c_f} \int_{\mathbb{R}_+} \frac{f(tA)}{t} dt. \quad (4.4)$$

Proof. First we note that for $a \in \mathbb{R} \setminus \{0\}$, the assumptions (A1)-(A2) imply (4.2), while for $a \in \mathbb{C}$ with $\Re a \neq 0$, (B1)-(B3) also yield (4.2).

In Case (A) we have $A = T D T^{-1}$, so that

$$f(tA) = T f(tD) T^{-1}. \quad (4.5)$$

Moreover, $\text{sign}(A) = T \text{sign}(D) T^{-1}$ holds and (4.2) implies the desired relation:

$$\frac{1}{c_f} \int_{\mathbb{R}_+} \frac{f(tA)}{t} dt = T \left(\frac{1}{c_f} \int_{\mathbb{R}_+} \frac{f(tD)}{t} dt \right) T^{-1} = T \text{sign}(D) T^{-1} = \text{sign}(A).$$

In Case (B), the analytic function $f : \mathbb{C} \rightarrow \mathbb{C}$ generates the family of matrix-valued functions $f(tA)$, $t \geq 0$, which can be represented by the Dunford-Cauchy integral

$$f(tA) = \frac{1}{2\pi i} \int_{\Gamma} f(tz) (zI - A)^{-1} dz, \quad (4.6)$$

where $\Gamma = \Gamma_+ \cup \Gamma_-$ is the union of two closed simply connected curves Γ_+ and Γ_- , each of which contains the respective part σ_{\pm} of the spectrum (cf. assumption (B2)). Note that Γ_{\pm} can be chosen in such a way that with some positive constant μ , the relation $|\Re z| > \mu > 0$ holds for $z \in \Gamma_{\pm}$. Now due to assumption (B3), we obtain

$$\|f(tA)\| \leq c \int_{\Gamma} |f(tz)| \|(zI - A)^{-1}\| |dz| \leq \frac{c}{1+t} \int_{\Gamma} \frac{1}{1+|z|} \|(zI - A)^{-1}\| |dz|,$$

which proves the existence of the integral in (4.4). Let us introduce the integrals

$$B_+ = \frac{1}{\pi i} \int_{\Gamma_+} (zI - A)^{-1} dz, \quad B_- = \frac{1}{\pi i} \int_{\Gamma_-} (zI - A)^{-1} dz.$$

By definition of Γ_+ and Γ_- we have

$$\frac{1}{2}(B_+ + B_-) = I, \quad \text{and thus } B_- = 2I - B_+. \quad (4.7)$$

We substitute the Dunford-Cauchy integral (4.6) into (4.4) and use (4.7) to derive

$$\begin{aligned} \frac{1}{c_f} \int_{\mathbb{R}_+} \frac{f(tA)}{t} dt &= \frac{1}{2\pi i} \int_{\Gamma} \left[\frac{1}{c_f} \int_{\mathbb{R}_+} \frac{f(tz)}{t} dt \right] (zI - A)^{-1} dz \\ &= \frac{1}{2\pi i} \int_{\Gamma_+ \cup \Gamma_-} \text{sign}(z) (zI - A)^{-1} dz \\ &= \frac{1}{2}(B_+ - B_-) = B_+ - I \\ &= \text{sign}(A), \end{aligned}$$

which completes the proof. ■

4.2 Analysis for the Integrands (4.3a) and (4.3b)

In both Cases (A) and (B), we derive an efficient quadrature for the choice $f = f_1(t)$ in (4.2). Let (1.2) be valid and let $\Re \lambda \geq 1$ for all $\lambda \in \sigma(A^2)$. We approximate the integral (4.4) with $f = f_1$ by applying an exponentially convergent quadrature rule to the integral

$$\int_0^\infty \exp(-t^2 A^2) dt = \frac{1}{2} \int_{\mathbb{R}} \exp(-t^2 A^2) dt$$

appearing in

$$\mathcal{F}(A) := A^{-1} \text{sign}(A) = \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} e^{-t^2 A^2} dt \approx \sum_{k=-M}^M c_k e^{-t_k^2 A^2} =: \mathcal{F}_M(A) \quad (4.8)$$

(we set $c_f = \sqrt{\pi}$) with c_k, t_k given in Appendix. Due to (1.2), we can use the same techniques as in §3.1 to represent each individual exponent in (4.8) in tensor-product form, which leads to the cost $\mathcal{O}(M d^2 n^2 \log^q n)$ with $M = \mathcal{O}(\log(\text{cond}(A^2)) \log \varepsilon^{-1})$ or with $M = \mathcal{O}(\log^2 \varepsilon^{-1})$ depending on the relation between ε^{-1} and $\text{cond}(A^2)$.

Note that in the case $f = f_2(t)$, we have

$$\text{sign}(A) = \frac{A}{c_f(\alpha)} \int_0^\infty (I + \alpha t^2 A^2)^{-1} dt,$$

which is similar to the familiar Robert's integral representation

$$\text{sign}(A) = \frac{2A}{\pi} \int_0^\infty (t^2 I + A^2)^{-1} dt.$$

This case can be reduced to the analysis of the matrix \mathcal{G} in (3.12), therefore, all the results in §3.2 can be applied as well.

4.3 Construction in Case (4.3c)

In Case (A), we may consider as well the generating functions $f = f_{3,n}$ from (4.3c). The spherical Bessel functions $j_n(z)$ (cf. [8]) have the asymptotical property

$$z^{-n} j_n(z) \rightarrow \frac{1}{1 \cdot 3 \cdot 5 \dots (2n-1)} \quad \text{as } z \rightarrow 0 \quad (n = 0, 1, 2, \dots).$$

We also use the integral representation

$$j_n(z) = \frac{z^n}{2^{n+1} n!} \int_0^\pi \cos(z \cos \theta) \sin^{2n+1} \theta d\theta \quad (n = 0, 1, 2, \dots). \quad (4.9)$$

Since the matrix A is diagonalisable, the error analysis of the quadrature rule is reduced to the scalar case.

Let us construct an exponentially convergent quadrature for (4.2) with $f = f_{3,n}$ and with $a \in \mathbb{R}$. In general, one can expect $a \in [1, \Lambda]$ with $1 \ll \Lambda$, so we deal with the integration of a highly oscillatory function with smooth weight. We recall that

$$j_n(z) = g_n(z) \sin z + (-1)^{n+1} g_{-n-1}(z) \cos z \quad (4.10)$$

(cf. [8]), where $g_0(z) = z^{-1}$, $g_1(z) = z^{-2}$, $g_{n-1}(z) + g_{n+1}(z) = (2n+1)z^{-1}g_n(z)$ for $n \in \mathbb{Z}$.

(4.10) yields the estimate $|j_n(t)| \leq C/t$, $t \rightarrow +\infty$, hence we have

$$\frac{f_{3,n}(at)}{t} = \frac{j_n(at)}{a^{n-1} t^n} \leq \frac{C}{at^2}, \quad t \rightarrow \infty. \quad (4.11)$$

The latter implies

$$\left| \int_R^\infty \frac{f_{3,n}(at)}{t} dt \right| \leq \frac{C}{aR}, \quad R > 0.$$

Moreover, due to (4.9), $\frac{f_{3,n}(az)}{z}$ is holomorphic at $z = 0$ (in fact, it is an entire function).

Now, given a tolerance $\varepsilon > 0$, we choose $R > 0$ such that $R^{-1} = a\varepsilon$, i.e., $R = (a\varepsilon)^{-1}$, and then construct a quadrature on the finite interval $[0, R]$.

Again, we split $[0, R]$ into the two parts $[0, a^{-1}]$ and $\omega := [a^{-1}, R]$. We further decompose the integration interval $\omega = \bigcup_{k=K_0}^{K_1} [b_k, b_{k+1}]$ by the points $b_k = 2^k$, $k = -K_0, \dots, 0, \dots, K_1$, where $K_0, K_1 \in \mathbb{N}$ are the minimal numbers, such that $2^{-K_0} \leq a^{-1}$ and $2^{K_1} \geq R$. Note that one can choose $K_1 = |\log \varepsilon| + \log(\text{cond}(A))$ and $K_0 = \log(\text{cond}(A))$, provided that $\min_{\lambda \in \sigma_+(A)} \lambda = \mathcal{O}(1)$.

Since $g_k(z)$ from (4.10) is a polynomial in z^{-1} , it can be approximated on each interval $\delta_k = [b_k, b_{k+1}]$ by a polynomial $\mathcal{P}_{p,k}$ of degree p such that

$$\max_{t \in \delta_k} |g_k(t) - \mathcal{P}_{p,k}(t)| \leq C e^{-cp}. \quad (4.12)$$

Next we use the integrals

$$\begin{aligned} \int_0^x t^m \sin(at) dt &= - \sum_{k=0}^m k! \binom{m}{k} \frac{x^{m-k}}{a^{k+1}} \cos\left(ax + \frac{1}{2}k\pi\right), \\ \int_0^x t^m \cos(at) dt &= \sum_{k=0}^m k! \binom{m}{k} \frac{x^{m-k}}{a^{k+1}} \sin\left(ax + \frac{1}{2}k\pi\right) \end{aligned}$$

(cf. [8]) to obtain the following approximation on the interval ω :

$$\frac{1}{c_f} \int_{\omega} \frac{f_{3,n}(at)}{t} dt \simeq \sum_{k=-K_0}^{K_1} \sum_{\ell=0}^p [\gamma_{k\ell} \sin(as_{k\ell}) + \mu_{k\ell} \cos(ac_{k\ell})],$$

which provides an exponential convergence of the order $\mathcal{O}(e^{-cp})$.

Due to (4.9), the integrand $\frac{f_{3,n}(az)}{z}$ is an entire function and, in particular, holomorphic in the Bernstein ellipse \mathcal{E}_ρ with $\rho > 1/(2a)$, corresponding to the interval $[0, a^{-1}]$ (cf. [16]). Furthermore, $\max_{z \in \mathcal{E}_\rho} \left| \frac{f_{3,n}(az)}{z} \right|$ can be estimated by a constant not depending on a . Therefore, the Gauss quadrature on $[0, \Lambda^{-1}]$ is exponentially convergent. This yields the approximation

$$\text{sign}(\lambda) \sim \text{sign}_M(\lambda) := \sum_{k=1}^M a_k \sin(s_k \lambda) + b_k \cos(c_k \lambda), \quad (4.13)$$

such that for $\lambda \in [1, \Lambda]$ there holds

$$|\text{sign}(\lambda) - \text{sign}_M(\lambda)| \leq C(K_0 + K_1) e^{-cp}$$

with

$$K_1 = |\log \varepsilon|, \quad K_0 = \log(\text{cond}(A)), \quad M := (K_0 + K_1) p. \quad (4.14)$$

Lemma 4.2 *Let A be symmetric with $\min_{\lambda \in \sigma_+(A)} \lambda = \mathcal{O}(1)$. Given $\varepsilon > 0$, then the quadrature points and weights from (4.13) and (4.14) fulfil*

$$\left\| \frac{1}{c_f} \int_0^\infty \frac{f_{3,n}(tA)}{t} dt - \mathcal{P}_n(A^{-1}) \sum_{k=1}^M [a_k \sin(s_k A) + b_k \cos(c_k A)] \right\|_2 \leq C(K_0 + K_1) e^{-cp}, \quad (4.15)$$

where p is defined by the choice of the polynomial $\mathcal{P}_{p,k}$ in (4.12) and M, K_0, K_1 are explained in (4.14).

Proof. Since $A = T^*DT$, we use the representation (4.5), where D has real entries, and derive

$$\begin{aligned}
& \left\| \frac{1}{c_f} \int_{\mathbb{R}_+} \frac{f_{3,n}(tA)}{t} dt - \sum_{k=1}^M [a_k \sin(s_k A) + b_k \cos(c_k A)] \right\|_2 \\
&= \left\| T^* \left(\frac{1}{c_f} \int_{\mathbb{R}_+} \frac{f_{3,n}(tD)}{t} dt - \sum_{k=1}^M [a_k \sin(s_k D) + b_k \cos(c_k D)] \right) T \right\|_2 \\
&\leq \max_{\lambda \in \sigma_+(A)} \left| \frac{1}{c_f} \int_{\mathbb{R}_+} \frac{f_{3,n}(t\lambda)}{t} dt - \sum_{k=1}^M [a_k \sin(s_k \lambda) + b_k \cos(c_k \lambda)] \right| \\
&\leq C[K_0 + K_1]e^{-c_p}.
\end{aligned}$$

Since $M = p(K_0 + K_1)$ (cf. (4.14)) the proof is complete. \blacksquare

Note that the simplest possible approximation can be constructed with the choice $f = f_{3,1}$.

To complete this section, we derive tensor-product representations of the matrices $\sin(s_k A)$ and $\cos(c_k A)$ involved in (4.15). For this purpose, we apply the following proposition which can be proved by induction. In the case $d = 2$, the assertion (4.16) is trivial.

Proposition 4.3 ([1]) *Let $d \geq 2$. The trigonometric identity*

$$\sin \left(\sum_{j=1}^d x_j \right) = \sum_{j=1}^d \sin(x_j) \prod_{k \in \{1, \dots, d\} \setminus \{j\}} \frac{\sin(x_k + \alpha_k - \alpha_j)}{\sin(\alpha_k - \alpha_j)} \quad (4.16)$$

holds for all choices of $\{\alpha_1, \dots, \alpha_d\}$ such that $\sin(\alpha_k - \alpha_j) \neq 0$ for all $j \neq k$.

Corollary 4.4 *Let $A = \sum_{j=1}^d A_j \in \mathbb{R}^{N \times N}$ with matrices A_j of the form (1.5), where $T_j \in \mathbb{R}^{n \times n}$ ($j = 1, \dots, d$) and $N = n^d$. Suppose that $\{\alpha_1, \dots, \alpha_d\} \subset \mathbb{R}$ are chosen in such a way that the representation (4.16) is valid. Then the following tensor-product representation with exactly d terms*

$$\sin(A) = \sum_{j=1}^d \bigotimes_{k=1}^d \beta_{kj} \sin(T_j + (\alpha_k - \alpha_j)I), \quad \beta_{kj} = \begin{cases} \frac{1}{\sin(\alpha_k - \alpha_j)}, & k \neq j, \\ 1, & k = j, \end{cases}$$

holds. A similar representation exists for the matrix $\cos(A)$.

Lemma 4.2 and Corollary 4.4 lead to the desired Kronecker tensor-product representation of the matrix $\text{sign}(A)$ having the complexity $\mathcal{O}(dMn \log^q n)$ provided that each T_j ($j = 1, \dots, d$) can be diagonalised with the cost $\mathcal{O}(n \log^q n)$.

4.4 Dunford-Cauchy Integral (4.1) Revisited

In the general case, one can apply the integral representation (4.1). The exponentially convergent quadrature

$$\text{sign}(A) \approx \sum_{k=1}^r c_k (z_k I - A)^{-1} - I, \quad r = \mathcal{O}(\log^2 \varepsilon + \log^2 \text{cond}(A)),$$

for the integral (4.1) provides the direct approximation of $\mathcal{F}(A) = \text{sign}(A)$ by a sum of matrix resolvents (cf. [6]). The quadrature points and weights can be chosen symmetrically. Using the results in §3 (cf. §2.2), this leads to the overall cost $\mathcal{O}(rd^2 n^2 \log^q n)$ in the multi-dimensional case. Again, the complexity is quadratic in d and n .

A Appendix

A.1 Separation by Integration

If a function of ρ can be written as the integral

$$\varphi(\rho) = \int_I e^{\rho F(t)} G(t) dt$$

over some $I \subset \mathbb{R}$ and if quadrature can be applied, one obtains $\varphi(\rho) \approx \sum_{\nu} e^{\rho F(x_{\nu})} G(x_{\nu})$. Setting $\rho = \sum_{i=1}^d x_i$, we get the separable approximation

$$\varphi(x_1 + \dots + x_d) \approx \sum_{\nu} G(x_{\nu}) \prod_{i=1}^d e^{x_i F(x_{\nu})}.$$

The above argument applies as well for the matrix-valued function $\varphi(A)$ with $A = \sum_{i=1}^d A_i$ and mutually commutable matrices A_i . In the sequel, we discuss the sinc quadrature in the case of $I = \mathbb{R}$ and study the quadrature error.

A.2 Sinc Interpolation and Quadratures

A.2.1 Definitions

In this section, we present *sinc* quadrature rules to compute the integral

$$I(f) = \int_{\omega} f(\xi) d\xi \quad (\omega = \mathbb{R} \text{ or } \omega = \mathbb{R}_+). \quad (\text{A.1})$$

In the case of $\omega = \mathbb{R}$, we introduce the family $\mathbf{H}^1(D_{\delta})$ of all matrix-valued functions, which are analytic in the strip $D_{\delta} := \{z \in \mathbb{C} : |\Im z| \leq \delta\}$, such that for each $f \in \mathbf{H}^1(D_{\delta})$ we have $N(f, D_{\delta}) < \infty$ with

$$N(f, D_{\delta}) := \int_{\partial D_{\delta}} |f(z)| |dz| = \int_{\mathbb{R}} (|f(x + i\delta)| + |f(x - i\delta)|) dx.$$

Let

$$S(k, \mathfrak{h})(x) = \frac{\sin[\pi(x - k\mathfrak{h})/\mathfrak{h}]}{\pi(x - k\mathfrak{h})/\mathfrak{h}} \quad (k \in \mathbb{Z}, \mathfrak{h} > 0, x \in \mathbb{R}) \quad (\text{A.2})$$

be the k -th *sinc* function with step size \mathfrak{h} , evaluated at x . Given $f \in \mathbf{H}^1(D_{\delta})$, $\mathfrak{h} > 0$, and $M \in \mathbb{N}_0$, the corresponding Sinc-interpolant (cardinal series representation) reads as $C(f, \mathfrak{h}) = \sum_{\nu=-\infty}^{\infty} S(\nu, \mathfrak{h})f(\nu\mathfrak{h})$. We use the conventional notations

$$\begin{aligned} C_M(f, \mathfrak{h}) &= \sum_{\nu=-M}^M S(\nu, \mathfrak{h})f(\nu\mathfrak{h}), & E_M(f, \mathfrak{h}) &= f - C_M(f, \mathfrak{h}), \\ T(f, \mathfrak{h}) &= \mathfrak{h} \sum_{k=-\infty}^{\infty} f(k\mathfrak{h}), & T_M(f, \mathfrak{h}) &= \mathfrak{h} \sum_{k=-M}^M f(k\mathfrak{h}), \\ \eta(f, \mathfrak{h}) &= I(f) - T(f, \mathfrak{h}), & \eta_M(f, \mathfrak{h}) &= I(f) - T_M(f, \mathfrak{h}). \end{aligned} \quad (\text{A.3})$$

Here $\eta(f, \mathfrak{h})$ represents the quadrature error via the Sinc-interpolant $C(f, \mathfrak{h})$,

$$\int_{\mathbb{R}} f(\xi) d\xi \approx \int_{\mathbb{R}} \sum_{\nu=-\infty}^{\infty} S(\nu, \mathfrak{h})f(\nu\mathfrak{h}) d\xi = T(f, \mathfrak{h}),$$

and $\eta_M(f, \mathfrak{h})$ includes in addition the corresponding truncation error $T(f, \mathfrak{h}) - T_M(f, \mathfrak{h})$, while $E_M(f, \mathfrak{h})$ describes the interpolation error by the truncated Sinc interpolant.

A.2.2 Standard Error Estimates

The error estimate of η_M is as follows. If $f \in \mathbf{H}^1(D_\delta)$ and

$$\|f(\xi)\|_\infty \leq C \exp(-b|\xi|) \quad \text{for all } \xi \in \mathbb{R} \text{ with } b, C > 0, \quad (\text{A.4})$$

then the quadrature error η_M from (A.3) satisfies

$$\|\eta_M(f, \mathfrak{h})\|_\infty \leq C \left[\frac{e^{-2\pi\delta/\mathfrak{h}}}{1 - e^{-2\pi\delta/\mathfrak{h}}} N(f, D_\delta) + \frac{1}{b} \exp(-b\mathfrak{h}M) \right]. \quad (\text{A.5})$$

Furthermore, under the same assumptions on $f \in \mathbf{H}^1(D_\delta)$, for the interpolation error we have

$$\|E_M(f, \mathfrak{h})\|_\infty \leq C \left[\frac{e^{-\pi\delta/\mathfrak{h}}}{2\pi\delta} N(f, D_\delta) + \frac{1}{b\mathfrak{h}} \exp(-b\mathfrak{h}M) \right]. \quad (\text{A.6})$$

Equalising both terms in (A.5), (A.6) leads to the following choice of the mesh parameter \mathfrak{h} .

In the case (A.5), the choice $\mathfrak{h} = \sqrt{2\pi\delta/bM}$ leads to the exponential convergence rate

$$\|\eta_M(f, \mathfrak{h})\|_\infty \leq C e^{-\sqrt{2\pi\delta bM}} \quad (\text{A.7})$$

with a positive constant C independent of M , depending only on f, δ, b (cf. [22, 5, 6]). Note that $2M + 1$ is the number of quadrature points. If f is an even function, the number of quadrature points reduces to $M + 1$.

In the case (A.6), the choice $\mathfrak{h} = \sqrt{\pi\delta/bM}$ implies

$$\|E_M(f, \mathfrak{h})\|_\infty \leq C M^{1/2} e^{-\sqrt{\pi\delta bM}} \quad (\text{A.8})$$

with a positive constant C depending only on f, δ, b (cf. [22]).

In the case $\omega = \mathbb{R}_+$ one has to substitute the integral (A.1) by $\xi = \varphi(z)$ such that $\varphi : \mathbb{R} \rightarrow \mathbb{R}_+$ is a bijection. This changes \mathcal{F} into $f_1 := \varphi' \cdot (f \circ \varphi)$. Assuming $f_1 \in \mathbf{H}^1(D_\delta)$, one can apply (A.4)-(A.7) to the transformed function.

A.2.3 Improved Quadrature Error Estimates

The error $\|\eta_M(f, \mathfrak{h})\|$ is estimated in (A.7) by an exponential expression involving \sqrt{M} . Under stronger assumptions it is possible to improve \sqrt{M} to $M/\log M$ (see [6] for a proof).

Proposition A.1 *Let $f \in \mathbf{H}^1(D_\delta)$. If, in addition, f satisfies the condition*

$$|f(\xi)| \leq C \exp(-be^{a|\xi|}) \quad \text{for all } \xi \in \mathbb{R} \text{ with } a, b, C > 0, \quad (\text{A.9})$$

then the error η_M of $T_M(f, \mathfrak{h})$ satisfies (with δ from $\mathbf{H}^1(D_\delta)$)

$$|\eta_M(f, \mathfrak{h})| \leq C \left[\frac{e^{-2\pi\delta/\mathfrak{h}}}{1 - e^{-2\pi\delta/\mathfrak{h}}} N(f, D_\delta) + \frac{1}{ab} \exp(-be^{a\mathfrak{h}M}) \right].$$

The choice $\mathfrak{h} = \log(\frac{2\pi aM}{b})/(aM)$ then leads to

$$|\eta_M(f, \mathfrak{h})| \leq C N(f, D_\delta) e^{-2\pi\delta aM/\log(2\pi aM/b)}. \quad (\text{A.10})$$

Remark A.2 *Note that the above quadratures apply also to matrix-valued functions $\mathcal{F}(\xi, A)$, in particular, to functions leading to A^{-1} (see numerics in Sect. A.4). In the case of diagonalisable matrices, the error analysis for the corresponding quadratures applied to $\mathcal{F}(\xi, A)$ is similar to the analysis of parameter dependent scalar functions $f(\xi, \rho)$, where $\rho \in \sigma(A)$ (cf. [6, 7]).*

In the following Sections A.3, A.4, we have a closer look to two integrals of real functions, which are transformed in such a way that either (A.7) or (A.10) apply.

A.2.4 Improved Interpolation Error Estimates

Again, under the same assumption on the decay rate of f on \mathbb{R} as in Section A.2.3, it is possible to improve (A.6) and (A.8).

Proposition A.3 *Let $f \in \mathbf{H}^1(D_\delta)$. If, in addition, f satisfies (A.9) then the error E_M of $C_M(f, \mathfrak{h})$ satisfies*

$$|E_M(f, \mathfrak{h})| \leq C \left[\frac{e^{-\pi\delta/\mathfrak{h}}}{2\pi\delta} N(f, D_\delta) + \frac{e^{-a\mathfrak{h}M}}{ab\mathfrak{h}} \exp(-be^{a\mathfrak{h}M}) \right]. \quad (\text{A.11})$$

The choice $\mathfrak{h} = \log(\frac{2\pi aM}{b}) / (aM)$ then leads to

$$|E_M(f, \mathfrak{h})| \leq C \frac{N(f, D_\delta)}{2\pi\delta} e^{-2\pi\delta aM / \log(2\pi aM/b)}. \quad (\text{A.12})$$

Proof. The error $E(f, \mathfrak{h}) := f - C(f, \mathfrak{h})$ allows the same estimate as in the standard case (see first term in the right-hand side of (A.6)),

$$|E(f, \mathfrak{h})| \leq C \frac{e^{-\pi\delta/\mathfrak{h}}}{2\pi\delta} N(f, D_\delta). \quad (\text{A.13})$$

The truncation error bound hinges only upon the decay rate in (A.9),

$$\|C(f, \mathfrak{h}) - C_M(f, \mathfrak{h})\|_\infty \leq \sum_{|k| \geq M+1} |f(k\mathfrak{h})| \leq 2C \sum_{k=M+1}^{\infty} e^{-be^{ak\mathfrak{h}}} \leq \frac{2C}{ba\mathfrak{h}e^{a\mathfrak{h}M}} e^{-be^{a\mathfrak{h}M}}, \quad (\text{A.14})$$

which proves the second term in the right-hand side of (A.11). For the present choice of \mathfrak{h} , the first term in the right-hand side in (A.11) dominates, hence (A.12) follows. \blacksquare

For further applications in FEM and BEM, we reformulate the previous result for parameter dependent functions $g(x, y)$ defined on the reference interval $x \in (0, 1]$. Following the approach in [20], we introduce the mapping

$$\zeta \in D_\delta \mapsto \phi(\zeta) = \frac{1}{\cosh(\sinh(\zeta))}, \quad \delta < \frac{\pi}{2}.$$

Clearly, $(0, 1] = \phi(\mathbb{R})$ and, moreover, $\phi(\zeta)$ decays twice exponentially,

$$|\phi(\zeta)| \leq 2 \exp\left(-\frac{\cos \delta}{2} e^{|\Re \zeta|}\right), \quad \zeta \in D_\delta.$$

In particular, we have $|\phi(\zeta)| \leq 2 \exp(-\frac{1}{2}e^{|\zeta|})$, $\zeta \in \mathbb{R}$. Let $D_\phi(\delta) := \{\phi(\zeta) : \zeta \in D_\delta\} \supset (0, 1]$ be the image of D_δ . One checks easily that $D_\phi(\delta) \subset \bar{S}_r(0) \setminus \{0\}$, where $S_r(0)$ is the disc around zero with a certain radius $r > 1$. Therefore, if a function g is holomorphic on $D_\phi(\delta)$, then

$$f(\zeta) := \phi^\alpha(\zeta)g(\phi(\zeta)) \quad \text{for any } \alpha > 0$$

is also holomorphic on D_δ .

Note that the finite Sinc interpolation $C_M(f(\cdot, y), \mathfrak{h}) = \sum_{k=-M}^M f(k\mathfrak{h}, y)S_{k, \mathfrak{h}}$ together with the back-transformation $\zeta = \phi^{-1}(x) = \operatorname{arsinh}(\operatorname{arcosh}(\frac{1}{x}))$ and multiplication by $x^{-\alpha}$ yield the separable approximation to the function of interest $g(x, y)$,

$$g_M(x, y) := \sum_{k=-M}^M \phi(k\mathfrak{h})^\alpha g(\phi(k\mathfrak{h}), y) \cdot x^{-\alpha} S_{k, \mathfrak{h}}(\phi^{-1}(x)) \approx g(x, y) \quad (\text{A.15})$$

with $x \in (0, 1] = \phi(\mathbb{R})$, $y \in Y$. The error analysis is given by the following statement.

Corollary A.4 *Let $Y \in \mathbb{R}^m$ be any parameter set and assume that for all $y \in Y$ the functions $g(\cdot, y)$ together with their transformed counterparts $f(\zeta, y) := \phi^\alpha(\zeta)g(\phi(\zeta), y)$ satisfy the following conditions: (a) $g(\cdot, y)$ is holomorphic on $D_\phi(\delta)$, and $N(f, D_\delta) < \infty$;*

(b) $f(\cdot, y)$ satisfies (A.9) with $a = 1$ and with certain C, b for all $y \in Y$.

Then the optimal choice $\mathfrak{h} := \frac{\log M}{M}$ of the step size yields the total pointwise error estimates

$$|E_M(f, \mathfrak{h})| = |f(\zeta, y) - C_M(f, \mathfrak{h})(\zeta)| \leq C \frac{N(f, D_\delta)}{2\pi\delta} e^{-\pi\delta M / \log M}, \quad (\text{A.16})$$

$$|g(x, y) - g_M(x, y)| \leq |x|^{-\alpha} |E_M(f(\cdot, y), \mathfrak{h})(\phi^{-1}(x))|. \quad (\text{A.17})$$

Proof. Due to the properties of $\phi : D_\delta \rightarrow D_\phi(\delta)$, condition (a) implies $f \in \mathbf{H}^1(D_\delta)$, hence, in view of (b), we can apply Proposition A.3. Now $\frac{N(f, D_\delta)}{2\pi\delta} e^{-\pi\delta M / \log M}$ corresponds to (A.13), while the evaluation of (A.14) for the present \mathfrak{h} yields the bound $\frac{2C}{b \log M} e^{-bM}$, which is asymptotically faster decaying when $M \rightarrow \infty$.

Now the approximant (A.15) implies the bound (A.17) for $g - g_M(x, y)$. \blacksquare

The blow-up at $x = 0$ is avoided by restricting x to $[h, 1]$ ($h > 0$). In applications with a discretisation step size h it suffices to apply this estimate for $|x| \geq \text{const} \cdot h$. Since usually $1/h = \mathcal{O}(n^\beta)$ for some β (and n the problem dimension), the factor $|x|^{-\alpha}$ is bounded by $\mathcal{O}(n^{\alpha\beta})$ and can be compensated by the exponential decay in (A.16) with respect to M .

Corollary A.4 and estimate (A.17) will be applied in section A.5.

A.2.5 Sinc Interpolation of Multi-Variate Functions

Given a multi-variate function $F : \mathbb{R}^d \rightarrow \mathbb{R}$, $d \geq 1$, we are interested in its approximation by a separable expansion

$$F_r(\zeta_1, \dots, \zeta_d) := \sum_{k=1}^r \Phi_k^1(\zeta_1) \cdots \Phi_k^d(\zeta_d) \approx F,$$

where the set of functions $\{\Phi_k^\ell(\zeta_\ell)\}$, $\ell = 1, \dots, d$ can be fixed or chosen adaptively (see the discussion in [1, 19, 24]). Here the key quantity is r , which is usually called the *separation rank* and which should be reasonably small.

Let us introduce the tensor-product Sinc interpolant with respect to the first $d - 1$ variables,

$$\mathbf{C}_M F := C_M^1 \cdots C_M^{d-1} F,$$

where $C_M^\ell F = C_M^\ell(F, \mathfrak{h})$ denotes the above defined univariate Sinc interpolation operator applied in the variable $\zeta_\ell \in I_\ell$, corresponding to the notation $[0, 1]^d = I_1 \times \dots \times I_d$ including d copies of the reference interval $[0, 1]$. For each fixed $\ell \leq d - 1$, choose the variable ζ_ℓ and consider the remaining variables as a parameter set $Y_\ell := I_1 \times \dots \times I_{\ell-1} \times I_{\ell+1} \times \dots \times I_d \in \mathbb{R}^{d-1}$. Define the univariate parameter dependent function $F_\ell(\cdot, y) : I_\ell \rightarrow \mathbb{R}$, which is the restriction of F onto the interval I_ℓ with any fixed parameter $y \in Y_\ell$. The estimation of the error $F - \mathbf{C}_M F$ of our tensor-product interpolant includes the so-called Lebesgue constant $\Lambda_M \in \mathbb{R}_{\geq 1}$ defined by

$$\|C_M(f, \mathfrak{h})\|_\infty \leq \Lambda_M \|f\|_\infty \quad \text{for all } f \in C(\mathbb{R}). \quad (\text{A.18})$$

We use the following estimate on Λ_M (cf. [22, p. 142]),

$$\Lambda_M = \max_{x \in \mathbb{R}} \sum_{k=-M}^M |S(k, \mathfrak{h})(x)| \leq \frac{2}{\pi} \left(\frac{3}{2} + \gamma + \log(M+1) \right) \leq \frac{2}{\pi} (3 + \log(M)) \quad (\text{A.19})$$

with Euler's constant $\gamma = 0.577\dots$. Note that we also have

$$\sum_{k=-\infty}^{\infty} |S(k, \mathfrak{h})(x)|^2 = 1, \quad x \in \mathbb{R}$$

(cf. [22, p. 142]), which indicates $\Lambda_M = 1$ with respect to the L^2 -norm. Now we are in the position to prove the counterpart of Proposition A.3 for the multi-variate interpolation error.

Proposition A.5 For each $\ell = 1, \dots, d-1$ we assume that for any fixed $y \in Y_\ell$ the functions $F_\ell(\zeta_\ell, y)$ satisfy the following conditions:

- (a) $F_\ell(\cdot, y) \in \mathbf{H}^1(D_\delta)$ with $N(F_\ell, D_\delta) < \infty$ uniform in y ;
- (b) $F_\ell(\cdot, y)$ satisfies (A.9) with $a = 1$ and with certain C, b for all $y \in Y_\ell$.

Then the optimal choice $\mathfrak{h} := \frac{\log M}{M}$ of the step size yields the total pointwise error estimate

$$|\mathbf{E}_M(F, \mathfrak{h})| = |F(\zeta, y) - \mathbf{C}_M(F, \mathfrak{h})(\zeta)| \leq C \Lambda_M^{d-1} \frac{N_0(F, D_\delta)}{2\pi\delta} e^{-\pi\delta M / \log M} \quad (\text{A.20})$$

with $N_0(F, D_\delta) = \max_{\ell=1, \dots, d-1} N(F_\ell, D_\delta)$ and Λ_M defined by (A.19).

Proof. The proof is based on a multiple use of (A.12) and triangle inequality combined with the bound on Lebesgue's constant (cf. [16, Prop. 4.3] related to the multi-variate polynomial interpolation). ■

In FEM/BEM applications we often deal with functions $G(x)$, $x \in \mathbb{R}^d$, defined in a hypercube. Specifically, we consider a function $G : [0, 1]^d \rightarrow \mathbb{R}$ that is holomorphic in each variable $x_\ell \in (0, 1)$, $\ell = 1, \dots, d-1$ within $D_\phi(\delta) \supset (0, 1]$ but it may have singularities at the end-point $x_\ell = 0$ of $(0, 1]$. In this case the polynomial interpolation is no longer efficient, however, we show that the Sinc interpolation method can be successfully applied. Given $\alpha \geq 0$, let us introduce a possibly modified function $F : \mathbb{R}^d \rightarrow \mathbb{R}$ by

$$F(\zeta_1, \dots, \zeta_d) = \left\{ \prod_{\ell=1}^{d-1} \phi^\alpha(\zeta_\ell) \right\} G(\phi(\zeta_1), \dots, \phi(\zeta_d)).$$

Now we can prove the counterpart of Corollary A.4 for the error by multi-variate Sinc interpolation.

Corollary A.6 For each $\ell = 1, \dots, d-1$, we assume that for any fixed $y \in Y_\ell$ the functions $G_\ell(\cdot, y) : I_\ell \rightarrow \mathbb{R}$ together with their transformed counterparts $F_\ell(\zeta_\ell, y)$ satisfy the following conditions:

- (a) $G_\ell(\cdot, y)$ is holomorphic on $D_\phi(\delta)$, and $N(F_\ell, D_\delta) < \infty$;
- (b) $F_\ell(\cdot, y)$ satisfies (A.9) with $a = 1$ and with certain C, b for all $y \in Y_\ell$.

Then the optimal choice $\mathfrak{h} := \frac{\log M}{M}$ of the step size yields the pointwise error estimate

$$|G(x) - G_M(x)| \leq \prod_{\ell=1}^{d-1} x_\ell^{-\alpha} |\mathbf{E}_M(F, \mathfrak{h})(\phi^{-1}(x))| \quad (x \in (0, 1]^d = \phi(\mathbb{R}^d)), \quad (\text{A.21})$$

where $\mathbf{E}_M(F, \mathfrak{h})$ is bounded by (A.20) and the corresponding interpolant $G_M(x)$ is given by

$$G_M(x) := \sum_{k=-M}^M G(\phi(k\mathfrak{h}), x_d) \prod_{\ell=1}^{d-1} \phi_\ell^\alpha(k\mathfrak{h}) \cdot x_\ell^{-\alpha} S_{k, \mathfrak{h}}(\phi^{-1}(x_\ell)) \approx G(x). \quad (\text{A.22})$$

Proof. Conditions (a), (b) ensure that the corresponding requirements in Proposition A.5 are valid. Then the assertion is a direct consequence of Proposition A.5. ■

The respective Kronecker rank can be calculated as $r = (2M+1)^{d-1}$, where $M = \mathcal{O}(\delta^{-1} |\log \varepsilon| \cdot \log |\log \varepsilon|)$. Note that in the BEM applications we typically have $\delta = \frac{\pi}{|\log h|}$, where h is the mesh parameter in the FE approximation (see example in §A.6.5).

The extension to the case of parameter dependent functions (cf. Corollary A.4) is now straightforward.

A.3 Integral $\int_0^\infty e^{-\rho t} dt$ and Applications

We consider the Laplace integral transform

$$\frac{1}{\rho} = \int_0^\infty e^{-\rho \xi} d\xi, \quad \rho > 0, \quad (\text{A.23})$$

with $f(\xi) = e^{-\rho \xi}$. We assume that ρ varies in $[R_{\min}, R_{\max}]$, where $R_{\min} > 0$ is required, while $R_{\max} = \infty$ is included. Since R_{\min} can be changed by a simple scaling, in the following we use the choice $R_{\min} = 1$.

A.3.1 Standard Quadrature

The substitution $\xi = \log(1 + e^u)$ results into

$$\frac{1}{\rho} = \int_{\mathbb{R}} \frac{e^{-\rho \log(1+e^u)}}{1 + e^{-u}} du = \int_{\mathbb{R}} f_1(u) du, \quad f_1(u) := \frac{e^{-\rho \log(1+e^u)}}{1 + e^{-u}}. \quad (\text{A.24})$$

For this integral, we are able to apply the standard quadrature.

Lemma A.7 *Let $\delta < \pi/2$. Then the function from (A.24) satisfies $f_1 \in \mathbf{H}^1(D_\delta)$ with a uniform bound $N(f, D_\delta) \leq C(R_{\min}) < \infty$ for all $\rho \geq R_{\min} > 0$. In particular, the behaviour is*

$$|f_1(u)| \leq e^{-\rho \Re u} \text{ for } \Re u \geq 0, \quad |f_1(u)| \leq e^{-|\Re u|} \text{ for } \Re u \leq 0 \quad (u \in D_\delta).$$

Under the condition $\rho \geq R_{\min} = 1$, (A.4) holds with $C = b = 1$ and the choice $\mathfrak{h} = \sqrt{2\pi\delta/M}$ yields the quadrature result $T_M(f_1, \mathfrak{h})$ with the error estimate (A.7) uniformly for all $\rho \geq R_{\min}$.

Proof. a) The zeros of $1 + e^u$ are $\pm ik\pi$ ($k \in \mathbb{Z}_{\text{odd}}$) and therefore outside of D_δ . Hence, $f_1(u)$ is analytic in D_δ .

b) For $u = \xi + i\eta \in D_\delta$ we claim that $\Re \log(1 + e^u) \geq \max(0, \xi)$. For a proof we use

$$\Re \log(1 + e^u) = \frac{1}{2} \log(|1 + e^u|^2) = \frac{1}{2} \log(1 + 2e^\xi \cos(\eta) + e^{2\xi}) \geq \frac{1}{2} \log(e^{2\xi}) = \xi$$

in the case of $\xi \geq 0$ and $\frac{1}{2} \log(1 + 2e^\xi \cos(\eta) + e^{2\xi}) \geq \frac{1}{2} \log(1) = 0$ otherwise.

c) Part b) together with $1/|1 + e^{-u}| \leq 1$ proves the inequality $|f_1(u)| \leq e^{-\rho \Re u}$ for $\Re u \geq 0$.

d) For $\Re u \leq 0$ we use $\Re \log(1 + e^u) \geq 0$ (i.e. $|e^{-\rho \log(1+e^u)}| \leq 1$) and $1/|1 + e^{-u}| \leq 1/|e^{-u}| = e^{\Re u} = e^{-|\Re u|}$.

e) If $\rho \geq R_{\min}$, the norm $N(f, D_\delta)$ is bounded independently of ρ and hence the error estimate (A.7) uniform in ρ . \blacksquare

The finite sum $T_M(f_1, \mathfrak{h})$ can be interpreted as a exponentially convergent quadrature for the integral (A.24). Lemma A.7 ensures that the tolerance ε can be achieved with $M = \mathcal{O}(|\log \varepsilon|^2)$ uniformly in R_{\max} .

A.3.2 Improved Quadrature

In order to apply the improved estimate (A.10), we apply a second substitution $u = \sinh(w)$ and obtain the integral

$$\frac{1}{\rho} = \int_{\mathbb{R}} f_2(w) dw \quad \text{with } f_2(w) = \cosh(w) f_1(\sinh(w)) = \frac{\cosh(w)}{1 + e^{-\sinh(w)}} e^{-\rho \log(1+e^{\sinh(w)})}. \quad (\text{A.25})$$

The decay of f_2 on the real axis is

$$f_2(w) \approx \frac{1}{2} e^{w - \frac{\rho}{2} e^w} \quad \text{as } w \rightarrow \infty, \quad f_2(w) \approx \frac{1}{2} e^{|w| - \frac{1}{2} e^{|w|}} \quad \text{as } w \rightarrow -\infty,$$

corresponding to $C = \frac{1}{2}$, $b = \min\{1, \rho\}/2$, $a = 1$ in (A.9). A particular difficulty is the behaviour of $f_2(w)$ in $w \in D_\delta$ for $\Re w < 0$, since the exponent $-\rho \log(1 + e^{\sinh(w)})$ may become positive. This effect requires the use of an ρ -dependent δ in the next lemma (note that the choice of δ does not change the quadrature, but only effects the error bound).

Conventionally, we denote $w = x + iy$, $x, y \in \mathbb{R}$. Note that $\sinh(w) = X + iY$ with

$$X = \sinh(x) \cos(y), \quad Y = \cosh(x) \sin(y) \quad \text{for } w \in D_\delta.$$

Given $\delta < \pi/2$, we introduce the constant $x_1 = x_1(\delta) = \text{arsinh}(\frac{1}{\cos \delta}) > 0$. Now set

$$A := \left(1 + \frac{\pi^2}{4} + \log^2(3\rho)\right) / 2, \quad B := \frac{\pi^2}{4} / \left(A + \sqrt{A^2 + \frac{\pi^2}{4}}\right),$$

and define

$$\delta(\rho) := \arcsin(\sqrt{B}), \quad x_0(\rho) := -\text{arsinh}\left(\frac{\log(3\rho)}{\cos(\delta(\rho))}\right) = -\mathcal{O}(\log \log(3\rho)). \quad (\text{A.26})$$

Lemma A.8 Let $\delta < \pi/2$. Then the following estimates of f_2 from (A.25) in D_δ cover all values of $x = \Re w$:

$$|f_2(w)| \leq \frac{\cosh(x)}{1 - e^{-X}} \exp(-\rho \log(e^X - 1)) \Big|_{X=\sinh(x) \cos(y)} \lesssim \frac{1}{2} e^{x - \rho \sinh(x) \cos(y)} \quad (\text{A.27a})$$

$$\lesssim \frac{1}{2} e^{x - \rho \frac{\cos(\delta)}{2} e^{|x|}} \quad \text{for } w \in D_\delta, x_1 < x \rightarrow +\infty,$$

$$|f_2(w)| \leq \sqrt{2} \quad \text{for } w \in D_\delta, 0 \leq x \leq x_1(\delta) \text{ and } \delta \leq 0.93 < \pi/2, \quad (\text{A.27b})$$

$$|f_2(w)| \leq \frac{1}{2} e^{x + \sinh(x) \cos(y)} \leq \frac{1}{2} e^{x - \frac{\cos(\delta)}{2} e^{|x|}} \quad \text{for } w \in D_\delta, x_0(\rho) \leq x \leq 0 \quad (\text{A.27c})$$

$$\text{with } 0 > x_0(\rho) = -\mathcal{O}(\log \log(3\rho)) \text{ and } \delta \leq \delta(\rho) = \mathcal{O}\left(\frac{\pi/2}{\log(3\rho)}\right),$$

$$|f_2(w)| \leq \frac{\sqrt{3}}{1 - 3^{-\cos(\delta)/\cos(\delta(\rho))}} \frac{1}{2} e^{-x + \sinh(x) \cos(y)} \leq C e^{|x| - \frac{\cos(\delta)}{2} e^{|x|}} \quad (\text{A.27d})$$

$$\text{for } w \in D_\delta, 0 > x_0(\rho) \geq x \rightarrow -\infty$$

with $x_0(\rho)$ and $\delta(\rho)$ described in (A.26). Hence $f_2 \in \mathbf{H}^1(D_\delta)$, while $\delta \in (0, \delta(\rho)]$ for $\rho \geq 1$ leads to the norm $N(f_2, D_\delta) < \infty$ independent of ρ . If ρ ranges in $[1, R]$, the choice $\delta = \delta(R)$, $a = 1$, $b = 1/2$ in Proposition A.1 implies the uniform quadrature error estimate

$$\|\eta_M(f_2, \mathbf{h})\| \leq C e^{-\frac{2\pi\delta(R)M}{\log(2\pi M)}} \lesssim C e^{-\frac{\pi^2 M}{\log(3R) \log(\pi^2 M)}}. \quad (\text{A.27e})$$

Proof. a) Again, the zeros of $1 + e^{\sinh(w)}$ are outside of D_δ , so that $f_2(u)$ is analytic in D_δ .

b) The assumption $x > x_1$ ensures $X > 1$. The same argument as in Lemma A.7 shows $\Re \log(1 + e^{X+iY}) = \frac{1}{2} \log(1 + 2e^X \cos(Y) + e^{2X})$. The worst case is $\cos(Y) = -1$ (which happens only for $x \geq x_0(\delta) > 0$) and yields $\Re \log(1 + e^{X+iY}) \geq \log(e^X - 1)$. This proves (A.27a).

c) $x \leq x_1$ implies $\sinh x \leq \frac{1}{\cos \delta}$. From $\cosh(x) = \sqrt{1 + \sinh^2 x}$ one concludes that $Y = \cosh(x) \sin(y) \leq \tan \delta \sqrt{1 + \cos^2 \delta}$. The restriction $\delta \leq 0.93$ guarantees $Y \in (-\pi/2, \pi/2)$ and thus $\Re e^{X+iY} > 0$. As a consequence $\Re(-\rho \log(1 + e^{\sinh(w)})) < 0$ and $|1 + e^{-\sinh(w)}| > 1$ show $|f_2(w)| \leq |\cosh(w)| \leq \sqrt{1 + \cos^{-2} \delta} \leq \sqrt{2}$.

d) By the definition (A.26) we have that¹ $\cosh(x) \leq \pi / (2 \sin(\delta(\rho)))$ and $|Y| \leq \frac{\pi}{2}$ holds for $x \geq x_0(\rho)$. This ensures $\Re(-\rho \log(1 + e^{\sinh(w)})) < 0$ and therefore

$$|f_2(w)| \leq \left| \frac{\cosh(w)}{1 + e^{-\sinh(w)}} \right| \leq \frac{e^x}{2(1 + e^{-X})} \leq \frac{1}{2} e^{x + \sinh(x) \cos(y)} \leq \frac{1}{2} e^{x + \sinh(x) \cos(\delta)}.$$

e) For $x \leq x_0(\rho)$ we have $X \leq -\frac{\log(3\rho)}{\cos(\delta(\rho))} \cos y \leq -\log(3\rho)$, so that as in Part b)

$$\begin{aligned} \Re\left(-\rho \log(1 + e^{\sinh(w)})\right) &= -\frac{\rho}{2} \log(1 + 2e^X \cos(Y) + e^{2X}) \\ &\leq -\frac{\rho}{2} \log(1 - 2e^X) \leq -\frac{\rho}{2} \log\left(1 - \frac{2}{3\rho}\right). \end{aligned}$$

The function $-\frac{\rho}{2} \log(1 - \frac{2}{3\rho})$ decreases with $\rho \geq 1$ leading to $-\frac{\rho}{2} \log(1 - \frac{2}{3\rho}) \leq \frac{\log 3}{2}$ and the bound $\exp(\frac{\log 3}{2}) = \sqrt{3}$ in (A.27d). Together with $1/|1 + e^{-\sinh(w)}| \leq 1/|1 - e^{-X}| = e^X / (1 - e^{-X}) \leq e^X / (1 - e^{\sinh(x_0(\rho)) \cos(\delta)}) = e^X / (1 - e^{-\frac{\log(3\rho)}{\cos(\delta(\rho))} \cos(\delta)}) = e^X / (1 - 3^{-\cos(\delta)/\cos(\delta(\rho))})$ and $e^X = e^{\sinh(x) \cos(\delta)}$ we obtain (A.27d). ■

Lemma A.8 ensures that the tolerance ε can be achieved with $M = \mathcal{O}(\log R_{max} |\log \varepsilon|)$ uniformly in $\rho \in [1, R_{max}]$.

¹With $x_0 = -\operatorname{arsinh}\left(\frac{\log(3\rho)}{\cos(\delta)}\right)$ it follows that $\cosh(x_0) = \sqrt{1 + \sinh^2(x_0)} = \sqrt{1 + \left(\frac{\log(3\rho)}{\cos(\delta)}\right)^2}$. Then the condition $|Y| = |\cosh(x_0) \sin(\delta)| \leq \frac{\pi}{2}$ is equivalent to $\sin^2(\delta) + \tan^2(\delta) \log^2(3\rho) \leq \frac{\pi^2}{4}$. Due to $\tan^2(\delta) = \frac{\sin^2(\delta)}{1 - \sin^2(\delta)}$, we obtain a quadratic equation in $\sin^2(\delta)$, whose solution is given by B .

A.3.3 Numerics

In the numerical example below, we apply the quadrature to f_2 using the simplified choice $\mathfrak{h} = C_{int} \frac{\log M}{M}$. All computations in this paper were performed in single precision arithmetic in MATLAB 5.3(R11).

Figures A.1 and A.2 illustrate the exponential convergence (semi-logarithmic scale) in the intervals $\rho \in [1, 1000]$ and $\rho \in [1, 18000]$, respectively. The numerical results indicate an almost linear dependence of the quadrature error on ρ , i.e., instead of the slower exponential factor in $e^{-\frac{\pi^2}{\log(3\rho)}M/\log(\frac{\pi^2 M}{\log(3\rho)})}$ predicted by (A.27e), we observe the behaviour $\mathcal{O}(\rho e^{-cM}) = \mathcal{O}(e^{-cM+\log \rho})$. If this would be true, we obtain a desired error bound ε by $M = \mathcal{O}(\log \frac{1}{\varepsilon} + \log \rho)$.

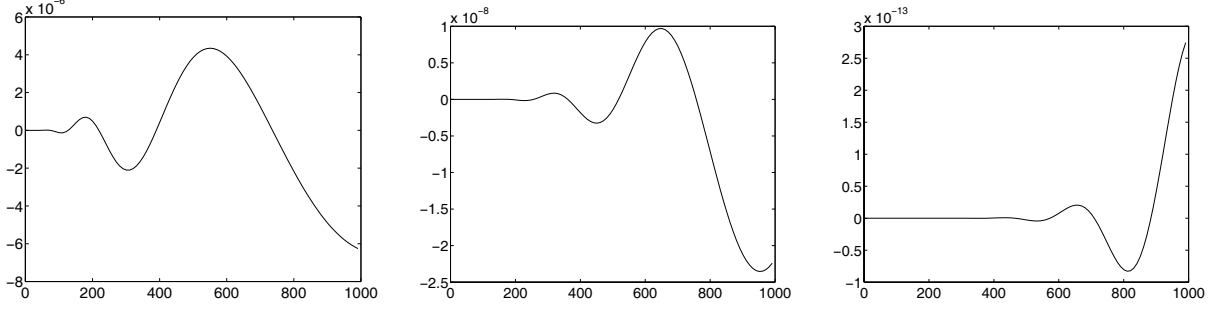


Figure A.1: The absolute quadrature error for (A.25) for $1 \leq \rho \leq 10^3$ with $M = 16$ (left), $M = 32$ (middle), $M = 64$ (right).

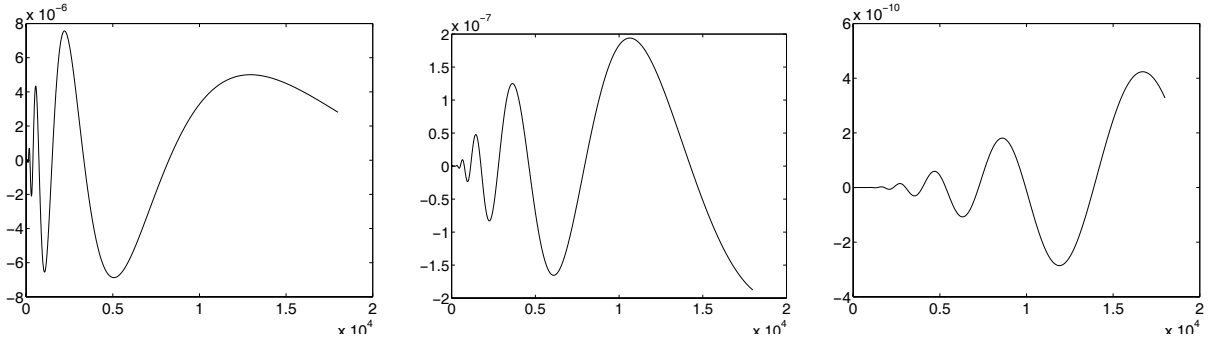


Figure A.2: The absolute quadrature error for (A.25) for $1 \leq r \leq 1.8 \cdot 10^4$ with $M = 16$ (left), $M = 32$ (middle), $M = 64$ (right).

Both the standard and the modified quadrature (cf. Lemma A.8) can be applied to matrices with spectrum in $[1, R]$. In the next examples, we illustrate the different quadratures approximating the inverse of the finite difference Laplacian $(-\Delta_h)^{-1}$ in \mathbb{R}^d . The first table represents the convergence of the standard quadrature for (A.23) of the order $e^{-c\sqrt{M}}$.

$e^{-c\sqrt{M}}$ - approximation to $(-\Delta_h)^{-1}$ in $[0, 1]^d$ with $N = n^d$, $n = 128$							
M	4	9	16	25	36	49	64
$d = 1$	2.1_{10-1}	1.8_{10-2}	5.6_{10-3}	1.5_{10-4}	8.5_{10-6}	3.6_{10-7}	1.8_{10-8}
$d = 2$	5.9_{10-2}	2.3_{10-3}	1.6_{10-3}	2.4_{10-5}	2.4_{10-6}	2.6_{10-7}	8.4_{10-9}
$d = 3$	3.1_{10-2}	2.8_{10-3}	3.6_{10-4}	1.5_{10-5}	1.2_{10-6}	1.3_{10-8}	1.6_{10-9}
$d = 4$	1.8_{10-1}	2.0_{10-2}	1.4_{10-4}	3.1_{10-6}	5.2_{10-7}	1.9_{10-8}	1.8_{10-10}

The next table gives the error for the quadrature (A.23) of the order $e^{-cM/\log[\text{cond}(A)]}$. We observe that the latter approximation shows faster exponential convergence for larger M , while the first version is more preferable for smaller M .

$e^{-cM/\log[\text{cond}(A)]}$ - approximation to $(-\Delta_h)^{-1}$ in $[0, 1]^d$ with $N = n^d$, $n = 128$							
M	4	9	16	25	36	49	64
$d = 1$	6.3 ₁₀ -1	7.9 ₁₀ -1	2.0 ₁₀ -1	2.0 ₁₀ -3	7.6 ₁₀ -6	7.9 ₁₀ -9	6.5 ₁₀ -12
$d = 2$	3.5 ₁₀ -2	4.3 ₁₀ -1	2.2 ₁₀ -1	1.9 ₁₀ -4	2.4 ₁₀ -6	4.8 ₁₀ -9	7.8 ₁₀ -12
$d = 3$	3.1 ₁₀ -1	1.5 ₁₀ -2	5.2 ₁₀ -4	1.9 ₁₀ -4	2.3 ₁₀ -7	2.1 ₁₀ -9	3.0 ₁₀ -13
$d = 4$	1.0 ₁₀ -2	1.0 ₁₀ -1	5.6 ₁₀ -4	1.0 ₁₀ -4	3.6 ₁₀ -6	6.4 ₁₀ -11	1.1 ₁₀ -13

The last table shows that the approximation error depends only weakly on n , confirming the theoretical predictions.

Approximation to $(-\Delta_h)^{-1}$ in $[0, 1]^d$ with $d = 3$, $M = 25$						
n	4	8	16	32	64	128
ε	2.5 ₁₀ -8	7.7 ₁₀ -8	4.2 ₁₀ -8	5.7 ₁₀ -7	8.5 ₁₀ -6	3.5 ₁₀ -6

A.4 Integral $\int_{-\infty}^{\infty} e^{-\rho^2 t^2} dt$

The integrand of the Gaussian integral

$$\frac{1}{\rho} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\rho^2 t^2} dt \quad (\text{A.28})$$

shows a fast decay if ρ is not too small. However, the results of §A.2.2 do not yield uniform error bounds with respect to $\rho \geq 1$. The reason is that $t = x + iy \in D_\delta$ results in $|e^{-\rho^2 t^2}| = \exp(-\rho^2(x^2 - \delta^2))$. For $-\delta \leq x \leq \delta$, the exponent is positive and therefore $N(f, D_\delta) \approx \mathcal{O}(e^{\rho^2})$ circumvents reasonable error estimates.

The same difficulty arises when the substitution $t = \sinh(w)$ is used to get the twice exponential decay of the integrand:

$$\frac{1}{\rho} = \int_{-\infty}^{\infty} f_3(w) dw \quad \text{with } f_3(w) = \cosh(w) \exp(-\rho^2 \sinh^2(w)). \quad (\text{A.29})$$

Lemma A.9 *For each $\rho > 0$, the symmetric $(M + 1)$ -point quadrature for the integral (A.29) converges exponentially (cf. (A.10)) with constants C , s depending on ρ .*

Proof. Clearly, for each $\rho > 0$, the function $f_3(w)$ defined above satisfies all the conditions in Proposition A.1. Thus, we choose $\mathfrak{h} = C_{\text{int}} \frac{\log M}{M}$ and obtain the exponential convergence as indicated in (A.10), where the constants C and s depend on the parameter ρ . ■

We conclude that the symmetric quadrature for the integral (A.29) is acceptable only in an interval $\rho \in I_{\text{ref}} = [R_{\text{min}}, R_{\text{max}}]$ with some fixed $R_{\text{min}} < 1$, $R_{\text{max}} > 1$ of order $\mathcal{O}(1)$ (see numerics below). Hence, an application of this quadrature in the larger range of the parameter $\rho \in [R_1, R_2]$ requires the proper re-scaling, thus, in general, we need p different quadratures, when $R_2/R_1 \approx Q^p$ with $Q = R_{\text{max}}/R_{\text{min}}$.

The following example illustrates the quadrature applied to f_3 . Fig. A.3 shows that stable convergence holds for the range $[0.2, 10]$ of ρ . However, for a fixed value ρ (considered as a constant), the quadrature applies and yields an accuracy ε with $M = \mathcal{O}(\log^2 \frac{1}{\varepsilon})$ (case of (A.28)) or $M = \mathcal{O}(\log \frac{1}{\varepsilon} \cdot \log \log \frac{1}{\varepsilon})$ (case of (A.29)).

To obtain robustness with respect to the parameter ρ , we propose another, nonsymmetric quadrature. For this purpose we rewrite the integral (A.28) as $\frac{1}{\rho} = \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-\rho^2 t^2} dt$ and then, similar to the previous section, substitute $t = \log(1 + e^u)$ and $u = \sinh(w)$:

$$\frac{1}{\rho} = \int_{\mathbb{R}} F(u) du = \int_{\mathbb{R}} f(w) dw \quad \text{with} \quad (\text{A.30})$$

$$F(u) := \frac{2}{\sqrt{\pi}} \frac{e^{-\rho^2 \log^2(1+e^u)}}{1 + e^{-u}}, \quad f(w) = \cosh(w) F(\sinh(w)). \quad (\text{A.31})$$

Lemma A.10 *Let $\delta < \pi/2$, then for the function f from (A.31) we have $f \in \mathbf{H}^1(D_\delta)$, and, in addition, the condition (A.9) is satisfied with $a = 1$. Let $\rho \geq 1$, then the improved $2M + 1$ -point quadrature (cf. Proposition A.1) with the choice $\delta(\rho) = \frac{\pi}{C + \log(\rho)}$ allows the error bound*

$$|\eta_M(f, \mathfrak{h})| \leq C_1 \exp\left(-\frac{\pi^2 M}{(C + \log(\rho)) \log M}\right). \quad (\text{A.32})$$

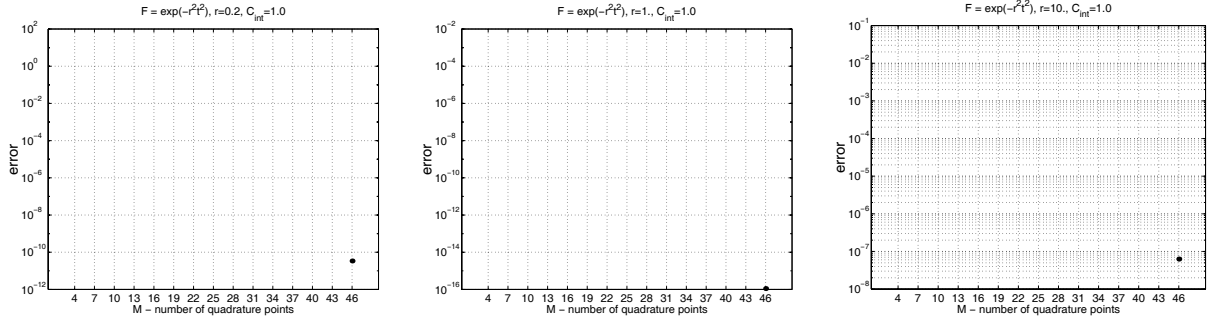


Figure A.3: Approximation to the Gaussian integral with $\rho \in \{0.2, 1, 10\}$, $C_{\text{int}} = 1.0$

Proof. It is easy to check that f is holomorphic in D_δ and $N(f, D_\delta) < \infty$ uniform in ρ . Further analysis is similar to that in Lemma A.8. \blacksquare

Numerical examples for this quadrature with values $\rho \in [1, R]$, $R \leq 5000$, are presented in Fig. A.4. Again, we observe almost linear error growth in ρ . Similar results are observed in the case $R > 5000$.

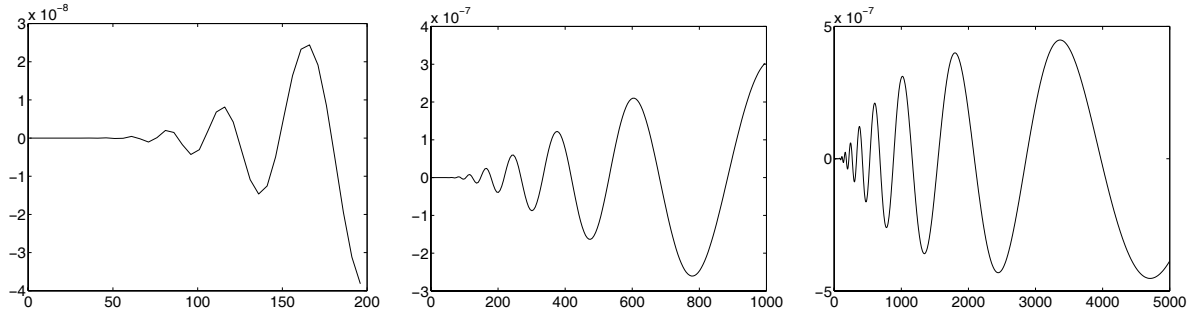


Figure A.4: The absolute quadrature error for $M = 64$ with $R = 200$ (left), $R = 1000$ (middle), $R = 5000$ (right).

A.5 Gaussian Charge Distribution

In some cases the precise scaling of the argument in the Newton potential, $1 \leq |x - y| \leq R$, might not be possible as in the following example. In fact, the supports of two Gaussian “basis functions” to be considered always have an overlap. Therefore, we are going to compare the accuracy of our quadrature from the previous section with that one derived for the explicit expression obtained by analytic spatial integration (often, this integration can be performed only numerically).

The energy of interaction between two spherical Gaussian distributions of unit charges centred at $P, P' \in \mathbb{R}^3$ is given by

$$V_{pp'} = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_p(\mathbf{x})\rho_{p'}(\mathbf{y})}{\|\mathbf{x} - \mathbf{y}\|} d\mathbf{x}d\mathbf{y},$$

$$\rho_p(\mathbf{x}) = (p/\pi)^{3/2} \exp(-p\|\mathbf{x} - P\|^2), \quad \rho_{p'}(\mathbf{y}) = (p'/\pi)^{3/2} \exp(-p'\|\mathbf{y} - P'\|^2),$$

(cf. [25]), where $p, p' \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$. The exact calculation using the incomplete gamma function F_0 , leads to

$$V_{pp'} = \sqrt{\frac{4\alpha}{\pi}} F_0(\alpha\|P - P'\|^2), \quad \alpha = \frac{pp'}{p + p'} (= 1) \quad \text{with } F_0(x) = \int_0^1 e^{-xt^2} dt \equiv \sqrt{\frac{\pi}{4x}} \operatorname{erf}(\sqrt{x}).$$

We use the integral representation on $[0, \infty)$,

$$F_0(x) = \int_0^\infty \exp\left(-x \frac{u^2}{1+u^2}\right) \frac{du}{(1+u^2)^{3/2}},$$

and derive the standard quadrature that converges as $\mathcal{O}(\exp(-s\sqrt{M}))$. Fig. A.5 gives numerical results for this quadrature with different $M = 25, 64, 121$ and for $x \in [0, R]$, $R = 100$. This example confirms that the exponential convergence is robust in R . Another important observation is that the error bound for the quadrature applied to the integrated in space variables expression (cf. the case $M = 64$) remains practically the same as that for Gaussian integral (cf. Fig. A.4).

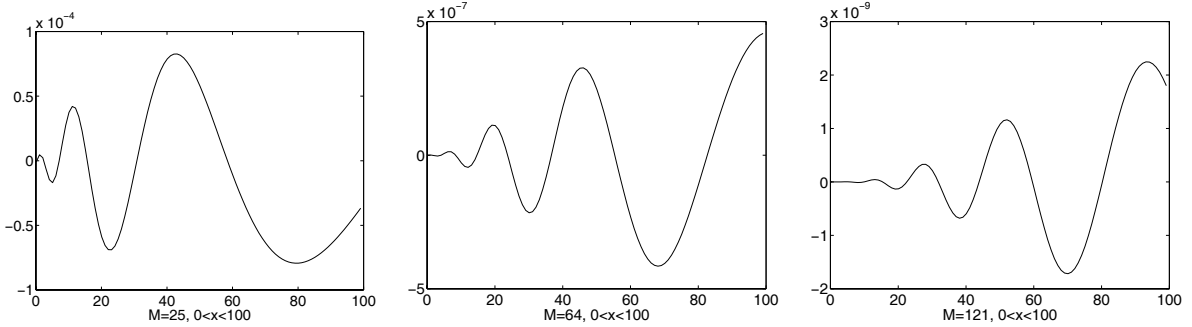


Figure A.5: The absolute quadrature error for $M = 25$ (left), $M = 64$ (middle), $M = 121$ (right).

A.6 On the Separable Approximation of Multi-Variate Functions

As a by-product, the sinc quadrature applied to the integrals (A.25) and (A.29) provides a separable approximation to the multi-variate functions

$$\frac{1}{x_1 + \dots + x_d} \quad \text{and} \quad \frac{1}{\sqrt{x_1^2 + \dots + x_d^2}} \quad (x_i > 0, i = 1, \dots, d).$$

In the case of $\frac{1}{x_1 + \dots + x_d}$, Lemma A.8 shows that the separation rank $k = 2M + 1$ depends only linear-logarithmically on both the tolerance $\varepsilon > 0$ and the upper bound R of $\rho = x_1 + \dots + x_d$. In the case of $1/\sqrt{x_1^2 + \dots + x_d^2}$, the dependence on ε and $\rho = \sqrt{x_1^2 + \dots + x_d^2}$ is similar, hence in both cases there is *no dependence on the dimension* d .

In the first case of $\frac{1}{x_1 + \dots + x_d}$, the estimate (A.27e) implies that an approximation of accuracy ε is obtainable with

$$M \leq \mathcal{O}\left(\log\left(\frac{1}{\varepsilon}\right) \cdot \log R\right), \quad (\text{A.33})$$

provided that $1 \leq x_1 + \dots + x_d \leq R$, which can be achieved by a proper scaling. The numerical results even support the better estimate $M \leq \mathcal{O}\left(\log\left(\frac{1}{\varepsilon}\right) + \log R\right)$ (cf. Fig. A.1, A.2). In the second case of $1/\sqrt{x_1^2 + \dots + x_d^2}$, we apply (A.32) and again obtain the bound (A.33), while our numerical results manifest a rather stable behaviour of the quadrature error with respect to R (cf. Fig. A.4).

A.6.1 Example: Newton Potential

Our separable representation to the function $\rho = 1/\sqrt{x_1^2 + \dots + x_d^2}$ directly results in a low Kronecker rank tensor-product approximation (cf. [19]) to the classical Newton potential $(\mathcal{A}u)(x) := \int_\Omega \frac{u(y)}{|x-y|} dy$ defined by the kernel function

$$\frac{1}{|x-y|} = \frac{1}{\sqrt{(x_1 - y_1)^2 + \dots + (x_d - y_d)^2}}, \quad x, y \in \mathbb{R}^d.$$

Indeed, with the Kronecker rank $r = 2M + 1$, where M satisfies (A.33), we readily obtain the separable approximation of accuracy ε ,

$$\frac{1}{|x - y|} \approx \sum_{k=1}^r C_k f_k^1(|x_1 - y_1|) \cdots f_k^d(|x_d - y_d|),$$

provided that $1 \leq |x - y| \leq R$. Note that the Kronecker rank r does not depend on d . In FEM/BEM applications by low order elements, one has $R = \mathcal{O}(h^{-1})$ (with a proper scaling), where h is the mesh parameter.

A.6.2 Example: $\log(x + y)$

In boundary element methods (BEM), one is interested in a low separation rank representation to the kernel function $s(x, y) = \log(x + y)$, $x \in [0, 1], y \in [h, 1]$ with some small parameter $h > 0$ (the mesh-size) (cf. [17, 19]). A representation like

$$\frac{1}{x + y} = \sum_{m=1}^k \Phi_m(x) \Psi_m(y) + \delta_k \quad \text{with } |\delta_k| \leq \varepsilon \quad (\text{A.34})$$

can be constructed by means of the quadrature applied to the integral (A.25) with $\rho = x + y$ and $k = 2M + 1$. Let ψ_m be the antiderivatives of Ψ_m . Integration of (A.34) yields

$$\begin{aligned} \log(x + y) &= \int_{1-x}^y \frac{dt}{x + t} = \int_{1-x}^y \left(\sum_{m=1}^k \Phi_m(x) \Psi_m(t) + \delta_k \right) dt \\ &= \sum_{m=1}^k \Phi_m(x) [\psi_m(y) - \psi_m(1 - x)] + S_k \\ &= \Phi_0(x) + \sum_{m=1}^k \Phi_m(x) \psi_m(y) + S_k \end{aligned}$$

with $\Phi_0(x) = - \sum_{m=1}^k \Phi_m(x) \psi_m(1 - x)$ and $|S_k| = \left| \int_{1-x}^y \delta_k dt \right| \leq \varepsilon$. This resulting representation of $\log(x + y)$ has the separation rank $k + 1$ and the same accuracy ε as (A.34).

In the next example we illustrate how to apply Corollary A.4.

A.6.3 Example: $1/(x + y)$

Interesting examples like $g(x, y) = \frac{1}{x+y}$ ($x, y \in \mathbb{R}_{>0}$) have a singularity at $x = -y$. In the following, we discuss the choice of δ such that g is holomorphic in the strip $D_\phi(\delta)$ (see §A.2.4). Solving the equation $\cosh(z) = -1/y$, we find the singularity points $z_m = \pm \operatorname{arccosh}(1/y) + (1 \pm 2m)\pi$, $m = 0, 1, \dots$. Now by solving the equation $\sinh(\zeta_m) = z_m$ we find $\zeta_m = \log(z_m + \sqrt{1 + z_m^2})$, $\Im m(\zeta_m) = \arg(z_m + \sqrt{1 + z_m^2})$. Suppose that $y \rightarrow 0$, hence $\min_{m=0,1,\dots} \{\Im m(\zeta_m)\} \geq C_0 \pi / \operatorname{arccosh}(1/y)$ is achieved with $m = 0$, where C_0 does not depend on y (we actually have $C_0 \approx 1$). Taking $m = 0$ and suppressing all the symmetric images, we come to the conclusion that $g(\phi(\zeta), y)$ is holomorphic in D_{δ_0} , where $\tan(\delta_0) = C_0 \pi / \operatorname{arccosh}(1/y)$ depends on $y \in Y$.

In the following, we fix $y > 0$ and, first, apply Corollary A.4 with

$$\delta = \delta_0 = \arctan(C_0 \pi / \operatorname{arccosh}(1/y)),$$

and next (A.17) to obtain

$$|g(x, y) - g_M(x, y)| \leq C |x|^{-\alpha} \frac{N(f, D_\delta)}{2\pi\delta_0} e^{-\pi\delta_0 M / \log M}. \quad (\text{A.35})$$

In BEM applications we have $x, y \geq h \rightarrow 0$, where $h > 0$ is the mesh parameter, so that $\delta_0 \approx \frac{\pi}{|\log h/2|}$ depends only mildly on h . Then (A.35) leads to

$$|g(x, y) - g_M(x, y)| \leq C |h|^{-\alpha} \frac{N(f, D_\delta) |\log h/2|}{2\pi^2} e^{-\pi^2 M / (|\log h/2| \log M)}. \quad (\text{A.36})$$

Hence, the tolerance ε can be achieved with $M = O(|\log h| |\log \varepsilon|)$ and with the Kronecker rank $r = 2M + 1$.

Note that the error estimate for the function $1/(x + y)$, $x, y \in [1, R]$, can be derived from (A.36) by the substitution $h = 1/R$.

Similar to the previous example, our Sinc approximation can be applied to the functions like $g(x, y) = \log(x + y)$, $g(x, y) = (x^2 + y^2) \log(x + y)$ (biharmonic kernel function) and to $H_0^{(1)}(x + y)$ (2D Helmholtz kernel).

A.6.4 Example: $\exp(-xy)$

In our next example we discuss the function $g(x, y) = \exp(-xy)$, $x \geq 0$, $y \in [1, \lambda_{max}] \subset \mathbb{R}_{>0}$ (see the proof of Lemma 3.4).

We consider the auxiliary function $f(x, y) = \frac{x}{1+x} \exp(-xy)$, $x \in \mathbb{R}_+$. This function satisfies all the conditions of [22, Example 4.2.11] with $\alpha = \beta = 1$ (see also §2.4.2 in [7]), and hence, with the corresponding choice of interpolation points $x_k := \log[e^{k\mathfrak{h}} + \sqrt{1 + e^{2k\mathfrak{h}}}] \in \mathbb{R}_+$, it can be approximated by

$$\sup_{0 < x < \infty} \left| f(x, y) - \sum_{k=-M}^M f(x_k, y) S(k, \mathfrak{h}) (\log\{\sinh(x)\}) \right| \leq CM^{1/2} e^{-cM^{1/2}}$$

with exponential convergence, where $S(k, \mathfrak{h})$ is the k -th Sinc function (cf. (A.2)) and $\mathfrak{h} = C_1/M^{1/2}$. The corresponding error estimate for the initial function $g(x, y)$ is given by (A.17) with $\alpha = 1$.

A.6.5 Example: Helmholtz Kernel in \mathbb{R}^d

We consider the singularity function corresponding to the Helmholtz operator in \mathbb{R}^d , $d \geq 2$. Specifically, given $\kappa \in \mathbb{R}$, define the Helmholtz kernel function

$$g(x, y) := \frac{\cos(\kappa|x - y|)}{|x - y|} = \Re e \frac{e^{i\kappa|x - y|}}{|x - y|} \quad \text{for } (x, y) \in [0, 1]^d \times [0, 1]^d$$

in Cartesian coordinates $x = (x_1, \dots, x_d), y = (y_1, \dots, y_d) \in \mathbb{R}^d$. We do not have the proper integral representation to this kernel, however, the Sinc method still provides an opportunity to construct a corresponding hierarchical tensor-product approximation. We mention that an analysis of polynomial approximations to the Helmholtz kernel function is presented in [16] in the context of the hierarchical matrix technique with standard admissibility criteria. The Sinc approximation below can be applied in the case of a weakly admissible block (cf. [17]) with respect to the transformed variables ζ_1, \dots, ζ_d .

For $(\zeta_1, \dots, \zeta_d) \in [0, 1]^d$, define

$$G(\zeta_1, \dots, \zeta_d) := g(x, y), \quad \zeta_\ell = |x_\ell - y_\ell|, \quad \ell = 1, \dots, d,$$

which implies

$$G(\zeta_1, \dots, \zeta_d) := \cos \left(\kappa \sqrt{\zeta_1^2 + \dots + \zeta_d^2} \right) / \sqrt{\zeta_1^2 + \dots + \zeta_d^2}.$$

We approximate the modified function

$$F(\zeta_1, \dots, \zeta_d) := (\zeta_1 \dots \zeta_{d-1})^{\alpha_0} G(\zeta_1, \dots, \zeta_d), \quad 0 < \alpha_0 < 1,$$

on the domain $\Omega_1 := [h, 1]^{d-1} \times [0, 1]$, where $h > 0$ is a small parameter, which can be associated with the mesh-size.

Now we apply Corollary A.6 with $\delta = 1/|\log h|$ to construct the corresponding interpolant $G_M(x)$ and obtain

$$\begin{aligned} |G(x) - G_M(x)| &\leq \prod_{\ell=1}^{d-1} x_\ell^{-\alpha_0} |\mathbf{E}_M(F, \mathfrak{h})(\phi^{-1}(x))| \\ &\leq Ch^{\alpha_0(1-d)} |\log h| \Lambda_M^{d-1} N_0(F, D_\delta) e^{-\pi M / (|\log h| \log M)} \end{aligned} \quad (\text{A.37})$$

with $\zeta \in (0, 1]^d$, where the corresponding interpolant $G_M(x)$ is given by (A.22).

Note that in this example $N_0(F, D_\delta) = \mathcal{O}(e^\kappa)$, while the Kronecker rank is given by $r = (2M + 1)^{d-1}$. Clearly, for the large parameter κ the bound (A.37) does not provide a satisfactory complexity.

The intrinsic alternative to the multi-variate Sinc interpolation would be the following two-step method: First, compute the polynomial interpolation to the entire function $\cos(\kappa \sqrt{\zeta_1^2 + \dots + \zeta_d^2})$ with $r = \mathcal{O}((\log R |\log \varepsilon|)^{d-1})$ and then multiply it with the HKT representation to the Newton potential as above. However, in this case the resulting Kronecker rank (obtained as a product of the corresponding ranks for the elementary factors) seems to be larger than for the Sinc interpolation method.

A.6.6 On the Optimality of Sinc-Quadrature Approximations and Generalisations

Our approximation theory to the Laplace and Gauss integrals is based on efficient Sinc quadratures which allow exponential convergence². Concerning this, we address the following questions:

- (a) How close are our Sinc quadratures to the optimal approximation by exponential sums which include the set of functions $\{\omega_\nu e^{-t_\nu x}\}$ or $\{\omega_\nu e^{-t_\nu x^2}\}$, $\omega_\nu, t_\nu \in \mathbb{R}$?
- (b) How can the approximation by exponential sums be applied to a more general class of functions which is no longer given by an explicit integral representation?

To answer these questions, we recall the basics of approximation theory by exponential sums for the class of so-called completely monotone functions f (cf. [2]). The existence result is based on the fundamental *Big Bernstein Theorem*: If f is completely monotone for $x \geq 0$, i.e.,

$$(-1)^n f^{(n)}(x) \geq 0, \quad n \geq 0, \quad x \geq 0,$$

then it is the restriction to the half-axis of the Laplace transform of a negative measure:

$$f(z) = \int_{\mathbb{R}_+} e^{-tz} d\mu(t).$$

For $n \geq 1$, consider the set E_n^0 of exponential sums and the extended set E_n :

$$\begin{aligned} E_n^0 &:= \left\{ u : u = \sum_{\nu=1}^n \omega_\nu e^{-t_\nu x}, \omega_\nu, t_\nu \in \mathbb{R} \right\}, \\ E_n &:= \left\{ u : u = \sum_{\nu=1}^l p_\nu(x) e^{-t_\nu x}, p_\nu \in \Pi_n, t_\nu \in \mathbb{R}, k := \sum_{\nu=1}^l (1 + \text{degree}(p_\nu)) \leq n \right\}, \end{aligned}$$

where Π_n is the set of polynomials of degree at most n . Now one can address the problem of finding the best approximation to f over the set E_n characterised by the best approximation error $d(f, E_n) := \inf_{v \in E_n} \|f - v\|_\infty$.

We recall the complete elliptic integral of the first kind with modulus κ ,

$$\mathbf{K}(\kappa) = \int_0^1 \frac{1}{\sqrt{(1-t^2)(1-\kappa^2 t^2)}}, \quad 0 < \kappa < 1,$$

and define $\mathbf{K}'(\kappa) = \mathbf{K}(\kappa')$ by $\kappa^2 + (\kappa')^2 = 1$. The following theorem is presented in [2].

²Note that generalised Gaussian quadratures for certain improper integrals were described in [26].

Theorem A.11 Assume that f is completely monotone and analytic for $\Re z > 0$, and let $0 < a < b$. Then³ for the uniform approximation on the interval $[a, b]$,

$$\lim_{n \rightarrow \infty} d(f, E_n)^{1/n} \leq \frac{1}{\omega^2}, \quad \text{where } \omega = \exp \frac{\pi \mathbf{K}(\kappa)}{\mathbf{K}'(\kappa)} \quad \text{with } \kappa = \frac{a}{b}.$$

In our particular case we have $\kappa = 1/R$. Applying the asymptotics of the complete elliptic integrals for $\kappa \rightarrow 0$ and for $\kappa \rightarrow 1$ (cf. [8]),

$$\begin{aligned} \mathbf{K}(\kappa') &= \ln \frac{4}{\kappa} + C_1 \kappa + \dots && \text{for } \kappa' \rightarrow 1, \\ \mathbf{K}(\kappa) &= \frac{\pi}{2} \left\{ 1 + \frac{1}{4} \kappa^2 + C_1 \kappa^4 + \dots \right\} && \text{for } \kappa \rightarrow 0, \end{aligned}$$

we obtain

$$\frac{1}{\omega^2} = \exp\left\{-\frac{2\pi \mathbf{K}(\kappa)}{\mathbf{K}(\kappa')}\right\} \approx \exp\left\{-\frac{\pi^2}{\ln(4R)}\right\} \approx 1 - \frac{\pi^2}{\ln(4R)}.$$

The latter indicates that the number n of different terms to achieve a tolerance ε is estimated by

$$n = \frac{|\log \varepsilon|}{|\log \omega^{-2}|} \approx \frac{|\log \varepsilon| \ln(4R)}{\pi^2}.$$

This result shows the same asymptotical convergence in n as the corresponding bounds in Lemmata A.8, A.10.

Concerning computational aspects of finding the best approximation in E_n^0 (same for E_n), we note that with a fixed interval $[a, b]$, we arrive at the strongly nonlinear minimisation problem $\inf_{v \in E_n^0} \|f - v\|_{L^\infty[a,b]}$, which involves $2n$ parameters $\{\omega_\nu, t_\nu\}_{\nu=1}^n$. The numerical algorithm is analogous to the determination of optimal rational approximations (see [14]). For our particular application with $f(x) = \frac{1}{x}$, we have the same asymptotical dependence $n = n(\varepsilon, R)$ as in Lemmata A.8, A.10, however the numerical results indicate a noticeable improvement compared with the quadrature method (cf. Lemma A.8) at least for small numbers $n \leq 15$. Numerical results for the best approximation of $\frac{1}{x}$ by sums of exponentials can be found in [2] and [14].

Acknowledgement. Discussions with Prof. D. Braess (Ruhr-Universität Bochum, Germany), Dr. H.-J. Flad (Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig) and with Prof. I. Gavriljuk (Berufsakademie Eisenach, Germany) are gratefully acknowledged.

References

- [1] G. Beylkin and M.J. Mohlenkamp: *Numerical operator calculus in higher dimensions*, University of Colorado, APPM Preprint No. 476, Aug. 2001
- [2] D. Braess: *Nonlinear approximation theory*. Springer-Verlag, Berlin, 1986.
- [3] H.-J. Flad, W. Hackbusch, D. Kolb, and R. Schneider: *Wavelet approximation of correlated wavefunctions. I. Basics*, J. Chem. Phys. **116**, (2002), 9641-9657.
- [4] I.P. Gavriljuk, W. Hackbusch, and B.N. Khoromskij: *\mathcal{H} -matrix approximation for the operator exponential with applications*, Numer. Math. **92** (2002), 83-111.
- [5] I.P. Gavriljuk, W. Hackbusch, and B.N. Khoromskij: *Data-sparse approximation to operator-valued functions of elliptic operators*. Math. Comp. **73** (2004), no. 247, 1297-1324 (electronic).
- [6] I.P. Gavriljuk, W. Hackbusch, and B.N. Khoromskij: *Data-sparse approximation to a hierarchy of operator-valued functions*. Preprint 20, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig 2003; Math. Comp. 2004 (to appear).

³The same result holds for E_n^0 , but the best approximation may belong to the closure E_n of E_n^0 .

- [7] I. P. Gavriilyuk, W. Hackbusch, and B. N. Khoromskij: *Tensor-product approximation to elliptic and parabolic solution operators in higher dimensions*. Preprint 83, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig 2003.
- [8] I.S. Gradshteyn and I.M. Ryzhik: *Table of Integrals, Series and Products* (6th edition), Academic Press, San Diego, 2000.
- [9] L. Grasedyck: *Existence and computation of a low Kronecker-rank approximation to the solution of a tensor system with tensor right-hand side*. Preprint 48, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, 2003.
- [10] L. Grasedyck and W. Hackbusch: *Construction and arithmetics of \mathcal{H} -matrices*. Computing **70** (2003), 295-334.
- [11] L. Grasedyck, W. Hackbusch, and B.N. Khoromskij: *Solution of large scale algebraic matrix Riccati equations by use of hierarchical matrices*. Computing **70** (2003), 121-165.
- [12] M. Griebel and S. Knapek: *Optimized tensor-product approximation spaces*. Constr. Approx. **16** (2000), 303-332.
- [13] W. Hackbusch: *A sparse matrix arithmetic based on \mathcal{H} -matrices. Part I: Introduction to \mathcal{H} -matrices*. Computing **62** (1999), 89-108.
- [14] W. Hackbusch: *Hierarchische Matrizen - Algorithmen und Analysis*. Manuskript, 2004
- [15] W. Hackbusch and B.N. Khoromskij: *A sparse \mathcal{H} -matrix arithmetic. Part II: Application to multi-dimensional problems*. Computing **64** (2000), 21-47.
- [16] W. Hackbusch and B.N. Khoromskij: *Towards \mathcal{H} -matrix approximation of the linear complexity*. Operator Theory: Advances and Applications **121**, Birkhäuser-Verlag, Basel, 2001, pp. 194-220.
- [17] W. Hackbusch, B.N. Khoromskij and R. Kriemann: *Hierarchical matrices based on a weak admissibility criterion*. Computing (to appear).
- [18] W. Hackbusch, B.N. Khoromskij and R. Kriemann: *Direct Schur complement method by domain decomposition based on \mathcal{H} -matrix approximation*. Preprint 25, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, 2004.
- [19] W. Hackbusch, B.N. Khoromskij, and E. Tyrtyshnikov: *Hierarchical Kronecker tensor-product approximation*, Preprint 35, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, 2003.
- [20] F. Keinert: *Uniform approximation to $|x|^\beta$ by Sinc functions*. J. Appr. Theory **66** (1991), 44-52.
- [21] H. Luo, D. Kolb, H.-J. Flad, W. Hackbusch, and T. Koprucki: *Wavelet approximation of correlated wavefunctions. II. Hyperbolic wavelets and adaptive approximation schemes*, J. Chem. Phys. **117**, (2002) pp. 3625-3638.
- [22] F. Stenger: *Numerical methods based on Sinc and analytic functions*. Springer-Verlag, 1993.
- [23] V.N. Temlyakov: *Approximation of functions with bounded mixed derivative*. Proc. Steklov Inst. Math. **178**, No.1 (1989), 275-293.
- [24] E.E. Tyrtyshnikov *Tensor approximations of matrices generated by asymptotically smooth functions*. (Russian) Mat. Sb. 194 (2003), no. 6, 147–160; translation in Sb. Math. 194 (2003), no. 5-6, 941–954.
- [25] D.R. Yarkony: *Modern electronic structure theory. Part II*. World Scientific, 1995.
- [26] N. Yarvin and V. Rokhlin: *Generalized Gaussian quadratures and singular value decompositions of integral operators*. SIAM J. Sci. Comput. **20** (1998), 699-718.
- [27] T. Zhang and G.H. Golub: *Rank-one approximation to high order tensors*. SIAM J. Matrix Anal. Appl. **23** (2001), 534-550.