FAST EVALUATION OF REAL AND COMPLEX EXPONENTIAL SUMS[∗]

STEFAN KUNIS[†] AND INES MELZER[†]

Abstract. Recently, the butterfly approximation scheme and hierarchical approximations have been proposed for the efficient computation of integral transforms with oscillatory or asymptotically smooth kernels. Combining both approaches, we propose a certain fast Fourier-Laplace transform, which in particular allows for the fast evaluation of polynomials at nodes in the complex unit disk. All theoretical results are illustrated by numerical experiments.

Key words. trigonometric approximation, nonharmonic Fourier series, fast Fourier transform, integral transforms, hierarchical matrices

AMS subject classifications. 65T50, 42A15, 30E10, 65D05, 65F30

1. Introduction. The fast Fourier transform (FFT) [9, 10, 13] belongs to the algorithms with large impact on science and engineering. Generalizations have been given for nonequispaced nodes; see [2, 8, 19, 20, 21, 27] for the recently suggested butterfly schemes and [6, 11, 12, 18, 23] together with its references for some wider survey on gridding type approximations. Moreover, structured low-rank approximations for integral transforms with smooth kernels have been developed as fast multipole methods [15, 25, 28] and hierarchical matrices [3, 4, 14, 16, 17]. One particular instance of a structured low-rank approximation for a smooth kernel is given in [22] for a discrete Laplace transform. In all cases, the concept in such schemes is to trade exactness for efficiency; instead of precise computations up to machine precision, the proposed methods guarantee a given target accuracy. Neglecting logarithmic factors in the problem size and the target accuracy, the computational complexity of all these algorithms scales linearly in the problems size.

Discrete Laplace transforms have been developed in [22, 24] based on polynomial interpolation and on approximations with Laguerre polynomials, respectively. In the first part of this paper, we present a matrix form of [22], including a small improvement in the error estimate, and develop a generalization to more general kernel functions. The main contribution of the paper is a combination of the matrix form of the discrete Laplace transform and a generalized FFT, where we use the decomposition of the Laplace transform explicitly and a small number of generalized FFTs as black box. In particular, this allows for the fast evaluation of a polynomial, given by its monomial coefficients, at many nodes in the complex unit disk. Alternatively, we might interpret this as an FFT with nonequispaced nodes in the upper half plane. For notational convenience, all ideas are presented for one space dimension but can be generalized in a straightforward manner to the multivariate case. Finally, the theoretical results on accuracy and computational complexity are illustrated by some numerical experiments.

2. Preliminaries. Let $q \in \mathbb{N}$ and the nodes $t_j = \cos \frac{2j+1}{2q}\pi$, $j = 0, \ldots, q-1$, be the zeros of the q-th Chebyshev polynomial of the first kind. Moreover, let $A := [a, b]$, $a < b$, be an interval with diameter $\mathrm{diam}(A) := b - a$, midpoint $c^A := \frac{a+b}{2}$, and Chebyshev nodes $y_j^A := c^A + \frac{\mathrm{diam}(A)}{2}t_j$, $j = 0, \ldots, q-1$. The corresponding Lagrange polynomials

---

[†]Institute of Mathematics, University Osnabrück, Osnabrück, Germany
({stefan.kunis,ines.melzer}@math.uos.de).

$L_r^A : A \to \mathbb{R}, r = 0, \ldots, q - 1$, are

$$L_r^A(y) := \prod_{\substack{j=0 \\ j \neq r}}^{q-1} \frac{y - y_j^A}{y_r^A - y_j^A} = \frac{\frac{\lambda_r}{y - y_r^A}}{\sum_{s=0}^{q-1} \frac{\lambda_s}{y - y_s^A}}, \qquad \lambda_r := (-1)^r \sin\left(\frac{2r + 1}{2q}\pi\right),$$

where the second identity is called the barycentric formula and allows for stable evaluation; cf. [5]. We denote by $C(A)$ the continuous functions on A and define the interpolation operator $\mathcal{I}_q^A : C(A) \to C(A)$,

$$\mathcal{I}_q^A g := \sum_{j=0}^{q-1} g(y_j^A) L_j^A,$$

which fulfills

$$(2.1) \qquad \|g - \mathcal{I}_q^A g\|_{C(A)} \leq \frac{\mathrm{diam}(A)^q}{2^{2q-1} q!} \|g^{(q)}\|_{C(A)}, \qquad \text{if } g^{(q)} \in C(A),$$

$$(2.2) \qquad \|\mathcal{I}_q^A\| := \sup_{\|g\|_{C(A)}=1} \|\mathcal{I}_q^A g\|_{C(A)} \leq 1 + \frac{2}{\pi} \log q.$$

Now let $A, B \subset \mathbb{R}$ be two intervals and $\kappa : A \times B \to \mathbb{R}$. We define the interpolation in both variables by

$$\mathcal{I}_q^{A \times B} := \mathcal{I}_q^A \otimes \mathcal{I}_q^B, \qquad \mathcal{I}_q^{A \times B} \kappa(y, \xi) := \sum_{s=0}^{q-1} \sum_{r=0}^{q-1} L_s^A(y) \kappa(y_s^A, \xi_r^B) L_r^B(\xi),$$

where ξ_r^B denote the Chebyshev nodes in the interval B. We note in passing that

$$\mathcal{I}_q^A \otimes \mathcal{I}_q^B = (\mathcal{I}_q^A \otimes \mathcal{I})(\mathcal{I} \otimes \mathcal{I}_q^B),$$

where $\mathcal{I}$ denotes the identity, and $\|\mathcal{I}_q^A \otimes \mathcal{I}\| \leq 1 + \frac{2}{\pi} \log q$.

3. Rokhlin's discrete Laplace transform and a generalization. We generalize [22] to a method for computing

$$(3.1) \qquad f(y_i) = \sum_{j=1}^N \hat{f}_j \kappa(y_i, \xi_j), \qquad i = 1, \ldots, N,$$

for specific kernels κ, given $N \in \mathbb{N}$, $y_1 > \ldots > y_N > 0$, $\xi_1 > \ldots > \xi_N > 0$, and $\hat{f}_k \in \mathbb{C}$.

Adopting the terms from the hierarchical matrices literature (see, e.g., [4, 17]), a function $\kappa : (0, \infty) \times (0, \infty) \to \mathbb{R}$ is said to be *asymptotically smooth* if there exist constants $\nu \in \mathbb{R}$ and $C, \mu, s \geq 0$ such that for all $q \in \mathbb{N}$ the conditions

$$\left| y^q \partial_y^q \kappa(y, \xi) \right| \leq C q! \mu^q q^\nu (y\xi)^{-s} \quad \text{and} \quad \left| \xi^q \partial_\xi^q \kappa(y, \xi) \right| \leq C q! \mu^q q^\nu (y\xi)^{-s}$$

are fulfilled for all $y, \xi \in (0, \infty)$. The parameter s characterizes the singularity of the kernel for $y\xi = 0$. Moreover, we call two intervals $A, B \subset [0, \infty)$ *admissible* if

$$\mathrm{diam}(A) \leq \mathrm{dist}(A, 0) := \min\{|x - 0| : x \in A\} \qquad \text{and} \qquad \mathrm{diam}(B) \leq \mathrm{dist}(B, 0).$$

THEOREM 3.1. *Let $q \in \mathbb{N}$, $q \geq 2$, $A, B \subset (0, \infty)$ be admissible and $\kappa \colon A \times B \to \mathbb{R}$ be asymptotically smooth with constants $\nu \in \mathbb{R}$ and $C, \mu, s \geq 0$. Then we have*

$$\left\| \kappa - \mathcal{I}_q^{A \times B} \kappa \right\|_{C(A \times B)} \leq \frac{C \mu^q q^\nu}{2^{2q-1}} \left(2 + \frac{2}{\pi} \log q \right) (\operatorname{dist}(A, 0) \operatorname{dist}(B, 0))^{-s}.$$

Proof. For fixed $\xi \in B$ and $g : A \to \mathbb{R}$, $g(y) := \kappa(y, \xi)$, we apply the error formula (2.1) and obtain

$$\| g - \mathcal{I}_q^A g \|_{C(A)} \leq \frac{\operatorname{diam}(A)^q}{2^{2q-1} q!} \| g^{(q)} \|_{C(A)}.$$

The asymptotic smoothness and the admissibility implies

$$\left| g^{(q)}(y) \right| \leq C q! \mu^q q^\nu y^{-q} (y\xi)^{-s} \leq C q! \mu^q q^\nu (\operatorname{dist}(A, 0))^{-q} (y\xi)^{-s}$$

and in conclusion

$$\| g - \mathcal{I}_q^A g \|_{C(A)} \leq C \mu^q q^\nu 2^{1-2q} \sup_{y \in A, \xi \in B} |(y\xi)^{-s}|.$$

The same estimate holds true with respect to $\xi \in B$. From this and together with the bound (2.2), we conclude

$$\| \kappa - \mathcal{I}_q^{A \times B} \kappa \|_{C(A \times B)} \leq \| \kappa - (\mathcal{I}_q^A \otimes \mathcal{I}) \kappa \|_{C(A \times B)} + \| \mathcal{I}_q^A \otimes \mathcal{I} \| \| \kappa - (\mathcal{I} \otimes \mathcal{I}_q^B) \kappa \|_{C(A \times B)}$$
$$\leq \frac{C \mu^q q^\nu}{2^{2q-1}} \left(2 + \frac{2}{\pi} \log q \right) \sup_{y \in A, \xi \in B} |(y\xi)^{-s}|.$$

The conditions $y\xi \geq \operatorname{dist}(A, 0) \operatorname{dist}(B, 0)$ and $s \geq 0$ imply the assertion. ☐

We start with the discrete Laplace transform from [22], i.e., $\kappa(y, \xi) = \mathrm{e}^{-\xi y}$. In matrix form, the computation of (3.1) reads as

$$(3.2) \qquad \mathbf{f} := \mathbf{K} \hat{\mathbf{f}} \quad \text{with} \quad \mathbf{K} := (\kappa(y_j, \xi_k))_{j,k=1,\ldots,N} \quad \text{and} \quad \hat{\mathbf{f}} := \left(\hat{f}_j \right)_{j=1,\ldots,N}.$$

Algorithm 1 below finally computes an approximation $\tilde{\mathbf{f}} \approx \mathbf{f}$, precisely defined in equation (A.1), and we refer the reader to Theorem A.3 for a shortened and slightly improved error estimate compared to [22]. Subsequently, we focus on the matrix partitioning of K into blocks of approximate low rank and the derivation of the computational costs of Algorithm 1. This will allow for the generalization to other kernels in the end of this section and for an application when evaluating polynomials in the unit disk in Section 4.

DEFINITION 3.2. *For a given accuracy $\varepsilon > 0$ and interval lengths $y_1, \xi_1 > 0$, we define*

$$q := \left\lceil \frac{1}{2} + \log_4 1/\varepsilon \right\rceil, \qquad\qquad M := \left\lceil \log_2 \frac{y_1 \xi_1}{\varepsilon} \right\rceil + 1,$$

$$\ell_m := \max(1, \lfloor \log_2(y_1 \xi_1) - m - \log_2(\log 1/\varepsilon) \rfloor + 1), \qquad L_m := M - m,$$

for $m = 1, \ldots, M - 1$ and set up the geometric partitioning (see Figure 3.1 for an illustration)

$$Y := [0, y_1], \qquad Y_M := \left[0, \frac{y_1}{2^{M-1}} \right], \qquad Y_m := \left(\frac{y_1}{2^m}, \frac{y_1}{2^{m-1}} \right],$$

$$\Omega := [0, \xi_1], \qquad \Omega_M := \left[0, \frac{\xi_1}{2^{M-1}} \right], \qquad \Omega_m := \left(\frac{\xi_1}{2^m}, \frac{\xi_1}{2^{m-1}} \right].$$
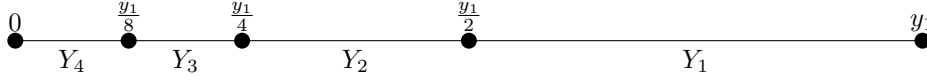
FIG. 3.1. *Decomposition of $Y = [0, y_1]$ for $M = 4$.*

For ease of notation in Algorithm 1, we moreover define

$$\mathbf{L}^{\Omega_l} := \left(L_r^{\Omega_l}(\xi_j) \right)_{\xi_j \in \Omega_l, r=0,\dots,q-1}, \qquad \mathbf{L}^{Y_m} := \left(L_s^{Y_m}(y_i) \right)_{y_i \in Y_m, s=0,\dots,q-1},$$

$$\hat{\mathbf{f}}^{\Omega_l} := \left(\hat{f}_j \right)_{\xi_j \in \Omega_l}, \qquad \mathbf{K}^{Y_m, \Omega_l} := \left(\kappa(y_s^{Y_m}, \xi_r^{\Omega_l}) \right)_{s,r=0}^{q-1,q-1},$$

for $m, l = 1, \dots, M-1$.

Regarding the computational costs, we first note that the computation of g_m in Algorithm 1 by means of a cumulative summation takes $\mathcal{O}(N)$ operations. The spatial partitions in Definition 3.2 yield a partition of the matrix $\mathbf{K}$ into admissible blocks. Considering only the case when the kernel is approximated by interpolation, Algorithm 1 factors out row- and column bases by

$$\tilde{\mathbf{K}} = \left(\mathbf{L}^{Y_m} \mathbf{K}^{Y_m, \Omega_\ell} \left(\mathbf{L}^{\Omega_\ell} \right)^\top \right)_{m=1, \ell=\ell_m}^{M-1, L_m}$$

$$= \operatorname{diag}(\mathbf{L}^{Y_1}, \dots, \mathbf{L}^{Y_{M-1}}) \left(\mathbf{K}^{Y_m, \Omega_\ell} \right)_{m=1, \ell=\ell_m}^{M-1, L_m} \operatorname{diag}(\mathbf{L}^{\Omega_{\ell_m}}, \dots, \mathbf{L}^{\Omega_{L_m}})^\top.$$

The application of the rightmost block-diagonal matrix takes at most $\sum_{\ell=\ell_m}^{L_m} q|\Omega_\ell| \leq qN$ operations. The second matrix has no more than $L_m - \ell_m \leq 2\log(1/\varepsilon)$ blocks in its m-th block row. In total its application takes $\mathcal{O}(Mq^2 \log \frac{1}{\varepsilon})$ operations, and the multiplication with the left block-diagonal matrix again takes at most qN operations. Note that we neglect the precomputation of the necessary matrices $\mathbf{L}^{\Omega_\ell}$, $\mathbf{L}^{Y_m}$, and $\mathbf{K}^{Y_m, \Omega_\ell}$. In total, Algorithm 1 takes

$$\mathcal{O}\left(N \log \frac{1}{\varepsilon} + \log^3 \frac{1}{\varepsilon} \log \frac{y_1 \xi_1}{\varepsilon} \right)$$

floating point operations.

This discrete Laplace transform can be generalized to asymptotically smooth kernels. The slightly modified versions of Algorithm 1 and Theorem A.3 read as follows.

THEOREM 3.3. *Let $\kappa : (0, \infty) \times (0, \infty) \to \mathbb{R}$ be asymptotically smooth with constants $\nu \in \mathbb{R}$, $s \geq 0$ and $\mu \in (0, 4)$. Furthermore, let $\varepsilon > 0$, $y_1, \xi_1 > 0$ be fixed for all $N \in \mathbb{N}$ and the nodes $y_1 > y_2 > \dots > y_N > 0$, $\xi_1 > \xi_2 > \dots > \xi_N > 0$ be locally quasi-uniformly such that the intervals $[0, y_1/N]$ and $[0, \xi_1/N]$, respectively, contain only a constant number of nodes for all $N \in \mathbb{N}$. We set $M := \left\lceil \log_2 \frac{y_1 \xi_1 N}{\varepsilon} \right\rceil + 1$ and $q \in \mathcal{O}(\log \frac{N}{\varepsilon})$ and define the approximation $\tilde{f} : Y \to \mathbb{C}$,*

$$\tilde{f}(y) := \begin{cases} f(y) & y \in Y_M, \\ \displaystyle\sum_{\ell=1}^{M-1} \sum_{\xi_j \in \Omega_\ell} \hat{f}_j \mathcal{I}^{Y_m \times \Omega_\ell} \kappa(y, \xi_j) + \sum_{\xi_k \in \Omega_M} \hat{f}_k \kappa(y, \xi_k) & y \in Y_m,\ 1 \leq m < M. \end{cases}$$

Then the error estimate $\|\mathbf{f} - \tilde{\mathbf{f}}\|_\infty \leq \varepsilon \|\hat{\mathbf{f}}\|_1$ is fulfilled and the modified Algorithm 1 computes this approximation in $\mathcal{O}(N \log \frac{N}{\varepsilon} + \log^4 \frac{N}{\varepsilon})$ floating point operations.

ALGORITHM 1 Discrete Laplace transform.

Input:

$\varepsilon \in (0,1),\ N \in \mathbb{N},\ \hat{\boldsymbol{f}} \in \mathbb{C}^N$ $\qquad \triangleright$ target accuracy, number of nodes, coefficients

$\xi_1 > \cdots > \xi_N > 0,\ y_1 > \cdots > y_N > 0$ $\qquad \triangleright$ nodes in frequency and spatial domain

Output:

$\tilde{\mathbf{f}} \in \mathbb{C}^N, \tilde{\mathbf{f}} \approx \boldsymbol{K}\hat{\boldsymbol{f}}$ $\qquad\qquad\qquad\qquad\quad \triangleright$ samples in spatial domain

---

$M = \left\lceil \log_2 \frac{y_1 \xi_1}{\varepsilon} \right\rceil + 1,\ q = \left\lceil \frac{1}{2} + \log_4 \frac{1}{\varepsilon} \right\rceil$ $\quad \triangleright$ decomposition number, approximation rank

$g_M = \sum_{\xi_j \in \Omega_M} \hat{f}_j$ $\qquad\qquad\qquad \triangleright$ compute result when the kernel is approximated by one

for $m = M-1, \dots, 1$ **do**

$\qquad g_m = g_{m+1} + \sum_{\xi_j \in \Omega_m} \hat{f}_j$

$\tilde{\boldsymbol{f}}^{Y_M} = g_1$

for $\ell = 1, \dots, M-1$ **do** $\qquad\qquad\qquad\qquad\qquad \triangleright$ adjoint interpolate to Chebyshev grid

$\qquad \mathbf{v}^{\Omega_\ell} = \left(\mathbf{L}^{\Omega_\ell}\right)^\top \hat{\mathbf{f}}^{\Omega_\ell}$

for $m = 1, \dots, M-1$ **do**

$\qquad \mathbf{h}^{Y_m} = \sum_{\ell=\ell_m}^{L_m} \mathbf{K}^{Y_m, \Omega_\ell} \mathbf{v}^{\Omega_\ell}$ $\qquad\qquad\qquad \triangleright$ apply Laplace kernel on Chebyshev grid

$\qquad \tilde{\boldsymbol{f}}^{Y_m} = \mathbf{L}^{Y_m} \mathbf{h}^{Y_m} + g_{L_m+1} \mathbf{1}$ $\qquad \triangleright$ interpolate to target nodes and compute final result

---

Proof. For $\mu \in (0,4)$, $C > 0$, and $\nu \in \mathbb{R}$, there exist constants $\tilde{C} > 0$ and $\frac{\mu}{4} < c < 1$, such that

$$2C \left(\frac{\mu}{4}\right)^q q^\nu \left(2 + \frac{2}{\pi} \log q\right) \leq \tilde{C} c^q$$

holds true for all $q \in \mathbb{N}$. Theorem 3.1 implies for $A, B \subset (0,\infty)$ a local approximation bound

$$(3.3) \qquad \left\| \kappa - \mathcal{I}_q^{A \times B} \kappa \right\|_{C(A \times B)} \leq \tilde{C} c^q (\mathrm{dist}(A,0)\, \mathrm{dist}(B,0))^{-s}.$$

Let now $A := Y_m$ and $B := \Omega_\ell$ with $m, \ell \neq M$ be given. Since $M - 1 \leq \log_2 \frac{y_1 \xi_1 N}{\varepsilon} + \frac{1}{2}$, it follows

$$\mathrm{dist}(A,0)\, \mathrm{dist}(B,0) \geq \frac{y_1}{2^{M-1}} \frac{\xi_1}{2^{M-1}} \geq \frac{\varepsilon^2}{2 y_1 \xi_1 N^2}.$$

Applying $q \geq \log_c \left(\frac{\tilde{C} \varepsilon^{2s+1}}{(2 y_1 \xi_1)^s N^{2s}}\right)$ implies

$$\left\| \kappa - \mathcal{I}_q^{A \times B} \kappa \right\|_{C(A \times B)} \leq \tilde{C} (2 y_1 \xi_1)^s \left(\frac{N}{\varepsilon}\right)^{2s} c^q \leq \varepsilon.$$

Finally, $M = \mathcal{O}(\log \frac{N}{\varepsilon})$ leads to a constant number of nodes in the near fields Y_M and Ω_M, where we apply direct computations. Of course the approximations of the kernel by either zero or one as in Lemma A.2 cannot be done in general, and thus we let $\ell = 1, \dots, M-1$ run in Algorithm 1. Since $q = \mathcal{O}(\log \frac{N}{\varepsilon})$, Algorithm 1 now takes $\mathcal{O}(N \log \frac{N}{\varepsilon} + \log^4 \frac{N}{\varepsilon})$ floating point operations. $\qquad \square$

EXAMPLE 3.4. Let the modified Bessel function of the second kind $K_\eta \colon (0,\infty) \to \mathbb{R}$,

$$K_\eta(x) := \int_0^\infty e^{-x \cosh(t)} \cosh(\eta t) \mathrm{d}t,$$

be given, set $\eta = \frac{1}{2}$, and consider the kernel $\kappa : (0, \infty) \times (0, \infty) \to \mathbb{R}$,

$$\kappa(y, \xi) := K_{\frac{1}{2}}(y\xi) = \sqrt{\frac{\pi}{2y\xi}} e^{-y\xi},$$

which possesses a singularity for $y\xi = 0$. Induction on $q \in \mathbb{N}_0$ shows the identity

$$\frac{\partial^q}{\partial y^q} \kappa(y, \xi) = \sqrt{\frac{\pi}{2}} (-\xi)^q e^{-y\xi} \sum_{k=0}^{q} \binom{q}{k} \frac{\prod_{j=0}^{k-1}(2j+1)}{2^k} (y\xi)^{-(2k+1)/2},$$

from which the asymptotic smoothness

$$\left| \frac{\partial^q}{\partial y^q} \kappa(y, \xi) \right| \leq \sqrt{\frac{\pi}{2y\xi}} \xi^q e^{-y\xi} \sum_{k=0}^{q} \binom{q}{k} k! (y\xi)^{-k} \leq \sqrt{\frac{\pi}{2}} q! (y\xi)^{-\frac{1}{2}} y^{-q}$$

with constants $C = \sqrt{\frac{\pi}{2}}$, $\mu = 1$, $\nu = 0$, and $s = \frac{1}{2}$ follows. Theorem 3.1 implies the corresponding local error estimate, and thus $\tilde{C} = 2\pi$ and $c = \frac{1}{3}$ can be chosen in (3.3).

4. Evaluation of polynomials in the unit disk. We are interested in evaluating the generalized polynomials $f : \mathbb{C} \to \mathbb{C}$,

$$(4.1) \qquad f(z) = \sum_{k=1}^{N} \hat{f}_k z^{\xi_k}$$

at nodes $z_j \in Z := \{z \in \mathbb{C} : |z| \leq 1\}$, $j = 1, \ldots, N$, and for exponents $\xi_k \in [1, N]$, $N \in \mathbb{N}$, where we exclude the nonpositive real axis $z_j \leq 0$ for noninteger exponents ξ_k.

The main idea now consists in a combination of the discrete Laplace and a nonequispaced Fourier transform [6, 11, 12, 18, 19, 23, 27]. We write

$$(4.2) \qquad z = e^{-y} e^{2\pi i x}, \qquad y \in [0, \infty), \ x \in [0, 1),$$

and note that the summation (4.1) is a matrix-vector multiplication with the matrix

$$\boldsymbol{C} := \boldsymbol{A} \odot \boldsymbol{K},$$

where the Fourier matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A} := \left(e^{2\pi i \xi_k x_j} \right)_{j,k=1,\ldots,N}$$

and the Laplace matrix $\boldsymbol{K}$ is given by (3.2). The symbol $\odot$ denotes the Hadamard (point-wise) product and we set $\|\boldsymbol{M}\|_{1 \to \infty} := \sup\{\|\boldsymbol{M}\boldsymbol{x}\|_\infty : \|\boldsymbol{x}\|_1 = 1\}$. We have the following result when approximating the factors as in the previous sections.

LEMMA 4.1. *Let $\varepsilon \in (0, 1)$ and the Fourier and the Laplace matrix be approximated by*

$$\|\boldsymbol{A} - \tilde{\boldsymbol{A}}\|_{1 \to \infty} \leq \frac{\varepsilon}{3}, \qquad \|\boldsymbol{K} - \tilde{\boldsymbol{K}}\|_{1 \to \infty} \leq \frac{\varepsilon}{3}.$$

Then

$$\|\boldsymbol{A} \odot \boldsymbol{K} - \tilde{\boldsymbol{A}} \odot \tilde{\boldsymbol{K}}\|_{1 \to \infty} \leq \varepsilon.$$

Proof. The estimate simply follows from $\|\boldsymbol{A}\|_{1 \to \infty} = 1$, $\|\tilde{\boldsymbol{K}}\|_{1 \to \infty} \leq 1 + \frac{\varepsilon}{3}$, and

$$\boldsymbol{A} \odot \boldsymbol{K} - \tilde{\boldsymbol{A}} \odot \tilde{\boldsymbol{K}} = \boldsymbol{A} \odot (\boldsymbol{K} - \tilde{\boldsymbol{K}}) + (\boldsymbol{A} - \tilde{\boldsymbol{A}}) \odot \tilde{\boldsymbol{K}}. \qquad \square$$

We use the hierarchical decomposition of the discrete Laplace transform and realize matrix-vector products with matrix blocks by the following technique.

LEMMA 4.2. *By slight abuse of notation, let* $\mathbf{K} = \mathbf{L}^Y \mathbf{K}^{Y,\Omega} \left(\mathbf{L}^\Omega\right)^\top$ *denote a single matrix block of the Laplace matrix and* $\mathbf{A}$ *the associated block of the Fourier matrix. Then*

$$(\mathbf{A} \odot \mathbf{K})\,\hat{\mathbf{f}} = \left(\mathbf{L}^Y \odot \mathbf{A} \left(\mathrm{diag}\,\hat{\mathbf{f}}\right) \mathbf{L}^\Omega \left(\mathbf{K}^{Y,\Omega}\right)^\top\right)\mathbf{1},$$

where $\mathbf{1} := (1,\ldots,1)^\top \in \mathbb{R}^q$.

Proof. The simplest case is $\mathbf{K} = \mathbf{l}^Y \left(\mathbf{l}^\Omega\right)^\top$ from which $a_{ij} \cdot k_{ij} = l_i^Y a_{ij} l_j^\Omega$ and the assertion follows. This implies the result since

$$\mathbf{K} = \mathbf{L}^Y \mathbf{K}^{Y,\Omega} \left(\mathbf{L}^\Omega\right)^\top = \sum_{r=1}^q \mathbf{l}_r^Y \left(\tilde{\mathbf{l}}_r^{Y,\Omega}\right)^\top,$$

where $\mathbf{l}_r^Y$ and $\tilde{\mathbf{l}}_r^{Y,\Omega}$ denote the columns of $\mathbf{L}^Y$ and $\mathbf{L}^\Omega \left(\mathbf{K}^{Y,\Omega}\right)^\top$, respectively. □

THEOREM 4.3. *Let* $N \in \mathbb{N}$, $\varepsilon > 0$, $z_j \in Z$, $j = 1,\ldots,N$, $\hat{\mathbf{f}} \in \mathbb{C}^N$ *and* $\mathbf{f} := \mathbf{C}\hat{\mathbf{f}}$ *with*

$$\mathbf{C} = (z_j^k)_{j,k=1,\ldots,N}$$

be given. Algorithm 2 takes $\mathcal{O}\left(N \log \frac{1}{\varepsilon} \log^2 \frac{N}{\varepsilon}\right)$ *or* $\mathcal{O}\left(N \log N \log \frac{1}{\varepsilon} \log^3 \frac{N}{\varepsilon}\right)$ *floating point operations using the FFT for nonequispaced nodes in time and frequency [12] or the butterfly sparse Fourier transform [19], respectively. Its output* $\tilde{\mathbf{f}} \in \mathbb{C}^N$ *fulfills the error estimate* $\|\mathbf{f} - \tilde{\mathbf{f}}\|_\infty \le \varepsilon \|\hat{\mathbf{f}}\|_1$.

Proof. We start by noting that $|z_j| < \varepsilon$ implies $|f(z_j)| < \varepsilon \|\hat{\boldsymbol{f}}\|_1$ in (4.1), i.e., $f(z_j)$ can be approximated by zero. We collect the associated nodes $Z^0 := \{z_j \in \mathbb{C} : |z_j| < \varepsilon\}$ and set the result $\tilde{\boldsymbol{f}} \in \mathbb{C}^N$, restricted to these nodes, to zero

$$\tilde{\boldsymbol{f}}^{Z_0} := \mathbf{0}.$$

Now write all nodes in polar form (4.2), where we can assume $0 \le y_N \le \ldots \le y_1 \le \log \frac{1}{\varepsilon}$ for the rest of the proof. The decomposition of the frequency nodes and the spatial nodes in dyadic intervals Ω_ℓ and Y_m for $\ell, m = 1,\ldots,M$, cf. Definition 3.2, induces a decomposition of the unit disk Z into concentric bands as well as a decomposition of the nodes $\{x_j : j = 1,\ldots,N\} = \dot{\cup}_{m=1,\ldots,M} X_m$ with $X_m := \{x_j : y_j \in Y_m\}$. Accordingly, we denote restrictions of the Fourier matrix, the Fourier coefficients, and the result vector by superscripts with these sets. In particular, spatial nodes are close to one in modulus $z_j \in Z_M$ if and only if $y_j \in Y_M$, and thus we set

$$\tilde{\boldsymbol{f}}^{Z_M} := \tilde{\boldsymbol{A}}^{X_M,[1,\xi_1]}\hat{\boldsymbol{f}}.$$

Here and subsequently, all multiplications with submatrices of $\boldsymbol{A}$ are realized by padding zeros to the input vector, multiplying with $\boldsymbol{A}$, and restricting to the desired results. Regarding the most interesting part of the approximation, the decomposition from Theorem A.3 yields

$$\boldsymbol{A} \odot \tilde{\boldsymbol{K}} = \left(\boldsymbol{A}^{X_m,\Omega_\ell} \odot \boldsymbol{L}^{Y_m} \boldsymbol{K}^{Y_m,\Omega_\ell} \left(\boldsymbol{L}^{\Omega_\ell}\right)^\top\right)_{m=1,\ell=\ell_m}^{M-1,L_m}.$$

For notational simplicity, we apply Lemma 4.2 to one block row m and two (artificial) block columns $\ell = 0, 1$, $\Omega := \Omega_0 \cup \Omega_1$, in

$$\left(\boldsymbol{A}^{X_m,\Omega_0} \odot \boldsymbol{L}^{Y_m} \boldsymbol{K}^{Y_m,\Omega_0} \left(\boldsymbol{L}^{\Omega_0} \right)^\top \quad \boldsymbol{A}^{X_m,\Omega_1} \odot \boldsymbol{L}^{Y_m} \boldsymbol{K}^{Y_m,\Omega_1} \left(\boldsymbol{L}^{\Omega_1} \right)^\top \right) \hat{\boldsymbol{f}}^\Omega$$

$$= \left(\boldsymbol{L}^{Y_m} \odot \boldsymbol{A}^{X_m,\Omega_0} \left(\operatorname{diag} \hat{\boldsymbol{f}}^{\Omega_0} \right) \boldsymbol{L}^{\Omega_0} \left(\boldsymbol{K}^{Y_m,\Omega_0} \right)^\top \right) \boldsymbol{1}$$

$$+ \left(\boldsymbol{L}^{Y_m} \odot \boldsymbol{A}^{X_m,\Omega_1} \left(\operatorname{diag} \hat{\boldsymbol{f}}^{\Omega_1} \right) \boldsymbol{L}^{\Omega_1} \left(\boldsymbol{K}^{Y_m,\Omega_1} \right)^\top \right) \boldsymbol{1}$$

$$= \left(\boldsymbol{L}^{Y_m} \odot \boldsymbol{A}^{X_m,\Omega} \left(\begin{array}{c} \left(\operatorname{diag} \hat{\boldsymbol{f}}^{\Omega_0} \right) \boldsymbol{L}^{\Omega_0} \left(\boldsymbol{K}^{Y_m,\Omega_0} \right)^\top \\ \left(\operatorname{diag} \hat{\boldsymbol{f}}^{\Omega_1} \right) \boldsymbol{L}^{\Omega_1} \left(\boldsymbol{K}^{Y_m,\Omega_1} \right)^\top \end{array} \right) \right) \boldsymbol{1}.$$

Now, the error estimate is a straightforward consequence of Lemma 4.1.

The complexity estimate follows when considering the dominant computation in the second last line $\boldsymbol{F} = \ldots$ in Algorithm 2. We have $M = \mathcal{O}(\log \frac{N}{\varepsilon})$ steps in the outer loop and $q = \mathcal{O}(\log \frac{1}{\varepsilon})$ right hand sides for the multiplication with the approximate Fourier matrix $\tilde{\boldsymbol{A}}$, for which the computational needs are given by $\mathcal{O}(N \log N \log^2 \frac{N}{\varepsilon})$ for the butterfly sparse Fourier transform [19] or by $\mathcal{O}(N \log \frac{N}{\varepsilon})$ floating point operations for the so-called fast Fourier transform for nonequispaced nodes in time and frequency (NNFFT); see [12]. Moreover, the constant in the $\mathcal{O}$-notation is improved for the special case $\xi_k = k$ by means of the nonequispaced fast Fourier transform (NFFT); see [18]. □

ALGORITHM 2 Evaluation of polynomials in the unit disk.

Input:
  $\varepsilon \in (0,1), \ N \in \mathbb{N}, \ \hat{\boldsymbol{f}} \in \mathbb{C}^N$ ▷ target accuracy, number of nodes, coefficients
  $\xi_1 > \cdots > \xi_N \geq 1, \ z_1, \ldots, z_N \in \{z \in \mathbb{C} : |z| \leq 1\}$ ▷ nodes in frequency, spatial domain
Output:
  $\tilde{\mathbf{f}} \in \mathbb{C}^N, \tilde{\mathbf{f}} \approx \boldsymbol{C} \hat{\boldsymbol{f}}$ ▷ samples in spatial domain

---

  $M = \left\lceil \log_2 \frac{N \log 1/\varepsilon}{\varepsilon} \right\rceil + 1, \ q = \lceil \frac{1}{2} + \log_4 1/\varepsilon \rceil$ ▷ decomposition number, approx. rank

  $\tilde{\boldsymbol{f}}^{Z_0} = \boldsymbol{0}$ ▷ Laplace kernel is approximated by zero

  $\Omega = [1, \xi_1], \ \tilde{\boldsymbol{f}}^{Z_M} = \tilde{\boldsymbol{A}}^{Y_M,\Omega} \hat{\boldsymbol{f}}$ ▷ Laplace kernel is approx. by one, compute NNFFT

  for $\ell = 1, \ldots, M-1$ **do**   ▷ adjoint interpolate to Chebyshev grid
     $\hat{\boldsymbol{F}}^{\Omega_\ell} = \left(\operatorname{diag} \hat{\boldsymbol{f}}^{\Omega_\ell} \right) \boldsymbol{L}^{\Omega_\ell} \in \mathbb{C}^{|\Omega_\ell| \times q}$

  for $m = 1, \ldots, M-1$ **do**
     $\Omega = \cup_{\ell > L_m} \Omega_\ell, \ \tilde{\boldsymbol{f}}^{Z_m} = \tilde{\boldsymbol{A}}^{X_m,\Omega} \hat{\boldsymbol{f}}^\Omega$ ▷ Laplace kernel is approx. by one, NNFFT
     $\Omega = \cup_{\ell = \ell_m}^{L_m} \Omega_\ell$ ▷ apply Laplace kernel on Chebyshev grid, NNFFT
     $\boldsymbol{F} = \tilde{\boldsymbol{A}}^{X_m,\Omega} \left(\hat{\boldsymbol{F}}^{\Omega_\ell} \left(\boldsymbol{K}^{Y_m,\Omega_\ell} \right)^\top \right)_{\ell = \ell_m, \ldots, L_m} \in \mathbb{C}^{|X_m| \times q}$

     $\tilde{\boldsymbol{f}}^{Z_m} = \tilde{\boldsymbol{f}}^{Z_m} + \left(\boldsymbol{L}^{Y_m} \odot \boldsymbol{F} \right) \boldsymbol{1}$ ▷ interpolate to target nodes and compute final result

---

We note in passing that minor improvements in the computational complexity are possible by considering the butterfly approximation scheme directly on the blocks $\boldsymbol{A}^{X_m,\Omega}$. Lemma 4.2

shows how to apply the Hadamard product of a low-rank matrix and a matrix that allows for a fast algorithm to a vector. The very same idea is used in [26] for a polynomial conversion matrix being a Hadamard product of an approximately low-rank Hankel matrix and a Toeplitz matrix, which of course allows for fast multiplication by means of FFTs. Moreover note that [1] suggests a fast algorithm for the multiplication with C when the nodes z_j are close to the unit circle. Regarding generalizations, we get a fast algorithm for the multiplication with the adjoint matrix C^* by simply using the adjoint algorithms for the matrices K and A. In particular, this allows to evaluate $g : \mathbb{C} \to \mathbb{C}$,

$$g(\xi) = \sum_{j=1}^{N} \hat{g}_j z_j^{\xi}$$

at nodes $\xi_k \in [1, N]$, $k = 1, \dots, N$ for given $z_j \in Z = \{z \in \mathbb{C} : |z| \leq 1\}$ and coefficients $\hat{g}_j \in \mathbb{C}$ efficiently. Possible applications include the fast evaluation of certain special functions when approximated as in [7] on the real line. The most general case with

$$e^{(\xi+i\eta)(x+iy)} = e^{\xi x} e^{-by} e^{i(\eta x + \xi y)}$$

as kernel allows for efficient treatment when (ξ, η) as well as (x, y) samples a smooth contour in $\mathbb{C}$ and are in appropriate ranges. Then the last term leads to a $2D$ sparse FFT [19, 27], and we might apply the Hadamard product idea twice.

5. Numerical results. The implementation of Algorithm 1 and Algorithm 2 is realized in MATLAB 2013a. We use one node of a Intel Xeon, 128GByte, 2.2GHz, Scientific Linux release 6.5 (Carbon), for all numerical experiments. We draw random uniformly distributed coefficients $\hat{f}_k$, equispaced frequencies $\xi_k = k$, $k = 1, \dots, N$, random nodes $x_j \in [0, 1)$, $j = 1, \dots, N$, and random nodes $0 \leq y_N \leq y_{N-1} \leq \cdots \leq y_1 \leq (2q - 1) \log 2$, which ensures $y_j \in [0, \log 1/\varepsilon]$. We consider the relative error

$$(5.1) \qquad\qquad E := \frac{\|\mathbf{f} - \tilde{\mathbf{f}}\|_\infty}{\|\hat{\mathbf{f}}\|_1},$$

where $\mathbf{f}, \tilde{\mathbf{f}} \in \mathbb{C}^N$ denote the exact result and its approximation, respectively; see Theorem A.3 and Theorem 4.3. Figure 5.1(a),(b),(c) shows the quantity E and the corresponding upper bound in dependence of the approximation rank $q = 1, \dots, 20$ for a fixed bandwidth $N = 2^{14}$. The error of Algorithm 1 is shown in Figure 5.1(a). The daggers represent the numerical errors, the solid line the theoretical estimate, cf. Lemma A.1, and the dashed line a least-squares fit $E \approx C_0 C^{-q}$, $C > 4$.

Figure 5.1(b),(c) illustrates the error of Algorithm 2 using the butterfly fast Fourier transform (BSFFT) and the nonequispaced fast Fourier transform (NFFT), respectively. The error of Algorithm 2 in combination with the BSFFT can be estimated by Lemma 4.1 and [19, Theorem 3.1], which supports the choice $p = \lceil q/2 \rceil + 8$ for the approximation rank of the BSFFT. Within Algorithm 2, in combination with the NFFT, we choose the Kaiser-Bessel window function, an approximation parameter $m = \lceil q/3 \rceil$, and an oversampling factor 2 for the NFFT; see [18]. In both cases, this leads to an upper bound $C_0 4^{-q}$ with some theoretical constant C_0 as shown by the solid line in Figure 5.1(b), (c), respectively.

In a second series of experiments, we compare the computational times, measured by the MATLAB functions `tic` and `toc`, of the naive evaluations and Algorithm 1 and 2 with respect to an increasing problem size N. Figure 5.2 displays the timings for the naive matrix-vector multiplication using entrywise and row-wise evaluations of the matrix, both shown as
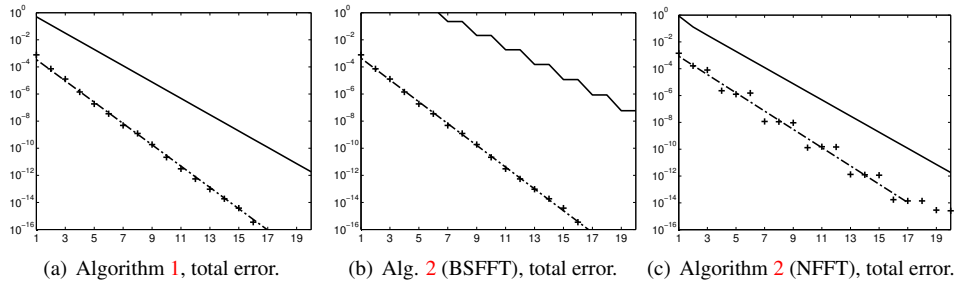
(a) Algorithm 1, total error.   (b) Alg. 2 (BSFFT), total error.   (c) Algorithm 2 (NFFT), total error.

FIG. 5.1. *Approximation error* (5.1) *with respect to the local expansion degree q.*

(a) Algorithm 1, computational times, $q = 8$.   (b) Algorithm 2, computational times, $q = 8$.
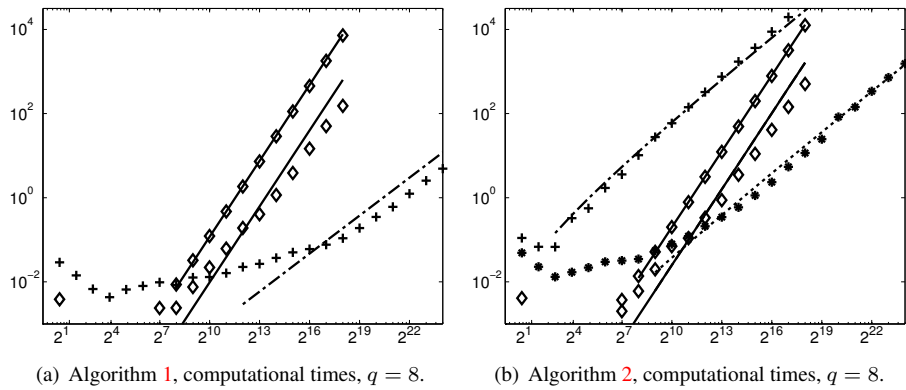
FIG. 5.2. *Computational times for fixed approximation rank $q = 8$ with respect to the problem size N.*

diamonds. Figure 5.2(a) presents the computational times for Algorithm 1 (symbol $+$), i.e., the fast multiplication with the matrix $\boldsymbol{K}$. Clearly, the complexity is linear in the problem size. Figure 5.2(b) displays the same for Algorithm 2 using the BSFFT (symbol $+$) and the NFFT ($*$). Both variants scale almost linear in N, but still the constant in the BSFFT is much larger than in the NFFT case.

6. Summary. Recently, the butterfly approximation scheme and hierarchical approximations have been proposed for the efficient computation of integral transforms with oscillatory or asymptotically smooth kernels. In the first part of this paper, we summarized and slightly improved the fast discrete Laplace transform [22]. We combined the matrix form of this discrete Laplace transform with a generalized fast Fourier transform in a purely algebraic fashion, where we used the decomposition of the Laplace transform explicitly and a small number of generalized FFTs as black box. In particular, this allows for the fast evaluation of a polynomial, given by its monomial coefficients, at many nodes in the complex unit disk. Alternatively, we might interpret this as an FFT with nonequispaced nodes in the upper half plane. In this situation, the butterfly Fourier transform could be replaced by the nonequispaced FFT which is faster both asymptotically as well as with respect to actual computation times.

Appendix A. To keep the paper self contained and to simplify the analysis of [22], we split the error analysis by first showing that the exponential kernel is asymptotically smooth with certain parameter. Together with the admissibility condition, which is purely geometric, this improves [22] slightly by getting rid of a logarithmic term.

LEMMA A.1. *[22, Eq. (36)] Let $q \in \mathbb{N}$, $A, B \subset [0, \infty)$ be admissible, and $\kappa(y, \xi) = e^{-\xi y}$. Then it holds that*

$$\left\| \kappa - \mathcal{I}_q^{A \times B} \kappa \right\|_{C(A \times B)} \leq 2^{1-2q}.$$

Proof. For $y > 0$, we have the necessary condition for a local maximum of $|\partial_y^q \kappa(y, \xi)|$,

$$\left| \partial_\xi \partial_y^q \kappa(y, \xi) \right| = \xi^{q-1} e^{-\xi y} |q - \xi y| = 0,$$

if and only if $\xi = q/y$. Using Stirling's approximation, we conclude the globally valid bound

$$\left| \partial_y^q \kappa(y, \xi) \right| = \xi^q e^{-\xi y} \leq \frac{q^q}{e^q y^q} \leq \frac{1}{\sqrt{2\pi q}} q! y^{-q},$$

i.e., the exponential kernel is asymptotically smooth with constants $C = 1/\sqrt{2\pi}$, $\mu = 1$, $s = 0$, $\nu = -1/2$. The result follows from Theorem 3.1 and $(2 + \frac{2}{\pi} \log q)/\sqrt{2\pi q} \leq 1$. □

LEMMA A.2. *[22, Sect. 4] Let $\varepsilon, y_1, \xi_1 > 0$ be given, use the notation of Definition 3.2 and Lemma A.1, and set $q := \lceil \frac{1}{2} + \log_4 1/\varepsilon \rceil$,*

$$\ell_m := \max(1, \lfloor \log_2(y_1 \xi_1) - m - \log_2(\log 1/\varepsilon) \rfloor + 1), \qquad L_m := M - m$$

for all $m = 1, \ldots, M - 1$. Then

  i) $y \in Y_M$ and $\xi \in \Omega$ (and analogously $y \in Y$ and $\xi \in \Omega_M$) implies $1 - e^{-y\xi} \leq \varepsilon$,
  ii) $y \in Y_m$, $1 \leq m < M$, and $\xi \in \Omega_l$, $l < \ell_m$, implies $e^{-y\xi} \leq \varepsilon$,
  iii) $y \in Y_m$, $1 \leq m < M$, $\xi \in \Omega_l$, $\ell_m \leq l \leq L_m$, implies $\left| e^{-y\xi} - \mathcal{I}_q^{Y_m \times \Omega_l} \kappa(y, \xi) \right| \leq \varepsilon$,
  iv) $y \in Y_m$, $1 \leq m < M$, and $\xi \in \Omega_l$, $l > L_m$, implies $1 - e^{-y\xi} \leq \varepsilon$.

Proof. The individual estimates can be proven as follows. At first, let $\xi \in [0, \xi_1]$ and $y \in Y_M = [0, y_1/2^{M-1}]$. Using $M \geq \log_2 \frac{y_1 \xi_1}{\varepsilon} + 1$, we obtain $0 \leq \xi y \leq y_1 \xi_1 / 2^{M-1} \leq \varepsilon$ and finally case i) since

$$1 \geq e^{-y\xi} \geq e^{-\varepsilon} = \sum_{k=0}^{\infty} \frac{(-\varepsilon)^k}{k!} \geq 1 - \varepsilon.$$

Now let $y \in Y_m$, $\xi \in \Omega_\ell$. The condition $\ell \leq \lfloor \log_2(y_1 \xi_1) - m - \log_2(\log 1/\varepsilon) \rfloor$ implies

$$e^{-\frac{y_1 \xi_1}{2^m 2^\ell}} \leq \varepsilon$$

and due to $y \geq \frac{y_1}{2^m}$, $\xi \geq \frac{\xi_1}{2^\ell}$ assertion ii). The third result follows from Lemma A.1 since the intervals Y_m, Ω_ℓ, $\ell, m = 1, \ldots, M - 1$, are admissible.

Finally, we have $y \in Y_m$, $\xi \in \Omega_\ell$, $\ell - 1 \geq \lceil \log_2(\xi_1 y_1) + \log_2(1/\varepsilon) \rceil - (m - 1)$, and thus

$$1 - e^{-y\xi} \leq 1 - e^{-\frac{y_1}{2^{m-1}} \frac{\xi_1}{2^{\ell-1}}} \leq 1 - e^{-\varepsilon} \leq \varepsilon. \qquad □$$

THEOREM A.3. *[22, Sect. 7] Let $N \in \mathbb{N}$, $\varepsilon > 0$, $\hat{\mathbf{f}} \in \mathbb{C}^N$, $y_1 > y_2 > \ldots > y_N > 0$, $\xi_1 > \xi_2 > \ldots > \xi_N > 0$, and $\kappa(y, \xi) := e^{-\xi y}$ be given. Set $M := \left\lceil \log_2 \frac{y_1 \xi_1}{\varepsilon} \right\rceil + 1$, $q := \lceil \frac{1}{2} + \log_4 1/\varepsilon \rceil$, and $\tilde{\mathbf{f}} = (\tilde{f}(y_i))_{i=1,\ldots,N}$ for the function $\tilde{f} : Y \to \mathbb{C}$,*

$$(A.1) \quad \tilde{f}(y) = \begin{cases} \sum_{j=1}^{N} \hat{f}_j & y \in Y_M, \\ \sum_{\ell=\ell_m}^{L_m} \sum_{\xi_j \in \Omega_\ell} \hat{f}_j \mathcal{I}^{Y_m \times \Omega_\ell} \kappa(y, \xi_j) + \sum_{\ell > L_m} \sum_{\xi_j \in \Omega_\ell} \hat{f}_j & y \in Y_m. \end{cases}$$
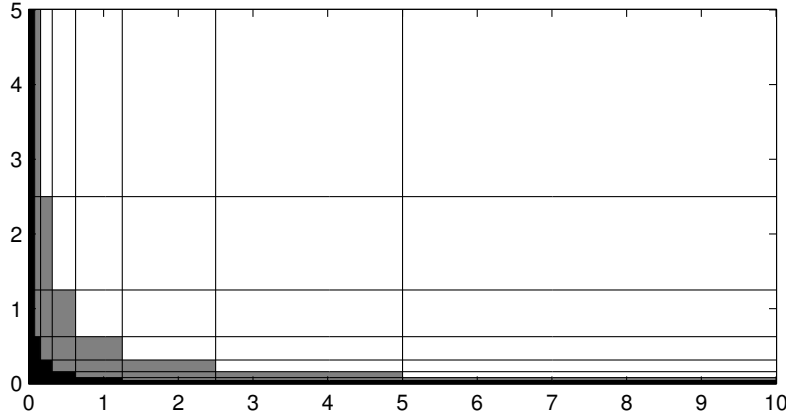
FIG. A.1. *Dyadic decomposition of the domain, the cases i),iv) in black, ii) in white, and iii) in gray, respectively.*

where $1 \leq m < M$. *Then the error estimate* $\|\mathbf{f} - \tilde{\mathbf{f}}\|_\infty \leq \varepsilon \|\hat{\mathbf{f}}\|_1$ *holds true.*

Proof. We start with the error estimate. For $y \in Y_M$, Lemma A.2 i) implies

$$\left| f(y) - \tilde{f}(y) \right| \leq \sum_{k=1}^{N} |\hat{f}_k| |\mathrm{e}^{-\xi_k y} - 1| \leq \varepsilon \sum_{k=1}^{N} |\hat{f}_k|.$$

Now let $m = 1, \ldots, M - 1$, $y \in Y_m$, and partition the function f into three parts

$$f(y) = \left(\sum_{\ell < \ell_m} + \sum_{\ell = \ell_m}^{L_m} + \sum_{\ell > L_m} \right) \sum_{\xi_j \in \Omega_\ell} \hat{f}_j \kappa(y, \xi_j).$$

The desired result follows by the application of Lemma A.2 ii)-iv) and the approximation of the kernel κ by zero, by interpolation, or by one, respectively. □

REFERENCES

[1] F. ANDERSSON, *Algorithms for unequally spaced fast Laplace transforms*, Appl. Comput. Harmon. Anal., 35 (2013), pp. 419–432.

[2] A. A. AYDINER, W. C. CHEW, J. SONG, AND T. J. CUI, *A sparse data fast Fourier transform (SDFFT)*, IEEE Trans. Antennas and Propagation, 51 (2003), pp. 3161–3170.

[3] M. BEBENDORF, *Approximation of boundary element matrices*, Numer. Math., 86 (2000), pp. 565–589.

[4] ———, *Hierarchical Matrices*, Lecture Notes in Computational Science and Engineering 63, Springer, Berlin, 2008.

[5] J.-P. BERRUT AND L. N. TREFETHEN, *Barycentric Lagrange interpolation*, SIAM Rev., 46 (2004), pp. 501–517.

[6] G. BEYLKIN, *On the fast Fourier transform of functions with singularities*, Appl. Comput. Harmon. Anal., 2 (1995), pp. 363–381.

[7] G. BEYLKIN AND L. MONZÓN, *On approximation of functions by exponential sums*, Appl. Comput. Harmon. Anal., 19 (2005), pp. 17–48.

[8] E. CANDÈS, L. DEMANET, AND L. YING, *A fast butterfly algorithm for the computation of Fourier integral operators*, Multiscale Model. Simul., 7 (2009), pp. 1727–1750.

[9] J. W. COOLEY AND J. W. TUKEY, *An algorithm for the machine calculation of complex Fourier series*, Math. Comp., 19 (1965), pp. 297–301.

[10] P. DUHAMEL AND M. VETTERLI, *Fast Fourier transforms: a tutorial review and a state of the art*, Signal Process., 19 (1990), pp. 259–299.

[11] A. DUTT AND V. ROKHLIN, *Fast Fourier transforms for nonequispaced data. II*, Appl. Comput. Harmon. Anal., 2 (1995), pp. 85–100.

[12] B. ELBEL AND G. STEIDL, *Fast Fourier transforms for nonequispaced data*, in Approximation Theory IX, Vol. 2, C. K. Chui and L. L. Schumaker, eds., Innov. Appl. Math., Vanderbilt Univ. Press, Nashville, 1998, pp. 39–46.

[13] M. FRIGO AND S. G. JOHNSON, *The design and implementation of FFTW3*, Proc. IEEE, 93 (2005), pp. 216–231.

[14] L. GRASEDYCK AND W. HACKBUSCH, *Construction and arithmetics of $\mathcal{H}$-matrices*, Computing, 70 (2003), pp. 295–334.

[15] L. GREENGARD AND V. ROKHLIN, *A fast algorithm for particle simulations*, J. Comput. Phys., 73 (1987), pp. 325–348.

[16] W. HACKBUSCH, *A sparse matrix arithmetic based on H-matrices. I. Introduction to H-matrices*, Computing, 62 (1999), pp. 89–108.

[17] ———, *Hierarchische Matrizen. Algorithmen und Analysis*, Springer, Berlin, 2009.

[18] J. KEINER, S. KUNIS, AND D. POTTS, *Using NFFT 3 – a software library for various nonequispaced fast Fourier transforms*, ACM Trans. Math. Software, 36 (2009), Art. 19 (30 pages).

[19] S. KUNIS AND I. MELZER, *A stable and accurate butterfly sparse Fourier transform*, SIAM J. Numer. Anal., 50 (2012), pp. 1777–1800.

[20] E. MICHIELSSEN AND A. BOAG, *A multilevel matrix decomposition algorithm for analyzing scattering from large structures*, IEEE Trans. Antennas and Propagation, 44 (1996), pp. 1086–1093.

[21] J. POULSON, L. DEMANET, N. MAXWELL, AND L. YING, *A parallel butterfly algorithm*, SIAM J. Sci. Comput., 36 (2014), pp. C49–C65.

[22] V. ROKHLIN, *A fast algorithm for the discrete Laplace transformation*, J. Complexity, 4 (1988), pp. 12–32.

[23] G. STEIDL, *A note on fast Fourier transforms for nonequispaced grids*, Adv. Comput. Math., 9 (1998), pp. 337–352.

[24] J. STRAIN, *A fast Laplace transform based on Laguerre functions*, Math. Comp., 58 (1992), pp. 275–283.

[25] X. SUN AND N. P. PITSIANIS, *A matrix version of the fast multipole method*, SIAM Rev., 43 (2001), pp. 289–300.

[26] A. TOWNSEND, M. WEBB, AND S. OLVER, *Fast polynomial transforms based on Toeplitz and Hankel matrices*, Preprint on arXiv, 2016. https://arxiv.org/abs/1604.07486

[27] L. YING, *Sparse Fourier transform via butterfly algorithm*, SIAM J. Sci. Comput., 31 (2009), pp. 1678–1694.

[28] L. YING, G. BIROS, AND D. ZORIN, *A kernel-independent adaptive fast multipole method in two and three dimensions*, J. Comput. Phys., 196 (2004), pp. 591–626.