Noname manuscript No.
(will be inserted by the editor)

Fast discrete convolution in $\mathbb{R}^2$ with radial kernels using Non-Uniform Fast Fourier Transform with nonequispaced frequencies

Martin Averseng

January 21, 2019

Abstract We introduce a new algorithm for the fast evaluation of discrete convolutions with radial kernels in $\mathbb{R}^2$ using the Non-Uniform Fast Fourier Transform. In contrast with other approaches, a Fourier representation of the kernel is obtained with frequency samples lying on concentric circles rather than on a Cartesian grid. As a consequence, we require much less terms to reach a given accuracy. This allows for a faster evaluation of the discrete convolution at the cost of longer precomputations. We provide a full analysis of the complexity and error of our method. Numerical results are reported.

Keywords Fast discrete convolution, Non-Uniform Fast Fourier Transform, Bessel decomposition, Radial functions.

Acknowledgements I wish to express my special gratitude to Professor François Alouges for his patient and very valuable help all along this work, and to Dr. Matthieu Aussal for introducing me to the subtleties of the SCSD algorithm.

Introduction

We describe a fast algorithm for computing discrete convolutions of the form

$$q_k = \sum_{l=1}^{N_z} G(z_k - z_l) f_l \,, \quad k \in \{1, \cdots, N_z\} \,. \tag{1}$$

when G is a radial function, i.e. there exists a function $g : \mathbb{R}^+ \to \mathbb{R}$ such that, for almost all $x \in \mathbb{R}^3$, $G(x) = g(|x|)$, where $|\cdot|$ denotes the Euclidean norm. To simplify the notation, we write in such a case $G(x) = G(r)$ where $r = |x|$. We assume that the nodes $z = (z_k)_{1 \leq k \leq N_z}$ lie inside a disk of radius $\delta_{\max}$ in $\mathbb{R}^2$. Finally $f = (f_k)_{1 \leq k \leq N_z}$ is

CMAP - Ecole Polytechnique, CNRS, Route de Saclay, 91128 Palaiseau Cedex France, E-mail: martin.averseng@polytechnique.edu

a complex vector. For example, in the resolution of the Laplace equation with Dirichlet boundary conditions by the boundary integral method with a Krylov one needs to compute many discrete convolutions like Equation (1) with $G(x) = -\frac{1}{2\pi}\log|x|$, the kernel of the single layer potential (when the kernel is singular, we take the convention $G(0) = 0$ in the convolution (1)). Their computation are a bottleneck in this method. Discrete convolutions also appear in multiple other fields of mathematics and physics, such as particle simulation, biochemistry, tomography etc.

In principle, the effective computation of the vector $q = (q_k)_{1\leq k\leq N_z}$ using Equation (1) requires N_z^2 evaluations of the kernel. However, several more efficient algorithms have emerged to compute an approximation of q with only quasilinear complexity in N_z. Among those is the celebrated Fast Multipole Method (FMM, see for example [8, 9, 14, 22, 23] and references therein). Although this algorithm is very efficient, it suffers from complicated and kernel-specific implementation. For a radial kernel and a 2-dimensional problem, Potts and Steidl [20] introduced an algorithm, which we call here *Fastsum*, that is both simpler and more versatile than the FMM, while achieving comparable performances.

This method relies on the Non-Uniform Fast Fourier Transform (NUFFT, see the seminal paper [12] and also [15, 16, 19, 21] and references therein for tutorials, numerical aspects and open source codes) to convert the convolution (1) into a product in frequency space. The NUFFT takes as inputs a set of N_z nodes z and N_ξ frequency samples ξ in $\mathbb{R}^2$, a complex vector $\alpha \in \mathbb{C}^{N_z}$ and returns the vector $u \in \mathbb{C}^{N_\xi}$ defined by:

$$u_\nu = \sum_{k=1}^{N_z} e^{\pm i z_k\cdot\xi_\nu}\alpha_k\,, \quad \nu \in \{1,\cdots,N_\xi\}\,.$$

We denote the resulting vector u by $\mathrm{NUFFT}_\pm[z,\xi](\alpha)$. This algorithm generalizes the classical Fast Fourier Transform (FFT, [10]) to nonequispaced data, preserving the quasi-linear complexity.

Fastsum exploits the NUFFT as follows. A trigonometric representation of the kernel G is derived:

$$G(x) \approx G_{\mathrm{trig}}(x) := \sum_{\nu=1}^{N_\xi} e^{ix\cdot\xi_\nu}\hat{\omega}_\nu\,, \tag{2}$$

where $(\xi_\nu)_{\nu=1\cdots N_\xi}$ are the frequency samples in $\mathbb{R}^2$ and $\hat{\omega}_\nu$ are complex numbers. This representation is replaced in (1) to yield the fast approximation

$$\begin{aligned} q_k &\approx \left(\sum_{\nu=1}^{N_\xi} e^{+iz_k\cdot\xi_\nu}\left[\hat{\omega}_\nu\sum_{l=1}^{N_z} e^{-iz_l\cdot\xi_\nu}f_l)\right]\right)_{1\leq k\leq N_z} \\ &= \mathrm{NUFFT}_+[z,\xi]\left(\hat{\omega}\odot\mathrm{NUFFT}_-[z,\xi](f)\right), \end{aligned} \tag{3}$$

where $\odot$ denotes the element wise product between vectors.

For 3 space dimensions, an analog method was developed by Alouges and Aussal, named the Sparse Cardinal Sine Decomposition (SCSD). The main difference lies in the choice of the frequency samples in the representation (2). Instead of choosing them on a Cartesian grid like *Fastsum*, the SCSD allows for nonequispaced frequencies and seeks a trigonometric representation with the smallest possible number of

terms. This non-linear and high-dimensional minimization problem is made tractable by exploiting the radial symmetry of the problem.

The aim of this work is to adapt the SCSD method to 2 space dimensions. The cardinal sines must be replaced by Bessel functions and the frequencies are chosen on a set of concentric circles centered at the origin. We obtain trigonometric representations with much fewer terms than in *Fastsum*. Since the frequencies are no longer equispaced, the slower NUFFT of type 3 must be used instead of type 1 (following the terminology of [17]). Nevertheless, the numerical tests, exposed in Section 7, show that the gain in number of frequency samples leads to a faster evaluation of (1), at the cost of longer precomputations.

The remainder of the paper is organized as follows. We first give a self-contained description of our algorithm in Section 1, and state its complexity in Theorem 1 in a special case. Then, Section 2 briefly presents the theory of Fourier-Bessel decomposition. In Section 3, we analyze the "Efficient Bessel Decomposition" (EBD), a new method to approximate singular kernels by Bessel series outside the origin with only a small number of terms. In Section 4, we estimate the number of terms required in the EBD to reach a given accuracy for the logarithmic kernel. We also describe how to deal with other kernels. In Section 5, we show how to convert a Bessel decomposition into a trigonometric polynomial through what we call "circular quadratures". In Section 6, we summarize the complexity of each step and prove Theorem 1. We finally give a numerical comparison of our algorithm with *Fastsum*.

A Matlab code of the method described here is available online [5].

1 Summary of the algorithm

1.1 Trigonometric representation

In this section, given a cut-off parameter $\delta_{\min} > 0$ (also recall $\delta_{\max}$ is the diameter of the set of nodes z), we describe how to select the frequency samples (ξ_ν) in the trigonometric representation (2).

Efficient Bessel Decomposition. We first reduce to a 1-dimensional approximation problem. We find a linear combination of Bessel functions approximating $G(r)$ on $[\delta_{\min}, \delta_{\max}]$:

$$G(r) \approx G(\delta_{\max}) + \sum_{p=1}^{P} \alpha_p J_0\left(\rho_p \frac{r}{\delta_{\max}}\right), \quad r \in [\delta_{\min}, \delta_{\max}], \tag{1.1}$$

where J_0 is the Bessel function of first kind and zero order and (ρ_p) is the sequence of its roots (see Definition 1 for more details). We choose $\alpha_1, \cdots, \alpha_P$ as the minimizers of the least square error

$$E(\alpha_1, \cdots, \alpha_P) = \int_{\delta_{\min}}^{\delta_{\max}} r \left[\frac{d}{dr}\left(G(r) - \sum_{p=1}^{P} \alpha_p J_0\left(\rho_p \frac{r}{\delta_{\max}}\right)\right)\right]^2 dr.$$

The number of terms P is chosen as the smallest integer for which the L^∞ error is below a given tolerance. This amounts to solving a series of linear systems where the matrix coefficients are explicit (see Theorem 2). The right-hand side can be approximated by numerical quadrature when an explicit formula is not available. In Section 3.2, we show how to compute the vector (α_p) with the minimal length without solving a full linear system for each candidate P.

Circular quadrature. In a second step, we write for each $p \in \{1, \cdots, P\}$

$$J_0(\rho_p |x|) \approx \frac{1}{M_p} \sum_{m=0}^{M_p-1} e^{i\rho_p \xi_p^m \cdot x}, \tag{1.2}$$

where $\xi_p^m = \left(\cos\left(\frac{2\pi m}{M_p}\right), \sin\left(\frac{2\pi m}{M_p}\right)\right)$. This is the trapezoidal rule applied to the identity

$$J_0(\rho_p |x|) = \frac{1}{2\pi} \int_{|\xi|=1} e^{i\rho_p x\cdot\xi} d\sigma(\xi).$$

The trigonometric representation (2) is obtained by combining Equations (1.1) and (1.2). A criterion for choosing the number of terms M_p is derived in Section 5.

1.2 Description of the algorithm and complexity

The full algorithm is split into two parts:

Offline part.

Inputs: A radial kernel G, a set of N_z nodes z in $\mathbb{R}^2$ of diameter $\delta_{\max}$, a cut-off parameter $\delta_{\min}$ and a tolerance $\varepsilon > 0$.
Trigonometric representation: Combine the Bessel decomposition and the circular quadrature to obtain the trigonometric representation G_{trig} of G, as in Equation (2). This approximation is valid for $\delta_{\min} \le r \le \delta_{\max}$ up to the tolerance ε.
Correction Matrix: Determine the set $\mathscr{P}$ of all the pairs (k,l) such that $|z_k - z_l| \le \delta_{\min}$ (fixed-radius neighbor search). Assemble the close correction sparse matrix:

$$D_{kl} = \delta_{(k,l)\in\mathscr{P}} \left(G(z_k - z_l) - G(\delta_{\max}) - \sum_{p=1}^{P} \alpha_p J_0\left(\rho_p \frac{|z_k - z_l|}{\delta_{\max}}\right)\right). \tag{1.3}$$

The last two terms inside the parentheses counteract the error introduced by the approximation of G in Bessel series near the origin. The last sum is computed efficiently using a piecewise polynomial interpolation of the function

$$r \mapsto \sum_{p=1}^{P} \alpha_p J_0(\rho_p r), \quad r \in (0, \delta_{\min}).$$

Outputs: The set of weights $(\hat{\omega}_\nu)$, the frequency samples (ξ_ν) and the (sparse) correction matrix D.

Online part.

Inputs: We take in input all outputs of the offline part, and a complex vector f of size N_z.
Far approximation: Compute

$$q_k^{\text{far}} = \sum_{l=1}^{N_z} G_{\text{trig}}(z_k - z_l) f_l, \quad 1 \le k \le N_z \tag{1.4}$$

in the following three steps:

(i) **Space → Fourier:** Compute $\hat{f} = \text{NUFFT}_-[z, \xi](f)$.

(ii) **Fourier multiply:** Perform an element-wise multiplication by $\hat{\omega}$:

$$\hat{g}_\nu := \hat{\omega}_\nu \hat{f}_\nu.$$

(iii) **Fourier → Space:** Compute $q^{\text{far}} = \text{NUFFT}_+[z, \xi](\hat{g})$.

Close correction: Compute the sparse matrix product:

$$q^{\text{close}} = Df\,.$$

Output: The vector $q = q^{\text{far}} + q^{\text{close}}$, with, for any $k \in \{1, \cdots, N_z\}$,

$$\left| q_k - \sum_{l=1}^{N_z} G(z_k - z_l) f_l \right| \le \varepsilon \sum_{l=1}^{N_z} |f_l|\,.$$

For the special case of the logarithmic kernel and for nodes uniformly distributed on a curve, the complexity of the algorithm is given by the following theorem, which is proved in Section 6.

Theorem 1 *Assume the nodes (z_k) are uniformly distributed on a regular curve, and $G(x) = \log|x|$. Let $\varepsilon > 0$ the desired accuracy of the method. Let*

$$a = \frac{1}{N_z^{2/3-\eta}}$$

for some $\eta \in \left[0, \frac{1}{6}\right]$, and choose

$$\delta_{\min} = a\delta_{\max}\,.$$

Then there exists a constant $C > 0$ independent of N_z, ε and η such that:

(i) The number of operations required for the computation of the trigonometric representation (2) *valid for $|x| \in [\delta_{\min}, \delta_{\max}]$ is bounded by*

$$C_{\text{off}}(N_z, \varepsilon, \eta) \le C\,|\log\varepsilon|^3\, N_z^{2-3\eta}\,,$$

(ii) The number of operations required for the assembling of the close correction matrix D is bounded by

$$C_{\text{assemble}}(N_z, \varepsilon, \eta) \le C N_z^{4/3+\eta}\,,$$

(iii) Once these two steps have been completed, the discrete convolution (1) *can be evaluated for any choice of vector f at a precision at least $\varepsilon \sum_l |f_l|$ in a number of operations bounded by*

$$C_{\text{on}}(N_z, \varepsilon, \eta) \leq C \left(N_z^{4/3+\eta} + N_z^{4/3-2\eta} \log(N_z) \left|\log \varepsilon\right|^4 \right).$$

The choice for the parameter a depends on the distribution of nodes z. For example, for data uniformly distributed on a disk, if we choose $a \propto \frac{1}{\sqrt{N_z}}$ we obtain complexities of $O(|\log \varepsilon|^3 N_z^{3/2})$ and $O(|\log \varepsilon|^4 N_z \log(N_z))$ for the offline and online parts respectively.

2 Series of Bessel functions and error estimates

In this section, we give a short introduction to Fourier-Bessel series as an orthonormal decomposition on the eigenfunctions of the Laplace operator, and study the truncation error for smooth functions. We suggest [3,18] and [26, chap. 18] for background. The main result needed for our purpose is Proposition 1.

Our method can be adapted for any space dimension. For example, in $\mathbb{R}^3$ the radial eigenvalues of the Laplace operator are proportional to

$$x \mapsto \frac{J_{1/2}(2\pi p\,|x|)}{|x|^{1/2}}, \quad p \in \mathbb{N}^*.$$

In other words, our approach generalizes [4] to any dimension.

2.1 Radial eigenfunctions of Laplace's operator with Dirichlet boundary conditions

In the following, B denotes the unit ball in $\mathbb{R}^2$ and ∂B its boundary. Let $L^2_{\text{rad}}(B)$ the space of those functions in $L^2(B)$ that are radial and

$$C^\infty_{c,\text{rad}}(B) = \{ \varphi \in C^\infty_c(B) \mid \varphi \text{ is radial } \}.$$

Similarly, let $H^1_{\text{rad}}(B) = H^1(B) \cap L^2_{\text{rad}}(B)$, where $H^1(B)$ is the usual Sobolev space of square integrable functions with square integrable weak derivatives. $H^1_{\text{rad}}(B)$ is a Hilbert space for the norm

$$\|u\|^2_{H^1_{\text{rad}}(B)} = \int_B |u(x)|^2 + |\nabla u(x)|^2 \, dx = 2\pi \int_0^1 r \left(u(r)^2 + u'(r)^2 \right) dr.$$

Finally, $H^1_{0,\text{rad}}(B)$ is the closure of $C^\infty_{c,\text{rad}}(B)$ in $H^1_{\text{rad}}(B)$, with the equivalent norm

$$\|u\|^2_{H^1_{0,\text{rad}}(B)} = 2\pi \int_0^1 r u'(r)^2 dr.$$

We now briefly recall some facts on Bessel functions.

Definition 1 The Bessel function of the first kind and order α, J_α is defined by the power series:

$$J_\alpha(r) := \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\,\Gamma(m+1+\alpha)} \left(\frac{r}{2}\right)^{2m+\alpha}. \tag{2.1}$$

The roots $(\rho_p)_{p\in\mathbb{N}^*}$ of J_0, behave for large p as

$$\rho_p \underset{p\to\infty}{\sim} \pi p\,.$$

For any $p \in \mathbb{N}^*$, we introduce:

$$e_p(x) = C_p J_0(\rho_p\,|x|)\,,$$

where the normalization constant C_p is chosen such that $\left\|e_p\right\|_{H_0^1(B)} = 1$, that is

$$C_p = \frac{1}{\left(2\pi \int_B r\rho_p^2 J_1(\rho_p r)^2\right)^{1/2}} = \frac{1}{\sqrt{\pi}\rho_p\left|J_1(\rho_p)\right|}\,.$$

One can check, using asymptotic expansions of Bessel functions, (see for example equations (10.17.1 - 10.17.3) in [18]) that

$$C_p = \frac{1}{\sqrt{2\pi p}} + O\left(\frac{1}{p^{3/2}}\right), \tag{2.2}$$

For any $p \in \mathbb{N}^*$, e_p satisfies:

$$-\Delta e_p = \rho_p^2 e_p\,. \tag{2.3}$$

In the sequel, we denote by $\mathscr{A}(a)$ the annulus

$$\mathscr{A}(a) = \left\{x \in \mathbb{R}^2 \;\middle|\; |x| \le a\right\}.$$

Theorem 2 *The family $\left\{e_p \mid p \in \mathbb{N}^*\right\}$ is a Hilbert basis of $H^1_{0,\mathrm{rad}}(B)$. For $0 \le a \le 1$, the H_0^1 scalar product on the annulus $\mathscr{A}(a)$ of two eigenfunctions is given by the following formulae:*

$$\int_{\mathscr{A}(a)} \nabla e_i \cdot \nabla e_j = \frac{2\pi C_i C_j \rho_i \rho_j}{\rho_j^2 - \rho_i^2}\left[F_{i,j}(1) - F_{j,i}(1) - F_{i,j}(a) + F_{j,i}(a)\right]$$

if $i \neq j$, where

$$F_{i,j}(r) = \rho_i r J_0(\rho_i r) J_0'(\rho_j r)\,,$$

while

$$\int_{\mathscr{A}(a)} |\nabla e_i|^2 = 2\pi C_i^2\left(F_i(1) - F_i(a)\right)$$

where

$$F_i(r) = \rho_i^2 r^2\left[\frac{1}{2}J_0(\rho_i r)^2 + \frac{1}{2}J_0'(\rho_i r)^2\right] + \rho_i r J_0(\rho_i r) J_0'(\rho_i r)\,.$$

Proof The Laplace operator is self-adjoint, positive definite on $H^1_0(B)$ thus its normalized eigenfunctions form a Hilbert basis of this space. Their expression is given in [13, Eq. (3.9)] in polar coordinates (r,φ) by

$$u_{nkl}(r,\varphi) = J_n(j_{nk}r)\begin{cases}\cos(n\varphi) & l=1,\\ \sin(n\varphi) & l=2\ (n\neq 0).\end{cases}$$

where j_{nk} is the k-th positive root of J_n. The fact that (e_p) is a Hilbert basis of $H^1_{0,\mathrm{rad}}(B)$ stems from the observation that the radial functions are H^1_0 orthogonal to u_{nkl} as soon as $n\neq 0$. The scalar product formulae can be verified by differentiating the functions $F_{i,j}$ and F_i. □

2.2 Truncation error for smooth functions

We now present the Fourier-Bessel series and prove a bound for the truncation error. First of all, Theorem 2 implies that any function $f\in H^1_{0,\mathrm{rad}}(B)$ can be expanded through its so-called Fourier-Bessel series as

$$f=\sum_{p\in\mathbb{N}^*} c_p(f)e_p\,,$$

where the "Fourier-Bessel" coefficients $c_p(f)$ are given by

$$c_p(f)=\int_B \nabla f(x)\cdot\nabla e_p(x)dx=\rho_p^2 C_p\int_0^1 rf(r)J_0(\rho_p r)dr\,.$$

See for example [26, Chap. 18, Eq. (2)]. Notice the second equality is obtained by integration by part. We will show that the coefficients decay rapidly when f is smooth. To this aim, we first introduce the following terminology:

Definition 2 A radial function f satisfies the multi-Dirichlet condition of order $n\in\mathbb{N}^*$ if f is H^{2n} in a neighborhood of ∂B and if for all $s\leq n-1$, the s-th iterate of $-\Delta$ on f, $(-\Delta)^s f$, vanishes on ∂B (with the convention $(-\Delta)^0 f=f$).

An equivalent statement of the following result can be found in [24, Chap. 8, Sec. 20, Thm. 1]. The proof is reproduced here for the reader's convenience.

Proposition 1 *If $f\in H^{2n}(B)$ satisfies the multi-Dirichlet condition of order n, then for any $p\in\mathbb{N}^*$:*

$$c_p(f)=\frac{1}{\rho_p^{2n-2}}\int_B(-\Delta)^n f(x)e_p(x)dx\,.$$

Proof We show this by induction. Let f satisfy the multi-Dirichlet condition of order $n=1$, then by integration by parts:

$$c_p(f)=\int_B(-\Delta)f(x)e_p(x)dx\,,$$

since e_p vanishes on ∂B. Assume the result is true for some $n \geq 1$ and let f satisfy the multi-Dirichlet condition of order $n+1$. Then, using $-\Delta e_p = \rho_p^2 e_p$, we get

$$c_p(f) = \frac{1}{\rho_p^{2n}} \int_B (-\Delta)^n f(x)\ (-\Delta) e_p(x) dx\,.$$

The result follows from integration by parts where we successively use that $(-\Delta)^n f$ and e_p vanish on ∂B. □

Corollary 1 *Let the remainder be defined as*

$$R_P(f) = \sum_{p=P+1}^{+\infty} c_p(f) e_p\,.$$

If $f \in H^{2n}(B)$ satisfies the multi-Dirichlet condition of order n, there exists a constant C independent of n and P such that:

$$\|R_P(f)\|_{H_0^1(B)} \leq C \frac{\|(-\Delta)^n f\|_{L^2(B)}}{(\pi P)^{2n}} \sqrt{\frac{P^3}{n}}\,.$$

Proof Parseval's identity yields $\|R_P(f)\|^2_{H^1_{0,\mathrm{rad}}(B)} = \sum_{p=P+1}^{+\infty} |c_p(f)|^2$. By Proposition 1 and using $\rho_p \sim \pi p$, we can find a constant C such that $\left|c_p(f)\right| \leq C \frac{\|(-\Delta)^n f\|_{L^2}}{(\pi p)^{2n-1}}$. Thus,

$$\|R_P(f)\|_{H^1_{0,\mathrm{rad}}(B)} \leq C \|(-\Delta)^n f\|_{L^2_{\mathrm{rad}}(B)} \sqrt{\sum_{p=P+1}^{+\infty} \frac{1}{(\pi p)^{4n-2}}}\,.$$

The announced result follows from $\displaystyle\sum_{p>P} \frac{1}{p^\alpha} \propto \frac{1}{(\alpha-1)P^{\alpha-1}}$ for $\alpha > 1$. □

2.3 Other boundary conditions

When we replace the Dirichlet boundary condition by the following Robin boundary conditions

$$\frac{\partial u}{\partial n} + Hu = 0 \tag{2.4}$$

for some constant $H \geq 0$, the same analysis can be conducted, leading to Dini series (also covered in [26]). This time, we construct a Hilbert basis of $H^1_{\mathrm{rad}}(B)$ with respect to the bilinear form

$$a_H(u,v) := \int_B \nabla u(x) \cdot \nabla v(x) dx + H \int_{\partial B} u(x) v(x) d\sigma(x).$$

The following result holds.

Theorem 3 *Let $(\rho_p^H)_{p\in\mathbb{N}_*}$ the sequence of positive solutions of*

$$rJ_0'(r) + HJ_0(r) = 0.$$

(i) If $H > 0$, the functions

$$e_p^H(r) = C_p J_0(\rho_p^H r),$$

with C_p such that $a_H(e_p^H, e_p^H) = 1$, form a Hilbert basis of $H^1_{rad}(B)$.

(ii) If $H = 0$, a constant function must be added to the previous family to form a complete set.

The truncation error estimates in Corollary 1 can be generalized to functions satisfying multi-Robin conditions of order $n \geq 1$, that is for all $s \leq n-1$, $(-\Delta)^s u$ satisfies (2.4).

3 Efficient Bessel Decomposition

In this section, we describe the method for approximating a singular function outside the origin by a Fourier-Bessel series with only a few terms.

3.1 Definition

Consider the kernel G involved in (1). We can assume up to rescaling G that the diameter $\delta_{\max}$ of the nodes z is 1, and therefore, we need to approximate the kernel only on the unit ball B. To apply the previous results, two kinds of complications may arise:

(i) When G is singular near the origin, it is not in $H^{2n}(B)$ (even for $n = 1$).
(ii) The multi-Dirichlet conditions may not be fulfilled up to a sufficient order.

The point (ii) is crucial in order to apply the error estimates of the previous section. The first two kernels that we will study (Laplace and Helmholtz) in the next section satisfy the favorable property:

$$\Delta G = \lambda G$$

for some $\lambda \in \mathbb{C}$, is helpful to ensure (ii) at any order. For more general kernels, we suggest a simple trick to enforce multi-Dirichlet conditions up to a given order in Section 4.3.

First, we deal with point (i). We avoid the singularity by approximating G only in the annulus $\mathscr{A}(a)$. We look for coefficients (α_p) such that

$$G(r) \approx \sum_{p=1}^{P} \alpha_p e_p(r), \quad a < r < 1 \tag{3.1}$$

for some $P > 0$, where we remind the reader that e_p are the radial eigenvalues of the Laplace operator on the unit ball defined in Section 2. We call this the Efficient Bessel Decomposition (EBD) of G. The EBD coefficients are chosen as the minimizers of the H^1_0 error on the annulus $\mathscr{A}(a)$:

$$E(\alpha_1, \alpha_2, ..., \alpha_P) = \int_{\mathscr{A}(a)} \left| \nabla \left(G(x) - \sum_{p=1}^{P} \alpha_p e_p(x) \right) \right|^2 dx.$$

Given a tolerance ε, we choose P as the smallest integer for which the L^∞ error of the approximation is below ε. Clearly, for any radial extension $\tilde{G}$ outside the annulus $\mathscr{A}(a)$, one has

$$E(\alpha_1,\cdots,\alpha_P) \leq \left\| \tilde{G} - \sum_{p=1}^{P} c_p(\tilde{G}) e_p(x) \right\|^2_{H_0^1(B)} \tag{3.2}$$

In particular, when $\tilde{G}$ is smooth up to the origin, this gives an immediate error estimate via Corollary 1.

Remark 1 In *Fastsum*, an explicit polynomial regularization and periodisation $\tilde{G}$ is constructed and the function G is replaced by the discrete Fourier expansion of $\tilde{G}$. In contrast here, the way we choose the coefficients α_p ensures that we get less terms than if we had constructed any particular extension $\tilde{G}$ and used $\alpha_p = c_p(\tilde{G})$.

Remark 2 In our context, the H_0^1 error controls the L^∞ norm on $\mathscr{A}(a)$, thus ruling out any risk of Gibb's phenomenon. Indeed, for a function u that vanishes on ∂B, one has

$$|u(x_0)| \leq \sqrt{\frac{-\log|x_0|}{2\pi}} \sqrt{\int_{\mathscr{A}(a)} |\nabla u(x)|^2 dx}, \quad \text{almost for all } x_0 \in \mathscr{A}(a),$$

as can be shown by applying the Cauchy-Schwarz inequality to $u(r) = -\int_r^1 u'(t)dt$.

3.2 Numerical computation of the EBD

Efficient computation of the coefficients. For a given kernel G and integer $P > 0$, the EBD coefficients $\alpha_1,\cdots,\alpha_P$ are found by solving the following linear system:

$$\begin{aligned} &\sum_{q=1}^{P} \rho_q C_q \left(\int_{\mathscr{A}(a)} J_1(\rho_p|x|) J_1(\rho_q|x|) dx \right) \alpha_q \\ &= -\int_{\mathscr{A}(a)} G'(x) J_1(\rho_p|x|) dx, \quad 1 \leq p \leq P, \end{aligned} \tag{3.3}$$

where J_1 is the Bessel function of first kind and order 1 (in fact, $J_0' = -J_1$). The explicit matrix entries are given in Theorem 2.

Using Cholesky decomposition, it is possible to compute efficiently the solution of the system for P' coefficients from the solution for a greater P. Indeed, for $P \in \mathbb{N}^*$, let us call A_P the system matrix and b_P the second member. If C_P is the first Cholesky factor, then $C_{P'}$ is obtained by extracting the top left $P' \times P'$ sub-matrix of C_P. Moreover, $b_{P'}$ is just the vector formed of the first P' entries of b_P. This way, once the tolerance is reached, we can efficiently remove as many terms as possible.

Numerical stability. Numerical evidence strongly suggests that the conditioning number (cond_2) of the linear system (3.3) only depends on the parameter $\gamma := Pa$. This is shown in Figure 1 where we plot the isovalues of the condition number of the system in function of P and γ. One can observe the condition number is almost insensitive to the value of P for constant γ.

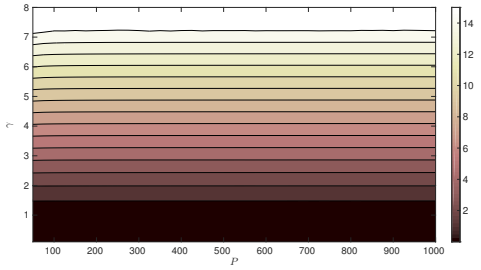

Fig. 1: Isovalues of the function $(P,\gamma) \mapsto \log_{10}(\mathrm{cond}_2(A))$ where A is the matrix of the linear system (3.3) and where $\gamma = Pa$.

4 Error estimates

4.1 Laplace kernel

When solving PDE's involving the Laplace operator (for example in heat conduction or electrostatic problems), one is led to (1) with the logarithmic kernel $G(r) = \log(r)$. Here we show that its EBD converges exponentially fast:

Theorem 4 *There exist two positive constants L_1 and l_2 such that*

$$\forall a \in (0,1), \forall P \in \mathbb{N}^*, \forall r \in (a,1), \quad \left| G(r) - \sum_{p=1}^{P} \alpha_p e_p(r) \right| \leq L_1 e^{-l_2 a P}$$

where $\alpha_1, \cdots, \alpha_P$ are the EBD coefficients of G of order P.

We prove this by exhibiting an extension $\tilde{G}$ of G outside the annulus $\mathscr{A}(a)$. We estimate the truncation error of the Fourier-Bessel series of $\tilde{G}$ in Lemma 2. Finally, Theorem 4 follows by combining Lemma 2 with the estimation (3.2) and using Remark 2.

For any $n \in \mathbb{N}^*$, we define extensions $\tilde{G}_n$ of G by

$$\tilde{G}_n = \begin{cases} r^{2n} \sum_{k=0}^{2n} \dfrac{a_{k,n}}{k!} (r-a)^k & \text{if } r \leq a, \\ G(r) & \text{otherwise,} \end{cases} \tag{4.1}$$

where the values $a_{k,n}$ are chosen so that $\tilde{G}_n$ has continuous derivatives up to order $2n$:

$$a_{k,n} = \frac{d^k}{dr^k} \left(\frac{\log(r)}{r^{2n}} \right) \Bigg|_{r=a}.$$

Observe that the r^{2n} term ensures the boundedness of $(-\Delta)^n \tilde{G}_n$ near the origin. Moreover, since $\tilde{G}_n \equiv G$ in the neighborhood of ∂B, we have $(-\Delta)^s \tilde{G}_n = 0$ on ∂B for any $s \in \mathbb{N}$.

Lemma 1 *There exists a constant C independent of n and a such that for $r \leq a$*

$$\left|\Delta^n \tilde{G}_n(r)\right| \leq C\left(\frac{16n}{e}\right)^{2n} \max_{k\in\{1,\cdots,2n\}} \left(\frac{|a_{k,n}|}{k!}a^k\right). \tag{4.2}$$

Proof For $r \leq a$, we have

$$|(-\Delta)^n \tilde{G}_n(r)| \leq \sum_{k=0}^{2n}\sum_{l=0}^{k}\binom{k}{l}\frac{|a_{k,n}|}{k!}a^{k-l}(2n+l)^2(2(n-1)+l)^2\times\ldots\times(2+l)^2 r^l.$$

Thus

$$\begin{aligned}|(-\Delta)^n \tilde{G}_n(r)| \leq\ & (4n)^2(4n-2)^2\times\ldots\times(2n+2)^2 \\ & \times \max_{k\in[\![0,2n]\!]}\left(\frac{|a_{k,n}|}{k!}a^k\right)\sum_{k=0}^{2n}\sum_{l=0}^{k}\binom{k}{l}a^{-l}r^l\,.\end{aligned}$$

Since $r \leq a$, the last sum is bounded by $\displaystyle\sum_{k=0}^{2n} 2^k = 2^{2n+1}-1 < 2^{2n+1}$, while

$$(4n)^2(4n-2)^2\times\ldots\times(2n+2)^2 \sim 2\left(\frac{8n}{e}\right)^{2n}$$

follows from Stirling formula. □

Lemma 2 *For any $P \in \mathbb{N}^*$ and $a \in (0,1)$, there exists a radial function $\tilde{G}$ which coincides with G on $\mathscr{A}(a)$ satisfying:*

$$\left\|\tilde{G} - \sum_{p=1}^{P} c_p(\tilde{G})e_p\right\|_{H_0^1(B)} \leq C\sqrt{P}\exp\left(-\frac{aP\pi}{32}\right),$$

where C is a positive constant independent of P and a.

Proof Let $n \in \mathbb{N}^*$. By the Leibniz formula, one has

$$a_{k,n} = \frac{(-1)^k k!}{a^{2n+k}}\left(-\sum_{j=0}^{k-1}\frac{\binom{2n+j-1}{j}}{k-j} + \binom{2n+k-1}{k}\log(a)\right).$$

Combining the identity $\sum_{j=0}^{k-1}\binom{j+2n-1}{j} = \frac{k}{2n}\binom{k+2n-1}{k}$ and the estimate

$$\binom{2n+k-1}{k} \leq \binom{4n-1}{2n} = \frac{1}{2}\binom{4n}{2n} \leq \frac{4^{2n}}{2\sqrt{2\pi n}} \quad \text{for } k\in\{1,\cdots,2n\},$$

one finds

$$\max_{0\leq k\leq 2n}\left(\frac{|a_{k,n}|}{k!}a^k\right) \leq \left(\frac{4}{a}\right)^{2n}\frac{1}{2\sqrt{2\pi n}}\left(\log\left(\frac{e}{a}\right)\right). \tag{4.3}$$

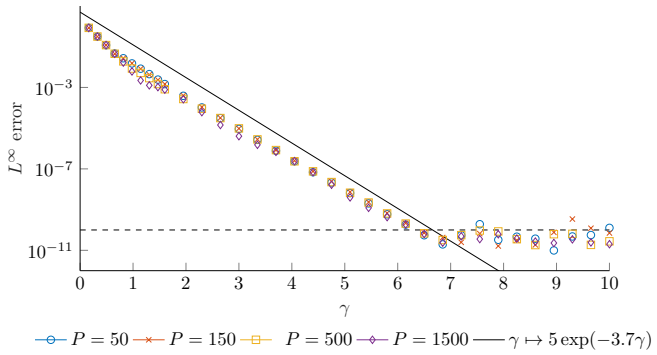

Fig. 2: Evolution of the L^∞ error over $[a,1]$ associated to the EBD for different values of P in function of γ. An exponential decay is indeed observed at the roughly estimated rate of $\propto \exp(-3.7\gamma)$. The error stops decreasing at $e_{\min} \approx 10^{-10}$, for a value of $\gamma \approx 6.7$ because of the ill conditioning of the linear system (3.3)

We apply Lemma 1 to get

$$|(-\Delta)^n \tilde{G}_n(r)| \leq \frac{C}{\sqrt{n}} \left(\frac{16n}{e}\right)^{2n} \left(\frac{4}{a}\right)^{2n} \log\left(\frac{e}{a}\right).$$

Since $(-\Delta)^n \tilde{G}_n(x) = (-\Delta)^n G(x) = 0$ outside $B(0,a)$, we obtain

$$\left\|(-\Delta)^n \tilde{G}_n\right\|_{L^2(B)} \leq \frac{Ca^2}{\sqrt{n}} \log\left(\frac{e}{a}\right) \left(\frac{64n}{ae}\right)^{2n}.$$

The conclusion now follows from Corollary 1. □

Remark 3 We see that the convergence rate is bounded by a function of the parameter $\gamma = Pa$. Figure 2 displays the decay of the L^∞ error with respect to γ for different values of P. We observe that the error consistently decreases at an exponential rate of about $l_2 \approx 3.7$ and stagnates at the minimal error $e_{\min} \approx 10^{-10}$. We believe this stagnation is due to rounding errors related to the increasing condition number of the linear system (3.3).

4.2 Helmholtz kernel

Let Y_0 the classical Bessel function of second kind and of order 0. For any frequency number $k > 0$, the Helmholtz kernel, $r \mapsto \frac{-i}{4} H_0^{(1)}(kr)$, where $H_0^{(1)}(r) = J_0(r) + iY_0(r)$, is the Green kernel associated to the harmonic wave operator $-\Delta - k^2$, that satisfies a Sommerfeld radiation condition at infinity. This kernel arises in various physical problems such as sound waves scattering (see for example [27]). To approximate $H_0^{(1)}(kr)$ as a sum of dilated J_0 functions, it is sufficient to study the function of $r \mapsto Y_0(kr)$. We now have three different cases:

- **When k is a root of Y_0**: In this case the multi-Dirichlet condition is satisfied at any order. Indeed, for any n,
$$(-\Delta)^n Y_0(kr)\big|_{r=1} = k^{2n} Y_0(k) = 0\,.$$
We thus compute the EBD decomposition of Y_0 on an interval $(a,1)$.
- **When k is greater than the first root of Y_0**: we compute the EBD for $r \mapsto Y_0(k'r)$ on $(a,1)$, where k' is the smallest root of Y_0 greater than k. This provides a decomposition for $r \mapsto Y_0(kr)$ valid on $(\frac{k'}{k}a, \frac{k'}{k})$.
- **When k is much smaller than the first root of Y_0**: Here, instead of rescaling, one can use the Bessel-Fourier series associated to the Robin condition (see Section 2.3):
$$\frac{\partial u}{\partial n} + Hu = 0\,,$$
with $H = -\dfrac{kY_0'(k)}{Y_0(k)} > 0$ in this region.

4.3 General kernel : enforcing the multi-Dirichlet condition

For a general radial kernel G, the multi-Dirichlet conditions may not be fulfilled. When applying the EBD method without any changes, this leads to errors near $r = 1$. In this case, we apply the EBD to a modified function $H = G - K$ where K is chosen to enforce the multi-Dirichlet condition. We can take for example
$$K(r) = \sum_{t=1}^{n} \mu_t J_0(\omega_t r)\,,$$
where ω_1 is the root of J_1 that is closest from the ratio $\sqrt{\left|\frac{G(1)}{\Delta G(1)}\right|}$ and $\omega_2, \cdots, \omega_n$, are the subsequent roots of J_1. The coefficients $(\mu_t)_{1\le t\le n}$ are found by solving
$$M\mu = \lambda\,,$$
where λ is the vector given by
$$\lambda_t = (-\Delta)^t G\big|_{r=1}\,, \quad t \in \{1, \cdots, n\}\,,$$
and
$$M = \begin{bmatrix} -\omega_1^2 & -\omega_2^2 & \cdots & -\omega_n^2 \\ \vdots & \vdots & \cdots & \vdots \\ (-1)^n\omega_1^{2n} & (-1)^n\omega_2^{2n} & \cdots & (-1)^n\omega_n^{2n} \end{bmatrix}.$$

In Figure 3, we show the efficiency of this method by applying the EBD with 100 terms to some highly oscillating function ($x \mapsto \log(x) + \sin(250x)$) and computing the maximal error of the decomposition in function of the parameter $\gamma = Pa$ (recall P is the number of terms in the Bessel decomposition and a is the inner radius of the annulus of approximation). We choose the frequencies $(\omega_t)_{1\le t\le n}$ as the roots of J_1 that are closest to 250. This figure illustrates the importance of the multi-Dirichlet condition for the fast decay of the error.

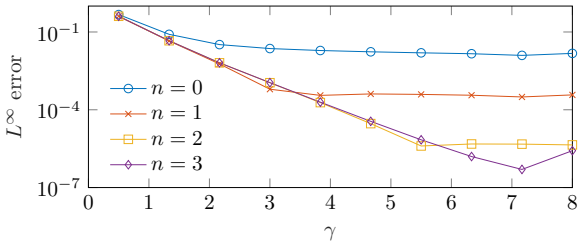

Fig. 3: Maximal error between the kernel $G_1(r) = \log(x) + \sin(250x)$ and its 100-terms EBD in function of $\gamma = Pa$ using the method described in this paragraph for several values of n. Without any change to the method ($n = 0$) the L^∞ error stops decreasing after some value of γ. Enforcing the multi-Dirichlet condition to higher order gradually improves the maximal reachable accuracy

5 Circular quadrature

In this section, we study the approximation

$$J_0(\rho_p|x|) \approx \frac{1}{M_p}\sum_{m=0}^{M_p-1} e^{i\rho_p \xi_m^p \cdot x},$$

where M_p is an integer and with

$$\xi_m^p = \left(\cos\left(\frac{2m\pi}{M_p}\right), \sin\left(\frac{2m\pi}{M_p}\right)\right), \quad 0 \le m \le M_p - 1\,. \tag{5.1}$$

This approximation is obtained by applying the trapezoidal rule to the identity

$$J_0(|x|) = \int_{\partial B} e^{ix\cdot\xi} d\xi\,.$$

Remark 4 In the SCSD method, the Bessel functions are replaced by cardinal sines because

$$\operatorname{sinc}(|x|) = \frac{1}{4\pi}\int_{|\xi|=1} e^{ix\cdot\xi} d\sigma(\xi)\,,$$

where the integral is now taken over $\mathbb{S}^2 \subset \mathbb{R}^3$.

Theorem 5 *There exists a constant K such that for any $r > 0$, for any integer $M \ge \frac{e}{2}r$, and for any $\varphi \in \mathbb{R}$*

$$\left| J_0(r) - \frac{1}{M}\sum_{m=0}^{M-1} e^{ir\sin\left(\frac{2m\pi}{M}-\varphi\right)} \right| \le K\left(\frac{er}{2M}\right)^M.$$

Proof Let $f : x \mapsto e^{ir\sin(x-\varphi)}$. For any $k \in \mathbb{Z}$, one has

$$J_k(r) = \int_0^{2\pi} e^{ir\sin(x)} e^{-ikx} dx = e^{-ik\varphi} \int_0^{2\pi} f(x) e^{-ikx} dx.$$

Thus, the (usual) Fourier coefficient of f are given by

$$c_k(f) = e^{ik\varphi} J_k(r).$$

Consequently, the aliasing formula yields

$$J_0(r) - \frac{1}{M} \sum_{j=0}^{M-1} e^{ir\sin\left(\frac{2j\pi}{M} - \varphi\right)} = - \sum_{k \in \mathbb{Z}^*} e^{iNk\varphi} J_{Nk}(r).$$

For large $|k|$, we have

$$J_k(r) \sim \left(\frac{er}{2|k|} \right)^{|k|}.$$

Therefore, there exists a constant C' such that:

$$\begin{aligned} \left| J_0(r) - \frac{1}{M} \sum_{m=0}^{M-1} e^{ir\sin\left(\frac{2m\pi}{M} - \varphi\right)} \right| &\leq C' \sum_{k \in \mathbb{Z}^*} \left(\frac{er}{2M|k|} \right)^{M|k|} \\ &\leq K \left(\frac{er}{2M} \right)^M \end{aligned}$$

for an appropriate choice of K. □

We conclude with the following result

Proposition 2 *Let $\varepsilon > 0$, $r > 0$, and assume $M > \frac{e}{2} r + \log\left(\frac{K}{\varepsilon}\right)$. Then*

$$\left| J_0(r) - \frac{1}{M} \sum_{m=0}^{M-1} e^{ir\sin\left(\frac{2m\pi}{M} - \varphi\right)} \right| \leq \varepsilon.$$

Proof This results from Theorem 5 and the following inequality:

$$\forall (A,B) \in \left(\mathbb{R}_+^*\right)^2, \quad \left(\frac{A}{A+B} \right)^{A+B} \leq e^{-B}.$$

which may be proved using $\left(1 + x^{-1}\right) \log(1+x) \geq 1$ valid for all $x \neq 0$. □

Our numerical tests suggest that the constant $\frac{e}{2}$ in the previous estimation is sharp.

6 Proof of the complexity

We now sum up all the results obtained in the previous sections to prove the complexity of the complete algorithm given in Section 1. We fix a free parameter $\eta \in [0, 1/6]$ and choose the inner radius of the annulus of approximation as

$$a = \frac{1}{N_z^{2/3-\eta}}. \tag{6.1}$$

We give a bound for the number of operations in function of the number of nodes N_z, the target tolerance ε and η. Let C_{EBD}, C_{circ}, C_{assemble}, C_{far} and C_{close} respectively denote the number of operations required to compute the EBD, the circular quadrature, to assemble the close correction matrix D (1.3), to compute the far approximation (1.4), and to apply D on a vector. In the sequel, C is used for any positive constant that is independent of N_z, ε and η.

6.1 Offline computations

6.1.1 Trigonometric representation

Recall that we compute a trigonometric representation of the form (2) for the kernel $G(r) = \log(r)$, by first computing its EBD:

$$G(x) \approx \sum_{p=1}^{P} \alpha_p e_p(x),$$

where $e_p(r) = C_p J_0(\rho_p r)$ are the normalized eigenfunctions of the Laplace operator and the number of terms P is chosen large enough so that the L^∞ error is below the required tolerance. In a second step, each function e_p is replaced by the circular quadrature analyzed in Section 5.

Efficient Bessel Decomposition. If the EBD is applied on the annulus $\{a < r < 1\}$, Theorem 4 shows that the tolerance $\frac{\varepsilon}{2}$ is reached for

$$P = O\left(\frac{|\log(\varepsilon)|}{a}\right). \tag{6.2}$$

Since the coefficients are obtained through the inversion of a $P \times P$ matrix, the computation of the EBD requires $O(P^3)$ computations. Therefore, with a defined as in (6.1), one has

$$C_{\text{EBD}}(N_z, \varepsilon, \eta) \leq C\,|\log \varepsilon|^3\, N_z^{2-3\eta}. \tag{6.3}$$

Circular quadrature. Using the notation of Section 5, we choose the number M_p of terms in the circular quadrature for each e_p such that

$$\left| e_p(|x|) - \frac{C_p}{M_p} \sum_{m=1}^{M_p} e^{i\rho_p x\cdot\xi_m^p} \right| \leq \frac{\varepsilon}{2P\left|\alpha_p\right|} \quad a<|x|<1\,. \tag{6.4}$$

By the triangle inequality, this ensures that the maximal error in the trigonometric representation of G is less than ε. According to Proposition 2, we may choose M_p such that

$$M_p > \frac{e}{2}\rho_p + \log\left(\frac{2KP|\alpha_p|C_p}{\varepsilon}\right). \tag{6.5}$$

Bessel's identity ensures the boundedness of α_p, and the constants C_p are bounded by estimation (2.2). Since $\rho_p = O(p)$, the total number of terms N_ξ in the trigonometric representation is of the order

$$N_\xi = \sum_{p=1}^{P} M_p = O(P^2) \tag{6.6}$$

and $C_{\text{circ}}(N_z,\varepsilon,\eta) \leq C\left|\log\varepsilon\right|^2 N_z^{4/3-2\eta}$. We thus see that the EBD dominates the complexity for computing the trigonometric representation, and the estimation (6.3) yields the first part of Theorem 1.

6.1.2 Close correction matrix:

Recall that $\delta_{\max}$ is the diameter of the cloud of nodes (z_k). Fix $\delta_{\min} := a\delta_{\max}$. We first determine the set $\mathscr{P}$ of all pairs (k,l) such that $|z_k - z_l| \leq \delta_{\min}$. This is the classical "fixed-radius near neighbors search", and can be solved in $O(N_z \log N_z + \#\mathscr{P})$ operations (see for example [6, 7, 11, 25]). In order to compute the close correction sparse matrix:

$$D_{kl} = \delta_{(k,l)\in\mathscr{P}}\left(G(z_k - z_l) - \sum_{p=1}^{P}\alpha_p J_0\left(\frac{\rho_p}{\delta_{\max}}|z_k - z_l|\right)\right),$$

we need $\#\mathscr{P}$ evaluations of log. For the computation of the second term, we use a piecewise polynomial interpolation of the smooth function.

$$f : r \mapsto \sum_{p=1}^{P}\alpha_p J_0(\rho_p r), \quad r\in(0,a)\,.$$

This involves only a small number of interpolation nodes, and the sum can be evaluated in $O(N_z)$ operations. For a set of nodes uniformly distributed on a curve, we have

$$\#\mathscr{P} = O\left(\frac{\delta_{\min}}{\delta_{\max}}N_z\right) = O(N_z^2 a)\,. \tag{6.7}$$

We thus get

$$C_{\text{close}}(N_z,\varepsilon,\eta) \leq C N_z^{4/3+\eta}\,. \tag{6.8}$$

This is the second part of Theorem 1.

6.2 Online Computations

Far approximation. Recall that for all $k \in \{1, \cdots, N_z\}$, the far approximation is defined by the following equation:

$$q_k^{\text{far}} = \sum_{l=1}^{N_z} G_{\text{trig}}(z_k - z_l) f_l,$$

where

$$G_{\text{trig}}(x) = \sum_{\nu=1}^{N_\xi} \hat{\omega}_\nu e^{ix\cdot\xi_\nu}.$$

To compute q^{far}, recall the following three steps:

(i) **Space** $\to$ **Fourier:** Compute $\hat{f} = \text{NUFFT}_-[z,\xi](f)$,
(ii) **Fourier multiply** Perform element wise multiplication by $\hat{\omega}$:

$$\hat{g}_\nu = \hat{\omega}_\nu \hat{f}_\nu,$$

(iii) **Fourier** $\to$ **Space:** Compute $q^{\text{far}} = \text{NUFFT}_+[z,\xi](\hat{g})$.

According to [16, Sec. 4.3], the complexity of the NUFFT of type 3 is

$$O\left(B_1 B_2 \log(B_1 B_2) + |\log\varepsilon|^2 (N_x + N_\xi)\right)$$

where B_1 and B_2 are the bandwidth parameters along each dimension. Here all frequencies lie in a circle of radius $\rho_P = O(P)$, thus $B_1 B_2 = O(P^2)$. On the other hand, equations (6.1) and (6.6) imply that $N_\xi \geq N_z$, and by (6.6), $N_\xi = O(P^2)$. Therefore, the complexity of the NUFFT is $O(P^2 \log(P) |\log\varepsilon|^2)$. By definition (6.1) and eq. (6.2), we conclude

$$C_{\text{far}}(N_z, \varepsilon, \eta) \leq C N_z^{4/3-2\eta} \log(N_z) |\log\varepsilon|^4. \tag{6.9}$$

Close correction. Since D has $\#\mathscr{P}$ non-zero entries, we get

$$C_{\text{close}}(N_z, \varepsilon, \eta) \leq C N_z^{4/3+\eta}. \tag{6.10}$$

Summing (6.9) and (6.10) yields the second part of Theorem 1.

Remark 5 The extreme cases $\eta = 0$ and $\eta = 1/6$ correspond respectively to the situations where one wishes to either minimize the total (offline + online) computation time or just the online time. The corresponding complexities are given in the table below (omitting the dependence in ε):

	$\eta = 0$	$\eta = 1/6$
Offline	$O(N_z^2)$	$O\left(N_z^{3/2} \log N_z\right)$
Online	$O(N_z^{4/3} \log(N_z))$	$O(N_z^{3/2})$

Table 1: Complexity of the algorithm (omitting dependence in ε) in the two extreme cases $\eta = 0$ and $\eta = 1/6$.

7 Performance of the method and comparison with *Fastsum*

In this section, we apply our method (which we call EBD) to compute the vector q given by

$$q_k = \sum_{l=1}^{N} G(X_k - Y_l) f_l, \quad k = 1, \cdots, N$$

where X and Y are two distinct clouds of points in $\mathbb{R}^2$ of size N. The performance is compared with *Fastsum*, which sources are downloaded from the toolbox [1]. We study three kernels, $G_1(x) = \log(x)$, $G_2(x) = x^2 \log(x)$ and $G_3(x) = \frac{1}{x^2}$. We measure the computing times for N from 10^3 to 10^6. The computer used for these tests is a laptop cadenced to 1.6 GHz and possessing 4GB of memory.

We measure the "offline" time as the time needed to set up the method and compute the first discrete convolution. The "online" time is the average time needed for the subsequent convolutions.

Parameter setting. For *Fastsum*, we use the parameter setting described in [20] for singular kernels. For the first two kernels, we use the cut-off parameter $m = 4$ and the degree of regularization $p = 3$. The approximate error level ε is measured by computing the maximal absolute error between $G(X_1 - Y_l)$ and its approximation by the trigonometric polynomial for $l = 1, \cdots, N$. This is taken as the (input) target tolerance in the EBD. For the third kernel, it is necessary to increase m and p in *Fastsum* to obtain an acceptable error. In such case, we take $p = m$ and choose the first value for which the error goes below 1. In our algorithm, the only free parameter is the inner radius of the annulus of approximation $\delta_{\min}$. Depending on the case, we set $\delta_{\min} = \frac{\lambda}{\sqrt{N}} \delta_{\max}$ with λ ranging from 2 to 20.

Programming language. Our program is written in Matlab except for the NUFFT routine which is the FORTRAN routine borrowed from the NYU website [2]. In contrast, Fastsum is fully written in C. This difference only plays a role in the offline computations, since in the online stage, most of the time is spent in the NUFFT for both methods.

Discussion. We observe that our method yields a representation of the radial kernel with much less frequency samples (a factor 100 is observed in most cases). The offline computations for EBD are always longer (by a factor 2 to 4 in computational intensive cases). After that, our method is faster than *Fastsum* for evaluating the convolution in every case, in general by a factor 2 to 3. This gain in online performance for EBD is particularly useful when one wishes to solve iteratively a large linear system with matrix entries of the form $G(X_k - Y_l)$.

In our algorithm, the 1D approximation allows for an accurate estimation of the error which can therefore be given and controlled as an input. Instead, in *Fastsum*, the parameters need a fine tuning "by hand" in order to achieve a given accuracy.

On the other hand, there is a free parameter in the EBD method, $\delta_{\min}$, that determines the performance. When the distribution of the nodes is known, the above analysis provides a way to select $\delta_{\min}$ to ensure a quasi-linear complexity. However,

Table 2: Computing times (s) for the kernel $G(x) = \log(x)$

	Fastsum				EBD			
N	N_ξ	Offline	Online	ε	N_ξ	Offline	Online	ε
10^3	1.2e4	2.1e-2	5.7e-3	1.3e-3	150	0.31	2.0e-3	1e-3
10^4	1.2e5	0.14	7.3e-2	1.3e-3	2.4e3	0.7	1.3e-2	1.2e-3
10^5	1.2e6	1.4	0.9	1.3e-3	2.8e4	3.8	0.15	1.2e-3
10^6	1.2e7	16	10	1.3e-3	1.2e6	30	4.6	1.3e-3

Table 3: Computing times (s) for the kernel $G(x) = x^2 \log(x)$

	Fastsum				EBD			
N	N_ξ	Offline	Online	ε	N_ξ	Offline	Online	ε
10^3	1.3e4	2.2e-2	5.9e-3	1e-5	166	0.92	2.6e-3	6.3e-6
10^4	1.2e5	0.15	7.4e-2	7.6e-7	1.7e3	0.94	2.5e-2	5.3e-7
10^5	1.2e6	1.5	0.91	2.1e-8	3.2e4	3.8	0.35	2e-8
10^6	1.2e7	16	10	1.5e8	3.7e5	28	8	1.5e8

Table 4: Computing times (s) for the kernel $G(x) = \frac{1}{x^2}$

	Fastsum				EBD			
N	N_ξ	Offline	Online	ε	N_ξ	Offline	Online	ε
10^3	2.1e4	2.5e-2	7e-3	6e-2	1.5e2	0.45	1.8e-3	5e-2
10^4	2.4e5	0.23	0.12	0.17	5.2e3	1.2	1.65e-2	0.14
10^5	2.8e6	3	1.5	0.2	1.5e5	7.8	0.36	0.2
10^6	3.2e7	45	25	0.6	5.2e6	145	16.5	0.5

it is not clear how to choose the optimal value for $\delta_{\min}$ automatically when the distribution of nodes is not known.

Conclusion

The method that we have presented (EBD) generalizes the SCSD method of [4] from 3 to 2 dimensions. Compared to SCSD, the cardinal sines must be replaced by Bessel functions, and the Fourier series appearing in 3D become Fourier-Bessel series in 2D. We have provided a complete analysis of the complexity and the error of our method. It is shown that the computational complexity is quasi-linear in terms of the number of nodes. We also provide numerical results that confirm the theoretical estimations. It has appeared that the structure of SCSD and EBD is very similar to that of *Fastsum*, differing only by the choice of the frequency samples. By exploiting the radial symmetry, our method allows to reduce drastically the number of those samples. As a result, after some precomputations, the EBD yields faster approximations of the discrete convolutions than *Fastsum*. Future work includes generalizing the present approach to n dimensions, and the automatic fine-tuning of the remaining free parameter of the method, $\delta_{\min}$.

References

1. NFFT project. See `https://www-user.tu-chemnitz.de/~potts/nfft/download.php`.
2. NUFFT package. See `https://cims.nyu.edu/cmcl/nufft/nufft.html`.
3. M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation, 1964.
4. F. Alouges and M. Aussal. The sparse cardinal sine decomposition and its application for fast numerical convolution. *Numerical Algorithms*, 70(2):427–448, 2015.
5. M. Averseng. EBD toolbox. See `https://github.com/MartinAverseng/EBD`.
6. J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517, 1975.
7. J. L. Bentley, D. F. Stanat, and E. H. Williams. The complexity of finding fixed-radius near neighbors. *Information processing letters*, 6(6):209–212, 1977.
8. H. Cheng, L. Greengard, and V. Rokhlin. A fast adaptive multipole algorithm in three dimensions. *Journal of computational physics*, 155(2):468–498, 1999.
9. R. Coifman, V. Rokhlin, and S. Wandzura. The fast multipole method for the wave equation: A pedestrian prescription. *IEEE Antennas and Propagation Magazine*, 35(3):7–12, 1993.
10. J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965.
11. M. T. Dickerson and R. S. Drysdale. Fixed-radius near neighbors search algorithms for points and segments. *Information Processing Letters*, 35(5):269–273, 1990.
12. A. Dutt and V. Rokhlin. Fast fourier transforms for nonequispaced data. *SIAM J. Sci. Comput.*, 14(6):1368–1393, November 1993.
13. D. S Grebenkov and B-T Nguyen. Geometrical structure of laplacian eigenfunctions. *SIAM Review*, 55(4):601–667, 2013.
14. L. Greengard. *The rapid evaluation of potential fields in particle systems*. MIT press, 1988.
15. L. Greengard and J. Y. Lee. Accelerating the nonuniform fast fourier transform. *SIAM review*, 46(3):443–454, 2004.
16. J. Keiner, S. Kunis, and D. Potts. Using nfft 3—a software library for various nonequispaced fast fourier transforms. *ACM Transactions on Mathematical Software (TOMS)*, 36(4):19, 2009.
17. June-Yub Lee and Leslie Greengard. The type 3 nonuniform fft and its applications. *Journal of Computational Physics*, 206(1):1–5, 2005.
18. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, and B. V. Saunders. *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.16 of 2017-09-18.
19. D. Potts, G. Pöplau, , and U. van Rienen. Calculation of 3d space-charge fields of bunches of charged particles by fast summation. In *Scientific Computing in Electrical Engineering*, volume 11, pages 241–246. Springer, 2006.
20. D. Potts, G. Steidl, and A. Nieslony. Fast convolution with radial kernels at nonequispaced knots. *Numerische Mathematik*, 98(2):329–351, 2004.
21. D. Potts, G. Steidl, and M. Tasche. Fast fourier transforms for nonequispaced data: A tutorial. In *Modern sampling theory*, pages 247–270. Springer, 2001.
22. V. Rokhlin. Rapid solution of integral equations of scattering theory in two dimensions. *Journal of Computational Physics*, 86(2):414–439, 1990.
23. V. Rokhlin. Diagonal forms of translation operators for the helmholtz equation in three dimensions. *Applied and Computational Harmonic Analysis*, 1(1):82–93, 1993.
24. G. P. Tolstov. *Fourier series*. Courier Corporation, 2012.
25. V. Turau. Fixed-radius near neighbors search. *Information processing letters*, 39(4):201–203, 1991.
26. G. N. Watson. *A treatise on the theory of Bessel functions*. Cambridge university press, 1995.
27. C. H. Wilcox. *Scattering theory for the d'Alembert equation in exterior domains*, volume 4. Springer Berlin, 1975.