


Fast and Accurate Approximation of High-Dimensional Radial Kernels via Slicing

Michael Quellmalz | TU Berlin | Workshop on Approximation Theory and Fast Algorithms, 9 April 2025
Joint work with Johannes Hertrich, Tim Jahn, Nicolaj Rux and Gabriele Steidl

The Problem

- Given: Points $x_n, y_m \in \mathbb{R}^d$ and coefficients $w_n \in \mathbb{R}$ for $n = 1, \dots, N$, $m = 1, \dots, M$
- Radial kernel $K\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$

$$K(x, y) = F(\|x - y\|)$$

  with basis function $F\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$

- We want: the kernel sums

$$s_m = \sum_{n=1}^{N} w_n K(x_n, y_m) \qquad \forall\, m = 1, \dots, M \tag{1}$$

Examples	$F(\|x\|)$
Gauss	$\exp(-\frac{\|x\|^2}{2\sigma^2})$
Matérn	$\frac{2^{1-\nu}}{\Gamma(\nu)}(\frac{\sqrt{2\nu}}{\beta}x)^\nu K_\nu(\frac{\sqrt{2\nu}}{\beta}x)$
Laplace	$\exp(-\alpha\|x\|)$
Riesz	$-\|x\|^r$

- Goal: Improve computational complexity $\mathcal{O}(MN)$

The Problem

Technische Universität Berlin

- Given: Points $x_n, y_m \in \mathbb{R}^d$ and coefficients $w_n \in \mathbb{R}$ for $n = 1, ..., N$, $m = 1, ..., M$
- Radial kernel $K\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$

$$K(x, y) = F(\|x - y\|)$$

  with basis function $F\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$

- We want: the kernel sums

$$s_m = \sum_{n=1}^{N} w_n K(x_n, y_m) \qquad \forall\, m = 1, ..., M \tag{1}$$

Examples	$F(\|x\|)$
Gauss	$\exp(-\frac{\|x\|^2}{2\sigma^2})$
Matérn	$\frac{2^{1-\nu}}{\Gamma(\nu)}(\frac{\sqrt{2\nu}}{\beta}x)^\nu K_\nu(\frac{\sqrt{2\nu}}{\beta}x)$
Laplace	$\exp(-\alpha\|x\|)$
Riesz	$-\|x\|^r$

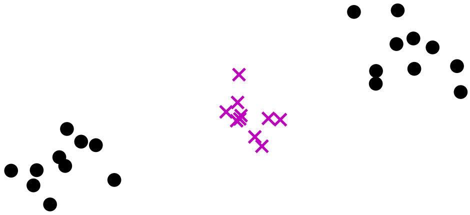
- Goal: Improve computational complexity $\mathcal{O}(MN)$

Motivational Example: Particle Flows

- Given: $\boldsymbol{y} = (y_1, \dots, y_M) \in (\mathbb{R}^d)^M$
- Find: $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$ that "well approximate" $\boldsymbol{y}$

$$\operatorname*{argmin}_{\boldsymbol{x}} \underbrace{-\frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\|x_i - x_j\|}_{\text{repulsion}} + \underbrace{\frac{1}{MN}\sum_{i=1}^{N}\sum_{m=1}^{M}\|x_i - y_m\|}_{\text{attraction}}$$

- Gradient descent approximation:

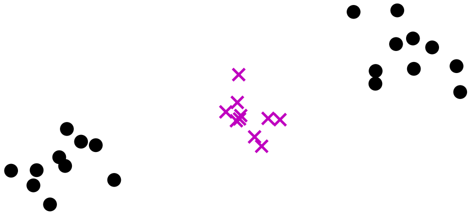


Motivational Example: Particle Flows

- Given: $\boldsymbol{y} = (y_1, \dots, y_M) \in (\mathbb{R}^d)^M$
- Find: $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$ that "well approximate" $\boldsymbol{y}$

$$\operatorname*{argmin}_{\boldsymbol{x}} \underbrace{-\frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\|x_i - x_j\|}_{\text{repulsion}} + \underbrace{\frac{1}{MN}\sum_{i=1}^{N}\sum_{m=1}^{M}\|x_i - y_m\|}_{\text{attraction}}$$

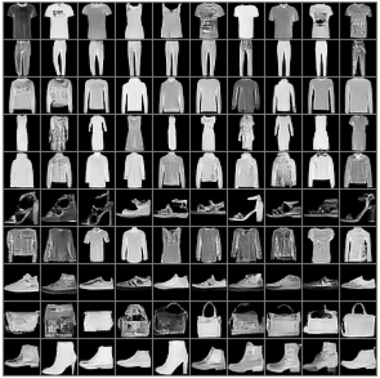
- Gradient descent approximation:



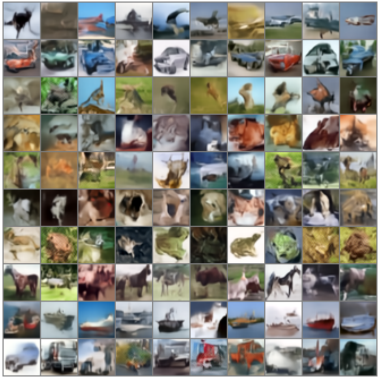
High-Dimensional Example: Neural MMD Gradient Flows

[Hagemann Hertrich Altekrüger Beinert Chemseddine Steidl 2024]

MNIST



FashionMNIST



CIFAR10

Figure: Class-conditional samples of MNIST, FashionMNIST and CIFAR10.
Each point $x_n \in \mathbb{R}^{784}$ corresponds to a 28×28 image (CIFAR10: $d = 3072$)

Applications of the Kernel Summation Problem

- Kernel density estimation [Parzen 1962] [Rosenblatt 1956]
- Radial basis function approximation [Buhmann 2003] [Wendland 2004]
- Classification via support vector machines [Steinwart Christmann 2008]
- Dimensionality reduction with kernelized principal component analysis [Schölkopf Smola 2002] [Shawe-Taylor Cristianini 2004]
- Dithering and halftoning of images [Teuber Steidl Gwosdek Schmaltz Weickert 2010] [Krahmer Veselovska 2024]
- Electrostatic particle simulation [Nestler Pippig Potts 2015]
- Maximum mean discrepancies or the energy distance on the space of probability measure [Gretton Borgwardt Rasch Schölkopf Smola 2006] [Székely 2002]
- Corresponding gradient flows [Arbel Korba Salim Gretton 2019] [Galashov Bortoli Gretton 2024] [Hagemann Hertrich Altekrüger Beinert Chemseddine Steidl 2024] [Kolouri Nadjahi Shahrampour Simsekli 2022]
- Methods for Bayesian inference like Stein variational gradient descent [Liu Wang 2016]

Approach 1: Fast Fourier Summation

[Kunis Potts 2006] [Potts Steidl Nieslony 2004]

- Fourier expansion of the kernel with degree $N_{\mathrm{ft}} \in \mathbb{N}$

$$F(\|x\|) \approx \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i} \langle x, k \rangle}$$

- Insert to (1)

$$s_m = \sum_{n=1}^{N} w_n F(\|x_n - y_m\|) \approx \sum_{n=1}^{N} w_n \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i} \langle x_n - y_m, k \rangle}$$

$$= \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{-2\pi \mathrm{i} \langle y_m, k \rangle} \underbrace{\sum_{n=1}^{N} w_n \mathrm{e}^{2\pi \mathrm{i} \langle x_n, k \rangle}}_{=: \hat{w}_k}$$

1. Inner sum $\hat{w}_k$ is **non-equispaced fast Fourier transform (NFFT)** of w
2. Outer sum s_m is NFFT of $c_k \hat{w}_k$ $\rightsquigarrow$ Complexity $\mathcal{O}(N + M + N_{\mathrm{ft}}^d \log N_{\mathrm{ft}})$

☺ Efficient for $d \in \{1, 2, 3\}$

☹ Becomes unfeasible for larger d

Approach 1: Fast Fourier Summation

[Kunis Potts 2006] [Potts Steidl Nieslony 2004]

Technische Universität Berlin

- Fourier expansion of the kernel with degree $N_{\mathrm{ft}} \in \mathbb{N}$

$$F(\|x\|) \approx \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i} \langle x, k \rangle}$$

- Insert to (1)

$$s_m = \sum_{n=1}^{N} w_n F(\|x_n - y_m\|) \approx \sum_{n=1}^{N} w_n \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i} \langle x_n - y_m, k \rangle}$$

$$= \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{-2\pi \mathrm{i} \langle y_m, k \rangle} \underbrace{\sum_{n=1}^{N} w_n \mathrm{e}^{2\pi \mathrm{i} \langle x_n, k \rangle}}_{=: \hat{w}_k}$$

1. Inner sum $\hat{w}_k$ is **non-equispaced fast Fourier transform (NFFT)** of w
2. Outer sum s_m is NFFT of $c_k \hat{w}_k$ $\rightsquigarrow$ Complexity $\mathcal{O}(N + M + N_{\mathrm{ft}}^d \log N_{\mathrm{ft}})$

☺ Efficient for $d \in \{1, 2, 3\}$

☹ Becomes unfeasible for larger d

Approach 1: Fast Fourier Summation

[Kunis Potts 2006] [Potts Steidl Nieslony 2004]

- Fourier expansion of the kernel with degree $N_{\mathrm{ft}} \in \mathbb{N}$

$$F(\|x\|) \approx \sum_{k \in \{-N_{\mathrm{ft}}, \ldots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i} \langle x, k \rangle}$$

- Insert to (1)

$$s_m = \sum_{n=1}^{N} w_n F(\|x_n - y_m\|) \approx \sum_{n=1}^{N} w_n \sum_{k \in \{-N_{\mathrm{ft}}, \ldots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i} \langle x_n - y_m, k \rangle}$$

$$= \sum_{k \in \{-N_{\mathrm{ft}}, \ldots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{-2\pi \mathrm{i} \langle y_m, k \rangle} \underbrace{\sum_{n=1}^{N} w_n \mathrm{e}^{2\pi \mathrm{i} \langle x_n, k \rangle}}_{=: \hat{w}_k}$$

1. Inner sum $\hat{w}_k$ is **non-equispaced fast Fourier transform (NFFT)** of w
2. Outer sum s_m is NFFT of $c_k \hat{w}_k$ $\rightsquigarrow$ Complexity $\mathcal{O}(N + M + N_{\mathrm{ft}}^d \log N_{\mathrm{ft}})$

☺ Efficient for $d \in \{1, 2, 3\}$

😐 Becomes unfeasible for larger d

Approach 1: Fast Fourier Summation

[Kunis Potts 2006] [Potts Steidl Nieslony 2004]

- Fourier expansion of the kernel with degree $N_{\mathrm{ft}} \in \mathbb{N}$

$$F(\|x\|) \approx \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i} \langle x, k \rangle}$$

- Insert to (1)

$$\begin{aligned} s_m = \sum_{n=1}^{N} w_n F(\|x_n - y_m\|) &\approx \sum_{n=1}^{N} w_n \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i} \langle x_n - y_m, k \rangle} \\ &= \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{-2\pi \mathrm{i} \langle y_m, k \rangle} \underbrace{\sum_{n=1}^{N} w_n \mathrm{e}^{2\pi \mathrm{i} \langle x_n, k \rangle}}_{=: \hat{w}_k} \end{aligned}$$

1. Inner sum $\hat{w}_k$ is **non-equispaced fast Fourier transform (NFFT)** of w
2. Outer sum s_m is NFFT of $c_k \hat{w}_k$ $\rightsquigarrow$ Complexity $\mathcal{O}(N + M + N_{\mathrm{ft}}^d \log N_{\mathrm{ft}})$

☺ Efficient for $d \in \{1, 2, 3\}$

😐 Becomes unfeasible for larger d

Approach 1: Fast Fourier Summation

[Kunis Potts 2006] [Potts Steidl Nieslony 2004]

- Fourier expansion of the kernel with degree $N_{\mathrm{ft}} \in \mathbb{N}$

$$F(\|x\|) \approx \sum_{k \in \{-N_{\mathrm{ft}},\dots,N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i}\langle x,k\rangle}$$

- Insert to (1)

$$s_m = \sum_{n=1}^{N} w_n F(\|x_n - y_m\|) \approx \sum_{n=1}^{N} w_n \sum_{k \in \{-N_{\mathrm{ft}},\dots,N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i}\langle x_n - y_m,k\rangle}$$

$$= \sum_{k \in \{-N_{\mathrm{ft}},\dots,N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{-2\pi \mathrm{i}\langle y_m,k\rangle} \underbrace{\sum_{n=1}^{N} w_n \mathrm{e}^{2\pi \mathrm{i}\langle x_n,k\rangle}}_{=:\hat{w}_k}$$

❶ Inner sum $\hat{w}_k$ is **non-equispaced fast Fourier transform (NFFT)** of w

❷ Outer sum s_m is NFFT of $c_k \hat{w}_k$ $\rightsquigarrow$ Complexity $\mathcal{O}(N + M + N_{\mathrm{ft}}^d \log N_{\mathrm{ft}})$

☺ Efficient for $d \in \{1,2,3\}$

☹ Becomes unfeasible for larger d

Approach 1: Fast Fourier Summation

[Kunis Potts 2006] [Potts Steidl Nieslony 2004]

- Fourier expansion of the kernel with degree $N_{\mathrm{ft}} \in \mathbb{N}$

$$F(\|x\|) \approx \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i} \langle x, k \rangle}$$

- Insert to (1)

$$s_m = \sum_{n=1}^{N} w_n F(\|x_n - y_m\|) \approx \sum_{n=1}^{N} w_n \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{2\pi \mathrm{i} \langle x_n - y_m, k \rangle}$$

$$= \sum_{k \in \{-N_{\mathrm{ft}}, \dots, N_{\mathrm{ft}}\}^d} c_k \mathrm{e}^{-2\pi \mathrm{i} \langle y_m, k \rangle} \underbrace{\sum_{n=1}^{N} w_n \mathrm{e}^{2\pi \mathrm{i} \langle x_n, k \rangle}}_{=: \hat{w}_k}$$

❶ Inner sum $\hat{w}_k$ is **non-equispaced fast Fourier transform (NFFT)** of w

❷ Outer sum s_m is NFFT of $c_k \hat{w}_k$ $\rightsquigarrow$ Complexity $\mathcal{O}(N + M + N_{\mathrm{ft}}^d \log N_{\mathrm{ft}})$

☺ Efficient for $d \in \{1, 2, 3\}$

😐 Becomes unfeasible for larger d

Approach 2: Random Fourier features (RFF)

[Rahimi Recht 2007]

Technische Universität Berlin

- Assumption: $F(\|\cdot\|)$ positive definite on $\mathbb{R}^d$ and $F(0)=1$
- By **Bochner's theorem**, $F(\|\cdot\|)$ is the Fourier transform of a probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$
- Draw iid samples $v_1, \dots, v_P$ from μ

$$F(\|x\|) = \int_{\mathbb{R}^d} \mathrm{e}^{-2\pi \mathrm{i}\langle x, v\rangle} \,\mathrm{d}\mu(v) \approx \frac{1}{P} \sum_{p=1}^{P} \mathrm{e}^{2\pi \mathrm{i}\langle x, v_p\rangle}$$

- Insert to (1)

$$s_m = \sum_{n=1}^{N} w_n F(\|x_n - y_m\|) \approx \sum_{n=1}^{N} w_n \frac{1}{P} \sum_{p=1}^{P} \mathrm{e}^{2\pi \mathrm{i}\langle x_n - y_m, v_p\rangle}$$

$$= \frac{1}{P} \sum_{p=1}^{P} \mathrm{e}^{-2\pi \mathrm{i}\langle y_m, v_p\rangle} \sum_{n=1}^{N} w_n \mathrm{e}^{2\pi \mathrm{i}\langle x_n, v_p\rangle}$$

☺ Complexity $\rightsquigarrow \mathcal{O}(P(N+M))$

☺ Simple to implement

😐 Error bound $\mathbb{E}[\|F - \tilde{F}\|_\infty] \in \mathcal{O}(1/\sqrt{P})$ [Sutherland Schneider 2015]

☹ Only for positive definite functions

Kernel	positive definite
Gaussian	✓
Matérn	✓
Laplacian	✓
Riesz	✗

Approach 2: Random Fourier features (RFF)

[Rahimi Recht 2007]

Technische Universität Berlin

- Assumption: $F(\|\cdot\|)$ positive definite on $\mathbb{R}^d$ and $F(0)=1$
- By **Bochner's theorem**, $F(\|\cdot\|)$ is the Fourier transform of a probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$
- Draw iid samples $v_1, \ldots, v_P$ from μ

$$F(\|x\|) = \int_{\mathbb{R}^d} \mathrm{e}^{-2\pi \mathrm{i}\langle x, v\rangle} \,\mathrm{d}\mu(v) \approx \frac{1}{P} \sum_{p=1}^{P} \mathrm{e}^{2\pi \mathrm{i}\langle x, v_p\rangle}$$

- Insert to (1)

$$\begin{aligned} s_m = \sum_{n=1}^{N} w_n F(\|x_n - y_m\|) &\approx \sum_{n=1}^{N} w_n \frac{1}{P} \sum_{p=1}^{P} \mathrm{e}^{2\pi \mathrm{i}\langle x_n - y_m, v_p\rangle} \\ &= \frac{1}{P} \sum_{p=1}^{P} \mathrm{e}^{-2\pi \mathrm{i}\langle y_m, v_p\rangle} \sum_{n=1}^{N} w_n \mathrm{e}^{2\pi \mathrm{i}\langle x_n, v_p\rangle} \end{aligned}$$

- ☺ Complexity $\rightsquigarrow \mathcal{O}(P(N+M))$
- ☺ Simple to implement
- 😐 Error bound $\mathbb{E}[\|F - \tilde{F}\|_\infty] \in \mathcal{O}(1/\sqrt{P})$ [Sutherland Schneider 2015]
- ☹ Only for positive definite functions

Kernel	positive definite
Gaussian	✓
Matérn	✓
Laplacian	✓
Riesz	✗

Approach 2: Random Fourier features (RFF)

[Rahimi Recht 2007]

- Assumption: $F(\|\cdot\|)$ positive definite on $\mathbb{R}^d$ and $F(0)=1$
- By **Bochner's theorem**, $F(\|\cdot\|)$ is the Fourier transform of a probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$
- Draw iid samples $v_1,\ldots,v_P$ from μ

$$F(\|x\|) = \int_{\mathbb{R}^d} \mathrm{e}^{-2\pi \mathrm{i}\langle x, v\rangle}\,\mathrm{d}\mu(v) \approx \frac{1}{P}\sum_{p=1}^{P} \mathrm{e}^{2\pi\mathrm{i}\langle x, v_p\rangle}$$

- Insert to (1)

$$\begin{aligned} s_m = \sum_{n=1}^{N} w_n F(\|x_n - y_m\|) &\approx \sum_{n=1}^{N} w_n \frac{1}{P}\sum_{p=1}^{P} \mathrm{e}^{2\pi\mathrm{i}\langle x_n - y_m, v_p\rangle} \\ &= \frac{1}{P}\sum_{p=1}^{P} \mathrm{e}^{-2\pi\mathrm{i}\langle y_m, v_p\rangle} \sum_{n=1}^{N} w_n \mathrm{e}^{2\pi\mathrm{i}\langle x_n, v_p\rangle} \end{aligned}$$

☺ Complexity $\rightsquigarrow \mathcal{O}(P(N+M))$

☺ Simple to implement

😐 Error bound $\mathbb{E}[\|F-\tilde{F}\|_\infty] \in \mathcal{O}(1/\sqrt{P})$ [Sutherland Schneider 2015]

☹ Only for positive definite functions

Kernel	positive definite
Gaussian	☑
Matérn	☑
Laplacian	☑
Riesz	✖

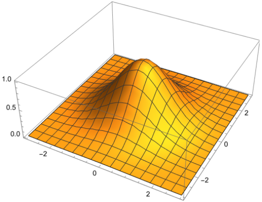
Approach 3: Slicing Summation

[Hertrich 2024]

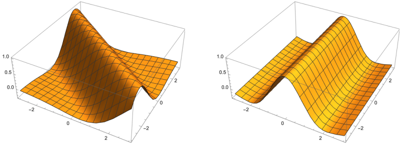
Technische Universität Berlin

- Idea: Construct $f\colon \mathbb{R} \to \mathbb{R}$ s.t.

$$F(\|x\|) = \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f(\langle x, \xi \rangle) \,\mathrm{d}\xi \qquad \forall\, x \in \mathbb{R}^d$$



radial function $F(\|x\|)$ (Gaussian)

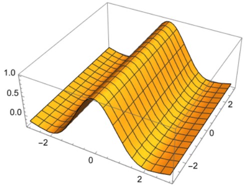


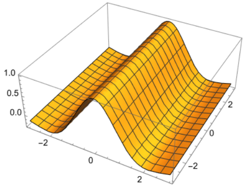
$x \mapsto f(\langle x, \xi \rangle)$ for 2 different ξ

[Hertrich 2024]

- Idea: Construct $f\colon \mathbb{R} \to \mathbb{R}$ s.t.

$$F(\|x\|) = \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f(\langle x, \xi \rangle) \,\mathrm{d}\xi \qquad\qquad \forall\, x \in \mathbb{R}^d$$





Top: Approximation of F
Bottom: $x \mapsto f(\langle x, \xi \rangle)$

Approach 3: Slicing Summation

- Idea: Construct $f\colon \mathbb{R} \to \mathbb{R}$ s.t.

$$F(\|x\|) = \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f(\langle x, \xi \rangle) \,\mathrm{d}\xi \qquad \forall\, x \in \mathbb{R}^d$$

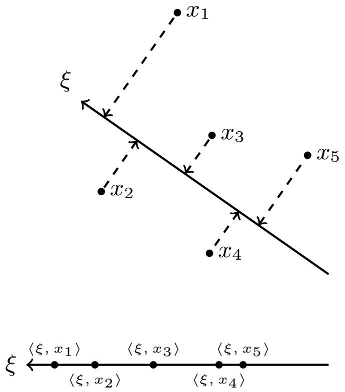
- Insert into the sum (1)

$$s_m = \sum_{n=1}^{N} w_n \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f(\langle x_n, \xi \rangle - \langle y_m, \xi \rangle) \,\mathrm{d}\xi$$

Sample $\xi_p \sim \mathcal{U}_{\mathbb{S}^{d-1}}$ (uniform measure on the sphere)

$$\approx \sum_{p=1}^{P} \underbrace{\sum_{n=1}^{N} w_n f(\langle x_n, \xi_p \rangle - \langle y_m, \xi_p \rangle)}_{\text{1D kernel sum}}$$

- P one-dimensional sums $\rightsquigarrow \mathcal{O}(P(N + M + N_{\mathrm{ft}} \log N_{\mathrm{ft}}))$



Projection of points $x_1, \ldots, x_5 \in \mathbb{R}^2$ onto the line in direction ξ

- Idea: Construct $f\colon \mathbb{R} \to \mathbb{R}$ s.t.

$$F(\|x\|) = \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f(\langle x, \xi \rangle)\,\mathrm{d}\xi \qquad \forall\, x \in \mathbb{R}^d$$

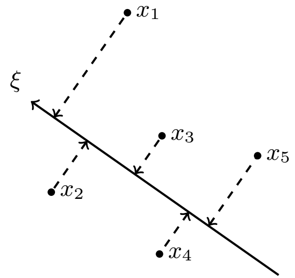
- Insert into the sum (1)

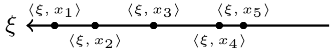
$$s_m = \sum_{n=1}^{N} w_n \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f(\langle x_n, \xi \rangle - \langle y_m, \xi \rangle)\,\mathrm{d}\xi$$

Sample $\xi_p \sim \mathcal{U}_{\mathbb{S}^{d-1}}$ (uniform measure on the sphere)

$$\approx \sum_{p=1}^{P} \underbrace{\sum_{n=1}^{N} w_n f(\langle x_n, \xi_p \rangle - \langle y_m, \xi_p \rangle)}_{\text{1D kernel sum}}$$

- P one-dimensional sums $\rightsquigarrow \mathcal{O}(P(N + M + N_{\mathrm{ft}} \log N_{\mathrm{ft}}))$





Projection of points $x_1, \ldots, x_5 \in \mathbb{R}^2$ onto the line in direction ξ

The Slicing Transform

Goal: Construct "sliced kernel function" $f \colon \mathbb{R} \to \mathbb{R}$ s.t.

$$F(\|x\|) = \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f(\langle x, \xi \rangle)\, \mathrm{d}\xi \qquad \forall\, x \in \mathbb{R}^d$$

Theorem ($f \mapsto F$) [Hertrich 2024] [Rubin 2003]

For $d \geq 2$, a pair (F, f) of basis functions in $L^\infty_{\mathrm{loc}}(\mathbb{R}_{\geq 0})$ fulfills this relation if and only if F is the **generalized Riemann–Liouville fractional integral**

$$F(t) = \frac{2\Gamma(\frac{d}{2})}{\sqrt{\pi}\Gamma(\frac{d-1}{2})} \int_0^1 f(ts)(1 - s^2)^{\frac{d-3}{2}}\, \mathrm{d}s.$$

Theorem (Smoothness of F) [Rux Q. Steidl 2025]

For $d \in \mathbb{N}$ with $d \geq 3$, let $f \in L^1_{\mathrm{loc}}([0, \infty))$ for odd d and $f \in L^p_{\mathrm{loc}}([0, \infty))$ with $p > 2$ for even d.
Then F is $\lfloor \frac{(d-2)}{2} \rfloor$-times continuously differentiable on $(0, \infty)$.

The Slicing Transform

Goal: Construct "sliced kernel function" $f\colon \mathbb{R} \to \mathbb{R}$ s.t.

$$F(\|x\|) = \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f(\langle x, \xi \rangle) \,\mathrm{d}\xi \qquad \forall\, x \in \mathbb{R}^d$$

Theorem ($f \mapsto F$) [Hertrich 2024] [Rubin 2003]

For $d \geq 2$, a pair (F, f) of basis functions in $L^\infty_{\mathrm{loc}}(\mathbb{R}_{\geq 0})$ fulfills this relation if and only if F is the **generalized Riemann–Liouville fractional integral**

$$F(t) = \frac{2\Gamma(\frac{d}{2})}{\sqrt{\pi}\Gamma(\frac{d-1}{2})} \int_0^1 f(ts)(1-s^2)^{\frac{d-3}{2}} \,\mathrm{d}s.$$

Theorem (Smoothness of F) [Rux Q. Steidl 2025]

For $d \in \mathbb{N}$ with $d \geq 3$, let $f \in L^1_{\mathrm{loc}}([0,\infty))$ for odd d and $f \in L^p_{\mathrm{loc}}([0,\infty))$ with $p > 2$ for even d. Then F is $\lfloor \frac{(d-2)}{2} \rfloor$-times continuously differentiable on $(0, \infty)$.

Inversion Formula for L^1 Functions

- **Rotation operator** $\mathcal{R}_d$ maps $F\colon [0,\infty) \to \mathbb{R}$ to the radial function

$$\mathcal{R}_d F(x) := F(\|x\|), \qquad x \in \mathbb{R}^d$$

- **Spherical averaging operator** maps $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ to

$$\mathcal{A}_d \Phi(r) := \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} \Phi(r\,\xi)\,\mathrm{d}\xi, \qquad r \in \mathbb{R}$$

- **Multiplication operator**

$$\mathcal{M}_d F(r) := F(|r|)|r|^{d-1}, \quad r \in \mathbb{R}$$

Theorem [Rux Q. Steidl 2025]

Let $d \geq 2$. Let $F\colon [0,\infty) \to \mathbb{R}$ fulfill $\mathcal{R}_d F \in L^1(\mathbb{R}^d)$ and $\mathcal{F}_d(\mathcal{R}_d F) \in L^1(\mathbb{R}^d)$, then

$$f = \frac{|\mathbb{S}^{d-1}|}{2} (\mathcal{F}_1 \circ \mathcal{M}_d \circ \mathcal{A}_d \circ \mathcal{F}_d \circ \mathcal{R}_d)[F].$$

☺ Inversion formula uses only one- and d-dimensional Fourier transform and simple operators

☹ Not applicable for the Riesz kernel: f exits, but the Fourier transform $\mathcal{F}_d\mathcal{R}_d[F]$ is only a distribution

Inversion Formula for L^1 Functions

- **Rotation operator** $\mathcal{R}_d$ maps $F\colon [0,\infty) \to \mathbb{R}$ to the radial function

$$\mathcal{R}_d F(x) := F(\|x\|), \qquad x \in \mathbb{R}^d$$

- **Spherical averaging operator** maps $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ to

$$\mathcal{A}_d \Phi(r) := \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} \Phi(r\,\xi)\,\mathrm{d}\xi, \qquad r \in \mathbb{R}$$

- **Multiplication operator**

$$\mathcal{M}_d F(r) := F(|r|)|r|^{d-1}, \quad r \in \mathbb{R}$$

Theorem [Rux Q. Steidl 2025]

Let $d \geq 2$. Let $F\colon [0,\infty) \to \mathbb{R}$ fulfill $\mathcal{R}_d F \in L^1(\mathbb{R}^d)$ and $\mathcal{F}_d(\mathcal{R}_d F) \in L^1(\mathbb{R}^d)$, then

$$f = \frac{|\mathbb{S}^{d-1}|}{2} (\mathcal{F}_1 \circ \mathcal{M}_d \circ \mathcal{A}_d \circ \mathcal{F}_d \circ \mathcal{R}_d)[F].$$

☺ Inversion formula uses only one- and d-dimensional Fourier transform and simple operators

☹ Not applicable for the Riesz kernel: f exits, but the Fourier transform $\mathcal{F}_d \mathcal{R}_d[F]$ is only a distribution

Inversion Formula for L^1 Functions

- **Rotation operator** $\mathcal{R}_d$ maps $F\colon [0,\infty) \to \mathbb{R}$ to the radial function

$$\mathcal{R}_d F(x) \coloneqq F(\|x\|), \qquad x \in \mathbb{R}^d$$

- **Spherical averaging operator** maps $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ to

$$\mathcal{A}_d \Phi(r) \coloneqq \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} \Phi(r\,\xi)\,\mathrm{d}\xi, \qquad r \in \mathbb{R}$$

- **Multiplication operator**

$$\mathcal{M}_d F(r) \coloneqq F(|r|)|r|^{d-1}, \quad r \in \mathbb{R}$$

Theorem [Rux Q. Steidl 2025]

Let $d \geq 2$. Let $F\colon [0,\infty) \to \mathbb{R}$ fulfill $\mathcal{R}_d F \in L^1(\mathbb{R}^d)$ and $\mathcal{F}_d(\mathcal{R}_d F) \in L^1(\mathbb{R}^d)$, then

$$f = \frac{|\mathbb{S}^{d-1}|}{2} (\mathcal{F}_1 \circ \mathcal{M}_d \circ \mathcal{A}_d \circ \mathcal{F}_d \circ \mathcal{R}_d)[F].$$

☺ Inversion formula uses only one- and d-dimensional Fourier transform and simple operators

☹ Not applicable for the Riesz kernel: f exits, but the Fourier transform $\mathcal{F}_d\mathcal{R}_d[F]$ is only a distribution

Inversion Formula for L^1 Functions

- **Rotation operator** $\mathcal{R}_d$ maps $F\colon [0,\infty) \to \mathbb{R}$ to the radial function

$$\mathcal{R}_d F(x) := F(\|x\|), \qquad x \in \mathbb{R}^d$$

- **Spherical averaging operator** maps $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ to

$$\mathcal{A}_d \Phi(r) := \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} \Phi(r\,\xi)\,\mathrm{d}\xi, \qquad r \in \mathbb{R}$$

- **Multiplication operator**

$$\mathcal{M}_d F(r) := F(|r|)|r|^{d-1}, \quad r \in \mathbb{R}$$

Theorem [Rux Q. Steidl 2025]

Let $d \geq 2$. Let $F\colon [0,\infty) \to \mathbb{R}$ fulfill $\mathcal{R}_d F \in L^1(\mathbb{R}^d)$ and $\mathcal{F}_d(\mathcal{R}_d F) \in L^1(\mathbb{R}^d)$, then

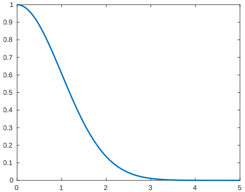
$$f = \frac{|\mathbb{S}^{d-1}|}{2} (\mathcal{F}_1 \circ \mathcal{M}_d \circ \mathcal{A}_d \circ \mathcal{F}_d \circ \mathcal{R}_d)[F].$$

☺ Inversion formula uses only one- and d-dimensional Fourier transform and simple operators

☹ Not applicable for the Riesz kernel: f exits, but the Fourier transform $\mathcal{F}_d \mathcal{R}_d[F]$ is only a distribution

Example: Gauss Kernel

Gauss kernel	$F(t) = \exp\left(-\frac{t^2}{2}\right)$
Radial function	$\mathcal{R}_d F(x) = \exp\left(-\frac{\|x\|^2}{2}\right)$
Fourier transform	$\mathcal{F}_d^{-1}\mathcal{R}_d F(x) = (2\pi)^{d/2}\exp(-2\pi^2\|x\|^2)$
radial part	$\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)$
Multiplication	$\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)\|t\|^{d-1}$
Sliced kernel	$f(t) = \mathcal{F}_1\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = {}_1F_1\left(\frac{d}{2}, \frac{1}{2}, -\frac{t^2}{2}\right)$ (confluent hypergeometric function)

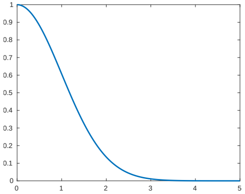


F

Figure: Gauss kernel F and the corresponding sliced kernel f for $d = 10$

Example: Gauss Kernel

Gauss kernel	$F(t) = \exp\left(-\frac{t^2}{2}\right)$
Radial function	$\mathcal{R}_d F(x) = \exp\left(-\frac{\|x\|^2}{2}\right)$
Fourier transform	$\mathcal{F}_d^{-1}\mathcal{R}_d F(x) = (2\pi)^{d/2}\exp(-2\pi^2\|x\|^2)$
radial part	$\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)$
Multiplication	$\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)\|t\|^{d-1}$
Sliced kernel	$f(t) = \mathcal{F}_1\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = {}_1F_1\left(\frac{d}{2}, \frac{1}{2}, -\frac{t^2}{2}\right)$ (confluent hypergeometric function)

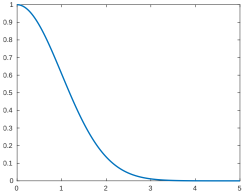


F

Figure: Gauss kernel F and the corresponding sliced kernel f for $d = 10$

Example: Gauss Kernel

Gauss kernel	$F(t) = \exp\left(-\frac{t^2}{2}\right)$
Radial function	$\mathcal{R}_d F(x) = \exp\left(-\frac{\|x\|^2}{2}\right)$
Fourier transform	$\mathcal{F}_d^{-1}\mathcal{R}_d F(x) = (2\pi)^{d/2}\exp(-2\pi^2\|x\|^2)$
radial part	$\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)$
Multiplication	$\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)\lvert t\rvert^{d-1}$
Sliced kernel	$f(t) = \mathcal{F}_1\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = {}_1F_1\left(\frac{d}{2}, \frac{1}{2}, -\frac{t^2}{2}\right)$ (confluent hypergeometric function)

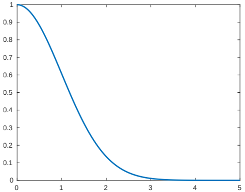


F

Figure: Gauss kernel F and the corresponding sliced kernel f for $d = 10$

Example: Gauss Kernel

Gauss kernel	$F(t) = \exp\left(-\frac{t^2}{2}\right)$
Radial function	$\mathcal{R}_d F(x) = \exp\left(-\frac{\|x\|^2}{2}\right)$
Fourier transform	$\mathcal{F}_d^{-1}\mathcal{R}_d F(x) = (2\pi)^{d/2}\exp(-2\pi^2\|x\|^2)$
radial part	$\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)$
Multiplication	$\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)\lvert t\rvert^{d-1}$
Sliced kernel	$f(t) = \mathcal{F}_1\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = {}_1F_1\left(\frac{d}{2}, \frac{1}{2}, -\frac{t^2}{2}\right)$ (confluent hypergeometric function)

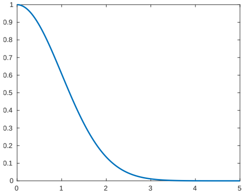


F

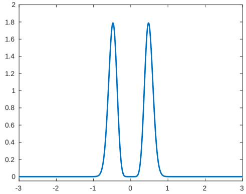
Figure: Gauss kernel F and the corresponding sliced kernel f for $d = 10$

Example: Gauss Kernel

Gauss kernel	$F(t) = \exp\left(-\frac{t^2}{2}\right)$
Radial function	$\mathcal{R}_d F(x) = \exp\left(-\frac{\|x\|^2}{2}\right)$
Fourier transform	$\mathcal{F}_d^{-1}\mathcal{R}_d F(x) = (2\pi)^{d/2}\exp(-2\pi^2\|x\|^2)$
radial part	$\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)$
Multiplication	$\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)\|t\|^{d-1}$
Sliced kernel	$f(t) = \mathcal{F}_1\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = {}_1F_1\left(\frac{d}{2}, \frac{1}{2}, -\frac{t^2}{2}\right)$ (confluent hypergeometric function)



F



$\mathcal{M}_d\rho = \mathcal{F}_1 f$

Figure: Gauss kernel F and the corresponding sliced kernel f for $d = 10$

Example: Gauss Kernel

Gauss kernel	$F(t) = \exp\left(-\frac{t^2}{2}\right)$
Radial function	$\mathcal{R}_d F(x) = \exp\left(-\frac{\|x\|^2}{2}\right)$
Fourier transform	$\mathcal{F}_d^{-1}\mathcal{R}_d F(x) = (2\pi)^{d/2}\exp(-2\pi^2\|x\|^2)$
radial part	$\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)$
Multiplication	$\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)\lvert t\rvert^{d-1}$
Sliced kernel	$f(t) = \mathcal{F}_1\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = {}_1F_1\left(\frac{d}{2}, \frac{1}{2}, -\frac{t^2}{2}\right)$ (confluent hypergeometric function)

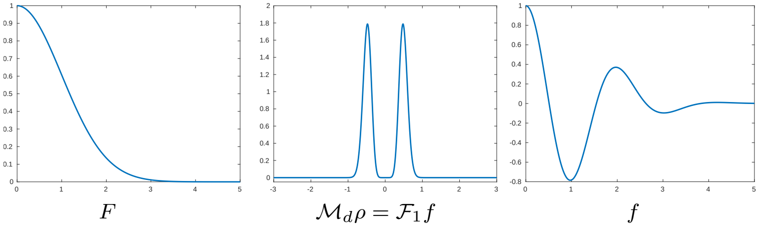


Figure: Gauss kernel F and the corresponding sliced kernel f for $d = 10$

Slicing via Tempered Distributions

- **Schwartz space** $\mathcal{S}(\mathbb{R}^d) = \{\varphi \in C^\infty(\mathbb{R}^d) : D^\alpha x^\beta \varphi(x) \in C_0(\mathbb{R}^d)\, \forall \alpha, \beta \in \mathbb{N}^d\}$
- **Tempered distributions** $\mathcal{S}'(\mathbb{R}^d)$ are continuous linear functionals on $\mathcal{S}(\mathbb{R}^d)$
- Fourier transform of $T \in \mathcal{S}'(\mathbb{R}^d)$ is defined as "adjoint"

$$\langle \mathcal{F}_d T, \varphi \rangle \coloneqq \langle T, \mathcal{F}_d \varphi \rangle \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

- Define **"adjoint operators"**

$$\mathcal{R}_d^\star \colon \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'_{\mathrm{rad}}(\mathbb{R}), \qquad \langle \mathcal{R}_d^\star T, \psi \rangle \coloneqq \langle T, (\mathcal{R}_d \circ \mathcal{A}_1)\psi \rangle \quad \forall \psi \in \mathcal{S}(\mathbb{R})$$
$$\mathcal{A}_d^* \colon \mathcal{S}'(\mathbb{R}) \to \mathcal{S}'_{\mathrm{rad}}(\mathbb{R}^d), \qquad \langle \mathcal{A}_d^* \tau, \varphi \rangle \coloneqq \langle \tau, \mathcal{A}_d \varphi \rangle \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

Theorem [Rux Q. Steidl 2025]

Let $f \in C(\mathbb{R})$ be slowly increasing and even. Set $F(\|x\|) = \mathbb{E}_{\xi \sim \mathcal{U}_{\mathbb{S}^{d-1}}}[f(\langle x, \xi \rangle)]$. Then f can be recovered from

$$f \coloneqq (\mathcal{F}_1 \circ \mathcal{R}_d^\star \circ \mathcal{F}_d^{-1})[\mathcal{R}_d F]. \tag{$\star$}$$

Theorem (Conditions on F) [Rux Q. Steidl 2025]

Let $d \geq 3$ and let the $\lfloor d/2 \rfloor$-th derivative of $F \in \mathcal{C}^{\lfloor d/2 \rfloor}([0,\infty))$ be slowly increasing. Then the sliced kernel f is well-defined and satisfies ($\star$).

Slicing via Tempered Distributions

- **Schwartz space** $\mathcal{S}(\mathbb{R}^d) = \{\varphi \in C^\infty(\mathbb{R}^d) : D^\alpha x^\beta \varphi(x) \in C_0(\mathbb{R}^d)\, \forall \alpha, \beta \in \mathbb{N}^d\}$
- **Tempered distributions** $\mathcal{S}'(\mathbb{R}^d)$ are continuous linear functionals on $\mathcal{S}(\mathbb{R}^d)$
- Fourier transform of $T \in \mathcal{S}'(\mathbb{R}^d)$ is defined as "adjoint"

$$\langle \mathcal{F}_d T, \varphi \rangle \coloneqq \langle T, \mathcal{F}_d \varphi \rangle \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

- Define **"adjoint operators"**

$$\mathcal{R}_d^\star \colon \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'_{\mathrm{rad}}(\mathbb{R}), \qquad \langle \mathcal{R}_d^\star T, \psi \rangle \coloneqq \langle T, (\mathcal{R}_d \circ \mathcal{A}_1)\psi \rangle \quad \forall \psi \in \mathcal{S}(\mathbb{R})$$

$$\mathcal{A}_d^* \colon \mathcal{S}'(\mathbb{R}) \to \mathcal{S}'_{\mathrm{rad}}(\mathbb{R}^d), \qquad \langle \mathcal{A}_d^* \tau, \varphi \rangle \coloneqq \langle \tau, \mathcal{A}_d \varphi \rangle \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

Theorem [Rux Q. Steidl 2025]

Let $f \in C(\mathbb{R})$ be slowly increasing and even. Set $F(\|x\|) = \mathbb{E}_{\xi \sim \mathcal{U}_{\mathbb{S}^{d-1}}}[f(\langle x, \xi \rangle)]$. Then f can be recovered from

$$f \coloneqq (\mathcal{F}_1 \circ \mathcal{R}_d^\star \circ \mathcal{F}_d^{-1})[\mathcal{R}_d F]. \qquad (\star)$$

Theorem (Conditions on F) [Rux Q. Steidl 2025]

Let $d \geq 3$ and let the $\lfloor d/2 \rfloor$-th derivative of $F \in C^{\lfloor d/2 \rfloor}([0, \infty))$ be slowly increasing. Then the sliced kernel f is well-defined and satisfies $(\star)$.

Slicing via Tempered Distributions

- **Schwartz space** $\mathcal{S}(\mathbb{R}^d) = \left\{\varphi \in C^\infty(\mathbb{R}^d) : D^\alpha x^\beta \varphi(x) \in C_0(\mathbb{R}^d)\, \forall \alpha, \beta \in \mathbb{N}^d\right\}$
- **Tempered distributions** $\mathcal{S}'(\mathbb{R}^d)$ are continuous linear functionals on $\mathcal{S}(\mathbb{R}^d)$
- Fourier transform of $T \in \mathcal{S}'(\mathbb{R}^d)$ is defined as "adjoint"

$$\langle \mathcal{F}_d T, \varphi \rangle \coloneqq \langle T, \mathcal{F}_d \varphi \rangle \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

- Define **"adjoint operators"**

$$\begin{aligned} \mathcal{R}_d^\star &\colon \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'_{\mathrm{rad}}(\mathbb{R}), & \langle \mathcal{R}_d^\star T, \psi \rangle &\coloneqq \langle T, (\mathcal{R}_d \circ \mathcal{A}_1)\psi \rangle \quad \forall \psi \in \mathcal{S}(\mathbb{R}) \\ \mathcal{A}_d^* &\colon \mathcal{S}'(\mathbb{R}) \to \mathcal{S}'_{\mathrm{rad}}(\mathbb{R}^d), & \langle \mathcal{A}_d^* \tau, \varphi \rangle &\coloneqq \langle \tau, \mathcal{A}_d \varphi \rangle \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d) \end{aligned}$$

Theorem [Rux Q. Steidl 2025]

Let $f \in C(\mathbb{R})$ be slowly increasing and even. Set $F(\|x\|) = \mathbb{E}_{\xi \sim \mathcal{U}_{\mathbb{S}^{d-1}}}[f(\langle x, \xi \rangle)]$. Then f can be recovered from

$$f \coloneqq (\mathcal{F}_1 \circ \mathcal{R}_d^\star \circ \mathcal{F}_d^{-1})[\mathcal{R}_d F]. \tag{$\star$}$$

Theorem (Conditions on F) [Rux Q. Steidl 2025]

Let $d \geq 3$ and let the $\lfloor d/2 \rfloor$-th derivative of $F \in C^{\lfloor d/2 \rfloor}([0, \infty))$ be slowly increasing. Then the sliced kernel f is well-defined and satisfies ($\star$).

Slicing via Tempered Distributions

- **Schwartz space** $\mathcal{S}(\mathbb{R}^d) = \left\{\varphi \in C^\infty(\mathbb{R}^d) : D^\alpha x^\beta \varphi(x) \in C_0(\mathbb{R}^d)\, \forall \alpha, \beta \in \mathbb{N}^d\right\}$
- **Tempered distributions** $\mathcal{S}'(\mathbb{R}^d)$ are continuous linear functionals on $\mathcal{S}(\mathbb{R}^d)$
- Fourier transform of $T \in \mathcal{S}'(\mathbb{R}^d)$ is defined as "adjoint"

$$\langle \mathcal{F}_d T, \varphi\rangle \coloneqq \langle T, \mathcal{F}_d \varphi\rangle \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

- Define **"adjoint operators"**

$$\mathcal{R}_d^\star \colon \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'_{\mathrm{rad}}(\mathbb{R}), \qquad \langle \mathcal{R}_d^\star T, \psi\rangle \coloneqq \langle T, (\mathcal{R}_d \circ \mathcal{A}_1)\psi\rangle \quad \forall \psi \in \mathcal{S}(\mathbb{R})$$
$$\mathcal{A}_d^* \colon \mathcal{S}'(\mathbb{R}) \to \mathcal{S}'_{\mathrm{rad}}(\mathbb{R}^d), \qquad \langle \mathcal{A}_d^* \tau, \varphi\rangle \coloneqq \langle \tau, \mathcal{A}_d \varphi\rangle \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

Theorem [Rux Q. Steidl 2025]

Let $f \in C(\mathbb{R})$ be slowly increasing and even. Set $F(\|x\|) = \mathbb{E}_{\xi\sim\mathcal{U}_{\mathbb{S}^{d-1}}}[f(\langle x, \xi\rangle)]$. Then f can be recovered from

$$f \coloneqq (\mathcal{F}_1 \circ \mathcal{R}_d^\star \circ \mathcal{F}_d^{-1})[\mathcal{R}_d F]. \qquad (\star)$$

Theorem (Conditions on F) [Rux Q. Steidl 2025]

Let $d \geq 3$ and let the $\lfloor d/2\rfloor$-th derivative of $F \in C^{\lfloor d/2\rfloor}([0,\infty))$ be slowly increasing. Then the sliced kernel f is well-defined and satisfies $(\star)$.

Slicing via Tempered Distributions

- **Schwartz space** $\mathcal{S}(\mathbb{R}^d) = \{\varphi \in C^\infty(\mathbb{R}^d) : D^\alpha x^\beta \varphi(x) \in C_0(\mathbb{R}^d)\, \forall \alpha, \beta \in \mathbb{N}^d\}$
- **Tempered distributions** $\mathcal{S}'(\mathbb{R}^d)$ are continuous linear functionals on $\mathcal{S}(\mathbb{R}^d)$
- Fourier transform of $T \in \mathcal{S}'(\mathbb{R}^d)$ is defined as "adjoint"

$$\langle \mathcal{F}_d T, \varphi \rangle \coloneqq \langle T, \mathcal{F}_d \varphi \rangle \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

- Define **"adjoint operators"**

$$\mathcal{R}_d^\star \colon \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'_{\mathrm{rad}}(\mathbb{R}), \qquad \langle \mathcal{R}_d^\star T, \psi \rangle \coloneqq \langle T, (\mathcal{R}_d \circ \mathcal{A}_1)\psi \rangle \quad \forall \psi \in \mathcal{S}(\mathbb{R})$$
$$\mathcal{A}_d^* \colon \mathcal{S}'(\mathbb{R}) \to \mathcal{S}'_{\mathrm{rad}}(\mathbb{R}^d), \qquad \langle \mathcal{A}_d^* \tau, \varphi \rangle \coloneqq \langle \tau, \mathcal{A}_d \varphi \rangle \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d)$$

Theorem [Rux Q. Steidl 2025]

Let $f \in C(\mathbb{R})$ be slowly increasing and even. Set $F(\|x\|) = \mathbb{E}_{\xi \sim \mathcal{U}_{\mathbb{S}^{d-1}}}[f(\langle x, \xi \rangle)]$. Then f can be recovered from

$$f \coloneqq (\mathcal{F}_1 \circ \mathcal{R}_d^\star \circ \mathcal{F}_d^{-1})[\mathcal{R}_d F]. \qquad (\star)$$

Theorem (Conditions on F) [Rux Q. Steidl 2025]

Let $d \geq 3$ and let the $\lfloor d/2 \rfloor$-th derivative of $F \in \mathcal{C}^{\lfloor d/2 \rfloor}([0, \infty))$ be slowly increasing. Then the sliced kernel f is well-defined and satisfies ($\star$).

Known Pairs (F, f)

Kernel	$F(x)$	$\mathcal{F}_d^{-1}[F(\|\cdot\|)](\|\omega\|)$	$f(x)$	$\mathcal{F}_1^{-1}[f](\lvert\omega\rvert)$
Gauss	$\exp(-\frac{x^2}{2\sigma^2})$	$(2\pi\sigma^2)^{d/2}\exp(-2\pi^2\sigma^2\omega^2)$	${}_1F_1(\frac{d}{2};\frac{1}{2};\frac{-x^2}{2\sigma^2})$	$\frac{\pi\sigma\exp(-2\pi^2\sigma^2\omega^2)(2\pi^2\sigma^2\omega^2)^{(d-1)/2}}{2\Gamma(\frac{d}{2})}$
Laplace	$\exp(-\alpha x)$	$\frac{\Gamma(\frac{d+1}{2})2^d\pi^{\frac{d-1}{2}}}{\alpha^d}(1+\frac{4\pi^2\omega^2}{\alpha^2})^{-\frac{d+1}{2}}$	$\sum_{n=0}^{\infty}\frac{(-\alpha)^n\sqrt{\pi}\Gamma(\frac{n+d}{2})}{n!\Gamma(\frac{d}{2})\Gamma(\frac{n+1}{2})}x^n$	$\frac{\Gamma(\frac{d+1}{2})2^d\pi^{\frac{d-1}{2}}\lvert\omega\rvert^{d-1}}{\Gamma(\frac{d}{2})\alpha^d}(1+\frac{4\pi^2\omega^2}{\alpha^2})^{-\frac{d+1}{2}}$
Sliced Laplace	$\frac{I_{\frac{d-2}{2}}(-\alpha t)+\mathbf{L}_{\frac{d-2}{2}}(-\alpha t)}{(\alpha t)^{\frac{d-2}{2}}}$	$\frac{2^{\frac{4-d}{2}}\Gamma(\frac{d}{2})}{\pi^{(d+1)/2}\Gamma(\frac{d-1}{2})\lvert\omega\rvert^{d-1}}\frac{2\alpha}{\alpha^2+4\pi^2\omega^2}$	$\frac{2^{\frac{4-d}{2}}}{\sqrt{\pi}\Gamma(\frac{d-1}{2})}\exp(-\alpha x)$	$\frac{2^{\frac{d-4}{2}}}{\sqrt{\pi}\Gamma(\frac{d-1}{2})}\frac{2\alpha}{\alpha^2+4\pi^2\omega^2}$
Matérn	$\frac{2^{1-\nu}}{\Gamma(\nu)}(\frac{\sqrt{2\nu}}{\beta}x)^\nu K_\nu(\frac{\sqrt{2\nu}}{\beta}x)$	$\frac{\Gamma(\frac{2\nu+d}{2})2^{\frac{d}{2}}\pi^{\frac{d}{2}}\beta^d}{\Gamma(\nu)\nu^{\frac{d}{2}}}(1+\frac{2\pi^2\beta^2\omega^2}{\nu})^{-\frac{2\nu+d}{2}}$	[Hertrich 2024, Appx C]	$\frac{\Gamma(\frac{2\nu+d}{2})2^{\frac{d}{2}}\pi^d\beta^d\lvert\omega\rvert^{d-1}}{\Gamma(\frac{d}{2})\Gamma(\nu)\nu^{\frac{d}{2}}}(1+\frac{2\pi^2\beta^2\omega^2}{\nu})^{-\frac{2\nu+d}{2}}$
Riesz for $r>-1$	$-x^r$	not a function	$-\frac{\sqrt{\pi}\Gamma(\frac{d+r}{2})}{\Gamma(\frac{d}{2})\Gamma(\frac{r+1}{2})}x^r$	not a function
Thin Plate Spline	$x^2\log(x)$	not a function	$dx^2\log(x)+C_dx^2$, with $C_d=\frac{d}{2}\left(H_{\frac{d}{2}}-2+\log(4)\right)$	not a function

Table: Basis functions F for different kernels $K(x,y)=F(\|x-y\|)$ and corresponding basis functions f from $\mathrm{k}(x,y)=f(\lvert x-y\rvert)$. We added the inverse Fourier transforms $\mathcal{F}_d^{-1}[F(\|\cdot\|)]$ and $\mathcal{F}_1^{-1}[f(\lvert\cdot\rvert)]$ to the table.

Slicing Error for Random Directions

Theorem [Hertrich Jahn Q. 2025]

The mean squared error is

$$\mathbb{E}_{\xi_1,\dots,\xi_P \sim \mathcal{U}_{\mathbb{S}^{d-1}}}\left[\left(\frac{1}{P}\sum_{p=1}^{P} f(\langle x, \xi_p\rangle - F(\|x\|)\right)^2\right] = \frac{\mathbb{V}_d[f](x)}{P}$$

with

Kernel	$\mathbb{V}_d[f](x)$
Any positive definite	$\leq F(0)^2 - F(\|x\|)^2$
Riesz $F(\|x\|) = -\|x\|^r$, $r > 0$	$\leq \left(\frac{\sqrt{\pi}\Gamma(r+\frac{1}{2})}{\Gamma(\frac{r+1}{2})^2} - 1\right) F(\|x\|)^2.$

- ☺ This gives the exact rate of the expected error
- 😐 Error rate $1/\sqrt{P}$ for $P \to \infty$

Quasi Monte Carlo (QMC) on the Sphere

Goal: Improve Error Rate $\mathcal{O}(1/\sqrt{P})$

Idea: Replace random points ξ_p by a quadrature rule on $\mathbb{S}^{d-1}$

Definition
[Brauchart Saff Sloan Womersley 2014]

A **QMC Design** for the Sobolev space $H^s(\mathbb{S}^{d-1})$ is a sequence $(\boldsymbol{\xi}^P)_P$ with $\boldsymbol{\xi}^P = (\xi_1^P, \dots, \xi_P^P) \in \mathbb{S}^{d-1}$ s.t.

$$\sup_{\substack{g \in H^s(\mathbb{S}^{d-1}) \\ \|g\|_{H^s(\mathbb{S}^{d-1})} \le 1}} \left| \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} g(\xi)\,\mathrm{d}\xi - \frac{1}{P}\sum_{p=1}^{P} g(\xi_p^P) \right| \le \frac{c(s,d)}{P^{s/(d-1)}}.$$

Application to slicing

1. Verify smoothness of the integrand
2. Find and compute QMC design

Quasi Monte Carlo (QMC) on the Sphere

Goal: Improve Error Rate $\mathcal{O}(1/\sqrt{P})$

Idea: Replace random points ξ_p by a quadrature rule on $\mathbb{S}^{d-1}$

Definition

[Brauchart Saff Sloan Womersley 2014]

A **QMC Design** for the Sobolev space $H^s(\mathbb{S}^{d-1})$ is a sequence $(\boldsymbol{\xi}^P)_P$ with $\boldsymbol{\xi}^P = (\xi_1^P, ..., \xi_P^P) \in \mathbb{S}^{d-1}$ s.t.

$$\sup_{\substack{g \in H^s(\mathbb{S}^{d-1}) \\ \|g\|_{H^s(\mathbb{S}^{d-1})} \leq 1}} \left| \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} g(\xi)\,\mathrm{d}\xi - \frac{1}{P} \sum_{p=1}^{P} g(\xi_p^P) \right| \leq \frac{c(s,d)}{P^{s/(d-1)}}.$$

Application to slicing

1. Verify smoothness of the integrand
2. Find and compute QMC design

Quasi Monte Carlo (QMC) on the Sphere

Technische Universität Berlin

Goal: Improve Error Rate $\mathcal{O}(1/\sqrt{P})$

Idea: Replace random points ξ_p by a quadrature rule on $\mathbb{S}^{d-1}$

Definition

[Brauchart Saff Sloan Womersley 2014]

A **QMC Design** for the Sobolev space $H^s(\mathbb{S}^{d-1})$ is a sequence $(\boldsymbol{\xi}^P)_P$ with $\boldsymbol{\xi}^P = (\xi_1^P, \dots, \xi_P^P) \in \mathbb{S}^{d-1}$ s.t.

$$\sup_{\substack{g \in H^s(\mathbb{S}^{d-1}) \\ \|g\|_{H^s(\mathbb{S}^{d-1})} \leq 1}} \left| \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} g(\xi)\, \mathrm{d}\xi - \frac{1}{P} \sum_{p=1}^{P} g(\xi_p^P) \right| \leq \frac{c(s,d)}{P^{s/(d-1)}}.$$

Application to slicing

1. Verify smoothness of the integrand
2. Find and compute QMC design

Smoothness of the Integrand

Recall: $F(\|x\|) = \mathbb{E}_{\xi \sim \mathcal{U}_{\mathbb{S}^{d-1}}}[f(\langle \xi, x \rangle)]$
Need to ensure smoothness of

$$g_x : \mathbb{S}^{d-1} \to \mathbb{R}, \quad g_x(\xi) = f(\langle \xi, x \rangle)$$

Question: Is $g_x \in H^s(\mathbb{S}^{d-1})$?

Theorem [Hertrich Jahn Q. 2025]

We have $g_x \in H^s(\mathbb{S}^{d-1})$ for

Kernel	s
Gaussian	for all $s \in \mathbb{R}_{\geq 0}$
Matérn (smoothness ν)	for $s < 2\nu + \frac{1}{2}$
Laplacian	for $s < \frac{3}{2}$
neg. distance (energy / Riesz)	for $s < \frac{3}{2}$

Smoothness of the Integrand

Recall: $F(\|x\|) = \mathbb{E}_{\xi \sim \mathcal{U}_{\mathbb{S}^{d-1}}}[f(\langle \xi, x \rangle)]$
Need to ensure smoothness of

$$g_x : \mathbb{S}^{d-1} \to \mathbb{R}, \quad g_x(\xi) = f(\langle \xi, x \rangle)$$

Question: Is $g_x \in H^s(\mathbb{S}^{d-1})$?

Theorem [Hertrich Jahn Q. 2025]

We have $g_x \in H^s(\mathbb{S}^{d-1})$ for

Kernel	s
Gaussian	for all $s \in \mathbb{R}_{\geq 0}$
Matérn (smoothness ν)	for $s < 2\nu + \frac{1}{2}$
Laplacian	for $s < \frac{3}{2}$
neg. distance (energy / Riesz)	for $s < \frac{3}{2}$

(Computable) QMC Designs

Spherical t-designs: $\boldsymbol{\xi}^P = (\xi_1^P, \dots, \xi_P^P) \in \mathbb{S}^d$ with

$$\frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f(\xi) \, \mathrm{d}\xi = \frac{1}{P} \sum_{p=1}^{P} f(\xi_p^P) \quad \forall \text{ polynomials } f \text{ of degree } \leq t$$

- ☺ Always exists and is a QMC design for $s = t$ whenever $s > \frac{d-1}{2}$
- ☺ Numerically known for $d = 3$ [Gräf Potts 2011] and $d = 4$ [Womersley 2018]
- ☹ Hard to compute for larger d

Distance points: Maximize

$$\mathcal{E}(\boldsymbol{\xi}^P) = \sum_{p,q=1}^{P} \|\xi_p^P - \xi_q^P\|$$

- ☺ Is a QMC design for $s = \frac{d+1}{2}$ [Brauchart Saff Sloan Womersley 2014]
- ☺ resulting worst case error rate $\mathcal{O}(1/P^{\frac{d}{2(d-1)}})$
  For $d = 3$: error rate $\mathcal{O}(1/P^{3/4})$

(Computable) QMC Designs

Spherical t-designs: $\boldsymbol{\xi}^P = (\xi_1^P, \ldots, \xi_P^P) \in \mathbb{S}^d$ with

$$\frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f(\xi)\,\mathrm{d}\xi = \frac{1}{P} \sum_{p=1}^{P} f(\xi_p^P) \quad \forall \text{ polynomials } f \text{ of degree } \leq t$$

- ☺ Always exists and is a QMC design for $s = t$ whenever $s > \frac{d-1}{2}$
- ☺ Numerically known for $d = 3$ [Gräf Potts 2011] and $d = 4$ [Womersley 2018]
- ☹ Hard to compute for larger d

Distance points: Maximize

$$\mathcal{E}(\boldsymbol{\xi}^P) = \sum_{p,q=1}^{P} \|\xi_p^P - \xi_q^P\|$$

- ☺ Is a QMC design for $s = \frac{d+1}{2}$ [Brauchart Saff Sloan Womersley 2014]
- ☺ resulting worst case error rate $\mathcal{O}(1/P^{\frac{d}{2(d-1)}})$
  For $d = 3$: error rate $\mathcal{O}(1/P^{3/4})$

Numerics: Evaluation of the Slicing Error

Mean error

$$\left| F(\|x\|) - \frac{1}{P} \sum_{p=1}^{P} f(\langle x, \xi_p \rangle) \right|$$

Slicing Directions ξ_p

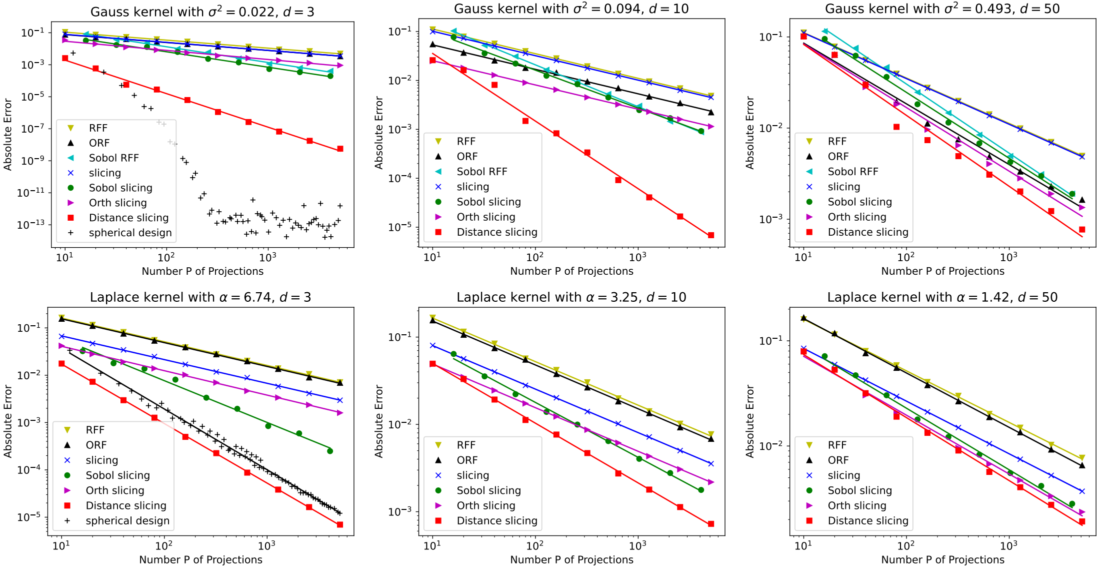
- Random directions
- Orthogonal directions: Randomly choose directions such that $\xi_1, \ldots, \xi_d$ are orthogonal
- Distance QMC Design

Compared Methods:

- Random Fourier features (RFF)
- Orthogonal Random Features (ORF) [Yu Suresh Choromanski Holtmann-Rice Kumar 2016]
- QMC-Random Fourier Features (Sobol RFF) for Gauss kernel [Avron Sindhwani Yang Mahoney 2016]

Slicing Error

Technische Universität Berlin



Technische Universität Berlin

Gauss kernel with median rule and scaling $\gamma = 1$

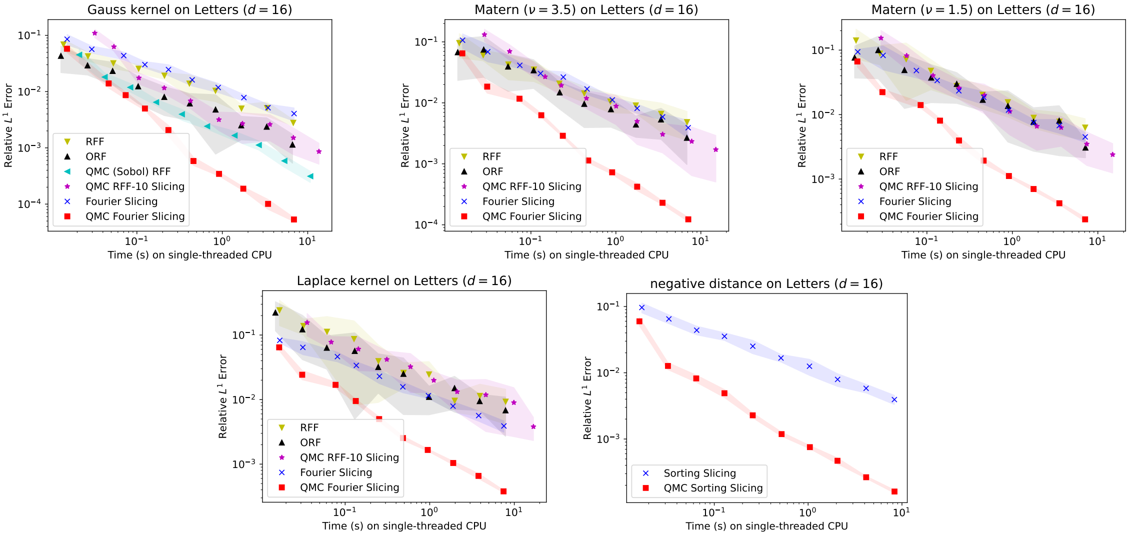
	RFF-based			Slicing-based			
Dimension	RFF	Sobol	ORF	Slicing	Sobol	Orth	Distance
$d = 3$	0.50	0.98	0.50	0.50	0.96	0.57	**2.10**
$d = 10$	0.50	0.86	0.50	0.50	0.78	0.50	**1.38**
$d = 50$	0.50	0.76	0.67	0.50	0.72	0.70	**0.78**

Laplace kernel with median rule and scaling $\gamma = 1$

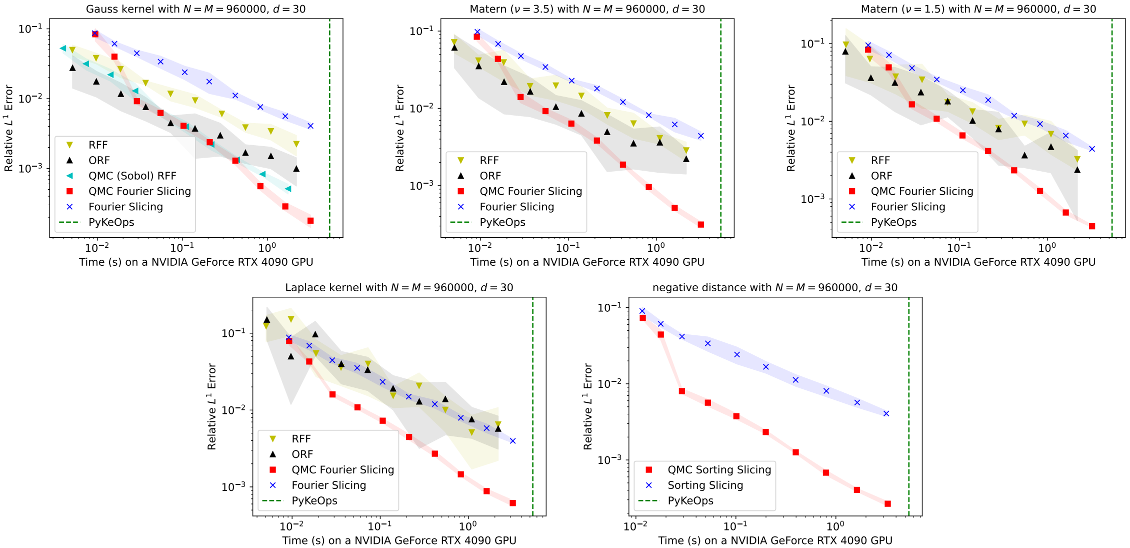
	RFF-based		Slicing-based				
Dimension	RFF	ORF	Slicing	Sobol	Orth	Distance	spherical design
$d = 3$	0.50	0.50	0.50	0.88	0.52	1.26	**1.28**
$d = 10$	0.50	0.50	0.50	0.63	0.50	**0.68**	-
$d = 50$	0.49	0.52	0.50	0.59	0.56	**0.60**	-

Table: Estimated convergence rates for the different methods. We estimate the rate r by fitting a regression line P^{-r} in the loglog plot. Larger values of r correspond to a faster convergence. The kernel parameters are determined via the median rule with scaling factor 1.

Slicing Summation on CPU (Letters dataset with $d = 16$ and $M = N = 20000$ points)

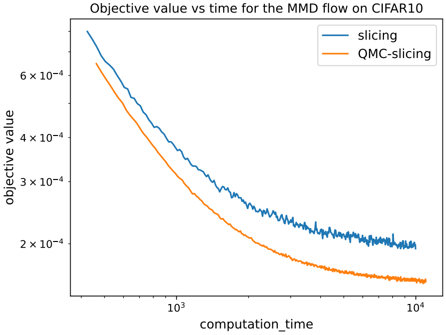


GPU Times


Gauss kernel with $N = M = 960000$, $d = 30$
Matern ($\nu = 3.5$) with $N = M = 960000$, $d = 30$
Matern ($\nu = 1.5$) with $N = M = 960000$, $d = 30$
Laplace kernel with $N = M = 960000$, $d = 30$
negative distance with $N = M = 960000$, $d = 30$
Relative L^1 Error
Time (s) on a NVIDIA GeForce RTX 4090 GPU
RFF
ORF
QMC (Sobol) RFF
QMC Fourier Slicing
Fourier Slicing
QMC Sorting Slicing
Sorting Slicing
PyKeOps
10^{-1}
10^{-2}
10^{-3}
10^{0}

Application: MMD (Maximum Mean Discrepancy) as Loss Function

- Minimize $\mathrm{Loss}(\mathbf{x}) = \mathrm{MMD}(\mathbf{x}, \mathbf{y})$
  via gradient descent
- $\mathbf{y}$ is CIFAR10 dataset ($M = 50000$ and $d = 3072$)
- use Riesz kernel $F(x) = -\|x\|$
  $\Rightarrow$ RFFs are not applicable
- Direct computation is extremely slow
  (one iteration takes $\approx 1\,\mathrm{h}$ vs. $0.3\,\mathrm{s}$ with slicing)



Conclusions

- Slicing is a powerful tool for fast kernel summation
- Applicable for many kernels, including the Riesz kernel
- Proven error rate $\mathcal{O}(1/\sqrt{P})$ for iid directions on the sphere
- Better convergence rates via QMC designs (but they require some precomputation)
- Numerical advantage QMC-slicing, especially for smooth kernels and dimensions $d \leq 100$

Future

- Improve our theoretical analysis on QMC slicing in order to match the convergence rate from the numerical section
- Study worst case errors for symmetric functions on the sphere, since g_x is always symmetric
- Use "up-slicing" to smoothen the Riesz kernel

Conclusions

- Slicing is a powerful tool for fast kernel summation
- Applicable for many kernels, including the Riesz kernel
- Proven error rate $\mathcal{O}(1/\sqrt{P})$ for iid directions on the sphere
- Better convergence rates via QMC designs (but they require some precomputation)
- Numerical advantage QMC-slicing, especially for smooth kernels and dimensions $d \leq 100$

Future

- Improve our theoretical analysis on QMC slicing in order to match the convergence rate from the numerical section
- Study worst case errors for symmetric functions on the sphere, since g_x is always symmetric
- Use “up-slicing” to smoothen the Riesz kernel

References

- J Hertrich, T Jahn, M Quellmalz. Fast Summation of Radial Kernels via QMC Slicing. *ICLR* 2025.
- N Rux, M Quellmalz, G Steidl. Slicing of Radial Functions: a Dimension Walk in the Fourier Space. *Sampling Theory, Signal Processing, and Data Analysis* 23(6), 2025.
- J Hertrich. Fast kernel summation in high dimensions via slicing and Fourier transforms. *SIAM Journal on Mathematics of Data Science* 6, 2024.
- P Hagemann, J Hertrich, F Altekrüger, R Beinert, J Chemseddine, G Steidl. Posterior sampling based on gradient flows of the MMD with negative distance kernel. *ICLR* 2024.

Thank you for your attention!