DFG Deutsche Forschungsgemeinschaft

SFB F68
Tomography Across the Scales

Technische Universität Berlin


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

Michael Quellmalz | TU Berlin | Chemnitz Summer School and Workshop on Applied Analysis, 22 September 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 $n = 1, ..., N$, $m = 1, ..., M$
- **Radial kernel** $K: \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$

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

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

- **We want:** Kernel sums

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

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 $\mathcal{O}(MN)$

The Problem

- 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: Kernel sums

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

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 $\mathcal{O}(MN)$

Approach 1: Fast Fourier Summation

[Kunis Potts 2006] [Potts Steidl Nieslony 2004]

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

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

- Insert to (1)

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

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

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 $\mathcal{O}(N + M + N_{\text{ft}}^d \log N_{\text{ft}})$

☺ Efficient for $d \in \{1, 2, 3\}$

😐 Becomes unfeasible for larger d

Technische Universität Berlin

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

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

- Insert to (1)

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

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

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 $\mathcal{O}(N + M + N_{\text{ft}}^d \log N_{\text{ft}})$

☺ 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)

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

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

1. Inner sum $\hat{w}_k$ is **non-equispaced fast Fourier transform (NFFT)** of w
2. Outer sum s_m is NFFT of $c_k \hat{w}_k$ $\quad\rightsquigarrow$ 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)

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

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

1. Inner sum $\hat{w}_k$ is **non-equispaced fast Fourier transform (NFFT)** of w
2. Outer sum s_m is NFFT of $c_k \hat{w}_k$ ⇝ 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 $c_k\hat{w}_k$ ⇝ 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)

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

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

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

❷ Outer sum s_m is NFFT of $c_k \hat{w}_k$ ⇝ 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 $v_1, \ldots, v_P$ 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)

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

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

☺ 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]

- 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]

- 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 3: Slicing Summation

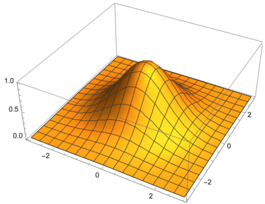
[Hertrich 2024]

Technische Universität Berlin

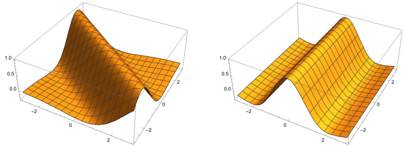
- Idea: Construct $f\colon \mathbb{R} \to \mathbb{R}$ s.t.

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

(Integral normalized to 1)



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



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

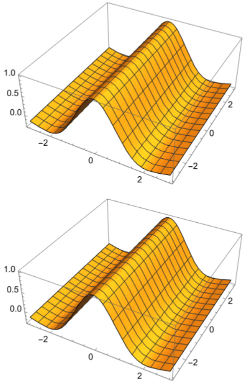
Approach 3: Slicing Summation

[Hertrich 2024]

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

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

(Integral normalized to 1)



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

Approach 3: Slicing Summation

[Hertrich 2024]

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

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

(Integral normalized to 1)

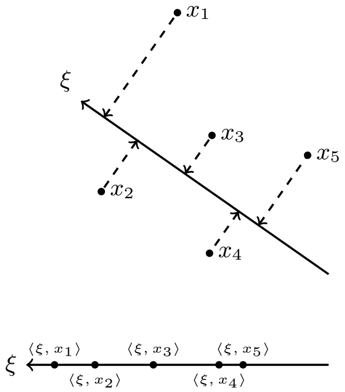
- Insert into the sum (1)

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

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

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

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



Projection of points $x_n \in \mathbb{R}^2$ to the line in direction ξ

Approach 3: Slicing Summation

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

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

(Integral normalized to 1)

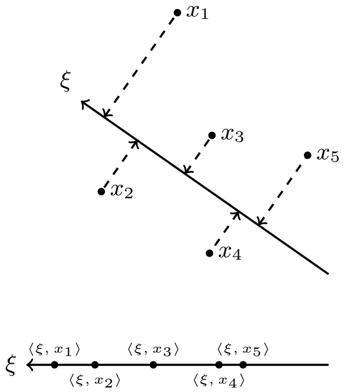
- Insert into the sum (1)

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

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

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

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



Projection of points $x_n \in \mathbb{R}^2$ to the line in direction ξ

The Slicing Transform

Technische Universität Berlin

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

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

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

Technische Universität Berlin

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

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

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 $F \mapsto f$ "Dimension Walk in Fourier Space"

Technische Universität Berlin

- **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

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

- **Multiplication operator**

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

Theorem (Inversion for L^1 functions) [Rux Q. Steidl 2025]

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

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

Theorem (Inversion via tempered distributions) [Rux Q. Steidl 2025]

Let $f \in C(\mathbb{R})$ is slowly increasing and even, it can be recovered via

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

Inversion $F \mapsto f$ "Dimension Walk in Fourier Space"

- **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

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

- **Multiplication operator**

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

Theorem (Inversion for L^1 functions) [Rux Q. Steidl 2025]

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

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

Theorem (Inversion via tempered distributions) [Rux Q. Steidl 2025]

Let $f \in C(\mathbb{R})$ is slowly increasing and even, it can be recovered via

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

Inversion $F \mapsto f$ "Dimension Walk in Fourier Space"

Technische Universität Berlin

- **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

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

- **Multiplication operator**

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

Theorem (Inversion for L^1 functions) [Rux Q. Steidl 2025]

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

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

Theorem (Inversion via tempered distributions) [Rux Q. Steidl 2025]

Let $f \in C(\mathbb{R})$ is slowly increasing and even, it can be recovered via

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

Inversion $F \mapsto f$ "Dimension Walk in Fourier Space"

- **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

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

- **Multiplication operator**

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

Theorem (Inversion for L^1 functions) [Rux Q. Steidl 2025]

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

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

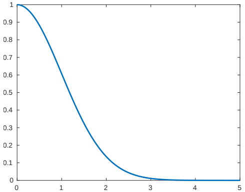
Theorem (Inversion via tempered distributions) [Rux Q. Steidl 2025]

Let $f \in C(\mathbb{R})$ is slowly increasing and even, it can be recovered via

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

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)

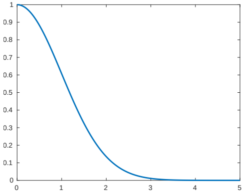


F

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)

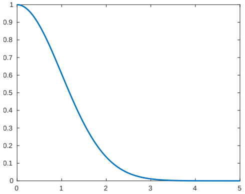


F

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	$\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)\|t\|^{d-1}$
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)

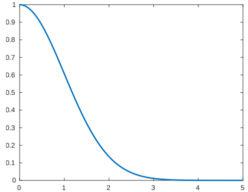


F

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)

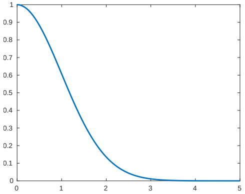


F

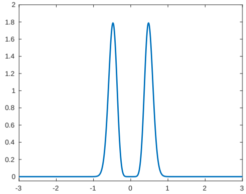
Figure: Gauss kernel F and the corresponding sliced kernel f for $d = 10$

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	$\mathcal{M}_d\mathcal{A}_d\mathcal{F}_d^{-1}\mathcal{R}_d F(t) = (2\pi)^{d/2}\exp(-2\pi^2 t^2)\|t\|^{d-1}$
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)

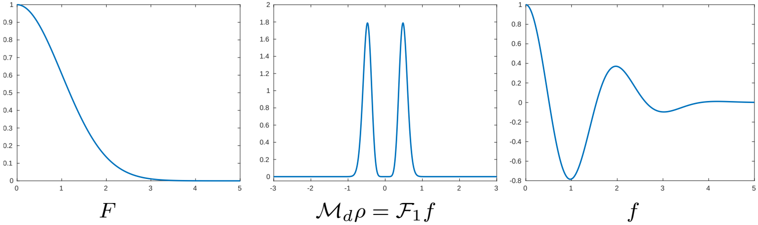


Figure: Gauss kernel F and the corresponding sliced kernel f for $d = 10$

Known Kernel 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 $C_d=\frac{d}{2}\big(H_{\frac{d}{2}}-2+\log(4)\big)$	not a function

Table: Basis functions F for different kernels $K(x,y)=F(\\|x-y\\|)$ and corresponding basis functions f.

Numerical computation of f ➔ Next talk by Nico

Slicing Error for Random Directions

Theorem [Hertrich Jahn Q. 2025]

The mean squared error is

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

with

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

- ☺ Error rate independent of dimension d
- 😐 Exact error rate $\sim 1/\sqrt{P}$ 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 $\boldsymbol{\xi}^P = (\xi_1^P, ..., \xi_P^P) \in (\mathbb{S}^{d-1})^P$ s.t.

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

Question: How smooth is our integrand $g_x : \mathbb{S}^{d-1} \to \mathbb{R}, \quad g_x(\xi) = f(\langle \xi, x \rangle)$?

Theorem [Hertrich Jahn Q. 2025]

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

Kernel	s
Gaussian	all $s \geq 0$
Matérn (smoothness ν)	$s < 2\nu + 1/2$
Laplacian	$s < 3/2$
neg. distance (energy / Riesz)	$s < 3/2$

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 $\boldsymbol{\xi}^P = (\xi_1^P, ..., \xi_P^P) \in (\mathbb{S}^{d-1})^P$ s.t.

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

Question: How smooth is our integrand $g_x : \mathbb{S}^{d-1} \to \mathbb{R}, \quad g_x(\xi) = f(\langle \xi, x \rangle)$?

Theorem [Hertrich Jahn Q. 2025]

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

Kernel	s
Gaussian	all $s \geq 0$
Matérn (smoothness ν)	$s < 2\nu + 1/2$
Laplacian	$s < 3/2$
neg. distance (energy / Riesz)	$s < 3/2$

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 $\boldsymbol{\xi}^P = (\xi_1^P, ..., \xi_P^P) \in (\mathbb{S}^{d-1})^P$ s.t.

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

Question: How smooth is our integrand $g_x : \mathbb{S}^{d-1} \to \mathbb{R}, \quad g_x(\xi) = f(\langle \xi, x \rangle)$?

Theorem [Hertrich Jahn Q. 2025]

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

Kernel	s
Gaussian	all $s \geq 0$
Matérn (smoothness ν)	$s < 2\nu + 1/2$
Laplacian	$s < 3/2$
neg. distance (energy / Riesz)	$s < 3/2$

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

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

- ☺ Always exists with $P \sim t^{d-1}$ and is a QMC design for $s > \frac{d-1}{2}$
- ☺ 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]
- ☺ Worst case error rate: $\mathcal{O}(1/P^{\frac{d}{2(d-1)}})$
  For $d = 3$: $\mathcal{O}(1/P^{3/4})$

(Computable) QMC Designs

Technische Universität Berlin

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

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

- ☺ Always exists with $P \sim t^{d-1}$ and is a QMC design for $s > \frac{d-1}{2}$
- ☺ 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]
- ☺ Worst case error rate: $\mathcal{O}(1/P^{\frac{d}{2(d-1)}})$
  For $d = 3$: $\mathcal{O}(1/P^{3/4})$

Technische Universität Berlin

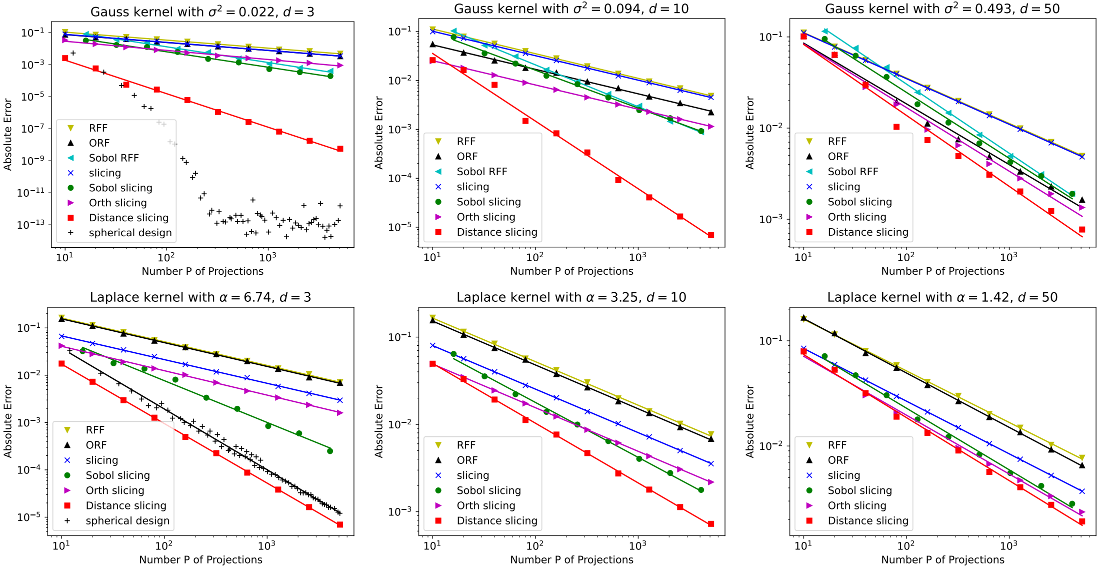
Slicing Directions ξ_p:

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



Estimated Convergence Rates

Gauss kernel

Dimension	RFF-based			Slicing-based			
	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**

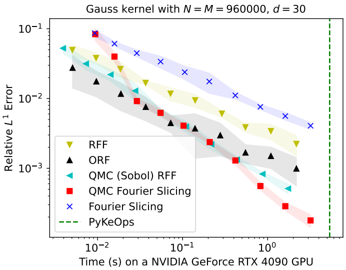
Laplace kernel

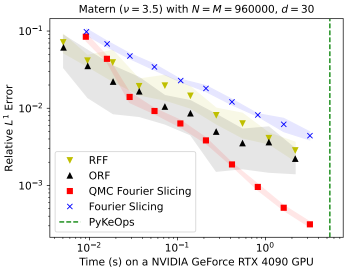
Dimension	RFF-based		Slicing-based				
	RFF	ORF	Slicing	Sobol	Orth	Distance	spherical design
$d = 3$	0.50	0.50	0.50	0.88	0.52	1.26	**1.28**
$d = 10$	0.50	0.50	0.50	0.63	0.50	**0.68**	-
$d = 50$	0.49	0.52	0.50	0.59	0.56	**0.60**	-

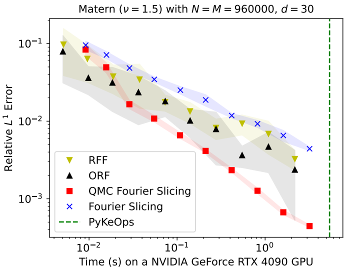
Table: Estimated convergence rates r computed by fitting a regression line P^{-r} in the loglog plot. Larger r corresponds to faster convergence. The kernel parameters are determined via the median rule with scaling factor 1.

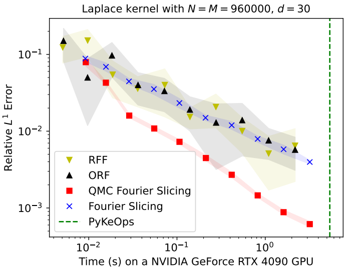
GPU Times for Slicing Summation

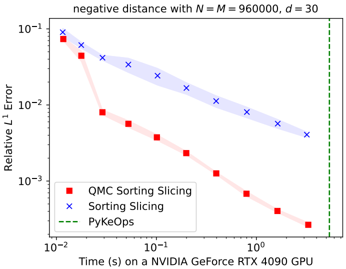
Technische Universität Berlin



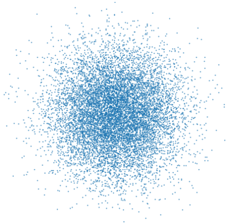




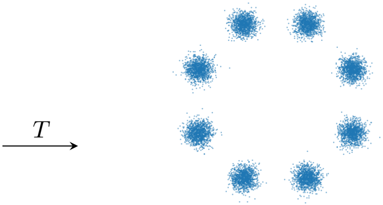




Application for MMD Flows in Generative Modeling



Latent distribution μ



Data distribution ν

- Given: Samples $y_1, \dots, y_M$ from a target probability measure ν
- Goal: Approximate ν
- Common Approach: Learn T such that $\nu \approx T_\# \mu = \mu \circ T^{-1}$ for a simple latent distribution μ

Examples: Generative Adversarial Nets (GANs), (Generalized) Normalizing Flows, etc.

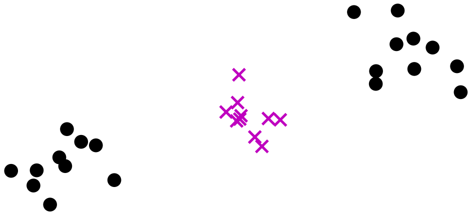
[Goodfellow et al. 2014] [Rezenede Mohamed 2015] [Hagemann Hertrich Steidl 2023]

Motivational Example

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

$$\min G_{\boldsymbol{y}}(\boldsymbol{x}) := \underbrace{-\frac{1}{2N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \|x_i - x_j\|}_{\text{repulsion}} + \underbrace{\frac{1}{MN} \sum_{i=1}^{N} \sum_{m=1}^{M} \|x_i - y_m\|}_{\text{attraction}}$$

- Gradient flow $\dot{\boldsymbol{x}}(t) = -\nabla G_{\boldsymbol{y}}(\boldsymbol{x}(t)),\ t > 0$, with $\boldsymbol{x}(0) = \boldsymbol{x}$

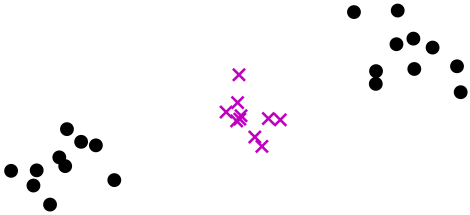


Motivational Example

- 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" $\boldsymbol{y}$

$$\min G_{\boldsymbol{y}}(\boldsymbol{x}) := \underbrace{-\frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\|x_i - x_j\|}_{\text{repulsion}} + \underbrace{\frac{1}{MN}\sum_{i=1}^{N}\sum_{m=1}^{M}\|x_i - y_m\|}_{\text{attraction}}$$

- Gradient flow $\dot{\boldsymbol{x}}(t) = -\nabla G_{\boldsymbol{y}}(\boldsymbol{x}(t)),\ t > 0$, with $\boldsymbol{x}(0) = \boldsymbol{x}$



Maximum Mean Discrepancy (MMD)

[Szekely 2002] [Gretton Borgwardt et al. 2006]

Technische Universität Berlin

- $\mathcal{P}_2(\mathbb{R}^d)$ space of probability measures μ with $\int_{\mathbb{R}^d} \|x\|^2 \,\mathrm{d}\mu(x) < \infty$
- $K\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ symmetric, positive definite kernel
- **Maximum mean discrepancy** (MMD) $\mathcal{D}_K\colon \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$

$$\mathcal{D}_K(\mu,\nu)^2 = \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K \,\mathrm{d}\mu\,\mathrm{d}\mu - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K\,\mathrm{d}\nu\,\mathrm{d}\mu + \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K\,\mathrm{d}\nu\,\mathrm{d}\nu$$

- If K is a so-called "characteristic" kernels, $\mathcal{D}_K$ is a metric on $\mathcal{P}_2(\mathbb{R}^d)$

- Fix target ν: $G(\mu) := \mathcal{D}_K(\mu,\nu)^2$
- **Wasserstein gradient flow**: "Analogue of gradient flow in $\mathcal{P}_2(\mathbb{R}^d)$ with Wasserstein distance"
  [Ambrosio Gigli Savaré 2008] [Arbel Korba Salim Gretton 2019]

- For $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$, $\nu = \frac{1}{M}\sum_{j=1}^M \delta_{y_j}$:

$$G(\mu) = \mathcal{D}_K^2(\mu,\nu) = -\frac{1}{2N^2}\sum_{i=1}^N\sum_{j=1}^N K(x_i,x_j) + \frac{1}{MN}\sum_{i=1}^N\sum_{j=1}^M K(x_i,y_j) + \mathrm{const}(\nu)$$

- **Fast computation via Slicing!**

Maximum Mean Discrepancy (MMD)

[Szekely 2002] [Gretton Borgwardt et al. 2006]

- $\mathcal{P}_2(\mathbb{R}^d)$ space of probability measures μ with $\int_{\mathbb{R}^d} \|x\|^2 \,\mathrm{d}\mu(x) < \infty$
- $K\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ symmetric, positive definite kernel
- **Maximum mean discrepancy** (MMD) $\mathcal{D}_K\colon \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$

$$\mathcal{D}_K(\mu,\nu)^2 = \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K \,\mathrm{d}\mu\,\mathrm{d}\mu - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K \,\mathrm{d}\nu\,\mathrm{d}\mu + \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K \,\mathrm{d}\nu\,\mathrm{d}\nu$$

- If K is a so-called "characteristic" kernels, $\mathcal{D}_K$ is a metric on $\mathcal{P}_2(\mathbb{R}^d)$

- Fix target ν: $G(\mu) := \mathcal{D}_K(\mu,\nu)^2$
- **Wasserstein gradient flow**: "Analogue of gradient flow in $\mathcal{P}_2(\mathbb{R}^d)$ with Wasserstein distance"
  [Ambrosio Gigli Savaré 2008] [Arbel Korba Salim Gretton 2019]

- For $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$, $\nu = \frac{1}{M}\sum_{j=1}^M \delta_{y_j}$:

$$G(\mu) = \mathcal{D}_K^2(\mu,\nu) = -\frac{1}{2N^2}\sum_{i=1}^N\sum_{j=1}^N K(x_i,x_j) + \frac{1}{MN}\sum_{i=1}^N\sum_{j=1}^M K(x_i,y_j) + \mathrm{const}(\nu)$$

- **Fast computation via Slicing!**

Maximum Mean Discrepancy (MMD)

[Szekely 2002] [Gretton Borgwardt et al. 2006]

Technische Universität Berlin

- $\mathcal{P}_2(\mathbb{R}^d)$ space of probability measures μ with $\int_{\mathbb{R}^d} \|x\|^2 \,\mathrm{d}\mu(x) < \infty$
- $K: \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ symmetric, positive definite kernel
- **Maximum mean discrepancy** (MMD) $\mathcal{D}_K: \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$

$$\mathcal{D}_K(\mu,\nu)^2 = \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K \,\mathrm{d}\mu\,\mathrm{d}\mu - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K \,\mathrm{d}\nu\,\mathrm{d}\mu + \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K \,\mathrm{d}\nu\,\mathrm{d}\nu$$

- If K is a so-called “characteristic” kernels, $\mathcal{D}_K$ is a metric on $\mathcal{P}_2(\mathbb{R}^d)$

- Fix target ν: $G(\mu) := \mathcal{D}_K(\mu,\nu)^2$
- **Wasserstein gradient flow**: “Analogue of gradient flow in $\mathcal{P}_2(\mathbb{R}^d)$ with Wasserstein distance”
  [Ambrosio Gigli Savaré 2008] [Arbel Korba Salim Gretton 2019]

- For $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$, $\nu = \frac{1}{M}\sum_{j=1}^M \delta_{y_j}$:

$$G(\mu) = \mathcal{D}_K^2(\mu,\nu) = -\frac{1}{2N^2}\sum_{i=1}^N\sum_{j=1}^N K(x_i,x_j) + \frac{1}{MN}\sum_{i=1}^N\sum_{j=1}^M K(x_i,y_j) + \mathrm{const}(\nu)$$

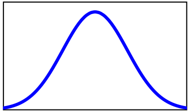
- **Fast computation via Slicing!**

What Kernels to use for MMD?

Technische Universität Berlin

Gauss

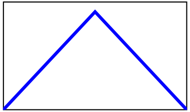
- ☺ Positive definite, characteristic kernel
- ☺ Defines a Wasserstein gradient flow
- ☹ Convergence depends heavily on a good parameter choice σ



$$F(t) = \mathrm{e}^{-\frac{t^2}{2\sigma^2}}$$

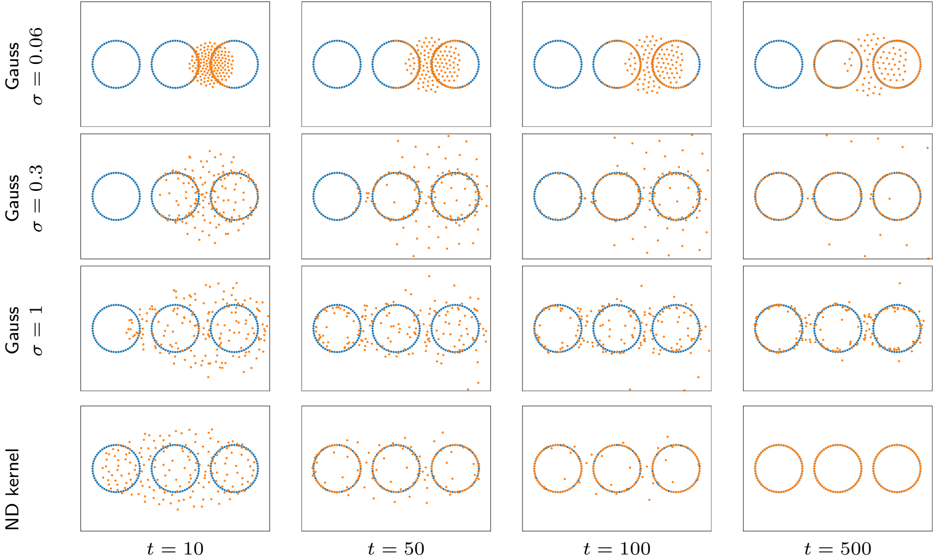
Negative Distance (ND)

- ☹ Not positive definite
- ☹ Not differentiable
- 😐 Wasserstein gradient flow of discrete measures can become continuous [Hertrich Gräf Beinert Steidl 2024]
- ☺ Often good numerical results
- ☺ Scale-independent, no parameter necessary



$$F(t) = -|t|$$

MMD Flows for Gauss and ND



Blue: target ν
Orange: Flow γ_t

Smooth Negative Distance (SND) Kernel

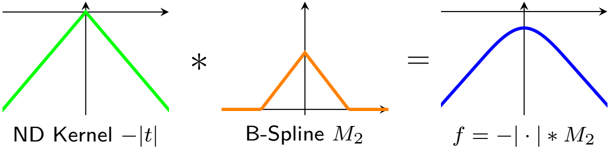
Goal: Smoothen the ND kernel $-\|x\|$ to retain the numeric behavior, and get theoretical properties

Smooth Negative Distance (SND) Kernel

Goal: Smoothen the ND kernel $-\|x\|$ to retain the numeric behavior, and get theoretical properties

Idea 1: Convolution of ND kernel function $\mathbb{R} \ni t \mapsto -|t|$ with a B–Spline M_n

☹ Not (conditionally) positive definite for $d \geq 2$

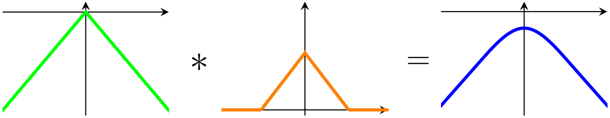


Smooth Negative Distance (SND) Kernel

Goal: Smoothen the ND kernel $-\|x\|$ to retain the numeric behavior, and get theoretical properties

Idea 1: Convolution of ND kernel function $\mathbb{R} \ni t \mapsto -|t|$ with a B–Spline M_n

☹ Not (conditionally) positive definite for $d \geq 2$



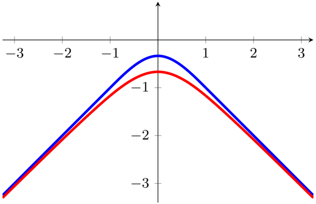
ND Kernel $-|t|$ B-Spline M_2 $f = -|\cdot| * M_2$

Idea 2 (Up-Slicing): Apply the slicing transform $\mathcal{I}_d$ to f

Theorem [Rux Q. Steidl 2025]

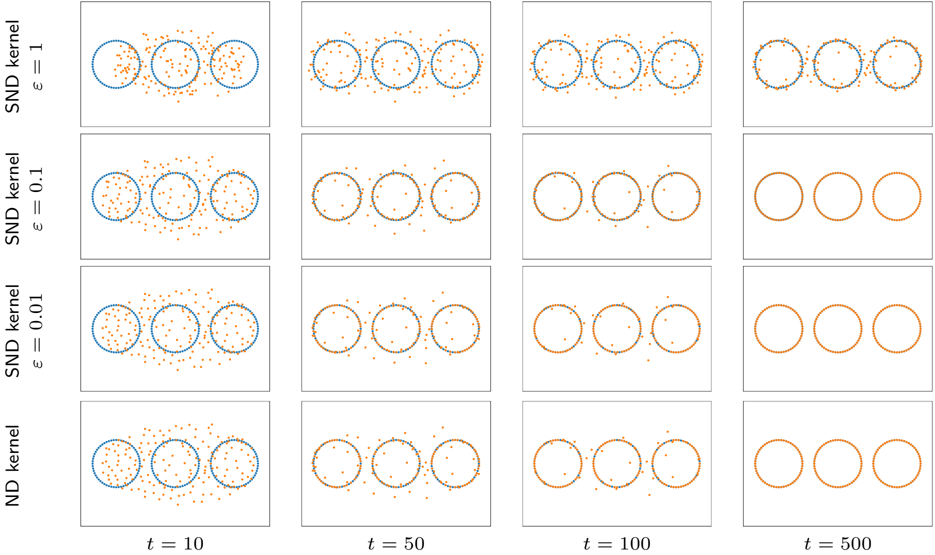
The kernel $K(x, y) = F(\|x - y\|) - F(\|x\|) - F(\|y\|)$ with $F := \mathcal{I}_d[|\cdot| * M_n]$ is positive definite on $\mathbb{R}^d$ and characteristic. There exists a Wasserstein gradient flow of the MMD functional G. It holds

$$v_t = -\int_{\mathbb{R}^d} \nabla_x F(\|x - y\|) \, \mathrm{d}(\gamma_t - \nu)(y).$$



Blue: $f = -|\cdot| * M_2$
Red: SND kernel $F = 2\mathcal{I}_3[f]$

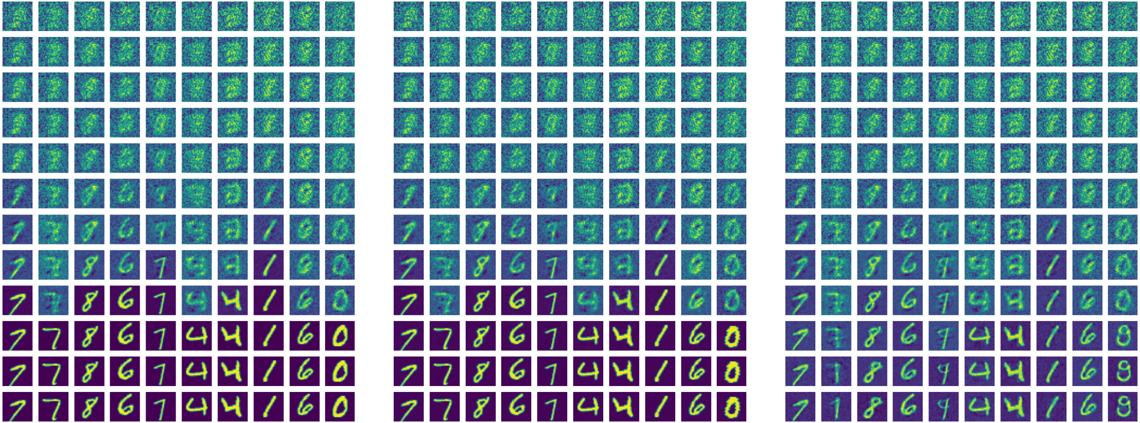
Smooth Negative Distance (SND) Kernel



MMD flow with step size 0.01. Our SND kernel with small ε is as good as the ND kernel and better than the Gaussian.

MMD Flows on MNIST

Technische Universität Berlin



(a) ND (b) SND ($\varepsilon = 0.01$). (c) SND ($\varepsilon = 0.1$).

Figure: MMD flow for MNIST target ν with different kernels. Initial distribution is uniform random. Each row shows the first 10 images $x_n \in \mathbb{R}^{28\times 28}$, the ℓ-th row corresponds to the iteration $k = 2^{3+\ell}$, $\ell = 1, \dots, 12$.

Conclusions

- Slicing is applicable for many kernels, including the Riesz kernel
- Proven error rate $\mathcal{O}(1/\sqrt{P})$ for iid directions on the sphere
- Better rates via QMC designs (but they require some precomputation)
- SND kernel retains desired numerical properties of the ND, but comes with well-defined gradient and theoretical convergence guarantees of the corresponding gradient flow

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.

N Rux, M Quellmalz, G Steidl. Smoothed Distance Kernels for MMDs and Applications in Wasserstein Gradient Flows. ArXiv prerpint 2504.07820.

Conclusions

- Slicing is applicable for many kernels, including the Riesz kernel
- Proven error rate $\mathcal{O}(1/\sqrt{P})$ for iid directions on the sphere
- Better rates via QMC designs (but they require some precomputation)
- SND kernel retains desired numerical properties of the ND, but comes with well-defined gradient and theoretical convergence guarantees of the corresponding gradient flow

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.

N Rux, M Quellmalz, G Steidl. Smoothed Distance Kernels for MMDs and Applications in Wasserstein Gradient Flows. ArXiv prerpint 2504.07820.

Thank you for your attention!