Kernel quadrature with determinantal point processes.

Ayoub Belhadji[*], Rémi Bardenet[*], Pierre Chainais[*]

[*]Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRIStAL, 59651 Villeneuve d'Ascq, France

Abstract

We study the design of nodes for kernel quadrature. We relate the approximation error for a given design to geometric quantities that translates relative positions of functional spaces. This geometric analysis allows to bound the approximation error when the nodes follow the distribution of a projection DPP tailored to the RKHS. This bound depends on the spectrum of the kernel and has been validated by simulations.

Kernel quadrature

Let $\mathcal{X}$ be a topological space equipped with a Borel measure $\mathrm{d}\omega$. Let $f, g \in \mathbb{L}_2(\mathrm{d}\omega)$, we are interested in the approximations

$$\int_{\mathcal{X}} f(x)g(x)\mathrm{d}\omega(x) \approx \sum_{j\in[N]} w_j f(x_j), \tag{1}$$

where $x_j \in \mathcal{X}$ are the nodes and the $w_j \in \mathbb{R}$ are the weights that depends only on g. We assume that f belongs to an RKHS $\mathcal{F}$represented by a kernel $k : \mathcal{X}\times\mathcal{X} \to \mathbb{R}_+$. We have

$$\left|\int_{\mathcal{X}} f(x)g(x)\mathrm{d}\omega(x) - \sum_{j\in[N]} w_j f(x_j)\right| \leq \left\|\mu_g - \sum_{j\in[N]} w_j k(x_j,.)\right\|_{\mathcal{F}} \|f\|_{\mathcal{F}},$$

with μ_g is the mean element of the distribution of density $x \mapsto g(x)\mathrm{d}\omega(x)$.

The problem of nodes design for kernel quadrature

How to construct a design of nodes $(x_j)_{j\in[N]}$, with a provable bound on the approximation error for the optimal weights

$$\|\mu_g - \Pi_{\mathcal{T}(x)}\mu_g\|_{\mathcal{F}} = \|\mu_g - \sum_{j\in[N]} \hat{w}_j k(x_j,.)\|_{\mathcal{F}}.$$

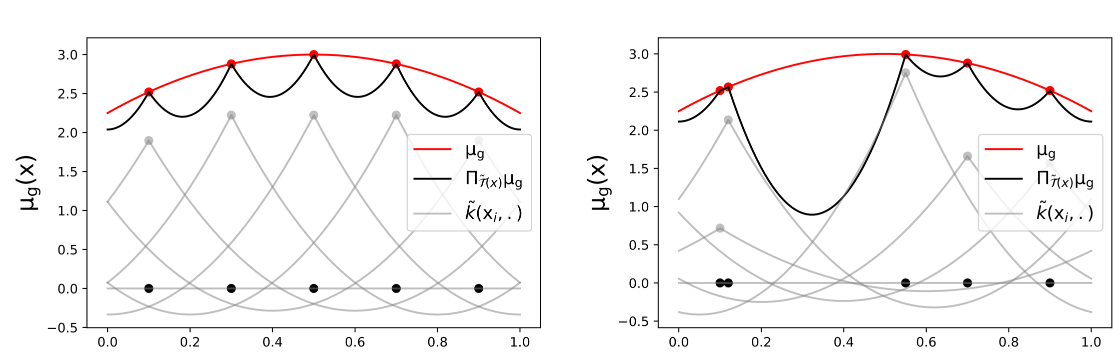


Figure 1:The influence of the design on the approximation quality.

We propose to sample the nodes according to a projection determinantal point process.

Projection Determinantal Point Process

We assume that the RKHS kernel satisfies the conditions for Mercer decompositions

$$k(x,y) = \sum_{m\in\mathbb{N}^*} \sigma_m e_m(x)e_m(y), \tag{2}$$

with the $(e_m)_{m\in\mathbb{N}^*}$ is an orthonormal basis in $\mathbb{L}_2(\mathrm{d}\omega)$ and $0 \leq \sigma_{m+1} \leq \sigma_m, \forall m \in \mathbb{N}^*$.

Definition

Define the repulsion kernel

$$\mathfrak{K}(x,y) = \sum_{n\in[N]} e_n(x)e_n(y). \tag{3}$$

The set $x = \{x_1, \dots x_N\}$ is said to be a projection DPP with reference measure $\mathrm{d}\omega$ and kernel $\mathfrak{K}$ if it follows the distribution of density

$$\frac{1}{N!}\mathrm{Det}(\mathfrak{K}(x_i,x_j)_{i,j\in[N]}) \prod_{i\in[N]} \omega(x_i). \tag{4}$$

A geometric characterization of "good" quadrature nodes

Define the saturated kernel

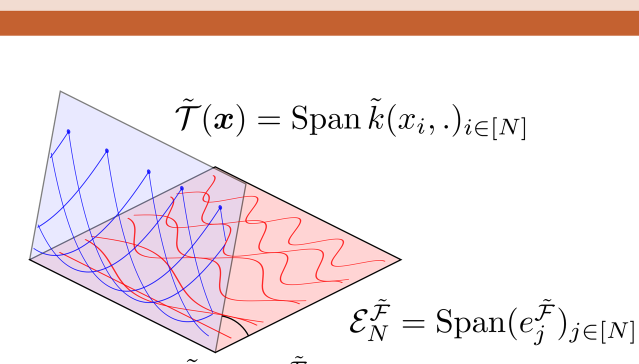
$$\tilde{k}(x,y) = \sum_{n\in[N]} \sigma_1 e_n(x)e_n(y) + \sum_{n\geq N+1} \sigma_n e_n(x)e_n(y). \tag{5}$$

Theorem

Define $\|g\|_{\mathrm{d}\omega,1} = \sum_{n\in[N]} |\langle e_n, g\rangle_{\mathrm{d}\omega}|$. Assume that $\|g\|_{\mathrm{d}\omega} \leq 1$ then

$$\|\mu_g - \Pi_{\mathcal{T}(x)}\mu_g\|^2_{\mathcal{F}} \leq 2\left(\sigma_{N+1} + \|g\|^2_{\mathrm{d}\omega,1}\tan^2\theta_N(\tilde{\mathcal{T}}(x), \mathcal{E}_N^{\tilde{\mathcal{F}}})\right). \tag{6}$$

- The "filtering" error $\|\Pi_{\mathcal{E}_N^{\mathcal{F}}}\mu_g\|^2_{\mathcal{F}}$
- The sparsity of the coefficients of g
- The angle between $\tilde{\mathcal{T}}(x)$ and $\mathcal{E}_N^{\tilde{\mathcal{F}}}$.



Main result

Theorem

Let $x = \{x_1, \dots, x_N\}$ be a projection DPP $(\mathrm{d}\omega, \mathfrak{K})$. Define $r_N = \sum_{m\geq N+1} \sigma_m$, then

$$\mathbb{E}_{\mathrm{DPP}}\left\|\mu_g - \Pi_{\mathcal{T}(x)}\mu_g\right\|^2_{\mathcal{F}} \leq 2\left(\sigma_{N+1} + \|g\|^2_{\mathrm{d}\omega,1}(Nr_N + o(Nr_N))\right). \tag{7}$$

Examples:

RKHS/$\mathrm{d}\omega$	Periodic Sobolev (s)/ Uniform	Korobov (d,s)/Uniform	Gaussian (d)/Gaussian
σ_{N+1}	$\mathcal{O}(s^{-2s})$	$\mathcal{O}((\log N)^{2s(d-1)}s^{-2s})$	$\mathcal{O}(e^{-\gamma d!^{1/d}N^{1/d}})$

The Christoffel function and the inclusion probability of projection DPP

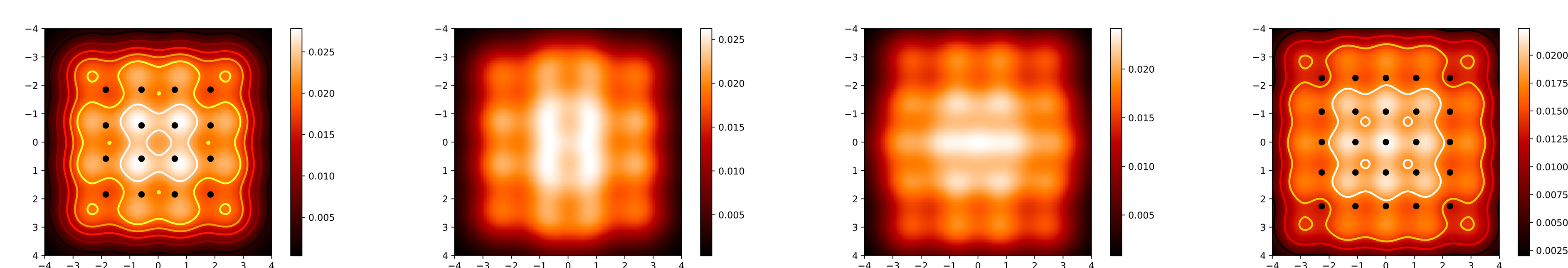


Figure 2:The inclusion probability $\mathbb{P}_{\mathrm{DPP}}(. \in x)$ in the Gaussian case $x \mapsto \frac{\mathfrak{K}(x,x)}{N}\mathrm{d}\omega(x)$ compared to the tensor product of the zeros of the Hermite polynomials.

Numerical experiments

A comparison of several kernel quadrature algorithms in three different RKHSs: periodic Sobolev spaces of order $s \in \{1, 3\}$, Korobov space in dimension $d = 2$ and for $s = 1$, Gaussian space in dimension $d = 1$. The square of the worst case error $\left\|\mu_g - \Pi_{\mathcal{T}(x)}\mu_g\right\|_{\mathcal{F}}$ is plotted with $g : x \mapsto 1$. The squared error is averaged on 50 samples for randomized algorithms.

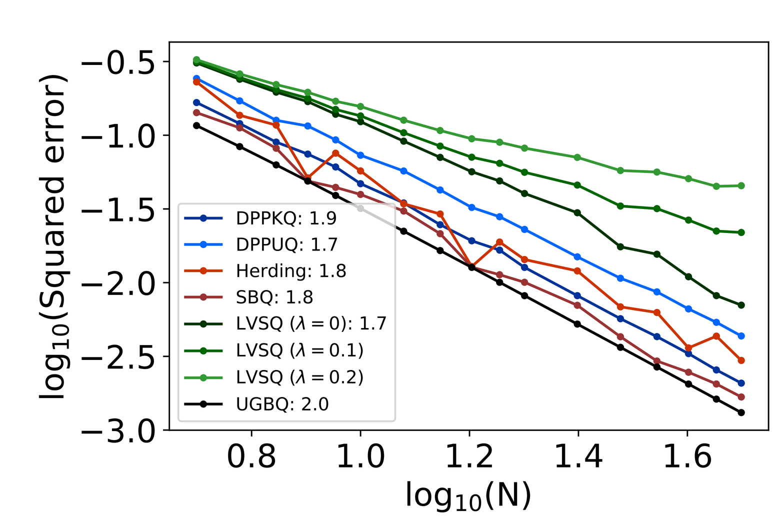


Figure 3: Sobolev space, $d = 1$, $s = 1$

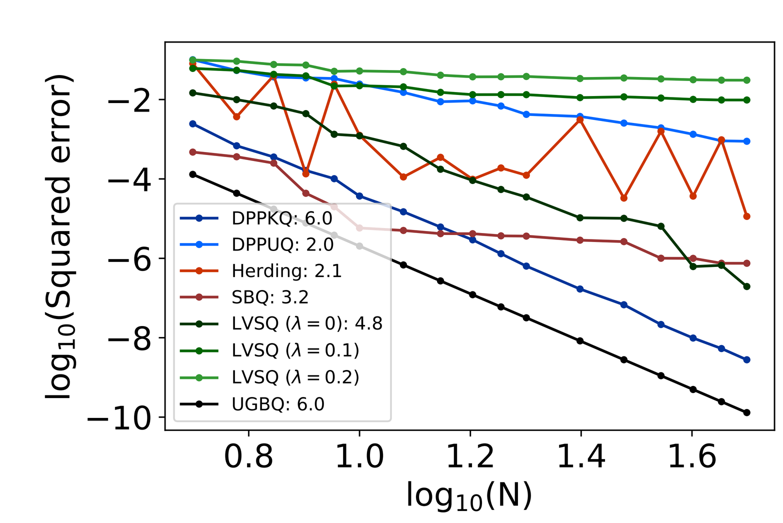


Figure 4: Sobolev space, $d = 1$, $s = 3$

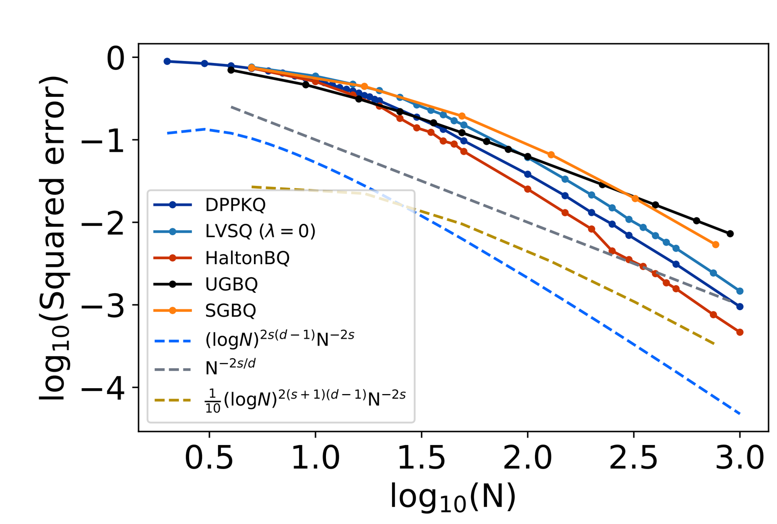


Figure 5: Korobov space, $d = 2$, $s = 1$

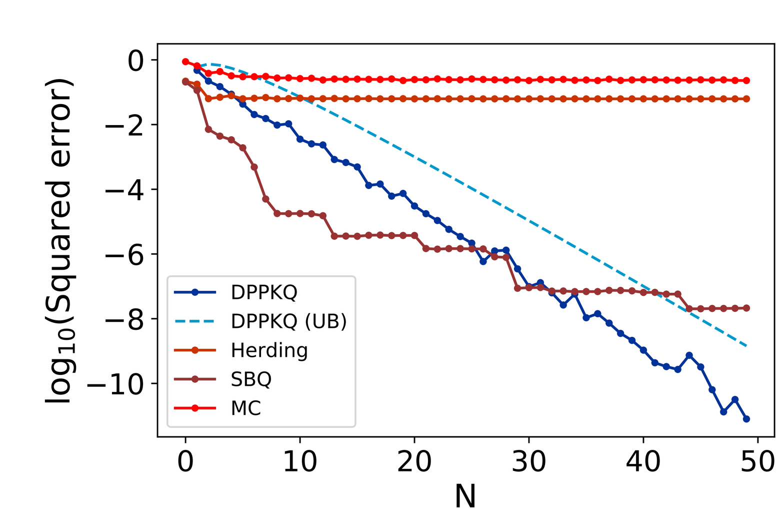


Figure 6: Gaussian space, $d = 1$

Take Home Messages

- A general analysis of kernel quadrature for nodes sampled according to a projection DPP,
- A new geometric interpretation of "good" quadrature nodes using principal angles
- Empirical validation on different RKHSs.

Perspectives

- Bridging the gap between the theoretical rates and the empirical rates for polynomially decreasing spectrum.
- Theoretical bounds under regularization.
- Lower bounds for the non regularized kernel quadrature.

We have an ERC grant coming up. If you are interested in DPPs and would be into a PhD/postdoc in France, feel free to contact Rémi Bardenet http://rbardenet.github.io.