This is the supplementary material for the paper: "Fast learning in reproducing kernel Kreĭn spaces via generalized measures", by Fanghui Liu, Xiaolin Huang, Yingyi Chen, and Johan A.K. Suykens. The supplementary material (Appendix) is organized as follows.

- Section A gives the proof of Theorem 3 on the studied open question.
- The approximation performance is theoretically demonstrated in Section B in terms of uniform approximation error bound and variance reduction.
- In Section (C), we demonstrate that polynomial kernels on the unit sphere by ℓ_2 normalization data is with finite total mass.
- Section D gives the proof of Theorem 4.
- The measure of arc-cosine kernels on the unit sphere by ℓ_2 normalization data is given in Section E.

A Proof of Theorem 3

Proof. We give the proof of the existence.
(i) Necessity.
An stationary indefinite kernel associated with RKKS admits the positive decomposition

$$k(\boldsymbol{x}-\boldsymbol{x}') = k_+(\boldsymbol{x}-\boldsymbol{x}') - k_-(\boldsymbol{x}-\boldsymbol{x}')\,, \quad \forall \boldsymbol{x}, \boldsymbol{x}' \in X\,,$$

where k_+ and k_- are two positive definite kernels. According to the Bochner's theorem [21], there exists two probability measures μ_+, μ_- such that

$$k(\boldsymbol{z}) = k_+(\boldsymbol{z}) - k_-(\boldsymbol{z}) = \int_\Omega \exp\left(\mathrm{i}\boldsymbol{\omega}^\top \boldsymbol{z}\right) \mu_+(\mathrm{d}\boldsymbol{\omega}) - \int_\Omega \exp\left(\mathrm{i}\boldsymbol{\omega}^\top \boldsymbol{z}\right) \mu_-(\mathrm{d}\boldsymbol{\omega})\,,$$

where $\boldsymbol{z} := \boldsymbol{x}-\boldsymbol{x}'$. Denote $\mu := \mu_+ - \mu_-$, it is clear that μ is a signed measure, and its total mass is finite because of $\|\mu\| = \|\mu_+\| + \|\mu_-\| = 2$.

(ii) Sufficiency.
Let $\Omega := \mathbb{R}^d$ and $\mathcal{A}$ be the smallest σ-algebra containing all open subsets of Ω, and $\mu : \mathcal{A} \to [-\infty, \infty]$

$$\mu(\boldsymbol{\omega}) = \int_\Omega \exp\left(-\mathrm{i}\boldsymbol{\omega}^\top \boldsymbol{z}\right) k(\boldsymbol{z}) \mathrm{d}\boldsymbol{z}\,.$$

Since we assume that μ has total mass $\|\mu\| < \infty$, i.e., μ is finite, μ can be regarded as a signed measure. By virtue of Jordan decomposition in Theorem 2, there exist two nonnegative finite measures μ_+ and μ_- such that $\mu = \mu_+ - \mu_-$. One intuitive implementation way is choosing $\mu_+ = \max\{\mu, 0\}$ and $\mu_- = \min\{0, \mu\}$. Then using the inverse Fourier transform and Plancherel's theorem [52], we have

$$\begin{aligned} k(\boldsymbol{z}) = \int_\Omega \exp\left(\mathrm{i}\boldsymbol{\omega}^\top \boldsymbol{z}\right) \mu(\mathrm{d}\boldsymbol{\omega}) &= \int_\Omega \exp\left(\mathrm{i}\boldsymbol{\omega}^\top \boldsymbol{z}\right) \mu_+(\mathrm{d}\boldsymbol{\omega}) - \int_\Omega \exp\left(\mathrm{i}\boldsymbol{\omega}^\top \boldsymbol{z}\right) \mu_-(\mathrm{d}\boldsymbol{\omega}) \\ &= \|\mu_+\| \int_\Omega \exp\left(\mathrm{i}\boldsymbol{\omega}^\top \boldsymbol{z}\right) \tilde{\mu}_+(\mathrm{d}\boldsymbol{\omega}) - \|\mu_-\| \int_\Omega \exp\left(\mathrm{i}\boldsymbol{\omega}^\top \boldsymbol{z}\right) \tilde{\mu}_-(\mathrm{d}\boldsymbol{\omega}) \\ &= \|\mu_+\| \tilde{k}_+(\boldsymbol{z}) - \|\mu_-\| \tilde{k}_-(\boldsymbol{z})\,, \end{aligned}$$

where $\tilde{\mu}_+ := \mu_+/\|\mu_+\|$ and $\tilde{\mu}_- := \mu_-/\|\mu_-\|$ are two nonnegative Borel measures, which correspond to two positive definite kernels $\tilde{k}_+$ and $\tilde{k}_-$, respectively. By defining $k_+ := \|\mu_+\|\tilde{k}_+$ and $k_- := \|\mu_+\|\tilde{k}_-$, we have

$$k(\boldsymbol{x}, \boldsymbol{x}') = k_+(\boldsymbol{x}, \boldsymbol{x}') - k_-(\boldsymbol{x}, \boldsymbol{x}'), \quad \forall \boldsymbol{x}, \boldsymbol{x}' \in X\,.$$

This completes the proof.

Based on the above analysis, we give a characterization of the RKHSs $\mathcal{H}_\pm$ through the given spectral density $\mu_\pm$. In [63], a RKHS can be characterized by its measure via Fourier transform. Therefore, in our model, the RKHSs $\mathcal{H}_\pm$ are represented by $\mu_\pm$. That is, for any $f \in \mathcal{H}_\pm$, the inner product is induced by the Hilbert norm

$$\|f\|^2_{\mathcal{H}_\pm} = \int_{\mathbb{R}^d} \frac{|F(\boldsymbol{\omega})|^2}{\mu_+(\boldsymbol{\omega})} \mathrm{d}\boldsymbol{\omega}\,,$$

where $F(\boldsymbol{\omega}) = \mathscr{F}(f) = \int_{\mathbb{R}^d} f(\boldsymbol{x}) e^{-2\pi \mathrm{i} \boldsymbol{\omega}^\top \boldsymbol{x}} \mathrm{d}\boldsymbol{x}$ is the Fourier transform of f.

□

B Proof of Proposition 2

The proof can be easily derived from [56, 57], and we briefly present here for completeness.

Proof. Proposition 1 in [56] demonstrates

$$\Pr\left[\sup_{\boldsymbol{x},\boldsymbol{x}'\in\mathcal{S}_R} |k_\pm(\boldsymbol{x},\boldsymbol{x}') - \tilde{k}_\pm(\boldsymbol{x},\boldsymbol{x}')| \geq \epsilon\right] \leq 66\left(\frac{\sigma_\pm R}{\epsilon}\right)^2 \exp\left(-\frac{s\epsilon^2}{8(d+2)}\right),$$

where $\sigma_\pm^2 = \mathbb{E}_{\boldsymbol{\omega}\sim\tilde{\mu}_\pm}[\boldsymbol{\omega}^\top\boldsymbol{\omega}] < \infty$. Since the indefinite kernel k admits

$$|k(\boldsymbol{x},\boldsymbol{x}') - \tilde{k}(\boldsymbol{x},\boldsymbol{x}')| \leq |k_+(\boldsymbol{x},\boldsymbol{x}') - \tilde{k}_+(\boldsymbol{x},\boldsymbol{x}')| + |k_-(\boldsymbol{x},\boldsymbol{x}') - \tilde{k}_-(\boldsymbol{x},\boldsymbol{x}')|, \quad \forall \boldsymbol{x},\boldsymbol{x}' \in \mathcal{S}_R,$$

then we have

$$\begin{aligned}
\Pr\left[\sup_{\boldsymbol{x},\boldsymbol{x}'\in\mathcal{S}_R} |k(\boldsymbol{x},\boldsymbol{x}') - \tilde{k}(\boldsymbol{x},\boldsymbol{x}')| \geq \epsilon\right] &\leq \Pr\left[\sup_{\boldsymbol{x},\boldsymbol{x}'\in\mathcal{S}_R} |k_+(\boldsymbol{x},\boldsymbol{x}') - \tilde{k}_+(\boldsymbol{x},\boldsymbol{x}')| \geq \frac{\epsilon}{2}\right] + \Pr\left[\sup_{\boldsymbol{x},\boldsymbol{x}'\in\mathcal{S}_R} |k_-(\boldsymbol{x},\boldsymbol{x}') - \tilde{k}_-(\boldsymbol{x},\boldsymbol{x}')| \geq \frac{\epsilon}{2}\right] \\
&\leq 66\left(\frac{2\sigma_+ R}{\epsilon}\right)^2 \exp\left(-\frac{s\epsilon^2}{32(d+2)}\right) + 66\left(\frac{2\sigma_- R}{\epsilon}\right)^2 \exp\left(-\frac{s\epsilon^2}{32(d+2)}\right) \\
&= 66\left(\frac{2\sigma R}{\epsilon}\right)^2 \exp\left(-\frac{s\epsilon^2}{32(d+2)}\right).
\end{aligned}$$

Then we study the variance reduction of the applied orthogonal Monte Carlo (OMC) sampling. Based on the definition of MSE, i.e., $\mathbb{E}[\tilde{k}(\boldsymbol{x},\boldsymbol{x}')] = \mathbb{E}[k(\boldsymbol{x},\boldsymbol{x}') - \tilde{k}(\boldsymbol{x},\boldsymbol{x}')]$, we conclude that $\mathbb{E}[\tilde{k}(\boldsymbol{x},\boldsymbol{x}')]$ is the variance of $\tilde{k}(\boldsymbol{x},\boldsymbol{x}')$, termed as $\mathrm{Var}[\tilde{k}(\boldsymbol{x},\boldsymbol{x}')]$ due to our unbiased estimator. According to Theorem 4.2 in [57], for sufficiently large d, we have

$$\mathrm{MSE}[\tilde{k}_+^{\mathrm{OMC}}(\boldsymbol{x},\boldsymbol{x}')] \leq \mathrm{MSE}[\tilde{k}_+^{\mathrm{MC}}(\boldsymbol{x},\boldsymbol{x}')] \quad \text{and} \quad \mathrm{MSE}[\tilde{k}_-^{\mathrm{OMC}}(\boldsymbol{x},\boldsymbol{x}')] \leq \mathrm{MSE}[\tilde{k}_-^{\mathrm{MC}}(\boldsymbol{x},\boldsymbol{x}')],$$

where $\boldsymbol{\omega} \sim \tilde{\mu}_+(\cdot)$ in $\tilde{k}_+$ and $\boldsymbol{\nu} \sim \tilde{\mu}_-(\cdot)$ in $\tilde{k}_-$ as indicated by Eq. (1). Since these two random vectors $\boldsymbol{\omega}$ and $\boldsymbol{\nu}$ are independent, we have

$$\mathrm{Var}[\tilde{k}^{\mathrm{OMC}}(\boldsymbol{x},\boldsymbol{x}')] = \mathrm{Var}[\tilde{k}_+^{\mathrm{OMC}}(\boldsymbol{x},\boldsymbol{x}')] + \mathrm{Var}[\tilde{k}_-^{\mathrm{OMC}}(\boldsymbol{x},\boldsymbol{x}')] \leq \mathrm{Var}[\tilde{k}_+^{\mathrm{MC}}(\boldsymbol{x},\boldsymbol{x}')] + \mathrm{Var}[\tilde{k}_-^{\mathrm{MC}}(\boldsymbol{x},\boldsymbol{x}')] = \mathrm{Var}[\tilde{k}^{\mathrm{MC}}(\boldsymbol{x},\boldsymbol{x}')].$$

which implies $\mathrm{MSE}[\tilde{k}^{\mathrm{OMC}}(\boldsymbol{x},\boldsymbol{x}')] \leq \mathrm{MSE}[\tilde{k}^{\mathrm{MC}}(\boldsymbol{x},\boldsymbol{x}')]$ for sufficiently large d.

□

C Polynomial kernels on the unit sphere with finite total mass

We consider the asymptotic properties of the Bessel function of the first kind $J_\alpha(x)$ under the large and small cases to study the $\|\mu\|$.

C.1 A small ω

Consider the asymptotic behavior for small ω. The Bessel function of the first kind is asymptotically equivalent to

$$J_\alpha(x) \sim \frac{1}{\Gamma(\alpha+1)}\left(\frac{x}{2}\right)^\alpha, \text{ when } 0 < x \ll \sqrt{\alpha+1}.$$

In this case, the measure μ is formulated as

$$\mu(\omega) \sim \sum_{i=0}^{p} \frac{p!}{(p-i)!}\left(1-\frac{4}{a^2}\right)^{p-i}\left(\frac{2}{a^2}\right)^i \frac{2^{d/2+i}}{\Gamma(d/2+i+1)}, \tag{5}$$

which can be regarded as a generalized version of a uniform distribution. Therefore, μ is absolutely integrable over a finite range $(0, c_1]$, where c_1 is some constant satisfying $c_1 \ll \sqrt{\frac{d}{2}-1}$.

C.2 A large ω

Consider the asymptotic behavior for large ω. The Bessel function of the first kind is asymptotically equivalent to

$$J_\alpha(x) \sim \sqrt{\frac{2}{\pi x}} \cos(x - \frac{\pi\alpha}{2} - \frac{\pi}{4}), \text{ when } x \gg |\alpha^2 - \frac{1}{4}| .$$

The Fourier transform of the polynomial kernel on the sphere, i.e., the measure μ, is hence given by [8]

$$\mu(\omega) \sim \frac{1}{\sqrt{\pi\omega}} \left(1 - \frac{4}{a^2}\right)^p \left(\frac{2}{\omega}\right)^{d/2} \cos\left((d+1)\frac{\pi}{4} - 2\omega\right) , \text{ for a large } \omega . \tag{6}$$

In this way, we have $\int_{c_2}^\infty |\mu(\omega)| \,\mathrm{d}\omega < \infty$ for a large ω, where c is some constant satisfying $c_2 \gg \frac{1}{4}|d^2 - 1|$.

Accordingly, combining Eq. (6) with Eq. (5), we conclude that

$$\|\mu\| := \int_0^\infty |\mu(\omega)| \,\mathrm{d}\omega = \int_0^{c_1} |\mu(\omega)| \,\mathrm{d}\omega + \int_{c_1}^{c_2} |\mu(\omega)| \,\mathrm{d}\omega + \int_{c_2}^\infty |\mu(\omega)| \,\mathrm{d}\omega < \infty ,$$

where we use $\int_{c_1}^{c_2} |\mu(\omega)| \,\mathrm{d}\omega$ is finite due to the continuous, bounded Bessel function $J_\alpha(x)$ on a finite region $[c_1, c_2]$.

D Proof of Theorem 4

To prove Theorem 4, we firstly derive its formulation on the unit sphere and then demonstrate that it is a shift-invariant but not positive definite kernel via *completely monotone* functions.

Definition 3. *(Completely monotone [64]) A function f is called completely monotone on $(0, +\infty)$ if it satisfies $f \in C^\infty(0, +\infty)$ and*

$$(-1)^r f^{(r)}(x) \geq 0 ,$$

for all $r = 0, 1, 2, \cdots$ and all $x > 0$. Moreover, f is called completely monotone on $[0, +\infty)$ if it is additionally defined in $C[0, +\infty)$.

Note that the definition of completely monotone functions can be also restricted to a finite interval, i.e., f is completely monotone on $[a, b] \subset \mathbb{R}$, see in [8].

Besides, we need the following lemma that demonstrates the connection between positive definite and completely monotone functions for the proof.

Lemma 1. *(Schoenberg's theorem [64]) A function f is completely monotone on $[0, +\infty)$ if and only if $f := g(\|\cdot\|_2^2)$ is radial and positive definite function on all $\mathbb{R}^d$ for every d.*

Now let us prove Theorem 4.

Proof. By virtue of $\langle \boldsymbol{x}, \boldsymbol{x}' \rangle = 1 - \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}'\|_2^2$ and $\|\boldsymbol{x}\|_2 = \|\boldsymbol{x}'\|_2 = 1$, we have $\|\boldsymbol{x} - \boldsymbol{x}'\|_2 \in [0, 2]$. Therefore, the standard NTK of a two-layer ReLU network can be formulated as

$$\begin{aligned} k(\boldsymbol{x}, \boldsymbol{x}') &= \langle \boldsymbol{x}, \boldsymbol{x}' \rangle \kappa_0(\langle \boldsymbol{x}, \boldsymbol{x}' \rangle) + \kappa_1(\langle \boldsymbol{x}, \boldsymbol{x}' \rangle) \\ &= \left(1 - \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}'\|_2^2\right) \kappa_0 \left(1 - \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}'\|_2^2\right) + \kappa_1 \left(1 - \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}'\|_2^2\right) \\ &= \frac{2 - \|\boldsymbol{x} - \boldsymbol{x}'\|_2^2}{\pi} \arccos(\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}'\|_2^2 - 1) + \frac{\|\boldsymbol{x} - \boldsymbol{x}'\|_2}{2\pi} \sqrt{4 - \|\boldsymbol{x} - \boldsymbol{x}'\|_2^2} \\ &= \frac{2 - z^2}{\pi} \arccos(\frac{1}{2}z^2 - 1) + \frac{z}{2\pi}\sqrt{4 - z^2} , z := \|\boldsymbol{x} - \boldsymbol{x}'\|_2 \in [0, 2] , \end{aligned}$$

which is shift-invariant.

Next, we prove that $k(z)$ is not a positive definite kernel, i.e., $g(\sqrt{z}) := k(z)$ is not a completely monotone function over $[0, \infty)$ by Lemma 1. In other words, there exist some value $x \in [0, \infty)$ such that $(-1)^l g^{(l)}(x) < 0$ for some l. To this end, the function g is given by

$$g(x) = \frac{2 - x}{\pi} \arccos(\frac{1}{2}x - 1) + \frac{1}{2\pi}\sqrt{4x - x^2} , \ x \in [0, 4] ,$$

and its first-order derivative is

$$g'(x) = \frac{4-2x}{4\pi\sqrt{4x-x^2}} - \frac{2-x}{2\pi\sqrt{1-\left(\frac{x}{2}-1\right)^2}} - \frac{\arccos\left(\frac{x}{2}-1\right)}{\pi}\,.$$

Since $g'(x)$ is continuous, and $\lim_{x\to 0} g'(x) = -\infty$ and $\lim_{x\to 4} g'(x) = \infty$, there exists a constant c such that $g'(x) < 0$ over $(0,c)$ and $g'(x) > 0$ over $(c,4)$. That is to say, $(-1)^l g^{(l)}(x) < 0$ holds for $x \in (c,4)$, which violates the definition of completely monotone functions. In this regard, $g(\sqrt{z}) := k(z)$ is not a completely monotone function over $[0,\infty)$ and thus $\{k(z), z\in[0,2]; 0, z>2\}$ is not positive definite. □

E The measure of arc-cosine kernels on the unit sphere

According to Appendix D, the zero/first-order arc-cosine kernel on the unit sphere is proven to be stationary but indefinite. In this section, we derive its measure μ.

E.1 The measure of the zero-order arc-cosine kernel

In this section, we derive the measure μ of the zero-order arc-cosine kernel on the unit sphere.

Proposition 3. *The measure μ of the zero-order arc-cosine kernel on the unit sphere: $\kappa_0(\boldsymbol{x},\boldsymbol{x}') := \kappa_0(z) = \frac{1}{\pi}\arccos(\frac{1}{2}z^2-1)$ is given by*

$$\mu(\omega) = (\frac{1}{\omega})(\frac{2}{\omega})^{\frac{d}{2}-1} J_{\frac{d}{2}}(2\omega) - \frac{1}{\pi}(\frac{1}{\omega})^{\frac{d}{2}-2}\sum_{j=0}^{\infty}\frac{(2j)!}{4^j(j!)^2(2j+1)}\int_0^2\left(\frac{1}{2}z^2-1\right)^{2j+1}\omega z^{d/2}J_{d/2-1}(z\omega)\mathrm{d}z\,,$$

where the integral $\int_0^2\left(\frac{1}{2}z^2-1\right)^{2j+1}\omega z^{d/2}J_{d/2-1}(z\omega)\mathrm{d}z$ can be computed by parts with the following simple recurrence formula

$$\int z^a J_{v+1}(z)\mathrm{d}z = 2v\int z^{a-1}J_v(z)\mathrm{d}z - \int z^a J_{v-1}(z)\mathrm{d}z\,. \tag{7}$$

Proof. According to the definition of $\kappa_0(z)$, we have

$$\mu(\omega) = \int_0^2 \frac{z}{\pi}\arccos(\frac{1}{2}z^2-1)(z/\omega)^{d/2-1}J_{d/2-1}(z\omega)\mathrm{d}z\,, \tag{8}$$

where $\kappa_0(\boldsymbol{z})$ is a radial function, i.e., $\kappa_0(\boldsymbol{z}) = \kappa_0(z)$ with $z := \|\boldsymbol{z}\|_2$, and thus its Fourier transform is also a radial function, i.e., $\mu(\omega) = \mu(\boldsymbol{\omega})$ with $\omega := \|\boldsymbol{\omega}\|_2$. Obviously, the integrand in Eq. (8) and the integration region are both bounded, and thus we have $\mu(\omega) < \infty$. Following the proof of $\|\mu\| < \infty$ for polynomial kernels on the unit sphere in Section C, we can also demonstrate that $\|\mu\| < \infty$ for the zero-order arc-cosine kernel on the unit sphere.

To compute the integration in Eq. (8), we take the Taylor expansion of $\arccos(\frac{1}{2}z^2-1)$ with t terms

$$\arccos(\frac{1}{2}z^2-1) = \frac{\pi}{2} - \sum_{j=0}^{t}\frac{(2j)!}{4^j(j!)^2(2j+1)}\left(\frac{1}{2}z^2-1\right)^{2j+1}\,,$$

and thus the integration in Eq. (8) can be integrated by each term regarding to Bessel functions. Moreover, by virtue of $\frac{\mathrm{d}z^v J_v(z\omega)}{\mathrm{d}z} = \omega z^v J_{v-1}(z\omega)$, the above integral can be computed by parts

$$\begin{aligned}\mu(\omega) &= \int_0^2 \frac{z}{\pi}\arccos(\frac{1}{2}z^2-1)(z/\omega)^{d/2-1}J_{d/2-1}(z\omega)\mathrm{d}z\\ &= \frac{1}{2}(\frac{1}{\omega})^{\frac{d}{2}-2}\int_0^2 \omega z^{\frac{d}{2}}J_{\frac{d}{2}-1}(z\omega)\mathrm{d}z - \frac{1}{\pi}(\frac{1}{\omega})^{\frac{d}{2}-2}\sum_{j=0}^{\infty}\frac{(2j)!}{4^j(j!)^2(2j+1)}\int_0^2\left(\frac{1}{2}z^2-1\right)^{2j+1}\omega z^{d/2}J_{d/2-1}(z\omega)\mathrm{d}z\,,\end{aligned} \tag{9}$$

where the first term equals to $(\frac{1}{\omega})(\frac{2}{\omega})^{\frac{d}{2}-1}J_{\frac{d}{2}}(2\omega)$. Accordingly, we can conclude our proof. □

It appears non-trivial to prove $\|\mu\| < \infty$ as Eq. (9) is quite complex. Here we choose $j = 0$ in Eq. (9) as an example, we have

$$\begin{aligned}\int_0^2 \left(\frac{1}{2}z^2 - 1\right) \omega z^{d/2} J_{d/2-1}(z\omega)\mathrm{d}z &= 2^{\frac{d}{2}} J_{\frac{d}{2}}(2\omega) - \int_0^2 z^{\frac{d}{2}+1} J_{\frac{d}{2}}(z\omega) \left(\frac{1}{2}z^2 - 1\right) \mathrm{d}z \\ &= 2^{\frac{d}{2}} J_{\frac{d}{2}}(2\omega) + \frac{1}{\omega} J_{\frac{d}{2}+1}(2\omega) - \frac{1}{2}\int_0^2 z^{\frac{d}{2}+3} J_{\frac{d}{2}}(z\omega)\mathrm{d}z\,,\end{aligned} \tag{10}$$

where $\int_0^2 z^{\frac{d}{2}+3} J_{\frac{d}{2}}(z\omega)\mathrm{d}z$ can be computed by parts

$$\int_0^2 z^{\frac{d}{2}+3} J_{\frac{d}{2}}(z\omega)\mathrm{d}z = 2^{\frac{d}{2}+3} J_{\frac{d}{2}}(2\omega) - \frac{1}{\omega^2} 2^{\frac{d}{2}+2} J_{\frac{d}{2}+2}(2\omega)\,. \tag{11}$$

Incorporating Eqs. (11), (10) into Eq. (9), we have

$$\mu(\omega) = \left(\frac{1}{\omega}\right)\left(\frac{2}{\omega}\right)^{\frac{d}{2}-1} J_{\frac{d}{2}}(2\omega) - \frac{1}{\pi}(\frac{1}{\omega})^{\frac{d}{2}-2}\left[(-3)2^{\frac{d}{2}} J_{\frac{d}{2}}(2\omega) + \frac{1}{\omega} J_{\frac{d}{2}+1}(2\omega) + \frac{1}{\omega^2} 2^{\frac{d}{2}+1} J_{\frac{d}{2}+2}(2\omega)\right]\,.$$

Following with the proof in Section C, we can demonstrate $\|\mu\| < \infty$ by the asymptotic equivalence of Bessel functions. Accordingly, in this case, μ can be decomposed into two nonnegative measures with $\mu(\omega) = \mu_+(\omega) - \mu_-(\omega)$, where $\mu_+(\omega) = \max\{0, \mu(\omega)\}$ and $\mu_-(\omega) = \max\{0, -\mu(\omega)\}$. As a consequence, Algorithm 1 is also suitable for this kernel.

E.2 the first-order arc-cosine kernel

In this subsection, we derive the measure μ of the zero-order arc-cosine kernel admitting $\kappa_1(\boldsymbol{x}, \boldsymbol{x}') = \frac{z}{2\pi}\sqrt{4-z^2}$.

Proposition 4. *The measure μ of the zero-order arc-cosine kernel on the unit sphere: $\kappa_0(\boldsymbol{x}, \boldsymbol{x}') := \kappa_1(z) = \frac{z}{2\pi}\sqrt{4-z^2}$ is given by*

$$\mu(\omega) = (\frac{1}{\omega})(\frac{2}{\omega})^{\frac{d}{2}-1} J_{\frac{d}{2}}(2\omega) - \frac{1}{\pi}(\frac{1}{\omega})^{\frac{d}{2}-2} \sum_{j=0}^{\infty} \frac{(2j)!}{4^j (j!)^2 (2j+1)} \int_0^2 \left(\frac{1}{2}z^2 - 1\right)^{2j+1} \omega z^{d/2} J_{d/2-1}(z\omega)\mathrm{d}z\,,$$

where the integral $\int_0^2 \left(\frac{1}{2}z^2 - 1\right)^{2j+1} \omega z^{d/2} J_{d/2-1}(z\omega)\mathrm{d}z$ can be computed by parts with the following simple recurrence formula (7)*.*

Proof. By fractional binomial theorem, we have

$$\binom{1/2}{j} = (-1)^{k-1}\frac{1}{2(2j-1)}\frac{(2j)!}{(2\cdot 4\cdot\cdot(2j))^2} = \frac{-1}{2(2j-1)}\left(-\frac{1}{4}\right)^j \binom{2j}{j}\,.$$

Then, according to the definition of $\kappa_1(z)$, we have

$$\sqrt{4-z^2} = 2\left(1 - \frac{z^2}{4}\right)^{\frac{1}{2}} = 2\sum_{j=0}^{\infty}\binom{1/2}{j}\left(-\frac{z^2}{4}\right)^j = \sum_{j=0}^{\infty}\frac{-1}{2j-1}\binom{2j}{j}\left(\frac{z}{4}\right)^{2j}\,.$$

Therefore, the measure μ of κ_1 is

$$\begin{aligned}\mu(\omega) &= \frac{1}{2\pi}\int_0^2 z^2\sqrt{4-z^2}(z/\omega)^{d/2-1} J_{d/2-1}(z\omega)\mathrm{d}z \\ &= \frac{1}{2\pi}\int_0^2 z^2 (z/\omega)^{d/2-1} J_{d/2-1}(z\omega)\sum_{j=0}^{\infty}\frac{-1}{2j-1}\binom{2j}{j}\left(\frac{z}{4}\right)^{2j}\mathrm{d}z.\end{aligned} \tag{12}$$

Accordingly, the above equation needs to compute the following integral

$$\int_0^2 z^{\frac{d}{2}+1+2j} J_{\frac{d}{2}-1}(z\omega)\mathrm{d}z\,,$$

which can be computed by Eq. (7). ∎

Similarly, it appears non-trivial to prove $\|\mu\| < \infty$ as Eq. (12) is quite complex. Here we choose $j = 0$ in Eq. (12) as an example, we have

$$\mu(\omega) = \frac{1}{2\pi} \int_0^2 z^2 (z/\omega)^{d/2-1} J_{d/2-1}(z\omega) \mathrm{d}z = \frac{\sqrt{2}-1}{2\pi} \left(\frac{1}{\omega}\right)^{\frac{d}{2}-2} 2^{\frac{d}{2}} J_{\frac{d}{2}}(2\omega)\,.$$

In this case, it is clear that $\|\mu\| < \infty$ and thus Algorithm 1 is also suitable for this kernel.