Journal of Multivariate Analysis 167 (2018) 306–318

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Equivalence and orthogonality of Gaussian measures on spheres

Ahmed Arafat [a,*], Emilio Porcu [b,c], Moreno Bevilacqua [d], Jorge Mateu [e]

[a] Department of Mathematics, Mansoura University, Mansoura, Egypt
[b] School of Mathematics and Statistics, University of Newcastle, Great Britain, United Kingdom
[c] Departamento de Matemática, Universidad Técnica Federico Santa María, 2360102 Valparaíso, Chile
[d] Instituto de Estadística, Universidad de Valparaíso, 2360102 Valparaíso, Chile
[e] Departament de Matemàtiques, Universitat Jaume I, Castellón de la Plana, Spain

A R T I C L E I N F O

Article history:
Received 23 May 2017
Available online 22 May 2018

AMS 2010 subject classifications:
28C20
42A82
42A16
86A32

Keywords:
Chordal distance
Equivalence of Gaussian measures
Great circle distance
Positive definite functions
Schoenberg coefficients

A B S T R A C T

The equivalence of Gaussian measures is a fundamental tool to establish the asymptotic properties of both prediction and estimation of Gaussian fields under fixed domain asymptotics. The paper solves Problem 18 in the list of open problems proposed by Gneiting (2013). Specifically, necessary and sufficient conditions are given for the equivalence of Gaussian measures associated to random fields defined on the d-dimensional sphere $\mathbb{S}^d$, and with covariance functions depending on the great circle distance. We also focus on a comparison of our result with existing results on the equivalence of Gaussian measures for isotropic Gaussian fields on $\mathbb{R}^{d+1}$ restricted to the sphere $\mathbb{S}^d$. For such a case, the covariance function depends on the chordal distance being an approximation of the true distance between two points located on the sphere. Finally, we provide equivalence conditions for some parametric families of covariance functions depending on the great circle distance. An important implication of our results is that all the parameters indexing some families of covariance functions on spheres can be consistently estimated. A simulation study illustrates our findings in terms of implications on the consistency of the maximum likelihood estimator under fixed domain asymptotics.

1. Introduction

1.1. Motivation

Statistical analysis of processes defined over the entire globe has attracted a lot of attention in recent years and as a result, the literature on Gaussian random fields defined over spheres is becoming ubiquitous in areas as diverse as mathematical analysis [3,26,40,41], probability theory [2,16,27,36], spatial point processes [42], spatial geostatistics [22,28,29], space–time geostatistics [4,15,44], and mathematical physics [32,37,38]. Global models, especially for climate data, are also finding applications in many fields; see, e.g., [9,10,12,44], and references therein.

The equivalence of Gaussian measures [30,47] represents an essential tool to establish the asymptotic properties of both prediction and estimation of Gaussian fields defined on Euclidean spaces, under fixed domain (also called infill) asymptotics. Such a framework typically applies when more and more data are collected by sampling densely in a bounded

* Corresponding author.
E-mail addresses: ah.arafat@mans.edu.eg (A. Arafat), emilio.porcu@newcastle.ac.uk (E. Porcu), moreno.bevilacqua@uv.cl (M. Bevilacqua), mateu@mat.uji.es (J. Mateu).

https://doi.org/10.1016/j.jmva.2018.05.005

set. Notable examples of application of the equivalence of Gaussian measures to fixed domain asymptotics for prediction can be found in [50,53,54]. The equivalence of Gaussian measures has been used in [57], and more recently in [5], for specific parametric families of covariance functions.

Fixed domain asymptotics is a natural framework for studying Gaussian random fields defined on d-dimensional spheres. However, the literature on conditions for the equivalence of Gaussian measures associated to stochastic processes over d-dimensional spheres has been sparse. We quote verbatim from Problem 18 in [24] as follows

> "Stein [53] provides a comprehensive discussion of the effects of misspecification of the covariance structure in spatial statistical models. Under infill asymptotics, the Euclidean case is well understood, and many technical details depend on the equivalence (or not) of the corresponding Gaussian measures. In this light, what are necessary and/or sufficient conditions for the equivalence of mean zero Gaussian measures that are indexed by d-Schoenberg sequences?"

This paper solves this problem and analyzes the implications in terms of equivalence of Gaussian measures for some parametric families of covariance functions.

1.2. Discussion of the literature on Euclidean spaces

Equivalence and orthogonality of probability measures for Gaussian fields defined over bounded sets of $\mathbb{R}^{d+1}$ and with a given stationary covariance function, have been studied by Gikhman et al. [21], Skorokhod et al. [47], and Da Prato et al. [17]. There has been substantial work on Gaussian measures, and we refer to notable contributions in [21,34,35,45,54–56]. Excellent treatises on the equivalence of Gaussian measures can be found in the textbooks by Bogachev [7] and Chatterji and Mandrekar [14]. An interesting bridge between equivalence of Gaussian measures and stochastic partial differential equations has been proposed in [31,48], and more recently in [11].

The results obtained in this direction motivated [50–54] to engage a "tour de force" in studying the effect of kriging prediction with a misspecified covariance function, and under fixed domain asymptotics.

A simple sufficient condition for the equivalence of two Gaussian measures is given in [47]. In the same paper, conditions are provided for the equivalence of Gaussian measures associated to isotropic Gaussian fields defined on $\mathbb{R}^{d+1}$ and restricted to the d-dimensional sphere. The result is limited to Gaussian processes with a covariance function that depends on the chordal distance.

There has been some criticism around the use of the chordal distance, and we refer the reader to [23] and to [44], where it is also argued that the correct metric for stochastic processes over spheres is the great circle distance, which describes an arc between any points located on the sphere. Hence, the need for studying conditions for equivalence of Gaussian measures associated to stochastic processes defined on the d-dimensional sphere.

1.3. Background

Let d be a positive integer. We consider the unit sphere $\mathbb{S}^d$ of $\mathbb{R}^{d+1}$, defined as $\mathbb{S}^d = \{\mathbf{x} \in \mathbb{R}^{d+1} : \|\mathbf{x}\|^2 = 1\}$, where $\|\cdot\|$ denotes the Euclidean norm. The natural distance on the sphere is the geodesic or great circle distance, defined as the mapping $\theta : \mathbb{S}^d \times \mathbb{S}^d \to [0, \pi]$ so that, for all $\mathbf{x}, \mathbf{y} \in \mathbb{S}^d$, $\theta(\mathbf{x}, \mathbf{y}) = \arccos(\langle \mathbf{x}, \mathbf{y} \rangle)$, with $\langle \cdot, \cdot \rangle$ denoting the classical dot product in $\mathbb{R}^d$. Thus, the geodesic distance describes an arc between any pair of points located on $\mathbb{S}^d$. Throughout, we shall equivalently use $\theta(\mathbf{x}, \mathbf{y})$ or its shortcut θ to denote the geodesic distance, whenever no confusion can arise.

An approximation of the true distance between any two points on the sphere is the chordal distance d_{CH} given, for all $\mathbf{x}, \mathbf{y} \in \mathbb{S}^d$, by

$$d_{\mathrm{CH}}(\mathbf{x}, \mathbf{y}) = 2 \sin \{\theta(\mathbf{x}, \mathbf{y})/2\} . \quad (1)$$

Thus, the chordal distance defines the segment "below" the arc joining two points on the sphere. We consider zero mean Gaussian fields $\{Z(\mathbf{x}) : \mathbf{x} \in \mathbb{S}^d\}$ with finite second order moment. The finite-dimensional distributions are completely specified by the covariance function $K : \mathbb{S}^d \times \mathbb{S}^d \to \mathbb{R}$ defined, for all $\mathbf{x}, \mathbf{y} \in \mathbb{S}^d$, by $K(\mathbf{x}, \mathbf{y}) = \mathrm{cov}\{Z(\mathbf{x}), Z(\mathbf{y})\}$. Covariance functions are positive definite: for any N distinct points $\{\mathbf{x}_i\}_{i=1}^N \subset \mathbb{S}^d$ and constants $c_1, \ldots, c_N \in \mathbb{R}$, we have

$$\sum_{i=1}^{N} \sum_{j=1}^{N} c_i K(\mathbf{x}_i, \mathbf{x}_j) c_j \geq 0;$$

see [6]. Porcu et al. [43] call K geodesically isotropic if $K(\mathbf{x}, \mathbf{y}) = \sigma^2 \psi\{\theta(\mathbf{x}, \mathbf{y})\}$, for some mapping $\psi : [0, \pi] \to \mathbb{R}$ such that $\psi(0) = 1$. Here, σ^2 denotes the variance of Z. The function ψ is called the geodesically isotropic part of K [19]. Henceforth, we shall refer to both K and ψ as covariance functions, in order to simplify exposition. For a characterization of geodesic isotropy, the reader is referred to [46] and the essay in [23].

From [39] and references therein, any zero mean isotropic Gaussian process on the d-dimensional sphere admits a series representation of the type

$$Z(\mathbf{x}) = \sum_{n=0}^{\infty} \sum_{\ell=1}^{h(n,d+1)} \alpha_{n\ell} Y_{n\ell}(\mathbf{x}),$$

Table 1
Parametric families of covariance functions depending on the great circle distance. The last column reports their d-Schoenberg coefficients. In all cases, σ is strictly positive.

Family	Expression	Parameters restrictions	d-Schoenberg coefficients
Multiquadric	$\psi_M(\theta; p, \tau, \sigma) = \sigma^2 \left\{ \frac{(1-\delta)^2}{1+\delta^2-2\delta\cos\theta} \right\}^{\tau}$	$\delta \in (0, 1), \tau = (d-1)/2$	$b_{n,d} = \sigma^2 \binom{2\tau+n-1}{n} \delta^n (1-\delta)^{2\tau}$
Sine power	$\psi_S(\theta; \alpha, \sigma) = \sigma^2\{1 - (\sin\theta/2)^{\alpha}\}$	$\alpha \in (0, 2)$	$b_{n,1} = \sigma^2\, \Gamma_{n+1}(\alpha/2)/\sqrt{2}$ $\Gamma_{n+1}(\alpha/2) = \frac{-1}{(n+1)!} \prod_{m=0}^{n} (m - \alpha/2)$
Exponential	$\psi_E(\theta; b, \sigma) = \sigma^2 \exp(-\theta/b)$	$b > 0$	$b_{n,1} = \frac{2b\sigma^2}{\pi(1+b^2n^2)} \{1 + (-1)^{n+1} \exp(-\pi/b)\}$
Askey	$\psi_A(\theta; c, \sigma) = \sigma^2(1-\theta/c)_+^3$	$c > 0$	$b_{0,1} = c\sigma^2\{1 - \exp(-\pi/c)\}/\pi$ $b_{n,1} = 240\sigma^2/(\pi c^5 n^6) \times$ $\{(cn)^2 + (cn)\sin(cn) + 4\cos(cn) - 4\}$
Møller	Unavailable	$\alpha, \beta, \kappa > 0$	$b_{0,1} = c\sigma^2/(4\pi)$ $b_{n,d} = \sigma^2\{1 + \beta \exp(n/\alpha)^{\kappa}\}^{-1}$

where $\mathbf{x} \in \mathbb{S}^d$ and $Y_{n\ell}$ are the spherical harmonics of degree n [18], with their uniquely determined cardinality $h(n, d+1)$, $n \in \mathbb{N} \cup \{0\}$; see Appendix A for details on spherical harmonics. Here, $\{\alpha_{n\ell} : n \in \mathbb{N} \cup \{0\},\ 1 \le \ell \le h(n, d+1)\}$ is a sequence of Gaussian random variables defined by

$$\alpha_{n\ell} = \frac{1}{\|\mathbb{S}^d\|} \int_{\mathbb{S}^d} Z(\mathbf{x}) Y_{n\ell}(\mathbf{x}) \mathrm{d}\omega_d(\mathbf{x}),$$

where $\|\mathbb{S}^d\|$ denotes the total mass of ω_d on $\mathbb{S}^d$ (see Appendix A), and ω_d being the surface measure of the d-dimensional sphere. The elements of the sequence $\alpha_{n\ell}$ satisfy $\mathrm{E}(\alpha_{n\ell}) = 0$ and

$$\mathrm{E}(\alpha_{n\ell}\, \alpha_{n'\ell'}) = \frac{\|\mathbb{S}^d\|\ c_{n,d}}{h(n, d+1)} \mathbf{1}_{(\{n=n'\} \bigcap \{\ell=\ell'\})}, \tag{2}$$

where $\mathrm{E}(\cdot)$ denotes mathematical expectation, and where $\mathbf{1}_A$ is the indicator function of a Borel set A. Here, $c_{n,d} \ge 0$, $n \in \mathbb{N} \cup \{0\}$, and the variance σ^2 of Z is given by

$$\sigma^2 = \sum_{n=0}^{\infty} c_{n,d}.$$

Indeed, arguments in [39] in concert with the addition theorem for spherical harmonics (see Corollary 1.2.8 in [18]) show that

$$K(\mathbf{x}, \mathbf{y}) = \mathrm{E}\{Z(\mathbf{x})Z(\mathbf{y})\} = \sigma^2\psi(\theta) = \sigma^2 \sum_{n=0}^{\infty} b_{n,d} \frac{C_n^{(d-1)/2}(\cos\theta)}{C_n^{(d-1)/2}(1)}, \tag{3}$$

where $\theta \in [0, \pi]$, and $\left(b_{n,d}\right)_{n=0}^{\infty}$ is a uniquely determined probability mass system; see [23].

Following [19], we call $\left(b_{n,d}\right)_{n=0}^{\infty}$ a sequence of d-Schoenberg coefficients, $b_{n,d}$. Here C_n^{λ} is the Gegenbauer polynomial [36] of degree n and order $\lambda > 0$. Note that $b_{n,d} = c_{n,d}/\sigma^2$, with $c_{n,d}$ as described in Eq. (2). A closed form for the d-Schoenberg coefficients is available as a by-product of the Funk–Hecke theorem; see Theorem 1.2.9 in [18]. Gneiting [23] classified the d-Schoenberg coefficients as follows. For $d \ge 2$, we have

$$b_{n,d} = \frac{2n+d-1}{2^{3-d}\pi} \frac{[\Gamma\{(d-1)/2\}]^2}{\Gamma(d-1)} \int_0^{\pi} \psi(\theta) C_n^{(d-1)/2}(\cos\theta)(\sin\theta)^{d-1} \mathrm{d}\theta. \tag{4}$$

For $d = 1$ and all $n \in \mathbb{N} \cup \{0\}$, we have

$$b_{0,1} = \frac{1}{\pi} \int_0^{\pi} \psi(\theta)\mathrm{d}\theta \quad \text{and} \quad b_{n,1} = \frac{2}{\pi} \int_0^{\pi} \cos(n\theta)\psi(\theta)\mathrm{d}\theta.$$

Examples of parametric families of geodesically isotropic covariances are reported in Table 1, together with their parameter restrictions and their d-Schoenberg coefficients. The Multiquadric family ψ_M has been introduced by [23]. For special cases, see also [13]. The Sine Power family was introduced by [49]. The validity of the Exponential and Askey families has been proved by [23] and the respective 1-Schoenberg coefficients have been given by [42]. Finally, we call Møller family the class of covariances specified directly through the d-Schoenberg coefficients, which does not admit an explicit closed form [42].

A Gaussian field defined on $\mathbb{R}^{d+1}$ is called isotropic in the Euclidean sense if its covariance function depends exclusively on Euclidean distance; see [19] for details. A different construction principle for a Gaussian field on the sphere is suggested in [47]. It is based on the restriction of an Euclidean isotropic Gaussian field on $\mathbb{R}^{d+1}$ to the d-dimensional sphere. In this case, the great circle distance is no longer the natural metric for the associated covariance function, which depends on the chordal distance $d_{\rm CH}$ defined in (1). In particular, we have

$$K(\mathbf{x},\mathbf{y}) = 2^{(d-1)/2}\Gamma\left(\frac{d+1}{2}\right)\int_0^\infty \frac{J_{(d-1)/2}\{\xi\ d_{\rm CH}(\mathbf{x},\mathbf{y})\}}{\{\xi\ d_{\rm CH}(\mathbf{x},\mathbf{y})\}^{(d-1)/2}}\,\mathrm{d}F(\xi), \tag{5}$$

where F is a positive and bounded measure, and where, for $\nu > 0$, J_ν denotes a Bessel function of order ν [25]. Thus, we have that, for any Euclidean isotropic covariance function on $\mathbb{R}^{d+1}$, the restriction of the covariance to the chordal distance $d_{\rm CH}$ is positive definite on $\mathbb{S}^d$.

We also need some background on equivalence and orthogonality of Gaussian probability measures. Denote by μ_1, μ_2 two probability measures defined on the same measurable space $(\Omega, \mathcal{B})$, where $\mathcal{B}$ denotes the Borel σ-algebra on Ω. Then μ_1 and μ_2 are said to be equivalent if $\mu_2(A) = 1$ for any $A \in \mathcal{B}$ implies $\mu_1(A) = 1$ and vice versa. Furthermore, μ_1 and μ_2 are said to be orthogonal if there exists an event A such that $\mu_2(A) = 1$ but $\mu_1(A) = 0$. Arguments in [17,30] and [34] show that Gaussian measures are either equivalent or orthogonal. For a real-valued Gaussian random field $Z = \{Z(\mathbf{x}) : \mathbf{x} \in \mathbb{S}^d\}$, to define previous concepts, we restrict the event A to the σ-algebra generated by Z. We emphasize this restriction by saying that the two measures are equivalent on the paths of Z.

1.4. Plan of the paper

The plan of the paper is the following. Section 2 states a characterization theorem for the equivalence of Gaussian measures on d-dimensional spheres in terms of d-Schoenberg coefficients. We then provide equivalence results for specific families of parametric covariances on spheres, described in Table 1. Section 3 illustrates our finding through simulations, and elucidates the statistical implications of our theoretical results in terms of ML estimation, under infill asymptotics. The paper ends with discussion. For neater exposition, Appendix A reports technical mathematical background that is not needed to read the main text. All the proofs of the new theoretical results are deferred to Appendix B. Finally, Appendix C generalizes the main results of Section 2 to two-point homogeneous spaces.

2. Results

We split this section into two parts. The former provides a characterization of equivalence in terms of d-Schoenberg coefficients as defined in Eq. (4), solving Problem 18 in [24]. The latter analyzes equivalence of Gaussian measures indexed by some parametric families of covariance functions listed in Table 1.

2.1. A characterization of equivalence on d-dimensional spheres

We start by stating this paper's main result.

Theorem 1. *Let μ_1, μ_2 be two zero mean Gaussian measures, such that, under μ_i, the random field $Z = \{Z(\mathbf{x}) : \mathbf{x} \in \mathbb{S}^d\}$ is Gaussian with covariance function uniquely determined through d-Schoenberg coefficients $b_{n,d,i}$, defined according to representation (3). Then, μ_1 and μ_2 are equivalent on the paths of Z if and only if*

$$\sum_{n=0}^{\infty} h(n, d+1)\left(\frac{b_{n,d,1}}{b_{n,d,2}} - 1\right)^2 < \infty. \tag{6}$$

Some comments are in order. Condition (6) is the analogue of the (only sufficient) condition provided in Theorem 4 of [47] in Euclidean spaces. In Theorem 6 of the same paper, Skorokhod and Yadrenko provide necessary and sufficient conditions for two Euclidean isotropic Gaussian random fields defined on $\mathbb{R}^{d+1}$ and restricted to $\mathbb{S}^d$. Namely, the Gaussian measures are equivalent if and only if

$$\sum_{n=0}^{\infty} h(n, d+1)\left(\frac{\breve{b}_{n,d,1}}{\breve{b}_{n,d,2}} - 1\right)^2 < \infty, \tag{7}$$

where, for $i \in \{1, 2\}$,

$$\breve{b}_{n,d,i} = 2^d\Gamma\left(\frac{d+1}{2}\right)\pi^{(d+1)/2}\int_0^\infty \frac{J^2_{n+d/2}(\lambda)}{\lambda^{d-1}}\,\mathrm{d}F_i(\lambda),$$

and F_i is a positive measure from the representation (5) of an isotropic covariance function depending on the chordal distance $d_{\rm CH}$. Thus, Theorem 6 of [47] and Theorem 1 are complementary since they are related to two different constructions of Gaussian fields over spheres, which in turn imply different geometries.

The consequences of Theorem 1 are not merely mathematical, as we are going to discuss below. The covariance models based on the great circle distance (see, e.g., Table 1) may not be, in general, valid when adapted to the chordal distance.

Table 2
Candidate functions C that can be adapted to the great circle or to the chordal distances, with their corresponding parameter restrictions. Here, $\mathcal{K}_\nu$ denotes a modified Bessel function of order $\nu > 0$. In all cases, b and σ are strictly positive.

Family	Expression	Parameters restrictions for $\psi(\theta) = C_{[0,\pi]}(\theta)$	Parameter restrictions for $C(d_{\mathrm{CH}})$
Matérn	$C_{\mathcal{M}}(t; b, \nu, \sigma) = \sigma^2 (t/b)^\nu \mathcal{K}_\nu\,(t/b)$	$\nu \in (0, 1/2]$	$\nu > 0$
Gen. Cauchy	$C_{\mathcal{C}}(t; b, \alpha, \beta, \sigma) = \sigma^2 \{1 + (t/b)^\alpha\}^{-\beta/\alpha}$	$\alpha \in (0, 1], \beta > 0$	$\alpha \in (0, 2], \beta > 0$
Dagum	$C_{\mathcal{D}}(t; b, \alpha, \tau, \sigma) = \sigma^2 \left[1 - [(t/b)^\tau/\{1 + (t/b)\}^\tau)^{\alpha/\tau}\right]$	$\tau \in (0, 1), \alpha \in (0, \tau)$	$\tau \in (0, 1), \alpha \in (0, \tau)$

Instead, any Euclidean isotropic model valid on $\mathbb{R}^{d+1}$ can be restricted to $\mathbb{S}^d$ by replacing the Euclidean distance with the chordal one.

Table 2 enriches the comparison of our Theorem 1 with Theorem 6 in [47]. Specifically, we consider candidate functions $C : [0, \infty) \to \mathbb{R}$ together with the corresponding adaptation to the spheres in terms of great circle or chordal distance. For the great circle, we have $\psi(\theta) = C_{[0,\pi]}(\theta)$, where $C_{[0,\pi]}$ denotes the restriction to the interval $[0, \pi]$. For the chordal distance, we just consider the composition $C(d_{\mathrm{CH}})$. A clear advantage of the chordal distance is that it allows for less restrictions on the parameters associated to the candidate functions C in Table 2. For instance, the Matérn family [53] has been very popular in geostatistics because of the parameter ν, which allows to govern the mean square differentiability of the associated Gaussian field. The severe restriction of the parameter space of the adaptation to the great circle [23] makes its use unpractical on the sphere. Instead, the corresponding adaptation to the chordal distance allows to attain positive definiteness on the sphere for any positive ν.

The use of the chordal distance has drawbacks as well. For instance, because the chordal distance underestimates the true distance between the points on the sphere, Porcu et al. [44] argue that this fact has a non-negligible impact on the estimation of the spatial scale. Moreover, Gneiting [23] argues that the chordal distance is counter to spherical geometry for larger values of the great circle distance, and thus may result in physically unrealistic distortions. Further, covariance functions based on the chordal distance inherit the limitations of isotropic models in Euclidean spaces in modeling covariances with negative values. For instance, a covariance based on the chordal distance on $\mathbb{S}^2$ does not allow for values lower than $-0.21\sigma^2$, with σ^2 being the variance as before. Instead, properties of Legendre polynomials imply that correlations based on the geodesic distance can attain any value between -1 and $+1$.

A final relevant remark is that verifying the Skorokhod– Yadrenko condition in Eq. (7) is prohibitive to say the least, because it involves the calculation of the coefficients $\breve{b}_{n,d}$, where the integral depends on Bessel functions. Instead, we show subsequently that our Theorem 1 can be verified at least for some parametric families.

2.2. Consequences of Theorem 1 for some parametric families

This section visits some classes of covariance functions that depend on the great circle distance. We start with the Multiquadric family $\psi_M(\cdot; \delta, \tau, \sigma)$, being the first entry in Table 1. Specifically, we provide the following result.

Proposition 1. *Let μ_1, μ_2 be two zero mean Gaussian measures on $\mathbb{S}^2$ such that for each $i \in \{1, 2\}$, under μ_i, the random field $Z = \{Z(\mathbf{x}) : \mathbf{x} \in \mathbb{S}^2\}$ is Gaussian with Multiquadric covariance function $\psi_i(\theta) = \psi_M(\theta; \delta_i, \tau, \sigma_i)$ with $\tau = (d-1)/2 = 1/2$, $\delta_i \in (0, 1)$ and $\sigma_i > 0$. Then, μ_1 and μ_2 are equivalent on the paths of Z if and only if $\sigma_1^2 = \sigma_2^2$ and $\delta_1 = \delta_2$.*

Proposition 1 has important implications both in terms of estimation and prediction of Gaussian fields under infill asymptotics. On the one hand, it shows that the parameters δ and σ^2 can be consistently estimated. On the other, Proposition 1 shows that misspecified kriging prediction has a nonnegligible impact even asymptotically.

We have been able to obtain a similar result for the case of two Gaussian measures with Sine Power covariance $\psi_S(\cdot; \alpha, \sigma)$ as defined through the second entry of Table 1.

Proposition 2. *Let μ_1, μ_2 be two zero mean Gaussian measures on $\mathbb{S}^1$ such that for each $i \in \{1, 2\}$, under μ_i, the random field $Z = \{Z(\mathbf{x}) : \mathbf{x} \in \mathbb{S}^1\}$ is Gaussian with covariance function $\psi_i(\theta) = \psi_S(\theta; \alpha_i, \sigma_i)$. Then, μ_1 and μ_2 are equivalent on the paths of Z if and only if $\alpha_1 = \alpha_2$ and $\sigma_1^2 = \sigma_2^2$.*

Similar comments apply here and the simulation results in Section 3 will confirm our findings. Working with condition (6) can be extremely difficult, because the d-Schoenberg coefficients might oscillate away from zero. This is the case, e.g., of the Exponential and Askey families [42], denoted ψ_E and ψ_A, respectively. Indeed, it has not been possible to obtain any analogue of Propositions 1 and 2 in these cases.

3. Simulation study

We now explore numerically the consequences of the results in Section 2 in terms of consistency of the ML estimator for some parametric families of covariance functions, under fixed domain asymptotics. To favor a neater exposition, we slightly abuse of notation when denoting σ_k^2, $k \in \{M, S, E\}$, the variances associated to the Multiquadric (ψ_M), the Sine Power (ψ_S) and the Exponential (ψ_E) models, being respectively the first three entries in Table 1.

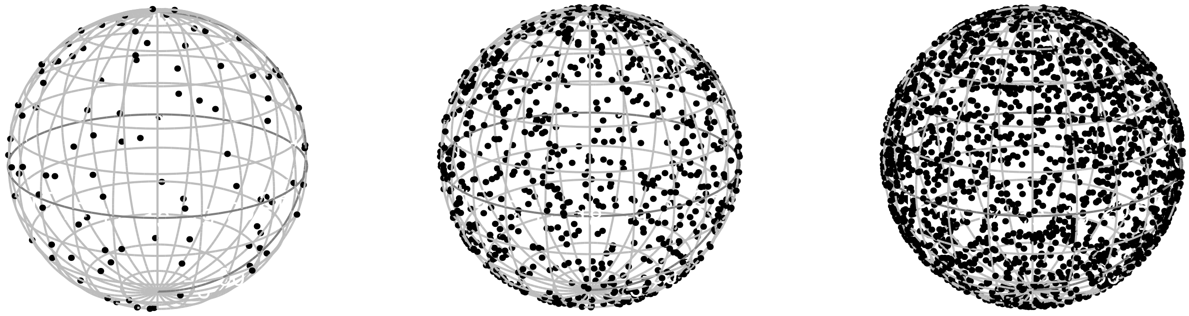

Fig. 1. Illustration of the increasing sequence of location sites on the unit sphere ($N \in \{100, 800, 2000\}$) considered in the simulation study.

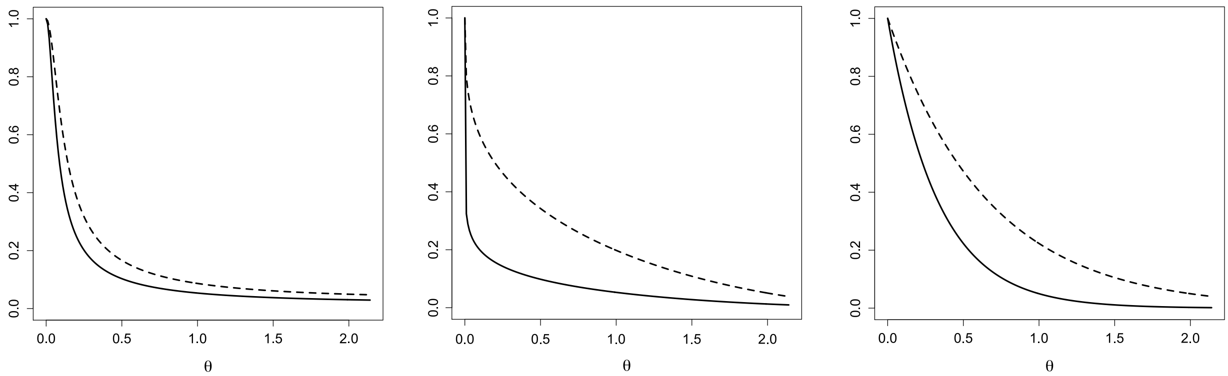

Fig. 2. From left to right: Multiquadric (ψ_M), Sine Power (ψ_S) and Exponential (ψ_E) correlation functions considered in the simulation study, with practical range equal to 1 (Scenario A, solid line) and equal to 2 (Scenario B, dashed line).

We first give an account on the computation of the great circle distance on the unit sphere $\mathbb{S}^2$ embedded in $\mathbb{R}^3$, where any point $\mathbf{x}$ has spherical coordinates $\mathbf{x} = (\varphi, \vartheta)^\top$, with $\varphi \in [0, \pi]$ and $\vartheta \in [0, 2\pi)$ being, respectively, the polar and the azimuthal angles (equivalently, latitude and longitude). Here, $^\top$ stands for the transpose operator. Then, the great circle distance is computed, for any $\mathbf{x}_1, \mathbf{x}_i \in \mathbb{S}^2$, through

$$\theta(\mathbf{x}_1, \mathbf{x}_2) = \arccos\left(\sin\varphi_1 \sin\varphi_2 + \cos\varphi_1 \cos\varphi_2 \cos(|\vartheta_1 - \vartheta_2|)\right).$$

The results in Propositions 1 and 2 are somehow counterintuitive if compared with those obtained for classes of Euclidean isotropic covariance functions. For instance, Zhang [57] proved that two zero mean Gaussian measures with Euclidean isotropic Exponential covariance functions and parameters $b_i, \sigma^2_{E,i}$ with $i \in \{1, 2\}$, are equivalent on any bounded set of $\mathbb{R}^d$, for $d \in \{1, 2, 3\}$, if and only if

$$\sigma^2_{E,1} b_1^{-1} = \sigma^2_{E,2} b_2^{-1}. \tag{8}$$

This result implies that the parameters σ^2_E and b cannot be estimated consistently, under fixed domain asymptotics, on $\mathbb{R}^d$, for $d \in \{1, 2, 3\}$. Instead, the so-called microergodic parameter [53], defined by $\sigma^2_E b^{-1}$, is consistently estimable. A brilliant approach in [1] shows instead that all parameters of the Exponential covariance can be estimated consistently when $d > 4$. Similar results are obtained in [5] for certain parametric classes of compactly supported correlation functions.

The results in Propositions 1 and 2 imply that all the parameters of the Multiquadric model (ψ_M) on $\mathbb{S}^2$, and of the Sine Power model (ψ_S), on the circle $\mathbb{S}^1$, can be estimated consistently under fixed domain asymptotics. To confirm this finding, we first consider 2000 points being uniformly distributed over the unit sphere. Then, we mimic a typical asymptotic setting, by considering the increasing sequence $N \in \{100, 200, 400, 800, 1200, 1600, 2000\}$ points, randomly chosen from the original set of 2000 points; see Fig. 1 for the cases $N \in \{100, 800, 2000\}$. For each N, 1000 replicates of zero mean Gaussian fields are simulated through Cholesky decomposition, using either the Multiquadric, the Sine Power, or the Exponential model. We set unit variance for all cases ($\sigma^2_k = 1, k \in \{M, S, E\}$), and consider the following scenarios:

Scenario A: $\delta = 0.95, \alpha = 0.074, b = 1/3$; Scenario B: $\delta = 0.92, \alpha = 0.30, b = 2/3$.

Scenarios A and B are depicted in Fig. 2. Under Scenario A, the three covariances have a common practical range, approximately equal to 1. The same criterion is used for Scenario B, under which the three covariances have a common

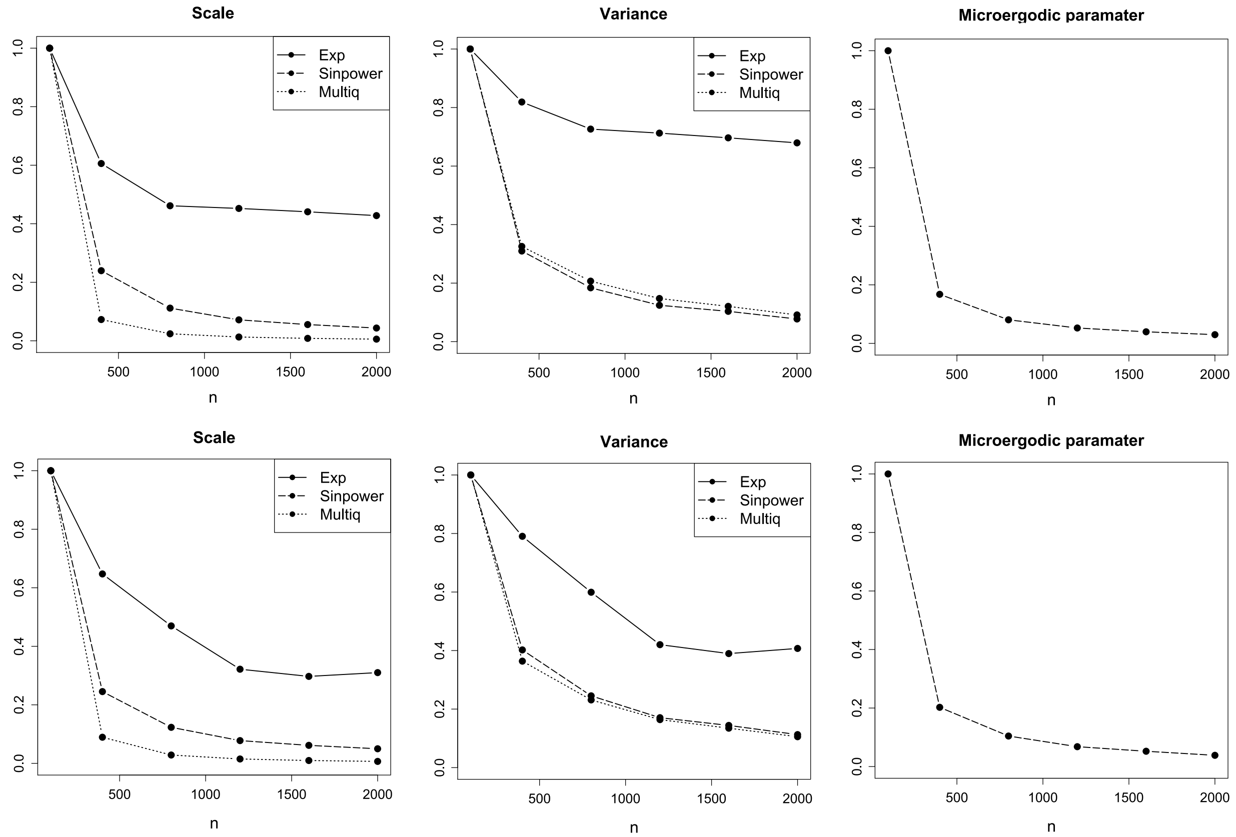

Fig. 3. Relative sample variance of the maximum likelihood estimates of the scale (first column) and variance (second column) parameters when increasing the number of location sites on the unit sphere for the Exponential, Sine Power and Multiquadric covariance models, under Scenarios A (first row) and B (second row). Relative sample variance of the maximum likelihood estimates of the microergodic parameter in the Exponential case (third column), under Scenarios A (first row) and B (second row).

practical range, equal to 2. For each N, model and simulation, we estimate through ML the parameters $(\sigma_M^2, \delta)^\top$, $(\sigma_S^2, \alpha)^\top$ and $(\sigma_E^2, b)^\top$ respectively for models ψ_M, ψ_S and ψ_E.

In order to check for consistency of the parameters we look at the behavior of the sample variance of the maximum likelihood estimates when increasing $N \in \{100, 400, \ldots, 2000\}$. To take into account the different orders of magnitude of the variances of each parameter, we consider a relative sample variance: for each N, we first consider the sample variances of the estimates. Then, the ratio between the sample variances and the maximum of them is computed.

As expected, for $N = 100$, the relative sample variance is overall equal to 1. Fig. 3 (first column) shows for each model how the sample relative variance of the ML estimates of the scale parameters δ, α and b decreases when the number of location sites increases under Scenarios A and B (from top to bottom). Fig. 3 (second column) shows for each model how the sample variance of the maximum likelihood estimates of σ_k^2, $k \in \{M, S, E\}$ decreases when increasing the number of location sites, under Scenarios A and B (from top to bottom). Finally, Fig. 3 (third column) shows the relative sample variance of the maximum likelihood estimates of the microergodic parameter $\sigma_E^2 b^{-1}$ for the Exponential model.

Note that we do not have any theoretical result for the Exponential model. Nevertheless from this example it becomes apparent that sampling more data on $\mathbb{S}^2$ may not improve the joint estimation of the scale and variance parameters when using the Exponential model. Instead, it is apparent from Fig. 3 (third column), that the microergodic parameter can be estimated consistently. These numerical results suggest that the equivalence condition in Eq. (8) is still valid at least on $\mathbb{S}^2$.

For the Multiquadric covariance model, as expected from Proposition 1, there is a clear pattern of decreasing relative sample variance for both parameters when jointly estimating the scale and variances parameters. For the Sine Power model, even if Proposition 2 is valid on $\mathbb{S}^1$, our simulation results suggest that orthogonality is still valid on $\mathbb{S}^2$.

4. Discussion

This work lays down the basis for studying the asymptotic properties of ML estimation and misspecified kriging prediction under fixed domain asymptotics, for covariance models depending on geodesic isotropy. A necessary step in terms of future research is to use the results of this paper as building blocks for more sophisticated constructions. For instance, using the stochastic expansion proposed by [33], it should be possible to obtain equivalence conditions under the hypothesis of axial symmetry. This should be made at the expense of relaxing Condition (2) and evoking again the constructive arguments used in Appendix B to prove Theorem 2.

Acknowledgments

Research work of Moreno Bevilacqua was partially supported by grant FONDECYT 1160280 from the Chilean government. Research work of Emilio Porcu was partially supported by grant FONDECYT 1170290 from the Chilean government. Jorge Mateu and Ahmed Arafat were partially supported by grant MTM2016-78917-R from the Spanish Ministry of Economy and Competitivity.

Appendix A. Mathematical background and auxiliary results

We need some background for a self-contained exposition and a neater illustration of the proofs coming subsequently. A spherical harmonic of degree n for $\mathbb{S}^d$ is the restriction to $\mathbb{S}^d$ of a real-valued harmonic homogeneous polynomial in $\mathbb{R}^{d+1}$ of degree n. Together with the zero function, the spherical harmonics of degree n form a finite dimensional vector space denoted $\mathcal{H}_n(\mathbb{S}^d)$. It is a subspace of the space of continuous functions on $\mathbb{S}^d$ [4]. For every $n \in \mathbb{N} \cup \{0\}$, we have

$$h(n, d+1) = \dim\{\mathcal{H}_n(\mathbb{S}^d)\} = \frac{(2n+d-1)\,(n+d-2)!}{n!\,(d-1)!},$$

with $h(0, d+1) = 1$ for all $d \geq 1$, and where $h(n, d+1) = \mathcal{O}(n^{d-1})$. The spaces $\mathcal{H}_n(\mathbb{S}^d)$ are mutually orthogonal subspaces of the Hilbert space $\mathcal{L}^2(\mathbb{S}^d)$ of real-valued squared-integrable functions on the sphere $\mathbb{S}^d$, endowed with the inner product

$$\langle f, g\rangle_{\mathcal{L}^2(\mathbb{S}^d)} = \frac{1}{\|\mathbb{S}^d\|} \int_{\mathbb{S}^d} f(\mathbf{x}) g(\mathbf{x}) \mathrm{d}\omega_d(\mathbf{x}),$$

for every $f, g \in \mathcal{L}^2(\mathbb{S}^d)$ and $\mathbf{x} \in \mathbb{S}^d$. Here, $\|\mathbb{S}^d\|$ denotes the total mass of the surface measure $\mathrm{d}\omega_d$ on $\mathbb{S}^d$, defined as

$$\|\mathbb{S}^d\| = \int_{\mathbb{S}^d} \mathrm{d}\omega_d = \frac{2\,\pi^{(d+1)/2}}{\Gamma\{(d+1)/2\}}.$$

Not only that, but the space $\mathcal{L}^2(\mathbb{S}^d)$ is generated by the spaces $\mathcal{H}_n(\mathbb{S}^d)$: any $f \in \mathcal{L}^2(\mathbb{S}^d)$ has an orthogonal expansion

$$f = \sum_{n=0}^{\infty} f_n,$$

where $f_n \in \mathcal{H}_n(\mathbb{S}^d)$. Subsequently we denote $\{Y_{n\ell} : \ n \in \mathbb{N} \cup \{0\},\ 1 \leq \ell \leq h(n, d+1)\}$ an orthonormal basis of $\mathcal{H}_n(\mathbb{S}^d)$, i.e.,

$$\langle Y_{n\ell}, Y_{n\ell'}\rangle_{\mathcal{L}^2(\mathbb{S}^d)} = \mathbf{1}_{\{\ell=\ell'\}}, \quad 1 \leq \ell, \ell' \leq h(n, d+1).$$

For $i \in \{1, 2\}$, let $\{Z_i(\mathbf{x}) : \mathbf{x} \in \mathbb{S}^d\}$ be two zero mean Gaussian random fields with geodesically isotropic covariance functions $\psi_i(\theta)$. Following [47], we have that for each $i \in \{1, 2\}$, Z_i induces, on the space $\mathcal{L}^2(\mathbb{S}^d)$, a Gaussian measure with covariance operators $\mathcal{C}_i$ defined, for all $f \in \mathcal{L}^2(\mathbb{S}^d)$, by

$$\mathcal{C}_i\big(f\big)(\mathbf{x}) = \int_{\mathbb{S}^d} \psi_i\{\theta(\mathbf{x}, \mathbf{y})\} f(\mathbf{y}) \mathrm{d}\omega_d(\mathbf{y}). \tag{A.1}$$

Let μ_1, μ_2 be two Gaussian measures, defined on the Borel σ-algebra $\mathcal{B}(\mathbb{S}^d)$, with corresponding covariance operators $\mathcal{C}_i$. An operator $D : \mathcal{L}^2(\mathbb{S}^d) \to \mathcal{L}^2(\mathbb{S}^d)$ is called a Hilbert–Schmidt operator if it is bounded and with finite norm $\|\cdot\|_{\mathcal{L}^2(\mathbb{S}^d)}$ defined by

$$\|D\|^2_{\mathcal{L}^2(\mathbb{S}^d)} = \sum_k \|De_k\|^2_{\mathcal{L}^2(\mathbb{S}^d)},$$

with (e_k) an orthonormal basis, and with $\|f\|_{\mathcal{L}^2(\mathbb{S}^d)} = \langle f, f\rangle_{\mathcal{L}^2(\mathbb{S}^d)}$.

Two important ingredients are needed to prove our results. We state them formally to have a self contained exposition. The former is known as the Feldman–Hájek theorem [17], and the latter can be found as Theorem 1 in Chapter 3 of [30]. Let us first introduce the entropy distance R [30], defined by

$$R = -\mathrm{E}_1(\ln L) + \mathrm{E}_2(\ln L), \tag{A.2}$$

where $L = \mathrm{d}\mu_2/\mathrm{d}\mu_1$ is the likelihood ratio between the two measures defined on $\mathbb{S}^d$. Here, E_i denotes mathematical expectation with respect to the measure μ_i, with $i \in \{1, 2\}$. Details on how to compute L will be given later.

Theorem 2 (*Feldman–Hájek Theorem*). *Let μ_1, μ_2 be two zero mean Gaussian measures defined on the Borel σ-algebra $\mathcal{B}(\mathbb{S}^d)$, with corresponding covariance operators $\mathcal{C}_i$ as defined in Eq. (A.1). Let I be the identity operator. Then, μ_1 and μ_2 are equivalent* *if and only if the operator*

$$D = \mathcal{C}_2^{-1/2}\mathcal{C}_1\mathcal{C}_2^{-1/2} \tag{A.3}$$

is a positive definite, invertible, and bounded, with $D - I$ a Hilbert–Schmidt operator. Moreover, the Radon–Nikodym derivative of μ_2 with respect to μ_1 is given by

$$\frac{\mathrm{d}\mu_2}{\mathrm{d}\mu_1} = \prod_{n=0}^{\infty}\prod_{\ell=1}^{h(n,d+1)} \sqrt{1+\lambda_n^\ell}\exp\left\{\frac{-\lambda_n^\ell}{2(1+\lambda_n^\ell)}\langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), e_n^\ell\rangle^2_{\mathcal{L}^2(\mathbb{S}^d)}\right\},$$

where e_n^ℓ are the eigenvectors of $D - I$, and λ_n^ℓ are their corresponding eigenvalues.

Theorem 3 (*Equivalence and Entropy*)**.** *Gaussian measures are either equivalent or orthogonal. Two Gaussian measures are equivalent if and only if their entropy distance R, defined at (A.2), is finite.*

Appendix B. Proofs of the results in Section 2

Proof of Theorem 1. We first provide a sketch of the proof, with the details discussed subsequently.

Necessity. For this part, we apply Theorem 3. Thus, we need to proceed as follows.

1. Compute the entropy distance R as in Eq. (A.2) between the two Gaussian measures on $(\mathbb{S}^d, \mathcal{B}(\mathbb{S}^d))$.
2. Show that R is a function of the respective d-Schoenberg coefficients $b_{n,d,1}$, $ib_{n,d,2}$, as defined in Eq. (4).
3. Sum up the two previous point and obtain condition (6).

Sufficiency. For sufficiency, we use Theorem 2 and show that the operator D defined through (A.3) is a function of the ratio $b_{n,d,1}/b_{n,d,2}$. The proof is completed by showing that $D - I$ is a Hilbert–Schmidt operator.

We now provide the details.

Necessity. Let μ_1 and μ_2 be two equivalent Gaussian measures and let L be the likelihood ratio defined around Eq. (A.2). Direct inspection, see also [53], shows that

$$\begin{aligned} L &= \prod_{n=0}^{\infty}\prod_{\ell=1}^{h(n,d+1)} \sqrt{1+\lambda_n^\ell}\exp\left\{\frac{-\lambda_n^\ell}{2(1+\lambda_n^\ell)}\langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), e_n^\ell\rangle^2_{\mathcal{L}^2(\mathbb{S}^d)}\right\} \\ &= \exp\left[-\frac{1}{2}\sum_{n=0}^{\infty}\sum_{\ell=1}^{h(n,d+1)}\left\{\langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), e_n^\ell\rangle^2_{\mathcal{L}^2(\mathbb{S}^d)}\frac{\lambda_n^\ell}{(1+\lambda_n^\ell)} - \ln(1+\lambda_n^\ell)\right\}\right], \end{aligned}$$

where $\mathcal{C}_2$ is the covariance operator, and where, according to Theorem 2, the e_n^ℓs and λ_n^ℓs are, respectively, the eigenvectors and eigenvalues associated to the operator $D - I$, with D defined through Eq. (A.3)), and I the identity operator. The log-likelihood is defined by

$$\ln L = -\frac{1}{2}\sum_{n=0}^{\infty}\sum_{\ell=1}^{h(n,d+1)}\left\{\langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), e_n^\ell\rangle^2_{\mathcal{L}^2(\mathbb{S}^d)}\frac{\lambda_n^\ell}{(1+\lambda_n^\ell)} - \ln(1+\lambda_n^\ell)\right\}. \tag{A.4}$$

To find a closed form of the entropy distance R in (A.2), we need to calculate the expectation of (A.4) with respect to μ_1 and μ_2. By assumption, μ_1 and μ_2 are equivalent. Thus, we invoke arguments in Theorem 1 of Skorokhod et al. [47] to assume that

$$\mathrm{E}_i\left\{\langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), e_n^\ell\rangle^2_{\mathcal{L}^2(\mathbb{S}^d)}\right\} = \begin{cases} 1 & \text{for } i = 1, \\ 1 + c_n & \text{for } i = 2, \end{cases}$$

where $c_n \geq 0$ and $\sum c_n^2 < \infty$. Hence, we conclude that the convergence of the series in (A.4) can be evaluated on the basis of the following argument:

$$\begin{aligned} \mathrm{E}_1\left\{\langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), e_n^\ell\rangle^2_{\mathcal{L}^2(\mathbb{S}^d)}\frac{\lambda_n^\ell}{(1+\lambda_n^\ell)} - \ln(1+\lambda_n^\ell)\right\} &= \frac{\lambda_n^\ell}{(1+\lambda_n^\ell)} - \ln(1+\lambda_n^\ell) = \mathcal{O}\{(\lambda_n^\ell)^2\}, \\ \mathrm{E}_2\left\{\langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), e_n^\ell\rangle^2_{\mathcal{L}^2(\mathbb{S}^d)}\frac{\lambda_n^\ell}{(1+\lambda_n^\ell)} - \ln(1+\lambda_n^\ell)\right\} &= (1+c_n)\frac{\lambda_n^\ell}{(1+\lambda_n^\ell)} - \ln(1+\lambda_n^\ell) = \mathcal{O}\{(\lambda_n^\ell)^2\}. \end{aligned}$$

Summing up, we have just shown that

$$R = \sum_{n=0}^{\infty}\sum_{\ell=1}^{h(n,d+1)} \mathcal{O}\{(\lambda_n^\ell)^2\}.$$

By unique decomposition of positive definite operators [17], and using the fact that the operators $\mathcal{C}_1$ and $\mathcal{C}_2$ are diagonal, we obtain $\lambda_n^\ell = b_{n,d,1}/b_{n,d,2} - 1$. We can now invoke Theorem 2 to obtain condition (6).

Sufficiency. We now assume that (6) holds for two Gaussian measures μ_1, μ_2, with corresponding geodesically isotropic covariance functions ψ_1, ψ_2 and covariance operators $\mathcal{C}_1, \mathcal{C}_2 : \mathcal{L}^2(\mathbb{S}^d) \to \mathcal{L}^2(\mathbb{S}^d)$. Let D be the operator defined by Eq. (A.3). Using the fact that the operators $\mathcal{C}_1, \mathcal{C}_2$ are diagonal and positive definite, we have that D is also positive definite, with eigenvalues system $\{\widetilde{\lambda}_n^\ell\}$ of strictly positive eigenvalues. Also, note that $\widetilde{\lambda}_n^\ell$ and λ_n^ℓ are obviously related and we have, for all ℓ and $n \in \mathbb{N} \cup \{0\}$, $\widetilde{\lambda}_n^\ell = b_{n,d,1}/b_{n,d,2}$. Furthermore, we have that the operator $D - I$ is diagonal, which in turn implies that its eigenvalues are strictly larger than -1. By (6), we conclude that $D - I$ is a Hilbert–Schmidt operator. By invoking Theorem 2, the proof is completed. □

Proof of Proposition 1. We work on the sphere $\mathbb{S}^2$, so that $h(n, 3) = 2n + 1$. Using the first entry in Table 1, we have that for each $i \in \{1, 2\}$, the Schoenberg coefficient $b_{n,2,i}$ associated to $\psi_i(\theta) = \psi_M(\theta; \delta_i, 1/2, \sigma_i)$ can be written as $b_{n,2,i} = \sigma_i^2 \delta_i^n (1-\delta_i)$. By Theorem 1, to prove the orthogonality of μ_1 and μ_2 we need to show that the series in (6) is always divergent. For our case, the series (6) is of the form

$$\sum_{n=0}^{\infty} h(n, 3)\left(\frac{b_{n,2,1}}{b_{n,2,2}} - 1\right)^2 = \sum_{n=0}^{\infty}(2n + 1)\left\{\left(\frac{\sigma_1}{\sigma_2}\right)^2\left(\frac{\delta_1}{\delta_2}\right)^n\left(\frac{1-\delta_1}{1-\delta_2}\right) - 1\right\}^2. \tag{A.5}$$

To check the convergence or divergence of this series, we use Raabe's test as follows. Set

$$a_n = (2n + 1)\left\{\left(\frac{\sigma_1}{\sigma_2}\right)^2\left(\frac{\delta_1}{\delta_2}\right)^n\left(\frac{1-\delta_1}{1-\delta_2}\right) - 1\right\}^2, \quad a_{n+1} = (2n + 3)\left\{\left(\frac{\sigma_1}{\sigma_2}\right)^2\left(\frac{\delta_1}{\delta_2}\right)^{n+1}\left(\frac{1-\delta_1}{1-\delta_2}\right) - 1\right\}^2,$$

and compute $n\,(a_n/a_{n+1} - 1)$. To find the limit of this expression, we consider two different cases. If $\delta_1 < \delta_2$, then $(\delta_1/\delta_2)^n$ and $(\delta_1/\delta_2)^{n+1}$ tend to zero as $n \to \infty$. Hence,

$$\lim_{n\to\infty} n\left(\frac{a_n}{a_{n+1}} - 1\right) = \lim_{n\to\infty} n\left(\frac{2n+1}{2n+3} - 1\right) = \lim_{n\to\infty}\left(\frac{-2n}{2n+3}\right) = -1 < 1.$$

Then Raabe's test yields divergence, meaning that the Gaussian measures μ_1 and μ_2 are orthogonal. If $\delta_1 > \delta_2$, then $(\delta_1/\delta_2)^{n+1}$ tends to infinity faster than $(\delta_1/\delta_2)^n$ as $n \to \infty$. Hence, $n\,(a_n/a_{n+1} - 1) \to -\infty$. Then the series (A.5) is always divergent. The proof is complete. □

Proof of Proposition 2. We now work on the circle $\mathbb{S}^1$ and assume covariances $\psi_i(\theta) = \psi_S(\theta; \alpha_i, \sigma_i)$ for $i \in \{1, 2\}$. Using the second entry in Table 1, we have, for each $i \in \{1, 2\}$,

$$b_{n,1,i} = \frac{\sigma_i^2}{\sqrt{2}}\,\Gamma_{n+1}\left(\frac{\alpha_i}{2}\right), \quad \Gamma_{n+1}\left(\frac{\alpha_i}{2}\right) = \frac{-1}{(n+1)!}\prod_{m=0}^{n}\left(m - \frac{\alpha_i}{2}\right).$$

According to Theorem 1, we need to check the convergence of the series in (6) to determine if the given Gaussian measures μ_1 and μ_2 are equivalent or orthogonal. So, for our case the series in (6) has expression

$$\sum_{n=0}^{\infty} h(n, 2)\left(\frac{b_{n,1,1}}{b_{n,1,2}} - 1\right)^2 = \sum_{n=0}^{\infty} 2\left\{\Lambda\prod_{m=0}^{n}\left(\frac{2m-\alpha_1}{2m-\alpha_2}\right) - 1\right\}^2, \tag{A.6}$$

where $\Lambda = (\sigma_1/\sigma_2)^2$. We use again Raabe's test to check the convergence of (A.6). Set

$$a_n = \left\{\Lambda\prod_{m=0}^{n}\left(\frac{2m-\alpha_1}{2m-\alpha_2}\right) - 1\right\}^2, \quad a_{n+1} = \left\{\Lambda\prod_{m=0}^{n+1}\left(\frac{2m-\alpha_1}{2m-\alpha_2}\right) - 1\right\}^2$$

and compute $n\,(a_n/a_{n+1} - 1)$. To find the limit of this expression, we consider two cases. If $\alpha_1 > \alpha_2$, then $(2m - \alpha_1)/(2m - \alpha_2) < 1$ and hence $n\,(a_n/a_{n+1} - 1) \to 0 < 1$. Thus, the series (A.6) is divergent. If $\alpha_1 < \alpha_2$, then $(2m - \alpha_2)/(2m - \alpha_1) < 1$. Moreover, we have

$$\begin{aligned}\lim_{n\to\infty} n\left(\frac{a_n}{a_{n+1}} - 1\right) &= \lim_{n\to\infty} n\left[\left\{\frac{\Lambda - \prod_{m=0}^{n}\left(\frac{2m-\alpha_2}{2m-\alpha_1}\right)}{\Lambda\frac{2n+1-\alpha_1}{2n+1-\alpha_2} - \prod_{m=0}^{n}\left(\frac{2m-\alpha_2}{2m-\alpha_1}\right)}\right\}^2 - 1\right] \\ &= \lim_{n\to\infty} n\left\{\left(\frac{2n+1-\alpha_2}{2n+1-\alpha_1}\right)^2 - 1\right\} = \alpha_1 - \alpha_2 < 1.\end{aligned}$$

Thus, the series (A.6) is divergent. Since the series is always divergent, we get that the Gaussian measures μ_1 and μ_2 are always orthogonal. □

Appendix C. Generalization to two-point homogeneous spaces

We now inspect necessary and sufficient conditions for equivalence of Gaussian measures on two-point homogeneous spaces (Riemannian symmetric spaces of rank one). A compact two-point homogeneous spaces of dimension d will be denoted by $\mathcal{M}^d$. Following [20], two-point homogeneous spaces $\mathcal{M}^d$ include, as special case, the following spaces:

(I) $\mathbb{S}^d$, for $d \in \{1, 2, 3, \ldots\}$;
(II) the real projective spaces $\mathbb{P}^d(\mathbb{R})$, for $d \in \{2, 3, 4, \ldots\}$;
(III) the complex projective spaces $\mathbb{P}^d(\mathbb{C})$, for $d \in \{4, 6, 8 \ldots\}$;
(IV) the quaternion projective spaces $\mathbb{P}^d(\mathbb{H})$, for $d \in \{8, 12, 16 \ldots\}$;
(V) the Cayley projective plane $\mathbb{P}^{16}(\mathrm{Cay})$.

The geodesic distance over $\mathcal{M}^d$ is the mapping $\theta_{\mathcal{M}^d} : \mathcal{M}^d \times \mathcal{M}^d \to [0, \pi]$ defined, for all $\mathbf{x}, \mathbf{y} \in \mathcal{M}^d$, as in [8] by $\theta_{\mathcal{M}^d} = \arccos(\langle \mathbf{x}, \mathbf{y} \rangle)$. According to [20], any isotropic positive definite function $\widehat{C}(\mathbf{x}, \mathbf{y}) = \widehat{\psi}\{\theta_{\mathcal{M}^d}(\mathbf{x}, \mathbf{y})\}$ on $\mathcal{M}^d$ has the spectral representation based on the Jacobi polynomials $P_n^{\alpha,\beta}$ [18] of degree $n \in \mathbb{N}$, $\alpha, \beta > -1$ with $\alpha = (d-2)/2$ as follows

$$\widehat{\psi}(\theta_{\mathcal{M}^d}) = \sum_{n=0}^{\infty} \Upsilon_n^{(d-2)/2,\beta} \frac{P_n^{(d-2)/2,\beta}(\cos \theta_{\mathcal{M}^d})}{P_n^{(d-2)/2,\beta}(1)}, \tag{9}$$

where $\{\Upsilon_n^{(d-2)/2,\beta}\}_{n=0}^{\infty}$ are non-negative coefficients and $\sum_n \Upsilon_n^{(d-2)/2,\beta} < \infty$; see [20]. For any fixed $n \in \mathbb{N}$, the coefficient $\Upsilon_n^{(d-2)/2,\beta}$ in the Jacobi expansion (9) of a positive definite function has the following explicit formula

$$\Upsilon_n^{(d-2)/2,\beta} = \frac{(2n+\beta+d/2)\, \Gamma(n+\beta+d/2)}{2^{\beta+d/2}\, \Gamma(n+\beta+1)\, \Gamma(d/2)} \times \int_0^{\pi} P_n^{(d-2)/2,\beta}(\cos \theta_{\mathcal{M}^d}) \widehat{\psi}(\theta_{\mathcal{M}^d}) \sin^{d-1} \theta_{\mathcal{M}^d} (1+\cos \theta_{\mathcal{M}^d})^{\beta-(d-2)/2} \mathrm{d}\theta_{\mathcal{M}^d}, \tag{10}$$

where

$$P_n^{(d-2)/2,\beta}(1) = \frac{\Gamma(n+d/2)}{n!\, \Gamma(d/2)}.$$

Also, in the spectral representation given by (9), $\beta = (d-2)/2, -1/2, 0, 1, 3$, for $\mathbb{S}^d$, $\mathbb{P}^d(\mathbb{R})$, $\mathbb{P}^d(\mathbb{C})$, $\mathbb{P}^d(\mathbb{H})$ and $\mathbb{P}^{16}(\mathrm{Cay})$, respectively.

Using the ingredients above and mimicking the proof of Theorem 1, it becomes fairly easy to prove the following.

Theorem 4. *Let μ_1, μ_2 be two zero mean Gaussian random measures on $\mathcal{M}^d$ with corresponding coefficients $\Upsilon_{n,1}^{(d-2)/2,\beta}$, $\Upsilon_{n,2}^{(d-2)/2,\beta}$ defined according to representation (9). Then, μ_1 and μ_2 are equivalent if and only if*

$$\sum_{n=0}^{\infty} \hbar(n, d, \beta) \left(\frac{\Upsilon_{n,1}^{(d-2)/2,\beta}}{\Upsilon_{n,2}^{(d-2)/2,\beta}} - 1 \right)^2 < \infty, \tag{11}$$

where $\hbar(n, d, \beta)$ denotes the dimension of the Laplacian eigenspace related to the eigenvalues $-n(n+d/2+\beta)$, and given by (see [8])

$$\hbar(n, d, \beta) = \frac{(2n+\beta+d/2)\, \Gamma(n+d/2)\, \Gamma(\beta+1)\, \Gamma(n+\beta+d/2)}{\Gamma(d/2)\, \Gamma(n+1)\, \Gamma(n+\beta+1)\, \Gamma(d/2+\beta+1)}.$$

We observe that Theorem 4 agrees with Theorem 1 when $\beta = (d-2)/2$. The proof of Theorem 4 comes exactly from the same arguments as in Theorem 1. We report it here for the sake of completeness.

Proof. *Necessity*. Let μ_1 and μ_2 be two equivalent Gaussian measures on $\mathcal{M}^d$. The likelihood ratio between μ_1 and μ_2 for that case is given by

$$\begin{aligned} L &= \prod_{n=0}^{\infty} \prod_{\ell=1}^{\hbar(n,d,\beta)} \sqrt{1+\hat{\lambda}_n^{\ell}} \exp\left\{ \frac{-\hat{\lambda}_n^{\ell}}{2(1+\hat{\lambda}_n^{\ell})} \langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), \hat{e}_n^{\ell} \rangle^2_{\mathcal{L}^2(\mathcal{M}^d)} \right\} \\ &= \exp\left[-\frac{1}{2} \sum_{n=0}^{\infty} \sum_{\ell=1}^{\hbar(n,d,\beta)} \left\{ \langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), \hat{e}_n^{\ell} \rangle^2_{\mathcal{L}^2(\mathcal{M}^d)} \frac{\hat{\lambda}_n^{\ell}}{(1+\hat{\lambda}_n^{\ell})} - \ln(1+\hat{\lambda}_n^{\ell}) \right\} \right], \end{aligned} \tag{12}$$

where $\mathcal{C}_2$ is the covariance operator, and where, according to Theorem 2, the e_n^{ℓ}s and $\hat{\lambda}_n^{\ell}$s are, respectively, the eigenvectors and eigenvalues associated to the operator $D - I$, with D defined through Eq. (A.3) and I the identity operator. The log-likelihood is defined by

$$\ln L = -\frac{1}{2} \sum_{n=0}^{\infty} \sum_{\ell=1}^{\hbar(n,d,\beta)} \left\{ \langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), \hat{e}_n^{\ell} \rangle^2_{\mathcal{L}^2(\mathcal{M}^d)} \frac{\hat{\lambda}_n^{\ell}}{(1+\hat{\lambda}_n^{\ell})} - \ln(1+\hat{\lambda}_n^{\ell}) \right\}. \tag{13}$$

To find a closed form of the entropy distance R in (A.2), we need to calculate the expectation of (13) with respect to μ_1 and μ_2. By assumption, μ_1 and μ_2 are equivalent. Thus, we invoke arguments in Theorem 1 of Skorokhod et al. [47] to assume that

$$\mathrm{E}_i\left(\langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), \hat{e}_n^{\ell}\rangle^2_{\mathcal{L}^2(\mathcal{M}^d)}\right) = \begin{cases} 1 & \text{for } i = 1, \\ 1 + \hat{c}_n & \text{for } i = 2, \end{cases}$$

where $\hat{c}_n \geq 0$ and $\sum \hat{c}_n^2 < \infty$. Hence, we conclude that the convergence of the series in (13) can be evaluated on the basis of the following argument

$$\mathrm{E}_1\left\{\langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), \hat{e}_n^{\ell}\rangle^2_{\mathcal{L}^2(\mathcal{M}^d)} \frac{\hat{\lambda}_n^{\ell}}{(1+\hat{\lambda}_n^{\ell})} - \ln(1+\hat{\lambda}_n^{\ell})\right\} = \frac{\hat{\lambda}_n^{\ell}}{(1+\hat{\lambda}_n^{\ell})} - \ln(1+\hat{\lambda}_n^{\ell}) = \mathcal{O}\{(\hat{\lambda}_n^{\ell})^2\},$$

$$\mathrm{E}_2\left\{\langle \mathcal{C}_2^{-1/2} f(\mathbf{x}), \hat{e}_n^{\ell}\rangle^2_{\mathcal{L}^2(\mathcal{M}^d)} \frac{\hat{\lambda}_n^{\ell}}{(1+\hat{\lambda}_n^{\ell})} - \ln(1+\hat{\lambda}_n^{\ell})\right\} = (1+\hat{c}_n)\frac{\hat{\lambda}_n^{\ell}}{(1+\hat{\lambda}_n^{\ell})} - \ln(1+\hat{\lambda}_n^{\ell}) = \mathcal{O}\{(\hat{\lambda}_n^{\ell})^2\}.$$

Summing up, we have just shown that

$$R = \sum_{n=0}^{\infty} \sum_{\ell=1}^{\hbar(n,d,\beta)} \mathcal{O}\{(\hat{\lambda}_n^{\ell})^2\}.$$

By unique decomposition of positive definite operators [17], and using the fact that the operators $\mathcal{C}_1$ and $\mathcal{C}_2$ are diagonal, we obtain

$$\hat{\lambda}_n^{\ell} = \Upsilon_{n,1}^{(d-2)/2,\beta} / \Upsilon_{n,2}^{(d-2)/2,\beta} - 1.$$

We can now invoke Theorem 2 to obtain condition (11).

Sufficiency. We now assume that (11) holds for two Gaussian measures μ_1, μ_2 with corresponding geodesically isotropic covariance function $\widehat{\psi}_1, \widehat{\psi}_2$ and covariance operators $\mathcal{C}_1, \mathcal{C}_2 : \mathcal{L}^2(\mathcal{M}^d) \to \mathcal{L}^2(\mathcal{M}^d)$. Let D be the operator defined by Eq. (A.3). Using the fact that the operators $\mathcal{C}_i$ are diagonal and positive definite, we have that D is also positive definite, with eigenvalues system $\{\hat{\lambda}_n^{\ell}\}$ of strictly positive eigenvalues. Also, note that for all ℓ and $n \in \mathbb{N} \cup \{0\}$,

$$\hat{\lambda}_n^{\ell} = \Upsilon_{n,1}^{(d-2)/2,\beta} / \Upsilon_{n,2}^{(d-2)/2,\beta}.$$

Furthermore, the operator $D - I$ is diagonal, which in turn implies that its eigenvalues are strictly larger than -1. By (11), we conclude that $D - I$ is a Hilbert–Schmidt operator. By invoking Theorem 2, the proof is completed. □

References

[1] E. Anderes, On the consistent separation of scale and variance for Gaussian random fields, Ann. Statist. 38 (2010) 870–893.
[2] P. Baldi, D. Marinucci, Some characterizations of the spherical harmonics coefficients for isotropic random fields, Statist. Probab. Lett. 77 (2006) 490–496.
[3] V.S. Barbosa, V.A. Menegatto, Strict positive definiteness on products of compact two-point homogeneous spaces, Integral Transforms Spec. Funct. 28 (2017) 56–73.
[4] C. Berg, E. Porcu, From Schoenberg coefficients to Schoenberg functions, Constr. Approx. 45 (2017) 217–241.
[5] M. Bevilacqua, T. Faouzi, R. Furrer, E. Porcu, Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics, Ann. Statist. (2018).
[6] N.H. Bingham, Positive definite functions on spheres, Proc. Cambridge Philos. Soc. 73 (1973) 145–156.
[7] V.I. Bogachev, Gaussian Measures, American Mathematical Society, Providence, RI, 1998.
[8] B. Bordin, A.K. Kushpel, J. Levesley, S.A. Tozoni, Estimates of n-widths of Sobolev's classes on compact globally symmetric spaces of rank one, J. Funct. Anal. 202 (2003) 307–326.
[9] S. Castruccio, M.G. Genton, Beyond axial symmetry: An improved class of models for global data, Statistics 3 (2014) 48–55.
[10] S. Castruccio, J. Guinness, An evolutionary spectrum approach to incorporate large-scale geographical descriptors on global processes, J. R. Stat. Soc. Ser. C Appl. Stat. 66 (2017) 329–344.
[11] S. Castruccio, E. Porcu, A Bridge between Equivalence of Gaussian Measures and SPDE Representations, Technical Report, Universidad Técnica Federico Santa María, Valparaíso, Chile, 2017.
[12] S. Castruccio, M.L. Stein, Global space–time models for climate ensembles, Ann. Appl. Stat. 7 (2013) 1593–1611.
[13] R. Cavoretto, A. De Rossi, Fast and accurate interpolation of large scattered data sets on the sphere, J. Comput. Appl. Math. 234 (2010) 1505–1521.
[14] S.D. Chatterji, V. Mandrekar, Equivalence and Singularity of Gaussian Measures and Applications, Technical Report, Academic Press, 1978.
[15] G. Christakos, On certain classes of spatiotemporal random fields with application to space–time data processing, Syst. Man Cybern. 21 (1991) 861–875.
[16] J. Clarke, A. Alegría, E. Porcu, Regularity properties and simulations of Gaussian random fields on the sphere cross time, Electron. J. Stat. 12 (2018) 399–426.
[17] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, Cambridge University Press, New York, 1992.
[18] F. Dai, Y. Xu, Approximation Theory and Harmonic Analysis on Spheres and Balls, Springer, New York, 2013.
[19] D.J. Daley, E. Porcu, Dimension walks and Schoenberg spectral measures, Proc. Amer. Math. Soc. 141 (2013) 1813–1824.
[20] R. Gangolli, Positive definite kernels on homogeneous spaces and certain stochastic processes related to Lévy's Brownian motion of several parameters, Ann. Inst. H. Poincaré 3 (1967) 121–226.
[21] I.I. Gikhman, A.V. Skorokhod, On the densities of probability measures in function spaces, Russian Math. Surveys 21 (1966) 83–156.
[22] T. Gneiting, Nonseparable, stationary covariance functions for space–time data, J. Amer. Statist. Assoc. 97 (2002) 590–600.

[23] T. Gneiting, Strictly and non-strictly positive definite functions on spheres, Bernoulli 19 (2013) 1327–1349.
[24] T. Gneiting, Strictly and non-strictly positive definite functions on spheres: Online supplement, 2013. Available at https://projecteuclid.org/download/suppdf_1/euclid.bj/1377612854.
[25] I.S. Gradshteyn, I.M. Ryzhik, Tables of Integrals, Series, and Products, seventh ed., Academic Press, Amsterdam, 2007.
[26] J.C. Guella, V.A. Menegatto, A.P. Peron, Strictly positive definite kernels on a product of spheres II, SIGMA 12 (2016) 15.
[27] L.V. Hansen, T.L. Thorarinsdottir, E. Ovcharov, T. Gneiting, Gaussian random particles with flexible Hausdorff dimension, Adv. Appl. Probab. 47 (2015) 307–327.
[28] M. Hitczenko, M.L. Stein, Some theory for anisotropic processes on the sphere, Stat. Methodol. 9 (2012) 211–227.
[29] C. Huang, H. Zhang, S. Robeson, A simplified representation of the covariance structure of axially symmetric processes on the sphere, Statist. Probab. Lett. 82 (2012) 1346–1351.
[30] I.A. Ibragimov, Y.A. Rozanov, Gaussian Random Processes, Springer, New York, 1978.
[31] K. Inoue, Equivalence of measures for some class of Gaussian random fields, J. Multivariate Anal. 6 (1976) 295–308.
[32] J. Istas, Spherical and hyperbolic fractional Brownian motion, Electron. Commun. Probab. 10 (2005) 254–262.
[33] R.H. Jones, Stochastic processes on a sphere, Ann. Math. Stat. 34 (1963) 213–218.
[34] T. Kühn, F. Liese, A short proof of the Hájek–Feldman theorem, Theory Probab. Appl. 23 (1979) 429–431.
[35] H. Kuo, Gaussian Measures in Banach Spaces, Springer, New York, 1975.
[36] A. Lang, C. Schwab, Isotropic random fields on the sphere: Regularity, fast simulation and stochastic partial differential equations, Ann. Appl. Probab. 25 (2013) 3047–3094.
[37] N. Leonenko, L. Sakhno, On spectral representation of tensor random fields on the sphere, Stoch. Anal. Appl. 31 (2012) 167–182.
[38] A. Malyarenko, Invariant Random Fields on Spaces with a Group Action, Springer, New York, 2013.
[39] D. Marinucci, G. Peccati, Random Fields on the Sphere, Representation, Limit Theorems and Cosmological Applications, Cambridge, New York, 2011.
[40] V.A. Menegatto, Strictly positive definite kernels on the Hilbert sphere, Appl. Anal. 55 (1994) 91–101.
[41] V.A. Menegatto, Strictly positive definite kernels on the circle, Rocky Mountain J. Math. 25 (1995) 1149–1163.
[42] J. Møller, M. Nielsen, E. Porcu, E. Rubak, Determinantal point process models on the sphere, Bernoulli 24 (2018) 1171–1201.
[43] E. Porcu, A. Alegría, R. Furrer, Modeling temporally evolving and spatially globally dependent data, 2017. arXiv preprint arXiv:1706.09233.
[44] E. Porcu, M. Bevilacqua, M.G. Genton, Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere, J. Amer. Statist. Assoc. 111 (2016) 888–898.
[45] H. Sato, On the equivalence of Gaussian measures, J. Math. Soc. Japan 19 (1967) 159–172.
[46] I.J. Schoenberg, Positive definite functions on spheres, Duke Math. J. 9 (1942) 96–108.
[47] A.V. Skorokhod, M.I. Yadrenko, On absolute continuity of measures corresponding to homogeneous Gaussian fields, Theory Probab. Appl. 8 (1973) 27–40.
[48] S.D. Sokolova, On the equivalence of Gaussian measures corresponding to the solutions of stochastic differential equations, Teor. Veroyatn. Primen. 28 (1983) 429–433.
[49] S. Soubeyrand, J. Enjalbert, I. Sache, Accounting for roughness of circular processes: Using Gaussian random processes to model the anisotropic spread of airborne plant disease, Theor. Popul. Biol. 73 (2008) 92–103.
[50] M.L. Stein, Asymptotically efficient prediction of a random field with a misspecified covariance function, Ann. Statist. 16 (1988) 55–63.
[51] M.L. Stein, Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure, Ann. Statist. 18 (1990) 850–872.
[52] M.L. Stein, A simple condition for asymptotic optimality of linear predictions of random fields, Statist. Probab. Lett. 17 (1993) 399–404.
[53] M.L. Stein, Statistical Interpolation of Spatial Data: Some Theory for Kriging, Springer, New York, 1999.
[54] M.L. Stein, Equivalence of Gaussian measures for some nonstationary random fields, J. Statist. Plann. Inference 123 (2004) 1–11.
[55] N. Vakhania, V. Tarieladze, On singularity and equivalence of Gaussian measures, in: Real and Stochastic Analysis: Recent Advances, CRC Press, Boca Raton, FL, 1997.
[56] D. Varberg, On equivalence of Gaussian measures, Pacific J. Math. 11 (1961) 751–762.
[57] H. Zhang, Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics, J. Amer. Statist. Assoc. 99 (2004) 250–261.