



CENTRE FOR **STOCHASTIC GEOMETRY**  
AND ADVANCED **BIOIMAGING**



Alfredo Alegria, Francisco Cuevas, Peter Diggle and Emilio Porcu

## **A Family of Covariance Functions for Random Fields on Spheres**

No. 08, August 2018

# A Family of Covariance Functions for Random Fields on Spheres

Alfredo Alegria<sup>1</sup>, Francisco Cuevas<sup>2,1</sup>, Peter Diggle<sup>3</sup>  
and Emilio Porcu<sup>\*,4,5</sup>

<sup>1</sup>Department of Mathematics, Technical University Federico Santa María

<sup>2</sup>Department of Mathematical Sciences, Aalborg University

<sup>3</sup>CHICAS, Lancaster Medical School, Lancaster University

<sup>4</sup>School Mathematics, Statistics, and Physics, Newcastle University

<sup>5</sup>Department of Mathematics, University of Atacama

## Abstract

The Matérn family of isotropic covariance functions has been central to the theoretical development and application of statistical models for geospatial data. For global data defined over the whole sphere representing planet Earth, the natural definition of the distance between two locations is the great-circle distance. In this setting, the Matérn family is no longer valid, and finding a suitable analogue for modelling data on the sphere has for some time been an open problem.

This paper proposes a new family of isotropic covariance functions for random fields defined on the sphere. The family has four parameters, one of which indexes the mean square differentiability of the corresponding Gaussian field. The new family also allows for any admissible range of fractal dimension.

We describe a simulation to show the behaviour of the maximum likelihood parameter estimation under fixed domain asymptotics, this being the relevant asymptotic regime for sampling a closed set. As expected, the results support the analogous result for planar processes that not all parameters can be estimated consistently under fixed domain asymptotics.

We apply the proposed model to a data-set of precipitable water content over a large portion of the Earth and show that the model gives more precise predictions of the underlying process at unsampled locations than does Matérn model using chordal distances. Technical details are given in an Appendix.

*Keywords:* Great-circle distance; Fractal Dimensions; Matérn covariance; Mean Square Differentiability

---

\*Partially supported by grant FONDECYT 1130647 from the Chilean government.

# 1 Introduction

The last decades have seen an unprecedented increase in the availability of georeferenced data-sets of global extent, for example in the form of environmental monitoring networks or climate model ensembles (Castruccio and Stein, 2013; Porcu et al., 2018a), and motivated by the increasing interest in climate-change (IPCC, 2013). This increase in data-availability has in turn motivated the mathematical and statistical communities to develop models for random fields defined on the two-dimensional surface of the sphere, representing our planet.

By far the most tractable class of random fields is a Gaussian field, whose properties are completely determined by its first two order moments. Thus, correct specification of the covariance function is crucial for parametric inference and optimal spatial prediction (Stein, 1999).

Valid covariance functions are positive definite functions. Proving positive-definiteness for a candidate covariance function is non-trivial. We refer the reader to Schoenberg (1942), Gneiting (2013), Berg and Porcu (2017) and Porcu et al. (2016) for the established theory about positive definite functions on  $d$ -dimensional spheres of  $\mathbb{R}^{d+1}$ . Also, comprehensive recent reviews can be found in Jeong et al. (2017) and Porcu et al. (2018a).

In spatial statistics, it is very common to assume the covariance function of a random field  $Z(\mathbf{x})$  to be isotropic, that is the covariance between  $Z(\mathbf{x}_1)$  and  $Z(\mathbf{x}_2)$  depends only on the distance between  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . For global data, the natural metric is the geodesic or great-circle distance, defined as the shortest arc joining two points located over the spherical shell. If the domain of interest covers a small portion of the Earth, then the curvature of the Earth has a negligible impact and Euclidean distance based on some map projection can be used, but this becomes increasingly inaccurate as the spatial coverage of the data increases.

The chordal distance has been used as an alternative to the geodesic (see Jun and Stein, 2007, with the references therein), but has obvious limitations when the data extend over a substantial proportion of the planet. A detailed critique on the use of chordal distance is provided by Gneiting (2013) and by Porcu et al. (2018a).

The Matérn covariance function (Stein, 1999) is widely considered as the default choice for modelling spatial covariance. Its main attractive feature is its inclusion of a parameter that allows the user to control the mean square differentiability of the associated Gaussian process. This is a key consideration because mean square differentiability is difficult to estimate empirically, but materially affects the properties of spatial prediction under infill asymptotics (see Stein, 1999, and the references therein) and of parameter estimation (Zhang, 2004). The Matérn covariance also has a nice closed form for the associated spectral density, which is convenient for theoretical analysis of the properties of maximum likelihood (ML) estimators (Zhang, 2004), approximate likelihood (Bevilacqua et al., 2012; Furrer et al., 2006; Kaufman and Shaby, 2013) and misspecified linear unbiased prediction (Stein, 1999) under infill asymptotics. A wealth of results is also available within SPDE's with Gaussian Markov approximations (Lindgren et al., 2011) as well as in the numerical analysis literature. We refer the reader to Scheuerer et al. (2013) for more details.

Gneiting (2013) and Porcu et al. (2018a) show that differentiable Matérn covari-

ance functions are no longer positive definite on the sphere when coupled with the geodesic distance. Jeong and Jun (2015a) and Guinness and Fuentes (2016) have addressed the problem of obtaining a spherical analogue of the Matérn function that allows different degrees of flexibility. Jeong and Jun (2015a) suggest smoothing a process with non-differentiable covariance over a spherical cap with given radius. Their approach does not allow for closed forms. Guinness and Fuentes (2016) propose two examples that use Bernoulli polynomials and linear combinations of hypergeometric polynomials. These models do not allow for a continuous parameterisation of differentiability. Also, the authors find that their performance is poorer than the Matérn coupled with chordal distance, both for estimation and prediction, leading them to claim that they *do not see any evidence that the use of chordal distance introduces any distortions*. This is in sharp contrast with Banerjee (2005), Jeong and Jun (2015b) and Porcu et al. (2016) who show exactly the opposite.

Spectral representations can be useful to understand the properties of the associated Gaussian field. Guinness and Fuentes (2016) proposed the so called *Circular Matérn* model based on Fourier series expansions with a parameter that controls mean square differentiability. A drawback of this construction is that no closed forms are in general available, so that in practice truncated versions of the series have to be used, in which case the resulting covariance function is no longer strictly positive definite (Porcu et al., 2018a). Furthermore, when the smoothing parameter is small, the convergence of the series is very slow. Despite these theoretical objections, the Circular Matérn model can perform well in practice, and we will use it as a valid competitor to our proposed model.

In conclusion, the search for covariance functions on spheres that allow for a continuous parameterisation of smoothness has proved elusive, and has been explicitly stated as an open problem in two collections of challenges posed by Gneiting (2013) and by Porcu et al. (2018a). We provide a solution to this problem.

The plan of the paper is the following. Section 2 contains preliminaries needed for the subsequent presentation. Section 3 introduces the  $\mathcal{F}$ -Family of covariance functions on spheres. We then study mean square differentiability and fractal dimension properties of spherical Gaussian fields with the new covariance function. Section 4 describes a simulation study to understand how well the parameters of the new covariance function can be estimated through ML. Our simulation study mimics the infill asymptotic framework (Stein, 1999) that is relevant for processes on a closed and bounded set. We especially focus on the estimation of scale, variance, and microergodic parameter when the smoothing parameter is fixed. Our simulations suggest, as expected, that consistent estimation of the scale and variance parameters is not achievable, but that the microergodic parameter can be estimated consistently. Section 5 analyses a dataset of precipitable water content over a large portion of the planet. We show that the new model of covariance functions has a better predictive performance on this data-set than the Matérn and the Circular Matérn. The paper concludes with discussion. Technical details and generalisations that might be useful for future research are given in Appendix A and B respectively.

## 2 Background

This section provides a background on random fields on spheres, their covariance functions and their spectral representation. We first introduce some notation. For a positive integer  $d$ ,  $\mathbb{S}^d = \{\mathbf{x} \in \mathbb{R}^{d+1}, \|\mathbf{x}\| = 1\}$  denotes the surface of the  $d$ -dimensional unit sphere embedded in  $\mathbb{R}^{d+1}$ , with  $\|\cdot\|$  denoting Euclidean distance. We shall sometimes refer to the Hilbert sphere  $\mathbb{S}^\infty = \{\mathbf{x} \in \mathbb{R}^N, \|\mathbf{x}\| = 1\}$ . The natural metric on  $\mathbb{S}^d$  is the *great-circle distance*,

$$\theta(\mathbf{x}_1, \mathbf{x}_2) = \arccos(\mathbf{x}_1^\top \mathbf{x}_2) \in [0, \pi],$$

for  $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}^d$ , where  $\top$  denotes transpose. The *chordal distance* on  $\mathbb{S}^d$  is

$$d_{\text{CH}}(\mathbf{x}_1, \mathbf{x}_2) = 2 \sin\left(\frac{\theta(\mathbf{x}_1, \mathbf{x}_2)}{2}\right), \quad \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}^d. \quad (2.1)$$

One can compute the Euclidean distance between any two points on a planar map projection of  $\mathbb{S}^d$ .

We denote by  $\{Z(\mathbf{x}), \mathbf{x} \in \mathbb{S}^d\}$  a stationary random field on  $\mathbb{S}^d$ , with constant mean and covariance function  $C(\mathbf{x}_1, \mathbf{x}_2) = \text{cov}\{Z(\mathbf{x}_1), Z(\mathbf{x}_2)\}$ , for  $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}^d$ . The requirement for validity of a candidate function  $C(\mathbf{x}_1, \mathbf{x}_2)$  is that for any positive integer  $n$ ,  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset \mathbb{S}^d$  and  $\{c_1, \dots, c_n\} \subset \mathbb{R}$ ,

$$\text{var}\left(\sum_{i=1}^n c_i Z(\mathbf{x}_i)\right) = \sum_{i,j=1}^n c_i c_j C(\mathbf{x}_i, \mathbf{x}_j) \geq 0. \quad (2.2)$$

Mappings  $C$  that satisfy Equation (2.2) are called positive definite, or strictly positive definite if the inequality is strict for any non-zero vector  $(c_1, \dots, c_n)^\top$ .

If in addition

$$C(\mathbf{x}_1, \mathbf{x}_2) = \sigma^2 \psi(\theta(\mathbf{x}_1, \mathbf{x}_2)), \quad \mathbf{x}_i \in \mathbb{S}^d, \quad i = 1, 2, \quad (2.3)$$

for some value  $\sigma^2 > 0$  and mapping  $\psi : [0, \pi] \rightarrow \mathbb{R}$  such that  $\psi(0) = 1$ , then  $C$  is called a geodesically isotropic covariance (Porcu et al., 2018a), and  $\sigma^2$  is the variance of  $Z(\mathbf{x})$ . Throughout, we use  $\theta$  to denote great-circle distance whenever no confusion can arise. Also, we shall not distinguish between positive and strict positive definiteness unless specifically required. We define  $\Psi_d$  as the class of continuous functions  $\psi$  with  $\psi(0) = 1$  associated to the covariance function  $C$  on  $\mathbb{S}^d$  through the identity (2.3). We also define  $\Psi_\infty = \bigcap_{d=1}^\infty \Psi_d$ , with the strict inclusion relation

$$\Psi_1 \supset \Psi_2 \supset \dots \supset \Psi_d \supset \dots \supset \Psi_\infty.$$

Spectral representations for positive definite functions on spheres are equivalent to Bochner and Schoenberg's theorems in Euclidean spaces (see Daley and Porcu, 2013, and references therein). Schoenberg (1942) showed that a mapping  $\psi : [0, \pi] \rightarrow \mathbb{R}$  belongs to the class  $\Psi_d$  if and only if it can be uniquely written as

$$\psi(\theta) = \sum_{n=0}^{\infty} b_{n,d} \frac{P_n^{(d-1)/2}(\cos \theta)}{P_n^{(d-1)/2}(1)}, \quad \theta \in [0, \pi], \quad (2.4)$$

**Table 1:** Parametric families of members of the classes  $\Psi_\infty$  whose Schoenberg coefficients are available in closed form.

Family	Analytic expression	Parameters range
Negative Binomial	$\psi(\theta) = \left(\frac{1-\delta}{1-\delta\cos\theta}\right)^\tau$	$\delta \in (0, 1), \tau > 0$
Multiquadric	$\psi(\theta) = \left(\frac{(1-p)^2}{1+p^2-2p\cos\theta}\right)^\tau$	$p \in (0, 1), \tau > 0$
Sine Power	$\psi(\theta) = 1 - 2^{-\alpha}(1 - \cos\theta)^{\alpha/2}$	$\alpha \in (0, 2]$
Poisson	$\psi(\theta) = \exp(\lambda(\cos\theta - 1))$	$\lambda > 0$

where  $P_n^\lambda$  denotes the  $\lambda$ -Gegenbauer polynomial of degree  $n$  (Abramowitz and Stegun, 1964), and  $\{b_{n,d}\}_{n=0}^\infty$  is a probability mass sequence. On the two dimensional sphere of  $\mathbb{R}^3$ , Gegenbauer polynomials simplify to Legendre polynomials (Abramowitz and Stegun, 1964).

Schoenberg (1942) also showed that  $\psi$  belongs to the class  $\Psi_\infty$  if and only if

$$\psi(\theta) = \sum_{n=0}^{\infty} b_n(\cos\theta)^n, \quad \theta \in [0, \pi], \quad (2.5)$$

with  $\{b_n\}_{n=0}^\infty$  being again a probability mass sequence. We follow Daley and Porcu (2013) in calling the sequences  $\{b_{n,d}\}_{n=0}^\infty$  in (2.4)  $d$ -Schoenberg sequences of coefficients, to emphasise the dependence on the index  $d$  in the class  $\Psi_d$ . Analogously, we call  $\{b_n\}_{n=0}^\infty$  Schoenberg coefficients. Fourier inversion allows for an explicit representation of the sequences  $\{b_{n,d}\}$ . Specifically,

$$b_{n,d} = \kappa(n, d) \int_0^\pi \psi(\theta) P_n^{(d-1)/2}(\cos\theta) (\sin\theta)^{d-1} d\theta, \quad \psi \in \Psi_d, \quad (2.6)$$

where  $\kappa(n, d)$  is a positive constant (see Berg and Porcu, 2017).

Lang and Schwab (2013) show that the rate of decay of the  $d$ -Schoenberg coefficients determines the regularity properties of the associated Gaussian field in terms of interpolation spaces and Hölder continuities of the sample paths. The  $d$ -Schoenberg coefficients are useful in contexts as diverse as spatial statistics (Guinness and Fuentes, 2016), equivalence of Gaussian measures and infill asymptotics (Arafat et al., 2018), approximation theory (Menegatto et al., 2006; Beatson et al., 2014; Ziegel, 2014; Massa et al., 2017) and spatial point processes (Møller et al., 2018).

Parametric families within the class  $\Psi_\infty$  are listed in Table 1. Each is obtained by evaluating the probability generating function associated with a particular probability mass sequence. The Schoenberg sequence of the last entry is provided in Porcu et al. (2016). For the Sine Power family, the Schoenberg coefficients have been obtained by Soubeyrand et al. (2008). Other parametric families whose Schoenberg coefficients are not available are listed in Gneiting (2013).

The first entry in Table 1 is called the Negative Binomial family, which we denote by  $\mathcal{N}_{\delta,\tau}$  for it. The name comes from the fact that it corresponds to using the negative binomial probability distribution,

$$b_n(\delta, \tau) = \binom{n+\tau-1}{n} \delta^n (1-\delta)^\tau, \quad \delta \in (0, 1), \tau > 0, \quad (2.7)$$

for the coefficient sequence  $b_n$  in (2.5). The Multiquadric family is obtained through the same Schoenberg sequence but under the change of variable  $\delta = 2p/(1 + p^2)$ .

The  $d$ -Schoenberg coefficients can be calculated explicitly using Theorem 4.2(b) in Møller et al. (2018). We refer the reader to Appendix A for details about  $d$ -Schoenberg coefficients and spectral representations over spheres.

One limitation of both the negative binomial and multiquadric families is that they are infinitely differentiable at the origin, making them not very appealing for spatial interpolation, as explained in Stein (1999). Nevertheless, we shall use the function  $\mathcal{N}_{\delta,\tau}$  as the starting point for the construction of our proposed more flexible family.

The Matérn class of covariance functions specifies that  $\text{cov}\{Z(\mathbf{x}_1), Z(\mathbf{x}_2)\} = \sigma^2 \mathcal{M}_{\nu,\alpha}(\|\mathbf{x}_1 - \mathbf{x}_2\|)$  where

$$\mathcal{M}_{\nu,\alpha}(t) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{t}{\alpha}\right)^\nu \mathcal{K}_\nu\left(\frac{t}{\alpha}\right), \quad t \geq 0, \quad (2.8)$$

$\alpha > 0$ ,  $\nu > 0$  and  $\mathcal{K}_\nu$  is a modified Bessel function of the second kind (Abramowitz and Stegun, 1964). The importance of the Matérn class stems from the parameter  $\nu$  that controls the differentiability (in the mean square sense) of the associated Gaussian field. Specifically, for any positive integer  $k$ , a Gaussian field with Matérn covariance function is  $k$ -times mean square differentiable if and only if  $\nu > k$ . Also the Matérn function converges to the Gaussian kernel as  $\nu \rightarrow \infty$ . When  $\nu = k + 1/2$ , for  $k$  a positive integer, the Matérn simplifies into the product of an exponential covariance with a polynomial of order  $k$ . For instance,  $\mathcal{M}_{1/2,1}(t) = \exp(-t)$  and  $\mathcal{M}_{3/2,1}(t) = \exp(-t)(1 + t)$ .

The Matérn covariance function is not in general a valid covariance function on  $\mathbb{S}^2$ . Lemma 1 in Gneiting (2013) shows that the restriction  $\mathcal{M}_{\nu,\alpha}(\theta)$ ,  $\theta \in [0, \pi]$ , is not a member of the class  $\Psi_1$  if  $\nu > 1/2$ . Thus, the Matérn class cannot be used to recover arbitrary degrees of differentiability on spheres. Note also that all the parametric families listed in Table 1 are either nondifferentiable or infinitely differentiable at the origin, and therefore cannot mimic the role of the Matérn function over spheres. The Matérn covariance function is valid on the sphere in conjunction with chordal distance, but for the reasons given above this is generally unsatisfactory.

For data on the surface of the sphere, Guinness and Fuentes (2016) have proposed the Circular Matérn covariance function,  $\mathcal{C}_{\nu,\alpha}$ , given by

$$\mathcal{C}_{\nu,\alpha}(\theta) = \sum_{n=0}^{\infty} b_{n,1} \cos(n\theta), \quad 0 \leq \theta \leq \pi. \quad (2.9)$$

In (2.9), the 1-Schoenberg coefficients are

$$b_{n,1} = \frac{1}{S(\alpha,\nu)} (n^2 + \alpha^2)^{-\nu-1/2},$$

with  $S(\alpha,\nu) = \sum_{n=0}^{\infty} (n^2 + \alpha^2)^{-(\nu+1/2)}$ . The function  $\mathcal{C}_{\alpha,\nu}$  belongs to the class  $\Psi_1$  for any positive  $\alpha$  and  $\nu$ . Further, arguments in Gneiting (2013) show that  $\mathcal{C}_{\alpha,\nu}$

belongs to the class  $\Phi_3$ . The 3-Schoenberg coefficients can be determined according to Lemma 1 in Gneiting (2013). This model is an adaptation of the classical spectral representation of the Matérn covariance on Euclidean spaces to the spherical case. Guinness and Fuentes (2016) show that the parameter  $\nu$  controls the mean square differentiability of the associated Gaussian field on  $\mathbb{S}^2$ , but no closed form expressions are available.

### 3 The $\mathcal{F}$ -Family of Covariance Functions

We now let

$${}_2F_1(a, b, c; z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}, \quad z \in \mathbb{R},$$

denote the Gauss Hypergeometric function, where  $(\cdot)_n$  denotes the Pochhammer symbol and  $B(\cdot)$  the Beta function (Abramowitz and Stegun, 1964). We now define the  $\mathcal{F} = \mathcal{F}_{\tau, \alpha, \nu}$  family of functions through the identity

$$\mathcal{F}_{\tau, \alpha, \nu}(\theta) = \frac{B(\alpha, \nu + \tau)}{B(\alpha, \nu)} {}_2F_1(\tau, \alpha, \alpha + \nu + \tau; \cos \theta), \quad \theta \in [0, \pi], \quad (3.1)$$

Where  $\tau, \alpha$  and  $\nu$  are strictly positive parameters. The main result of this section is the following Theorem.

**Theorem 3.1.** *Let  $\tau, \alpha$  and  $\nu$  be strictly positive. Then, the function  $\mathcal{F}_{\tau, \alpha, \nu}(\cdot)$  defined through Equation (3.1) is a member of the class  $\Psi_{\infty}$ .*

*Proof.* We give a constructive proof on the basis of the following criterion that can be found in Lemma 1 in Gneiting (2013), adapted to our notation.

**Lemma 3.2.** *Let  $A$  be a subset of the Euclidean space and let  $\mu$  be a Borel probability measure on  $A$ . Let  $\psi_c : [0, \pi] \rightarrow \mathbb{R}$  be an element of the class  $\Psi_{\infty}$  for any  $c \in A$ . Then the function  $\psi : [0, \pi] \rightarrow \mathbb{R}$  defined by*

$$\psi(\theta) = \int_A \psi_c(\theta) \mu(\mathrm{d}c), \quad \theta \in [0, \pi], \quad (3.2)$$

*belongs to the class  $\Psi_{\infty}$ .*

We now consider the Negative Binomial family  $\mathcal{N}_{\delta, \tau}$  and the Beta probability measure

$$\mu_{\alpha, \nu}(\mathrm{d}\delta) = \frac{1}{B(\alpha, \nu)} \delta^{\alpha-1} (1 - \delta)^{\nu-1} \mathrm{d}\delta, \quad \delta \in (0, 1), \quad \alpha, \nu > 0. \quad (3.3)$$

We now invoke Lemma 3.2 to claim that  $\mathcal{F}_{\tau, \alpha, \nu} \in \Psi_{\infty}$  because

$$\mathcal{F}_{\tau, \alpha, \nu}(\theta) = \int_{(0,1)} \mathcal{N}_{\delta, \tau}(\theta) \mu_{\alpha, \nu}(\mathrm{d}\delta), \quad \theta \in [0, \pi]. \quad (3.4)$$

In fact, direct inspection shows that, for  $\theta \in [0, \pi]$ ,

$$\begin{aligned}
& \frac{1}{B(\alpha, \nu)} \int_0^1 \left( \frac{1-\delta}{1-\delta \cos \theta} \right)^\tau \delta^{\alpha-1} (1-\delta)^{\nu-1} d\delta \\
&= \frac{1}{B(\alpha, \nu)} \int_0^1 \left( \sum_{n=0}^{\infty} \binom{n+\tau-1}{n} (1-\delta)^\tau \delta^n (\cos \theta)^n \right) \delta^{\alpha-1} (1-\delta)^{\nu-1} d\delta \\
&= \frac{B(\alpha, \nu + \tau)}{B(\alpha, \nu)} \sum_{n=0}^{\infty} \left( \int_0^1 \delta^n \frac{\delta^{\alpha-1} (1-\delta)^{\tau+\nu-1}}{B(\alpha, \nu + \tau)} d\delta \right) \frac{(\tau)_n}{n!} (\cos \theta)^n, \tag{3.5}
\end{aligned}$$

where the second equality comes from (2.7), and the last equality comes from bounded convergence. Note that the integral of the last expression corresponds to the  $n$ -th moment of a Beta distribution with shape parameters  $\alpha$  and  $\tau + \nu$ , which is given by  $(\alpha)_n / (\alpha + \tau + \nu)_n$  (Johnson et al., 1995). Thus, it follows that (3.5) is identically equal to

$$\begin{aligned}
& \frac{B(\alpha, \nu + \tau)}{B(\alpha, \nu)} \sum_{n=0}^{\infty} \frac{(\alpha)_n (\tau)_n}{(\alpha + \nu + \tau)_n} \frac{(\cos \theta)^n}{n!} \\
&= \frac{B(\alpha, \nu + \tau)}{B(\alpha, \nu)} {}_2F_1(\tau, \alpha, \alpha + \nu + \tau; \cos \theta), \tag{3.6}
\end{aligned}$$

$\theta \in [0, \pi]$ , which shows that (3.4) and (3.1) agree as asserted. The proof is completed by invoking the Schoenberg (1942) theorem together with Equation (2.5).  $\square$

The proof of Theorem 3.1 also shows that the Schoenberg coefficients related to the spectral expansion of  $\mathcal{F}_{\tau, \alpha, \nu}$  on the Hilbert sphere  $\mathbb{S}^\infty$  are uniquely defined as

$$b_n(\tau, \alpha, \nu) = \frac{B(\alpha, \nu + \tau)}{B(\alpha, \nu)} \frac{(\alpha)_n (\tau)_n}{(\alpha + \nu + \tau)_n n!}, \quad n = 0, 1, \dots \tag{3.7}$$

Theorem 4.2(b) in Møller et al. (2018) can be used to calculate the associated  $d$ -Schoenberg coefficients  $b_{n,d}(\tau, \alpha, \nu)$ . The detailed calculations are given in Appendix A.

### 3.1 Mean Square Differentiability

Let  $Z(\mathbf{x})$ , for  $\mathbf{x} \in \mathbb{S}^d$ , be a Gaussian random field with geodesically isotropic covariance function  $C$  specified, as in Equation (2.3), by some member  $\psi$  within the class  $\Psi_d$ . Let  $\mathcal{H}_\psi$  be the Hilbert space of linear combinations of  $Z(\mathbf{x})$  with finite variance. A great-circle of  $\mathbb{S}^d$  is the intersection between  $\mathbb{S}^d$  with any plane that passes through the origin. We let  $\mathbb{X}$  denote the set of great circles. Because  $\mathbb{X}$  is isometrically isomorphic to  $\mathbb{S}^1$ , there exists a distance-preserving mapping  $\phi : \mathbb{X} \mapsto [0, 2\pi)$  that associates each point in  $\mathbb{X}$  with an angle. Next, let  $Z_{\mathbb{X}}(\phi(\mathbf{x})) = Z(\mathbf{x})$  be the restriction of  $Z$  to  $\mathbb{X}$ . Then, we say that  $Z_{\mathbb{X}}$  is mean squared differentiable at  $\mathbf{x}$  if the limit

$$Z_{\mathbb{X}}^{(1)}(\phi(\mathbf{x})) = \lim_{\epsilon \rightarrow 0} \frac{Z_{\mathbb{X}}(\phi(\mathbf{x}) + \epsilon) - Z_{\mathbb{X}}(\phi(\mathbf{x}))}{\epsilon}, \quad \mathbf{x} \in \mathbb{S}^d,$$

exists in  $\mathcal{H}_\psi$ ,  $Z$  is mean square differentiable at  $\mathbf{x}$  if  $Z_{\mathbb{X}}^{(1)}(\phi(\mathbf{x}))$  exists for every  $\mathbb{X}$  that contains  $\mathbf{x}$ , and the entire field  $Z$  is mean square differentiable if  $Z$  is mean square differentiable at every  $\mathbf{x} \in \mathbb{S}^d$ . As detailed in Guinness and Fuentes (2016), for isotropic random fields, mean square differentiability at one point along one great-circle implies mean square differentiability of the entire field. Higher order mean square differentiability is defined analogously.

We denote as  $\tilde{\psi}$  the even extension of  $\psi$  to the interval  $[-\pi, \pi]$  and write  $\tilde{\psi}^{(n)}$  for the  $n$ -th order derivative of  $\tilde{\psi}$ . Similarly to the Euclidean case, Guinness and Fuentes (2016) show that  $Z(\mathbf{x})$  with covariance  $\psi$  is  $n$  times mean square differentiable if and only if  $\tilde{\psi}^{(2n)}(0)$  exists. They provide an analogous characterisation in the spectral domain, which we do not use in this paper.

The following result shows that the parameter  $\nu$  controls the smoothness of a random field with covariance  $\mathcal{F}_{\tau, \alpha, \nu}(\theta)$ . In what follows,  $\lfloor x \rfloor$  denotes the largest integer less than or equal to  $x \in \mathbb{R}$ .

**Proposition 3.3.** *Let  $d$  and  $n$  be positive integers. Let  $Z(\mathbf{x})$ ,  $\mathbf{x} \in \mathbb{S}^d$ , be a stationary and isotropic Gaussian random field with covariance function given by  $\mathcal{F}_{\tau, \alpha, \nu}(\theta)$  as in Equation (3.1). Then,  $Z(\mathbf{x})$  is  $n$  times mean square differentiable if and only if  $\lfloor \frac{\nu}{2} \rfloor > n$ .*

*Proof.* We give a constructive proof. Consider the even extension of  $\mathcal{F}_{\tau, \alpha, \nu}(\theta)$  to the interval  $[-\pi, \pi]$  and denote this by  $\tilde{\mathcal{F}}_{\tau, \alpha, \nu}(\theta)$ . The  $n$ -th order differentiability in the mean square sense of  $Z(\mathbf{x})$ ,  $\mathbf{x} \in \mathbb{S}^d$  with covariance  $\mathcal{F}_{\tau, \alpha, \nu}$  (or, equivalently,  $\tilde{\mathcal{F}}_{\tau, \alpha, \nu}$ ) is equivalent to  $2n$ -th order differentiability of  $\tilde{\mathcal{F}}_{\tau, \alpha, \nu}(\theta)$  at  $\theta = 0$ . Using properties of the Gauss hypergeometric function  ${}_2F_1$  (Prudnikov et al., 1983) we have

$$\frac{\partial^n {}_2F_1(a, b, c; \xi)}{\partial \xi^n} = \frac{(a)_n (b)_n}{(c)_n} {}_2F_1(a+n, b+n, c+n; \xi), \quad |\xi| \leq 1, \quad (3.8)$$

where the right hand side of (3.8) is well defined for  $|\xi| = 1$  provided

$$\operatorname{Re}(c - a - b - n) > 0. \quad (3.9)$$

This fact, together with Faà-di Bruno's formula (Mortini, 2013), gives the  $n$ -th derivative of  $\tilde{\mathcal{F}}_{\tau, \alpha, \nu}(\theta)$  as

$$\begin{aligned} \tilde{\mathcal{F}}_{\tau, \alpha, \nu}^{(n)}(\theta) &= \frac{B(\alpha, \nu + \tau)}{B(\alpha, \nu)} \frac{\partial^n \mathcal{F}_{\tau, \alpha, \nu}(\theta)}{\partial \theta^n} \\ &= \sum_{k=1}^n \frac{\partial^n}{\partial \xi^n} {}_2F_1(\tau, \alpha, \tau + \alpha + \nu; \xi) \Big|_{\xi = \cos \theta} B_{n,k}(\cos^{(1)} \theta, \dots, \cos^{(n-k+1)} \theta), \end{aligned} \quad (3.10)$$

where  $B_{n,k}$  are the Bell polynomials (Abramowitz and Stegun, 1964) and  $\cos^{(n)}$  is the  $n$ -th derivative of the cosine function. Hence, according to the condition (3.9), Equation (3.10) is well defined at  $\theta = 0$  if and only if  $\nu > n$ .  $\square$

To prove that  $\mathcal{F}_{\tau, \alpha, \nu}$  allows for a continuous parameterisation of smoothness through the parameter  $\nu$  we need to assess the limiting behaviour of  $\mathcal{F}_{\tau, \alpha, \nu}$  when  $\nu \rightarrow \infty$ . This needs some care as discussed at the end of Section 3.3.

## 3.2 Fractal Dimensions

As noted by Hansen et al. (2015) the roughness or smoothness of a surface at an infinitesimal scale is quantified by the Hausdorff or fractal dimension,  $D$ , which for a surface in  $\mathbb{R}^3$  must lie in the interval  $[2, 3)$ , attaining the lower limit when the surface is differentiable. Hansen et al. (2015) investigate the properties of Gaussian surfaces under several choices of geodesically isotropic covariance functions.

An isotropic random field  $Z(\mathbf{x})$  on the sphere  $\mathbb{S}^2$  with correlation function  $C(\mathbf{x}, \mathbf{y}) = \psi(\theta(\mathbf{x}, \mathbf{y}))$ , for  $\psi \in \Psi_2$ , has fractal index  $a$  if there exists a constant  $b > 0$  such that

$$\lim_{\theta \searrow 0} \frac{\psi(0) - \psi(\theta)}{\theta^a} = b, \quad (3.11)$$

where  $\lim_{\theta \searrow 0}$  is taken from the right. The fractal index exists for most parametric families of correlation functions, in which case the fractal dimension  $D$  and fractal index  $a$  are related by  $D = 3 - a/2$ , so that  $a = 2$  and  $a \rightarrow 0$  correspond to extreme smoothness and roughness, respectively.

**Proposition 3.4.** *Let  $Z(\mathbf{x})$ ,  $\mathbf{x} \in \mathbb{S}^2$ , be a Gaussian random field on the sphere with geodesically isotropic covariance function given by  $\mathcal{F}_{\alpha, \tau, \nu}(\theta)$ ,  $\theta \in [0, \pi]$ . Then, the fractal index of  $Z$  is  $2\nu$  if  $\nu < 1$ ,  $2$  if  $\nu > 1$  and does not exist if  $\nu = 1$ .*

Before providing a formal proof, some comments are in order. According to Proposition 3.4, a realisation of a Gaussian field with covariance function belonging to the  $\mathcal{F}$ -Family will be smooth when  $\nu \geq 1$  and becomes rough when  $\nu$  is smaller than one. In fact,  $D = 3 - \nu$  when  $\nu < 1$ , and  $D = 2$  whenever  $\nu \geq 1$ . The fact that we can characterise the fractal dimension through the parameter  $\nu$  gives an additional way to interpret the effect of the value of this parameter on the properties of the process  $Z(x)$ .

A Gaussian field with Matérn covariance function has fractal dimension  $D = \min(\nu, 1)$ . Amongst the examples proposed by Hansen et al. (2015), only the power kernel allows the fractal dimension to vary; for the other kernels the fractal dimension is constant whatever the parameter setting. The covariances proposed by Guinness and Fuentes (2016) also do not allow different fractal dimensions. We now prove formally the assertion above.

*Proof.* We give a proof of the constructive type. First, application of De L'Hôpital's rule gives

$$\begin{aligned} & \lim_{\theta \searrow 0} \frac{\mathcal{F}_{\alpha, \tau, \nu}(0) - \mathcal{F}_{\alpha, \tau, \nu}(\theta)}{\theta^a} \\ &= \frac{B(\alpha, \tau + \nu)}{B(\alpha, \nu)} \lim_{\theta \rightarrow 0} \frac{\sin \theta {}_2F_1(\alpha + 1, \tau + 1, \alpha + \tau + \nu + 1, \cos \theta)}{a\theta^{a-1}}. \end{aligned} \quad (3.12)$$

Let us first suppose that  $\nu > 1$ . Then (3.12) is well defined if  $a = 2$ . When  $\nu < 1$ ,

we have

$$\begin{aligned}
& \lim_{\theta \searrow 0} \frac{\mathcal{F}_{\alpha, \tau, \nu}(0) - \mathcal{F}_{\alpha, \tau, \nu}(\theta)}{\theta^a} \\
&= \frac{B(\alpha, \tau + \nu)}{aB(\alpha, \nu)} \lim_{\theta \searrow 0} \frac{\sin \theta}{\theta} \frac{{}_2F_1(\alpha + 1, \tau + 1, \alpha + \tau + \nu + 1, \cos \theta)}{(1 - \cos \theta)^{\nu-1}} \frac{(1 - \cos \theta)^{\nu-1}}{\theta^{a-2}} \\
&= \frac{B(\alpha, \tau + \nu)}{aB(\alpha, \nu)} \frac{\Gamma(\alpha + \tau + \nu + 1)\Gamma(1 - \nu)}{\Gamma(\alpha + 1)\Gamma(\tau + 1)} \lim_{\theta \searrow 0} \left( \frac{1 - \cos \theta}{\theta^{(a-2)/(\nu-1)}} \right)^{\nu-1}, \tag{3.13}
\end{aligned}$$

where the last statement is justified in Prudnikov et al. (1983). Also,

$$\lim_{x \searrow 0} \frac{{}_2F_1(\alpha + 1, \tau + 1, \alpha + \tau + \nu + 1, x)}{(1 - x)^{\nu-1}} = \frac{\Gamma(\alpha + \tau + \nu + 1)\Gamma(1 - \nu)}{\Gamma(\alpha + 1)\Gamma(\tau + 1)}.$$

Hence,

$$\lim_{\theta \searrow 0} \left( \frac{1 - \cos \theta}{\theta^{(a-2)/(\nu-1)}} \right)^{\nu-1} = \begin{cases} \frac{1}{2^{\nu-1}} & \text{if } a = 2\nu, \\ 0 & \text{if } a < 2\nu, \\ \infty & \text{if } a > 2\nu. \end{cases}$$

Finally, for  $\nu = 1$  we have that

$$\begin{aligned}
& \lim_{\theta \searrow 0} \frac{\mathcal{F}_{\alpha, \tau, \nu}(0) - \mathcal{F}_{\alpha, \tau, \nu}(\theta)}{\theta^a} \\
&= \frac{B(\alpha, \tau + \nu)}{aB(\alpha, \nu)} \lim_{\theta \searrow 0} \frac{\sin \theta}{\theta} \frac{{}_2F_1(\alpha + 1, \tau + 1, \alpha + \tau + \nu + 1, \cos \theta) - \log(1 - \cos \theta)}{-\log(1 - \cos \theta)} \frac{1}{\theta^{a-1}},
\end{aligned}$$

which exists and is identically equal to zero for  $a = 1$ .  $\square$

### 3.3 Special Cases

We now show that the  $\mathcal{F}$ -family admits closed form expressions when the smoothing parameter is of the form  $\nu = 1/2 + k$ , for positive integers  $k$ . We use the following recurrence formula (Prudnikov et al., 1983); for  $0 < |z| \leq 1$ ,

$$\begin{aligned}
{}_2F_1(a, b, c + 1; z) &= \frac{1}{(c - a)(c - b)z} \left( c(1 - c)(z - 1){}_2F_1(a, b, c - 1; z) \right. \\
&\quad \left. - c(c - 1 - \{2c - a - b - 1\}z){}_2F_1(a, b, c; z) \right). \tag{3.14}
\end{aligned}$$

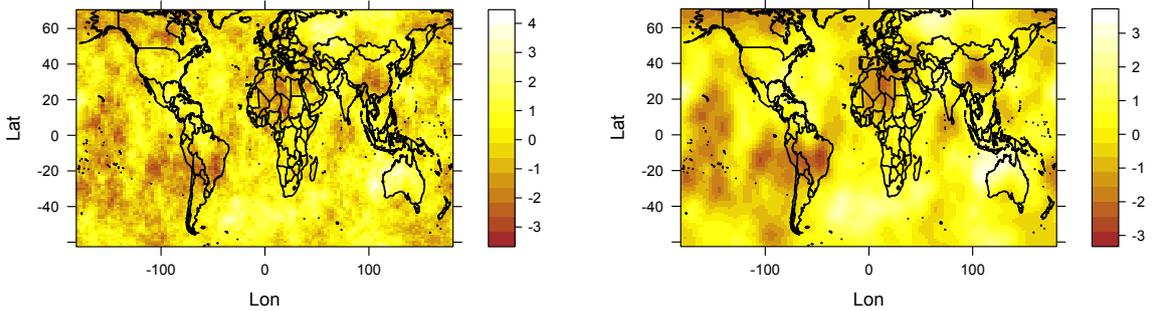
Note that Equation (3.14) is not defined at  $z = 0$ . However, by definition, the Gauss Hypergeometric function is identically equal to 1 at  $z = 0$ .

To iterate the recurrence in Equation (3.14), we need to provide two initial conditions on the right hand side of the equation. For  $|z| \leq 1$  let  $g(z) = \frac{1}{2} + \frac{1}{2}(1 - z)^{1/2}$ . Also, for  $\alpha > 0$  let  $a = \alpha$  and  $b = \alpha + 1/2$ . We then have the following identities (Prudnikov et al., 1983)

$${}_2F_1(a, b, a + b - 1/2; z) = \frac{1}{(1 - z)^{1/2}} g(z)^{1-2\alpha} \tag{3.15}$$

$${}_2F_1(a, b, a + b + 1/2; z) = g(z)^{-2\alpha}. \tag{3.16}$$

In particular, Equation (3.16) provides a covariance function that is continuous but not differentiable at the origin. Thus, any Gaussian random field with such a covariance would be mean square continuous but non-differentiable.



**Figure 1:** Simulated data sets from the  $\mathcal{F}$ -family with  $\sigma^2 = 1$  and an approximated practical range of 0.75 radians. We consider  $\nu = 1/2$  (Left) and  $\nu = 5/2$  (Right). We have used the same random seed for both realisations.

To obtain special cases with higher degrees of differentiability at the origin, we can combine Equation (3.14) with the special cases (3.15) and (3.16):

$${}_2F_1(a, b, a + b + 3/2; z) = \frac{(2\alpha + 1)g(z)^{-2\alpha}p(z)}{(\alpha + 1)(\alpha + 1/2)z},$$

where  $p(z) = -(\alpha + 1/2)(1 - z) + \alpha(1 - z)^{1/2} + 1/2$ . Iterating the formula one more time, we obtain a covariance function generating once mean square differentiable Gaussian random fields:

$${}_2F_1(a, b, a + b + 5/2; z) = \frac{(2\alpha + 2)g(z)^{-2\alpha}}{(\alpha + 2)(\alpha + 3/2)z^2} \cdot \left( \frac{(-2\alpha - 1 + \{2\alpha + 5/2\}z)(2\alpha + 1)p(z)}{(\alpha + 1)(\alpha + 1/2)} + (2\alpha + 1)(1 - z)z \right). \quad (3.17)$$

Figure 1 depicts two realisations from the  $\mathcal{F}$ -family, with  $\nu = 1/2$  and  $\nu = 5/2$ . In order to control the variance, we have used rescaled versions of Equations (3.16) and (3.17). We choose  $\alpha$  such that both covariance functions have an approximate practical range (the great circle distance at which the correlation reaches 0.05) of 0.75 radians. We shall use this parameterisation in Sections 5 and 6.

### 3.4 Limit Cases

Note how the proof of Theorem 1 emphasises that  $\mathcal{F}_{\tau, \alpha, \nu}$  is the scale mixture of  $\mathcal{N}_{\delta, \tau}$  with a Beta distribution with parameters  $\alpha$  and  $\nu$ , where the scale mixing is taken with respect to  $\delta \in (0, 1)$ . A similar argument is used to show that a reparameterised version of  $\mathcal{M}_{\alpha, \nu}(\|\cdot\|)$  converges to  $\exp(-\|\cdot\|^2/\alpha)$  when  $\nu \rightarrow \infty$ , uniformly on any compact set of  $\mathbb{R}^d$ . In particular, when  $\nu \rightarrow \infty$ , the Beta density function tends to a delta measure with single atom at the origin. Therefore, the corresponding limiting covariance is identically one. To proceed formally, we consider the reparameterisation from  $(\tau, \alpha, \nu)$  to  $(\tau, \alpha^*, \nu)$  where  $\alpha^* = \alpha\nu$  and  $\tau, \alpha^*$  and  $\nu$  are all positive. The next result illustrates the limiting behaviour of the  $\mathcal{F}$  covariance when  $\nu \rightarrow \infty$ .

**Proposition 3.5.** *Let  $\mathcal{F}_{\tau,\alpha,\nu}$  be the family defined through (3.1). Let  $\alpha^* = \alpha\nu$ , for  $\alpha$  and  $\tau$  be fixed positive parameters. Let  $\mathcal{N}_{\delta,\tau}$  be the Negative Binomial family, being the first entry in Table 1, with  $\delta \in (0, 1)$  and  $\tau > 0$ . Then, for each  $\theta \in [0, \pi]$ ,*

$$\lim_{\nu \rightarrow \infty} \mathcal{F}_{\tau,\alpha^*,\nu}(\theta) = \mathcal{N}_{\alpha/(\alpha+1),\tau}(\theta).$$

*Proof.* To prove our assertion, we invoke again the scale mixture argument as in Lemma 1 of Gneiting (2013), i.e.

$$\mathcal{F}_{\tau,\alpha^*,\nu}(\theta) = \int_0^1 \mathcal{N}_{\delta,\tau}(\theta) \frac{\delta^{\alpha\nu-1} (1-\delta)^{\nu-1}}{B(\alpha\nu, \nu)} d\delta, \quad \theta \in [0, \pi]. \quad (3.18)$$

The proof follows from the fact that when  $\nu \rightarrow \infty$ , the Beta distribution converges to a delta measure, with a single atom at its expected value  $\alpha/(\alpha + 1)$ .  $\square$

## 4 Simulation Study

The ML method is generally considered to be an efficient method for estimating the parameters of statistical models, although the theoretical justification for this stems primarily from the asymptotic properties of ML estimators. In the present context, the study of asymptotic properties of ML estimators is complicated by the fact that the only physically sensible asymptotic regime for a process on the unit sphere is fixed domain asymptotics, i.e. increasingly dense sampling of  $Z(x)$  on its fixed domain, the unit sphere.

It is generally the case that for spatially continuous processes under fixed domain asymptotics, prediction is consistent but parameter estimation is not (Stein, 1999). The rationale for this simulation study is therefore to explore the finite sample behaviour of ML estimates for the  $\mathcal{F}$ -Family of covariance functions.

A key theoretical result for fixed domain asymptotics is the equivalence of Gaussian measures associated with random fields defined over bounded sets of  $\mathbb{R}^d$  (Skorokhod and Yadrenko, 1973). Equivalence of Gaussian measures has specific consequences for both ML estimation and for kriging predictions. Firstly, equivalence implies that the ML estimates of the parameters of a given class of covariance functions cannot be estimated consistently. Secondly, the miss-specified minimum mean square error predictor under the wrong covariance model is asymptotically equivalent to the kriging predictor under the true covariance. For the Matérn covariance function, using Euclidean distance and assuming the smoothing parameter to be fixed, Zhang (2004) shows that the scale and the variance cannot be estimated consistently. Instead, a specific function of the variance and the scale (called microergodic parameter – see below) can be estimated consistently.

### 4.1 Maximum likelihood estimates

We first study the influence of the correlation range and differentiability on the variability of the ML estimators. Intuitively, we expect the performance of the estimators to deteriorate as the correlation range increases. For the simulations we

parameterise the  $\mathcal{F}$ -Family of covariance functions as

$$\begin{aligned} & \mathcal{F}_{1/\alpha, 1/\alpha+0.5, \nu}(\theta) \\ &= \sigma^2 \frac{\Gamma(\frac{1}{\alpha} + \frac{1}{2} + \nu) \Gamma(\frac{1}{\alpha} + \nu)}{\Gamma(\frac{2}{\alpha} + \frac{1}{2} + \nu) \Gamma(\nu)} {}_2F_1\left(\frac{1}{\alpha}, \frac{1}{\alpha} + \frac{1}{2}, \frac{2}{\alpha} + \frac{1}{2} + \nu, \cos \theta\right), \end{aligned} \quad (4.1)$$

with  $0 \leq \theta \leq \pi$ , so that increasing  $\alpha$  corresponds to increasing correlation range. We also set  $\sigma^2 = 1$  and consider four scenarios for  $\alpha$  and  $\nu$ : (a)  $(\alpha, \nu) = (0.3, 0.5)$ , (b)  $(\alpha, \nu) = (0.6, 0.5)$ , (c)  $(\alpha, \nu) = (0.3, 2.5)$ , and (d)  $(\alpha, \nu) = (0.6, 2.5)$ . Scenarios (a) and (b) correspond to a continuous, non-differentiable random field, whereas Scenarios (c) and (d) to a once differentiable random field. Each simulated realisation generates  $N = 256$  data-values on a 14 by 14 grid of longitudes and latitudes.

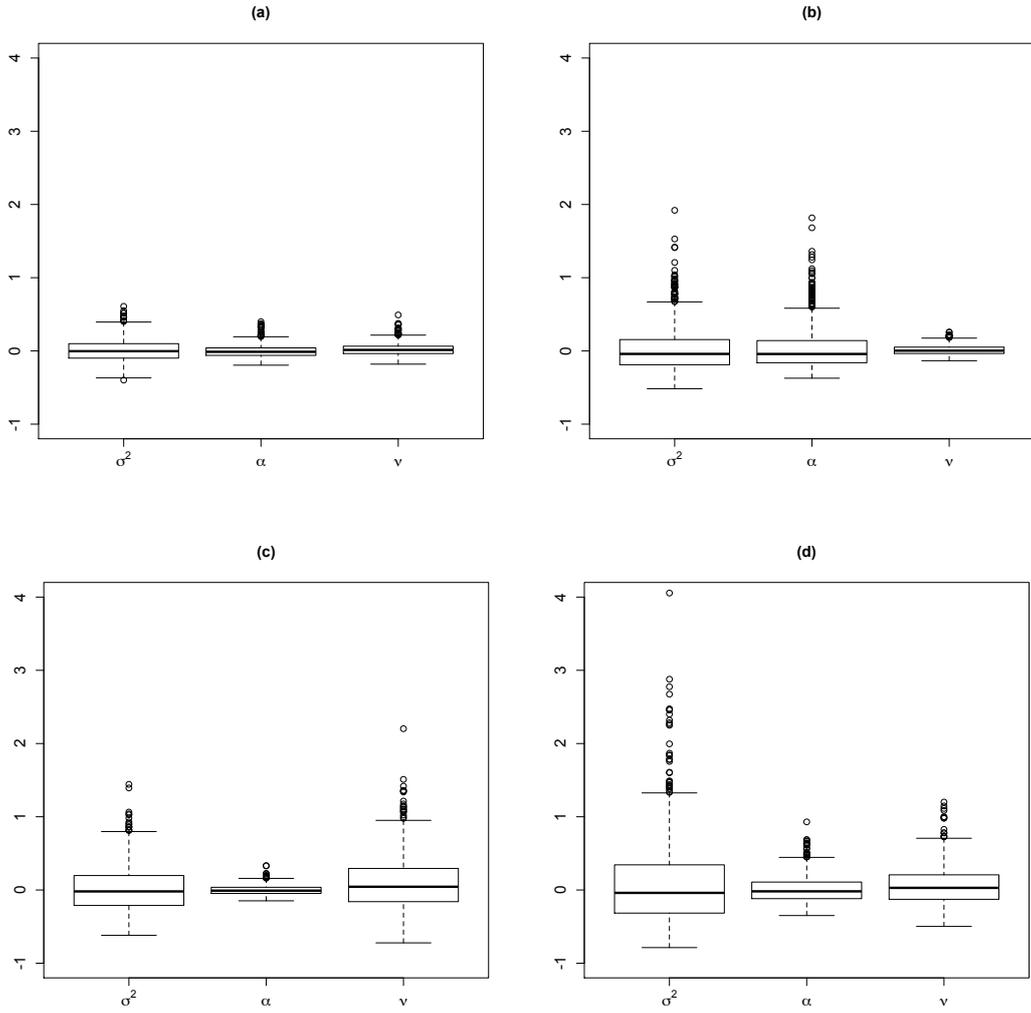
Figure 2 reports the centered boxplots of the ML estimates under Scenarios (a)–(d), based on 1000 independent replications. Larger values of  $\alpha$  or  $\nu$  correspond to higher variability, but there is no evidence of significant bias.

## 4.2 Microergodic parameter

Zhang (2004) has shown that, for the Matérn class of covariance functions as in Equation (2.8), not all parameters can be estimated consistently under infill asymptotics. However, using the parameterisation analogous to ours, in which  $\sigma^2$  is the variance and  $\nu$  determines the degree of differentiability of  $Z(x)$ , the ML estimator of the *microergodic parameter*  $\phi = \sigma^2/\alpha^{2\nu}$  is consistent.

To mimic an infill asymptotic scheme, for each scenario we now generate 1000 realisations of  $Z(x)$  at  $N = 300, 600, 900, 1200, 1500, 1800, 2100$  and 2400 locations uniformly distributed on the unit sphere, with parameter values  $\sigma^2 = 1$ ,  $\alpha = 0.2$  and  $\nu = 1/2$ . Table 2 summarises the properties of the ML estimates by their empirical bias and *relative variance*, *i.e.* the ratio between their sample variance at each value of  $N$  and their sample variance when  $N = 300$ . The biases are again negligible. The standard asymptotic result for parameter estimation is that the variance,  $v$  say, of an ML estimator is proportional to  $N^{-1}$ , hence the log of the relative variance is linear in  $\log(N)$  with slope  $-1$ . Figure 3 shows the empirical relationship between log-transformed relative variance and sample size from our simulation experiment. For the microergodic parameter  $\phi$ , the relationship is close to linear, with estimated slope  $-1.048$ , whereas for  $\sigma^2$  and  $\alpha$ , the estimated slopes are  $-0.872$  and  $-0.913$ , respectively. Also, as  $N$  approaches 2400, there is at least a hint that the linearity is breaking down.

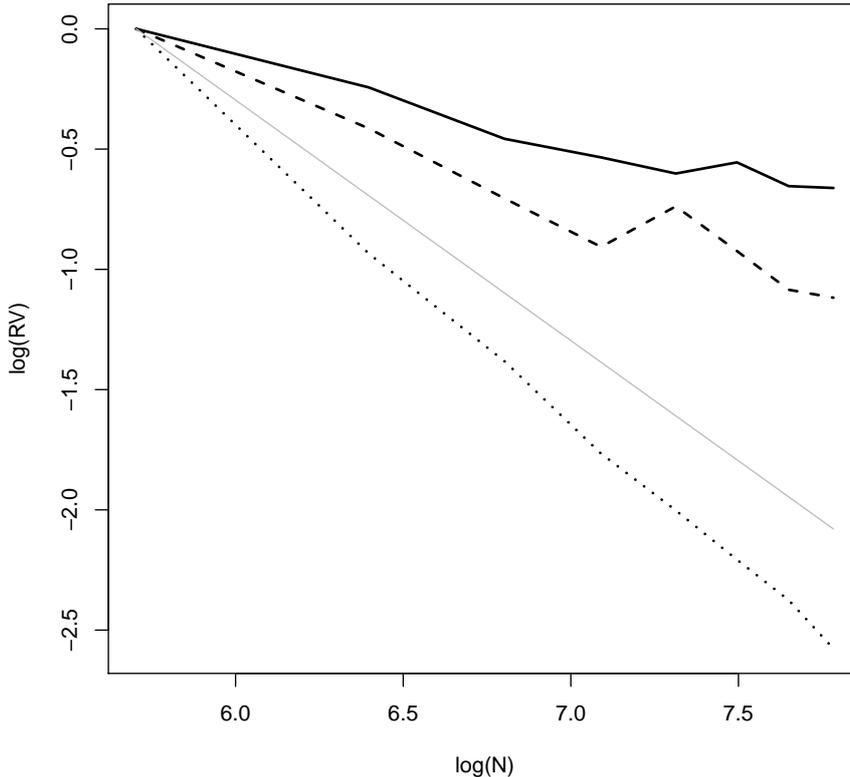
In summary, the experiment suggests that ML estimates for the parameters of the  $\mathcal{F}$ -family behave similarly to those of the planar Matérn model under infill asymptotics.



**Figure 2:** Centered boxplots of the Maximum Likelihood estimates for the  $\mathcal{F}$ -Family under Scenarios (a)–(d), based on 1000 independent repetitions.

**Table 2:** Bias and Relative Variance (RV) for the maximum likelihood estimates of  $\sigma^2$ ,  $\alpha$  and  $\phi = \sigma^2/\alpha^{2\nu}$  against sample size,  $N$ .

Sample Size	$\sigma^2$		$\alpha$		$\phi$	
	Bias	RV	Bias	RV	Bias	RV
300	-0.00511	1	-0.00041	1	0.04908	1
600	-0.00044	0.784	0.00145	0.661	-0.00914	0.393
900	-0.00446	0.633	-0.00005	0.494	-0.00183	0.251
1200	-0.00147	0.586	0.00042	0.404	-0.00516	0.171
1500	0.00029	0.548	0.00052	0.378	-0.00054	0.135
1800	0.00047	0.574	0.00073	0.397	-0.00537	0.110
2100	0.00174	0.520	0.00097	0.338	-0.00740	0.093
2400	-0.00142	0.516	0.00009	0.327	-0.00228	0.076



**Figure 3:** Relationship between log-transformed relative variance and sample size from our simulation experiment.

## 5 Data Illustration

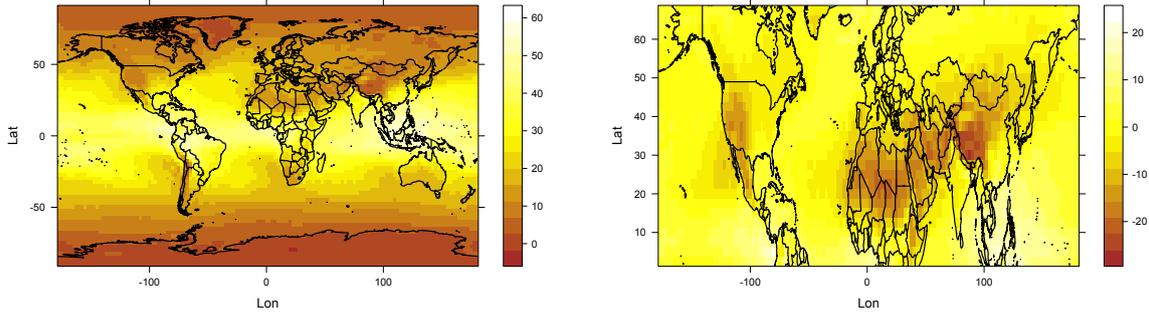
We illustrate the predictive performance of the  $\mathcal{F}$  class of covariance functions on a dataset of Precipitable Water Content (PWC) in  $\text{kg/m}^2$ , downloadable at [www.esrl.noaa.gov](http://www.esrl.noaa.gov). This data-product is considered to be representative of the state of the Earth system (Kalnay et al., 1996) and has been used in regional studies of seasonal stream flow and water scarcity (Müller et al., 2014; Müller and Thompson, 2016).

Here, we analyse the 2017 average of PWC on a grid with spacing  $2.5^\circ$  degrees of longitude and latitude. Figure 4 (left panel) shows that the mean structure depends strongly on latitude. We remove the spatial trend through simple harmonic regression model,

$$\text{mean}(Z(L, \ell)) = \eta_0 + \eta_1 \cos\left(\frac{\pi L}{90^\circ}\right) + \eta_2 \sin\left(\frac{\pi L}{90^\circ}\right),$$

where  $L$  and  $\ell$  denote latitude and longitude, respectively, in degrees. We then analyse the residual spatial variation, restricted to latitudes between  $0^\circ$  and  $70^\circ$  North in order to mitigate the effect of non-stationarities over southern latitudes (Stein, 2007) and to avoid numerical instabilities around the North Pole (Castruccio and Stein, 2013). These residuals are shown in the right panel of Figure 4.

We now compare the performance of the  $\mathcal{F}$  class of covariance functions in Equation (3.1) with respect to ML estimation and kriging predictions. The comparator



**Figure 4:** (Left) Average for Precipitable Water Content for 2017. (Right) Precipitable Water Content for the northern hemisphere, after removing its trend through simple harmonic regression.

families of correlation functions are the following.

1. The  $\mathcal{F}$  correlation function, defined according to Equation (4.1)
2. The Circular-Matérn correlation function (Guinness and Fuentes, 2016) as given in Equation (2.9).

As explained in Section 2, in practice we need to truncate the series expansion in Equation (2.9). Here, we truncate the sum after 1000 terms; see Lang and Schwab (2013), who adopt the same strategy and give bounds for the approximation in the mean square sense.

3. A Matérn correlation function  $\mathcal{M}_{\nu,\alpha}(d_{\text{CH}})$ , where  $d_{\text{CH}}$  denotes the chordal distance.

Following Stein (2007), to control the latitude-dependent variance we augment the three families of correlation functions with multiplicative terms. For the correlations based on the great-circle distance, we therefore consider models of the form

$$\sigma(L_1)\sigma(L_2)\psi(\theta(\mathbf{x}_1, \mathbf{x}_2)),$$

where  $\mathbf{x}_i \in \mathbb{S}^2$  has associated latitude  $L_i$ , for  $i = 1, 2$ . We use an analogous structure for the Matérn correlation function based on chordal distance. Here,

$$\sigma(L) = \sum_{k=0}^M \varrho_k P_k(\cos L), \quad (5.1)$$

where  $P_k$  is the Legendre polynomial of degree  $k$  (Dai and Xu, 2013) and the coefficients  $\varrho_i$  are unknown parameters. To alleviate computational burden and numerical instabilities we fix  $M = 1$  in Equation (5.1). The resulting parameter vector for each of the three models is  $(\varrho_0, \varrho_1, \alpha, \nu)^\top$ .

We use the following two-step procedure:

- Step 1: We first sample 200 data-locations independently at random over the region delimited by latitudes  $0^\circ$  to  $70^\circ$  and longitudes  $-180^\circ$  to  $180^\circ$ . We use this data as a training set to calculate the ML estimates for each of the three models.

Step 2: Then, we reserve the residuals at 20 sampled locations in the region delimited by  $0^\circ$  to  $70^\circ$  latitude and  $120^\circ$  to  $180^\circ$  in longitude as a validation set. We repeat this step 100 times in order to evaluate the predictive performance of the models.

A similar experiment has been carried out by Jeong and Jun (2015b). They note that this scenario may arise in practice when the interest is in predicting over the ocean but most observations are on land.

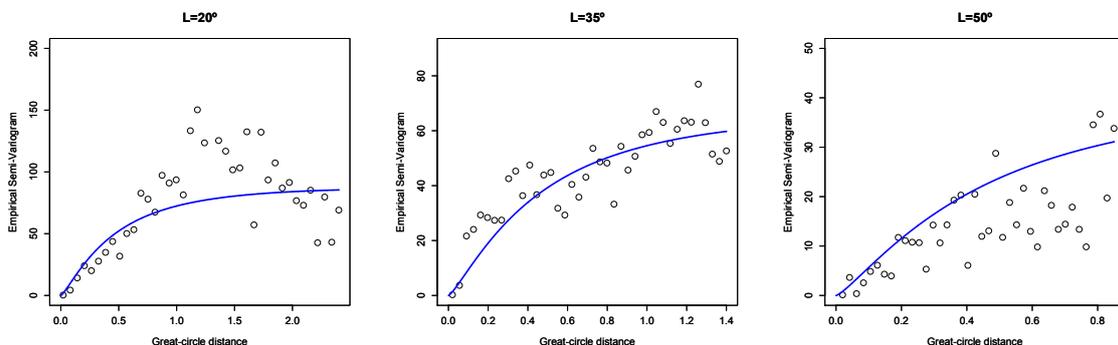
Table 3 reports the ML estimates for each model, their empirical standard errors and the maximised log-likelihood. In all three cases, the estimate  $\hat{\nu}$  corresponds to a continuous but non-differentiable field. The values of the maximised log-likelihood are very similar.

**Table 3:** ML estimates and Log-Likelihood value attained at the optimum, for each model. Standard errors are specified in parentheses.

Model	$\hat{\varrho}_0$	$\hat{\varrho}_1$	$\hat{\alpha}$	$\hat{\nu}$	Log-Likelihood
$\mathcal{F}$ -Family	-0.275 (0.942)	10.503 (2.831)	0.573 (0.361)	0.675 (0.112)	-522.18
Circular-Matérn	-0.279 (0.935)	10.466 (2.376)	0.384 (0.173)	0.644 (0.094)	-521.83
Matérn Chordal	-0.271 (0.880)	9.876 (2.159)	0.388 (0.189)	0.646 (0.095)	-521.67

Figure 5 shows the empirical semi-variogram using points within a horizontal window at latitudes  $20^\circ$ ,  $35^\circ$  and  $50^\circ$ , together with the fitted  $\mathcal{F}$  and Matérn Chordal models, using a tolerance region of  $10^\circ$ ; the fitted variogram for the Matérn and Circular-Matérn model is not shown, because it is very similar to the curves shown here. From this perspective, all three models fit equally well.

We now compare the models in terms of their predictive performance. From the



**Figure 5:** Empirical semi-variogram at latitudes  $L = 20^\circ$ ,  $35^\circ$  and  $50^\circ$ , and the fitted curves for the  $\mathcal{F}$  covariance function.

results for repetitions  $i = 1, \dots, 100$  we calculate

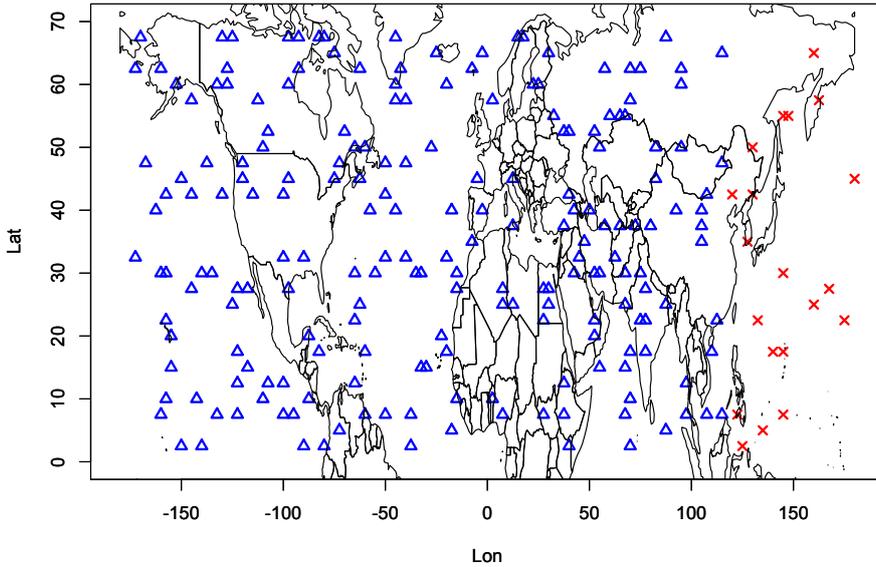
$$\text{RMSE}_i = \left( \frac{1}{n_i} \sum_{j=1}^{n_i} (Z(\mathbf{x}_{ij}) - \widehat{Z}(\mathbf{x}_{ij}))^2 \right)^{1/2},$$

and

$$\text{MAE}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} |Z(\mathbf{x}_{ij}) - \widehat{Z}(\mathbf{x}_{ij})|,$$

where  $\mathbf{x}_{ij} : j = 1, \dots, n_i$  is the validation set from repetition  $i$ , and  $Z(\mathbf{x})$  and  $\widehat{Z}(\mathbf{x})$  denote the true and predicted values, respectively, of the field at the site  $\mathbf{x}$ .

Figure 7 depicts, for each model, the distribution of RMSE and MAE over the 100 repetitions of the experiment, whilst Table 4 gives their average values. In each case, the  $\mathcal{F}$  model gives the lowest of the three values, although the improvements over the other two models are modest: 2.2% and 8% in terms of relative RMSE, 1.8% and 6.9% in terms of relative MAE, compared with the Circular-Matérn and the Matérn Chordal models, respectively.



**Figure 6:** Training ( $\Delta$ ) and validation ( $X$ ) sets for the prediction experiment. We illustrate only over 100 repetitions of this experiment.

**Table 4:** RMSE and MAE averages, for each model, based on 100 repetitions.

	$\mathcal{F}$ -Family	Circular-Matérn	Matérn chordal
RMSE	<b>3.535</b>	3.613	3.844
MAE	<b>2.910</b>	2.964	3.128

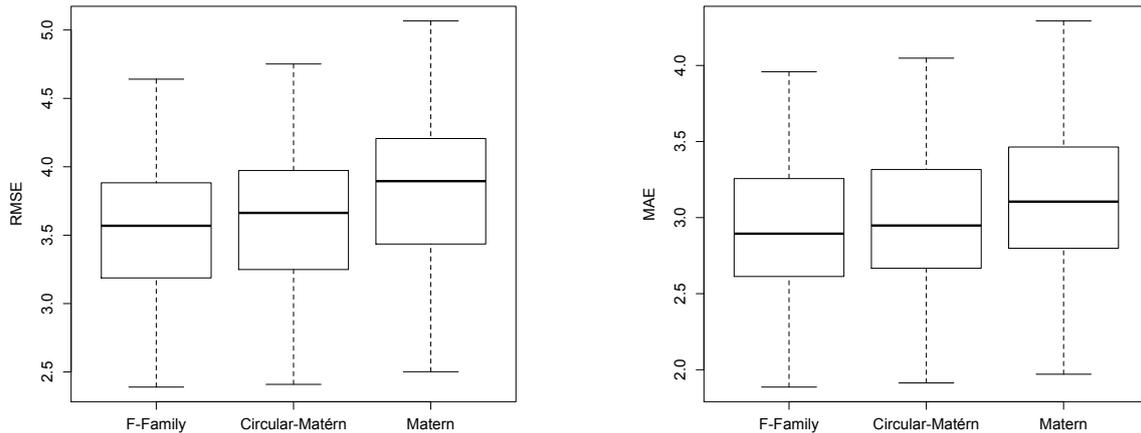


Figure 7: Boxplots for RMSE and MAE, for each model.

## 6 Discussion

Our simulation experiment supports the view that the consistency properties of ML estimation under our proposed model mirror known results for the analogous parameterisation of the planar Matérn model. Proving this is a challenging problem. A possible line of attack would be to use the theory developed by Arafat et al. (2018) about equivalence of Gaussian measures on spherical spaces. This would amount to use the  $d$ -Schoenberg sequences associated to the  $\mathcal{F}$  class, which are provided in Appendix A.

Our illustrative analysis based on the PWC data-set shows the limitations of using a model based on chordal distances. These limitations become more apparent as data-locations cover an increasing proportion of the sphere, and when prediction locations become more remote from data-locations. In additional simulation experiments not reported here, we have found that within the span of the data-locations, predictions based on the  $\mathcal{F}$  and Matérn Chordal models are often almost identical, which is as expected because spatial interpolations based on Gaussian models are driven primarily by local properties of the covariance function (Stein, 1999).

The reported differences in predictive performance for the PWC data-set are similar to those reported in other comparative studies. For example, in Gneiting et al. (2010), the multivariate Matérn model and the Linear Model of Coregionalisation show a discrepancy of approximately 1% in terms of MAE, whilst in Jeong et al. (2017) reported relative differences in MAE of 6% to 8% when comparing several parametric non-stationary models for fields on spheres.

One thing that might be worth being considered is that Stein (2007) proposes to control the latitude-dependent variance and the measurement error by augmenting the correlation function with an additive nugget effect. In our experiments, we always found that our estimates for the nugget were identically equal to zero, so that we excluded this from the presentation in Section 5.

One important thing to be mentioned is that often climate data are not isotropic on the sphere. In particular, Stein (2007) evokes Jones (1963) to call those covariance functions defined over  $S^2$  that are nonstationary over latitudes but stationary over longitudes *axially symmetric*. Appendix B shows that the  $\mathcal{F}$ -Family introduced in

this paper can be used as a building block to create models that satisfy Jones hypothesis.

## Acknowledgments

Emilio Porcu is partially supported by *Proyecto Fondecyt* number 1170290 from the Chilean Commission for Scientific and Technological Research. Francisco Cuevas is supported by The Danish Council for Independent Research | Natural Sciences, grant DFF-7014-00074 “Statistics for point processes in space and beyond”, and by the “Centre for Stochastic Geometry and Advanced Bioimaging”, funded by grant 8721 from the Villum Foundation.

## Appendices

### A $d$ -Schoenberg Coefficients for the $\mathcal{F}$ Covariance

Let  $p$  and  $q$  be positive integers. The generalised hypergeometric functions (Abramowitz and Stegun, 1964) are defined as

$${}_pF_q \left[ \begin{matrix} a_1 & \cdots & a_p \\ b_1 & \cdots & b_q \end{matrix}; z \right] = \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{z^n}{n!}, \quad z \in \mathbb{R}.$$

The special case  ${}_2F_1 \left[ \begin{matrix} a & b \\ c \end{matrix}; z \right] = {}_2F_1(a, b, c; z)$  has been used in Section 3 to introduce the  $\mathcal{F}$ -Family in Equation (3.1).

Let  $d$  be a positive integer. We now provide detailed calculations for the  $d$ -Schoenberg coefficients associated to (3.1) through the identity (2.6). Theorem 3.1 has provided an expression for the Schoenberg coefficients related to the Hilbert sphere  $\mathbb{S}^\infty$ .

**Property A.1.** *Let  $\mathcal{F}_{\tau, \alpha, \nu}$  be the family of functions defined through Equation (3.1). Let  $d$  be a positive integer. Then, the  $d$ -Schoenberg sequence of coefficients  $\{b_{n,d}\}_{n=0}^{\infty}$  related to  $\mathcal{F}_{\tau, \alpha, \nu}$  through Equation (2.6) are uniquely determined by*

$$b_{n,d} = \begin{cases} K_{\text{even}}(\alpha, \tau, \nu, n, d) {}_5F_4 \left[ \begin{matrix} \frac{\alpha}{2} + n & \frac{\alpha+1}{2} + n & \frac{\tau}{2} + n & \frac{\tau+1}{2} + n & 1 \\ \frac{\alpha+\nu+\tau}{2} + n & \frac{\alpha+\nu+\tau+1}{2} + n & \frac{n}{2} + 1 & \frac{3n+d+1}{2} \end{matrix}; 1 \right] & \text{if } n \text{ is even,} \\ K_{\text{odd}}(\alpha, \tau, \nu, n, d) {}_5F_4 \left[ \begin{matrix} \frac{\alpha}{2} + n + 1 & \frac{\alpha+1}{2} + n & \frac{\tau+1}{2} + n & \frac{\tau}{2} + n + 1 & 1 \\ \frac{\alpha+\nu+\tau+1}{2} + n & \frac{\alpha+\nu+\tau}{2} + n + 1 & \frac{n+3}{2} & \frac{3n+d+2}{2} \end{matrix}; 1 \right] & \text{if } n \text{ is odd,} \end{cases}$$

where

$$\begin{aligned} & K_{\text{even}}(\alpha, \tau, \nu, n, d) \\ &= \frac{B(\alpha, \nu + \tau)}{B(\alpha, \nu)} \frac{(\alpha)_{2n} (\tau)_{2n} (2n + d - 1) \Gamma(\frac{d-1}{2})}{(\alpha + \nu + \tau)_{2n} 2^{2n+1} \Gamma(\frac{n+2}{2}) \Gamma(\frac{3n+d+1}{2})} \binom{n + d - 2}{n}, \end{aligned}$$

and

$$\begin{aligned} & K_{\text{odd}}(\alpha, \tau, \nu, n, d) \\ &= \frac{B(\alpha, \nu + \tau)}{B(\alpha, \nu)} \frac{(\alpha)_{2n+1} (\tau)_{2n+1} (2n + d - 1) \Gamma(\frac{d-1}{2})}{(\alpha + \nu + \tau)_{2n+1} 2^{2n+2} \Gamma(\frac{n+3}{2}) \Gamma(\frac{3n+d+2}{2})} \binom{n + d - 2}{n}. \end{aligned}$$

*Proof.* The proof of Theorem 1 has shown that the Schoenberg coefficients related to the  $\Psi_\infty$  representation of the family  $\mathcal{F}$  are defined by

$$b_n = \frac{B(\alpha, \nu + \tau)}{B(\alpha, \nu)} \frac{(\alpha)_n(\tau)_n}{(\alpha + \nu + \tau)_n(n!)}. \quad (\text{A.1})$$

Theorem 4.1 in Møller et al. (2018) shows that, since the function  $\mathcal{F}_{\tau, \alpha, \nu}$  belongs to the class  $\Psi_\infty$ , its  $d$ -Schoenberg coefficients  $b_{n,d}$  are uniquely determined through

$$b_{n,d} = \sum_{\substack{l=n \\ n-l \equiv 0 \pmod{2}}}^{\infty} b_n \gamma_{n,l}^{(d)} \quad (\text{A.2})$$

where  $b_n$  is defined in (A.1) and

$$\gamma_{n,l}^{(d)} = \frac{(2n + d - 1)(l!) \Gamma(\frac{d-1}{2})}{2^{l+1} \{(\frac{l-n}{2})!\} \Gamma(\frac{l+n+d+1}{2})} \binom{n + d - 2}{n}.$$

We can now plugin (A.1) into (A.2) to get

$$\begin{aligned} b_{n,d} &= K_1(\alpha, \tau, \nu, n, d) \sum_{\substack{l=n \\ n-l \equiv 0 \pmod{2}}}^{\infty} \frac{(\alpha)_l(\tau)_l}{(\alpha + \nu + \tau)_l 2^{l+1} \{(\frac{l-n}{2})!\} \Gamma(\frac{l+n+d+1}{2})} \\ &= K_1(\alpha, \tau, \nu, n, d) S_n(\alpha, \tau, \nu, n, d), \end{aligned}$$

where

$$K_1(\alpha, \tau, \nu, n, d) = (2n + d - 1) \Gamma\left(\frac{d-1}{2}\right) \binom{n + d - 2}{n} \frac{B(\alpha, \nu + \tau)}{B(\alpha, \nu)}.$$

When  $n$  is an even positive integer, we get that

$$S_n(\alpha, \tau, \nu, n, d) = \sum_{j=n}^{\infty} \frac{(\alpha)_{2j}(\tau)_{2j}}{(\alpha + \nu + \tau)_{2j} 2^{2j+1} \{(\frac{2j-n}{2})!\} \Gamma(\frac{2j+n+d+1}{2})}.$$

We now define  $i = j - n$  to obtain

$$S_n(\alpha, \tau, \nu, n, d) = \sum_{i=0}^{\infty} \frac{(\alpha)_{2i+2n}(\tau)_{2i+2n}}{(\alpha + \nu + \tau)_{2i+2n} 2^{2i+2n+1} \{(i + \frac{n}{2})!\} \Gamma(i + \frac{3n+d+1}{2})}.$$

We can now make use of the following factorisation for Pochhammer symbols (Prudnikov et al., 1983)

$$(\alpha)_{2i+2n} = (\alpha)_{2n} (a + n)_{2i},$$

so to obtain

$$\begin{aligned} &S_n(\alpha, \tau, \nu, n, d) \\ &= \frac{(\alpha)_{2n}(\tau)_{2n}}{(\alpha + \nu + \tau)_{2n} 2^{2n+1}} \sum_{i=0}^{\infty} \frac{(\alpha + n)_{2i}(\tau + n)_{2i}}{(\alpha + \nu + \tau + n)_{2i} 2^{2i+1} \{(i + \frac{n}{2})!\} \Gamma(i + \frac{3n+d+1}{2})}. \end{aligned}$$

Using the dimidiation formula for the Pochhammer symbol (Prudnikov et al., 1983)

$$(\alpha)_{2i} = 2^{2i}(\alpha/2)_i((\alpha + 1)/2)_i,$$

and completing terms, the series  $S_n$  is

$$\begin{aligned} S_n(\alpha, \tau, \nu, n, d) &= K_2(\alpha, \tau, \nu, n, d) \\ &\cdot \sum_{i=0}^{\infty} \frac{(\frac{\alpha+1}{2} + n)_i (\frac{\alpha}{2} + n)_i (\frac{\tau}{2} + n)_i (\frac{\tau+1}{2} + n)_i (1)_i}{(\frac{\alpha+\nu+\tau}{2} + n)_i (\frac{\alpha+\nu+\tau+1}{2} + n)_i (\frac{n}{2} + 1)_i (\frac{3n+d+1}{2})_i i!}, \\ &= K_2(\alpha, \tau, \nu, n, d) {}_5F_4 \left[ \begin{matrix} \frac{\alpha}{2} + n & \frac{\alpha+1}{2} + n & \frac{\tau}{2} + n & \frac{\tau+1}{2} + n & 1 \\ \frac{\alpha+\nu+\tau}{2} + n & \frac{\alpha+\nu+\tau+1}{2} + n & \frac{n}{2} + 1 & \frac{3n+d+1}{2} \end{matrix}; 1 \right], \end{aligned}$$

where

$$\begin{aligned} K_2(\alpha, \tau, \nu, n, d) &= \frac{(\alpha)_{2n}(\tau)_{2n}}{(\alpha + \nu + \tau)_{2n} 2^{2n+1} \Gamma(\frac{n}{2}) \Gamma(\frac{3n+d+1}{2})}. \\ K_{\text{even}}(\alpha, \tau, \nu, n, d) &= K_1(\alpha, \tau, \nu, n, d) K_2(\alpha, \tau, \nu, n, d). \end{aligned}$$

When  $n$  is an odd positive integer, the proof works *mutatis mutandis* through similar calculations.  $\square$

## B Axially Symmetric Version of the $\mathcal{F}$ Class

For phenomena covering a big portion of our planet, isotropy is a questionable assumption. On the one hand, isotropy might be expected for microscale meteorology on a sufficiently temporally aggregated level for many physical quantities. On the other hand, mesoscale and synoptic scale meteorology are not even approximately isotropic, due to the highly nonlinear nature of the Earth's system. Indeed, Stein (2007) shows that total column ozone data show significant changes over latitude. Castruccio and Stein (2013) argued that both inter and intra annual variability for surface temperature is dependent on latitude. For the sequel, we refer to the unit sphere  $\mathbb{S}^2$  of  $\mathbb{R}^3$  with coordinates  $\mathbf{x} = (L, \ell)^\top$ , with  $L \in [0, \pi]$  denoting latitude and  $\ell \in [0, 2\pi)$  denoting longitude. In particular, Stein (2007) resorts to the results in Jones (1963) to call the covariance  $C$  axially symmetric when

$$C(\mathbf{x}_1, \mathbf{x}_2) = \mathcal{C}(L_1, L_2, \ell_1 - \ell_2), \quad (L_i, \ell_i) \in [0, \pi] \times [0, 2\pi), i = 1, 2.$$

Axially symmetric processes have a well understood spectral representation that includes as a special case the geodesic isotropy illustrated through Equations (2.3) and (2.4). For details, the reader is referred to Jones (1963) and more recently to Stein (2007).

The literature on axially symmetric models is sparse, with the attempt in Porcu et al. (2018b) being a notable exception. Let  $d_{\text{CH}}(\ell_1, \ell_2)$  denote the chordal distance between two longitudes  $\ell_1$  and  $\ell_2$ . Let  $\mathcal{M}_{\alpha, \nu}$  denote the Matérn class defined at (2.8). Then, Porcu et al. (2018b) propose an axially symmetric model of the type

$$\mathcal{C}(L_1, L_2, \ell_1 - \ell_2) = \sigma(L_1, L_2) \mathcal{M}_{\alpha(L_1, L_2), \nu(L_1, L_2)}(d_{\text{CH}}(\ell_1, \ell_2)),$$

$(L_i, \ell_i) \in [0, \pi] \times [0, 2\pi)$ ,  $i = 1, 2$ , where  $\sigma$ ,  $\alpha$  and  $\nu$  are strictly positive functions that must be carefully chosen in order to preserve positive definiteness. The interpretation of these functions is very intuitive, as they indicate how, respectively, variance, scale and smoothness can vary across latitudes. Usually  $\sigma$  is modeled through a linear combination of Legendre polynomial (Jun and Stein, 2007). To illustrate the new model, we need to define a stochastic process  $\{X(L), L \in [0, \pi]\}$  and we call variogram the quantity  $\text{var}(X(L_2) - X(L_1))/2$ ,  $L_1, L_2 \in [0, \pi]$  (see Chiles and Delfiner, 1999, with the references therein).

**Theorem B.1.** *Let  $\mathcal{F}$  be the family of functions defined at Equation (3.1). Let  $\tau > 0$ . Let  $\sigma : [0, \pi]^2 \rightarrow \mathbb{R}_+$  be positive definite and let  $\alpha, \nu : [0, \pi]^2 \rightarrow \mathbb{R}_+$  be continuous functions such that the functions  $(L_1, L_2) \mapsto \alpha(L_1, L_2)$  and  $(L_1, L_2) \mapsto \nu(L_1, L_2)$  define two variograms on  $[0, \pi]^2$ . Then, the function*

$$\mathcal{C}(L_1, L_2, \ell_1 - \ell_2) = \sigma(L_1, L_2) \mathcal{F}_{\tau, \alpha(L_1, L_2), \nu(L_1, L_2)}(\theta(\ell_1, \ell_2)),$$

for  $(L_i, \ell_i) \in [0, \pi] \times [0, 2\pi)$ , is positive definite.

*Proof.* We give a constructive proof. We consider the scale mixture

$$\int_0^1 \frac{(1 - \delta)^\tau}{(1 - \delta \cos \theta(\ell_1, \ell_2))^\tau} \delta^{\alpha(L_1, L_2) - 1} (1 - \delta)^{\nu(L_1, L_2) - 1} d\delta.$$

Clearly, the function  $(\ell_1, \ell_2) \mapsto (1 - \delta)^\tau / (1 - \delta \cos \theta(\ell_1, \ell_2))^\tau$  is positive definite for any  $\tau > 0$  and  $\delta \in (0, 1)$ . Invoking Schoenberg's theorem (Schoenberg, 1942), it is easy to show that both functions  $a^\alpha$  and  $a^\nu$  are positive definite on  $[0, \pi]^2$  provided  $0 \leq a \leq 1$ . Since the scale mixture above is well defined, the proof is completed by using the same arguments as in Theorem 3.1.  $\square$

## References

- Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, volume 55. Courier Corporation.
- Arafat, A., Porcu, E., Bevilacqua, M., and Mateu, J. (2018). Equivalence and orthogonality of Gaussian measures on spheres. *Journal of Multivariate Analysis*, 267:306–318.
- Banerjee, S. (2005). On geodetic distance computations in spatial modeling. *Biometrics*, 61:617–625.
- Beatson, R. K., zu Castell, W., and Xu, Y. (2014). Pólya criterion for (strict) positive definiteness on the sphere. *IMA Journal of Numerical Analysis*, 34:550–568.
- Berg, C. and Porcu, E. (2017). From Schoenberg coefficients to Schoenberg functions. *Constructive Approximation*, 45:217–241.
- Bevilacqua, M., Gaetan, C., Mateu, J., and Porcu, E. (2012). Estimating space and space-time covariance functions: a weighted composite likelihood approach. *Journal of the American Statistical Association*, 107:268–280.
- Castruccio, S. and Stein, M. L. (2013). Global Space-Time Models for Climate Ensembles. *Annals of Applied Statistics*, 7:1593–1611.

- Chiles, J. and Delfiner, P. (1999). *Geostatistics: Modeling Spatial Uncertainty*. New York: Wiley.
- Dai, F. and Xu, Y. (2013). *Approximation Theory and Harmonic Analysis on Spheres and Balls*. Springer.
- Daley, D. J. and Porcu, E. (2013). Dimension walks and Schoenberg spectral measures. *Proceedings of the American Mathematical Society*, 141:1813–1824.
- Furrer, R., Genton, M. G., and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 15:502–523.
- Gneiting, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli*, 19:1327–1349.
- Gneiting, T., Kleiber, W., and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105(491):1167–1177.
- Guinness, J. and Fuentes, M. (2016). Isotropic covariance functions on spheres: some properties and modeling considerations. *Journal of Multivariate Analysis*, 143:143–152.
- Hansen, L. V., Thorarinsdottir, T. L., Ovcharov, E., Gneiting, T., and Richards, D. (2015). Gaussian random particles with flexible Hausdorff dimension. *Advances in Applied Probability*, 47(2):307–327.
- IPCC (2013). Climate change 2013: The physical science basis. In Stocker, T. et al., editors, *Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, New York.
- Jeong, J., Castruccio, S., Crippa, P., and Genton, M. G. (2017). Statistics-based compression of global wind fields. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. in press.
- Jeong, J. and Jun, M. (2015a). A class of Matérn-like covariance functions for smooth processes on a sphere. *Spatial Statistics*, 11:1–18.
- Jeong, J. and Jun, M. (2015b). Covariance models on the surface of a sphere: when does it matter? *STAT*, 4:167–182.
- Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions, vol. 2*. Wiley, New York,.
- Jones, R. H. (1963). Stochastic processes on a sphere. *Annals of Mathematical Statistics*, 34:213–218.
- Jun, M. and Stein, M. L. (2007). An approach to producing space-time covariance functions on spheres. *Technometrics*, 49:468–479.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society*, 77(3):437–471.
- Kaufman, C. and Shaby, B. (2013). The role of the range parameter for estimation and prediction in geostatistics. *Biometrika*, 100:473–484.
- Lang, A. and Schwab, C. (2013). Isotropic random fields on the sphere: regularity, fast simulation and stochastic partial differential equations. *Annals of Applied Probability*, 25:3047–3094.

- Lindgren, F., Rue, H., and Lindstroem, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B*, 73:423–498.
- Massa, E., Perón, A., and Porcu, E. (2017). Positive definite functions on complex spheres, and their walks through dimensions. *SIGMA*, 13.
- Menegatto, V. A., Oliveira, C. P., and Perón, A. P. (2006). Strictly positive definite kernels on subsets of the complex plane. *Computational Mathematics and Applications*, 51:1233–1250.
- Møller, J., Nielsen, M., Porcu, E., and Rubak, E. (2018). Determinantal point process models on the sphere. *Bernoulli*, 24(2):1171–1201.
- Mortini, R. (2013). The Faa di Bruno formula revisited. *arXiv:1303.2023*.
- Müller, M. and Thompson, S. (2016). Comparing statistical and process-based flow duration curve models in ungauged basins and changing rain regimes. *Hydrology and Earth System Sciences*, 20(2):669.
- Müller, M. F., Dralle, D. N., and Thompson, S. E. (2014). Analytical model for flow duration curves in seasonally dry climates. *Water Resources Research*, 50(7):5510–5531.
- Porcu, E., Alegría, A., and Furrer, R. (2018a). Modeling temporally evolving and spatially globally dependent data. *International Statistical Review*. To Appear.
- Porcu, E., Bevilacqua, M., and Genton, M. G. (2016). Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. *Journal of the American Statistical Association*, 111(514):888–898.
- Porcu, E., Castruccio, S., Alegría, A., and Crippa, P. (2018b). Axially symmetric models for global data: a journey between geostatistics and stochastic generators. *Submitted*.
- Prudnikov, A., A., B. Y., and OI, M. (1983). *Integrals and Series. Special Functions*.
- Scheuerer, M., Schlather, M., and Schaback, R. (2013). Interpolation of spatial data – a stochastic or a deterministic problem? *European Journal of Applied Mathematics*, 24:601–609.
- Schoenberg, I. J. (1942). Positive definite functions on spheres. *Duke Math. Journal*, 9:96–108.
- Skorokhod, A. V. and Yadrenko, M. I. (1973). On absolute continuity of measures corresponding to homogeneous Gaussian fields. *Theory of Probability and Its Applications*, 18:27–40.
- Soubeyrand, S., Enjalbert, J., and Sache, I. (2008). Accounting for roughness of circular processes: using Gaussian random processes to model the anisotropic spread of airborne plant disease. *Theoretical Population Biology*, 73:92–103.
- Stein, M. L. (1999). *Statistical Interpolation of Spatial Data: Some Theory for Kriging*. Springer, New York.
- Stein, M. L. (2007). Spatial variation of total column ozone on a global scale. *Annals of Applied Statistics*, 1:191–210.
- Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99:250–261.
- Ziegel, J. (2014). Convolution roots and differentiability of isotropic positive definite functions on spheres. *Proceedings of the American Mathematical Society*, 142:2053–2077.