

# Gaussian random fields on the product of spheres: Theory and applications

Alfredo Alegría\*

*Departamento de Matemática, Universidad Técnica Federico Santa María, Chile*  
e-mail: [alfredo.alegria@usm.cl](mailto:alfredo.alegria@usm.cl)

Galatia Cleanthous

*Department of Mathematics and Statistics, National University of Ireland, Maynooth*  
e-mail: [galatia.cleanthous@mu.ie](mailto:galatia.cleanthous@mu.ie)

Athanasios G. Georgiadis

*School of Computer Science and Statistics, Trinity College of Dublin, Ireland*  
e-mail: [georgiaa@tcd.ie](mailto:georgiaa@tcd.ie)

Emilio Porcu

*Department of Mathematics, Khalifa University, Abu Dhabi, United Arab Emirates*  
*ADIA Lab, Abu Dhabi, United Arab Emirates*  
e-mail: [emilio.porcu@ku.ac.ae](mailto:emilio.porcu@ku.ac.ae)

Philip A. White

*Department of Statistics, Brigham Young University, USA*  
e-mail: [pwhite@stat.byu.edu](mailto:pwhite@stat.byu.edu)

**Abstract:** We consider Gaussian random fields on the product of spheres. We study the regularity and the Hölder continuity of such random fields via their covariance function. Moreover, we approximate the Gaussian random fields using truncations of the Karhunen-Loève expansion and conduct simulation experiments to illustrate our approximation results. Using hourly wind speed and global space-time cloud cover datasets, we discuss modelling data in a Bayesian framework using Gaussian random fields over the product of spheres with covariance approximations through truncated series expansions.

**MSC2020 subject classifications:** Primary 60G60; secondary 62F15, 62P12, 60G15, 60G20, 53A35.

**Keywords and phrases:** Approximation, covariance kernel, directional data, Karhunen-Loève, Sobolev regularity, Hölder continuity, seasonal time series, simulations.

Received October 2023.

---

\*Corresponding author

## 1. Introduction

This paper aims to provide some basic theory for random fields defined continuously over the product of hyperspheres, together with simulations and direct applications in environmental science.

A wealth of applications from applied science motivates interest in this geometry. The recent contribution by [Porcu and White \(2022\)](#) contains several examples where the statistical analysis of real-life datasets benefit from covariance models defined over such products. Specifically, the authors emphasize two situations. The former happens when detrending the random field to account for cyclic patterns. The latter happens when one embeds spatial or space-time domains into the product of circles; relevant applications in environmental, atmospheric, oceanographic, and earth sciences include datasets collected over a large section of the Earth where the measured variable exhibits a strong seasonality or is directionally dependent (e.g., the wind speed or the ocean current velocity). Depending on the setting, one or the other approach can be used.

As noted by [Emery et al. \(2022\)](#), complex sources of seasonality are crucial to space-time modeling, and this fact is witnessed by recent applications to climate studies (see [Emery et al., 2022](#), with the references therein). Capturing seasonality through a covariance function using time embedded in a circle requires distance calculations that depend on a seasonal period. However, the assumption of a fixed period can be relaxed by integrating models on the product of spheres with warping approaches. To this point, we refer to [Porcu et al. \(2020\)](#), with the references therein, as well as to [White and Porcu \(2019b\)](#). The idea of embedding time periodicities inside the covariance structure is not new. In particular, [Shirota and Gelfand \(2017a\)](#) has applied such methodologies to model crime events using log-Gaussian Cox processes. Continuous-time monitoring of ground-level ozone concentrations has instead been proposed by [White and Porcu \(2019a\)](#). In either case, a random field with an index set of  $\mathbb{S}^1 \times \mathbb{S}^2$  (product of the unit circle with the unit two-dimensional sphere) would be a natural choice to represent phenomena in  $\text{TIME} \times \text{SPACE}$  or  $\text{DIRECTION} \times \text{SPACE}$  ([Mastrantonio et al., 2016](#)).

We recognize alternative approaches to model space-time seasonality stochastically, allowing for unknown seasons. Among those, some approaches are discussed in [West and Harrison \(2006\)](#). [Held and Paul \(2012\)](#) extend classical time series approaches to model seasonality in spatiotemporal surveillance of infectious diseases. [Lanfredi et al. \(2020\)](#) take an approach based on space-time coherence to show that seasonality under the current climate can be synthesized in the form of a progressive deformation process of the annual cycle, which starts from the northernmost areas with maximum values in summer and ends in the south, where maximum values are recorded in winter. The role of seasonality in space-time prediction is instead discussed in [Atkinson et al. \(2003\)](#). Notable approaches address seasonality to achieve a better understanding of space-time variability can be found in [Celleri et al. \(2007\)](#).

A wealth of different approaches to seasonality in time series is available in the literature. To mention a few of those, we cite [West and Harrison \(2006\)](#),

Hylleberg (1992), and Franses (1991). Indeed, in some settings more traditional time series approaches would be preferable, but, as demonstrated in Porcu and White (2022), modelling the data using covariance functions over the product of spheres outperforms more traditional time series methods in some cases in terms of prediction. Moreover, the models presented herein can easily handle uneven time sampling.

Unlike in Porcu and White (2022), we emphasize that our primary target with this paper is not to outperform a set of well-established methods. Instead, in this paper, we focus on providing theoretical properties of random fields defined over this new geometry. In particular, we emphasize the fact that properties of a random field may be completely different in space and time. For instance, the spatial and temporal regularities (which shall be rigorously defined in subsequent sections) may differ substantially in real applications. Typically, regularity is expressed in terms of function spaces (see Lang and Schwab, 2015; Clarke et al., 2018; Cleanthous et al., 2020, 2021, for similar approaches in other geometries). In turn, it is well known that covariance functions are crucial to understanding continuity and regularity properties of Gaussian random fields (see Kerkycharian et al., 2018, and the references therein). Simply stated, the geometric properties of a Gaussian random field are closely related to the behaviour of its stationary covariance function in the neighborhood of the origin. For instance, mean square continuity and mean square differentiability have a one-to-one correspondence with, respectively, continuity and differentiability of the covariance function at the origin. In turn, the behaviour of a stationary covariance function at the origin has a one-to-one correspondence with the behaviour of its spectrum away from the origin. Thus, in this paper, we present regularity in terms of the spectral coefficients associated with a covariance function. Moreover, we explore the continuity properties of Gaussian random fields over the product of spheres.

Using harmonic analysis, we discuss approximations of Gaussian random fields and prove their theoretical convergence rates. We explore our theoretical results via simulations. Then, using two real datasets, we illustrate how a truncated spectral representation of the covariance function can be used for modelling. As a part of this, we describe how to incorporate seasonalities, directional quantities, or global characteristics into the second-order structure, as well as assessing model performance as function of the truncation point of the spectral representation.

We illustrate the use of Gaussian random fields on the product of spheres by discussing data modeling using covariance approximations through truncated series expansions. For our first illustration, we use hourly wind speed data that exhibit seasonal and directional autocorrelation. To account for the seasonal and directional covariance patterns, we model winds speeds using a Gaussian random field with covariance defined over the product of two circles. In our second example, we consider a global space-time cloud cover dataset with strong annual seasonality. To account for spatial and seasonal trends, we use a covariance model defined over the product of a sphere and a circle. For both datasets, we discuss modeling and model fitting in a Bayesian framework. We discuss

model performance as a function of the number of terms in the covariance approximation and present results for both data analyses.

Here, we summarize the main *contributions* of this work:

(a) We study the *regularity* of Gaussian random fields on the product of spheres, measured in terms of function spaces that count mixed-smoothness. Precisely, we prove necessary and sufficient conditions on the summability of the spectral coefficients so that the covariance kernel has certain (mixed) smoothness (see Theorem 3.3 and Corollary 3.4).

(b) We obtain the *Hölder continuity* of the covariance kernels under the summability of the spectral coefficients and Hölder bounds for the moments of the differences of Gaussian random fields (Theorems 4.2 and 4.3).

(c) We derive fast *approximations* of Gaussian random fields, by truncation of their Karhunen-Loève expansions, by assuming a specific decay for the spectrum (Theorem 5.1).

(d) The rates of convergence obtained in (c) are verified via *simulation* examples. In addition, simulated Gaussian random fields on the torus, for different spectral coefficients, are shown.

(e) We discuss *modeling data* using Gaussian random fields on the product of spheres through covariance approximations obtained by truncated spectral representations. Our first example considers directionally-indexed wind speed data with seasonal autocorrelation. In the second example, we present a global space-time dataset with strong seasonality.

As is customary in scientific research, we have developed a theory that draws inspiration from previous discoveries, particularly advancements in purely spherical geometries. In the context of the product of spheres, we have adapted these concepts while being mindful of the need to operate in product spaces, where all mathematical components require adjustments. In particular, our approach involves the manipulation of joint spectra and necessitates the utilization of bivariate mathematical techniques. Similar challenges to those encountered in the transition from space to space-time processes are prevalent in this work. For instance, the models examined throughout the manuscript are non-separable in their coordinates, meaning that covariance cannot be factorized into individual covariances, each defined on a single sphere. Additionally, notions of anisotropy (i.e., different scales in each coordinate) and mixed smoothness may arise, and, in general, the dynamics and statistical characteristics can vary across each sphere forming the product space.

The structure of the paper is as follows: In Section 2, we provide the general theory of isotropic random fields on the product of spheres. Sections 3 and 4 are dedicated to the (a) and (b) from above respectively. Items (c) and (d) are the objective of Section 5, while (e) is explored in Section 6. Some final remarks are presented in Section 7 and the proofs of all the above Theorems have been placed in the Appendix.

*Notation:* Before we proceed, we fix here some necessary notation. We denote by  $c$  positive constants that may vary at every occurrence. The dependence of a

constant to some parameter  $p$ , would be indicated by  $c_p$ . For two non-negative functions  $f$  and  $g$  we denote by  $f \sim g$  the existence of a global constant  $c \geq 1$  such that  $c^{-1}g \leq f \leq cg$ . The sets of positive and non-negative integers are denoted by  $\mathbb{N}$  and  $\mathbb{N}_0$  respectively. For every real number  $x$  we define

$$(x)_+ := \max(x, 0). \tag{1.1}$$

## 2. Gaussian random fields on the product of spheres

Let  $d_1, d_2$  be positive integers. We denote by  $\mathbb{T}^{d_1, d_2}$  the *product of spheres*

$$\mathbb{T}^{d_1, d_2} := \mathbb{S}^{d_1} \times \mathbb{S}^{d_2} = \{\mathbf{x} = (x_1, x_2) : x_i \in \mathbb{S}^{d_i}, i = 1, 2\},$$

where  $\mathbb{S}^{d_i} := \{x_i \in \mathbb{R}^{d_i+1} : \|x_i\|_{d_i+1} = 1\}$  denotes the  $d_i$ -dimensional unit sphere embedded in  $\mathbb{R}^{d_i+1}$  and  $\|\cdot\|_{d_i+1}$  is the Euclidean norm on  $\mathbb{R}^{d_i+1}$ ,  $i = 1, 2$ . Note that the classical torus  $\mathbb{T}$  is isomorphic with the product of two circles:  $\mathbb{T} \simeq \mathbb{S}^1 \times \mathbb{S}^1 = \mathbb{T}^{1,1}$ , is included in our study.

We consider a Gaussian random field  $\{Z(\mathbf{x}), \mathbf{x} \in \mathbb{T}^{d_1, d_2}\}$  on the product of spheres, which is assumed to be (i) real-valued and (ii) of zero-mean.

Its covariance function  $K_Z : \mathbb{T}^{d_1, d_2} \times \mathbb{T}^{d_1, d_2} \rightarrow \mathbb{R}$ , takes the form

$$K_Z(\mathbf{x}, \mathbf{y}) := \text{cov}(Z(\mathbf{x}), Z(\mathbf{y})) = \mathbb{E}(Z(\mathbf{x})Z(\mathbf{y})), \quad \mathbf{x}, \mathbf{y} \in \mathbb{T}^{d_1, d_2}. \tag{2.1}$$

Following [Guella et al. \(2015\)](#), we will say that a random field  $\{Z(\mathbf{x}), \mathbf{x} \in \mathbb{T}^{d_1, d_2}\}$  as above, is *isotropic* when there exists a continuous mapping  $K_{\text{iso}} : [-1, 1]^2 \rightarrow \mathbb{R}$ , such that for every  $\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2) \in \mathbb{T}^{d_1, d_2}$ ,

$$K_Z(\mathbf{x}, \mathbf{y}) = K_{\text{iso}}(\langle x_1, y_1 \rangle_1, \langle x_2, y_2 \rangle_2), \tag{2.2}$$

where  $\langle \cdot, \cdot \rangle_i$  denotes the classical dot product on  $\mathbb{R}^{d_i+1}$   $i = 1, 2$ . Note that for every  $x_i, y_i \in \mathbb{S}^{d_i}$ , the angular distance on  $\mathbb{S}^{d_i}$  is  $\rho_{\mathbb{S}^{d_i}}(x_i, y_i) = \arccos(\langle x_i, y_i \rangle_i)$ ,  $i = 1, 2$ . That is the covariance between  $Z(\mathbf{x})$  and  $Z(\mathbf{y})$ , depends only on the spherical distances  $\rho_{\mathbb{S}^{d_i}}(x_i, y_i)$ ,  $i = 1, 2$ .

**Remark 2.1.** *We list a few remarks on the isotropy and covariance.*

(a) *It is well-known the covariance function is positive definite. Moreover, the regularity of the Gaussian random field depends heavily on the corresponding regularity of the covariance function; see [Kerkyacharian et al. \(2018\)](#).*

(b) *For any isotropic random field  $Z(\mathbf{x})$  on the product of spheres, the “marginal” random fields  $Z(\cdot, x_2)$  (for a fixed  $x_2$ ) and  $Z(x_1, \cdot)$  (for a fixed  $x_1$ ) are isotropic random fields on  $\mathbb{S}^{d_1}$  and  $\mathbb{S}^{d_2}$ , respectively.*

(c) *An isotropic random field is of constant variance;  $\text{Var}(Z(\mathbf{x})) = \mathbb{E}(Z(\mathbf{x})^2) = K_{\text{iso}}(1, 1) \in [0, \infty)$ , for every  $\mathbf{x} \in \mathbb{T}^{d_1, d_2}$ .*

The functions  $K_{\text{iso}}$  above can be expressed uniquely by the following series expansion ([Guella et al., 2015](#))

$$K_{\text{iso}}(t_1, t_2) = \sum_{\mathbf{k}=(k_1, k_2) \in \mathbb{N}_0^2} b_{\mathbf{k}} C_{k_1}^{\frac{d_1-1}{2}}(t_1) C_{k_2}^{\frac{d_2-1}{2}}(t_2), \quad t_1, t_2 \in [-1, 1], \tag{2.3}$$

where the coefficients  $b_{\mathbf{k}}$  are non-negative and satisfy

$$\sum_{\mathbf{k} \in \mathbb{N}_0^2} b_{\mathbf{k}} C_{k_1}^{\frac{d_1-1}{2}}(1) C_{k_2}^{\frac{d_2-1}{2}}(1) < \infty, \tag{2.4}$$

where we denoted by  $C_k^\lambda$  the Gegenbauer polynomial of degree  $k$  and order  $\lambda > 0$  and by  $C_k^0$  the Chebyshev polynomials (Szegő, 1939) and  $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ .

Properties of orthogonal polynomials, necessary for our study, have been deferred to the Appendix. At this point we just recall that Chebyshev and Gegenbauer polynomials maximize at 1. In particular, the Chebyshev polynomials satisfy  $|C_k^0(t)| \leq C_k^0(1) = 1$ , for every  $k \in \mathbb{N}_0$  and  $t \in [-1, 1]$ . Also, for every  $d \geq 2$ , see equation (4.7.3) in (Szegő, 1939)

$$\|C_k^{\frac{d-1}{2}}\|_\infty = C_k^{\frac{d-1}{2}}(1) = \binom{k+d-2}{k} \sim (k+1)^{d-2}, \quad \text{for every } k \in \mathbb{N}_0. \tag{2.5}$$

Consequently

$$\|C_k^{\frac{d-1}{2}}\|_\infty = C_k^{\frac{d-1}{2}}(1) \sim (k+1)^{(d-2)_+}, \quad \text{for every } k \in \mathbb{N}_0, d \in \mathbb{N}, \tag{2.6}$$

where  $(d-2)_+$  is as in (1.1).

Therefore the convergence in (2.3) is uniform on the cube  $[-1, 1]^2$ , thanks to the two-dimensional  $M$ -Weirstrass criterion. Finally, the convergence in (2.4) takes the equivalent form

$$\sum_{\mathbf{k} \in \mathbb{N}_0^2} b_{\mathbf{k}} (k_1 + 1)^{(d_1-2)_+} (k_2 + 1)^{(d_2-2)_+} < \infty. \tag{2.7}$$

### 3. Regularity properties

Regularity properties of random objects are of fundamental importance in their estimation and/or approximation. Regularity spaces have been employed for decades in density estimation, regression, image compression, signal processing, neural networks and many more. Our first target is to explore the regularity of Gaussian random fields on the product of spheres, and this will be achieved by studying their covariance function, in the spirit of Lang and Schwab (2015).

#### 3.1. Sobolev spaces

In the framework of the product of spheres, the domain of the covariance kernels is the square  $[-1, 1]^2$ ; see (2.3). Therefore we shall work with the following class of weighted square integrable functions:

We define the weighted Lebesgue space  $L^2_{(d_1, d_2)} := L^2_{(d_1, d_2)}([-1, 1]^2)$ ,  $d_1, d_2 \in \mathbb{N}$ , as the class of all functions  $u : [-1, 1]^2 \rightarrow \mathbb{R}$  such that

$$\|u\|_{L^2_{(d_1, d_2)}}^2 := \int_{-1}^1 \int_{-1}^1 |u(t_1, t_2)|^2 (1-t_1^2)^{\frac{d_1-2}{2}} (1-t_2^2)^{\frac{d_2-2}{2}} dt_1 dt_2 < \infty. \tag{3.1}$$

Every function  $u \in L^2_{(d_1, d_2)}$  can be uniquely represented as

$$u = \sum_{\mathbf{k} \in \mathbb{N}_0^2} u_{\mathbf{k}} C_{\mathbf{k}}^{d_1, d_2} \quad (\text{convergence in } L^2_{(d_1, d_2)}), \quad (3.2)$$

where

$$C_{\mathbf{k}}^{d_1, d_2}(t_1, t_2) := C_{k_1}^{\frac{d_1-1}{2}}(t_1) C_{k_2}^{\frac{d_2-1}{2}}(t_2), \quad (3.3)$$

for every  $\mathbf{k} = (k_1, k_2) \in \mathbb{N}_0^2$  and  $t_1, t_2 \in [-1, 1]$  (more details may be found in the Appendix). The sequence of coefficients  $\{u_{\mathbf{k}}\}_{\mathbf{k} \in \mathbb{N}_0^2}$  is referred as the *spectrum* or the *Fourier coefficients* of the function  $u$ .

Since our functions are defined on the product domain  $[-1, 1] \times [-1, 1]$ , the two different variables may enjoy completely different regularities. This can be easily understood since the phenomenon under study may behave completely differently in space or direction and time. For this purpose we will study not only *classical Sobolev spaces*, but also *anisotropic Sobolev spaces* and *Sobolev spaces with dominating mixed smoothness*. We proceed to their definitions, which involve partial derivatives and weighted Lebesgue norms as in (3.1).

**Definition 3.1.** Let  $u : [-1, 1]^2 \rightarrow \mathbb{R}$  and some non-negative integers  $N, n$  and  $m$ . We say that  $u$  belongs to the

(i) *(Classical) Weighted Sobolev space*  $W^N := W^N([-1, 1]^2)$  when

$$\|u\|_{W^N}^2 := \|u\|_{L^2_{(d_1, d_2)}}^2 + \sum_{n_1+n_2=N} \|\partial^{(n_1, n_2)} u\|_{L^2_{(d_1+2n_1, d_2+2n_2)}}^2 < \infty. \quad (3.4)$$

(ii) *Weighted Sobolev space with dominating mixed smoothness*  $DW^{(n, m)} := DW^{(n, m)}([-1, 1]^2)$  when

$$\begin{aligned} \|u\|_{DW^{(n, m)}}^2 &:= \|u\|_{L^2_{(d_1, d_2)}}^2 + \|\partial_1^n u\|_{L^2_{(d_1+2n, d_2)}}^2 \\ &\quad + \|\partial_2^m u\|_{L^2_{(d_1, d_2+2m)}}^2 + \|\partial^{(n, m)} u\|_{L^2_{(d_1+2n, d_2+2m)}}^2 < \infty. \end{aligned} \quad (3.5)$$

(iii) *Anisotropic weighted Sobolev space*  $W^{(n, m)} := W^{(n, m)}([-1, 1]^2)$  when

$$\|u\|_{W^{(n, m)}}^2 := \|u\|_{L^2_{(d_1, d_2)}}^2 + \|\partial_1^n u\|_{L^2_{(d_1+2n, d_2)}}^2 + \|\partial_2^m u\|_{L^2_{(d_1, d_2+2m)}}^2 < \infty. \quad (3.6)$$

Note that in the above definition, which is inspired by (Schmeisser and Triebel, 1987) and (Xu, 2018), all the derivatives should be understood in the weak (distributional) sense. For this notion and for further information on Sobolev spaces, we refer the reader to the book Adams and Fournier (2003).

A classical property of Sobolev spaces (see Adams and Fournier, 2003; Schmeisser and Triebel, 1987) is

$$\|u\|_{W^N}^2 \sim \sum_{n_1+n_2 \leq N} \|\partial^{(n_1, n_2)} u\|_{L^2_{(d_1+2n_1, d_2+2n_2)}}^2, \quad (3.7)$$

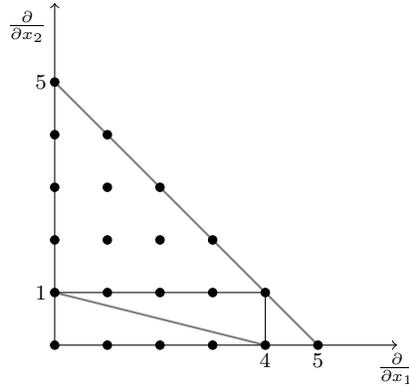


FIG 3.1.  $W^5$ : All the derivatives of order  $\leq 5$ , so all the vertices inside the triangle with edges  $(0,0)$ ,  $(0,5)$  and  $(5,0)$ .  $DW^{(4,1)}$ : The vertices in the area of the orthogonal  $(0,0)$ ,  $(0,1)$ ,  $(4,1)$  and  $(4,0)$ .  $W^{(4,1)}$ : The vertices in the area of the triangle  $(0,0)$ ,  $(0,1)$  and  $(4,0)$ .

which gives the following inclusions between the three spaces:

$$W^{n+m} \subset DW^{(n,m)} \subset W^{(n,m)}, \quad \text{for every } n, m \in \mathbb{N}_0. \tag{3.8}$$

What makes these spaces different is the way that they count the regularity of a function. In Figure 3.1 we indicate the derivatives that are counted for the spaces in a specific example.

### 3.2. Sequence spaces

We proceed by introducing the following sequence spaces which will be crucial in our study. We will extensively use the following function

$$\tau(d) := \begin{cases} d - 3, & d > 1 \\ 0, & d = 1 \end{cases}, \tag{3.9}$$

which naturally appears because the Chebyshev polynomials —that correspond to the coordinate sphere  $S^1$  (i.e. the circle)— present different behaviour than the Gegenbauer polynomials  $C_k^{\frac{d-1}{2}}$ , for  $d > 1$ ; see Section 8.1.

**Definition 3.2.** Let  $n, m \in \mathbb{N}_0$ . Then  $\ell^{(n,m)} := \ell_{(d_1, d_2)}^{(n,m)}$  is defined as the class of all sequences  $b_{\mathbf{k}} : \mathbb{N}_0^2 \rightarrow \mathbb{R}$  such that.

$$\sum_{\mathbf{k}=(k_1, k_2) \in \mathbb{N}_0^2} b_{\mathbf{k}}^2 (k_1 + 1)^{\tau(d_1)+2n} (k_2 + 1)^{\tau(d_2)+2m} < \infty. \tag{3.10}$$

### 3.3. Equivalence of kernel regularity and summability of spectrum coefficients

We are now ready to present the main result of this section. This is a necessary and sufficient condition for a covariance kernel  $K_{\text{iso}}$  to belong in a regularity space, in terms of the coefficients  $u_{\mathbf{k}}$ . This follows as a consequence of a very broad result that we present first.

**Theorem 3.3.** *Let  $u \in L^2_{(d_1, d_2)}$  and  $N, n, m \in \mathbb{N}_0$ . Then,  $u$  belongs (either) to the spaces  $W^N$ ,  $DW^{(n, m)}$  or  $W^{(n, m)}$  if and only if, respectively, the sequence  $\{u_{\mathbf{k}}\}$  of Fourier coefficients in the expansion (3.2), belongs to*

(i) *the space  $\bigcap_{\nu=0}^N \ell^{(\nu, N-\nu)}$  and precisely*

$$\|u\|_{W^N}^2 \sim \sum_{\nu=0}^N \sum_{\mathbf{k} \in \mathbb{N}_0^2} u_{\mathbf{k}}^2 (k_1 + 1)^{2\nu + \tau(d_1)} (k_2 + 1)^{2(N-\nu) + \tau(d_2)}; \quad (3.11)$$

(ii) *the space  $\ell^{(n, m)}$  and precisely*

$$\|u\|_{DW^{(n, m)}}^2 \sim \sum_{\mathbf{k} \in \mathbb{N}_0^2} u_{\mathbf{k}}^2 (k_1 + 1)^{2n + \tau(d_1)} (k_2 + 1)^{2m + \tau(d_2)}; \quad (3.12)$$

(iii) *the space  $\ell^{(n, 0)} \cap \ell^{(0, m)}$  and precisely*

$$\begin{aligned} \|u\|_{W^{(n, m)}}^2 &\sim \sum_{\mathbf{k} \in \mathbb{N}_0^2} u_{\mathbf{k}}^2 (k_1 + 1)^{2n + \tau(d_1)} (k_2 + 1)^{\tau(d_2)} \\ &\quad + \sum_{\mathbf{k} \in \mathbb{N}_0^2} u_{\mathbf{k}}^2 (k_1 + 1)^{\tau(d_1)} (k_2 + 1)^{2m + \tau(d_2)}, \end{aligned} \quad (3.13)$$

where the function  $\tau$  is defined in (3.9).

The result above narrows as follows for kernels  $K_{\text{iso}} : [-1, 1] \times [-1, 1] \rightarrow \mathbb{R}$  of the form (2.3).

**Corollary 3.4.** *Let  $\{Z(\mathbf{x}) : \mathbf{x} \in \mathbb{T}^{d_1, d_2}\}$  be an isotropic random field on  $\mathbb{T}^{d_1, d_2}$ , with covariance kernel  $K_{\text{iso}}$  be as in (2.3) and  $n, m \in \mathbb{N}_0$ . Then, the function*

$$(1 - t_1^2)^{\frac{n}{2}} (1 - t_2^2)^{\frac{m}{2}} \frac{\partial^n}{\partial t_1^n} \frac{\partial^m}{\partial t_2^m} K_{\text{iso}}(t_1, t_2), \quad (t_1, t_2) \in [-1, 1]^2,$$

belongs to  $L^2_{(d_1, d_2)}$  if and only if the sequence

$$\{b_{\mathbf{k}} (k_1 + 1)^{(2n + \tau(d_1))/2} (k_2 + 1)^{(2m + \tau(d_2))/2}\}_{\mathbf{k} \in \mathbb{N}_0^2},$$

with  $b_{\mathbf{k}}$  determined from (2.3), belongs to  $\ell^2(\mathbb{N}_0^2)$ .

To streamline the presentation of the study, the proofs of Theorem 3.3 and Corollary 3.4 are provided in Section 8.2, following all the necessary background and some auxiliary lemmata.

### 4. Hölder continuity

Hölder continuity asserts that short variations on the index domain, imply short changes in the values of the function. In this section we study this property for random fields on the product of spheres, in the spirit of [Cleanthous et al. \(2020\)](#) and [Lang and Schwab \(2015\)](#). We start by obtaining such properties on the covariance kernel and we transfer the property on the original random field.

As was the case for Sobolev spaces, we have different norms for measuring the continuity. We use classical and mixed Hölder spaces.

**Definition 4.1.** *Let  $u : [-1, 1]^2 \rightarrow \mathbb{R}$  and  $\delta, \delta_1, \delta_2 \in (0, 1]$ . We say that  $u$  belongs to the*

(i) *(Classical) Hölder space  $\mathcal{H}^\delta := \mathcal{H}^\delta([-1, 1]^2)$  when*

$$\|u\|_{\mathcal{H}^\delta} := \sup_{t \neq s} \frac{|u(t) - u(s)|}{|t - s|^\delta} < \infty, \tag{4.1}$$

where  $|r| = \sqrt{r_1^2 + r_2^2}$ , for every  $r = (r_1, r_2) \in [-1, 1]^2$ . A function  $u \in \mathcal{H}^\delta$  is used to be referred as Hölder continuous of order  $\delta$ .

(ii) *Mixed Hölder space  $\mathcal{H}^{(\delta_1, \delta_2)} := \mathcal{H}^{(\delta_1, \delta_2)}([-1, 1]^2)$  when*

$$\begin{aligned} \|u\|_{\mathcal{H}^{(\delta_1, \delta_2)}} := & \sup_{t_2 \in [-1, 1]} \sup_{t_1 \neq s_1} \frac{|u(t_1, t_2) - u(s_1, t_2)|}{|t_1 - s_1|^{\delta_1}} \\ & + \sup_{t_1 \in [-1, 1]} \sup_{t_2 \neq s_2} \frac{|u(t_1, t_2) - u(t_1, s_2)|}{|t_2 - s_2|^{\delta_2}} < \infty. \end{aligned} \tag{4.2}$$

We will say that the function  $u \in \mathcal{H}^{(\delta_1, \delta_2)}$  is Hölder continuous of order  $(\delta_1, \delta_2)$ .

We present here the following result containing conditions on the spectrum, sufficient for the kernel  $K_{iso}$  to be Hölder continuous. This time we need the  $\ell^1$ -summability of the spectrum. The proof can be found in [Section 8.3](#).

**Theorem 4.2.** *Let  $Z$  be an isotropic random field on  $\mathbb{T}^{d_1, d_2}$  and let  $K_{iso}$  be as in [\(2.3\)](#). Let  $\delta, \delta_1, \delta_2 \in (0, 1]$ .*

(i) *If*

$$\sum_{\mathbf{k} \in \mathbb{N}_0^2} b_{\mathbf{k}} (k_1 + 1)^{(d_1 - 2)_+} (k_2 + 1)^{(d_2 - 2)_+} ((k_1 + 1)^2 + (k_2 + 1)^2)^\delta < \infty, \tag{4.3}$$

then  $K_{iso} \in \mathcal{H}^\delta$ .

(ii) *If*

$$\begin{aligned} & \sum_{\mathbf{k} \in \mathbb{N}_0^2} b_{\mathbf{k}} (k_1 + 1)^{(d_1 - 2)_+ + 2\delta_1} (k_2 + 1)^{(d_2 - 2)_+} \\ & + \sum_{\mathbf{k} \in \mathbb{N}_0^2} b_{\mathbf{k}} (k_1 + 1)^{(d_1 - 2)_+} (k_2 + 1)^{(d_2 - 2)_+ + 2\delta_2} < \infty, \end{aligned} \tag{4.4}$$

then  $K_{iso} \in \mathcal{H}^{(\delta_1, \delta_2)}$ .

### 4.1. Moments of $|Z(\mathbf{x}) - Z(\mathbf{y})|$

We use Theorem 4.2 to obtain Hölder bounds for the  $p$ -moments of the difference  $|Z(\mathbf{x}) - Z(\mathbf{y})|$ , in terms of the distance on the product of spheres, which is defined as follows: for  $\mathbf{x} = (x_1, x_2)$  and  $\mathbf{y} = (y_1, y_2)$  in  $\mathbb{T}^{d_1, d_2}$ , we set

$$\rho(\mathbf{x}, \mathbf{y}) := \sqrt{\rho_{\mathbb{S}^{d_1}}^2(x_1, y_1) + \rho_{\mathbb{S}^{d_2}}^2(x_2, y_2)},$$

where  $\rho_{\mathbb{S}^{d_i}} = \arccos(\langle x_i, y_i \rangle_i)$  is the standard metric in  $\mathbb{S}^{d_i}$ , for  $i = 1, 2$ .

Of course, other choices (like the maximum or the sum of  $\rho_{\mathbb{S}^{d_i}}$ ) are equivalent metrics with  $\rho$ , as it holds for the Cartesian product of metric spaces in general (Folland, 2009, page 13). We present here the next theorem, and its proof can be found in Section 8.3:

**Theorem 4.3.** *Let  $Z$  be an isotropic Gaussian random field on  $\mathbb{T}^{d_1, d_2}$ , with covariance isotropic kernel function  $K_{iso}$  as in (2.3), and assume that the sequence  $\{b_{\mathbf{k}}\}_{\mathbf{k}}$  satisfies (4.3) for some  $\delta \in (0, 1]$ . Then, for every  $p \in \mathbb{N}$ , there exists a constant  $c = c_{\delta, p} > 0$  such that for every  $\mathbf{x}, \mathbf{y} \in \mathbb{T}^{d_1, d_2}$*

$$\mathbb{E}(|Z(\mathbf{x}) - Z(\mathbf{y})|^{2p}) \leq c\rho(\mathbf{x}, \mathbf{y})^{2p\delta}. \tag{4.5}$$

## 5. Approximation of Gaussian random fields on $\mathbb{T}^{d_1, d_2}$

In this section we study Gaussian random fields on the product of spheres with a Karhunen-Loève type expansion, with covariance satisfying (2.3). We also provide an approximation method for such random fields and study its level of accuracy. The theoretical findings are illustrated by simulations.

### 5.1. Harmonic analysis on the sphere $\mathbb{S}^d$

We recall some standard background on analysis on the sphere; see for example Marinucci and Peccati (2011) and Yadrenko (1983).

Let  $\Delta$  be the Laplace-Beltrami operator defined over  $\mathbb{S}^d$ . The spectrum of  $\Delta$  is discrete, real and non-positive, with eigenvalues given by  $\lambda_k = -k(k + d - 1)$  for  $k \geq 0$  (Szegő, 1939). Let  $\mathcal{H}_k$  be the eigenspace that corresponds to the eigenvalue  $\lambda_k$ , then (see Yadrenko (1983))

$$L^2(\mathbb{S}^d) = \bigoplus_{k=0}^{\infty} \mathcal{H}_k.$$

Let  $\{S_{k, \ell}^d : k \in \mathbb{N}_0, \ell = 1, 2, \dots, D_k(d)\}$  be an orthonormal basis of  $\mathcal{H}_k$  whose dimension,  $D_k(d)$ , equals  $D_k(1) = 2$  and

$$D_k(d) = \dim \mathcal{H}_k = (2k + d - 1) \frac{(k + d - 2)!}{k!(d - 1)!}, \quad d \geq 2. \tag{5.1}$$

From (5.1) it becomes apparent that there exists a constant  $c_d \geq 1$  such that

$$c_d^{-1}(k+1)^{d-1} \leq D_k(d) \leq c_d(k+1)^{d-1}. \tag{5.2}$$

The last can be easily verified for  $d \leq 2$ , since  $D_k(2) = 2k + 1$ . For  $d \geq 3$ , we have  $2(k+1) \leq 2k + d - 1 \leq (d-1)(k+1)$ . Also  $\frac{(k+d-2)!}{k!} = \prod_{j=1}^{d-2} (k+j)!$ , therefore

$$(k+1)^{d-2} \leq \prod_{j=1}^{d-2} (k+j)! \leq (k+(d-2))^{d-2} \leq (d-2)^{d-2} (k+1)^{d-2},$$

which completes the argument.

The functions  $S_{k,\ell}^d$  are the well-known spherical harmonics on  $\mathbb{S}^d$ , and they satisfy the following addition formula; [Yadrenko \(1983\)](#) page 72:

$$\sum_{\ell=1}^{D_k(d)} S_{k,\ell}^d(x) \overline{S_{k,\ell}^d(y)} = \frac{D_k(d)}{\omega_d} c_k(d; \langle x, y \rangle) \tag{5.3}$$

where  $\omega_d = 2\pi^{(d+1)/2} / \Gamma((d+1)/2)$  is the total area of  $\mathbb{S}^d$ , with

$$c_k(d; r) = \frac{C_k^{\frac{d-1}{2}}(r)}{C_k^{\frac{d-1}{2}}(1)}, \quad r \in [-1, 1].$$

**5.2. Gaussian random fields on the product of spheres**

We consider real valued, Gaussian random fields satisfying the following expansion of Karhunen-Loève type:

$$Z(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{N}_0^2} \sum_{m_1=1}^{D_{k_1}(d_1)} \sum_{m_2=1}^{D_{k_2}(d_2)} \alpha_{\mathbf{k},\mathbf{m}} S_{k_1,m_1}^{d_1}(x_1) S_{k_2,m_2}^{d_2}(x_2), \tag{5.4}$$

for every  $\mathbf{x} = (x_1, x_2) \in \mathbb{T}^{d_1, d_2}$ , where  $\alpha_{\mathbf{k},\mathbf{m}}$  is a sequence of random variables satisfying

$$\mathbb{E}(\alpha_{\mathbf{k},\mathbf{m}}) = 0, \tag{5.5}$$

$$\mathbb{E}(\alpha_{\mathbf{k},\mathbf{m}} \overline{\alpha_{\mathbf{k}',\mathbf{m}'}}) = \delta_{k_1,k_1'} \delta_{k_2,k_2'} \delta_{m_1,m_1'} \delta_{m_2,m_2'} B_{\mathbf{k}}, \tag{5.6}$$

for some sequence  $(B_{\mathbf{k}})_{\mathbf{k} \in \mathbb{N}_0^2}$  of non-negative numbers, which we refer to as the angular power spectrum and

$$\sum_{\mathbf{k} \in \mathbb{N}_0^2} B_{\mathbf{k}} \frac{D_{k_1}(d_1)}{\omega_{d_1}} \frac{D_{k_2}(d_2)}{\omega_{d_2}} < \infty. \tag{5.7}$$

Combining the addition formula (5.3) for the spherical harmonics together with relations (5.5) and (5.6), we see that the random field is isotropic and

$K_Z(\mathbf{x}, \mathbf{y}) = K_{\text{iso}}(\langle x_1, y_1 \rangle_1, \langle x_2, y_2 \rangle_2)$ , for every  $\mathbf{x} = (x_1, x_2)$ ,  $\mathbf{y} = (y_1, y_2) \in \mathbb{T}^{d_1, d_2}$ , where

$$K_{\text{iso}}(t_1, t_2) = \sum_{\mathbf{k} \in \mathbb{N}_0^2} B_{\mathbf{k}} \frac{D_{k_1}(d_1)}{\omega_{d_1}} \frac{D_{k_2}(d_2)}{\omega_{d_2}} c_{k_1}(d_1; t_1) c_{k_2}(d_2; t_2). \tag{5.8}$$

It is apparent by (2.3) and (5.8) that  $B_{\mathbf{k}}$  and  $b_{\mathbf{k}}$  are related through the identity

$$B_{\mathbf{k}} = b_{\mathbf{k}} \frac{\omega_{d_1}}{D_{k_1}(d_1)} \frac{\omega_{d_2}}{D_{k_2}(d_2)} C_{k_1}^{\frac{d_1-1}{2}}(1) C_{k_2}^{\frac{d_2-1}{2}}(1). \tag{5.9}$$

We will study the approximation of random fields as in (5.4). For doing this, let  $\{X_{\mathbf{k}, \mathbf{m}} : \mathbf{k} \in \mathbb{N}_0^2, m_i = 1, \dots, D_{k_i}(d_i), i = 1, 2\}$  be the sequence of independent, standard normally distributed random variables given by  $X_{\mathbf{k}, \mathbf{m}} = \frac{\alpha_{\mathbf{k}, \mathbf{m}}}{\sqrt{B_{\mathbf{k}}}}$ . The random field (5.4) can be rewritten as

$$Z(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{N}_0^2} \sqrt{B_{\mathbf{k}}} \sum_{m_1=1}^{D_{k_1}(d_1)} \sum_{m_2=1}^{D_{k_2}(d_2)} X_{\mathbf{k}, \mathbf{m}} S_{k_1, m_1}^{d_1}(x_1) S_{k_2, m_2}^{d_2}(x_2).$$

We introduce now the following random fields which consist of a truncated version of the above series, by summing only a certain finite number of their first terms. For  $N \in \mathbb{N}$ , we set

$$Z^N(\mathbf{x}) := \sum_{|\mathbf{k}| \leq N} \sqrt{B_{\mathbf{k}}} \sum_{m_1=1}^{D_{k_1}(d_1)} \sum_{m_2=1}^{D_{k_2}(d_2)} X_{\mathbf{k}, \mathbf{m}} S_{k_1, m_1}^{d_1}(x_1) S_{k_2, m_2}^{d_2}(x_2), \tag{5.10}$$

which is the truncated version of the expansion (5.4) of the initial random field  $Z$ .

The following theorem characterizes the rate of convergence of the error between  $Z$  and  $Z^N$  in  $L^p$  and  $\mathbb{P}$ -almost sure sense in terms of the angular power spectrum. For this we need  $\ell^\infty$ -conditions (decay for the angular power spectrum).

**Theorem 5.1.** *Let  $Z$  be a Gaussian random field on  $\mathbb{T}^{d_1, d_2}$  given by (5.4). Assume that the angular power spectrum  $(B_{\mathbf{k}})_{\mathbf{k} \in \mathbb{N}_0^2}$  decays as follows: there exist  $\varepsilon > 0$ ,  $c_* > 0$  and  $N_0 \in \mathbb{N}$ , such that*

$$B_{\mathbf{k}} \leq c_* (1 + |\mathbf{k}|^2)^{-\varepsilon - \frac{d_1 + d_2}{2}}, \quad \text{for all } |\mathbf{k}| \geq N_0. \tag{5.11}$$

*Then the series of approximate random fields  $\{Z^N\}_{N \in \mathbb{N}}$  converges to the random field  $Z$  in the following senses:*

*(i) in  $L^2(\Omega, L^2(\mathbb{T}^{d_1, d_2}))$  and precisely there exists a constant  $c = c(d_1, d_2, p) > 0$  such that the truncation's error to be bounded by*

$$\|Z - Z^N\|_{L^2(\Omega, L^2(\mathbb{T}^{d_1, d_2}))} \leq c \sqrt{\frac{c_*}{\varepsilon}} N^{-\varepsilon}, \quad \text{for every } N > N_0; \tag{5.12}$$

(ii) in  $L^p(\Omega, L^2(\mathbb{T}^{d_1, d_2}))$  for every  $p > 0$  and precisely there exists a constant  $c = c(d_1, d_2, p) > 0$  such that the truncation's error to be bounded by

$$\|Z - Z^N\|_{L^p(\Omega, L^2(\mathbb{T}^{d_1, d_2}))} \leq c \sqrt{\frac{c_*}{\varepsilon}} N^{-\varepsilon}, \quad \text{for every } N > N_0. \quad (5.13)$$

(iii)  $\mathbb{P}$ -almost surely and the truncation's error is asymptotically bounded by

$$\|Z - Z^N\|_{L^2(\mathbb{T}^{d_1, d_2})} \leq N^{-\gamma}, \quad \mathbb{P}\text{-a.s.}, \quad \text{for every } \gamma < \varepsilon. \quad (5.14)$$

For streamlining the presentation of the article, we postpone the proof until the Appendix.

**Remark 5.2.** *Some remarks on the above theorem are listed here.*

(a) *The relation between  $B_{\mathbf{k}}$  and  $b_{\mathbf{k}}$ , by combining relations (5.9), (5.2) and (2.6), is as follows:*

$$B_{\mathbf{k}} \sim b_{\mathbf{k}} \prod_{i=1}^2 (k_i + 1)^{(d_i - 2)_+ - d_i + 1}. \quad (5.15)$$

*For later use we mention that the sufficient condition on the decay of  $b_{\mathbf{k}}$ 's for the case of  $\mathbb{T}^{1,1}$  takes the form:*

$$b_{\mathbf{k}} \leq c(1 + |\mathbf{k}|^2)^{-\delta}, \quad \text{for some } \delta > 1. \quad (5.16)$$

(b) *Similar truncated approximations have been obtained on the sphere [Lang and Schwab \(2015\)](#), [Alegria et al. \(2021\)](#), the ball [Cleanthous \(2023\)](#), compact two-point homogeneous spaces [Cleanthous et al. \(2020\)](#), and their products with real numbers [Cleanthous et al. \(2021\)](#). The corresponding series in the above papers are series of some index  $k$  from 0 to  $\infty$  and can be truncated by just limiting  $k$  by some natural number  $N$ , which we later send to infinity. Here we need to truncate a series of  $\mathbf{k} \in \mathbb{N}_0^2$ . Therefore, we truncate in the above ‘‘polar’’ way, stopping  $|\mathbf{k}| \leq N$ . The proof will of course follows the spirit of the existing works, but technically it requires bi-variate treatments.*

(c) *The above (quasi-)norms  $L^p(\Omega, L^2(\mathbb{T}^{d_1, d_2}))$  are the natural measures for the expected risk, expressed in the case of product of spheres. For reasons of completeness and brevity, we present them in the Appendix (8.41).*

### 5.3. Simulations

We illustrate the theoretical developments using simulations. This section focuses on the special case of random fields  $Z(\mathbf{x})$  indexed by  $\mathbf{x} = (x_1, x_2) \in \mathbb{T}^{1,1}$ . Note that both  $x_1$  and  $x_2$  are fully characterized by angles  $\vartheta_1$  and  $\vartheta_2$  in  $[0, 2\pi)$ , respectively. Hence, here we are using an abuse of notation  $Z(\boldsymbol{\vartheta})$ , where  $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2)$ . The Karhunen-Loève expansion of  $Z(\boldsymbol{\vartheta})$  is given by a double Fourier

series, which is the natural extension of the representation of a field on a circle (see, e.g., Equation (14) in [Yadrenko, 1983](#), page 74):

$$Z(\boldsymbol{\vartheta}) = \sum_{\mathbf{k} \in \mathbb{Z}^2} \alpha_{\mathbf{k}} \exp \{i \mathbf{k}^\top \boldsymbol{\vartheta}\}, \tag{5.17}$$

where  $i$  is the complex unit and  $\alpha_{\mathbf{k}}$  is a sequence of random coefficients. In order to obtain a real-valued random field, we formulate the following conditions on the coefficients: (C1)  $\alpha_{\mathbf{0}}$  is real valued; (C2)  $\alpha_{-\mathbf{k}} = \overline{\alpha_{\mathbf{k}}}$ . We then write (5.17) in the following manner:

$$\begin{aligned} Z(\boldsymbol{\vartheta}) = & \alpha_{\mathbf{0}} + \sum_{k_1 \in \mathbb{Z} \setminus \{0\}} \alpha_{k_1,0} \exp\{ik_1\vartheta_1\} + \sum_{k_2 \in \mathbb{Z} \setminus \{0\}} \alpha_{0,k_2} \exp\{ik_2\vartheta_2\} \\ & + \sum_{\mathbf{k} \in \mathbb{N}^2} \alpha_{\mathbf{k}} \exp \{i \mathbf{k}^\top \boldsymbol{\vartheta}\} + \sum_{-\mathbf{k} \in \mathbb{N}^2} \alpha_{\mathbf{k}} \exp \{i \mathbf{k}^\top \boldsymbol{\vartheta}\} \\ & + \sum_{(-k_1,k_2) \in \mathbb{N}^2} \alpha_{\mathbf{k}} \exp \{i \mathbf{k}^\top \boldsymbol{\vartheta}\} + \sum_{(k_1,-k_2) \in \mathbb{N}^2} \alpha_{\mathbf{k}} \exp \{i \mathbf{k}^\top \boldsymbol{\vartheta}\} \end{aligned}$$

Under conditions (C1)-(C2), and using the fact that  $ab + \overline{a}b = 2\text{Re}(ab) = 2(\text{Re}(a)\text{Re}(b) - \text{Im}(a)\text{Im}(b))$ , for every  $a, b \in \mathbb{C}$ , straightforward calculation shows that  $Z(\boldsymbol{\vartheta})$  can be explicitly written as

$$\begin{aligned} Z(\boldsymbol{\vartheta}) = & X_{\mathbf{0}} + 2 \sum_{k_1=1}^{\infty} \left\{ X_{k_1,0} \cos(k_1\vartheta_1) + Y_{k_1,0} \sin(k_1\vartheta_1) \right\} \\ & + 2 \sum_{k_2=1}^{\infty} \left\{ X_{0,k_2} \cos(k_2\vartheta_2) + Y_{0,k_2} \sin(k_2\vartheta_2) \right\} \\ & + 2 \sum_{\mathbf{k} \in \mathbb{N}^2} \left\{ X_{k_1,k_2} \cos(k_1\vartheta_1 + k_2\vartheta_2) + Y_{k_1,k_2} \sin(k_1\vartheta_1 + k_2\vartheta_2) \right\} \\ & + 2 \sum_{\mathbf{k} \in \mathbb{N}^2} \left\{ \tilde{X}_{k_1,k_2} \cos(-k_1\vartheta_1 + k_2\vartheta_2) + \tilde{Y}_{k_1,k_2} \sin(-k_1\vartheta_1 + k_2\vartheta_2) \right\}, \tag{5.18} \end{aligned}$$

where  $X_{\mathbf{k}}, Y_{\mathbf{k}}, \tilde{X}_{\mathbf{k}}$  and  $\tilde{Y}_{\mathbf{k}}$  are the following real-valued random variables:

$$X_{k_1,k_2} = \text{Re}(\alpha_{k_1,k_2}), Y_{k_1,k_2} = -\text{Im}(\alpha_{k_1,k_2}), \tilde{X}_{k_1,k_2} = \text{Re}(\alpha_{-k_1,k_2}) \text{ and } \tilde{Y}_{k_1,k_2} = -\text{Im}(\alpha_{-k_1,k_2}) \text{ and } X_{\mathbf{0}} = \alpha_{\mathbf{0}}.$$

Assume that all these random variables are uncorrelated, with  $\text{var}(X_{\mathbf{0}}) = b_{\mathbf{0}}$ , whereas for positive indices  $k_1$  and  $k_2$ ,  $\text{var}(X_{k_1,0}) = \text{var}(Y_{k_1,0}) = b_{k_1,0}/4$ ,  $\text{var}(X_{0,k_2}) = \text{var}(Y_{0,k_2}) = b_{0,k_2}/4$ , and  $\text{var}(X_{\mathbf{k}}) = \text{var}(Y_{\mathbf{k}}) = \text{var}(\tilde{X}_{\mathbf{k}}) = \text{var}(\tilde{Y}_{\mathbf{k}}) = b_{\mathbf{k}}/8$ .

The absence of correlation among the coefficients, together with standard trigonometric identities implies that the random field (5.18), has a covariance function given by

$$\text{cov}\{Z(\boldsymbol{\vartheta}), Z(\boldsymbol{\vartheta}')\} = \sum_{\mathbf{k} \in \mathbb{N}_0^2} b_{\mathbf{k}} \cos(k_1 s) \cos(k_2 t), \tag{5.19}$$

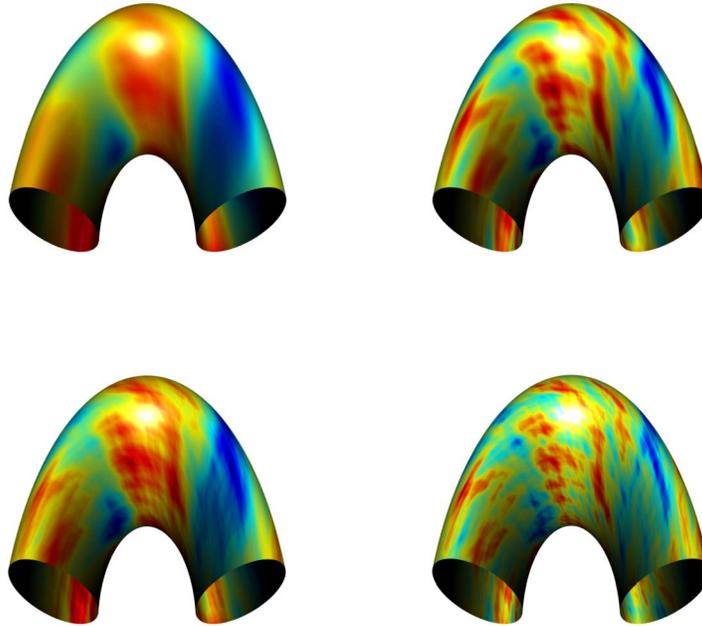


FIG 5.1. Simulated random fields on  $\mathbb{T}^{1,1}$ , over a grid of size  $500 \times 500$ , with  $N = 200$ , and angular power spectrum  $b_{\mathbf{k}} = (\tau^2 + |\mathbf{k}|^2)^{-\delta}$ . We have considered four cases:  $(\tau^2, \delta) = (10, 3)$ ,  $(\tau^2, \delta) = (100, 3)$ ,  $(\tau^2, \delta) = (10, 2)$ , and  $(\tau^2, \delta) = (100, 2)$  (from top left to bottom right).

where  $s$  is the geodesic distance between  $\vartheta_1$  and  $\vartheta'_1$ , and  $t$  is the geodesic distance between  $\vartheta_2$  and  $\vartheta'_2$ . It worth mentioning that our obtained expression (5.19) consists of the bi-variate analogous of expression (12) in [Yadrenko \(1983\)](#), page 74.

A truncation of (5.18), as in (5.10), can be used to simulate random fields with a prescribed angular power spectrum. Figure 5.1 shows simulated random fields on  $\mathbb{T}^{1,1}$ , over a grid of size  $500 \times 500$ , with  $N = 200$ , using the angular power spectrum  $b_{\mathbf{k}} = (\tau^2 + |\mathbf{k}|^2)^{-\delta}$ , for different values of  $\tau^2 > 0$  and  $\delta > 1$ . Here,  $\tau^2$  controls the range of the field, i.e., it regulates the distance at which the spatial correlation is negligible, whereas  $\delta$  is a parameter responsible for the smoothness of the sample paths. We have only reported half of the surface of  $\mathbb{T}^{1,1}$  in order to obtain a better visualization.

We turn to a numerical validation of the truncation error derived in the previous subsection. We simulate a random field over 100 uniformly sampled coordinates, from the same angular power spectrum described above with  $\tau^2 = 10$  and two values for  $\delta = 2, 3$ . Note that the condition (5.11) of Theorem 5.1 is satisfied with  $\varepsilon = \delta - 1$ . The “true” random field is taken as the truncated expansion with  $N = 500$  according to (5.10). We gradually truncate the expansion at different values of  $N$  and measure the discrepancy between the truncated and “true” fields, in a similar fashion to the works of [Lang and](#)

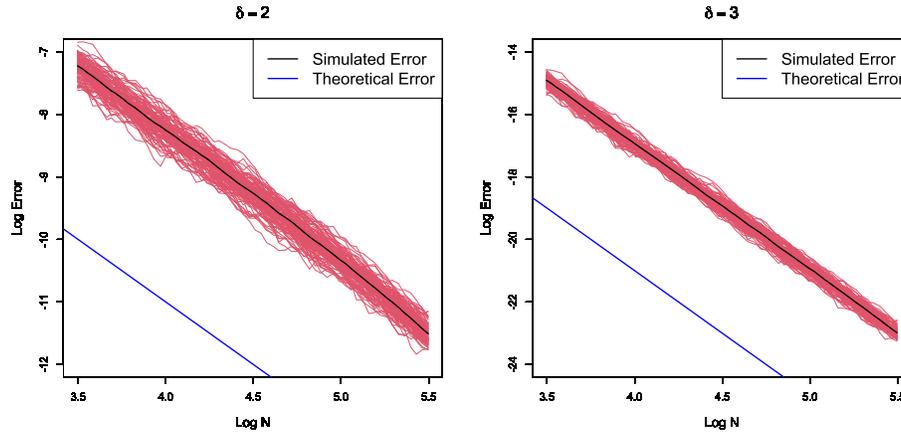


FIG 5.2. Convergence rates in terms of  $N$  in a logarithmic scale, for  $\delta = 2$  (left) and  $\delta = 3$  (right). The empirical convergence rate is given by the black line, whereas the theoretical one is given by the blue line. The red lines indicate the 100 independent repetitions that were averaged to approximate the integral over the probability space and construct the black line.

Schwab (2015) and Clarke et al. (2018). The uniformly distributed coordinates and 100 independent repetitions of this experiment allow us to approximate the integrals involved in  $\|Z - Z^N\|_{L^2(\Omega, L^2(\mathbb{T}^{1,1}))}^2$  by means of a Monte Carlo argument. Figure 5.2 shows the results. Specifically, we display the decay of the empirical (mean squared) errors in terms of  $N$ , in a logarithmic scale, and draw artificial lines with the expected theoretical slopes. According to (5.12),  $\log \|Z - Z^N\|_{L^2(\Omega, L^2(\mathbb{T}^{d_1, d_2}))}^2 \leq c - 2\varepsilon \log N$ , i.e. decreases as a linear function of  $\log N$ , with slope  $-2\varepsilon = -2(\delta - 1)$ . Figure 5.2 clearly shows that the empirical error matches with the theoretical one for  $\delta \in \{2, 3\}$ , as expected.

## 6. Data illustrations and methods

In this section we illustrate modeling data using truncated expansion of (2.3). We consider two datasets: wind speed data from a U.S. Climate Reference Network (USCRN) monitoring station near Provo, UT, USA (Diamond et al., 2013) and National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) global cloud cover reanalysis data (Kalnay et al., 1996). Our focus of this analysis is to (1) discuss a modeling framework for these data, (2) discuss how directional, seasonal, and global data can be viewed as quantities on the product of spheres, and (3) assess model performance as a function of number of terms included in the truncation of (2.3). For these examples, as in Porcu and White (2022), we use covariance models over  $\mathbb{R} \times \mathbb{T}^{d_1, d_2}$ , where  $\mathbb{R}$  accounts for temporally decaying autocovariance patterns. The code and datasets supporting these results are available on [https://github.com/philawhite/Torus\\_code2](https://github.com/philawhite/Torus_code2).

To make renewable energy sources more effective and economical, forecasting wind speed has been given particular attention in the statistics literature (see, e.g., Gneiting et al., 2006; Hering and Genton, 2010; Ezzat et al., 2019). In contrast to these forecasting models, we present a descriptive analysis of wind speed using a Gaussian process model with covariance over  $\mathbb{R} \times \mathbb{T}^{1,1}$  to account for temporal, seasonal, and directional autocorrelation. Using covariance functions over the circle have been used to capture seasonality (see Shirota and Gelfand, 2017b; Mastrantonio et al., 2019; White and Porcu, 2019a). Similarly, directional outcomes have been modeled using projections on the circle (Jona-Lasinio et al., 2012; Wang and Gelfand, 2014). In addition to our covariance model, we use maximum five-minute air temperature, maximum relative humidity, and total solar radiation as predictors (covariates) to explain wind speed.

We also consider a random subset of global monthly-averaged cloud coverage data derived from the 2017 NCEP/NCAR reanalysis data (Kalnay et al., 1996). Cloud coverage has been affected by climatic changes (see, e.g., Wylie et al., 2005), where cloud coverage is the fraction of the sky covered with visible clouds. Because cloud coverage is closely related to solar radiation levels (see, e.g., Wang et al., 2018; Gandoman et al., 2018), rainfall (see Karbalaee et al., 2017), and ecosystem stability (see Pounds et al., 1999), cloud coverage changes have widespread effects. These cloud coverage data are irregularly-sampled, in both space and time, and globally distributed, showing temporal, spatial, and seasonal autocorrelation. However, unlike most seasonal data, we only observe a single cycle. Together, these attributes motivate a covariance model over time, the globe, and season ( $\mathbb{R} \times \mathbb{T}^{2,1}$ ). Cloud coverage is nonlinearly connected to latitude and dependent on whether the location is over water or land. For these reasons, we include a cubic basis spline for latitude (with knots at  $-30^\circ$  and  $30^\circ$ ) and absolute latitude as covariates. We also include interactions between these latitude functions and an indicator variable for whether the location is land-bound or over water. Moreover, the relationships between these covariates and cloud coverage are modeled dynamically.

For both datasets, we define  $\mathbf{y}$  to be the outcome of interest, and we let  $\mathbf{X}$  be a design matrix, including a column of ones, where the rows of  $\mathbf{X}$  are indexed by the same quantities (e.g., time and location) as the outcome. Our model is

$$\mathbf{y} \sim \mathcal{N}(\mathbf{X}\beta, \Sigma + \tau^2\mathbf{I}), \quad (6.1)$$

where  $\beta$  are regression coefficients, including an intercept term, and  $\tau^2$  is a variance term accounting for residual error. The  $i$ th row and  $j$ th column of the covariance matrix  $\Sigma$  is defined using the product of (1) an exponential correlation function of time difference  $|t_i - t_j|$  and (2) a truncated expansion of (2.3), using Gegenbauer polynomials of order up to and including  $K_1, K_2 \in \mathbb{N}$ ,

$$e^{-\phi|t_i - t_j|} \sum_{k_1=0}^{K_1} \sum_{k_2=0}^{K_2} b_{k_1, k_2} C_{k_1}^{\frac{d_1-1}{2}}(\cos r_{1ij}) C_{k_2}^{\frac{d_2-1}{2}}(\cos r_{2ij}), \quad (6.2)$$

where  $r_{1ij}$  and  $r_{2ij}$  are great circle distances over the appropriate domains. The

parameters  $b_{k_1, k_2}$ , for  $k_1 \in \{0, \dots, K_1\}$  and  $k_2 \in \{0, \dots, K_2\}$ , account for the scale of covariance over the product of spheres and time.

We use the following prior distributions on covariance parameters:

$$\begin{aligned} b_{k_1, k_2} &\stackrel{iid}{\sim} \text{Half-Cauchy}(0, 1), & \text{for } k_1, k_2 \in \{0, \dots, K\}, \\ \tau^2 &\sim \text{Half-Cauchy}(0, 1), \\ \phi &\sim \text{Gamma}(a_\phi, b_\phi), \end{aligned} \tag{6.3}$$

where the  $\text{Gamma}(a, b)$  distribution has mean  $a/b$ . We choose Half-Cauchy prior distributions because it is a well-behaved default prior distribution for variance components that may be close to 0. Gelman (2006) and Polson and Scott (2012) recommend a half-Cauchy prior distribution because its thick tail is approximately uniform in the tail, yet it is weakly informative near the origin. Thus, this prior distribution avoids some of the poor behavior of the more standard inverse-gamma prior distribution. Moreover, because the datasets have low sample variance, 0.5 and 1, respectively, the Half-Cauchy prior distributions are very diffuse. We choose the prior distributions for  $\phi$  based on the time scale and our exploratory analyses. We use Normal prior distribution on regression coefficients  $\beta$ . For both  $\phi$  and  $\beta$ , our specifications are different in the two data analyses and are discussed in more detail in their respective sections.

We fit models using Markov chain Monte Carlo (MCMC). We run the sampler for 60,000 iterations, discarding the first 30,000 iterations. Because the posterior conditional distribution for  $\beta$  is Normal, we sample its conditional distribution directly. No other parameters have closed-form posterior conditional distributions. Therefore, we sample all  $b_{k_1, k_2}$ ,  $\tau^2$  and  $\phi$  jointly on the log-scale using multivariate Normal proposals for Metropolis-Hastings updates. We tune the candidate covariance of the multivariate normal using previously accepted samples, an adaptive MCMC approach described in Haario et al. (2001). We also tune the scaling factor of the covariance to obtain an acceptance rate between 0.1 and 0.6 (Roberts and Rosenthal, 2009). We initialize  $\beta$  using least-squares estimates from a multiple regression model,  $b_{k_1, k_2}$  using least-squares fits to empirical covariances, and  $\phi$  based on our exploratory analyses. All code is available at [https://github.com/philawhite/Torus\\_code2](https://github.com/philawhite/Torus_code2).

### 6.1. Wind speed data analysis

This dataset consists of  $n = 2205$  average hourly wind speed measurements in m/s from March-May of 2019. The data are missing three measurements, and, for the model we pose, there is no need to impute missing observations. However, this approach could be used to infill missing values. We model the  $\log(\cdot + 1)$  of the observed values to stabilize the mean-variance relationship and call this quantity log wind speed.

In the top row of Figure 6.1, we plot the average log wind speed over hour-of-day and wind direction. These data show strong seasonality over the time of day and clear nonlinear trends over wind direction. Even after accounting

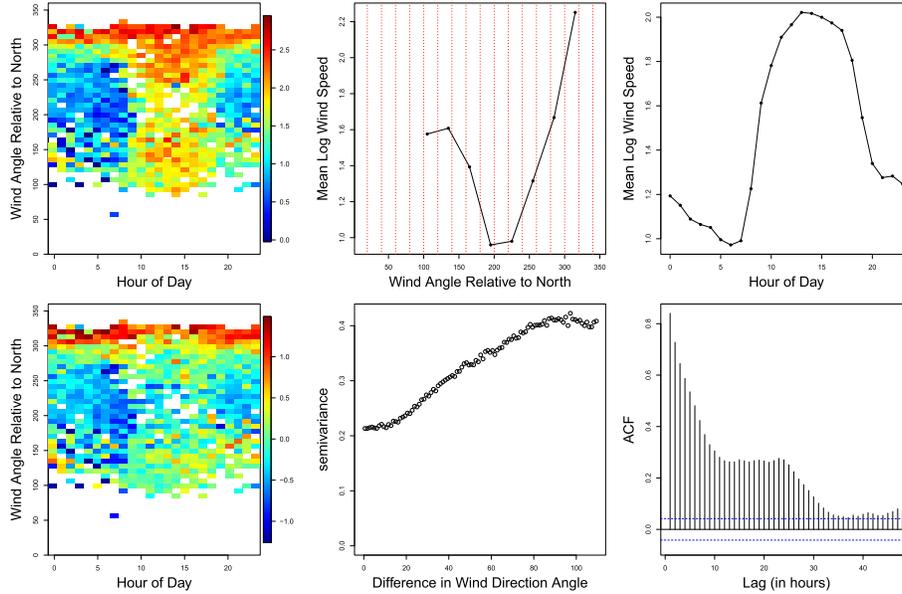


FIG 6.1. (Top-Left) Average log wind speed over hour of day and wind angle, (Top-Center) Average log wind speed over wind angle, (Top-Right) Log wind speed averaged over hour of day, (Bottom-Left) Residuals of log wind speed averaged hour of day and wind angle, (Bottom-Center) Empirical binned semivariogram as a function of difference in wind direction angle, (Bottom-Right) Residual autocorrelation as a function of lagged hours.

for temperature, relative humidity, and total solar radiation, there are strong residual trends in wind direction and weak daily seasonality (See the bottom row of Figure 6.1), motivating a covariance model over  $\mathbb{T}^{1,1}$ . Lastly, the decaying temporal autocorrelation pattern in Figure 6.1 indicates the need to account for autocorrelation decay, including  $\mathbb{R}$  in the covariance function.

Representing time  $t_i$  in units of hours from the beginning of January 2019, the  $i$ th element of the observation vector  $\mathbf{y}$  is the log-transformed wind speed  $y(t_i) \in \mathbb{R}$ . We define the design matrix  $\mathbf{X}$  using centered-and-scaled temperature, relative humidity, and total solar radiation, with rows  $\mathbf{x}(t_i)$  observed at time  $t_i$ . We let wind direction (in radians) be  $a(t_i) \in (0, 2\pi)$ , and define distances for (6.2) as

$$r_{1ij} = \arccos \left[ \cos \left( \frac{\pi |t_i - t_j|}{12} \right) \right],$$

$$r_{2ij} = \arccos [\cos (a(t_i) - a(t_j))],$$

and consider expansions of (6.2) where  $K_1 = K_2 = K$ . For this dataset, we use a weakly-informative prior distribution for the regression coefficients,

$$\beta \sim \mathcal{N}(0, 100\mathbf{I}).$$

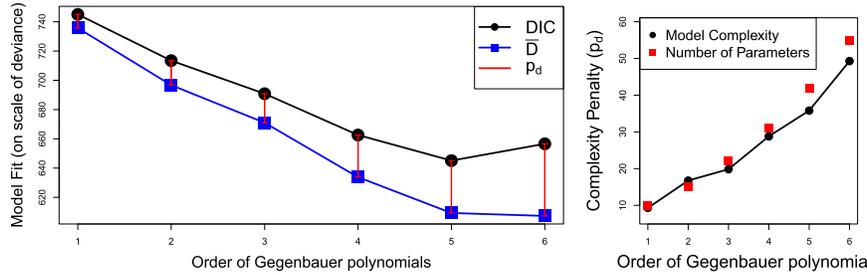


FIG 6.2. (Left)  $\bar{D}$ ,  $p_D$ , and DIC as a function of the order of Gegenbauer polynomials, (Right) Estimated model complexity  $p_D$  compared to number of parameters.

Lastly, we choose a Gamma(2,72) prior distribution for  $\phi$  because, from our exploratory analysis, the autocorrelation of residuals persists for approximately 36 hours.

We explore the fit of the model in terms of order  $K$  of Gegenbauer polynomials used (i.e., the truncation point). After fitting the model using MCMC, we calculate the deviance information criterion (DIC), mean deviance ( $\bar{D}$ ), estimated model complexity ( $p_d$ , i.e., the effective number of parameters), and number of parameters (See Figure 6.2). We see clear improvements in fit  $\bar{D}$  up to  $K = 5$ ; however, there is little improvement increasing to  $K = 6$ , as we see an increase (worsening) in DIC. Figure 6.2 shows that the effective number of parameters  $p_d$  tracks closely with the number of parameters.

For the model with the lowest DIC,  $K = 5$ , we analyze the output of the model. First, we give posterior summaries for regression coefficients, the decay parameter  $\phi$ , and the estimated residual variance divided by the sample variable in Table 1. Because the covariates were centered and scaled, we can interpret the relative importance of the covariates by the magnitude of the corresponding coefficients. All covariates have 95% credible intervals that exclude 0, and maximum temperature has the strongest estimated relationship with wind speed. Aligning with our exploratory analyses, the estimated  $\phi$  suggests that temporal autocovariance in log wind speed persists for one or two days (24-48 hours). Lastly, the ratio of estimated residual variance to sample variance is 0.03, suggesting that this model explains the vast majority of the variability in the data.

We calculate the covariance over a grid of wind direction differences and time differences for each posterior sample and take the posterior mean over this grid (See Figure 6.3). This posterior covariance surface shows that the wind direction and temporal decay are the most evident covariance patterns after accounting for covariates. The seasonal autocorrelation is, however, apparent (See gray vertical lines in Figure 6.3). When interpreted with respect to the estimated ratio of residual variance and sample variance, as well as our exploratory analyses, it is evident that the covariance model posed here effectively captures the autocovariance present in the data.

TABLE 1  
 Posterior summaries for regression coefficients and estimated ratio of unexplained variance to sample variance.

|                                  | Mean   | Std. Dev. | 2.5%   | 97.5%  |
|----------------------------------|--------|-----------|--------|--------|
| Intercept                        | 1.505  | 0.104     | 1.307  | 1.709  |
| Maximum Temperature              | 0.380  | 0.031     | 0.317  | 0.440  |
| Maximum Relative Humidity        | -0.103 | 0.019     | -0.139 | -0.069 |
| Total Shortwave Radiation        | 0.055  | 0.017     | 0.020  | 0.087  |
| Temporal Covariance Decay $\phi$ | 0.023  | 0.003     | 0.018  | 0.029  |
| $\tau^2$ /(sample variance)      | 0.026  | 0.003     | 0.020  | 0.032  |

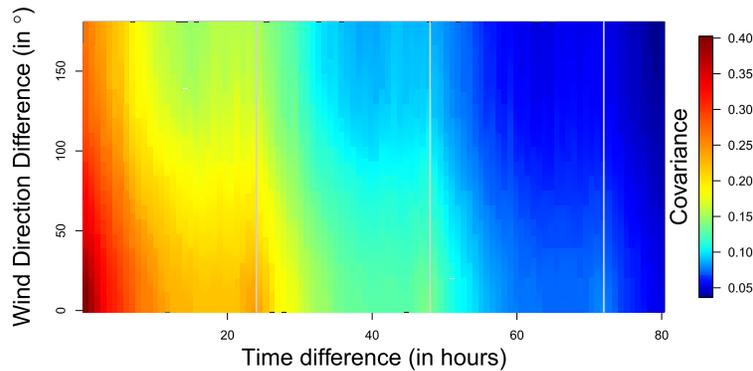


FIG 6.3. Posterior mean covariance as a function of wind direction and time differences.

## 6.2. NCEP/NCAR monthly-averaged cloud coverage analysis

We examine a dataset of 3,600 monthly-averaged cloud coverage observations over 2017 coming from 1,727 unique locations. Most sites are unobserved for most months. We center and scale the observed cloud coverages so that the dataset has a mean of zero and variance of one.

As a preliminary analysis, we fit a model using absolute latitude and cubic basis splines of latitude with knots at  $-30^\circ$  and  $30^\circ$  for each month. In addition, we include the interaction between these latitude functions and an indicator for land or sea. We then assess spatial and temporal patterns in the residuals to motivate our covariance modeling approach. First, we calculate an empirical binned semivariogram and plot the time-averaged (over all months) semivariogram in Figure 6.4. In addition, we examine autocorrelation as a function of difference in months for small spatial windows (See Figure 6.4). These plots show clear spatial and seasonal autocorrelation patterns, as well as some temporal decay over time.

The  $i$ th element of the observation vector  $\mathbf{y}$  is the scaled monthly-averaged cloud-coverage  $y(t_i, \mathbf{s}_i) \in \mathbb{R}$ , where the associated time  $t_i$  is in units of months from the beginning of January 2017 and location is  $\mathbf{s}_i$ . We define the required distances for (6.2) as

$$r_{1ij} = d(\mathbf{s}_i, \mathbf{s}_j),$$

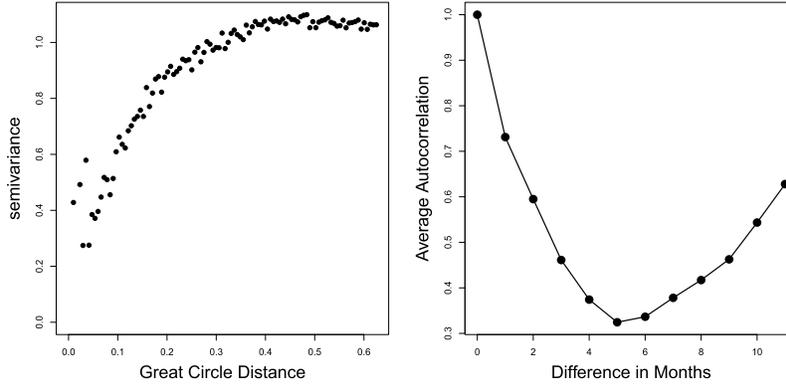


FIG 6.4. (Left) Time-average semivariogram as a function of great-circle distance, averaged over all months, (Right) Average temporal autocorrelation as a difference in months.

$$r_{2ij} = \arccos \left[ \cos \left( \frac{\pi |t_i - t_j|}{6} \right) \right],$$

where  $d(\mathbf{s}_i, \mathbf{s}_j) \in [0, \pi]$  is the great-circle distance. We consider expansions of (6.2), varying  $K_1$  and fixing  $K_2 = 3$ . We found no benefit to including higher order seasonal terms ( $K_2$ ) in (6.2). Because we consider dynamic coefficients for latitudinal trends (as discussed previously), we let  $\mathbf{X}$  have a block-diagonal structure, where the non-zero blocks consist of month-specific design matrices. We use a prior distribution for regression coefficients that accounts for temporal and between-predictor covariance structure:

$$\beta \sim \mathcal{N}(\mathbf{1} \otimes \mu_\beta, \mathbf{R} \otimes \mathbf{V}),$$

where  $\mathbf{V} \sim \text{Inverse-Wishart}(\mathbf{I}, p + 1)$ ,  $p = 14$  is the number of covariates specifying the land and sea latitude functions, and  $i$ th row and  $j$ th row of  $\mathbf{R}$  is  $\exp(-\phi_\beta |i - j|)$ . We choose a Gamma(2,24) prior distribution for  $\phi$  because cloud coverage autocorrelation is high after 11 months (Figure 6.4). Lastly, we use a Gamma(1,2) prior distribution for  $\phi_\beta$  to allow rapid or slow autocorrelation decay in the regression coefficients.

We explore the fit of the model for various orders of the spatial autocovariance  $K_1 + 1 = 4, 8, 12, 15, 20$ , where, after fitting each model using MCMC, we calculate DIC,  $\bar{D}$ , and  $p_d$  (See Figure 6.5). We see clear improvements in fit  $\bar{D}$  up to  $K_1 = 19$ . Model fitting for more covariance terms  $K_1$  was not practically feasible.

For the model with  $K_1 = 19$  and  $K_2 = 3$ , we present summaries of the estimated latitudinal trends and space-time covariance. We plot the posterior mean effect of latitude for all month in Figure 6.6. For land-bound sites, the estimated effect of latitude appears nearly quadratic but is asymmetric about  $0^\circ$ . The estimated effect of latitude over water appears more complicated and is also asymmetric about  $0^\circ$ . On average, land-bound observations have lower cloud

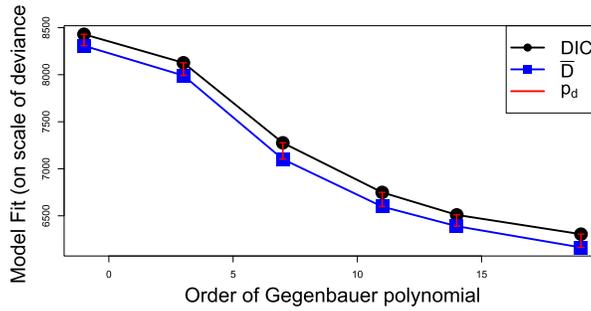


FIG 6.5.  $\bar{D}$ ,  $p_D$ , and DIC as a function of the order of Gegenbauer polynomials ( $K_1$ ). Here,  $K_1 = -1$  indicates that this model was an independent errors model.

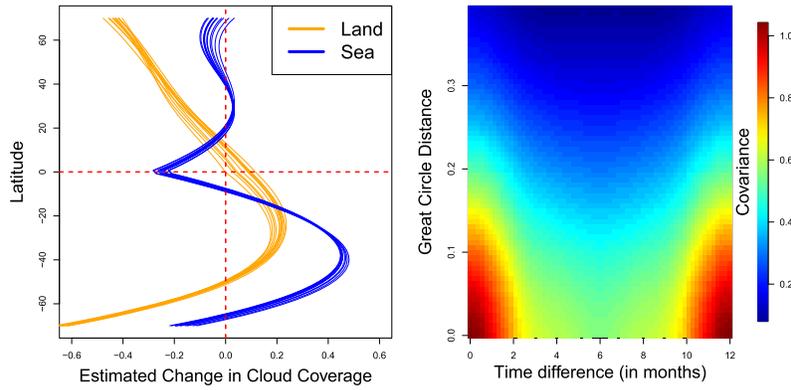


FIG 6.6. (Left) Latitude trends for cloud coverage for land and water. Each line represents the posterior mean for one month. (Right) Posterior mean covariance as a function of time difference and great circle distance.

coverage than observations over water, except close to the equator. Extreme latitudes have low cloud coverage over land, while this pattern is much weaker over water.

We calculate the covariance over a grid of great-circle distances and time differences (in months) for each posterior sample. We use the posterior mean over this grid and plot this in Figure 6.6. This posterior covariance surface shows persistent spatial covariance patterns, as well as strong seasonal patterns. As in Section 6.1, we calculate the estimated ratio of residual variance and sample variance. We find the posterior mean of this ratio to be 0.135 with 95% credible interval (0.125,0.145). This indicates that our model explains approximately 86%-87% of the variance in the data.

## 7. Final remarks

The framework put forward in this paper is ideal for the study of periodic phenomena in space and for directional data. For the theoretical point of view, our study tried to be as general as possible by considering the product of two arbitrary dimensional spheres. We hope that the setting will find its place in future developments from the side of applications, theory or both. Open directions such as the study of multivariate random fields, SPDEs, Bayesian analysis' problems and simulations could be performed in the current framework following Lang and Schwab (2015), Alegría et al. (2021) and Clarke et al. (2018). The process has already been started and some first multivariate study can be found in Bachoc et al. (2022). The exploration of similar research in the context of random fields on groups and connected compact two-point homogeneous spaces is an intriguing subject worthy of attention. Relevant works by Malyarenko and Olenko (1992), Malyarenko (2004) and Ma and Malyarenko (2020) can provide valuable insights in this regard.

## 8. Appendix

We place here the proofs of our results, together with the necessary background and other auxiliary properties.

### 8.1. Weighted Lebesgue spaces and orthogonal polynomials

We start with some functional analysis' properties that guide our study.

Our first target is to construct an *orthonormal* basis for  $L^2_{(d_1, d_2)}([-1, 1]^2)$ ,  $d_1, d_2 \in \mathbb{N}$ , involving the Gegenbauer and Chebyshev polynomials  $C_{k_i}^{\frac{d_i-1}{2}}$ ,  $i = 1, 2$  aligned with the expression (2.3) established by (Guella et al., 2015).

Let  $d \in \mathbb{N}$ . The space  $L^2_{(d)} = L^2_{(d)}[-1, 1]$  is defined as the set of functions  $u : [-1, 1] \rightarrow \mathbb{R}$  which are square integrable with respect to the weight  $\omega_d(t) := (1 - t^2)^{\frac{d-2}{2}}$ ; i.e.

$$\|u\|_{L^2_{(d)}}^2 := \int_{-1}^1 |u(t)|^2 (1 - t^2)^{\frac{d-2}{2}} dt < \infty. \quad (8.1)$$

$L^2_{(d)}$  is a Hilbert space associated with the inner product

$$\langle u, v \rangle_d := \int_{-1}^1 u(t)v(t)(1 - t^2)^{\frac{d-2}{2}} dt, \quad \text{for every } u, v \in L^2_{(d)}.$$

Let  $\alpha > -1$ . Then, (see for example (Szegő, 1939)) the Jacobi polynomials  $P_k^\alpha := P_k^{(\alpha, \alpha)}$ ,  $k \in \mathbb{N}_0$ , form an orthogonal basis for  $L^2_{(d)}$  for  $\alpha = \frac{d-2}{2}$ .

Jacobi and Gegenbauer polynomials are connected via the following identity, which can be found in (Szegő, 1939); equation (4.7.1):

$$C_k^\lambda = \frac{\Gamma(\lambda + 1/2)\Gamma(k + 2\lambda)}{\Gamma(2\lambda)\Gamma(k + \lambda + 1/2)} P_k^{\lambda - \frac{1}{2}} =: \gamma_k(\lambda) P_k^{\lambda - \frac{1}{2}}, \quad k \in \mathbb{N}_0, \quad \lambda > 0, \quad (8.2)$$

where  $\Gamma$  is the well-known Gamma function. It is very important to mention that because  $\Gamma(2\lambda) \rightarrow \infty$ , when  $\lambda \rightarrow 0^+$ , the above representation cannot be valid for the case  $C_k^0$ . In the case of our interest the exponent  $\lambda$  corresponds to  $\frac{d-1}{2}$ ; see (2.3). Thus  $\lambda = 0$  exactly for  $d = 1$ , therefore we must distinguish the studies of the cases  $d > 1$  and  $d = 1$ .

We start with  $d > 1$  and we keep using  $\alpha = \frac{d-2}{2}$ . By (4.3.3) in (Szegő, 1939) we have

$$\int_{-1}^1 P_k^\alpha(t) P_{k'}^\alpha(t) (1-t^2)^\alpha dt = \delta_{kk'} r_k(\alpha), \quad (8.3)$$

where  $\delta_{kk'}$  is Kronecker's delta and  $r_k(\alpha)$ , is given by

$$r_k(\alpha) = \begin{cases} \frac{2^{2\alpha+1}}{2k+2\alpha+1} \frac{\Gamma(k+\alpha+1)^2}{k! \Gamma(k+2\alpha+1)}, & k \geq 1 \\ 2^{2\alpha+1} \frac{\Gamma(\alpha+1)^2}{\Gamma(2\alpha+2)}, & k = 0, \end{cases}$$

and has been estimated in Lemma 4.3 in (Cleantous et al., 2020) by

$$r_k(\alpha) \sim (k+1)^{-1}, \quad \text{for every } k \in \mathbb{N}_0, \alpha > -1/2. \quad (8.4)$$

Using the following basic property of the Gamma function,

$$\Gamma(\nu + 1 + r) \sim \nu!(\nu + 1)^r, \quad \text{for every } \nu \in \mathbb{N}_0 \text{ and } r \notin \{-1, -2, \dots\}, \quad (8.5)$$

we receive the estimate below for the  $\gamma_k(\lambda)$  appearing in (8.2):

$$\gamma_k(\lambda) \sim (k+1)^{\lambda-1/2}, \quad \text{for every } k \in \mathbb{N}_0, \lambda > 0. \quad (8.6)$$

With all the above in hand we establish the following orthonormal basis for  $L^2_{(d)}$ , when  $d > 1$  that involves the Gegenbauer polynomials:

$$\{\tilde{P}_k^d(t) := \beta_k(d) C_k^{\frac{d-1}{2}}(t) : k \in \mathbb{N}_0\}, \quad (8.7)$$

where

$$\beta_k(d) := r_k \left(\frac{d-2}{2}\right)^{-1/2} \gamma_k \left(\frac{d-1}{2}\right)^{-1} \sim (k+1)^{-\frac{d-3}{2}} \quad d \geq 2, k \in \mathbb{N}_0, \quad (8.8)$$

in the light of (8.4) and (8.6).

We now turn our attention to the case  $d = 1$ . Here we work with the Chebyshev polynomials, as can be found in Section 1.7 of (Szegő, 1939) (referred as

Chebyshev polynomials). For every  $k \in \mathbb{N}_0$  the Chebyshev polynomial of order  $k$  is defined by

$$C_k^0(t) := T_k(t) := \cos(k \arccos(t)), \quad t \in [-1, 1], \quad (8.9)$$

it is a polynomial of degree  $k$  and evidently  $|C_k^0(t)| \leq 1 = C_k^0(1)$ . In the spirit of (8.3) we have (see e.g. page 60 of (Szegő, 1939), or use a change of variables)

$$\int_{-1}^1 C_k^0(t) C_{k'}^0(t) \frac{dt}{\sqrt{1-t^2}} = \delta_{kk'} r_k \left( -\frac{1}{2} \right), \quad \text{where } r_k \left( -\frac{1}{2} \right) := \begin{cases} \pi, & k = 0 \\ \frac{\pi}{2}, & k \geq 1 \end{cases}. \quad (8.10)$$

Thus the sequence  $\{C_k^0\}_{k \in \mathbb{N}_0}$  is an orthogonal basis for  $L^2_{(1)}$  and can be normalized to

$$\tilde{P}_k^1(t) := r_k \left( -\frac{1}{2} \right)^{-1/2} C_k^0(t) =: \beta_k(1) C_k^0(t), \quad \text{with } \beta_k(1) \sim 1. \quad (8.11)$$

As a summary of (8.7), (8.8) and (8.11), for every  $d \in \mathbb{N}$ , the sequence  $\{\tilde{P}_k^d := \beta_k(d) C_k^{\frac{d-1}{2}}\}_{k \in \mathbb{N}_0}$  forms an orthonormal basis for  $L^2_{(d)}$  and

$$\beta_k(d) \sim (k+1)^{-\tau(d)/2}, \quad \text{where } \tau(d) := \begin{cases} d-3, & d \geq 2 \\ 0, & d = 1 \end{cases}. \quad (8.12)$$

We are now ready to pass to the product domain  $[-1, 1] \times [-1, 1]$ .

For every  $d_1, d_2 \in \mathbb{N}$ ,  $L^2_{(d_1, d_2)}$  introduced in (3.1) is a Hilbert spaces and the associated inner product is given by

$$\langle u, v \rangle_{d_1, d_2} := \int_{-1}^1 \int_{-1}^1 u(t_1, t_2) v(t_1, t_2) (1-t_1^2)^{\frac{d_1-2}{2}} (1-t_2^2)^{\frac{d_2-2}{2}} dt_1 dt_2, \quad (8.13)$$

for every  $u, v \in L^2_{(d_1, d_2)}$ .

Then —see for example (Reed and Simon, 1980), page 51— the sequence

$$\{\tilde{P}_{\mathbf{k}}^{d_1, d_2} := \tilde{P}_{k_1}^{d_1} \otimes \tilde{P}_{k_2}^{d_2} : \mathbf{k} = (k_1, k_2) \in \mathbb{N}_0^2\}, \quad (8.14)$$

where  $\otimes$  is the tensor product, is an orthonormal basis for the product space  $L^2_{(d_1, d_2)}([-1, 1] \times [-1, 1])$ . Moreover the elements of the basis in (8.14) can be expressed in terms of Gegenbauer/Chebyshev polynomials, as we want, thanks to (8.7) and (8.11):

$$\tilde{P}_{\mathbf{k}}^{d_1, d_2}(t_1, t_2) = \prod_{i=1}^2 \beta_{k_i}(d_i) C_{k_i}^{\frac{d_i-1}{2}}(t_i), \quad (t_1, t_2) \in [-1, 1]^2, \mathbf{k} = (k_1, k_2) \in \mathbb{N}_0^2. \quad (8.15)$$

Hence, every  $u \in L^2_{(d_1, d_2)}$  can be represented in the form

$$u = \sum_{\mathbf{k} \in \mathbb{N}_0^2} \tilde{u}_{\mathbf{k}} \tilde{P}_{\mathbf{k}}^{d_1, d_2} \quad (\text{convergence in } L^2_{(d_1, d_2)}), \quad (8.16)$$

where the sequence of the coefficients  $\{\tilde{u}_{\mathbf{k}}\}$  is given by  $\tilde{u}_{\mathbf{k}} = \langle u, \tilde{P}_{\mathbf{k}}^{d_1, d_2} \rangle_{d_1, d_2}$ . The last combined with (8.15) implies that every function  $u$  belonging to the space  $L^2_{(d_1, d_2)}$  can be represented through the expansion

$$u = \sum_{\mathbf{k} \in \mathbb{N}_0^2} u_{\mathbf{k}} C_{k_1}^{\frac{d_1-1}{2}} \otimes C_{k_2}^{\frac{d_2-1}{2}} \quad (\text{convergence in } L^2_{(d_1, d_2)}), \quad (8.17)$$

where

$$u_{\mathbf{k}} = \prod_{i=1}^2 \beta_{k_i}(d_i)^2 \times \int_{-1}^1 \int_{-1}^1 u(t_1, t_2) C_{k_1}^{\frac{d_1-1}{2}}(t_1) C_{k_2}^{\frac{d_2-1}{2}}(t_2) (1-t_1^2)^{\frac{d_1-2}{2}} (1-t_2^2)^{\frac{d_2-2}{2}} dt_1 dt_2. \quad (8.18)$$

### 8.2. Proofs of the results in Section 3

We aim to prove Theorem 3.3. We start by proving the next auxiliary lemmata which we find that could be of some independent interest.

**Lemma 8.1.** (α) For every  $n \in \mathbb{N}$  and  $\lambda > 0$ ,

$$\frac{d^n}{dt^n} C_k^\lambda(t) = \delta(n, \lambda) C_{k-n}^{\lambda+n}(t), \quad (8.19)$$

for every  $k \geq n$  and  $t \in [-1, 1]$  where  $\delta(n, \lambda) := 2^n \lambda(\lambda+1) \cdots (\lambda+n-1)$ .

(β) For every  $n \in \mathbb{N}$ , the  $n$ -th derivative of the Chebyshev polynomial  $C_k^0$  can be expressed as

$$\frac{d^n}{dt^n} C_k^0(t) = \zeta(k, n) C_{k-n}^n(t), \quad (8.20)$$

for every  $k \geq n$  and  $t \in [-1, 1]$  where

$$\zeta(k, n) := \frac{2 \cdot 4 \cdots (2k)}{1 \cdot 3 \cdots (2k-1)} \frac{k}{2} \gamma_{k-1}(1)^{-1} \cdot \begin{cases} \delta(n-1, 1), & n > 1 \\ 1, & n = 1 \end{cases} \sim (k+1). \quad (8.21)$$

*Proof.* (α) This claim can be confirmed by recursive applications of the equation (4.7.14) in (Szegő, 1939):

$$\frac{d}{dt} C_k^\lambda(t) = 2\lambda C_{k-1}^{\lambda+1}(t). \quad (8.22)$$

(β) Chebyshev and Jacobi polynomials are connected via the following equation in page 60 of (Szegő, 1939):

$$C_k^0(t) = \frac{2 \cdot 4 \cdots (2k)}{1 \cdot 3 \cdots (2k-1)} P_k^{-\frac{1}{2}}(t) =: \varepsilon_k P_k^{-\frac{1}{2}}(t), \quad k \geq 1. \quad (8.23)$$

The sequence  $\varepsilon_k$  satisfies the behaviour below, as we can evidently verify by induction:

$$\sqrt{k+1} < \sqrt{3k+1} \leq \varepsilon_k \leq \sqrt{4k+1} < 2\sqrt{k+1}, \quad k \in \mathbb{N}. \tag{8.24}$$

On the other hand, by (4.21.7) in (Szegő, 1939) and (8.2) we extract

$$(C_k^0(t))' = \varepsilon_k \left( P_k^{-\frac{1}{2}}(t) \right)' = \varepsilon_k \frac{k}{2} P_{k-1}^{\frac{1}{2}}(t) = \varepsilon_k \frac{k}{2} \gamma_{k-1}(1)^{-1} C_{k-1}^1(t) \tag{8.25}$$

$$= \zeta(k, 1) C_{k-1}^1(t), \tag{8.26}$$

which is of course (8.21) for  $n = 1$ .

Let now  $n > 1$ . We apply (8.26) and the claim  $(\alpha)$  of this Lemma to obtain

$$\frac{d^n}{dt^n} C_k^0(t) = \zeta(k, 1) \frac{d^{n-1}}{dt^{n-1}} C_{k-1}^1(t) = \zeta(k, 1) \delta(n-1, 1) C_{k-n}^n(t), \tag{8.27}$$

as in (8.20), with the estimates in (8.21) to be confirmed by (8.6) and (8.24).  $\square$

The next ingredient of the proof can be summarized in the following lemma.

**Lemma 8.2.** *Let  $d \in \mathbb{N}$ ,  $n \in \mathbb{N}_0$  and  $k, k' \in \mathbb{N}_0$  such that  $k, k' \geq n$ . Then*

$$\begin{aligned} I_{k,k'}^{d,n} &:= \int_{-1}^1 C_{k-n}^{\frac{d-1}{2}+n}(t) C_{k'-n}^{\frac{d-1}{2}+n}(t) (1-t^2)^{\frac{d-2}{2}+n} dt \\ &\sim \delta_{kk'} (k+1)^{\tau(2n+d)} \end{aligned} \tag{8.28}$$

and

$$\begin{aligned} \tilde{I}_{k,k'}^{d,n} &:= \int_{-1}^1 \frac{d^n}{dt^n} C_k^{\frac{d-1}{2}}(t) \frac{d^n}{dt^n} C_{k'}^{\frac{d-1}{2}}(t) (1-t^2)^{\frac{d-2}{2}+n} dt \\ &\sim \delta_{kk'} (k+1)^{2n+\tau(d)}, \end{aligned} \tag{8.29}$$

where the function  $\tau(d)$  is as in (3.9).

*Proof.* We start proving (8.28). By (8.7) and (8.11) we can express

$$C_{k-n}^{\frac{d-1}{2}+n} = \beta_{k-n} (d+2n)^{-1} \tilde{P}_{k-n}^{d+2n}.$$

Then by combining (8.7), (8.11) and (8.12) we obtain

$$\begin{aligned} I_{k,k'}^{d,n} &= \beta_{k-n} (d+2n)^{-1} \beta_{k'-n} (d+2n)^{-1} \langle \tilde{P}_{k-n}^{d+2n}, \tilde{P}_{k'-n}^{d+2n} \rangle_{d+2n} \\ &= \beta_{k-n} (d+2n)^{-1} \beta_{k'-n} (d+2n)^{-1} \delta_{k-n, k'-n} = \delta_{kk'} \beta_{k-n} (d+2n)^{-2} \\ &\sim \delta_{kk'} (k+1)^{\tau(d+2n)}. \end{aligned}$$

For the proof of (8.29) we split the following three cases.

( $\alpha$ ) For  $d = 1$  and  $n = 0$ ; by (8.10)  $\tilde{I}_{k,k'}^{1,0} \sim \delta_{kk'}$ ; confirming (8.29).

( $\beta$ ) For  $d = 1$  and  $n \in \mathbb{N}$ ; by ( $\beta$ ) of Lemma 8.1 and (8.28) we have

$$\begin{aligned} \tilde{I}_{k,k'}^{1,n} &:= \zeta(k,n)\zeta(k',n)I_{k,k'}^{1,n} \sim \delta_{kk'}(k+1)^2(k+1)^{\tau(2n+1)} \\ &= \delta_{kk'}(k+1)^{2+2n+1-3} = \delta_{kk'}(k+1)^{2n+\tau(1)}. \end{aligned}$$

( $\gamma$ ) For  $d > 1$  and  $n \in \mathbb{N}_0$ ; by ( $\alpha$ ) of Lemma 8.1 and (8.28) we obtain

$$\begin{aligned} \tilde{I}_{k,k'}^{d,n} &:= \delta\left(n, \frac{d-1}{2}\right)^2 I_{k,k'}^{d,n} \sim \delta_{kk'}(k+1)^{\tau(2n+d)} \\ &= \delta_{kk'}(k+1)^{2n+d-3} = \delta_{kk'}(k+1)^{2n+\tau(d)}. \end{aligned}$$

□

We proceed to prove Theorem 3.3.

*Proof.* We start by estimating the  $L^2_{(d_1+2n_1, d_2+2n_2)}$ -norm of the distribution  $\partial^{(n_1, n_2)}u$ . By (3.2) we derive

$$\partial^{(n_1, n_2)}u = \sum_{\mathbf{k} \in \mathbb{N}_0^2} u_{\mathbf{k}} (C_{k_1}^{\frac{d_1-1}{2}})^{(n_1)} (C_{k_2}^{\frac{d_2-1}{2}})^{(n_2)}.$$

Then, (8.13) and (8.15) imply

$$\begin{aligned} \|\partial^{(n_1, n_2)}u\|_{L^2_{(d_1+2n_1, d_2+2n_2)}}^2 &= \langle \partial^{(n_1, n_2)}u, \partial^{(n_1, n_2)}u \rangle_{d_1+2n_1, d_2+2n_2} \\ &= \sum_{k_1, k'_1=n_1}^{\infty} \sum_{k_2, k'_2=n_2}^{\infty} u_{\mathbf{k}} u_{\mathbf{k}'} I_{\mathbf{k}, \mathbf{k}'}, \end{aligned} \tag{8.30}$$

where

$$I_{\mathbf{k}, \mathbf{k}'} := \left\langle (C_{k_1}^{\frac{d_1-1}{2}})^{(n_1)} (C_{k_2}^{\frac{d_2-1}{2}})^{(n_2)}, (C_{k'_1}^{\frac{d_1-1}{2}})^{(n_1)} (C_{k'_2}^{\frac{d_2-1}{2}})^{(n_2)} \right\rangle_{d_1+2n_1, d_2+2n_2}.$$

By (8.19) and (8.29) we get the expression

$$\begin{aligned} I_{\mathbf{k}, \mathbf{k}'} &= \prod_{i=1}^2 \int_{-1}^1 (C_{k_i}^{\frac{d_i-1}{2}}(t_i))^{(n_i)} (C_{k'_i}^{\frac{d_i-1}{2}}(t_i))^{(n_i)} (1-t_i^2)^{\frac{d_i-2}{2}+n_i} dt_i \\ &= \prod_{i=1}^2 \tilde{I}_{k_i, k'_i}^{d_i, n_i} \sim \prod_{i=1}^2 \delta_{k_i, k'_i} (k_i+1)^{2n_i+\tau(d_i)}. \end{aligned} \tag{8.31}$$

We now replace (8.31) in (8.30) and we conclude

$$\|\partial^{(n_1, n_2)}u\|_{L^2_{(d_1+2n_1, d_2+2n_2)}}^2 \sim \sum_{k_1=n_1}^{\infty} \sum_{k_2=n_2}^{\infty} u_{\mathbf{k}}^2 (k_1+1)^{2n_1+\tau(d_1)} (k_2+1)^{2n_2+\tau(d_2)}. \tag{8.32}$$

We will first prove Assertion (ii). Equations (3.5) and (8.32) imply

$$\begin{aligned} \|u\|_{DW^{(n,m)}}^2 &\sim \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} u_{\mathbf{k}}^2 (k_1 + 1)^{\tau(d_1)} (k_2 + 1)^{\tau(d_2)} \\ &+ \sum_{k_1=n}^{\infty} \sum_{k_2=0}^{\infty} u_{\mathbf{k}}^2 (k_1 + 1)^{2n+\tau(d_1)} (k_2 + 1)^{\tau(d_2)} \\ &+ \sum_{k_1=0}^{\infty} \sum_{k_2=m}^{\infty} u_{\mathbf{k}}^2 (k_1 + 1)^{\tau(d_1)} (k_2 + 1)^{2m+\tau(d_2)} \\ &+ \sum_{k_1=n}^{\infty} \sum_{k_2=m}^{\infty} u_{\mathbf{k}}^2 (k_1 + 1)^{2n+\tau(d_1)} (k_2 + 1)^{2m+\tau(d_2)} \\ &=: s_1 + \dots + s_4. \end{aligned}$$

Thus,

$$\|u\|_{DW^{(n,m)}}^2 \leq 4 \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} u_{\mathbf{k}}^2 (k_1 + 1)^{2n+\tau(d_1)} (k_2 + 1)^{2m+\tau(d_2)} =: 4s_0$$

and this proves the sufficient part of the assertion, since the right hand side is bounded when  $\{u_{\mathbf{k}}\} \in \ell^{(n,m)}$ .

We now prove the necessity. We decompose  $s_0$  as follows:

$$\begin{aligned} s_0 &= \left( \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{m-1} + \sum_{k_1=n}^{\infty} \sum_{k_2=0}^{m-1} + \sum_{k_1=0}^{n-1} \sum_{k_2=m}^{\infty} + \sum_{k_1=n}^{\infty} \sum_{k_2=m}^{\infty} \right) \\ &\times u_{\mathbf{k}}^2 (k_1 + 1)^{2n+\tau(d_1)} (k_2 + 1)^{2m+\tau(d_2)} \\ &=: s_5 + \dots + s_8. \end{aligned}$$

By (8.32) we derive the following:

$$s_5 \leq n^{2n} m^{2m} \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} u_{\mathbf{k}}^2 (k_1 + 1)^{\tau(d_1)} (k_2 + 1)^{\tau(d_2)} \leq c \|u\|_{L^2_{(d_1, d_2)}}^2.$$

$$s_6 \leq m^{2m} \sum_{k_1=n}^{\infty} \sum_{k_2=0}^{\infty} u_{\mathbf{k}}^2 (k_1 + 1)^{2n+\tau(d_1)} (k_2 + 1)^{\tau(d_2)} \leq c \|\partial_1^n u\|_{L^2_{(d_1+2n, d_2)}}^2$$

and similarly  $s_7 \leq c \|\partial_2^m u\|_{L^2_{(d_1, d_2+2m)}}^2$ . Directly from (8.32) it holds that

$$s_8 \leq c \|\partial^{(n,m)} u\|_{L^2_{(d_1+2n, d_2+2m)}}^2.$$

The above inequalities, in concert with (3.5) complete the proof of claim (ii).

(i) We start by noting the following connection between Sobolev-type spaces:

$$W^N = \bigcap_{\nu=0}^N DW^{(\nu, N-\nu)}.$$

Thus, a  $L^2$  function  $u$  belongs to  $W^N$  if and only if it belongs to the intersection of  $W^{(\nu, N-\nu)}$ , for every  $\nu = 0, \dots, N$ , which equivalently means that its Fourier coefficients belong to  $\bigcap_{\nu=0}^N \ell^{(\nu, N-\nu)}$ , in the light of claim (ii).

(iii) By (3.5) and (8.32) it is true that

$$\begin{aligned} \|u\|_{W^{(n,m)}}^2 &\sim \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} u_{\mathbf{k}}^2 (k_1+1)^{\tau(d_1)} (k_2+1)^{\tau(d_2)} \\ &\quad + \sum_{k_1=n}^{\infty} \sum_{k_2=0}^{\infty} u_{\mathbf{k}}^2 (k_1+1)^{2n+\tau(d_1)} (k_2+1)^{\tau(d_2)} \\ &\quad + \sum_{k_1=0}^{\infty} \sum_{k_2=m}^{\infty} u_{\mathbf{k}}^2 (k_1+1)^{\tau(d_1)} (k_2+1)^{2m+\tau(d_2)} \\ &= s_1 + s_2 + s_3. \end{aligned}$$

Thus,

$$\begin{aligned} &\|u\|_{W^{(n,m)}}^2 \\ &\leq c \sum_{k_1, k_2=0}^{\infty} u_{\mathbf{k}}^2 \left( (k_1+1)^{2n+\tau(d_1)} (k_2+1)^{\tau(d_2)} + (k_1+1)^{\tau(d_1)} (k_2+1)^{2m+\tau(d_2)} \right) \\ &=: c(\tau_1 + \tau_2). \end{aligned}$$

Therefore, if  $\{u_{\mathbf{k}}\}$  belongs to the intersection  $\ell^{(n,0)} \cap \ell^{(0,m)}$ , then  $u \in W^{(n,m)}$ .

On the other hand, let  $u \in W^{(n,m)}$ . Then

$$\begin{aligned} \tau_1 &= \sum_{k_1, k_2=0}^{\infty} u_{\mathbf{k}}^2 (k_1+1)^{2n+\tau(d_1)} (k_2+1)^{\tau(d_2)} \\ &= \left( \sum_{k_1=0}^{n-1} + \sum_{k_1=n}^{\infty} \right) \sum_{k_2=0}^{\infty} u_{\mathbf{k}}^2 (k_1+1)^{2n+\tau(d_1)} (k_2+1)^{\tau(d_2)} \\ &\leq n^{2n} \|u\|_{L^2_{(d_1, d_2)}}^2 + \|\partial_1^n u\|_{L^2_{(d_1+2n, d_2)}}^2 \end{aligned}$$

and similarly  $\tau_2 \leq m^{2m} \|u\|_{L^2_{(d_1, d_2)}}^2 + \|\partial_2^m u\|_{L^2_{(d_1, d_2+2m)}}^2$ .

All the above imply  $\tau_1 + \tau_2 \leq c \|u\|_{W^{(n,m)}}^2$  and the proof is complete.  $\square$

Proof of Corollary 3.4.

*Proof.* The result follows directly by Theorem 3.3, provided that  $K_{\text{iso}}$  is on  $L^2_{(d_1, d_2)}$ . To show it, we make use of (8.15) and (2.3) to derive

$$K_{\text{iso}} = \sum_{\mathbf{k}=(k_1, k_2) \in \mathbb{N}_0^2} b_{\mathbf{k}} \prod_{i=1}^2 \beta_{k_i}(d_i)^{-1} \tilde{P}_{\mathbf{k}}^{d_1, d_2}.$$

Since  $\{\tilde{F}_{\mathbf{k}}^{d_1, d_2}\}$  is an orthonormal basis for  $L^2_{(d_1, d_2)}$ , Parseval's theorem asserts that

$$\|K_{\text{iso}}\|_{L^2_{(d_1, d_2)}}^2 = \sum_{\mathbf{k}=(k_1, k_2) \in \mathbb{N}_0^2} b_{\mathbf{k}}^2 \prod_{i=1}^2 \beta_{k_i}(d_i)^{-2} \sim \sum_{k_1, k_2=0}^{\infty} b_{\mathbf{k}}^2 \prod_{i=1}^2 (k_i + 1)^{\tau(d_i)},$$

thanks to (8.12).

By (2.7), there exists a constant  $c_0 \in (0, \infty)$  such that

$$b_{\mathbf{k}}(k_1 + 1)^{(d_1-2)_+} (k_2 + 1)^{(d_2-2)_+} \leq c_0.$$

Then for every  $\mathbf{k} \in \mathbb{N}_0^2$ ,

$$b_{\mathbf{k}}^2 \prod_{i=1}^2 (k_i + 1)^{\tau(d_i)} \leq c_0 b_{\mathbf{k}} \prod_{i=1}^2 (k_i + 1)^{\tau(d_i) - (d_i-2)_+} \leq c_0 b_{\mathbf{k}} \prod_{i=1}^2 (k_i + 1)^{(d_i-2)_+}.$$

Combining all the above together with (2.7) we justify that

$$\|K_{\text{iso}}\|_{L^2_{(d_1, d_2)}}^2 \leq c \sum_{k_1, k_2=0}^{\infty} b_{\mathbf{k}} (k_1 + 1)^{(d_1-2)_+} (k_2 + 1)^{(d_2-2)_+} < \infty,$$

i.e.  $K_{\text{iso}} \in L^2_{(d_1, d_2)}$ . □

### 8.3. Proofs of the results in Section 4

#### 8.3.1. Proof of Theorem 4.2

*Proof.* (i) By (2.3) and for  $t = (t_1, t_2)$ ,  $s = (s_1, s_2)$  we derive

$$K_{\text{iso}}(t) - K_{\text{iso}}(s) = \sum_{k_1, k_2=0}^{\infty} b_{k_1, k_2} \Delta_{k_1, k_2}(t, s), \tag{8.33}$$

where

$$\begin{aligned} \Delta_{k_1, k_2}(t, s) &:= C_{k_1}^{\frac{d_1-1}{2}}(t_1) C_{k_2}^{\frac{d_2-1}{2}}(t_2) - C_{k_1}^{\frac{d_1-1}{2}}(s_1) C_{k_2}^{\frac{d_2-1}{2}}(s_2) \\ &=: g(t) - g(s). \end{aligned}$$

We apply Taylor's expansion to the function of two variables  $g(t) = g(t_1, t_2)$  (see e.g. Marsden and Tromba (2011)) and we derive

$$\Delta_{k_1, k_2}(t, s) = R_{(1,0)}(t_1 - s_1) + R_{(0,1)}(t_2 - s_2),$$

where  $R_{\alpha}$  is given by

$$R_{\alpha} := R_{\alpha} g(t, s) := \int_0^1 \partial^{\alpha} g(s + r(t - s)) dr,$$

for  $\alpha = (1, 0)$  either  $\alpha = (0, 1)$ .

Let us estimate  $R_{(1,0)}$ . We distinguish the cases  $d_1 > 1$  and  $d_1 = 1$ .

When  $d_1 > 1$ : By (8.19) and (2.6) we derive

$$\begin{aligned} |\partial^{(1,0)}g(t)| &= \left| \left( C_{k_1}^{\frac{d_1-1}{2}}(t_1) \right)' C_{k_2}^{\frac{d_2-1}{2}}(t_2) \right| = c \left| C_{k_1-1}^{\frac{d_1+1}{2}}(t_1) C_{k_2}^{\frac{d_2-1}{2}}(t_2) \right| \\ &\leq c(k_1 + 1)^{d_1} (k_2 + 1)^{(d_2-2)_+}, \end{aligned}$$

then of course  $|R_{(1,0)}| \leq c(k_1 + 1)^{d_1} (k_2 + 1)^{(d_2-2)_+}$ .

When  $d_1 = 1$ : By  $(\beta)$  of Lemma 8.1 and (2.6)

$$\begin{aligned} |\partial^{(1,0)}g(t)| &= \left| \left( C_{k_1}^0(t_1) \right)' C_{k_2}^{\frac{d_2-1}{2}}(t_2) \right| = \zeta(k_1, 1) \left| C_{k_1-1}^1(t_1) C_{k_2}^{\frac{d_2-1}{2}}(t_2) \right| \\ &\leq c(k_1 + 1)^2 (k_2 + 1)^{(d_2-2)_+}, \end{aligned}$$

and consequently  $|R_{(1,0)}| \leq c(k_1 + 1)^2 (k_2 + 1)^{(d_2-2)_+}$ .

Joining the above we have proved that for every  $d_1, d_2 \in \mathbb{N}$

$$|R_{(1,0)}| \leq c(k_1 + 1)^{(d_1-2)_++2} (k_2 + 1)^{(d_2-2)_+}.$$

In a similar way it turns out that  $|R_{(0,1)}| \leq c(k_1 + 1)^{(d_1-2)_+} (k_2 + 1)^{(d_2-2)_++2}$ . Summarizing the above we have the following upper bound for  $|\Delta_{k_1, k_2}(t, s)|$ :

$$|\Delta_{k_1, k_2}(t, s)| \leq c|t-s|(k_1+1)^{(d_1-2)_+} (k_2+1)^{(d_2-2)_+} ((k_1+1)^2 + (k_2+1)^2). \quad (8.34)$$

On the other hand, obviously by (2.6) we infer

$$|\Delta_{k_1, k_2}(t, s)| \leq 2 \left\| C_{k_1}^{\frac{d_1-1}{2}} \right\|_{\infty} \left\| C_{k_2}^{\frac{d_2-1}{2}} \right\|_{\infty} \leq c(k_1+1)^{(d_1-2)_+} (k_2+1)^{(d_2-2)_+}. \quad (8.35)$$

Let  $\delta \in (0, 1]$ . We interpolate (8.34) and (8.35) as follows

$$\begin{aligned} |\Delta_{k_1, k_2}(t, s)| &= |\Delta_{k_1, k_2}(t, s)|^{\delta} |\Delta_{k_1, k_2}(t, s)|^{1-\delta} \\ &\leq c|t-s|^{\delta} (k_1+1)^{(d_1-2)_+} (k_2+1)^{(d_2-2)_+} ((k_1+1)^2 + (k_2+1)^2)^{\delta}. \end{aligned}$$

The last replaced in (8.33) implies that

$$\|K_{\text{iso}}\|_{\mathcal{H}^{\delta}} \leq c \sum_{k_1, k_2=0}^{\infty} b_{k_1, k_2} (k_1 + 1)^{(d_1-2)_+} (k_2 + 1)^{(d_2-2)_+} ((k_1 + 1)^2 + (k_2 + 1)^2)^{\delta},$$

which is finite thanks to (4.3).

(ii) Let us now consider the mixed case. By (2.3) we have the expression

$$K_{\text{iso}}(t_1, t_2) - K_{\text{iso}}(s_1, t_2) = \sigma^2 \sum_{k_1, k_2=0}^{\infty} b_{k_1, k_2} C_{k_2}^{\frac{d_2-1}{2}}(t_2) \tilde{\Delta}_{k_1}(t_1, s_1), \quad (8.36)$$

where

$$\tilde{\Delta}_{k_1}(t_1, s_1) := C_{k_1}^{\frac{d_1-1}{2}}(t_1) - C_{k_1}^{\frac{d_1-1}{2}}(s_1).$$

By (2.6)  $|C_{k_2}^{\frac{d_2-1}{2}}(t_2)| \leq c(k_2 + 1)^{(d_2-2)_+}$  and  $\|\tilde{\Delta}_{k_1}\|_\infty \leq c(k_1 + 1)^{(d_1-2)_+}$ .  
 On the other hand

$$|\tilde{\Delta}_{k_1}(t_1, s_1)| = \left| \int_{s_1}^{t_1} \left(C_{k_1}^{\frac{d_1-1}{2}}\right)'(t) dt \right| \leq |t_1 - s_1| \left\| \left(C_{k_1}^{\frac{d_1-1}{2}}\right)' \right\|_\infty. \tag{8.37}$$

Repeating the arguments used for proving the claim (i), by using  $(\alpha)$  or  $(\beta)$  of Lemma 8.1 for  $d_1 > 1$  and  $d_1 = 1$  respectively and applying (2.6) we get

$$\left\| \left(C_{k_1}^{\frac{d_1-1}{2}}\right)' \right\|_\infty \leq c(k_1 + 1)^{(d_1-2)_++2}. \tag{8.38}$$

Let  $\delta_1 \in (0, 1]$ . By interpolation and using (8.37), (8.38) and (8.36) we arrive at

$$\begin{aligned} & |K_{\text{iso}}(t_1, t_2) - K_{\text{iso}}(s_1, t_2)| \\ & \leq c|t_1 - s_1|^{\delta_1} \sum_{k_1, k_2=0}^{\infty} b_{k_1, k_2} (k_1 + 1)^{(d_1-2)_++2\delta_1} (k_2 + 1)^{(d_2-2)_+}. \end{aligned}$$

The analogous estimate for the difference  $|K_{\text{iso}}(t_1, t_2) - K_{\text{iso}}(t_1, s_2)|$  and  $\delta_2 \in (0, 1]$  can be extracted similarly. Then by (4.2) we conclude that  $K_{\text{iso}} \in \mathcal{H}^{(\delta_1, \delta_2)}$ , thanks to the assumption (4.4) and the proof is complete.  $\square$

8.3.2. Proof of Theorem 4.3

We proceed to the proof of theorem 4.3:

*Proof.* Let  $\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2) \in \mathbb{T}^{d_1, d_2}$ . Then  $W_{\mathbf{xy}} := Z(\mathbf{x}) - Z(\mathbf{y})$  is a centred Gaussian random variable. Then  $\sigma^2 := \text{Var}(W_{\mathbf{xy}}) = \mathbb{E}(|Z(\mathbf{x}) - Z(\mathbf{y})|^2)$  and  $W_{\mathbf{xy}} = \sigma X$ , for  $X$  standard normally distributed. Consequently for every  $p \in \mathbb{N}$

$$\begin{aligned} \mathbb{E}(|Z(\mathbf{x}) - Z(\mathbf{y})|^{2p}) &= \mathbb{E}(|W_{\mathbf{xy}}|^{2p}) = \mathbb{E}(|\sigma X|^{2p}) \\ &= \sigma^{2p} \mathbb{E}(|X|^{2p}) = \sigma^{2p} c_{2p}, \end{aligned} \tag{8.39}$$

where  $c_{2p}$  the (known)  $2p$ -th moment of the standard normal  $X$ .

We can easily derive that the variance of  $W_{\mathbf{xy}}$  equals,

$$\begin{aligned} \sigma^2 &= \text{Var}(W_{\mathbf{xy}}) = \text{Var}Z(\mathbf{x}) + \text{Var}Z(\mathbf{y}) - 2\text{cov}(Z(\mathbf{x}), Z(\mathbf{y})) \\ &= 2\left(K_{\text{iso}}(1, 1) - K_{\text{iso}}(\langle x_1, y_1 \rangle_1, \langle x_2, y_2 \rangle_2)\right), \end{aligned}$$

thanks to the Remark 2.1 and (2.3).

Let now  $\delta \in (0, 1]$ . As  $\{b_{\mathbf{k}}\}$  satisfy (4.3), Theorem 4.2 asserts that  $K_{\text{iso}}$  belongs to  $\mathcal{H}^\delta$ , that is

$$\begin{aligned} |K_{\text{iso}}(1, 1) - K_{\text{iso}}(\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle)| &\leq c_\delta |(1, 1) - (\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle)|^\delta \\ &= c_\delta \sqrt{(1 - \langle x_1, y_1 \rangle)^2 + (1 - \langle x_2, y_2 \rangle)^2}^\delta \\ &= c_\delta \sqrt{(1 - \cos \rho_{\mathbb{S}^{d_1}}(x_1, y_1))^2 + (1 - \cos \rho_{\mathbb{S}^{d_2}}(x_2, y_2))^2}^\delta \\ &\leq c_\delta \sqrt{(\rho_{\mathbb{S}^{d_1}}(x_1, y_1))^4 + (\rho_{\mathbb{S}^{d_2}}(x_2, y_2))^4}^\delta \\ &\leq c_\delta (\rho_{\mathbb{S}^{d_1}}(x_1, y_1))^2 + (\rho_{\mathbb{S}^{d_2}}(x_2, y_2))^2)^\delta \\ &= c_\delta \rho(\mathbf{x}, \mathbf{y})^{2\delta}, \end{aligned} \tag{8.40}$$

where for the second inequality we used that  $1 - \cos \theta \leq \theta^2$  and for the third that  $\sqrt{\alpha^2 + \beta^2} \leq \alpha + \beta$ , for every  $\alpha, \beta \geq 0$ .

Combining (8.39) with (8.40) we get

$$\mathbb{E}(|Z(\mathbf{x}) - Z(\mathbf{y})|^{2p}) \leq c_{2p} 2^p c_\delta^p \rho(\mathbf{x}, \mathbf{y})^{2p\delta},$$

and the proof is complete. □

**Remark 8.3.** *From the above prove, it is apparent that the assumption of  $Z$  being Gaussian is necessary for getting Hölder bounds for the moments of  $|Z(\mathbf{x}) - Z(\mathbf{y})|$ . Note that such an assumption was not necessary for the previous results.*

### 8.4. Proof of Theorem 5.1

Before we prove the aforementioned result, we state the (quasi-)norms used to count the estimated risk.

Let  $(\Omega, \mathbb{F}, \mathbb{P})$  be a probability space and  $p > 0$ . We say that a (random) function  $f : \mathbb{T}^{d_1, d_2} \times \Omega \rightarrow \mathbb{R}$  belongs to the space  $L^p(\Omega, L^2(\mathbb{T}^{d_1, d_2}))$  when

$$\begin{aligned} \|f\|_{L^p(\Omega, L^2(\mathbb{T}^{d_1, d_2}))} &:= \left( \mathbb{E}(\|f\|_{L^2(\mathbb{T}^{d_1, d_2})}^p) \right)^{1/p} \\ &= \left( \int_\Omega \left( \int_{\mathbb{T}^{d_1, d_2}} |f(\mathbf{x}, \omega)|^2 d\mathbf{x} \right)^{p/2} d\mathbb{P}(\omega) \right)^{1/p} < \infty. \end{aligned} \tag{8.41}$$

Note further that

$$\|f\|_{L^p(\Omega, L^2(\mathbb{T}^{d_1, d_2}))} \leq \|f\|_{L^q(\Omega, L^2(\mathbb{T}^{d_1, d_2}))}, \quad \text{for every } 0 < p < q < \infty. \tag{8.42}$$

To verify the last we just use Hölder's inequality for the exponent  $r = \frac{q}{p} > 1$  on the probability space  $(\Omega, \mathbb{F}, \mathbb{P})$ . Indeed

$$\|f\|_{L^p(\Omega, L^2(\mathbb{T}^{d_1, d_2}))}^p = \int_\Omega \|f(\cdot, \omega)\|_{L^2(\mathbb{T}^{d_1, d_2})}^p d\mathbb{P}(\omega)$$

$$\begin{aligned}
&\leq \left( \int_{\Omega} \|f(\cdot, \omega)\|_{L^2(\mathbb{T}^{d_1, d_2})}^q d\mathbb{P}(\omega) \right)^{p/q} \left( \int_{\Omega} 1 d\mathbb{P}(\omega) \right)^{1-p/q} \\
&= \|f\|_{L^q(\Omega, L^2(\mathbb{T}^{d_1, d_2}))}^p.
\end{aligned}$$

*Proof.* Since  $X_{\mathbf{k}, \mathbf{m}}$  is a sequence of independent, standard normally distributed random variables and  $\{S_{k_i, m_i}^{d_i} : k_i \in \mathbb{N}_0, 1 \leq m_i \leq D_{k_i}(d_i)\}$ , is an orthonormal basis for  $L^2(\mathbb{S}^{d_i})$ ,  $i = 1, 2$ , we apply Fubini-Tonelli theorem to derive

$$\begin{aligned}
&\|Z - Z^N\|_{L^2(\Omega, L^2(\mathbb{T}^{d_1, d_2}))}^2 \\
&= \sum_{|\mathbf{k}| > N} B_{\mathbf{k}} \sum_{m_1=1}^{D_{k_1}(d_1)} \sum_{m_2=1}^{D_{k_2}(d_2)} \|S_{k_1, m_1}^{d_1}\|_{L^2(\mathbb{S}^{d_1})}^2 \|S_{k_2, m_2}^{d_2}\|_{L^2(\mathbb{S}^{d_2})}^2 \mathbb{E}(X_{\mathbf{k}, \mathbf{m}}^2) \\
&\leq c_{d_1} c_{d_2} c_* \sum_{|\mathbf{k}| > N} (1 + |\mathbf{k}|^2)^{-\varepsilon - \frac{d_1 + d_2}{2}} (1 + k_1)^{d_1 - 1} (1 + k_2)^{d_2 - 1} \\
&\leq c c_* \sum_{|\mathbf{k}| > N} (1 + |\mathbf{k}|^2)^{-\varepsilon - 1},
\end{aligned}$$

where for the first inequality we used (5.2) together with assumption (5.11) and for the second one the trivial inequalities  $1 + k_i \leq \sqrt{2}(1 + k_i^2)^{1/2} \leq c(1 + |\mathbf{k}|^2)^{1/2}$ ,  $i = 1, 2$ .

The last series, for the given range of the value of  $\varepsilon$ , can be bounded from above by the double integral

$$I_N := \iint_{|\mathbf{x}| \geq N-1} (1 + |\mathbf{x}|^2)^{-\varepsilon - 1} d\mathbf{x}.$$

We change to polar coordinates and we infer

$$I_N = \pi \int_{N-1}^{\infty} (1 + \rho^2)^{-\varepsilon - 1} 2\rho d\rho = \pi \int_{(N-1)^2}^{\infty} (1 + \varrho)^{-\varepsilon - 1} d\varrho,$$

using a second change to the variable  $\varrho = \rho^2$ .

Since now  $\varepsilon > 0$ , the last non proper integral equals

$$\int_{(N-1)^2}^{\infty} (1 + \varrho)^{-\varepsilon - 1} d\varrho = \frac{(1 + (N-1)^2)^{-\varepsilon}}{\varepsilon} \leq \frac{cN^{-2\varepsilon}}{\varepsilon},$$

which completes the proof of (5.12).

To show (5.13), we begin with the case  $p < 2$ . Applying the inequality (8.42) we derive

$$\|Z - Z^N\|_{L^p(\Omega, L^2(\mathbb{T}^{d_1, d_2}))} \leq \|Z - Z^N\|_{L^2(\Omega, L^2(\mathbb{T}^{d_1, d_2}))} \leq c \sqrt{\frac{c_*}{\varepsilon}} N^{-\varepsilon}.$$

We consider now  $p > 2$ . We can choose  $\nu = \nu_p \in \mathbb{N}$  such that  $2(\nu - 1) < p \leq 2\nu$ . The inequality (8.42) yields

$$\|Z - Z^N\|_{L^p(\Omega, L^2(\mathbb{T}^{d_1, d_2}))} \leq \|Z - Z^N\|_{L^{2\nu}(\Omega, L^2(\mathbb{T}^{d_1, d_2}))}. \tag{8.43}$$

By Corollary 2.17 in Da Prato and Zabczyk (1992) there exists a constant  $c = c_\nu > 0$  such that

$$\|Z - Z^N\|_{L^{2\nu}(\Omega, L^2(\mathbb{T}^{d_1, d_2}))} \leq c_\nu \|Z - Z^N\|_{L^2(\Omega, L^2(\mathbb{T}^{d_1, d_2}))}. \tag{8.44}$$

The combination of (8.43), (8.44) and (5.12) leads to (5.13).

For the last claim, we have to prove that

$$\mathbb{P}(\|Z - Z^N\|_{L^2(\mathbb{T}^{d_1, d_2})} \geq N^{-\gamma}, \text{ for infinity many } N \in \mathbb{N}) = 0. \tag{8.45}$$

This will be a consequence of the Borel-Cantelli lemma. It suffices to prove that

$$\sum_{N=1}^{\infty} \mathbb{P}(\|Z - Z^N\|_{L^2(\mathbb{T}^{d_1, d_2})} \geq N^{-\gamma}) < \infty. \tag{8.46}$$

By (the general) Chebyshev’s inequality, as in page 193 of Folland (2009) applied in the measure space  $(\Omega, \mathbb{F}, \mathbb{P})$ , and (5.13) we derive

$$\begin{aligned} \mathbb{P}(\|Z - Z^N\|_{L^2(\mathbb{T}^{d_1, d_2})} \geq N^{-\gamma}) &\leq N^{\gamma p} \mathbb{E}(\|Z - Z^N\|_{L^2(\mathbb{T}^{d_1, d_2})}^p) \\ &\leq c_p^p \left(\frac{CC_*}{\varepsilon}\right)^{p/2} N^{-(\varepsilon-\gamma)p}, \end{aligned} \tag{8.47}$$

for every  $p > 0$  and every  $N \geq N_0$ .

We choose now  $p > 1/(\varepsilon - \gamma)$ , this is allowed since  $\gamma < \varepsilon$ . For this choice we obtain

$$\sum_{N=N_0}^{\infty} N^{-(\varepsilon-\gamma)p} < \infty,$$

which together with (8.47) leads to (8.46) and completes the proof. □

### Funding

Alfredo Alegría acknowledges the funding of the National Agency for Research and Development of Chile, through grant ANID/FONDECYT/INICIACIÓN/No. 11190686.

Galatia Cleanthous has been supported by the individual grant “New function spaces in harmonic analysis and their applications in Statistics”, from the University of Cyprus.

## References

- Adams, R. A. and Fournier, J. J. F.: 2003, *Sobolev spaces*, Vol. 140 of *Pure and Applied Mathematics (Amsterdam)*, second edn, Elsevier/Academic Press, Amsterdam. [MR2424078](#)
- Alegría, A., Bissiri, P. G., Cleanthous, G., Porcu, E. and White, P.: 2021, Multivariate isotropic random fields on spheres: nonparametric bayesian modeling and  $l^p$  fast approximations, *Electronic Journal of Statistics* **15**(1), 2360–2392. [MR4255327](#)
- Atkinson, S., Sivapalan, M., Viney, N. and Woods, R.: 2003, Predicting space–time variability of hourly streamflow and the role of climate seasonality: Mahurangi catchment, new zealand, *Hydrological Processes* **17**(11), 2171–2193.
- Bachoc, F., Peron, A. and Porcu, E.: 2022, Multivariate gaussian random fields over generalized product spaces involving the hypertorus. [MR4511141](#)
- Celleri, R., Willems, P., Buytaert, W. and Feyen, J.: 2007, Space–time rainfall variability in the paute basin, ecuadorian andes, *Hydrological Processes: An International Journal* **21**(24), 3316–3327.
- Clarke, J., Alegría, A. and Porcu, E.: 2018, Regularity properties and simulations of Gaussian random fields on the sphere cross time, *Electronic Journal of Statistics* **12**, 399–426. [MR3763911](#)
- Cleanthous, G.: 2023, On the properties of multivariate isotropic random fields on the ball, *Manuscript. Submitted for publication.* [MR4255327](#)
- Cleanthous, G., Georgiadis, A. G., Lang, A. and Porcu, E.: 2020, Regularity, continuity and approximation of isotropic Gaussian random fields on compact two-point homogeneous spaces, *Stochastic Process. Appl.* **130**(8), 4873–4891. **URL:** <https://doi.org/10.1016/j.spa.2020.02.003> [MR4108475](#)
- Cleanthous, G., Porcu, E. and White, P.: 2021, Regularity and approximation of gaussian random fields evolving temporally over compact two-point homogeneous spaces, *TEST* pp. 1–25. **URL:** <https://doi.org/10.1007/s11749-021-00755-1> [MR4346808](#)
- Da Prato, G. and Zabczyk, J.: 1992, *Stochastic equations in infinite dimensions*, Vol. 44 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge. **URL:** <https://doi.org/10.1017/CBO9780511666223> [MR1207136](#)
- Diamond, H. J., Karl, T. R., Palecki, M. A., Baker, C. B., Bell, J. E., Leeper, R. D., Easterling, D. R., Lawrimore, J. H., Meyers, T. P. and Helfert, M. R.: 2013, US climate reference network after one decade of operations: Status and assessment, *Bulletin of the American Meteorological Society* **94**(4), 485–498.
- Emery, X., Peron, P. and Porcu, E.: 2022, A catalogue of covariance models on hypertori, *Manuscript. Submitted for publication.* [MR0961862](#)
- Ezzat, A. A., Jun, M. and Ding, Y.: 2019, Spatio-temporal short-term wind forecast: A calibrated regime-switching method, *The Annals of Applied Statistics* **13**(3), 1484–1510. [MR4019147](#)
- Folland, G. B.: 2009, *Real Analysis: Modern Techniques and Their Applications*, John Wiley & Sons. [MR1681462](#)

- Franses, P. H.: 1991, Seasonality, non-stationarity and the forecasting of monthly time series, *International Journal of forecasting* **7**(2), 199–208.
- Gandoman, F. H., Aleem, S. H. A., Omar, N., Ahmadi, A. and Alenezi, F. Q.: 2018, Short-term solar power forecasting considering cloud coverage and ambient temperature variation effects, *Renewable Energy* **123**, 793–805.
- Gelman, A.: 2006, Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper), *Bayesian Analysis* **1**(3), 515–534. [MR2221284](#)
- Gneiting, T., Larson, K., Westrick, K., Genton, M. G. and Aldrich, E.: 2006, Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space-time method, *Journal of the American Statistical Association* **101**(475), 968–979. [MR2324108](#)
- Guella, J., Menegatto, V. and Peron, A.: 2015, An extension of a theorem of Schoenberg to products of spheres, *Banach Journal of Mathematical Analysis* **435**, 286–301. [MR3543906](#)
- Haario, H., Saksman, E. and Tamminen, J.: 2001, An adaptive Metropolis algorithm, *Bernoulli* **7**(2), 223–242. [MR1828504](#)
- Held, L. and Paul, M.: 2012, Modeling seasonality in space-time infectious disease surveillance data, *Biometrical Journal* **54**(6), 824–843. [MR2993630](#)
- Hering, A. S. and Genton, M. G.: 2010, Powering up with space-time wind forecasting, *Journal of the American Statistical Association* **105**(489), 92–104. [MR2757195](#)
- Hylleberg, S.: 1992, *Modelling seasonality*, Oxford University Press.
- Jona-Lasinio, G., Gelfand, A. and Jona-Lasinio, M.: 2012, Spatial analysis of wave direction data using wrapped Gaussian processes, *The Annals of Applied Statistics* **6**(4), 1478–1498. [MR3058672](#)
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G. and Woollen, J.: 1996, The NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American meteorological Society* **77**(3), 437–471.
- Karbalae, N., Hsu, K., Sorooshian, S. and Braithwaite, D.: 2017, Bias adjustment of infrared-based rainfall estimation using passive microwave satellite rainfall data, *Journal of Geophysical Research: Atmospheres* **122**(7), 3859–3876.
- Kerkyacharian, G., Ogawa, S., Petrushev, P. and Picard, D.: 2018, Regularity of Gaussian processes on Dirichlet spaces, *Constr. Approx.* **47**(2), 277–320. [URL: https://doi.org/10.1007/s00365-018-9416-8](https://doi.org/10.1007/s00365-018-9416-8) [MR3769279](#)
- Lanfredi, M., Coluzzi, R., Imbrenda, V., Macchiato, M. and Simoniello, T.: 2020, Analyzing space–time coherence in precipitation seasonality across different european climates, *Remote Sensing* **12**(1), 171.
- Lang, A. and Schwab, C.: 2015, Isotropic gaussian random fields on the sphere: Regularity, fast simulation and stochastic partial differential equations, **25**(6), 3047–3094. [MR3404631](#)
- Ma, C. and Malyarenko, A.: 2020, Time-varying isotropic vector random fields on compact two-point homogeneous spaces, *Journal of Theoretical Probability* **33**, 319–339. [MR4064303](#)

- Malyarenko, A.: 2004, Abelian and tauberian theorems for random fields on two-point homogeneous spaces, *Theory of Probability and Mathematical Statistics* **69**, 115–127. [MR2110910](#)
- Malyarenko, A. and Olenko, A. Y.: 1992, Multidimensional covariant random fields on commutative locally compact groups, *Ukrainian Mathematical Journal* **44**(11), 1384–1389. [MR1213893](#)
- Marinucci, D. and Peccati, G.: 2011, *Random Fields on the Sphere, Representation, Limit Theorems and Cosmological Applications*, Cambridge, New York. [MR2840154](#)
- Marsden, J. and Tromba, A.: 2011, *Vector Calculus*, W. H. Freeman. **URL:** <https://books.google.com.cy/books?id=b3oVDAEACAAJ>
- Mastrantonio, G., Jona Lasinio, G. and Gelfand, A.: 2016, Spatio-temporal circular models with non-separable covariance structure, *Test* **25**, 331–350. [MR3493522](#)
- Mastrantonio, G., Lasinio, G. J., Pollice, A., Capotorti, G., Teodonio, L., Genova, G. and Blasi, C.: 2019, A hierarchical multivariate spatio-temporal model for clustered climate data with annual cycles, *Annals of Applied Statistics* **13**(2), 797–823. [MR3963553](#)
- Polson, N. G. and Scott, J. G.: 2012, On the half-Cauchy prior for a global scale parameter, *Bayesian Analysis* **7**(4), 887–902. [MR3000018](#)
- Porcu, E., Furrer, R. and Nychka, D.: 2020, 30 years of space–time covariance functions, *Wiley Interdisciplinary Reviews: Computational Statistics* p. e1512. [MR4218945](#)
- Porcu, E. and White, P. A.: 2022, Random fields on the hypertorus: Covariance modeling and applications, *Environmetrics* p. e2701. [MR4376811](#)
- Pounds, J. A., Fogden, M. P. and Campbell, J. H.: 1999, Biological Response to Climate Change on a Tropical Mountain, *Nature* **398**(6728), 611.
- Reed, M. and Simon, B.: 1980, *Methods of Modern Mathematical Physics: Functional analysis*, number  $\tau$ . 1 in *Methods of Modern Mathematical Physics*, Academic Press. **URL:** <https://books.google.com.cy/books?id=bvuRuwuFBWwC>
- Roberts, G. O. and Rosenthal, J. S.: 2009, Examples of adaptive MCMC, *Journal of Computational and Graphical Statistics* **18**(2), 349–367. [MR2749836](#)
- Schmeisser, H.-J. and Triebel, H.: 1987, *Topics in Fourier Analysis and Function Spaces*, A Wiley-Interscience Publication, John Wiley & Sons, Ltd., Chichester. [MR0891189](#)
- Shirota, S. and Gelfand, A.: 2017a, Space and circular time log Gaussian Cox processes with application to crime event data, *Annals of Applied Statistics* **11**(2), 481–503. [MR3693535](#)
- Shirota, S. and Gelfand, A. E.: 2017b, Space and circular time log Gaussian Cox processes with application to crime event data, *The Annals of Applied Statistics* **11**(2), 481–503. [MR3693535](#)
- Szegő, G.: 1939, *Orthogonal Polynomials*, Vol. XXIII of *COLLOQUIUM PUBLICATIONS*, American Mathematical Society. [MR0000077](#)
- Wang, F. and Gelfand, A. E.: 2014, Modeling space and space-time directional data using projected Gaussian processes, *Journal of the American Statistical Association* **109**(508), 1565–1580. [MR3293610](#)

- Wang, Y., Wang, C., Shi, C. and Xiao, B.: 2018, Short-term cloud coverage prediction using the arima time series model, *Remote Sensing Letters* **9**(3), 274–283.
- West, M. and Harrison, J.: 2006, *Bayesian forecasting and dynamic models*, Springer Science & Business Media. [MR1020301](#)
- White, P. and Porcu, E.: 2019a, Nonseparable covariance models on circles cross time: A study of Mexico City ozone, *Environmetrics* p. e2558. [MR3999535](#)
- White, P. and Porcu, E.: 2019b, Towards a complete picture of stationary covariance functions on spheres cross time, *Electronic Journal of Statistics* **13**, 2566–2594. [MR3988087](#)
- Wylie, D., Jackson, D. L., Menzel, W. P. and Bates, J. J.: 2005, Trends in Global Cloud Cover in Two Decades of HIRS Observations, *Journal of Climate* **18**(15), 3021–3031.
- Xu, Y.: 2018, Approximation by polynomials in sobolev spaces with jacobi weight, *Journal of Fourier Analysis and Applications* **24**(6), 1438–1459. [MR3881837](#)
- Yadrenko, M. u.: 1983, *Spectral theory of random fields*, Translation Series in Mathematics and Engineering, Optimization Software, Inc., Publications Division, New York. Translated from the Russian. [MR0697386](#)