# Semiparametric Estimation of Cross-covariance Functions for Multivariate Random Fields

Ghulam A. Qadir<sup>1</sup> and Ying Sun<sup>1</sup>

November 7, 2019

#### Abstract

The prevalence of spatially referenced multivariate data has impelled researchers to develop a procedure for the joint modeling of multiple spatial processes. This ordinarily involves modeling marginal and cross-process dependence for any arbitrary pair of locations using a multivariate spatial covariance function. However, building a flexible multivariate spatial covariance function that is nonnegative definite is challenging. Here, we propose a semiparametric approach for multivariate spatial covariance function estimation with approximate Matérn marginals and highly flexible cross-covariance functions via their spectral representations. The flexibility in our cross-covariance function arises due to B-spline based specification of the underlying coherence functions, which in turn allows us to capture non-trivial cross-spectral features. We then develop a likelihood-based estimation procedure and perform multiple simulation studies to demonstrate the performance of our method, especially on the coherence function estimation. Finally, we analyze particulate matter concentrations  $(PM_{2.5})$  and wind speed data over the North-Eastern region of the United States, where we illustrate that our proposed method outperforms the commonly used full bivariate Matérn model and the linear model of coregionalization for spatial prediction.

Some key words: Coherence, co-kriging, Matérn covariance, nonnegative definite, multivariate spatial data.

<sup>&</sup>lt;sup>1</sup> CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia. E-mail: ghulam.qadir@kaust.edu.sa; ying.sun@kaust.edu.sa

### 1 Introduction

Recent technological advances have led to the exposition of spatially indexed multivariate data in a wide range of applications, such as, for instance, in geophysical, environmental and atmospheric sciences, to name but a few (Sain et al., 2011; Greasby and Sain, 2011). This has motivated and facilitated researchers to jointly model multiple spatial processes for gaining scientific insights into the dynamics within each variable and between distinct variables. Modeling spatial data conventionally involves quantifying spatial dependence through valid covariance functions, which call for marginal and cross-covariance functions in the case of multivariate spatial data. Let  $\mathbf{X}(\mathbf{s}) = \begin{pmatrix} X_1(\mathbf{s}), \dots, X_p(\mathbf{s}) \end{pmatrix}^T$  be a p-variate zero mean Gaussian random field defined on a spatial domain  $\mathcal{D} \subset \mathbb{R}^d, d \geq 1$ . Under the assumption of second-order stationarity, the covariance functions associated with  $\mathbf{X}(\mathbf{s})$  are defined as:

$$C_{ij}(\mathbf{s}_1 - \mathbf{s}_2) = \mathbb{E}[X_i(\mathbf{s}_1)X_j(\mathbf{s}_2)], \ i, j = 1, \dots, p, \ \mathbf{s}_1, \mathbf{s}_2 \in \mathcal{D},$$

where  $C_{ii}(\cdot)$ ,  $i=1,\ldots,p$  are the marginal covariance functions that describe the spatial dependence of the  $i^{th}$  process component  $\{X_i(\mathbf{s}): \mathbf{s} \in \mathcal{D}\}$ , whereas  $C_{ij}(\cdot)$ ,  $1 \leq i \neq j \leq p$ , often termed as the cross-covariance function, describes the spatial dependence between  $i^{th}$  and  $j^{th}$  process components. If the covariance function depends on the spatial lag  $\mathbf{s}_1 - \mathbf{s}_2$  only through its Euclidean norm, i.e.,  $\|\mathbf{s}_1 - \mathbf{s}_2\|$ , then the random field  $\mathbf{X}$  is said to be isotropic. The assumptions of stationarity and isotropy state that the covariances are invariant under rigid transformations of the coordinates, and hence may seem unrealistic for many applications. However, this class of models is important, as they form the basic ingredients for more complex and sophisticated non-stationary and anisotropic models. Construction of a valid and flexible model for multivariate covariances entails the difficulty of guaranteeing the nonnegative definiteness, or the nonnegative definite covariance matrix  $\Sigma$  for the random

vector  $(\mathbf{X}(\mathbf{s}_1)^{\mathrm{T}}, \dots, \mathbf{X}(\mathbf{s}_n)^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{np}$ . Specifically, the main challenge is to build a flexible model for  $C_{ij}(\cdot)$  that yields  $\Sigma$ , and ensures  $\mathbf{c}^{\mathrm{T}}\Sigma\mathbf{c} \geq 0$  for any nonzero vector  $\mathbf{c} \in \mathbb{R}^{np}$ , any set of spatial coordinates  $\mathbf{s}_1, \dots, \mathbf{s}_n$ , and any positive integer n.

The growing interest in building models for multivariate spatial fields has led to the development of a fairly rich literature in the last few decades, and a comprehensive summary of the existing approaches can be found in the review paper Genton and Kleiber (2015). Many of these models have their genesis in combining univariate covariance functions. Perhaps the most rudimentary modeling approach is to introduce separability by setting  $C_{ij}(\mathbf{s}_1 - \mathbf{s}_2) = \mathbf{A}C(\mathbf{s}_1 - \mathbf{s}_2)$ , where  $\mathbf{A}$  is a  $p \times p$  nonnegative definite matrix, and  $C(\cdot)$  is any valid univariate covariance function (Mardia and Goodall, 1993; Helterbrand and Cressie, 1994; Bhat et al., 2010). Such a specification enforces the same shape of covariance function for all the marginal and cross components, which inhibits its use for modeling complex dependencies. The linear model of coregionalization (LMC) is another univariate covariance function based model, which decomposes the multivariate random field as a linear combination of independent univariate random fields (Goulard and Voltz, 1992; Schmidt and Gelfand, 2003; Wackernagel, 2003; Zhang, 2007). The roughest underlying univariate field in the LMC governs the smoothness of all the components of a multivariate random field, making it inflexible for modeling distinct smoothness in components. Apanasovich and Genton (2010) introduced an approach that can produce flexible multivariate models with distinct smoothnesses in each component while controlling nonseparability. However, this approach involves representing a multivariate random field as a univariate random field in a higher dimensional Euclidean space, which in turn requires the estimation of latent dimensions for each component. Moreover, kernel convolution (Ver Hoef and Barry, 1998; Ver Hoef et al., 2004) and covariance convolution (Gaspari and Cohn, 1999; Gaspari et al., 2006; Majumdar and Gelfand, 2007) methods are other popular univariate covariance function based approaches for building valid cross-covariance functions.

In the context of univariate random fields, the Matérn class (Matérn, 1986; Guttorp and Gneiting, 2006) has become a preferred choice for modeling covariances, primarily due to its smoothness controlling parameter that governs the correlations at small distances. Gneiting et al. (2010) extended this class for multivariate random fields and introduced a matrix-valued covariance function such that both marginal and cross-covariances are of the Matérn type. For the bivariate case (p = 2), these authors provided full characterization of the parameter values that lead to a valid full bivariate Matérn model, whereas for p > 2, they specified a parsimonious multivariate Matérn model that admits only common spatial scale parameters and constrained smoothness parameters. Further generalization of this idea in Apanasovich et al. (2012) provided sufficient validity conditions on the parameter space for any p > 1 and introduced the flexible multivariate Matérn model.

Recently Kleiber (2017) analyzed the spectral properties of a number of existing multivariate spatial models, and pointed out that many of them are not sufficiently flexible to
capture non-trivial coherence between components. For instance, separable, kernel convolution and the parsimonious multivariate Matérn model impose constant coherence between
components. The full bivariate Matérn model although is quite flexible as its parameters can
control the decay rate of coherence at high frequency, as well as supervise the frequency of
the greatest coherence, its flexibility is limited to its parametric form of coherence function
that can capture only certain shapes of coherence and not beyond that. For example, the
full bivariate Matérn model cannot comprehend a multivariate process with an underlying
coherence function that shows oscillations or multiple peaks. In this article, we propose
a semiparametric multivariate spatial covariance model with highly flexible underlying coherence functions. The proposed model specifies an approximate Matérn marginal for each
component and highly flexible cross-covariances for every pair of components. We specify

the coherence functions as a linear combination of cubic splines (B-splines of order 4). Such a specification enables our coherence functions to represent a wide range of smooth curves and allows us to model non-trivial coherence between every pair of process components. The flexibility of our coherence functions is also reflected in the corresponding cross-covariances in the space domain. Additionally, we enact the exact likelihood based inference method jointly for both the parametric marginal and nonparametric coherence function in the proposed model, for both the regularly and irregularly spaced multivariate spatial data.

The rest of our paper is organized as follows. In Section 2, we describe the construction of our model and its properties. We also provide sufficient conditions on B-spline coefficients to ensure the validity of our model. We perform multiple simulation studies to explore the performance of our model in Section 3. In particular, we estimate the coherence of the processes generated from the full bivariate Matérn model and the LMC, using our model with maximum likelihood estimation (MLE). In Section 4, we illustrate the application of our proposed model on a bivariate dataset of particulate matter concentrations (PM<sub>2.5</sub>) and wind speed over the North-Eastern region of the United States. We compare our model with the full bivariate Matérn model and the LMC on the basis of commonly used prediction scores. We conclude in Section 5 with a discussion and potential future extension.

# 2 Multivariate Spatial Model

In this section, we introduce our proposed semiparametric model through its origin in the spectral domain, and provide sufficient conditions to ensure its validity. We revisit some notions and concepts of spectral domain in Section 2.1 that are crucial to our model construction in Section 2.2.

#### 2.1 Spectral Representation

Let  $\mathbf{X}(\mathbf{s}) = (X_1(\mathbf{s}), \dots, X_p(\mathbf{s}))^{\mathrm{T}}$  be a p-variate weakly stationary random field defined on a spatial domain  $\mathcal{D} \subset \mathbb{R}^d$ ,  $d \geq 1$ , and  $\mathbf{C}(\mathbf{h}) = \{C_{ij}(\mathbf{h})\}_{i,j=1}^p$  be a matrix valued covariance function for  $\mathbf{X}$  such that  $C_{ij}(\mathbf{h}) = \mathrm{Cov}(X_i(\mathbf{s}), X_j(\mathbf{s} + \mathbf{h}))$ . The validity of  $\mathbf{C}(\cdot)$  is generally ensured by using the Cramér's Theorem (Cramér, 1940) in its spectral density version (Wackernagel, 2003, p. 215; Kleiber, 2017) which states that:

The necessary and sufficient condition for the matrix valued function  $\mathbf{C}: \mathbb{R}^d \to \mathbb{C}^{p \times p}$ ,  $\mathbf{C}(\mathbf{h}) = \{C_{ij}(\mathbf{h})\}_{i,j=1}^p$  to be nonnegative definite is its representation as

$$C_{ij}(\mathbf{h}) = \int_{\mathbb{R}^d} \exp(i\mathbf{u}^{\mathrm{T}}\mathbf{h}) g_{ij}(\mathbf{u}) d\mathbf{u}, \quad (i = \sqrt{-1}),$$
(1)

for i, j = 1, ..., p such that the matrix  $\mathbf{g}(\mathbf{u}) = \{g_{ij}(\mathbf{u})\}_{i,j=1}^p$  is nonnegative definite for all  $\mathbf{u} \in \mathbb{R}^d$ .

Here the functions  $g_{ij}: \mathbb{R}^d \to \mathbb{C}$ , such that  $g_{ij}(\mathbf{u}) = \overline{g_{ji}(\mathbf{u})}$ , are the spectral densities for marginal and cross-covariance functions, that admit the d-dimensional frequencies  $\mathbf{u}$  as an argument and return a complex or real value. Under the assumption of isotropy,  $g_{ij}(\mathbf{u}_1) = g_{ij}(\mathbf{u}_2) \; \forall \; i, j = 1, \ldots, p \; \text{whenever} \; \|\mathbf{u}_1\| = \|\mathbf{u}_2\| \; \text{and therefore} \; (1) \; \text{can be reduced to} \; \text{a one dimensional integral (Stein, 1999, p. 42-44):}$ 

$$C_{ij}(\mathbf{h}) = \int_0^\infty \|\mathbf{h}\| \left(\frac{2\pi\omega}{\|\mathbf{h}\|}\right)^{\kappa+1} J_{\kappa}(\omega \|\mathbf{h}\|) f_{ij}(\omega) d\omega, \tag{2}$$

where  $\omega = \|\mathbf{u}\| \ge 0$ ,  $\kappa = \frac{d}{2} - 1$ ,  $J_{\kappa}(\cdot)$  is a Bessel function of the first kind of order  $\kappa$  (Watson, 1944) and  $f_{ij} : \mathbb{R} \to \mathbb{C}$  are the isotropic spectral densities such that  $g_{ij}(\mathbf{u}) = f_{ij}(\|\mathbf{u}\|)$ ,  $\forall \mathbf{u} \in \mathbb{R}^d$ ,  $i, j = 1, \ldots, p$ .

For given spectral densities  $\{g_{ij}(\cdot), i, j = 1, \dots, p\}$ , the coherence between the  $i^{th}$  and

 $j^{th}$  components of the process **X** at a particular frequency **u** is defined as:

$$\gamma_{ij}(\mathbf{u}) = \frac{g_{ij}(\mathbf{u})}{\sqrt{g_{ii}(\mathbf{u})g_{jj}(\mathbf{u})}} \,\forall \, 1 \le i \ne j \le p.$$
(3)

Coherence functions in general can be complex-valued depending on the codomain of the spectral densities  $\{g_{ij}(\cdot), i, j = 1, ..., p\}$ , and therefore absolute coherence functions  $|\gamma_{ij}(\cdot)|$  are examined in practice. The isotropic version of the coherence function can be obtained trivially by replacing the argument  $\mathbf{u}$  by  $\omega$  and functions  $g_{ij}$  by  $f_{ij}$  in (3). For a more detailed account on coherence functions in spatial case, we refer readers to Kleiber (2017). In the subsequent sections, we develop our semiparametric multivariate covariance functions using the above-mentioned notions.

#### 2.2 Semiparametric Multivariate Spatial Model

We consider the isotropic spectral densities  $\{f_{ij}(\cdot), i, j = 1, ..., p\}$  up to a certain sufficiently large threshold frequency  $\omega_t$ . We choose the marginal spectral densities  $\{f_{ii}(\cdot), i = 1, ..., p\}$  to be of Matérn type (Gneiting et al., 2010, A.1), truncated for frequencies greater than  $\omega_t$ , i.e,

$$f_{ii}(\omega|\sigma_i,\nu_i,a_i) = \sigma_i^2 \frac{\Gamma(\nu_i + d/2)a_i^{2\nu_i}}{\Gamma(\nu_i)\pi^{d/2}(a_i^2 + \omega^2)^{\nu_i + d/2}}, \ 0 \le \omega \le \omega_t, \ \sigma_i,\nu_i,a_i > 0.$$
 (4)

The untruncated version of (4) corresponds to the spectral density of the isotropic Matérn covariance function (Matérn, 1986; Guttorp and Gneiting, 2006):

$$M(\mathbf{h}|\sigma,\nu,a) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} (a\|\mathbf{h}\|)^{\nu} K_{\nu}(a\|\mathbf{h}\|),$$

where  $\sigma > 0$  is the marginal standard deviation, a > 0 represents a spatial scale parameter,  $\nu > 0$  is a smoothness parameter and  $K_{\nu}$  is a modified Bessel function of the second kind of order  $\nu$ .

For given marginal spectral densities in (4), we specify the cross-spectral densities

 $\{f_{ij}(\cdot), 1 \leq i < j \leq p\}$  using the linear combination of B-splines as follows:

$$f_{ij}(\omega|f_{ii}, f_{jj}, \mathbf{S}_{ij}, K) = \sum_{k=-3}^{K} b_k^{(ij)} B_k(\omega) \sqrt{f_{ii}(\omega) f_{jj}(\omega)}, \ 0 \le \omega \le \omega_t,$$
 (5)

where  $B_k$ 's are the cubic splines (B-splines of order 4) (De Boor, 2001, chapter IX; Im et al., 2006), for a sequence of uniform knots  $(-3\Delta, \ldots, 0, \Delta, 2\Delta, \ldots, (K+1)\Delta)$  such that  $\omega_t \in (K\Delta, (K+1)\Delta]$ , and  $\{b_k^{(ij)}, k=-3,\ldots,K, 1 \leq i < j \leq p\}$  are the B-spline coefficients. We begin the B-splines combinations from k=-3 to k=K in order to include all the B-splines that have support on the interval  $[0,\omega_t]$ . Im et al. (2007) used a similar B-spline representation for defining the nonparametric part of their univariate semiparametric spectral density. Here K supervises the number of knots,  $\Delta$  represents its uniform spacing and  $\mathbf{S}_{ij}$  constitutes the set of coefficients required to fully specify the B-spline part of (5), i.e.,  $\mathbf{S}_{ij} = \{b_k^{(ij)}, k=-3,\ldots,K\}, 1 \leq i < j \leq p$ . Note that the cross-spectral densities specified in (5) are real valued, therefore  $f_{ij}(\cdot) = f_{ji}(\cdot)$ ,  $\forall 1 \leq i < j \leq p$ , and consequently  $\mathbf{S}_{ij} = \mathbf{S}_{ji}, \ \forall 1 \leq i < j \leq p$ . We choose the B-spline of order 4, however, a higher order B-spline can also be incorporated in (5) with only slight modifications.

Following the definition in (3), the coherence between  $i^{th}$  and  $j^{th}$  process components at frequency  $\omega$  for the spectral densities specified in (4) and (5) is given as:

$$\gamma_{ij}(\omega) = \sum_{k=-3}^{K} b_k^{(ij)} B_k(\omega), \ 0 \le \omega \le \omega_t.$$
 (6)

Here, our specified spectral densities lead to fully nonparametric coherence functions based on the linear combination of B-splines that can accommodate a wide range of smooth functions, and therefore induces a great deal of flexibility in our proposed coherence model that can be controlled by the value of  $\Delta$ . The smaller values of  $\Delta$  produce more flexible coherence functions, however, it makes the estimation computationally challenging due to a large number of B-spline coefficients, whereas the large values of  $\Delta$  generate relatively less flexible

coherence functions, but the estimation is computationally more feasible due to a smaller number of B-spline coefficients. For an appropriate choice of  $\Delta$ , our proposed approach can model coherence functions that are beyond the comprehension of existing multivariate models.

In order to obtain the multivariate covariance functions from any given isotropic marginal and cross spectral densities, we resort to the integral (2), also known as the Hankel transform of the order  $\kappa$ . However, in our proposed framework, integral (2) cannot be computed for the spectral densities defined in (4) and (5) because of their truncation to  $\omega_t$  and unknown closed form solutions. Consequently, we choose a small value of  $\delta$  to define a discrete set of frequencies  $\mathcal{F} = \{\delta, \dots, m\delta\}$  such that  $m\delta = \omega_t$ , and then we compute the following finite sum approximation of (2) to obtain the multivariate spatial covariance function:

$$\begin{cases}
C_{ii}(\mathbf{h}) = \sum_{\omega \in \mathcal{F}} \frac{(2\pi\omega)^{\kappa+1}}{\|\mathbf{h}\|^{\kappa}} J_{\kappa}(\omega \|\mathbf{h}\|) f_{ii}(\omega | \sigma_{i}, \nu_{i}, a_{i}) \delta, & i = 1, \dots, p \\
C_{ij}(\mathbf{h}) = \sum_{\omega \in \mathcal{F}} \frac{(2\pi\omega)^{\kappa+1}}{\|\mathbf{h}\|^{\kappa}} J_{\kappa}(\omega \|\mathbf{h}\|) f_{ij}(\omega | f_{ii}, f_{jj}, \mathbf{S}_{ij}, K) \delta, & 1 \leq i \neq j \leq p,
\end{cases}$$
(7)

where  $\{f_{ij}(\cdot), i, j = 1, ..., p\}$  corresponds to the spectral densities defined in (4) and (5). The finite sum based approach has been commonly used to propose nonparametric univariate covariance functions (Shapiro and Botha, 1991; Genton and Gorsich, 2002; Gorsich and Genton, 2004), however, its extension to a multivariate setting is not very popular yet. For a reasonably small value of  $\delta$  (or large value of m), a large value of  $\omega_t$  and an appropriate normalization of finite sums, the marginal covariance functions  $C_{ii}(\cdot)$  in (7) are numerically equivalent to the corresponding exact Matérn covariance functions, and hence the parameters  $(\sigma_i, \nu_i, a_i, i = 1, ..., p)$  retain their interpretations of the exact Matérn. In order to ensure the validity of the cross-covariances  $C_{ij}(\cdot)$  in (7), we need to impose certain constraints on the set of B-spline coefficients  $\mathbf{S}_{ij}$ ,  $1 \le i \ne j \le p$ . In Theorem 1, we provide sufficient conditions for the validity of our proposed multivariate covariance function  $\mathbf{C}(\mathbf{h}) = \{C_{ij}(\mathbf{h})\}_{i,j=1}^p$  in (7):

**Theorem 1.** Let  $\boldsymbol{\beta}_k = \{b_k^{(ij)}\}_{i,j=1}^p$ ,  $k = -3, -2, \ldots, K$  be the  $p \times p$  symmetric matrices with diagonal elements  $\{b_k^{(ii)} = 1 \ \forall \ i = 1, 2, \ldots, p, \ k = -3, -2, \ldots, K\}$ , then the matrix-valued covariance function  $\boldsymbol{C}(\boldsymbol{h}) = \{C_{ij}(\boldsymbol{h})\}_{i,j=1}^p$  in (7) is valid if the matrices  $\{\boldsymbol{\beta}_k, \ k = -3, \ldots, K\}$  are nonnegative definite.

Figure 1 shows a realization of a trivariate zero mean Gaussian random field X, simulated from our proposed model (7) with threshold frequency  $\omega_t = 4.5$ , and m = 990for discretization of frequencies. The coherence functions (shown in Figure 1(a)) are generated from suitably selected  $\mathbf{S}_{ij}$ ,  $1 \leq i \neq j \leq p$ , such that  $X_3$  has the highest coherence with  $X_2$  and lowest coherence with  $X_1$ , at all frequencies. The marginal parameters  $(\sigma_1 = \sigma_2 = \sigma_3 = 1, \nu_1 = 1, a_1 = 1, \nu_2 = 2, a_2 = 0.5, \nu_3 = 2.5, a_3 = 0.4)$  induce distinct features in the process components, varying from lowest smoothness  $(\nu_1)$  and correlation range  $(1/a_1)$  in  $X_1$  (shown in Figure 1(b)), moderate in  $X_2$  (shown in Figure 1(c)) to the highest smoothness  $(\nu_3)$  and correlation range  $(1/a_3)$  in  $X_3$  (shown in Figure 1(d)). The interpretation of the coherence functions become clearer when we look at the filtered signal  $\check{\mathbf{X}}^{fb}$  of the simulated trivariate dataset  $\mathbf{X}$  at a frequency band fb. We apply a low-pass and a high-pass filter to obtain the filtered signals at low frequency (lf) and high frequency (hf)bands. In particular, we consider  $lf=0\leq\omega\leq1$  and  $hf=3.25\leq\omega\leq4.25$  to asses the signal behavior in low frequencies and high frequencies, respectively. Figures 2(a)-2(f) show the filtered signals for the chosen frequency bands. The empirical correlation between filtered signal pairs  $(\check{X_1}^{lf}, \check{X_2}^{lf})$ ,  $(\check{X_2}^{lf}, \check{X_3}^{lf})$  and  $(\check{X_1}^{lf}, \check{X_3}^{lf})$  are 0.46, 0.55 and 0.14, respectively, and for the pairs  $(\check{X_1}^{hf}, \check{X_2}^{hf})$ ,  $(\check{X_2}^{hf}, \check{X_3}^{hf})$  and  $(\check{X_1}^{hf}, \check{X_3}^{hf})$  the correlations are 0.54, 0.65 and 0.17, respectively. The empirical correlations mimic the underlying coherence function as the pair  $(\check{X}_2^{fb}, \check{X}_3^{fb})$  exhibits the highest correlation and the pair  $(\check{X}_1^{fb}, \check{X}_3^{fb})$  shows the weakest correlation, at both frequency bands  $fb = \{lf, hf\}$ . Moreover, similar to the under-

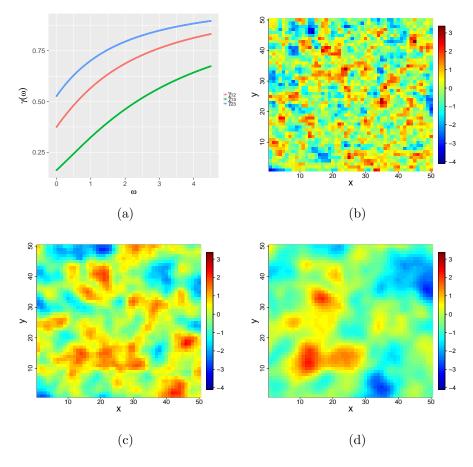


Figure 1: (a) Coherence functions for each pair of variables. (b) Simulated realization for  $X_1$  ( $\sigma_1 = 1, \nu_1 = 1, a_1 = 1$ ). (c) Simulated realization for  $X_2$  ( $\sigma_2 = 1, \nu_2 = 2, a_2 = 0.5$ ). (d) Simulated realization for  $X_3$  ( $\sigma_3 = 1, \nu_3 = 2.5, a_3 = 0.4$ ).

lying coherence function, all the pairwise correlations at hf are stronger than those at lf.

The sufficient conditions stated in Theorem 1 can be corroborated during model estimation by further parameterizing  $\boldsymbol{\beta}_k's$ , such that  $\{\boldsymbol{\beta}_k = \boldsymbol{\Lambda}_{\boldsymbol{\theta}_k}, \ k = -3, \dots, K\}$  where  $\boldsymbol{\Lambda}_{\boldsymbol{\theta}_k}'s$  essentially are the correlation matrices of size  $p \times p$  that allows for both the negative and nonnegative off-diagonal entries that can be derived from any valid correlation function that depends on the set of parameters  $\boldsymbol{\theta}_k$ . For example, let  $\boldsymbol{\theta}_k = \{t_{ij,k} \in \mathbb{R}, \ t_{ii,k} = 1, \ i = 1, \dots, p, \ 1 \leq i < j \leq p\}$ , then  $\boldsymbol{\Lambda}_{\boldsymbol{\theta}_k} = \{\frac{\sum_{l=j}^p t_{il,k} t_{jl,k}}{\sqrt{\sum_{u=i}^p (t_{iu,k})^2} \sqrt{\sum_{v=j}^p (t_{jv,k})^2}}\}_{i,j=1}^p$  is one valid and flexible parameterization that requires the total  $(K+4)\binom{p}{2}$  parameters to define  $\{\boldsymbol{\beta}_k, \ k = -3, \dots, K\}$ . Alternatively,

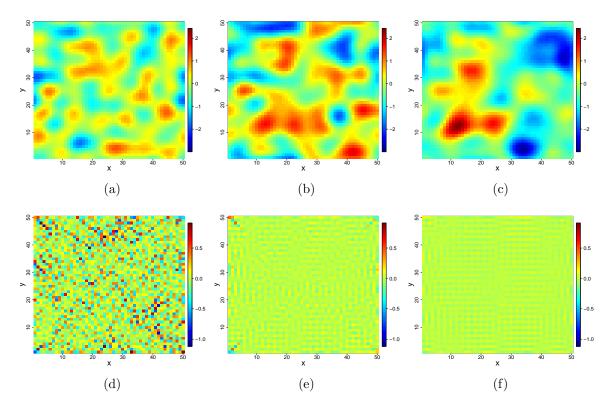


Figure 2: (a)  $\check{X}_1^{lf}$  ( $X_1$  filtered at  $0 \le \omega \le 1$ ). (b)  $\check{X}_2^{lf}$  ( $X_2$  filtered at  $0 \le \omega \le 1$ ). (c)  $\check{X}_3^{lf}$  ( $X_3$  filtered at  $0 \le \omega \le 1$ ). (d)  $\check{X}_1^{hf}$  ( $X_1$  filtered at  $3.25 \le \omega \le 4.25$ ). (e)  $\check{X}_2^{hf}$  ( $X_2$  filtered at  $3.25 \le \omega \le 4.25$ ). (f)  $\check{X}_3^{hf}$  ( $X_3$  filtered at  $3.25 \le \omega \le 4.25$ ).

we can consider a smaller set  $\boldsymbol{\theta}_k = \{t_{i,k} \in \mathbb{R}, i = 1, ..., p\}$  and define the parameterization as  $\boldsymbol{\Lambda}_{\boldsymbol{\theta}_k} = \{\exp(-|t_{i,k} - t_{j,k}|)\}_{i,j=1}^p$ , in which case the total number of parameters required to define  $\{\boldsymbol{\beta}_k, k = -3, ..., K\}$  is (K+4)p, which is much less than  $(K+4)\binom{p}{2}$ . However, this is a relatively less flexible parameterization as it will lead to only positive values of spline coefficients that will produce only positive coherence functions and positive cross-covariance functions, and therefore, should be considered only when the coherence functions are known to be positive for all frequencies. In the case p=2, a bivariate random field, the sufficient conditions are

$$-1 \le b_k^{(12)} \le 1, \ k = -3, \dots, K. \tag{8}$$

Thus, the B-spline coefficients should lie between -1 to 1 in a bivariate case to ensure that

the absolute coherence never exceeds unity at any frequency band.

The advantage of B-spline based specification (6) of the coherence functions is that our proposed model (7) approximately accommodates many existing classes of cross-covariance models that are constructed from the Matérn family, e.g., Multivariate Matérn, Separable models with Matérn components, etc. For a sufficiently large value of  $\omega_t$  and m, and appropriately specified B-splines, our proposed method can almost exactly reproduce those multivariate cross-covariances. For instance, the three examples of coherence functions shown in Figures 3(a), 3(b) and 3(c) are generated from our coherence model (6) for suitably selected spline coefficients  $S_{12}$ . They closely match with the coherence functions of the full bivariate Matérn model for three settings listed as Model 1-3 in Table 1. Figures 3(d), 3(e), and 3(f) show the computed cross-covariances from our model (7) corresponding to the coherence functions in Figures 3(a), 3(b) and 3(c) and the marginal parameter values of Model 1-3 from Table 1, respectively. The computed cross-covariances from our model are numerically equivalent to the corresponding full bivariate Matérn cross-covariances, thus exemplifying the generality of our proposed model. Furthermore, for a specific setting of parameters, the so-called parsimonious multivariate Matérn model is a special case in our proposed construction:

**Proposition 1.** For a common spatial scale parameter  $a_i = a, i = 1,...,p, K \to \infty$ ,  $\omega_t \to \infty$ , and common spline coefficients  $b_k^{(ij)} = \tau_{ij}, k = -3,...,K, 1 \le i \ne j \le p$  (or equivalently constant coherence function  $\gamma_{ij}(\omega) = \tau_{ij}, \forall \omega \ge 0$ ) satisfying the sufficient conditions of Theorem 1, the closed form solution of the integral (2) for the spectral densities in (4) and (5) exists, and is equal to the parsimonious multivariate Matérn model.

Various choices of spline coefficients and marginal parameters  $(\sigma_i, a_i, \nu_i, i = 1, ..., p)$  in our model (7) can imply the oscillation of coherence functions and cross-covariance functions

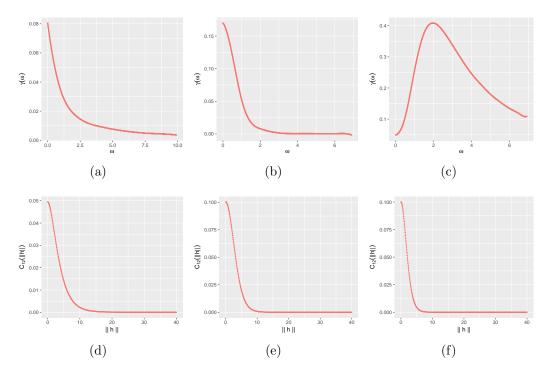


Figure 3: Example of bivariate coherence function for case-1 (a), case-2 (b) and case-3 (c). Corresponding cross-covariance function for case-1 (d), case-2 (e) and case-3 (f).

Table 1: Three parameter settings of full bivariate Matérn

Model settings	$\sigma_1$	$a_1$	$ u_1$	$\sigma_2$	$a_2$	$ u_2$	$a_{12}$	$ u_{12}$	$ ho_{12}$
Model 1	1	0.5	1	1	0.5	1	0.5	1.5	0.05
Model 2	1	1	2	1	1	3	1.1	5	0.1
Model 3	1	0.6	3	1	1.4	3	1.5	4	0.1

between negative and positive values. Figure 4 reflects one such example where we set the marginal parameters ( $\sigma_i = a_i = \nu_i = 1$ , i = 1, 2), threshold frequency  $\omega_t = 4.5$  and m = 990. We choose  $\Delta = 1$  (K = 4) and  $\mathbf{S}_{12} = \{-0.99, -0.99, 0.99, 0.99, 0.99, 0.99, -0.99, -0.99, -0.99\}$  to produce negative coherence at low frequencies and positive coherence at higher frequencies (shown in Figure 4(a)). The corresponding cross-covariance function from our model (7) (shown in Figure 4(b)) exhibits a transition from positive dependence to negative dependence with increasing distance, and eventually decays to zero at large distances. Figure 4(c) and 4(d) shows one realization of a zero mean bivariate Gaussian process  $\mathbf{Y}$  simulated with the

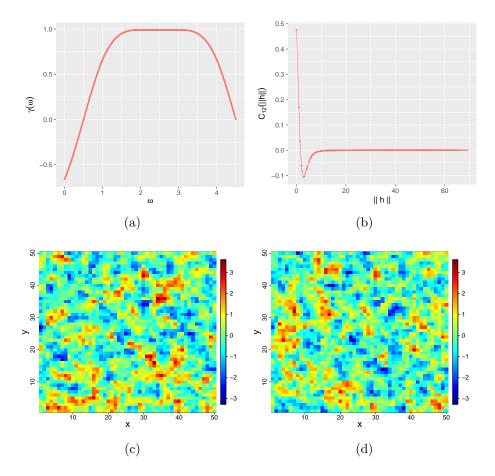


Figure 4: (a) Coherence function. (b) Cross-covariance function. (c) Simulated realization for  $Y_1$  ( $\sigma_1 = 1, \nu_1 = 1, a_1 = 1$ ). (d) Simulated realization for  $Y_2$  ( $\sigma_2 = 1, \nu_2 = 1, a_2 = 1$ ).

chosen marginal and cross-covariance function. The filtered signal  $\check{\mathbf{Y}}^{fb}$  for the simulated dataset  $\mathbf{Y}$  at the low-frequency band  $lf = 0 \le \omega \le 0.2$  and the high-frequency band  $hf = 3 \le \omega \le 4$  are shown in Figure 5. While the empirical correlation for the filtered signal pair  $(\check{Y}_1^{lf}, \check{Y}_2^{lf})$  is -0.5, i.e., negatively correlated, the empirical correlation for the pair  $(\check{Y}_1^{hf}, \check{Y}_2^{hf})$  is 0.94, i.e., positively correlated. This change of sign from negative to positive while going from lf to hf is to be expected due to the oscillatory nature of the underlying coherence function. Our proposed construction provides a potential working covariance model for real multivariate datasets, which exhibits such cross-process behavior.

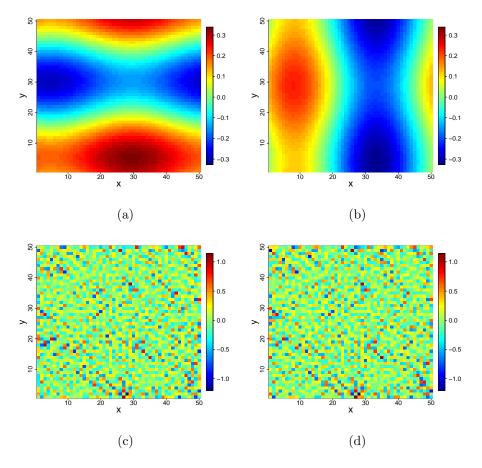


Figure 5: (a)  $\check{Y}_1^{lf}$  ( $Y_1$  filtered at  $0 \le \omega \le 0.2$ ). (b)  $\check{Y}_2^{lf}$  ( $Y_2$  filtered at  $0 \le \omega \le 0.2$ ). (c)  $\check{Y}_1^{hf}$  ( $Y_1$  filtered at  $0 \le \omega \le 0.2$ ). (d)  $\check{Y}_2^{hf}$  ( $Y_2$  filtered at  $0 \le \omega \le 0.2$ ).

### 2.3 Maximum Likelihood Estimation

Let  $\tilde{\mathbf{X}} = (\mathbf{X}(\mathbf{s}_1)^T, \dots, \mathbf{X}(\mathbf{s}_n)^T)^T$  be a realization from a zero mean stationary multivariate Gaussian process where  $\mathbf{X}(\mathbf{s}) = (X_1(\mathbf{s}), \dots, X_p(\mathbf{s}))^T$ . Let  $\Sigma_{\boldsymbol{\theta}_{SP}}$  denote the  $np \times np$  covariance matrix for  $\tilde{\mathbf{X}}$  where  $\{C_{ij}(\mathbf{s}_q - \mathbf{s}_r)\}_{i,j=1}^p \in \mathbb{R}^{p \times p}$  defined in (7) constitutes the  $(q,r)^{th}$ ,  $q,r = 1, \dots, n$  block entry of  $\Sigma_{\boldsymbol{\theta}_{SP}}$ , and  $\boldsymbol{\theta}_{SP}$  denote the set of parameters in our semiparametric model (7). Then  $\tilde{\mathbf{X}} \sim MVN_{np}(0, \Sigma_{\boldsymbol{\theta}_{SP}})$ , and the log-likelihood is given as:

$$\ell(\boldsymbol{\theta}_{\mathcal{SP}}|\tilde{\mathbf{X}}) = -\frac{1}{2}(\log \det \Sigma_{\boldsymbol{\theta}_{\mathcal{SP}}} + \tilde{\mathbf{X}}^{\mathrm{T}} \Sigma_{\boldsymbol{\theta}_{\mathcal{SP}}}^{-1} \tilde{\mathbf{X}} + np \log 2\pi)$$
(9)

For an appropriately chosen large value of  $\omega_t$  and m, and suitably specified uniform knot spacing  $\Delta$ , our semiparametric model (7) entirely depends on the set of parameters  $\boldsymbol{\theta}_{\mathcal{SP}}$ . Here the set  $\boldsymbol{\theta}_{\mathcal{SP}}$  consists of 3p marginal parameters  $(\sigma_i, \nu_i, a_i, i = 1, \dots, p)$  and  $(K+4)\binom{p}{2}$  spline coefficients  $\{b_k^{(ij)}, k = -3, -2, \dots, K, 1 \leq i < j \leq p.\}$ . In our implementation, we perform joint numerical maximization of the log-likelihood over the elements of the set  $\boldsymbol{\theta}_{\mathcal{SP}}$ , while ensuring the sufficient conditions of validity in Theorem 1 by further parameterizing the B-spline coefficients as discussed in Section 2.1. In the case of p=2, the estimation procedure is straightforward, as restricting the values of B-spline coefficients to lie between -1 to 1 would suffice for the validity, and therefore does not require tricky parameterizations.

# 3 Simulation Study

In this section, we explore the performance of our proposed semiparametric model (7) by evaluating the maximum likelihood estimates of its marginal parameters and the underlying coherence function for bivariate processes simulated from different multivariate models. In particular, we simulate the Gaussian random field from the full bivariate Matérn model (see Section 3.1) and the LMC with latent Matérn fields (see Section 3.2), and excercise our semiparametric model to estimate the marginal and cross-process behaviour from simulated datasets.

#### 3.1 Simulation 1: Full Bivariate Matérn Model

We consider a zero mean bivariate Gaussian random field  $\mathbf{X}(\mathbf{s}) = (X_1(\mathbf{s}), X_2(\mathbf{s}))^{\mathrm{T}}$  on a grid of coordinates  $\{(i, j)\}_{i,j=1}^{30}$ , with marginal and cross-covariances defined by the full bivariate Matérn model:

$$C_{ii}(\mathbf{h}) = M(\mathbf{h}|\sigma_i, \nu_i, a_i), i = 1, 2,$$

$$C_{ij}(\mathbf{h}) = \rho_{ij} M(\mathbf{h}|\sqrt{\sigma_i \sigma_j}, \nu_{ij}, a_{ij}), \ 1 \le i \ne j \le 2,$$

where  $\rho_{ij}$  refers to the co-located correlation coefficient that requires to satisfy the necessary and sufficient condition provided in Theorem 3 of Gneiting et al. (2010). The full bivariate Matérn model implies the following isotropic coherence function in a bivariate process defined over a spatial domain  $\mathcal{D} \in \mathbb{R}^d$  (d = 2 in our case):

$$\gamma_{12}(\omega) = \rho_{12} \frac{\Gamma(\nu_{12} + d/2)\Gamma(\nu_1)^{\frac{1}{2}}\Gamma(\nu_2)^{\frac{1}{2}}a_{12}^{2\nu_{12}}(a_1^2 + \omega^2)^{\frac{\nu_1}{2} + \frac{d}{4}}(a_2^2 + \omega^2)^{\frac{\nu_2}{2} + \frac{d}{4}}}{\Gamma(\nu_1 + d/2)^{\frac{1}{2}}\Gamma(\nu_2 + d/2)^{\frac{1}{2}}\Gamma(\nu_{12})a_1^{\nu_1}a_2^{\nu_2}(a_{12}^2 + \omega^2)^{\nu_{12} + \frac{d}{2}}}.$$

We simulate 50 realizations of **X**, for three cases of parameter settings listed as Model 1-3 in Table 2. An example of simulated bivariate processes from these models is shown in Figure 6. These three models simulate bivariate processes with contrasting coherence features, broadly covering all the shapes of a coherence function that a full bivariate Matérn model can generate. Whereas Model 1 and 2 lead to monotonically increasing and monotonically decreasing coherence functions, respectively, Model 3 leads to a bump in the coherence function at some frequency band.

We fit our semiparametric model (7) on the simulated realizations, using the method of maximum likelihood to investigate its efficiency. For estimation in each of the three cases of simulation, we specify the threshold frequency  $\omega_t = 4.5$ , and m = 380 for the discretization of the frequency interval  $[0, \omega_t]$ . Furthermore, we set  $\Delta = 1$  (or equivalently K = 4) to completely specify the coherence function, which in turn requires the estimation of eight B-spline coefficients  $\{\mathbf{S}_{12} = b_k^{(12)}, k = -3, -2, \dots, 4\}$ . We also assume that the marginal smoothness parameters  $\{\nu_i, i = 1, 2\}$  are known, and therefore are fixed to their true value in our model, to avoid possible identifiability issues (Zhang, 2004). Thus, in each of the three cases, we estimate 12 parameters in total, including the 4 marginal parameters  $\{a_i, \sigma_i^2, i = 1, 2\}$  and a set of 8 B-spline coefficients  $\mathbf{S}_{12} = \{b_k^{(12)}, k = -3, -2, \dots, 4\}$ .

Figure 7 shows a comparison of the true coherence function and the averaged estimated coherence function with 95% pointwise intervals for the three cases of monotonically in-

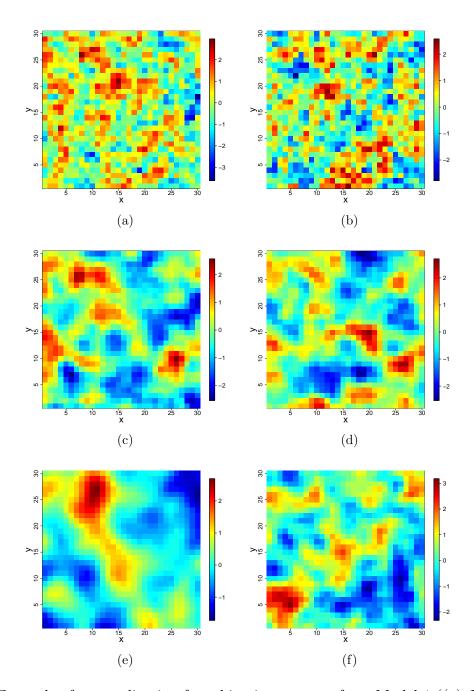


Figure 6: Example of one realization for a bivariate process from Model 1 ((a)  $X_1$ , (b)  $X_2$ ), Model 2 ((c)  $X_1$ , (d)  $X_2$ ) and Model 3 ((d)  $X_1$ , (e)  $X_2$ ).

creasing coherence (Figure 7(a)), monotonically decreasing coherence (Figure 7(b)) and the coherence function with a bump (Figure 7(c)). For all the three cases, the averaged estimated coherence function overlaps the true coherence function at almost all frequencies, thus indicating the efficiency of our model in adequately capturing the cross-spectral behaviour of

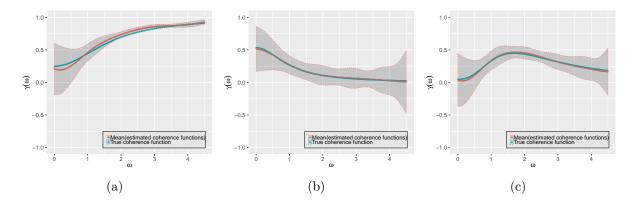


Figure 7: Comparison of the average estimate of the coherence function (95% pointwise intervals in grey) and the true coherence functions for the processes generated from Model 1 (a), Model 2 (b) and Model 3 (c).

the processes. Additionally, it also implies sufficiently reasonable fit of the cross-covariances, due to the complementary translation of coherence functions in the frequency domain to the cross-covariances in the space domain. Table 2 reports the average estimates of marginal parameters with their standard errors in parenthesis, to draw a comparison between the true parameters of the exact marginal Matérn and the estimated parameters from our model with approximately Matérn marginals. The remarkable closeness of the estimated spatial scales  $\{a_i, i = 1, 2\}$  and the variances  $\{\sigma_i^2, i = 1, 2\}$  of our model to the true parameter values demonstrates satisfactory marginal fits. Although our semiparametric model requires a slightly higher number of parameters as compared to the true full bivariate Matérn model, the validity conditions are much simpler to implement, and leads to a noticeably good fit for both the marginal and cross-process relationships.

### 3.2 Simulation 2: Linear Model of Coregionalization

In this section, we consider a zero mean bivariate Gaussian random field  $\mathbf{X}(\mathbf{s}) = (X_1(\mathbf{s}), X_2(\mathbf{s}))^T$  on 500 irregularly spaced locations in the domain  $[0, 40]^2$  with cross and marginal spatial

Table 2: Simulation summary for marginal parameter estimates. The true values under Model 1-3 corresponds to the parameter values for the full bivariate Matérn model chosen for simulations. The average estimate and standard error values under Model 1-3 corresponds the mean and standard error of the marginal parameter estimates from the semiparametric model over 50 runs. Note that average estimate and standard error entries for the last three columns are left blank since the cross-covariance part the semiparametric model is non-parametric and has been shown as comparison of coherence functions in Figure 7

Models	Parameters	$a_1$	$\sigma_1^2$	$ u_1$	$a_2$	$\sigma_2^2$	$ u_2$	$a_{12}$	$ u_{12}$	$ ho_{12}$
	True value	1	1	1	1	1	1	$\sqrt{2}$	1	0.5
Model 1	Average estimate	$\bar{1}.\bar{1}\bar{3}$	0.99		1.13	0.98	-			
	Standard error	$(\bar{0}.\bar{0}\bar{6})$	[0.08]		(0.06)	(0.08)	-	-	 -	1
	True value	1	1	3	1	1	3	1	4	0.4
Model 2	Average estimate	$\bar{1}.\bar{0}\bar{1}$	0.99		1.02	0.98	-	-		
	Standard error	$(\bar{0}.\bar{0}\bar{3})$	(0.12)	-	(0.03)	(0.12)	-	-		1
	True value	0.5	1	3	1	1	3	1.2	4	0.1
Model 3	Average estimate	$\bar{0.51}$	$\begin{bmatrix} - & \overline{1} & - & \overline{1} \end{bmatrix}$	_	1.01	0.99	-	-		
	Standard error	$(\bar{0}.\bar{0}\bar{2})$	(0.22)		(0.03)	(0.11)	-			

dependence described by the LMC:

$$\mathbf{X}(\mathbf{s}) = \begin{pmatrix} X_1(\mathbf{s}) \\ X_2(\mathbf{s}) \end{pmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{pmatrix} Z_1(\mathbf{s}) \\ Z_2(\mathbf{s}) \end{pmatrix} = \mathbf{BZ}(\mathbf{s}),$$

where **B** is the coregionalization matrix that supervises the magnitude of dependencies on the uncorrelated latent processes  $\mathbf{Z}(\mathbf{s})$ . We specify the independent processes  $Z_1(\mathbf{s})$  and  $Z_2(\mathbf{s})$  to marginally admit Matérn covariance functions  $\mathbf{M}(\mathbf{h}|\sigma_1, \nu_1, a_1)$  and  $\mathbf{M}(\mathbf{h}|\sigma_2, \nu_2, a_2)$ , respectively. The coherence function for the bivariate process  $\mathbf{X}(\mathbf{s})$  is then given as:

$$\gamma_{12}(\omega) = \frac{b_{11}b_{21}f_1(\omega) + b_{12}b_{22}f_2(\omega)}{\sqrt{b_{11}^2f_1(\omega) + b_{12}^2f_2(\omega)}\sqrt{b_{21}^2f_1(\omega) + b_{22}^2f_2(\omega)}},$$

where  $f_1(\omega)$  and  $f_2(\omega)$  are the Matérn spectral densities corresponding to  $\mathbf{M}(\mathbf{h}|\sigma_1, \nu_1, a_1)$  and  $\mathbf{M}(\mathbf{h}|\sigma_2, \nu_2, a_2)$ , respectively.

We consider the marginal Matérn parameters for  $\mathbf{Z}(\mathbf{s})$  to be  $(\sigma_1, \nu_1, a_1) = (1, 1, 0.5)$  and  $(\sigma_2, \nu_2, a_2) = (1, 2, 0.5)$ , and we set the entries of the coregionalization matrix  $\mathbf{B}$  as  $b_{11} = 1, b_{12} = 0.4, b_{21} = 0.9$  and  $b_{22} = 7.5$ . The coherence function for a bivariate process

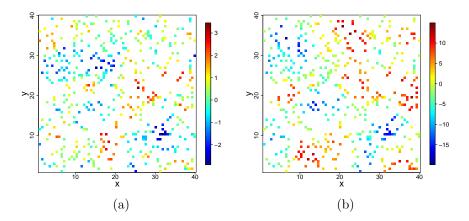


Figure 8: Example of one realization for a bivariate process from the specified LMC ((a) Variable 1, (b) Variable 2)).

with this choice of parameters shows a decreasing trend at lower frequencies, followed by an increasing trend at higher frequencies. We simulate 50 realizations of the specified bivariate process  $\mathbf{X}(\mathbf{s})$ , and fit our semiparametric model (7) using MLE, to model the coherence function as well as the marginal and cross-process dependence. An example realization for the simulated bivariate process from the specified LMC is shown in Figure 8.

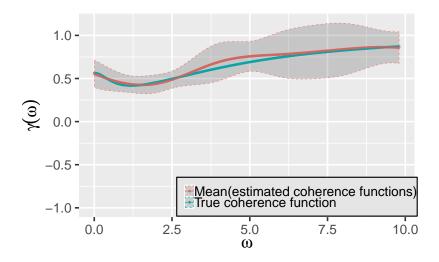


Figure 9: Comparison of the average estimate of the coherence functions (95% pointwise intervals in grey) and the true coherence function for the process generated from the specified LMC.

Prior to the estimation of our semiparametric model (7) from the simulated datasets,

we are required to specify the value of m and the threshold frequency  $\omega_t$ , which we set at 299 and 9.8, respectively. For the specification of our B-spline based coherence function, we set  $\Delta=2$  (or equivalently K=4), which in turn would require the estimation of 8 B-spline coefficients. Therefore, for this simulation study, we estimate a total of 14 parameters, that include 6 marginal Matérn parameters  $(\sigma_i, \nu_i, a_i, i=1, 2)$  and 8 B-spline coefficients  $\mathbf{S}_{12}=\{b_k^{(12)}, k=-3, -2, \ldots, 4\}.$ 

Table 3: Average estimates and standard error of marginal parameters from the semiparametric model

Model Parameters	$\sigma_1^2$	$a_1$	$ u_1$	$\sigma_2^2$	$a_2$	$ u_2$
Average estimates	1.21	0.45	0.96	61.63	0.46	1.84
Standard error	(0.19)	(0.076)	(0.11)	(1.58)	(0.04)	(0.10)

The averaged estimated coherence functions with 95% pointwise interval and the true underlying coherence function shown in Figure 9 display conspicuous comparability. Our semiparametric model efficiently recovers the true shape of the underlying coherence function, which, moreover, signals toward decent fit of the cross-covariance function. Table 3 reports the estimates and standard errors of marginal parameters from our semiparametric model. Note that the estimates reported in Table 3 correspond to the marginal parameter estimates of our semiparametric model that describes the marginal spatial dependences of the process  $\mathbf{X}(\mathbf{s})$ , and therefore its direct comparison with the true Matérn parameters  $(\sigma_i, \nu_i, a_i, i = 1, 2)$  of  $\mathbf{Z}(\mathbf{s})$  is not straightforward. However, the true marginal variances for the processes  $X_1(\mathbf{s})$  and  $X_2(\mathbf{s})$  are  $b_{11}^2 \mathbf{M}(\|0\||1, 1, 0.5) + b_{12}^2 \mathbf{M}(\|0\||1, 2, 0.5) = 1.16$  and  $b_{21}^2 \mathbf{M}(\|0\||1, 1, 0.5) + b_{22}^2 \mathbf{M}(\|0\||1, 2, 0.5) = 57.06$ , respectively, and are comparable with the estimated marginal variances of our semiparametric model reported in Table 3.

# 4 Applications to PM<sub>2.5</sub> and Wind Speed Data

We now illustrate the flexibility of our proposed semiparametric approach by applying our method to an atmospheric dataset consisting of a bivariate spatial field of particulate matter concentrations ( $PM_{2.5}$ ) and wind speed.  $PM_{2.5}$  is one of the principle indicators of air pollution level and represents the concentration of fine particulate matter with diameter less than  $2.5\mu m$  suspended in the atmosphere. Its major constituent components include nitrate, sulfate, organic carbon and elemental carbon, which in high concentrations, have hazardous effects on human health (Dominici et al., 2006; Pope III and Dockery, 2006; Samoli et al., 2008; Chang et al., 2011). While various meteorological variables such as regional stagnation, humidity, precipitation, etc., impact the concentration of  $PM_{2.5}$  in polluted regions, here we focus on  $PM_{2.5}$ 's association with wind speed, which generally tends to be negatively correlated in nature (Jacob and Winner, 2009). We explore the marginal and cross-spatial dependence of  $PM_{2.5}$  and wind speed by fitting various multivariate spatial models. Moreover, we perform spatial prediction to draw a comparison between the performance of our semiparametric model and other traditionally used multivariate models such as full bivariate Matérn and the LMC.

We study the dynamics of PM<sub>2.5</sub> and wind speed over the North-Eastern climatic region of the United States which comprises 11 states, namely, Maine, New Hampshire, Vermont, New York, Massachusetts, Connecticut, Rhode Island, Pennsylvania, New jersey, Delaware and Maryland. The data for PM<sub>2.5</sub> is sourced from the Environmental Protection Agency (EPA) which provides the daily average values that are generated via Community Multiscale Air Quality Modeling System (CMAQ, https://www.epa.gov/cmaq). The wind speed data is obtained from North American Regional Reanalysis (NARR, https://www.esrl.noaa.gov/psd) which provides the monthly mean values of various meteorological variables. The

raw datasets for our two variables differ in their spatial and temporal resolution, which we adjust by averaging the  $PM_{2.5}$  data. We average the daily  $PM_{2.5}$  values over each month to comply with monthly mean wind speed data, and in addition we spatially average the monthly mean  $PM_{2.5}$  data over the vicinity of 481 wind speed data locations to prepare a colocated bivariate  $PM_{2.5}$ /wind speed dataset.

For our application, we consider the bivariate PM<sub>2.5</sub>/wind speed data for the month of January 2013 (shown in Figure 10). Whereas the wind speed exhibits approximately Gaussian distribution, the distribution of PM<sub>2.5</sub> shows positive skewness, which prompts us to log transform PM<sub>2.5</sub> to more closely satisfy the assumption of a bivariate Gaussian random field. Here, we primarily focus on modeling the second-order dependence structure of the log (PM<sub>2.5</sub>) and wind speed; therefore, we detach the mean component by subtracting their respective empirical marginal means. Furthermore, we compute the empirical marginal variances and exercise componentwise standardization to bring (1) uniformity in the order of magnitude of process components and (2) numerical stability. Now, let us assume  $\mathbf{X}(\mathbf{s}) = (X_{PM_{2.5}}(\mathbf{s}), X_{WS}(\mathbf{s}))^{\mathrm{T}}$  to be a bivariate Gaussian random field, where components  $X_{PM_{2.5}}$  and  $X_{WS}$  represent the standardized log (PM<sub>2.5</sub>) and wind speed, respectively. Then, for the set of 481 observed locations  $\{\mathbf{s}_1, \ldots, \mathbf{s}_{481}\}$  (Shown in Figure 10),  $\mathbf{X} \sim MVN_{982}(0, \Sigma_{982 \times 982})$ , where  $\Sigma_{982 \times 982}$  is the covariance matrix and our primary object of interest that we model using various bivariate spatial models.

Prior to modeling the covariance matrix  $\Sigma_{982\times982}$ , we divide our data into a training set of 381 randomly selected locations and a validation set of the remaining 100 locations. We then proceed to fit various bivariate covariance models, augmented with nugget effects to capture the measurement errors, on 381 training locations, using the method of maximum likelihood. In particular, we consider six candidate models; an independent Matérn model that serves as our baseline performance standard due to its complete incomprehension of the

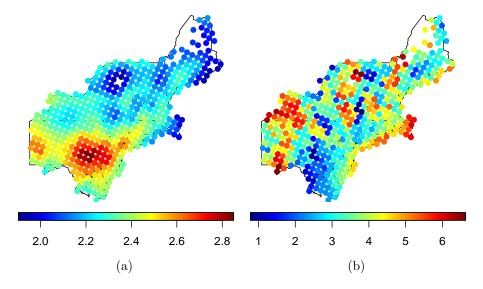


Figure 10: (a)  $\log (PM_{2.5})$  data over the North-Eastern climatic region of the United States. (b) Wind speed data over the North-Eastern climatic region of the United States.

cross-covariances between  $X_{PM_{2.5}}$  and  $X_{WS}$ ; the commonly used full bivariate Matérn model; full LMC with two latent Matérn fields; and our proposed semiparametric model with three different choices of uniform knot spacing  $\Delta$ .

For our semiparametric model, we specify the threshold frequency  $\omega_t = 9$ , and set m = 499 for the discretization of the frequency interval [0, 9]. We consider three values of the uniform knot spacing  $\Delta \in (2, 4, 5)$ , which allows for varying degrees of flexibility in the underlying coherence function of the semiparametric model. The model with  $\Delta = 2$  enjoys the most flexible underlying coherence function relative to the models with  $\Delta = 4$  and  $\Delta = 5$ , having a slightly tighter construct for the shape of the underlying coherence functions. The semiparametric models with  $\Delta \in (2, 4, 5)$  require the estimation of 8,6 and 5 B-spline coefficients, respectively, in addition to 6 marginal parameters and 2 parameters representing the nugget effect of each process component.

Table 4 reports the maximized log-likelihood values and the Akaike information criterion (AIC) values along with the number of parameters for the six candidate models. Strikingly,

Table 4: Model fit summary for different candidate models. The highest log-likelihood value (shown as bold) is achieved by the semiparametric ( $\Delta = 2$ )+Nugget model and the lowest AIC (shown as bold) is achieved by the semiparametric ( $\Delta = 4$ )+Nugget model

Candidate Models	No. of parameters	Log-likelihood	AIC
Independent Matérn + Nugget	8	-331.179	678.357
Full bivariate Matérn + Nugget	11	-331.429	684.857
LMC + Nugget	12	-312.226	648.452
Semiparametric ( $\Delta = 2$ ) + Nugget	16	-307.989	647.977
Semiparametric ( $\Delta = 4$ ) + Nugget	14	-308.092	644.184
Semiparametric ( $\Delta = 5$ ) + Nugget	13	-309.123	644.246

Table 4 points out the comparable performance of the full bivariate Matérn model and the independent Matérn model in terms of maximized log-likelihood, and, in fact, identifies the full bivariate Matérn as the most inferior model in terms of the AIC values. While this result seems unrealistic and misleading at first glance due to the theoretically desired properties that the full bivariate Matérn model enjoys, it actually indicate towards the problems associated with its inefficient parameter estimation. We use the function RFfit from the R-package RANDOMFIELDS (Schlather et al., 2015) to fit the full bivariate Matérn model, which in our case provides reasonably good estimates for the marginal parameters, but gives a noticeably substandard estimate for cross-covariance parameters. The estimated co-located correlation coefficient  $\widehat{\rho_{12}} = -6.70 \times 10^{-09}$  is numerically equivalent to 0, and is indeed far from its empirical value of -0.39. The estimate  $\widehat{\rho_{12}} = -6.70 \times 10^{-09}$  reduces the full bivariate Matérn model to almost independent Matérn model, thus, producing similar loglikelihood values, but a higher AIC value due to its 3 additional cross-covariance parameters. We observe a significant improvement in the log-likelihood value and the AIC value for the full LMC model as compared to the baseline independent Matérn case, which is not surprising because the full LMC takes into account the cross-process spatial dependence between  $X_{PM_{2.5}}$ and  $X_{WS}$ , unlike the independent Matérn model. Our semiparametric model in all three cases of  $\Delta \in \{2, 4, 5\}$  outperforms all other candidate models as it achieves the highest loglikelihood values and the lowest AIC values, which is to be expected because of the flexible specification of underlying coherence function. Even the most restricted semiparametric model corresponding to  $\Delta=5$  demonstrates a superior fit than all the other candidate models.

Figure 11 reveals the estimated coherence functions from all the candidate models. The independent Matérn model exhibits zero coherence at all frequency bands, which is obvious due to its assumed independence between  $X_{PM_{2.5}}$  and  $X_{WS}$ . The co-located correlation coefficient  $\rho_{12}$  in the full bivariate Matérn model acts as the scaling parameter for its coherence function, which being estimated close to zero, puts the coherence practically at 0 for all the frequency bands. The estimated coherence function for the full LMC model acquires a shape similar to the one we studied in Section 3.2, but lies in the negative axis and puts the lowest coherence (highest in magnitude) at  $\omega_t \approx 0.98$ . The most restricted semiparametric model with  $\Delta = 5$  shares the common shape with the LMC; however, it puts the lowest coherence at  $\omega_t \approx 3.95$ . The other two relatively flexible semiparametric models with  $\Delta = 2$  and  $\Delta = 4$  exhibit slightly oscillating coherence functions, and are even favoured by the log-likelihood and AIC values to represent the best fit for the true underlying coherence that cannot be captured by any existing multivariate models.

Table 5: Prediction scores for different candidate models. The semiparametric ( $\Delta = 5$ )+Nugget model shows best prediction performance in terms of RMSE, NMSE and mCRPS (shown as bold) and the semiparametric ( $\Delta = 4$ )+Nugget model shows best prediction performance in terms of MAE and mLogS (shown as bold)

Model	RMSPE	MAE	NMSE	mCRPS	mLogS
Independent Matérn + Nugget	0.533	0.333	0.746	0.242	0.232
Full bivariate Matérn + Nugget	0.534	0.333	0.745	0.243	0.236
LMC + Nugget	0.522	0.329	0.757	0.238	0.220
Semiparametric ( $\Delta = 2$ ) + Nugget	0.520	0.327	0.758	0.237	0.220
Semiparametric ( $\Delta = 4$ ) + Nugget	0.519	0.327	0.760	0.236	0.218
Semiparametric ( $\Delta = 5$ ) + Nugget	0.518	0.327	0.760	0.236	0.221

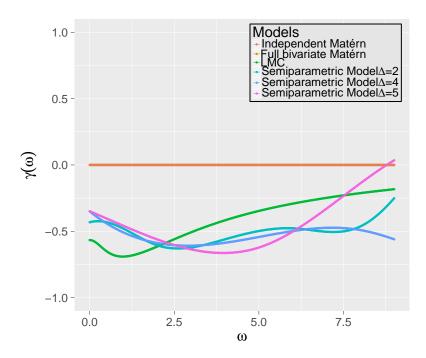


Figure 11: Coherence between  $\log (PM_{2.5})$  and wind speed estimated using different candidate models.

Here, we perform spatial predictions over the 100 left out validation locations for both the  $X_{PM_{2.5}}$  and  $X_{WS}$  to achieve a cross validation analysis for all the candidate models. In Table 5, we list some frequently used prediction scores combined for both the  $X_{PM_{2.5}}$  and  $X_{WS}$ , computed over 100 validation locations. The smaller values of the root mean squared prediction error (RMSPE), mean absolute error (MAE), mean continuous ranked probability score (mCRPS) and the mean logarithmic score (mLogS) (Gneiting and Raftery, 2007) are suggestive of better predictions, whereas the normalised-mean-squared error (NMSE) indicates a better prediction for the value closer to unity. Here, the computed prediction scores identify the independent Matérn model and the full bivariate Matérn model as the worst among the candidate models. While this is expected for the independent Matérn model because the spatial predictions with the independent Matérn model correspond to the independent univariate kriging, which is generally inferior to the co-kriging, the poor performance of the full bivariate Matérn is due to its poor model estimation, and not because

of its inflexibility. The LMC shows improvement in spatial prediction over the independent Matérn and full bivariate Matérn model, which is obvious as it utilizes correlations across the process components, however, due to its inflexible cross-covariance specification, its performance is not the best. Our proposed semiparametric models outperformed all the other candidate models in terms of spatial prediction, over nearly all cross-validation diagnostics combined for  $X_{PM_{2.5}}$  and  $X_{WS}$ , which empirically substantiate the importance of flexibly modeling coherence functions for spatial predictions.

### 5 Discussion

In this article, we introduced a semiparametric multivariate spatial covariance function via its spectral representation, that can flexibly model the coherence functions between the pair of components of a multivariate process. The B-spline based specification of the coherence function allows for more data-driven estimation of cross-covariances, relative to the available parametric models. We have presented simulation studies to demonstrate the performance of our proposed model through efficient maximum likelihood estimation of the multivariate spatial dependence, especially the underlying coherence function. The application of the proposed semiparametric model has been illustrated on a bivariate atmospheric dataset of particulate matter concentrations (PM<sub>2.5</sub>) and wind speed over the North-Eastern region of the United States. We have shown that our semiparametric model outperformed the conventionally used full bivariate Matérn model and the LMC, by producing lower AIC values and prediction scores.

The choice of uniform knot spacing ( $\Delta$ ) is crucial, as it governs the possible shapes that the coherence function can achieve. While we tried a number of different adhoc values for  $\Delta$  in our application section to choose the best model fit, the careful examination of the empirical coherence function can guide for the choice of  $\Delta$  in case of complete data on a regularly spaced grid of location. However, when the spatial data is not located on grid points, we suggest to try different sensible values of  $\Delta$  that maintain the trade-off between flexibility of coherence and the computational feasibility, and choose the best value based on cross-validation scores or some model selection criterion such as AIC.

In our proposed framework, we specified Matérn marginal, which makes our approach directly comparable with the full bivariate Matérn and the parsimonious multivariate Matérn models. However, any other choice of parametric or nonparametric spectral densities can be plugged in straightforwardly to specify marginal spatial dependence, and that would still lead to a valid multivariate model with exactly the same validity conditions provided in Theorem 1, thus leaving the door open for any future improvements.

Our model specifies the spectral densities and coherence functions only up to a threshold frequency  $\omega_t$ ; therefore, extending the proposed model to characterize spectral features for all frequencies  $\omega \geq 0$  is one potential direction for future research. This can be done by following the approach of Im et al. (2007) to add a parametric tail part in the coherence function, which would further finding validity conditions on the tail part.

### References

Apanasovich, T. V. and Genton, M. G. (2010). Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika*, 97:15–30.

Apanasovich, T. V., Genton, M. G., and Sun, Y. (2012). A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. *Journal of the American Statistical Association*, 107:180–193.

Bhat, K., Haran, M., and Goes, M. (2010). Computer model calibration with multivariate spatial output: A case study. In Chen, M.-H., Müller, P., Sun, D., Ye, K., and Dey, D. K.,

- editors, Frontiers of statistical decision making and Bayesian analysis, pages 168–184. Springer, New York.
- Chang, H. H., Reich, B. J., and Miranda, M. L. (2011). Time-to-event analysis of fine particle air pollution and preterm birth: Results from north carolina, 2001–2005. *American Journal of Epidemiology*, 175(2):91–98.
- Cramér, H. (1940). On the theory of stationary random processes. *Annals of Mathematics*, 41:215–230.
- De Boor, C. (2001). A practical guide to splines. New York: Springer.
- Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L., and Samet, J. M. (2006). Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA*, 295:1127–1134.
- Gaspari, G. and Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125:723–757.
- Gaspari, G., Cohn, S. E., Guo, J., and Pawson, S. (2006). Construction and application of covariance functions with variable length-fields. Quarterly Journal of the Royal Meteorological Society, 132:1815–1838.
- Genton, M. G. and Gorsich, D. J. (2002). Nonparametric variogram and covariogram estimation with fourier-bessel matrices. *Computational Statistics & Data Analysis*, 41:47 57. Special issue on Matrix Computations and Statistics.
- Genton, M. G. and Kleiber, W. (2015). Cross-covariance functions for multivariate geostatistics. *Statistical Science*, 30:147–163.

- Gneiting, T., Kleiber, W., and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105:1167–1177.
- Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102:359–378.
- Gorsich, D. J. and Genton, M. G. (2004). On the discretization of nonparametric isotropic covariogram estimators. *Statistics and Computing*, 14:99–108.
- Goulard, M. and Voltz, M. (1992). Linear coregionalization model: Tools for estimation and choice of cross-variogram matrix. *Mathematical Geology*, 24:269–286.
- Greasby, T. A. and Sain, S. R. (2011). Multivariate spatial analysis of climate change projections. *Journal of agricultural, biological, and environmental statistics*, 16:571–585.
- Guttorp, P. and Gneiting, T. (2006). Studies in the history of probability and statistics XLIX: On the Matérn correlation family. *Biometrika*, 93:989–995.
- Helterbrand, J. D. and Cressie, N. (1994). Universal cokriging under intrinsic coregionalization. *Mathematical Geology*, 26:205–226.
- Horn, R. A. and Johnson, C. R. (2013). *Matrix analysis*. Cambridge University Press, Cambridge, 2nd edition.
- Im, H. K., Stein, M. L., and Zhu, Z. (2006). Semiparametric estimation of spectral densities with scattered data. Technical report, University of Chicago, Center for Integrating Statistical and Environmental Sciences.
- Im, H. K., Stein, M. L., and Zhu, Z. (2007). Semiparametric estimation of spectral density with irregular observations. *Journal of the American Statistical Association*, 102:726–735.

- Jacob, D. J. and Winner, D. A. (2009). Effect of climate change on air quality. *Atmospheric Environment*, 43:51 63.
- Kleiber, W. (2017). Coherence for multivariate random fields. *Statistica Sinica*, 27:1675–1697.
- Majumdar, A. and Gelfand, A. E. (2007). Multivariate spatial modeling for geostatistical data using convolved covariance functions. *Mathematical Geology*, 39:225–245.
- Mardia, K. V. and Goodall, C. R. (1993). Spatial-temporal analysis of multivariate environmental monitoring data. In *Multivariate Environmental Statistics. North-Holland Series* in *Statistics and Probability*, volume 6, pages 347–386. North-Holland, Amsterdam.
- Matérn, B. (1986). Spatial Variation. Berlin:Springer-Verlag, 2nd edition.
- Pope III, C. A. and Dockery, D. W. (2006). Health effects of fine particulate air pollution: Lines that connect. *Journal of the Air & Waste Management Association*, 56:709–742.
- Sain, S. R., Furrer, R., and Cressie, N. (2011). A spatial analysis of multivariate output from regional climate models. *The Annals of Applied Statistics*, 5:150–175.
- Samoli, E., Peng, R., Ramsay, T., Pipikou, M., Touloumi, G., Dominici, F., Burnett, R., Cohen, A., Krewski, D., Samet, J., and Katsouyanni, K. (2008). Acute effects of ambient particulate matter on mortality in Europe and North America: Results from the APHENA study. *Environmental health perspectives*, 116:1480–1486.
- Schlather, M., Malinowski, A., Menck, P. J., Oesting, M., and Strokorb, K. (2015). Analysis, simulation and prediction of multivariate random fields with package randomfields. *Journal of Statistical Software*, 63:1–25.

- Schmidt, A. M. and Gelfand, A. E. (2003). A bayesian coregionalization approach for multivariate pollutant data. *Journal of Geophysical Research: Atmospheres*, 108.
- Shapiro, A. and Botha, J. (1991). Variogram fitting with a general class of conditionally nonnegative definite functions. *Computational Statistics & Data Analysis*, 11:87 96.
- Stein, M. L. (1999). Interpolation of spatial data: Some theory for kriging. Springer-Verlag New York.
- Ver Hoef, J. M. and Barry, R. P. (1998). Constructing and fitting models for cokriging and multivariable spatial prediction. *Journal of Statistical Planning and Inference*, 69:275–294.
- Ver Hoef, J. M., Cressie, N., and Barry, R. P. (2004). Flexible spatial models for kriging and cokriging using moving averages and the fast Fourier transform (FFT). *Journal of Computational and Graphical Statistics*, 13:265–282.
- Wackernagel, H. (2003). Multivariate geostatistics: An Introduction with Applications.

  Berlin: Springer, 3rd edition.
- Watson, G. N. (1944). A treatise on the theory of Bessel functions. Cambridge university press, 2nd edition.
- Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99:250–261.
- Zhang, H. (2007). Maximum-likelihood estimation for multivariate spatial linear coregionalization models. *Environmetrics*, 18:125–139.

# Appendix

### A Proof of Theorem 1

The spectral matrix for the spectral densities in (4) and (5) is given as:

$$\mathbf{f}(\omega) = \begin{bmatrix} f_{11}(\omega) & \dots & f_{1p}(\omega) \\ \vdots & \ddots & \vdots \\ f_{p1}(\omega) & \dots & f_{pp}(\omega) \end{bmatrix}, \omega \leq \omega_t$$

$$= \begin{bmatrix} \sqrt{f_{11}(\omega)} & & & \\ & \ddots & & \vdots \\ \sqrt{f_{pp}(\omega)} \end{bmatrix} \begin{bmatrix} 1 & \dots & \gamma_{1p}(\omega) \\ \vdots & \ddots & \vdots \\ \gamma_{p1}(\omega) & \dots & 1 \end{bmatrix} \begin{bmatrix} \sqrt{f_{11}(\omega)} & & \\ & \ddots & \\ & & \sqrt{f_{pp}(\omega)} \end{bmatrix}$$

$$= \operatorname{Diag}(\sqrt{f_{ii}(\omega)})_{i=1}^{p} \begin{bmatrix} 1 & \dots & \gamma_{1p}(\omega) \\ \vdots & \ddots & \vdots \\ \gamma_{p1}(\omega) & \dots & 1 \end{bmatrix} \operatorname{Diag}(\sqrt{f_{ii}(\omega)})_{i=1}^{p}$$

The spectral matrix  $\mathbf{f}(\omega)$  is then nonnegative definite if the matrix

$$\Gamma(\omega) = \begin{bmatrix} 1 & \dots & \gamma_{1p}(\omega) \\ \vdots & \ddots & \vdots \\ \gamma_{p1}(\omega) & \dots & 1 \end{bmatrix}$$

is nonnegative definite (:: if a nonnegative definite matrix  $\mathbf{M}$  is pre and post-multiplied by a full rank square matrix  $\mathbf{N}$  and its transpose  $\mathbf{N}^{\mathrm{T}}$ , the resulting matrix  $\mathbf{N}\mathbf{M}\mathbf{N}^{\mathrm{T}}$  is nonnegative definite (Horn and Johnson, 2013, Observation 7.1.8, p. 431)).

$$\Gamma(\omega) = \begin{bmatrix} 1 & \dots & \sum_{k=-3}^{K} b_k^{(1p)} B_k(\omega) \\ \vdots & \ddots & \vdots \\ \sum_{k=-3}^{K} b_k^{(p1)} B_k(\omega) & \dots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=-3}^{K} B_k(\omega) & \dots & \sum_{k=-3}^{K} b_k^{(1p)} B_k(\omega) \\ \vdots & \ddots & \vdots \\ \sum_{k=-3}^{K} b_k^{(p1)} B_k(\omega) & \dots & \sum_{k=-3}^{K} B_k(\omega) \end{bmatrix} (\because \sum_{k=-3}^{K} B_k(\omega) = 1, \ \forall \omega \in [0, (K+1)\Delta)$$

$$= \sum_{k=-3}^{K} B_k(\omega) \beta_k$$

where  $\beta_k = \{b_k^{(ij)}\}_{i,j=1}^p$  are the  $p \times p$  symmetric matrices with diagonal elements  $\{b_k^{(ii)} = 1\}$   $\forall i = 1, 2, \ldots, p, \quad k = -3, -2, \ldots, K\}$ . The quantity  $\sum_{k=-3}^K B_k(\omega)$  is nonnegative  $\forall \omega \leq \omega_t$ . Therefore the matrix  $\Gamma(\omega)$  is nonnegative definite  $\forall \omega \leq \omega_t$  if the matrices  $\{\beta_k, k = -3, \ldots, K\}$  are nonnegative definite (: the linear combination of nonnegative definite matrices with nonnegative coefficients is a nonnegative definite matrix (Horn and Johnson, 2013, Observation 7.1.3, p. 430)). Consequently, following the Cramér's Theorem in its spectral density version, the matrix-valued covariance function  $\mathbf{C}(\mathbf{h}) = \{\mathbf{C}_{ij}(\mathbf{h})\}_{i,j=1}^p$  in (7) is valid if the matrices  $\{\beta_k, k = -3, \ldots, K\}$  are non-negative definite.

# B Proof for Proposition 1

For  $\omega_t \to \infty$  and common spatial scale parameters  $a_i = a > 0$ , i = 1, ..., p, the marginal spectral densities in (4) becomes the untruncated Matérn spectral densities:

$$f_{ii}(\omega|\sigma_i,\nu_i,a) = \sigma_i^2 \frac{\Gamma(\nu_i + d/2)a_i^{2\nu_i}}{\Gamma(\nu_i)\pi^{d/2}(a_i^2 + \omega^2)^{\nu_i + d/2}}, \ \omega \ge 0, \ \sigma_i,\nu_i,a_i > 0, \ i = 1,\dots,p$$

and the corresponding marginal covariance functions are of the Matérn type with common spatial scales a, distinct smoothness  $\nu_i$ ,  $i=1,\ldots,p$  and distinct variances  $\sigma_i^2$ ,  $i=1,\ldots,p$ :

$$C_{ii}(\mathbf{h}) = \int_0^\infty \|\mathbf{h}\| \left(\frac{2\pi\omega}{\|\mathbf{h}\|}\right)^{\kappa+1} J_{\kappa}(\omega\|\mathbf{h}\|) f_{ii}(\omega|\sigma_i,\nu_i,a) d\omega = M(\mathbf{h}|\sigma_i,\nu_i,a), \ i = 1,\ldots,p.$$

For  $K \to \infty$  and common B-spline coefficients  $b_k^{(ij)} = \tau_{ij}, \ k = -3, \dots, K, \ 1 \le i \ne j \le p$ , the coherence function for the  $(i,j)^{th}$  pair of components is given as:

$$\gamma_{ij}(\omega) = \tau_{ij} \sum_{k=-3}^{\infty} B_k(\omega) = \tau_{ij}, \ \omega \ge 0, \ 1 \le i \ne j \le p.$$

The cross spectral densities in (5) then becomes:

$$f_{ij}(\omega|f_{ii}, f_{jj}, \mathbf{S}_{ij}, K) = \tau_{ij}\mathcal{C}(\nu_i, \nu_j, d)\sigma_i\sigma_j \frac{\Gamma((\nu_i + \nu_j)/2 + d/2)a^{(\nu_i + \nu_j)}}{\Gamma((\nu_i + \nu_j)/2)\pi^{d/2}(a^2 + \omega^2)^{(\nu_i + \nu_j)/2 + d/2}}, \omega \ge 0, \ 1 \le i \ne j \le p,$$

where

$$C(\nu_i, \nu_j, d) = \frac{\Gamma(\nu_i + d/2)^{\frac{1}{2}} \Gamma(\nu_j + d/2)^{\frac{1}{2}} \Gamma((\nu_i + \nu_j)/2)}{\Gamma(\nu_i)^{\frac{1}{2}} \Gamma(\nu_i)^{\frac{1}{2}} \Gamma((\nu_i + \nu_j)/2 + d/2)}.$$

The corresponding cross-covariances is then given as;

$$C_{ij}(\mathbf{h}) = \int_0^\infty \|\mathbf{h}\| \left(\frac{2\pi\omega}{\|\mathbf{h}\|}\right)^{\kappa+1} J_{\kappa}(\omega \|\mathbf{h}\|) f_{ij}(\omega | f_{ii}, f_{jj}, \mathbf{S}_{ij}, K) d\omega$$

=

$$= M(\mathbf{h}|\sqrt{\tau_{ij}\mathcal{C}(\nu_i,\nu_j,d)\sigma_i\sigma_j}, (\nu_i + \nu_j)/2, a), \ 1 \le i \ne j \le p.$$

which is a parsimonious multivariate Matérn cross-covariance function with the colocated correlation coefficient  $\rho_{ij} = \tau_{ij} \mathcal{C}(\nu_i, \nu_j, d)$ .