

# Asymptotically equivalent prediction in multivariate geostatistics

FRANÇOIS BACHOC<sup>1,a</sup>, EMILIO PORCU<sup>2,b</sup>, MORENO BEVILACQUA<sup>3,c</sup>,  
REINHARD FURRER<sup>4,d</sup> and TARIK FAOUZI<sup>5,e</sup>

<sup>1</sup>*Institut de mathématique, UMR5219; Université de Toulouse; CNRS, UPS IMT, F-31062 Toulouse Cedex 9, France.* <sup>a</sup>[francois.bachoc@math.univ-toulouse.fr](mailto:francois.bachoc@math.univ-toulouse.fr)

<sup>2</sup>*Department of Mathematics, Khalifa University, Abu Dhabi, The United Arab Emirates; School of Computer Science and Statistics, Trinity College, Dublin.* <sup>b</sup>[porcue@tcd.ie](mailto:porcue@tcd.ie)

<sup>3</sup>*Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Viña del Mar, Chile; ANID-Millennium Science Initiative Program-Millennium Nucleus Center for the Discovery of Structures in Complex Data.*

<sup>c</sup>[moreno.bevilacqua@uai.cl](mailto:moreno.bevilacqua@uai.cl)

<sup>4</sup>*Department of Mathematics and Department of Computational Science, University of Zurich, 8057, Zurich, Switzerland.* <sup>d</sup>[reinhard.furrer@math.uzh.ch](mailto:reinhard.furrer@math.uzh.ch)

<sup>5</sup>*Department of Statistics, Universidad del BíoBío, 4081112, Concepcion, Chile.* <sup>e</sup>[tfaouzi@ubiobio.cl](mailto:tfaouzi@ubiobio.cl)

Cokriging is the common method of spatial interpolation (best linear unbiased prediction) in multivariate geostatistics. While best linear prediction has been well understood in univariate spatial statistics, the literature for the multivariate case has been elusive so far. The new challenges provided by modern spatial datasets, being typically multivariate, call for a deeper study of cokriging. In particular, we deal with the problem of misspecified cokriging prediction within the framework of fixed domain asymptotics. Specifically, we provide conditions for equivalence of measures associated with multivariate Gaussian random fields, with index set in a compact set of a  $d$ -dimensional Euclidean space. Such conditions have been elusive for over about 50 years of spatial statistics.

We then focus on the multivariate Matérn and Generalized Wendland classes of matrix valued covariance functions, that have been very popular for having parameters that are crucial to spatial interpolation, and that control the mean square differentiability of the associated Gaussian process. We provide sufficient conditions, for equivalence of Gaussian measures, relying on the covariance parameters of these two classes. This enables to identify the parameters that are crucial to asymptotically equivalent interpolation in multivariate geostatistics. Our findings are then illustrated through simulation studies.

*Keywords:* Cokriging; equivalence of Gaussian measures; fixed domain asymptotics; functional analysis; Generalized Wendland; Matérn; spectral analysis

## 1. Introduction

### 1.1. Context

Our paper deals with equivalence of Gaussian measures and asymptotically equivalent cokriging prediction in multivariate geostatistics. We consider a multivariate ( $p$ -variate) stationary Gaussian field  $\mathbf{Z} = \{\mathbf{Z}(s) = (Z_1(s), \dots, Z_p(s))^\top, s \in D\}$ , where  $D$  is a fixed bounded subset of  $\mathbb{R}^d$  with non-empty interior. Throughout, the integers  $d$  and  $p$  are fixed. The assumption of Gaussianity implies that modeling, inference and prediction depend exclusively on the mean of  $\mathbf{Z}$ , which is constant and assumed to be zero, and on the multivariate covariance function, being a  $p \times p$  matrix function  $\mathbf{R} = [R_{ij}]_{i,j=1}^p$ , defined in  $\mathbb{R}^d$ , such that

$$R_{ij}(\mathbf{h}) = \text{Cov}(Z_i(\mathbf{t}), Z_j(\mathbf{s})), \quad \mathbf{h} = \mathbf{t} - \mathbf{s},$$

for  $t, s \in D$  and  $i, j = 1, \dots, p$ . Throughout, the diagonal elements  $R_{ii}$  are called marginal covariances, whereas the off-diagonal members  $R_{ij}$  are called cross-covariances. The mapping  $\mathbf{R}$  must be positive definite, which means that

$$\sum_{\ell=1}^n \sum_{k=1}^n \mathbf{a}_\ell^\top \mathbf{R}(s_\ell - s_k) \mathbf{a}_k \geq 0, \tag{1}$$

for all positive integers  $n$ ,  $\{s_1, \dots, s_n\} \subset D$  and  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \subset \mathbb{R}^P$ .

Spatial prediction in multivariate geostatistics is known as cokriging, which is the analogue of best linear unbiased prediction in classical regression. Cokriging is based on the conditional expectation of a component of  $\mathbf{Z}$  at a target point  $s_0 \in D$ , given observed values of all its components at given observation locations  $\{s_1, \dots, s_n\} \subset D$ . We put emphasis on the following problems: How important is the multivariate covariance function? Which covariance parameters are important?

We provide answers to these two questions and, in doing so, we obtain general sufficient conditions for equivalence of multivariate Gaussian measures, that are of independent interest. Notice that [Zhang and Cai \(2015\)](#) have previously also addressed these two questions and provided partial answers only.

## 1.2. Literature review

### Multivariate covariance functions

Multivariate covariance functions in  $d$ -dimensional Euclidean spaces have become ubiquitous and we refer the reader to [Genton and Kleiber \(2015\)](#) for a detailed account. Recently, there has been some work on multivariate covariance functions on non planar surfaces, and the reader is referred to [Alegria and Porcu \(2017\)](#), [Alegria et al. \(2019\)](#), [Porcu, Bevilacqua and Genton \(2016\)](#) and [Bevilacqua et al. \(2019\)](#).

As for constructive methods to provide new models, the linear model of coregionalization ([Wackernagel, 2003](#)) is based on representing any component of the multivariate field  $\mathbf{Z}$  as a linear combination of latent, uncorrelated fields. Such a technique has been constructively criticized by [Gneiting, Kleiber and Schlather \(2010\)](#) and [Daley, Porcu and Bevilacqua \(2015\)](#) as the smoothness of any component of the multivariate field amounts to that of the roughest underlying univariate process. Moreover, the number of parameters can quickly become massive as the number of components increases. Scale mixture techniques as in [Porcu and Zastavnyi \(2011\)](#), as well as latent dimension approaches ([Apanasovich and Genton, 2010](#), [Porcu, Gregori and Mateu, 2006](#), [Porcu and Zastavnyi, 2011](#)) have been largely used to propose new multivariate models. [Bevilacqua, Hering and Porcu \(2015\)](#) call the following construction principle *multivariate parametric adaptation*: let  $\{R(\cdot; \lambda) : [0, \infty) \rightarrow \mathbb{R}, \lambda \in \mathbb{R}^k\}$  be a parametric family of continuous functions, such that  $R(\|\cdot\|; \lambda)$  is a correlation function in  $\mathbb{R}^d$  ( $R(0; \lambda) = 1$ ), indexed by a parameter vector  $\lambda = (\lambda_1, \dots, \lambda_k)^\top$ . Call  $\lambda_{ij} = (\lambda_{ij,1}, \dots, \lambda_{ij,k})^\top$ ,  $i, j = 1, \dots, p$  a collection of parameter vectors in  $\mathbb{R}^k$ . Then, define  $\mathbf{R} : [0, \infty) \rightarrow \mathbb{R}^{p \times p}$  through

$$\mathbf{R}(x) = [R_{ij}(x)]_{i,j=1}^p, \quad x \in [0, \infty),$$

with elements  $R_{ij}$  defined as

$$R_{ij}(x) = \sigma_{ii} \sigma_{jj} \rho_{ij} R(x; \lambda_{ij}), \quad x \in [0, \infty), \tag{2}$$

where  $\sigma_{ii}^2$  is the variance of the  $i$ th component of the multivariate random field and where  $\rho_{ii} = 1$  and  $\rho_{ij}$ ,  $i \neq j$ , is the collocated correlation coefficient. Thus, the problem is finding the restriction on the parameters  $\lambda_{ij}$  such that  $\mathbf{R}(\|\cdot\|)$  is positive definite as in (1).

A crucial benefit of this strategy, by comparison with the linear model of coregionalization, is a clear physical interpretation of the parameters (Bevilacqua, Hering and Porcu, 2015, Vallejos, Osorio and Bevilacqua, 2020). For example, for a bivariate random field ( $p = 2$ ), the colocated correlation parameter,  $\rho_{12}$ , expresses the marginal correlation between the components, since  $R_{12}(0) = R_{21}(0) = \rho_{12}$  if  $\sigma_{11}^2 = \sigma_{22}^2 = 1$ . In Euclidean spaces this strategy has been adopted by Apanasovich, Genton and Sun (2012), Gneiting, Kleiber and Schlather (2010) and by Daley, Porcu and Bevilacqua (2015).

### *Misspecified kriging predictions under infill asymptotics*

The study of asymptotic properties of (co)kriging predictors is complicated by the fact that more than one asymptotic framework can be considered when observing a single realization from a (multivariate) Gaussian field. Under infill asymptotics (also called fixed domain asymptotics), the typical assumption is that the sampling domain is bounded and that the sampling set becomes increasingly dense. Under increasing domain asymptotics, the sampling domain increases with the number of observed data, and the distance between any two observation locations is bounded away from zero (Bachoc, 2014, Mardia and Marshall, 1984).

The focus of this paper is on infill asymptotics. In this case, in the univariate case, a key concept is the equivalence of Gaussian measures (Ibragimov and Rozanov, 1978, Skorohod and Jadrenko, 1973). Furthermore, a long-standing object of attention is asymptotically optimal prediction when using a misspecified covariance function (the predictor is then called pseudo BLUP by Michael Stein Stein, 1999a). In the univariate case, Michael Stein has shown that, when the Gaussian measures obtained from the true and misspecified covariance function are equivalent, then the predictions under the misspecified covariance function are asymptotically efficient, and mean square errors are asymptotically equivalent to their targets (Stein, 1988, 1990, 1993, 1999b, 2004).

When working with specific covariance models, it is thus crucial to know which conditions on the parameters imply the equivalence of Gaussian measures. Specific results have been provided for the Matérn (Zhang, 2004) and Generalized Wendland (Bevilacqua et al., 2019) classes of covariance functions, associated with scalar valued random fields. These results themselves follow from earlier works on general conditions for equivalence of univariate Gaussian measures, in particular based on spectral densities (Skorohod and Jadrenko, 1973). Nevertheless, multivariate extensions of these various results are lacking. They are provided in the present paper.

## 1.3. Outline

The new challenges provided by modern spatial datasets, being typically multivariate, call for a deeper study of cokriging than is provided by the current literature discussed above. This is the object of this paper, where we deal with the problem of misspecified cokriging prediction within the framework of infill asymptotics. We provide general sufficient conditions for equivalence of Gaussian measures in the multivariate case, complementing contributions that are limited, since the early 70ies, to scalar valued random fields, as discussed above.

We then focus on the multivariate Matérn and Generalized Wendland classes of matrix valued covariance functions, that have been very popular in spatial statistics for having parameters that are crucial to spatial interpolation, and that control the mean square differentiability of the associated Gaussian process. We show parametric conditions ensuring these matrix valued covariance models to be compatible, that is, to yield equivalent Gaussian measures. Hence, we provide sufficient conditions for asymptotic equivalence of misspecified cokriging predictions. We confirm and illustrate this asymptotic equivalence numerically.

The outline of the paper is the following: Section 2 contains the necessary mathematical and probabilistic background. Section 3 contains general results about compatible matrix valued covariance functions. Section 4 relates on the compatibility between the Matérn and Generalized Wendland parametric classes of multivariate covariance functions. Section 5 inspects the problem of cokriging predictions through these models. Our findings are then illustrated through a simulation study in Section 6. Section 7 provides the proofs for Section 3. The remaining proofs are given in the supplementary material Bachoc *et al.* (2021).

## 2. Background and notation

### 2.1. Multivariate covariance functions and function spaces

Let  $d, p$  be positive integers. Let  $\mathbf{R} : \mathbb{R}^d \rightarrow \mathbb{R}^{p \times p}$  be positive definite. We let the elements  $R_{ij}$  of  $\mathbf{R}$  be continuous in  $\mathbb{R}^d$ . The matrix spectral density of  $\mathbf{R}$  is the  $p \times p$  matrix function  $\mathbf{F} = [F_{ij}]_{i,j=1}^p$  defined by

$$R_{ij}(\mathbf{h}) = \int_{\mathbb{R}^d} F_{ij}(\lambda) e^{i\mathbf{h}^\top \lambda} d\lambda,$$

for  $\mathbf{h} \in \mathbb{R}^d$  and  $i, j = 1, \dots, p$ . Here  $i$  is the complex number satisfying  $i^2 = -1$ . Note that a sufficient condition for  $\mathbf{F}$  to be well defined is that  $\mathbf{R}$  has elements  $R_{ij}$  that are pointwise absolutely integrable in  $\mathbb{R}^d$ , and that the same holds for the Fourier transforms of these elements.

For  $a = 0, 1$ , we consider a stationary matrix covariance function  $\mathbf{R}^{(a)} = [R_{ij}^{(a)}]_{i,j=1}^p$  on  $\mathbb{R}^d$ . We assume that, for  $a = 0, 1$  and  $i, j = 1, \dots, p$ , the function  $R_{ij}^{(a)}$  is summable on  $\mathbb{R}^d$  and that  $\mathbf{R}^{(a)}$  has matrix spectral density  $\mathbf{F}^{(a)} = [F_{ij}^{(a)}]_{i,j=1}^p$ .

We further assume that for  $a = 0, 1$  and  $j = 1, \dots, p$ ,  $F_{jj}^{(a)}$  is real-valued, strictly positive on  $\mathbb{R}^d$  and summable on  $\mathbb{R}^d$ . We remark that for  $a = 0, 1$  and  $i, j = 1, \dots, p$ ,  $i \neq j$ ,  $F_{ij}^{(a)}$  is complex-valued and we also assume that  $|F_{ij}^{(a)}|$  is summable on  $\mathbb{R}^d$ , with  $|z|$  the modulus of  $z \in \mathbb{C}$ . Cramér’s theorem shows that, for any  $\lambda \in \mathbb{R}^d$  and  $a = 0, 1$ , the matrix  $\mathbf{F}^{(a)}(\lambda)$  is Hermitian with non-negative eigenvalues.

For a  $p \times p$  Hermitian matrix  $\mathbf{M}$ , we let  $\lambda_1(\mathbf{M}) \leq \dots \leq \lambda_p(\mathbf{M})$  be its  $p$  eigenvalues. If  $\mathbf{M}$  is non-negative definite, we let  $\mathbf{M}^{1/2}$  be its unique Hermitian non-negative definite square root. For a square complex matrix  $\mathbf{N}$ , we let  $\|\mathbf{N}\|$  be its largest singular value. For a complex column vector  $\mathbf{v}$ , we let  $\bar{\mathbf{v}}$  be composed of the conjugates of  $\mathbf{v}$  and  $\|\mathbf{v}\|^2 = \bar{\mathbf{v}}^\top \mathbf{v}$ . For two  $p \times p$  Hermitian matrices  $\mathbf{M}$  and  $\mathbf{N}$  we write  $\mathbf{M} \geq \mathbf{N}$  when for all  $\mathbf{v} \in \mathbb{C}^p$ ,  $\bar{\mathbf{v}}^\top \mathbf{M} \mathbf{v} \geq \bar{\mathbf{v}}^\top \mathbf{N} \mathbf{v}$ . We let  $\mathbf{e}_1, \dots, \mathbf{e}_q$  be the  $q$  basis column vectors of  $\mathbb{R}^q$  for  $q \in \mathbb{N}$ .

For a summable function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , we let the Fourier transform  $\mathcal{F}(f)$  of  $f$  be defined by, for  $\lambda \in \mathbb{R}^d$ ,

$$\mathcal{F}(f)(\lambda) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} f(\mathbf{t}) e^{-i\lambda^\top \mathbf{t}} d\mathbf{t}.$$

For a bounded subset  $S$  of  $\mathbb{R}^d$  with non-empty interior, we let  $\mathcal{W}_S$  be the set of functions from  $\mathbb{R}^d$  to  $\mathbb{C}^p$  of the form  $(f_1, \dots, f_p)^\top$ , where for  $i = 1, \dots, p$ ,  $f_i = \mathcal{F}(g_i)$  for a function  $g_i$  in  $L^2(\mathbb{R}^d)$  that is zero outside of  $S$ . As observed in Skorohod and Jadrenko (1973), a function  $(f_1, \dots, f_p)^\top$  in  $\mathcal{W}_S$  satisfies  $\int_{\mathbb{R}^d} (|f_1(\lambda)|^2 + \dots + |f_p(\lambda)|^2) d\lambda < \infty$ . Consider a matrix function  $\lambda \mapsto \mathbf{F}(\lambda)$  with  $\lambda \in \mathbb{R}^d$  and with  $\mathbf{F}(\lambda)$  a  $p \times p$  Hermitian strictly positive definite matrix and assume that  $\|\mathbf{F}\|$  and  $\lambda_p(\mathbf{F})/\lambda_1(\mathbf{F})$  are bounded

on  $\mathbb{R}^d$ . Then we define  $\mathcal{W}_S(\mathbf{F})$  as the closure of  $\mathcal{W}_S$  in the metric

$$\|f\|_{\mathcal{W}_S(\mathbf{F})}^2 = \int_{\mathbb{R}^d} \bar{f}(\lambda)^\top \mathbf{F}(\lambda) f(\lambda) d\lambda.$$

We remark that  $\mathcal{W}_S(\mathbf{F})$  is a (complex) separable Hilbert space, with inner-product given by

$$(f_1, f_2)_{\mathcal{W}_S(\mathbf{F})} = \int_{\mathbb{R}^d} \bar{f}_1(\lambda)^\top \mathbf{F}(\lambda) f_2(\lambda) d\lambda.$$

Indeed, for  $f = (f_1, \dots, f_p)^\top \in \mathcal{W}_S(\mathbf{F})$ , for  $i = 1, \dots, p$ ,  $f_i$  is included in the space of square integrable functions w.r.t. the measure  $\|\mathbf{F}(\lambda)\| d\lambda$  which is separable and complete.

For  $S \subset \mathbb{R}^d$ , we let  $L_S^{2,p}$  be the Hilbert space of the vectors of functions of the form  $(f_1, \dots, f_p)^\top$ , with  $f_i : S \rightarrow \mathbb{C}$  square summable for  $i = 1, \dots, p$ , endowed with the inner product

$$((f_1, \dots, f_p)^\top, (g_1, \dots, g_p)^\top)_{L_S^{2,p}} = \int_S \bar{f}_1(t) g_1(t) dt + \dots + \int_S \bar{f}_p(t) g_p(t) dt.$$

## 2.2. The univariate Matérn and Generalized Wendland covariance functions

We start by describing the two univariate classes of covariance functions that will be used throughout as building blocks for matrix valued covariance functions.

1. The Matérn function (Stein, 1999a) is defined as:

$$\mathcal{M}_{\nu, \alpha}(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{r}{\alpha}\right)^\nu \mathcal{K}_\nu\left(\frac{r}{\alpha}\right), \quad r \geq 0, \tag{3}$$

where  $r = \|\mathbf{h}\|$ ,  $\mathbf{h} \in \mathbb{R}^d$ . The Matérn covariance function is positive definite in  $\mathbb{R}^d$  for all positive  $\alpha$  and for any value of  $d$ . Here,  $\mathcal{K}_\nu$  is a modified Bessel function of the second kind of order  $\nu$ . The parameter  $\nu > 0$  characterizes the differentiability at the origin and, as a consequence, the differentiability of the associated sample paths. In particular for a positive integer  $k$ , the sample paths are  $k$  times differentiable, in any direction, if and only if  $\nu > k$ . When  $\nu = 1/2 + m$  and  $m$  is a nonnegative integer, the Matérn function simplifies to the product of a negative exponential with a polynomial of degree  $m$ , and for  $\nu$  tending to infinity, a rescaled version of the Matérn converges to a squared exponential model being infinitely differentiable at the origin. Thus, the Matérn function allows for a continuous parameterization of its associated Gaussian field in terms of smoothness.

2. The Generalized Wendland function (Gneiting, 2002b, Zastavnyĭ, 2006) is defined, for  $\kappa, \beta > 0$ , as

$$\mathcal{W}_{\mu, \kappa, \beta}(r) := \begin{cases} \frac{1}{B(2\kappa, \mu + 1)} \int_{r/\beta}^1 u(u^2 - (r/\beta)^2)^{\kappa-1} (1-u)^\mu du, & 0 \leq r/\beta < 1, \\ 0, & r/\beta \geq 1, \end{cases} \tag{4}$$

with  $B$  denoting the beta function, and where  $r = \|\mathbf{h}\|$ ,  $\mathbf{h} \in \mathbb{R}^d$ . The function  $\mathcal{W}_{\mu, \kappa, \beta}(r)$  is positive definite in  $\mathbb{R}^d$  if and only if

$$\mu \geq (d + 1)/2 + \kappa. \tag{5}$$

Note that  $\mathcal{W}_{\mu,0,\beta}$  is not defined because  $\kappa$  must be strictly positive. In this special case we consider the Askey function (Askey, 1973)

$$\mathcal{A}_{\mu,\beta}(r) := \begin{cases} (1 - r/\beta)^\mu, & 0 \leq r/\beta < 1, \\ 0, & r/\beta \geq 1. \end{cases}$$

Arguments in Golubov (1981) show that  $\mathcal{A}_{\mu,\beta}$  is positive definite if and only if  $\mu \geq (d + 1)/2$  and we define  $\mathcal{W}_{\mu,0,\beta} := \mathcal{A}_{\mu,\beta}$ .

Closed form solution of the integral in (4) can be obtained when  $\kappa = k$ , a positive integer. In this case,  $\mathcal{W}_{\mu,k,\beta}(r) = \mathcal{A}_{\mu+k,\beta}(r)P_k(r)$ , with  $P_k$  a polynomial of order  $k$ . These functions, termed (original) Wendland functions, were originally proposed by Wendland (1995).

Other closed form solutions of integral (4) can be obtained when  $\kappa = k + 1/2$ , using some results in Schaback (2011). Such solutions are called *missing Wendland* functions.

As noted by Gneiting (2002a), Generalized Wendland and Matérn functions exhibit the same behavior at the origin, with the smoothness parameters of the two covariance models related by the equation  $\nu = \kappa + 1/2$ .

Here, for a positive integer  $k$ , the sample paths of a Gaussian field with Generalized Wendland function are  $k$  times differentiable, in any direction, if and only if  $\kappa > k - 1/2$ .

### 2.3. Multivariate Matérn and Generalized Wendland models

We now consider the multivariate parametric adaptation, illustrated through Equation (2), as a construction principle for multivariate covariance functions.

We follow Gneiting, Kleiber and Schlather (2010) and Daley, Porcu and Bevilacqua (2015) to couple construction (2) with, respectively, the Matérn model (3) and the Generalized Wendland model (4), to obtain:

**1. The Multivariate Matérn model**, denoted  $\mathcal{MM}_\theta$ , and defined as

$$\mathcal{MM}_\theta(r) = \left[ \rho_{ij} \sigma_{ii} \sigma_{jj} \mathcal{M}_{\nu, \alpha_{ij}}(r) \right]_{i,j=1}^p, \quad \rho_{ii} = 1, \rho_{ij} = \rho_{ji}, \alpha_{ij} = \alpha_{ji}, i, j = 1, \dots, p, \quad (6)$$

where  $\theta = ([\sigma_{ii}]_{i=1, \dots, p}, [\rho_{ij}]_{i,j=1, \dots, p, i < j}, \nu, [\alpha_{ij}]_{i,j=1, \dots, p, i \leq j})$ ;

**2. The Multivariate Generalized Wendland model**, denoted  $\mathcal{MW}_\lambda$ , and defined as

$$\mathcal{MW}_\lambda(r) = \left[ \rho_{ij} \sigma_{ii} \sigma_{jj} \mathcal{W}_{\mu, \kappa, \beta_{ij}}(r) \right]_{i,j=1}^p, \quad \rho_{ii} = 1, \rho_{ij} = \rho_{ji}, \beta_{ij} = \beta_{ji}, i, j = 1, \dots, p, \quad (7)$$

where  $\lambda = ([\sigma_{ii}]_{i=1, \dots, p}, [\rho_{ij}]_{i,j=1, \dots, p, i < j}, \mu, \kappa, [\beta_{ij}]_{i,j=1, \dots, p, i \leq j})$ .

Note that, in principle, the smoothness parameters  $\nu$  and  $\kappa$  for both models can change through the components. Nevertheless, in this paper we assume common smoothness parameters.

Henceforth, for the multivariate Matérn, we assume the following condition on the parameter  $\theta$ :

$$\inf_{z \geq 0} \lambda_1 \left( \left[ \rho_{ij} \sigma_{ii} \sigma_{jj} \frac{\Gamma(\nu + d/2)}{\pi^{d/2} \Gamma(\nu)} \frac{\alpha_{ij}^d (1+z)^{2\nu+d}}{(1 + \alpha_{ij}^2 z^2)^{\nu+d/2}} \right]_{i,j=1}^p \right) > 0. \quad (8)$$

The condition (8) is interpreted as follows (see the supplementary material Bachoc *et al.*, 2021 for more details). The matrix spectral density of the Matérn model is real-valued and its diagonal elements are of order  $(1 + \|\lambda\|)^{-2\nu-d}$  for large frequency  $\lambda \in \mathbb{R}^d$ . After normalization by  $(1 + \|\lambda\|)^{2\nu+d}$  yielding lower and upper bounded diagonal elements, we ask the resulting symmetric matrix to have its smallest

eigenvalue bounded from below by zero. This enables the satisfaction of Condition 1 in Section 3 below. This requirement is only slightly more than asking for the matrix spectral density to be positive definite, which is necessary. Hence, the condition (8) is arguably mild.

Similarly, for the multivariate Generalized Wendland model we assume that

$$\inf_{z \geq 0} \lambda_1 \left( \left[ \rho_{ij} \sigma_{ii} \sigma_{jj} \beta_{ij}^d (1+z)^{d+1+2\kappa} {}_1F_2 \left( \zeta; \zeta + \frac{\mu}{2}, \zeta + \frac{\mu}{2} + \frac{1}{2}; -\frac{(z\beta_{ij})^2}{4} \right) \right]_{i,j=1}^p \right) > 0, \tag{9}$$

where  $\zeta = (d + 1)/2 + \kappa$  and

$${}_1F_2(a; b, c; z) = \sum_{k=0}^{\infty} \frac{(a)_k z^k}{(b)_k (c)_k k!}, \quad z \in \mathbb{R}, \tag{10}$$

is a special case of the generalized hypergeometric functions  ${}_qF_p$  (Abramowitz and Stegun, 1966), with  $(q)_k = \Gamma(q + k)/\Gamma(q)$  for  $k \in \mathbb{N} \cup \{0\}$ , being the Pochhammer symbol. The discussion of the condition (9) is similar to that of the condition (8).

**Remark 1.** In the bivariate case, the condition (8) is equivalent to

$$\rho_{12}^2 < \frac{\alpha_{12}^{4\nu}}{\alpha_{11}^{2\nu} \alpha_{22}^{2\nu}} \inf_{z \geq 0} \frac{(\alpha_{12}^{-2} + z^2)^{2\nu+d}}{(\alpha_{11}^{-2} + z^2)^{\nu+d/2} (\alpha_{22}^{-2} + z^2)^{\nu+d/2}} \tag{11}$$

and the condition (9) is equivalent to

$$\rho_{12}^2 < \frac{\beta_{11}^d \beta_{22}^d}{\beta_{12}^{2d}} \inf_{z \geq 0} \frac{{}_1F_2 \left( \zeta; \zeta + \frac{\mu}{2}, \zeta + \frac{\mu}{2} + \frac{1}{2}; -\frac{(z\beta_{11})^2}{4} \right) {}_1F_2 \left( \zeta; \zeta + \frac{\mu}{2}, \zeta + \frac{\mu}{2} + \frac{1}{2}; -\frac{(z\beta_{22})^2}{4} \right)}{\left( {}_1F_2 \left( \zeta; \zeta + \frac{\mu}{2}, \zeta + \frac{\mu}{2} + \frac{1}{2}; -\frac{(z\beta_{12})^2}{4} \right) \right)^2}. \tag{12}$$

These two latter conditions are obtained by writing that the determinants of the (normalized)  $2 \times 2$  matrix spectral densities of the Matérn and Generalized Wendland models are bounded from below by 0. The proof of the two equivalences is given in the supplementary material Bachoc et al. (2021). The condition (11), with “ $\leq$ ” instead of “ $<$ ”, is necessary and sufficient for the bivariate Matérn model to be valid (to indeed be positive definite), see Gneiting, Kleiber and Schlather (2010, Theorem 3).

### 2.4. Equivalence of Gaussian measures and cokriging

Equivalence and orthogonality of probability measures are useful tools when assessing the asymptotic properties of both prediction and estimation for Gaussian fields. We denote with  $P^{(a)}$ ,  $a = 0, 1$ , two probability measures defined on the same measurable space  $\{\Omega, \mathcal{F}\}$ . The measures  $P^{(0)}$  and  $P^{(1)}$  are called equivalent (denoted  $P^{(0)} \equiv P^{(1)}$ ) if, for any  $A \in \mathcal{F}$ ,  $P^{(1)}(A) = 1$  implies  $P^{(0)}(A) = 1$  and vice versa. On the other hand,  $P^{(0)}$  and  $P^{(1)}$  are orthogonal if there exists an event  $A$  such that  $P^{(1)}(A) = 1$  but  $P^{(0)}(A) = 0$ . For a  $p$ -variate Gaussian random field  $\mathbf{Z} : \Omega \times D \rightarrow \mathbb{R}^p$ , to define previous concepts, we restrict the event  $A$  to the  $\sigma$ -algebra generated by  $\mathbf{Z}$  and we emphasize this restriction by saying that the two measures are equivalent on the paths of  $\mathbf{Z}$ . It is well known that two Gaussian measures (that is two measures on  $\Omega$  such that  $\mathbf{Z}$  is Gaussian) are either equivalent or orthogonal on the paths of  $\mathbf{Z}$  (Ibragimov and Rozanov, 1978).

Since a Gaussian measure is completely characterized by the mean function and matrix covariance function, we write  $P(\mathbf{R})$  for a Gaussian measure on  $(\Omega, \mathcal{F})$  such that  $\mathbf{Z}$  has zero mean and matrix covariance function  $\mathbf{R}$ . We also write  $P(\mathbf{R}^{(0)}) \equiv P(\mathbf{R}^{(1)})$  on the paths of  $\mathbf{Z}$ , if two Gaussian measures with mean zero and the matrix covariance functions  $\mathbf{R}^{(0)}$  and  $\mathbf{R}^{(1)}$  are equivalent on the paths of  $\mathbf{Z}$ .

For the remainder of the paper, we call  $\mathbf{R}^{(0)}$  and  $\mathbf{R}^{(1)}$  *compatible* when  $P(\mathbf{R}^{(0)}) \equiv P(\mathbf{R}^{(1)})$  on the paths of  $\mathbf{Z}$ .

A direct implication of the celebrated result by Blackwell and Dubins (1962) is that, if two matrix valued covariance functions are compatible, then the two cokriging predictors are asymptotically equivalent (under fixed domain asymptotics).

### 3. General results

Let us consider two matrix covariance functions  $\mathbf{R}^{(0)}$  and  $\mathbf{R}^{(1)}$  with associated matrix spectral densities  $\mathbf{F}^{(0)}$  and  $\mathbf{F}^{(1)}$  and let  $P(\mathbf{R}^{(0)})$  and  $P(\mathbf{R}^{(1)})$  be the associated Gaussian measures. The next condition is our general technical requirement on  $\mathbf{R}^{(0)}$  and  $\mathbf{R}^{(1)}$ . As shown in the supplementary material Bachoc et al. (2021), this condition holds for the Matérn and Generalized Wendland matrix covariance functions.

**Condition 1.** There exist two constants  $0 < c_1 < c_2 < \infty$  and a function  $\gamma : \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $(\gamma, \dots, \gamma)^\top$  is in  $\mathcal{W}_{[-b, b]^d}$  for some fixed  $0 < b < \infty$  and, for all  $\lambda \in \mathbb{R}^d$ ,

$$c_1 \gamma^2(\lambda) \mathbf{I}_p \leq \mathbf{F}^{(0)}(\lambda) \leq c_2 \gamma^2(\lambda) \mathbf{I}_p,$$

$$c_1 \gamma^2(\lambda) \mathbf{I}_p \leq \mathbf{F}^{(1)}(\lambda) \leq c_2 \gamma^2(\lambda) \mathbf{I}_p.$$

Condition 1 can be interpreted as follows. First, for the two matrix covariance functions, the (univariate) spectral densities of the  $p$  components of the  $p$ -variate Gaussian field have the same asymptotic behavior (given by the function  $\gamma^2$ ) as  $\|\lambda\| \rightarrow \infty$ . Second, again for the two matrix covariance functions, the matrix spectral density is asymptotically well-conditioned as  $\|\lambda\| \rightarrow \infty$ .

We now provide a fundamental result for this paper. It relates about a sufficient condition for the compatibility of  $\mathbf{R}^{(0)}$  and  $\mathbf{R}^{(1)}$ . It is an extension of Theorem 1 in Skorohod and Jadrenko (1973) from the univariate to the multivariate case.

**Theorem 1.** Assume that Condition 1 holds, and that there exists a matrix-valued function  $\mathbf{B}$  on  $(\mathbb{R}^d)^2$ , such that for  $\lambda, \mu \in \mathbb{R}^d$ ,  $\mathbf{B}(\lambda, \mu) = [b_{ij}(\lambda, \mu)]_{i,j=1}^p$  is a  $p \times p$  complex matrix. Let  $\mathbf{B}(\lambda, \mu) = \bar{\mathbf{B}}(\mu, \lambda)^\top$  for all  $\lambda, \mu \in \mathbb{R}^d$ . Assume also that for  $i, j = 1, \dots, p$ , we have

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |b_{ij}(\lambda, \mu)|^2 \|\mathbf{F}^{(0)}(\lambda)\| \|\mathbf{F}^{(0)}(\mu)\| d\lambda d\mu < +\infty. \tag{13}$$

Assume then that we have, for  $\mathbf{t}, \mathbf{s} \in D$  and  $\mathbf{h} = \mathbf{t} - \mathbf{s}$ ,

$$\mathbf{R}^{(1)}(\mathbf{h}) - \mathbf{R}^{(0)}(\mathbf{h}) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-i\lambda^\top \mathbf{t} + i\mu^\top \mathbf{s}} \mathbf{F}^{(0)}(\lambda) \mathbf{B}(\lambda, \mu) \mathbf{F}^{(0)}(\mu) d\lambda d\mu. \tag{14}$$

Then,  $P(\mathbf{R}^{(0)}) \equiv P(\mathbf{R}^{(1)})$  on the paths of  $\mathbf{Z}$ .



Notice that the function in the integral (14) is summable because of Cauchy–Schwarz inequality in concert with (13) and Condition 1.

We now provide a second fundamental result for this paper. The proof of Theorem 2 relies on Theorem 1. Theorem 2 is particularly well applicable to specific models of covariance functions, as it just requires to show that two matrix spectral densities are sufficiently close for large frequencies. This enables us to address the Matérn and Generalized Wendland models in Section 4. Theorem 2 is an extension of Theorem 4 in Skorohod and Jadrenko (1973) from the univariate to the multivariate case.

**Theorem 2.** *Assume that Condition 1 holds, and that*

$$\int_{\mathbb{R}^d} \frac{1}{\gamma(\lambda)^4} \|F^{(0)}(\lambda) - F^{(1)}(\lambda)\|^2 d\lambda < \infty.$$

*Then,  $P(\mathbf{R}^{(0)}) \equiv P(\mathbf{R}^{(1)})$  on the paths of  $\mathbf{Z}$ .*

As written just above, the sufficient condition in Theorem 2 is that the two matrix spectral densities have the same asymptotic behavior for large frequency. This can be interpreted as the two multivariate covariance functions having the same smoothness and local behavior at zero. This is well interpreted, since under fixed domain asymptotics, the smoothness and local behavior of a univariate covariance function at zero are of central importance, as is for instance highlighted by the vast literature on their estimation Anderes (2010), Azais et al. (2020), Istas and Lang (1997), Lang and Roueff (2001), Loh (2015).

## 4. Compatible multivariate Matérn and Generalized Wendland correlation models

Let us consider the parameter vectors

$$\theta^{(a)} = ([\sigma_{ii}^{(a)}]_{i=1,\dots,p}, [\rho_{ij}^{(a)}]_{i,j=1,\dots,p,i < j}, \nu, [\alpha_{ij}^{(a)}]_{i,j=1,\dots,p,i \leq j})$$

and

$$\lambda^{(a)} = ([\sigma_{ii}^{(a)}]_{i=1,\dots,p}, [\rho_{ij}^{(a)}]_{i,j=1,\dots,p,i < j}, \mu, \kappa, [\beta_{ij}^{(a)}]_{i,j=1,\dots,p,i \leq j}),$$

for  $a = 0, 1$ .

Our first result gives sufficient conditions for the compatibility of two multivariate Matérn models with a common smoothness parameter.

**Theorem 3.** *Let  $\nu > 0$ . If*

$$\frac{\sigma_{ii}^{(0)} \sigma_{jj}^{(0)} \rho_{ij}^{(0)}}{(\alpha_{ij}^{(0)})^{2\nu}} = \frac{\sigma_{ii}^{(1)} \sigma_{jj}^{(1)} \rho_{ij}^{(1)}}{(\alpha_{ij}^{(1)})^{2\nu}}, \quad i, j = 1, \dots, p, \tag{15}$$

*then for  $d = 1, 2, 3$ , the matrix valued covariance models  $\mathcal{MM}_{\theta^{(0)}}$  and  $\mathcal{MM}_{\theta^{(1)}}$  are compatible.*

Some comments are in order. For each pair of covariance or cross covariance functions, the equality condition (15) is the same as in the univariate case (Zhang, 2004). In the bivariate case, Zhang and Cai (2015) provide conditions for compatibility of  $\mathcal{MM}_{\theta^{(0)}}$  and  $\mathcal{MM}_{\theta^{(1)}}$  for a very special case, where

$\alpha_{ij}^{(a)} = \alpha^{(a)} > 0$  for all  $i, j = 1, 2$  and  $a = 0, 1$ . Hence, the authors consider two separable models. Thus, Theorem 3 allows for a considerable improvement with respect to Zhang and Cai (2015) as it allows for different range parameters among the covariance and cross covariance functions. In the case where  $p = 2$  and  $\alpha_{11}^{(a)} = \alpha_{22}^{(a)} = \alpha_{12}^{(a)}$  for  $a = 0, 1$ , Theorem 3 coincides with Zhang and Cai (2015).

Our second result gives sufficient conditions for the compatibility of two multivariate Generalized Wendland models with a common smoothness parameter.

**Theorem 4.** For a given  $\kappa \geq 0$ , let  $\mu > d + 1/2 + \kappa$ . If

$$\frac{\sigma_{ii}^{(0)} \sigma_{jj}^{(0)} \rho_{ij}^{(0)}}{(\beta_{ij}^{(0)})^{1+2\kappa}} = \frac{\sigma_{ii}^{(1)} \sigma_{jj}^{(1)} \rho_{ij}^{(1)}}{(\beta_{ij}^{(1)})^{1+2\kappa}}, \quad i, j = 1, \dots, p, \tag{16}$$

then for  $d = 1, 2, 3$ , the matrix valued covariance models  $\mathcal{M}\mathcal{W}_{\lambda^{(0)}}$  and  $\mathcal{M}\mathcal{W}_{\lambda^{(1)}}$  are compatible.

Our third result gives sufficient conditions for the compatibility of a multivariate Matérn model with a multivariate Generalized Wendland. To simplify notation, we let  $\alpha_{ij}^{(0)} = \alpha_{ij}$  and  $\beta_{ij}^{(1)} = \beta_{ij}$  for  $i, j = 1, \dots, p$ .

**Theorem 5.** For given  $\nu \geq 1/2$  and  $\kappa \geq 0$ , if  $\nu = \kappa + 1/2$ ,  $\mu > d + 1/2 + \kappa$ , and

$$\frac{\sigma_{ii}^{(0)} \sigma_{jj}^{(0)} \rho_{ij}^{(0)}}{\alpha_{ij}^{2\nu}} = C_{\kappa, \mu} \frac{\sigma_{ii}^{(1)} \sigma_{jj}^{(1)} \rho_{ij}^{(1)}}{\beta_{ij}^{1+2\kappa}}, \quad i, j = 1, \dots, p, \tag{17}$$

$C_{\kappa, \mu} = \mu\Gamma(2\kappa + \mu + 1) / \Gamma(\mu + 1)$  then for  $d = 1, 2, 3$ , the matrix valued covariance models  $\mathcal{M}\mathcal{M}_{\theta^{(0)}}$  and  $\mathcal{M}\mathcal{W}_{\lambda^{(1)}}$  are compatible.

Theorems 4 and 5 have no existing counterpart, even in the restricted setting where  $\alpha_{ij}^{(a)} = \alpha^{(a)}$  and  $\beta_{ij}^{(a)} = \beta^{(a)}$  for all  $i, j = 1, \dots, p$  and  $a = 0, 1$ . Again, for each pair of covariance or cross covariance functions, the conditions (16) and (17) on the covariance parameters are the same as in the univariate case in Bevilacqua et al. (2019) (see the proof of Theorem 5 that relates  $C_{\kappa, \mu}$  to the constants used in Bevilacqua et al., 2019).

To conclude Section 4, we note that in Theorem 3, the smoothness parameter  $\nu$  takes the same value between the two models  $\mathcal{M}\mathcal{M}_{\theta^{(0)}}$  and  $\mathcal{M}\mathcal{M}_{\theta^{(1)}}$ . If two different values of the smoothness parameter  $\nu$  were considered for the two models  $\mathcal{M}\mathcal{M}_{\theta^{(0)}}$  and  $\mathcal{M}\mathcal{M}_{\theta^{(1)}}$ , then these models would not be compatible. Indeed, the univariate covariance functions  $(\mathcal{M}\mathcal{M}_{\theta^{(0)}})_{11}$  and  $(\mathcal{M}\mathcal{M}_{\theta^{(1)}})_{11}$  would not be compatible, as follows for instance from the fact that the two fixed values of  $\nu$  are consistently estimable in input dimension one (see for instance Loh, 2015) and that  $D$  has non-empty interior. Similar discussions apply to Theorems 4 and 5.

## 5. Cokriging predictions with multivariate Generalized Wendland and Matérn models

We now consider prediction at a new target location  $s_0 \in D$  given a realization of a zero mean multivariate Gaussian field, using the multivariate Matérn and Generalized Wendland model, under fixed domain asymptotics. Specifically, we focus on two properties: asymptotic efficiency of prediction and

asymptotically correct estimation of prediction variance. Stein (1988), in the univariate case, shows that both asymptotic properties hold when kriging prediction is performed with two equivalent Gaussian measures.

Let  $\{s_{i,1}, \dots, s_{i,n_i} \in D \subset \mathbb{R}^d\}, i = 1, \dots, p$ , be any  $p$  sets of two-by-two distinct observation locations. Let  $\mathbf{Z}_{n_1, \dots, n_p} = (\mathbf{Z}_{1;n_1}^\top, \dots, \mathbf{Z}_{p;n_p}^\top)^\top$  be the observation vector obtained from a  $p$ -variate Gaussian field  $\{\mathbf{Z}(s) = (Z_1(s), \dots, Z_p(s))^\top, s \in D\}$ , where  $\mathbf{Z}_{i;n_i} = (Z_i(s_{i,1}), \dots, Z_i(s_{i,n_i}))^\top, i = 1, \dots, p$ . Remark that we do not necessarily assume collocated observation locations, that is the  $p$  sets  $\{s_{i,1}, \dots, s_{i,n_i}\}, i = 1, \dots, p$ , can be different.

Suppose we want to predict the first of the  $p$  components of the multivariate random field at  $s_0$  that is  $Z_1(s_0), s_0 \in D$ , using a misspecified multivariate model. For simplicity we only consider prediction for the first as symmetrical arguments hold for the other  $p - 1$  components. Specifically, we denote with  $MC^{(a)} = [C_{ij}^{(a)}]_{i,j=1}^p, a = 0, 1$  the true and misspecified matrix covariance function respectively.

Let  $\mathbf{c}_{1;n_1, \dots, n_p}^{(a)} = (\mathbf{c}_{1;1;n_1}^{(a)\top}, \dots, \mathbf{c}_{p;1;n_p}^{(a)\top})^\top$  with  $\mathbf{c}_{i;1;n_i}^{(a)} = [C_{1i}^{(a)}(\|s_0 - s_{i,\ell}\|)]_{\ell=1}^{n_i}, i = 1, \dots, p$ , the vector covariances between the location to predict and  $\mathbf{Z}_{n_1, \dots, n_p}$ . Let also  $\mathbf{C}_{n_1, \dots, n_p}^{(a)}$  be the  $(n_1 + \dots + n_p) \times (n_1 + \dots + n_p)$  matrix, with block  $i, j$ , of size  $n_i \times n_j$ , given by  $[C_{ij}^{(a)}(\|s_{i,\ell} - s_{j,k}\|)]_{\ell=1, k=1}^{n_i, n_j}, i, j = 1, \dots, p$ , the variance-covariance matrix associated with  $\mathbf{Z}_{n_1, \dots, n_p}$ .

The (misspecified) optimal predictor of  $Z_1(s_0)$ , using  $MC^{(0)}$  and  $MC^{(1)}$ , is given by,

$$\widehat{\mathbf{Z}}_{1;n_1, \dots, n_p}^{MC^{(a)}}(s_0) = \mathbf{c}_{1;n_1, \dots, n_p}^{(a)\top} (\mathbf{C}_{n_1, \dots, n_p}^{(a)})^{-1} \mathbf{Z}_{n_1, \dots, n_p}.$$

Under the correct model  $MC^{(0)}$ , the mean squared prediction error based on  $MC^{(1)}$  is given by

$$\begin{aligned} & \text{Var}_{MC^{(0)}} \left[ \widehat{\mathbf{Z}}_{1;n_1, \dots, n_p}^{MC^{(1)}}(s_0) - Z_1(s_0) \right] \\ &= (\sigma_{11}^{(0)})^2 - 2 \mathbf{c}_{1;n_1, \dots, n_p}^{(1)\top} (\mathbf{C}_{n_1, \dots, n_p}^{(1)})^{-1} \mathbf{c}_{1;n_1, \dots, n_p}^{(0)} \\ & \quad + \mathbf{c}_{1;n_1, \dots, n_p}^{(1)\top} (\mathbf{C}_{n_1, \dots, n_p}^{(1)})^{-1} \mathbf{C}_{n_1, \dots, n_p}^{(0)} (\mathbf{C}_{n_1, \dots, n_p}^{(1)})^{-1} \mathbf{c}_{1;n_1, \dots, n_p}^{(1)} \end{aligned} \tag{18}$$

and if the true and misspecified models coincide then (18) simplifies for  $a = 0, 1$  to

$$\text{Var}_{MC^{(a)}} \left[ \widehat{\mathbf{Z}}_{1;n_1, \dots, n_p}^{MC^{(a)}}(s_0) - Z_1(s_0) \right] = (\sigma_{11}^{(a)})^2 - \mathbf{c}_{1;n_1, \dots, n_p}^{(a)\top} (\mathbf{C}_{n_1, \dots, n_p}^{(a)})^{-1} \mathbf{c}_{1;n_1, \dots, n_p}^{(a)}.$$

The following theorem follows directly from the arguments in (Stein, 1999a, Section 4.3) extended to the multivariate case. Hence the proof is omitted. We remark that these arguments indeed do not require collocated observation locations.

**Theorem 6.** For  $i = 1, \dots, p$ , let  $\{s_{i,1}, \dots, s_{i,n_i}\}$  be dense in  $D$  as  $n_i \rightarrow \infty$ . For all  $s_0 \in D$ , if  $P(MC^{(0)}) \equiv P(MC^{(1)})$  on the paths of  $\mathbf{Z}$  then:

1. As  $n_1, \dots, n_p \rightarrow \infty$

$$\frac{\text{Var}_{MC^{(0)}} \left[ \widehat{\mathbf{Z}}_{1;n_1, \dots, n_p}^{MC^{(1)}}(s_0) - Z_1(s_0) \right]}{\text{Var}_{MC^{(0)}} \left[ \widehat{\mathbf{Z}}_{1;n_1, \dots, n_p}^{MC^{(0)}}(s_0) - Z_1(s_0) \right]} \rightarrow 1. \tag{19}$$

2. As  $n_1, \dots, n_p \rightarrow \infty$

$$\frac{\text{Var}_{\mathcal{MC}^{(1)}} [\widehat{Z}_{1;n_1, \dots, n_p}^{\mathcal{MC}^{(1)}}(\mathbf{s}_0) - Z_1(\mathbf{s}_0)]}{\text{Var}_{\mathcal{MC}^{(0)}} [\widehat{Z}_{1;n_1, \dots, n_p}^{\mathcal{MC}^{(1)}}(\mathbf{s}_0) - Z_1(\mathbf{s}_0)]} \rightarrow 1. \tag{20}$$

Then we can apply Theorem 6 using the results on the equivalence of Gaussian measures given in Section 4 between two multivariate Matérn models, two multivariate Generalized Wendland models and between a multivariate Matérn and a multivariate Generalized Wendland model.

The last is probably the most interesting case. With this goal in mind, we consider the cases  $\mathcal{MC}^{(0)} = \mathcal{MM}_{\theta^{(0)}}$  and  $\mathcal{MC}^{(1)} = \mathcal{MW}_{\lambda^{(1)}}$  defined in (6) and (7).

**Theorem 7.** For given  $\nu \geq 1/2$  and  $\kappa \geq 0$ , consider  $P(\mathcal{MM}_{\theta^{(0)}})$  and  $P(\mathcal{MW}_{\lambda^{(1)}})$ . Let for simplicity  $\alpha_{ij}^{(0)} = \alpha_{ij}$  and  $\beta_{ij}^{(1)} = \beta_{ij}$  for  $i, j = 1, \dots, p$ . Assume that  $\nu = \kappa + 1/2$ ,  $\mu > d + 1/2 + \kappa$  and that (17) holds. For  $i = 1, \dots, p$ , let  $\{s_{i,1}, \dots, s_{i,n_i}\}$  be dense in  $D$  as  $n_i \rightarrow \infty$ . Then for  $d = 1, 2, 3$ :

1. As  $n_1, \dots, n_p \rightarrow \infty$

$$\frac{\text{Var}_{\mathcal{MM}_{\theta^{(0)}}} [\widehat{Z}_{1;n_1, \dots, n_p}^{\mathcal{MW}_{\lambda^{(1)}}}(\mathbf{s}_0) - Z_1(\mathbf{s}_0)]}{\text{Var}_{\mathcal{MM}_{\theta^{(0)}}} [\widehat{Z}_{1;n_1, \dots, n_p}^{\mathcal{MM}_{\theta^{(0)}}}(\mathbf{s}_0) - Z_1(\mathbf{s}_0)]} \rightarrow 1. \tag{21}$$

2. As  $n_1, \dots, n_p \rightarrow \infty$

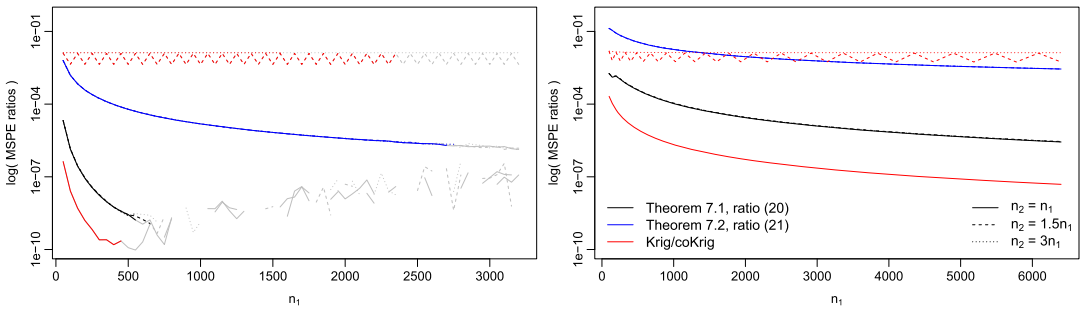
$$\frac{\text{Var}_{\mathcal{MW}_{\lambda^{(1)}}} [\widehat{Z}_{1;n_1, \dots, n_p}^{\mathcal{MW}_{\lambda^{(1)}}}(\mathbf{s}_0) - Z_1(\mathbf{s}_0)]}{\text{Var}_{\mathcal{MM}_{\theta^{(0)}}} [\widehat{Z}_{1;n_1, \dots, n_p}^{\mathcal{MW}_{\lambda^{(1)}}}(\mathbf{s}_0) - Z_1(\mathbf{s}_0)]} \rightarrow 1. \tag{22}$$

An important implication of (21) and (22) is that, if  $\nu = \kappa + 1/2$ ,  $\mu > d + 1/2 + \kappa$ , and under condition (17), asymptotic cokriging prediction efficiency and asymptotically correct estimates of error variance are achieved using a multivariate Generalized Wendland model when the true model is multivariate Matérn. This result has important practical implications, since the Generalized Wendland matrix covariance functions, unlike the Matérn ones, are compactly supported. Hence, using a Generalized Wendland model provides important computational benefits, by enabling to exploit sparse matrix structures (Bevilacqua *et al.*, 2019, Furrer, Genton and Nychka, 2006, Kaufman, Schervish and Nychka, 2008), with a typically negligible loss of statistical accuracy if the true matrix covariance function is in the Matérn class.

To conclude Section 5, remark that the asymptotic properties of predictions from fixed covariance models are considered. The case of sample-size dependent or even estimated covariance models (for instance an estimated smoothness parameter  $\nu$  for the Matérn model) would be technically very different, and has been much less explored, even in the univariate case (Putter and Young, 2001).

## 6. Numerical illustration

In this section, illustrating Section 5, we present a numerical illustration of the rates of convergence of the ratios (21) and (22). We consider the bivariate case, with  $p = 2$ . The mean square prediction error (MSPE) for kriging and cokriging can be interpreted as a statistic of the observation locations in relation to the prediction location, i.e., the MSPE essentially depends on the distance to the nearest



**Figure 1.** Log of ratios (21), (22) and log of MSPE of kriging versus cokriging as a function of  $n_1$  in one (left) and two (right) dimensions. Gray lines indicate numerical instabilities.

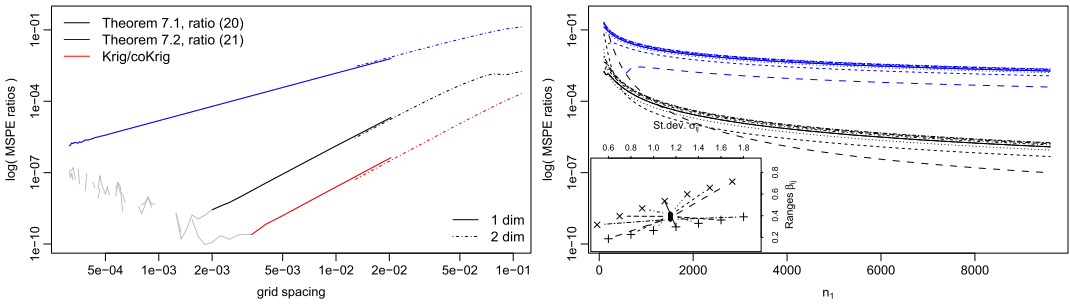
observation location(s). Thus we work with regular grids specified as follows. In one dimension, the observation locations for the primary variable (component of the multivariate random field) are  $(k - 1)/(n_1 - 1)$ ,  $k = 1, \dots, n_1$ , with  $n_1$  even. For the secondary variable we select  $(\ell - 1)/(n_2 - 1)$  with  $n_2 = n_1, 1.5n_1, 3n_1$ ,  $\ell = 1, \dots, n_2$ . In two dimensions we take  $n_1 = n_x^2$  observation locations at  $((k - 1)/(n_x - 1), (k' - 1)/(n_x - 1))$ , with  $n_x$  even,  $k, k' = 1, \dots, n_x$ . For the secondary variable we select a similar grid with  $n_2 = n_x^2, (1.5n_x)^2, (3n_x)^2$ . Prediction for the first variable is at the center of the domain, i.e.,  $s_0 = 0.5$  and  $s_0 = (0.5, 0.5)$ , respectively.

We consider a bivariate Matérn model with  $\theta^{(0)} = (\sigma_{11}^{(0)}, \sigma_{22}^{(0)}, \rho_{12}^{(0)}, \nu, \alpha_{11}, \alpha_{22}, \alpha_{12})^\top = (1.2, 1.1, 0.2, 3/2, 0.05, 0.09, 0.07)^\top$ . In the first illustration we keep the same marginal variances and the same correlation parameter for the bivariate Generalized Wendland model with  $\kappa = 1$  and  $\mu = 5$ . The range parameters are chosen according to the equivalence condition (17) and yield for  $\kappa = 1$  the parameter vector  $\lambda^{(1)} = (\sigma_{11}^{(1)}, \sigma_{22}^{(1)}, \rho_{12}^{(1)}, \mu, \kappa, \beta_{11}, \beta_{22}, \beta_{12})^\top = (1.2, 1.1, 0.2, 5, 1, 0.297, 0.535, 0.416)^\top$ .

Figure 1 illustrates the ratios (21), (22) and the ratio of MSPE of kriging (prediction of  $Z_1(s_0)$  based on  $Z_1$  only) versus cokriging in one and two dimensions. The convergence of the ratios is fast, and numerical instabilities are observed in one dimension for quite small  $n_1$ . Except for the kriging/cokriging ratio, increasing the number of location points for the secondary variable has only a very minor effect and can hardly be distinguished visually. The saw-tooth shape of the dashed red line is due to the alternating even/odd number of observations  $n_2$ . For a fixed  $n_1$ , there is of course a nonlinear relation between the ratios and where we exactly place the point to predict within the observed grid. In one dimension, the log-ratio can be reduced by roughly a factor of two if we move the prediction location from 0.5 towards the nearest right observation location  $n_1/(2n_1 - 2)$ . The left panel of Figure 2 illustrates the log ratios as a function of the grid spacing and emphasizes again that the MSPE is essentially driven by the locations of the nearby observations. The convergence rate for (21) is slightly higher compared to (22) but equivalent to the ratio kriging versus cokriging in case of  $n_1 = n_2$ .

To study the effect of different ranges we modify the variance parameters of the Generalized Wendland by  $(\sigma_{11}^{(1)} + \delta, \sigma_{22}^{(1)} - \delta)$ ,  $\delta = -0.6, -0.4, \dots, 0.6$ . The range parameters are updated according to (17), leading to shorter ranges for smaller standard deviations. The right panel of Figure 2 shows that the ratios are quite stable with respect to different ranges. Increasing the range parameter of the secondary variable reduces the ratio. Hence, it is possible to choose a range parameter  $\beta_{ij}$  tailored to available computing and memory amount with a bearable cost in terms of MSPE.

Note that for other values of the smoothness parameter  $\nu$  the rates themselves change but the conclusions remain the same. Similarly, scaling the ranges of the covariance parameters of the Matérn model



**Figure 2.** Left panel: Log of ratios (21), (22) and log of MSPE of kriging versus cokriging in one (solid) and two dimensions (dash-dotted) as a function of grid spacing. Transparent lines indicate numerical instabilities. Right panel: Log of ratios (21), (22) for different parameter settings of the bivariate Wendland model. The inset figure shows the different configurations with varying  $\sigma_{11}$  (+) and  $\sigma_{22}$  (x) and induced range parameters  $\beta_{11}$ ,  $\beta_{12}$  and  $\beta_{22}$ . ( $\beta_{12}$  is plotted at fixed x-axis value 1.15.) In both panels  $n_2 = n_1$ .

has no effect on the asymptotic results as the scaling is essentially equivalent to adapting the number of observation points.

### 7. Proofs for Section 3

**Sketch of proof for Theorem 1.** We sketch the proof prior to a formal exposition.

- The crux of the proof is to prove that identity (37) holds, with the operators  $\mathcal{B}^{(0)}$  and  $\mathcal{B}^{(1)}$  as being defined in (34).
- To do so, we start by considering the operator  $V$  in (23). We show properties of  $V$  that allow to decompose it into a sequence of eigenfunctions. Using then approximation arguments, we can obtain the relation (29), that relates the matrix spectral densities of the two covariance models.
- We then define the operator  $A$  through (30),  $\Delta := I - A^*A$ , and show that (32) is true.
- The difference in the right-hand side of (32) is then shown to depend on  $\mathcal{B}^{(0)}$  and  $\mathcal{B}^{(1)}$ , which eventually enables us to obtain (37). □

**Proof of Theorem 1.** Let  $\mu \in \mathbb{R}^d$ . We consider the integral operator  $V$  on  $\mathcal{W}_D(F^{(0)})$  defined by

$$(Vf)(\mu) = \int_{\mathbb{R}^d} \mathbf{B}(\mu, \lambda) F^{(0)}(\lambda) f(\lambda) d\lambda. \tag{23}$$

We note that (13) in concert with Cauchy–Schwarz inequality imply

$$\int_{\mathbb{R}^d} \|\mathbf{B}(\mu, \lambda)\| \|F^{(0)}(\lambda)\| \|f(\lambda)\| d\lambda \leq \sqrt{\int_{\mathbb{R}^d} \|\mathbf{B}(\mu, \lambda)\|^2 \|F^{(0)}(\lambda)\| d\lambda} \sqrt{\int_{\mathbb{R}^d} \|F^{(0)}(\lambda)\| \|f(\lambda)\|^2 d\lambda}. \tag{24}$$

The first integral in (24) is finite for almost all  $\mu \in \mathbb{R}^d$  from (13). The second integral in (24) is smaller than a constant times  $\int_{\mathbb{R}^d} \tilde{f}(\lambda)^\top F^{(0)}(\lambda) f(\lambda) d\lambda < \infty$  from Condition 1. Hence,  $Vf$  is well-defined as a function from  $\mathbb{R}^d$  to  $\mathbb{C}^p$ .

Let us now check that  $Vf$  belongs to  $\mathcal{W}_D(\mathbf{F}^{(0)})$  when  $f$  belongs to  $\mathcal{W}_D(\mathbf{F}^{(0)})$ . We use  $f \in \mathcal{W}_D(\mathbf{F}^{(0)})$  and repeatedly apply Cauchy-Schwarz inequality, so that, for a finite constant  $c$ ,

$$\begin{aligned} \|Vf\|_{\mathcal{W}_D(\mathbf{F}^{(0)})}^2 &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \bar{f}(\lambda)^\top \mathbf{F}^{(0)}(\lambda) \bar{\mathbf{B}}(\lambda, t) \mathbf{F}^{(0)}(t) \mathbf{B}(t, \mu) \mathbf{F}^{(0)}(\mu) f(\mu) dt d\lambda d\mu \\ &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \|f(\lambda)\| \|\mathbf{B}(\lambda, t)\| \|\mathbf{F}^{(0)}(\lambda)\| \|\mathbf{F}^{(0)}(t)\| \|\mathbf{B}(\mu, t)\| \\ &\quad \|\mathbf{F}^{(0)}(\mu)\| \|f(\mu)\| dt d\lambda d\mu \\ &= \int_{\mathbb{R}^d} dt \|\mathbf{F}^{(0)}(t)\| \left( \int_{\mathbb{R}^d} d\lambda \|\mathbf{F}^{(0)}(\lambda)\| \|f(\lambda)\| \|\mathbf{B}(\lambda, t)\| \right) \\ &\quad \left( \int_{\mathbb{R}^d} d\mu \|\mathbf{F}^{(0)}(\mu)\| \|f(\mu)\| \|\mathbf{B}(\mu, t)\| \right) \\ &\leq \int_{\mathbb{R}^d} dt \|\mathbf{F}^{(0)}(t)\| \sqrt{\left( \int_{\mathbb{R}^d} d\lambda \|\mathbf{F}^{(0)}(\lambda)\| \|f(\lambda)\|^2 \right) \left( \int_{\mathbb{R}^d} d\lambda \|\mathbf{F}^{(0)}(\lambda)\| \|\mathbf{B}(\lambda, t)\|^2 \right)} \\ &\quad \sqrt{\left( \int_{\mathbb{R}^d} d\mu \|\mathbf{F}^{(0)}(\mu)\| \|f(\mu)\|^2 \right) \left( \int_{\mathbb{R}^d} d\mu \|\mathbf{F}^{(0)}(\mu)\| \|\mathbf{B}(\mu, t)\|^2 \right)} \\ &\leq c \left( \int_{\mathbb{R}^d} d\lambda \bar{f}(\lambda)^\top \mathbf{F}^{(0)}(\lambda) f(\lambda) \right) \int_{\mathbb{R}^d} dt \|\mathbf{F}^{(0)}(t)\| \\ &\quad \sqrt{\int_{\mathbb{R}^d} d\lambda \|\mathbf{F}^{(0)}(\lambda)\| \|\mathbf{B}(\lambda, t)\|^2} \sqrt{\int_{\mathbb{R}^d} d\mu \|\mathbf{F}^{(0)}(\mu)\| \|\mathbf{B}(\mu, t)\|^2} \\ &= c \|f\|_{\mathcal{W}_D(\mathbf{F}^{(0)})}^2 \int_{\mathbb{R}^d} dt \|\mathbf{F}^{(0)}(t)\| \int_{\mathbb{R}^d} d\lambda \|\mathbf{F}^{(0)}(\lambda)\| \|\mathbf{B}(\lambda, t)\|^2 < +\infty. \end{aligned}$$

In the strict inequality below, we have used (13). In the second to last “ $\leq$ ” we have used Condition 1.

Hence  $V$  maps  $\mathcal{W}_D(\mathbf{F}^{(0)})$  to  $\mathcal{W}_D(\mathbf{F}^{(0)})$ . Let us check that  $V$  is Hermitian. For any  $f_1, f_2 \in \mathcal{W}_D(\mathbf{F}^{(0)})$ , we have

$$\begin{aligned} (Vf_1, f_2)_{\mathcal{W}_D(\mathbf{F}^{(0)})} &= \int_{\mathbb{R}^d} d\lambda \overline{\left( \int_{\mathbb{R}^d} d\mu \mathbf{B}(\lambda, \mu) \mathbf{F}^{(0)}(\mu) f_1(\mu) \right)^\top} \mathbf{F}^{(0)}(\lambda) f_2(\lambda) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} d\lambda d\mu \bar{f}_1(\mu)^\top \mathbf{F}^{(0)}(\mu) \bar{\mathbf{B}}(\lambda, \mu)^\top \mathbf{F}^{(0)}(\lambda) f_2(\lambda) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} d\lambda d\mu \bar{f}_1(\mu)^\top \mathbf{F}^{(0)}(\mu) \mathbf{B}(\mu, \lambda) \mathbf{F}^{(0)}(\lambda) f_2(\lambda) \\ &= \int_{\mathbb{R}^d} d\mu \bar{f}_1(\mu)^\top \mathbf{F}^{(0)}(\mu) \left( \int_{\mathbb{R}^d} d\lambda \mathbf{B}(\mu, \lambda) \mathbf{F}^{(0)}(\lambda) f_2(\lambda) \right) \\ &= (f_1, Vf_2)_{\mathcal{W}_D(\mathbf{F}^{(0)})}, \end{aligned}$$

where we have used  $\bar{\mathbf{B}}(\lambda, \mu)^\top = \mathbf{B}(\mu, \lambda)$ .

Let us now show that  $V$  is an Hilbert–Schmidt operator. Let  $(\phi_k)_{k \in \mathbb{N}}$  be an orthonormal basis of  $\mathcal{W}_D(\mathbf{F}^{(0)})$ . Let, for  $n \in \mathbb{N}$ ,  $\lambda, \mu \in \mathbb{R}^d$ ,

$$\overline{\mathbf{B}_n(\lambda, \mu)} = \sum_{a=1}^n \left( \int_{\mathbb{R}^d} \mathbf{B}(\lambda, z) \mathbf{F}^{(0)}(z) \phi_a(z) dz \right) \phi_a(\mu)^\top.$$

Let also  $\overline{\mathbf{E}_n(\lambda, \mu)} = \overline{\mathbf{B}_n(\lambda, \mu)} - \mathbf{B}(\lambda, \mu)$ . Then, we have

$$\text{tr} \left( \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (\mathbf{F}^{(0)})^{1/2}(\lambda) \overline{\mathbf{E}_n(\lambda, \mu)} \mathbf{F}^{(0)}(\mu) \mathbf{E}_n(\lambda, \mu)^\top (\mathbf{F}^{(0)})^{1/2}(\lambda) d\lambda d\mu \right) \tag{25}$$

$$\begin{aligned} &\leq p \int_{\mathbb{R}^d} d\lambda \|\mathbf{F}^{(0)}(\lambda)\| \int_{\mathbb{R}^d} d\mu \left\| \overline{\mathbf{E}_n(\lambda, \mu)} \mathbf{F}^{(0)}(\mu) \mathbf{E}_n(\lambda, \mu)^\top \right\| \\ &\leq p \sum_{a=1}^p \int_{\mathbb{R}^d} d\lambda \|\mathbf{F}^{(0)}(\lambda)\| \int_{\mathbb{R}^d} d\mu \left( e_a^\top \overline{\mathbf{E}_n(\lambda, \mu)} \right) \mathbf{F}^{(0)}(\mu) \left( e_a^\top \mathbf{E}_n(\lambda, \mu) \right)^\top. \end{aligned} \tag{26}$$

We note that  $e_a^\top \overline{\mathbf{B}_n(\lambda, \mu)}$  is the orthogonal projection in  $\mathcal{W}_D(\mathbf{F}^{(0)})$  of the row  $a$  of  $\mu \mapsto \overline{\mathbf{B}(\lambda, \mu)}$  on the linear space spanned by  $\phi_1^\top, \dots, \phi_n^\top$ . For almost all  $\lambda \in \mathbb{R}^d$ , the norm of this row in  $\mathcal{W}_D(\mathbf{F}^{(0)})$  is finite from (13). Hence, for almost all  $\lambda \in \mathbb{R}^d$ , the inner most integral in (26) goes to zero as  $n \rightarrow \infty$ . Furthermore, this inner most integral is bounded by

$$\int_{\mathbb{R}^d} d\mu \left( e_a^\top \overline{\mathbf{B}(\lambda, \mu)} \right) \mathbf{F}^{(0)}(\mu) \left( e_a^\top \mathbf{B}(\lambda, \mu) \right)^\top,$$

which satisfies

$$\int_{\mathbb{R}^d} d\lambda \|\mathbf{F}^{(0)}(\lambda)\| \int_{\mathbb{R}^d} d\mu \left( e_a^\top \overline{\mathbf{B}(\lambda, \mu)} \right) \mathbf{F}^{(0)}(\mu) \left( e_a^\top \mathbf{B}(\lambda, \mu) \right)^\top < \infty$$

from (13). Hence, by the dominated convergence theorem, (25) goes to zero as  $n \rightarrow \infty$ . Now, consider the application

$$\mathbf{A}, \mathbf{C} \mapsto \text{tr} \left( \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (\mathbf{F}^{(0)})^{1/2}(\lambda) \mathbf{A}(\lambda, \mu) \mathbf{F}^{(0)}(\mu) \overline{\mathbf{C}(\lambda, \mu)}^\top (\mathbf{F}^{(0)})^{1/2}(\lambda) d\lambda d\mu \right),$$

for functions  $\mathbf{A}, \mathbf{C}$  from  $\mathbb{R}^d \times \mathbb{R}^d$  to  $\mathbb{C}^{p^2}$  satisfying (13) with  $\mathbf{B}$  replaced by  $\mathbf{A}$  or  $\mathbf{C}$  there. One can check that this application is a scalar product. Hence, using the triangle inequality, it follows that

$$\text{tr} \left( \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (\mathbf{F}^{(0)})^{1/2}(\lambda) \overline{\mathbf{B}_n(\lambda, \mu)} \mathbf{F}^{(0)}(\mu) \mathbf{B}_n(\lambda, \mu)^\top (\mathbf{F}^{(0)})^{1/2}(\lambda) d\lambda d\mu \right)$$

is bounded as  $n \rightarrow \infty$ . This bounded quantity is equal to, using the orthogonality of  $\phi_1, \dots, \phi_n$ ,

$$\begin{aligned} &\sum_{i,j=1}^n \text{tr} \left( \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (\mathbf{F}^{(0)})^{1/2}(\lambda) \left( \int_{\mathbb{R}^d} \mathbf{B}(\lambda, z) \mathbf{F}^{(0)}(z) \phi_i(z) dz \right) \phi_i(\mu)^\top \mathbf{F}^{(0)}(\mu) \right. \\ &\quad \left. \overline{\left( \int_{\mathbb{R}^d} \mathbf{B}(\lambda, z) \mathbf{F}^{(0)}(z) \phi_j(z) dz \right)^\top} (\mathbf{F}^{(0)})^{1/2}(\lambda) d\lambda d\mu \right) \end{aligned}$$



$$\begin{aligned}
 &= \sum_{i=1}^n \operatorname{tr} \left( \int_{\mathbb{R}^d} (\mathbf{F}^{(0)})^{1/2}(\lambda) \left( \int_{\mathbb{R}^d} \mathbf{B}(\lambda, z) \mathbf{F}^{(0)}(z) \boldsymbol{\phi}_i(z) dz \right) \right. \\
 &\quad \left. \overline{\left( \int_{\mathbb{R}^d} \mathbf{B}(\lambda, z) \mathbf{F}^{(0)}(z) \boldsymbol{\phi}_i(z) dz \right)}^\top (\mathbf{F}^{(0)})^{1/2}(\lambda) d\lambda \right) \\
 &= \sum_{i=1}^n \operatorname{tr} \left( \int_{\mathbb{R}^d} \overline{\left( \int_{\mathbb{R}^d} \mathbf{B}(\lambda, z) \mathbf{F}^{(0)}(z) \boldsymbol{\phi}_i(z) dz \right)}^\top (\mathbf{F}^{(0)})^{1/2}(\lambda) \right. \\
 &\quad \left. (\mathbf{F}^{(0)})^{1/2}(\lambda) \left( \int_{\mathbb{R}^d} \mathbf{B}(\lambda, z) \mathbf{F}^{(0)}(z) \boldsymbol{\phi}_i(z) dz \right) d\lambda \right) \\
 &= \sum_{i=1}^n \int_{\mathbb{R}^d} \overline{\left( \int_{\mathbb{R}^d} \mathbf{B}(\lambda, z) \mathbf{F}^{(0)}(z) \boldsymbol{\phi}_i(z) dz \right)}^\top \mathbf{F}^{(0)}(\lambda) \left( \int_{\mathbb{R}^d} \mathbf{B}(\lambda, z) \mathbf{F}^{(0)}(z) \boldsymbol{\phi}_i(z) dz \right) d\lambda \\
 &= \sum_{i=1}^n \|\mathbf{V} \boldsymbol{\phi}_i\|_{\mathcal{W}_D(\mathbf{F}^{(0)})}^2.
 \end{aligned}$$

This implies that  $\sum_{i=1}^\infty \|\mathbf{V} \boldsymbol{\phi}_i\|_{\mathcal{W}_D(\mathbf{F}^{(0)})}^2 < \infty$  and thus  $V$  is Hilbert–Schmidt.

Hence, there exists a sequence  $(\mathbf{g}_k)_{k \in \mathbb{N}}$  of eigenfunctions of  $V$ . For  $k \in \mathbb{N}$ , we let  $\mathbf{g}_k = (g_{k,1}, \dots, g_{k,p})^\top$  from  $\mathbb{R}^d$  to  $\mathbb{C}^p$  and we remark that we have  $(\mathbf{g}_k, \mathbf{g}_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} = \delta_{k,j}$  for  $k, j \in \mathbb{N}$ . We let  $(\lambda_k)_{k \in \mathbb{N}}$  be the corresponding sequence of eigenvalues of  $V$ , such that we have  $V \mathbf{g}_k = \lambda_k \mathbf{g}_k$  for  $k \in \mathbb{N}$  and  $\sum_{k=1}^\infty \lambda_k^2 < \infty$ .

Let  $k, j \in \mathbb{N}$  be fixed. By definition of  $\mathcal{W}_D(\mathbf{F}^{(0)})$ , there exists a sequence  $(\boldsymbol{\phi}_{k,n})_{n \in \mathbb{N}}$  such that  $\boldsymbol{\phi}_{k,n} : D \rightarrow \mathbb{C}^p$  for  $n \in \mathbb{N}$  and such that, with  $\mathbf{u}_{k,n} = (u_{k,n,1}, \dots, u_{k,n,p})^\top$  from  $\mathbb{R}^d$  to  $\mathbb{C}^p$  defined by

$$u_{k,n,i}(\lambda) = \int_D e^{-\lambda^\top t} \boldsymbol{\phi}_{k,n,i}(t) dt,$$

for  $i = 1, \dots, p$  and  $\lambda \in \mathbb{R}^d$ , we have  $\mathbf{u}_{k,n} \rightarrow \mathbf{g}_k$  in  $\mathcal{W}_D(\mathbf{F}^{(0)})$ . There also exists a sequence  $(\boldsymbol{\phi}_{j,n})_{n \in \mathbb{N}}$  that is defined similarly for  $\mathbf{g}_j$  instead of  $\mathbf{g}_k$ .

We have, using (14),

$$\int_D \int_D \bar{\boldsymbol{\phi}}_{k,n}(t)^\top \mathbf{R}^{(1)}(t-s) \boldsymbol{\phi}_{j,n}(s) dt ds - \int_D \int_D \bar{\boldsymbol{\phi}}_{k,n}(t)^\top \mathbf{R}^{(0)}(t-s) \boldsymbol{\phi}_{j,n}(s) dt ds \tag{27}$$

$$\begin{aligned}
 &= \int_D \int_D \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-\lambda^\top t + \mu^\top s} \bar{\boldsymbol{\phi}}_{k,n}(t)^\top \mathbf{F}^{(0)}(\lambda) \mathbf{B}(\lambda, \mu) \mathbf{F}^{(0)}(\mu) \boldsymbol{\phi}_{j,n}(s) dt ds d\lambda d\mu \\
 &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \bar{\mathbf{u}}_{k,n}(\lambda)^\top \mathbf{F}^{(0)}(\lambda) \mathbf{B}(\lambda, \mu) \mathbf{F}^{(0)}(\mu) \mathbf{u}_{j,n}(\mu) d\lambda d\mu.
 \end{aligned} \tag{28}$$

Let us find the limit of the two terms in (27) as  $n \rightarrow \infty$ . We have

$$\begin{aligned}
 \int_D \int_D \bar{\boldsymbol{\phi}}_{k,n}(t)^\top \mathbf{R}^{(1)}(t-s) \boldsymbol{\phi}_{j,n}(s) dt ds &= \int_D \int_D \int_{\mathbb{R}^d} e^{i\lambda^\top (t-s)} \bar{\boldsymbol{\phi}}_{k,n}(t)^\top \mathbf{F}^{(1)}(\lambda) \boldsymbol{\phi}_{j,n}(s) dt ds d\lambda \\
 &= \int_{\mathbb{R}^d} \bar{\mathbf{u}}_{k,n}(\lambda)^\top \mathbf{F}^{(1)}(\lambda) \mathbf{u}_{j,n}(\lambda) d\lambda
 \end{aligned}$$

$$\xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^d} \bar{\mathbf{g}}_k(\boldsymbol{\lambda})^\top \mathbf{F}^{(1)}(\boldsymbol{\lambda}) \mathbf{g}_j(\boldsymbol{\lambda}) d\boldsymbol{\lambda}$$

using Lemma 1 and the triangle inequality. Similarly we have

$$\int_D \int_D \bar{\boldsymbol{\phi}}_{k,n}(\mathbf{t})^\top \mathbf{R}^{(0)}(\mathbf{t} - \mathbf{s}) \boldsymbol{\phi}_{j,n}(\mathbf{s}) dt ds \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^d} \bar{\mathbf{g}}_k(\boldsymbol{\lambda})^\top \mathbf{F}^{(0)}(\boldsymbol{\lambda}) \mathbf{g}_j(\boldsymbol{\lambda}) d\boldsymbol{\lambda}.$$

Let us find the limit of (28) as  $n \rightarrow \infty$ . We have, with a finite constant  $c$ ,

$$\begin{aligned} & \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \bar{\mathbf{u}}_{k,n}(\boldsymbol{\lambda})^\top \mathbf{F}^{(0)}(\boldsymbol{\lambda}) \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \mathbf{F}^{(0)}(\boldsymbol{\mu}) \mathbf{u}_{j,n}(\boldsymbol{\mu}) d\boldsymbol{\lambda} d\boldsymbol{\mu} \\ & - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \bar{\mathbf{g}}_k(\boldsymbol{\lambda})^\top \mathbf{F}^{(0)}(\boldsymbol{\lambda}) \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \mathbf{F}^{(0)}(\boldsymbol{\mu}) \mathbf{u}_{j,n}(\boldsymbol{\mu}) d\boldsymbol{\lambda} d\boldsymbol{\mu} \\ & = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [\bar{\mathbf{u}}_{k,n}(\boldsymbol{\lambda}) - \bar{\mathbf{g}}_k(\boldsymbol{\lambda})]^\top \mathbf{F}^{(0)}(\boldsymbol{\lambda}) \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \mathbf{F}^{(0)}(\boldsymbol{\mu}) \mathbf{u}_{j,n}(\boldsymbol{\mu}) d\boldsymbol{\lambda} d\boldsymbol{\mu} \\ & \leq c \sqrt{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \|\mathbf{F}^{(0)}(\boldsymbol{\lambda})\| \|\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\mu})\|^2 \|\mathbf{F}^{(0)}(\boldsymbol{\mu})\| d\boldsymbol{\lambda} d\boldsymbol{\mu}} \\ & \quad \sqrt{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \|\mathbf{F}^{(0)}(\boldsymbol{\lambda})\| \|\mathbf{u}_{k,n}(\boldsymbol{\lambda}) - \mathbf{g}_k(\boldsymbol{\lambda})\|^2 \|\mathbf{F}^{(0)}(\boldsymbol{\mu})\| \|\mathbf{u}_{j,n}(\boldsymbol{\mu})\|^2 d\boldsymbol{\mu} d\boldsymbol{\lambda}} \\ & = c \sqrt{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \|\mathbf{F}^{(0)}(\boldsymbol{\lambda})\| \|\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\mu})\|^2 \|\mathbf{F}^{(0)}(\boldsymbol{\mu})\| d\boldsymbol{\lambda} d\boldsymbol{\mu}} \sqrt{\int_{\mathbb{R}^d} \|\mathbf{F}^{(0)}(\boldsymbol{\lambda})\| \|\mathbf{u}_{k,n}(\boldsymbol{\lambda}) - \mathbf{g}_k(\boldsymbol{\lambda})\|^2 d\boldsymbol{\lambda}} \\ & \quad \sqrt{\int_{\mathbb{R}^d} \|\mathbf{F}^{(0)}(\boldsymbol{\mu})\| \|\mathbf{u}_{j,n}(\boldsymbol{\mu})\|^2 d\boldsymbol{\mu}} \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

where we have used the Cauchy–Schwarz inequality and the fact that  $\mathbf{u}_{k,n}$  converges to  $\mathbf{g}_k$  in  $\mathcal{W}_D(\mathbf{F}^{(0)})$ , together with Condition 1 and (13).

We show similarly

$$\begin{aligned} & \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \bar{\mathbf{g}}_k(\boldsymbol{\lambda})^\top \mathbf{F}^{(0)}(\boldsymbol{\lambda}) \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \mathbf{F}^{(0)}(\boldsymbol{\mu}) \mathbf{u}_{j,n}(\boldsymbol{\mu}) d\boldsymbol{\lambda} d\boldsymbol{\mu} \\ & - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \bar{\mathbf{g}}_k(\boldsymbol{\lambda})^\top \mathbf{F}^{(0)}(\boldsymbol{\lambda}) \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \mathbf{F}^{(0)}(\boldsymbol{\mu}) \mathbf{g}_j(\boldsymbol{\mu}) d\boldsymbol{\lambda} d\boldsymbol{\mu} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Hence, (28) converges to

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \bar{\mathbf{g}}_k(\boldsymbol{\lambda})^\top \mathbf{F}^{(0)}(\boldsymbol{\lambda}) \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \mathbf{F}^{(0)}(\boldsymbol{\mu}) \mathbf{g}_j(\boldsymbol{\mu}) d\boldsymbol{\lambda} d\boldsymbol{\mu}$$

as  $n \rightarrow \infty$ . Thus, from (27) and (28), we obtain

$$(\mathbf{g}_k, \mathbf{g}_j)_{\mathcal{W}_D(\mathbf{F}^{(1)})} - (\mathbf{g}_k, \mathbf{g}_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} = \lambda_k \delta_{jk}. \tag{29}$$

Let  $A$  be the operator from  $\mathcal{W}_D(\mathbf{F}^{(0)})$  to  $\mathcal{W}_D(\mathbf{F}^{(1)})$  defined by

$$A\phi = \phi, \quad \phi \in \mathcal{W}_D(\mathbf{F}^{(0)}). \tag{30}$$

From Lemma 1,  $A$  is well-defined and bounded. Let  $A^*$  from  $\mathcal{W}_D(\mathbf{F}^{(1)})$  to  $\mathcal{W}_D(\mathbf{F}^{(0)})$  be the adjoint operator to  $A$ . We remark that we have for  $\phi_1, \phi_2 \in \mathcal{W}_D(\mathbf{F}^{(0)})$ ,

$$(A^*A\phi_1, \phi_2)_{\mathcal{W}_D(\mathbf{F}^{(0)})} = (A\phi_1, A\phi_2)_{\mathcal{W}_D(\mathbf{F}^{(1)})} = (\phi_1, \phi_2)_{\mathcal{W}_D(\mathbf{F}^{(1)})}. \tag{31}$$

Consider the operator  $\Delta = I - A^*A$  from  $\mathcal{W}_D(\mathbf{F}^{(0)})$  to  $\mathcal{W}_D(\mathbf{F}^{(0)})$  where  $I$  is the identity operator. Then from (31) we have, for  $\phi_1, \phi_2 \in \mathcal{W}_D(\mathbf{F}^{(0)})$ ,

$$(\Delta\phi_1, \phi_2)_{\mathcal{W}_D(\mathbf{F}^{(0)})} = (\phi_1, \phi_2)_{\mathcal{W}_D(\mathbf{F}^{(0)})} - (\phi_1, \phi_2)_{\mathcal{W}_D(\mathbf{F}^{(1)})}. \tag{32}$$

Let now  $\phi_1, \dots, \phi_n$  be any orthonormal functions in  $\mathcal{W}_D(\mathbf{F}^{(0)})$ . From (29) and (32), we have, using Bessel’s inequality and Parseval’s identity,

$$\begin{aligned} +\infty &> \sum_{k=1}^{\infty} \lambda_k^2 = \sum_{k,j=1}^{\infty} \left( (\Delta\mathbf{g}_k, \mathbf{g}_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} \right)^2 = \sum_{k=1}^{\infty} \left( \|\Delta\mathbf{g}_k\|_{\mathcal{W}_D(\mathbf{F}^{(0)})} \right)^2 \\ &\geq \sum_{k=1}^{\infty} \sum_{j=1}^n \left( (\Delta\mathbf{g}_k, \phi_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} \right)^2 = \sum_{k=1}^{\infty} \sum_{j=1}^n \left( (\mathbf{g}_k, \Delta\phi_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} \right)^2 \\ &= \sum_{j=1}^n \left( \|\Delta\phi_j\|_{\mathcal{W}_D(\mathbf{F}^{(0)})} \right)^2 \geq \sum_{k,j=1}^n \left( (\phi_k, \Delta\phi_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} \right)^2. \end{aligned}$$

Hence, from (32) we have

$$\sum_{k,j=1}^n \left( (\phi_k, \phi_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} - (\phi_k, \phi_j)_{\mathcal{W}_D(\mathbf{F}^{(1)})} \right)^2 \leq \sum_{k=1}^{\infty} \lambda_k^2 < +\infty. \tag{33}$$

Let for  $a = 0, 1$ ,  $\mathcal{B}^{(a)}$  be the operator on  $L_D^{2,p}$  defined by

$$\mathcal{B}^{(a)}(f)(t) = \int_D \mathbf{R}^{(a)}(t - u) f(u) du. \tag{34}$$

Let  $(\mathbf{h}_k)_{k \in \mathbb{N}}$  be the orthonormal basis of  $L_D^{2,p}$  composed of the eigenfunctions of  $\mathcal{B}^{(0)}$ , with eigenvalues  $(\rho_k)_{k \in \mathbb{N}}$  (the existence can be proved as for the proof that  $V$  is Hermitian and Hilbert–Schmidt above). Let, for  $k \in \mathbb{N}$ ,  $\phi_k = (\phi_{k,1}, \dots, \phi_{k,p})$  with  $\phi_{k,i}(\lambda) = \int_D h_{k,i}(t) e^{-i\lambda^\top t} dt$  for  $i = 1, \dots, p$ . Then  $\phi_k \in \mathcal{W}_D(\mathbf{F}^{(0)})$  for  $k \in \mathbb{N}$  and we have, for  $k, j \in \mathbb{N}$ ,

$$\begin{aligned} (\phi_k, \phi_j)_{\mathcal{W}_D(\mathbf{F}^{(1)})} &= \int_D \int_D \int_{\mathbb{R}^d} \bar{\mathbf{h}}_k(t)^\top e^{i\lambda^\top t} \mathbf{F}^{(1)}(\lambda) \mathbf{h}_j(u) e^{-i\lambda^\top u} dt du d\lambda \\ &= \int_D \int_D \bar{\mathbf{h}}_k(t)^\top \mathbf{R}^{(1)}(t - u) \mathbf{h}_j(u). \end{aligned} \tag{35}$$

Similarly,

$$(\boldsymbol{\phi}_k, \boldsymbol{\phi}_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} = \int_D \int_D \bar{\mathbf{h}}_k(\mathbf{t})^\top \mathbf{R}^{(0)}(\mathbf{t} - \mathbf{u}) \mathbf{h}_j(\mathbf{u}). \tag{36}$$

In particular  $(\boldsymbol{\phi}_k, \boldsymbol{\phi}_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} = \delta_{k,j} \rho_k$ . Since the  $(\boldsymbol{\phi}_k)_{k \in \mathbb{N}}$  are not almost surely equal to zero, it follows that  $\rho_k > 0$  for  $k \in \mathbb{N}$  from Condition 1.

Hence, from (33), (35) and (36), we obtain, with  $\tilde{\boldsymbol{\phi}}_k = \boldsymbol{\phi}_k / \rho_k^{1/2}$ ,

$$\begin{aligned} \infty > \sum_{k=1}^{\infty} \lambda_k^2 &\geq \sum_{k,j=1}^n \left( (\tilde{\boldsymbol{\phi}}_k, \tilde{\boldsymbol{\phi}}_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} - (\tilde{\boldsymbol{\phi}}_k, \tilde{\boldsymbol{\phi}}_j)_{\mathcal{W}_D(\mathbf{F}^{(1)})} \right)^2 \\ &= \sum_{k,j=1}^n \left( \frac{1}{\sqrt{\rho_k}} \frac{1}{\sqrt{\rho_j}} (\boldsymbol{\phi}_k, \boldsymbol{\phi}_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} - \frac{1}{\sqrt{\rho_k}} \frac{1}{\sqrt{\rho_j}} (\boldsymbol{\phi}_k, \boldsymbol{\phi}_j)_{\mathcal{W}_D(\mathbf{F}^{(1)})} \right)^2 \\ &= \sum_{k,j=1}^n \left( \frac{1}{\sqrt{\rho_k}} \frac{1}{\sqrt{\rho_j}} (\mathbf{h}_k, \mathcal{B}^{(0)} \mathbf{h}_j)_{L_D^{2,p}} - \frac{1}{\sqrt{\rho_k}} \frac{1}{\sqrt{\rho_j}} (\mathbf{h}_k, \mathcal{B}^{(1)} \mathbf{h}_j)_{L_D^{2,p}} \right)^2. \end{aligned}$$

Let now  $f \in L_D^{2,p}$ . We can write  $f = \sum_{i=1}^{\infty} \alpha_i \mathbf{h}_i$  with  $\sum_{i=1}^{\infty} \alpha_i^2 < \infty$ . If  $(f, \mathcal{B}^{(0)} f)_{L_D^{2,p}} = 1$ , we have  $\sum_{i=1}^{\infty} \alpha_i^2 \rho_i = 1$  and

$$\begin{aligned} (f, \mathcal{B}^{(1)} f)_{L_D^{2,p}} - 1 &\leq \left| (f, \mathcal{B}^{(0)} f)_{L_D^{2,p}} - (f, \mathcal{B}^{(1)} f)_{L_D^{2,p}} \right| \\ &\leq \sum_{i,j=1}^{\infty} |\alpha_i \alpha_j| \left| (\mathbf{h}_i, \mathcal{B}^{(0)} \mathbf{h}_j)_{L_D^{2,p}} - (\mathbf{h}_i, \mathcal{B}^{(1)} \mathbf{h}_j)_{L_D^{2,p}} \right| \\ &= \sum_{i,j=1}^{\infty} |\alpha_i \sqrt{\rho_i} \alpha_j \sqrt{\rho_j}| \left| \frac{1}{\sqrt{\rho_i}} \frac{1}{\sqrt{\rho_j}} (\mathbf{h}_i, \mathcal{B}^{(0)} \mathbf{h}_j)_{L_D^{2,p}} - \frac{1}{\sqrt{\rho_i}} \frac{1}{\sqrt{\rho_j}} (\mathbf{h}_i, \mathcal{B}^{(1)} \mathbf{h}_j)_{L_D^{2,p}} \right| \\ &\leq \sqrt{\sum_{i,j=1}^{\infty} \alpha_i^2 \rho_i \alpha_j^2 \rho_j} \sqrt{\sum_{i,j=1}^{\infty} \left( \frac{1}{\sqrt{\rho_i}} \frac{1}{\sqrt{\rho_j}} (\mathbf{h}_i, \mathcal{B}^{(0)} \mathbf{h}_j)_{L_D^{2,p}} - \frac{1}{\sqrt{\rho_i}} \frac{1}{\sqrt{\rho_j}} (\mathbf{h}_i, \mathcal{B}^{(1)} \mathbf{h}_j)_{L_D^{2,p}} \right)^2} \leq \sum_{k=1}^{\infty} \lambda_k^2. \end{aligned}$$

Hence, with  $(\mathcal{B}^{(a)})^{1/2}$  the unique operator square root of  $\mathcal{B}^{(a)}$  for  $a = 0, 1$ , there exists a finite constant  $c$  such that for any  $f \in L_D^{2,p}$ ,  $((\mathcal{B}^{(1)})^{1/2} f, (\mathcal{B}^{(1)})^{1/2} f)_{L_D^{2,p}} \leq c((\mathcal{B}^{(0)})^{1/2} f, (\mathcal{B}^{(0)})^{1/2} f)_{L_D^{2,p}}$ . Similarly, there exists a finite constant  $c'$  such that for any  $f \in L_D^{2,p}$ ,  $((\mathcal{B}^{(0)})^{1/2} f, (\mathcal{B}^{(0)})^{1/2} f)_{L_D^{2,p}} \leq c'((\mathcal{B}^{(1)})^{1/2} f, (\mathcal{B}^{(1)})^{1/2} f)_{L_D^{2,p}}$ . Hence from Proposition B.1 in Da Prato and Zabczyk (2014), the image spaces of  $(\mathcal{B}^{(0)})^{1/2}$  and  $(\mathcal{B}^{(1)})^{1/2}$  are the same.

Let  $(\mathcal{B}^{(0)})^{-1/2}$  be the pseudo inverse of  $(\mathcal{B}^{(0)})^{1/2}$  (see Da Prato and Zabczyk, 2014). Let also  $\boldsymbol{\psi}_k = (\mathcal{B}^{(0)})^{1/2} \mathbf{h}_k / \rho_k^{1/2}$ . Then  $(\boldsymbol{\psi}_k)_{k \in \mathbb{N}}$  is an orthonormal basis of the image of  $\mathcal{B}^{(0)}$  in  $L_D^{2,p}$ . We obtain, recalling that the  $\tilde{\boldsymbol{\phi}}_k$  for  $k = 1, \dots, p$  are orthonormal in  $\mathcal{W}_D(\mathbf{F}^{(0)})$ , from (33), (35) and (36),

$$\infty > \sum_{k=1}^{\infty} \lambda_k^2 \geq \sum_{k,j=1}^n \left( (\tilde{\boldsymbol{\phi}}_k, \tilde{\boldsymbol{\phi}}_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} - (\tilde{\boldsymbol{\phi}}_k, \tilde{\boldsymbol{\phi}}_j)_{\mathcal{W}_D(\mathbf{F}^{(1)})} \right)^2$$

$$\begin{aligned}
 &= \sum_{k,j=1}^n \left( \frac{1}{\sqrt{\rho_k}} \frac{1}{\sqrt{\rho_j}} (\phi_k, \phi_j)_{\mathcal{W}_D(\mathbf{F}^{(0)})} - \frac{1}{\sqrt{\rho_k}} \frac{1}{\sqrt{\rho_j}} (\phi_k, \phi_j)_{\mathcal{W}_D(\mathbf{F}^{(1)})} \right)^2 \\
 &= \sum_{k,j=1}^n \left( \frac{1}{\sqrt{\rho_k}} \frac{1}{\sqrt{\rho_j}} (\mathbf{h}_k, \mathcal{B}^{(0)} \mathbf{h}_j)_{L_D^{2,p}} - \frac{1}{\sqrt{\rho_k}} \frac{1}{\sqrt{\rho_j}} (\mathbf{h}_k, \mathcal{B}^{(1)} \mathbf{h}_j)_{L_D^{2,p}} \right)^2 \\
 &= \sum_{k,j=1}^n \left( (\psi_k, \psi_j)_{L_D^{2,p}} - ((\mathcal{B}^{(0)})^{-1/2} \psi_k, \mathcal{B}^{(1)} (\mathcal{B}^{(0)})^{-1/2} \psi_j)_{L_D^{2,p}} \right)^2 \\
 &= \sum_{k,j=1}^n \left( (\psi_k, [I - (\mathcal{B}^{(0)})^{-1/2} \mathcal{B}^{(1)} (\mathcal{B}^{(0)})^{-1/2}] \psi_j)_{L_D^{2,p}} \right)^2.
 \end{aligned}$$

Hence, with the orthonormal basis  $(\psi_k)_{k \in \mathbb{N}}$  of the image of  $\mathcal{B}^{(0)}$  in  $L_D^{2,p}$ , we have

$$\sum_{k,j=1}^{\infty} \left( (\psi_k, [I - (\mathcal{B}^{(0)})^{-1/2} \mathcal{B}^{(1)} (\mathcal{B}^{(0)})^{-1/2}] \psi_j)_{L_D^{2,p}} \right)^2 < \infty. \tag{37}$$

Hence,  $[I - (\mathcal{B}^{(0)})^{-1/2} \mathcal{B}^{(1)} (\mathcal{B}^{(0)})^{-1/2}]$  is a Hilbert–Schmidt operator from the image of  $\mathcal{B}^{(0)}$  to  $L_D^{2,p}$ . Since we have seen also that the images of  $(\mathcal{B}^{(0)})^{1/2}$  and  $(\mathcal{B}^{(1)})^{1/2}$  are the same, from Theorem 2.25 in Da Prato and Zabczyk (2014) (see also Chapter 1 of Maniglia and Rhandi, 2004),  $P(\mathbf{R}^{(0)}) \equiv P(\mathbf{R}^{(1)})$  on the paths of  $\mathbf{Z}$ . □

**Proof of Theorem 2.** It is convenient to consider two real-valued stationary  $p$ -variate Gaussian random fields  $\{\mathbf{Z}^{(0)}(s) = (Z_1^{(0)}(s), \dots, Z_p^{(0)}(s))^T, s \in D\}$  and  $\{\mathbf{Z}^{(1)}(s) = (Z_1^{(1)}(s), \dots, Z_p^{(1)}(s))^T, s \in D\}$ , where  $\mathbf{Z}^{(0)}$  and  $\mathbf{Z}^{(1)}$  have continuous sample paths, and where, for  $a = 0, 1$ ,  $\mathbf{Z}^{(a)}$  has zero mean and matrix covariance function  $\mathbf{R}^{(a)} = [R_{ij}^{(a)}]_{i,j=1}^p$ .

Let us first assume that for all  $\lambda \in \mathbb{R}^d$ ,  $\mathbf{F}^{(1)}(\lambda) \geq \mathbf{F}^{(0)}(\lambda)$ . Let, for  $\lambda \in \mathbb{R}^d$ , with  $c_1$  as in Condition 1,

$$\begin{aligned}
 \tilde{\mathbf{F}}^{(0)}(\lambda) &= c_1 \gamma(\lambda)^2 \mathbf{I}_p, \\
 \tilde{\mathbf{F}}^{(1)}(\lambda) &= c_1 \gamma(\lambda)^2 \mathbf{I}_p + \mathbf{F}^{(1)}(\lambda) - \mathbf{F}^{(0)}(\lambda), \quad \text{and} \\
 \tilde{\mathbf{F}}^{(2)}(\lambda) &= \mathbf{F}^{(0)}(\lambda) - c_1 \gamma(\lambda)^2 \mathbf{I}_p.
 \end{aligned}$$

Let  $\tilde{\mathbf{Z}}^{(0)}$ ,  $\tilde{\mathbf{Z}}^{(1)}$  and  $\tilde{\mathbf{Z}}^{(2)}$  be three independent  $p$ -variate Gaussian processes with mean function zero and respective matrix spectral densities  $\tilde{\mathbf{F}}^{(0)}$ ,  $\tilde{\mathbf{F}}^{(1)}$  and  $\tilde{\mathbf{F}}^{(2)}$ . Then, in distribution, for  $a = 0, 1$ ,  $\mathbf{Z}^{(a)} = \tilde{\mathbf{Z}}^{(a)} + \tilde{\mathbf{Z}}^{(2)}$ . Hence, in order to prove the theorem, it is sufficient to show that the Gaussian measures given by  $\tilde{\mathbf{Z}}^{(0)}$  and  $\tilde{\mathbf{Z}}^{(1)}$  are equivalent. Let us do this.

Let  $(\tilde{\mathbf{g}}_k)_{k \in \mathbb{N}}$  be an orthonormal basis in  $\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})$ . Let  $n \in \mathbb{N}$ . Since  $\mathcal{W}_D$  is dense everywhere in  $\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})$ , from Lemma 2, we can find  $\mathbf{g}_1, \dots, \mathbf{g}_n$  in  $\mathcal{W}_D$  for which  $(\mathbf{g}_k, \mathbf{g}_l)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} = \delta_{k,l}$  for  $k, l = 1, \dots, n$  and

$$\sum_{k=1}^n \left[ \|\tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})}^2 - \|\tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}^2 \right]^2 \leq 1 + \sum_{k=1}^n \left[ \|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})}^2 - \|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}^2 \right]^2. \tag{38}$$

We have, using Cauchy–Schwarz inequality,

$$\begin{aligned} \sum_{k=1}^n \left[ \|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})}^2 - \|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}^2 \right]^2 &= \sum_{k=1}^n \left[ \int_{\mathbb{R}^d} \tilde{\mathbf{g}}_k(\boldsymbol{\lambda})^\top \left( \tilde{\mathbf{F}}^{(1)}(\boldsymbol{\lambda}) - \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda}) \right) \mathbf{g}_k(\boldsymbol{\lambda}) d\boldsymbol{\lambda} \right]^2 \\ &\leq \sum_{k=1}^n \left[ \int_{\mathbb{R}^d} \|\mathbf{g}_k(\boldsymbol{\lambda})\| \gamma(\boldsymbol{\lambda}) \|\tilde{\mathbf{F}}^{(1)}(\boldsymbol{\lambda}) - \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda})\| \frac{1}{\gamma(\boldsymbol{\lambda})} \|\mathbf{g}_k(\boldsymbol{\lambda})\| d\boldsymbol{\lambda} \right]^2 \\ &\leq \sum_{k=1}^n \left( \int_{\mathbb{R}^d} \|\mathbf{g}_k(\boldsymbol{\lambda})\|^2 \gamma(\boldsymbol{\lambda})^2 d\boldsymbol{\lambda} \right) \left( \int_{\mathbb{R}^d} \|\tilde{\mathbf{F}}^{(1)}(\boldsymbol{\lambda}) - \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda})\|^2 \frac{1}{\gamma(\boldsymbol{\lambda})^2} \|\mathbf{g}_k(\boldsymbol{\lambda})\|^2 d\boldsymbol{\lambda} \right) \\ &\leq \frac{p2^dT^d}{c_1^3(2\pi)^d} \int_{\mathbb{R}^d} \|\tilde{\mathbf{F}}^{(1)}(\boldsymbol{\lambda}) - \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda})\|^2 \frac{1}{\gamma(\boldsymbol{\lambda})^4} d\boldsymbol{\lambda}, \end{aligned}$$

from Lemma 3, Condition 1, and since  $\|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} = 1$ . The above display is finite and does not depend on  $n$ . We thus have

$$\sum_{k=1}^\infty \left[ \|\tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})}^2 - \|\tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}^2 \right]^2 < \infty.$$

Let us define  $V$  as a symmetric operator on  $\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})$  by  $(V\mathbf{g}, \mathbf{h})_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} = \int_{\mathbb{R}^d} \tilde{\mathbf{g}}(\boldsymbol{\lambda})^\top \tilde{\mathbf{F}}^{(1)}(\boldsymbol{\lambda}) \mathbf{h}(\boldsymbol{\lambda}) d\boldsymbol{\lambda}$ , for  $\mathbf{g}, \mathbf{h} \in \mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})$  (its existence follows from Riesz representation theorem, since for any fixed  $\mathbf{g} \in \mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})$ , the function  $\mathbf{h} \mapsto \int_{\mathbb{R}^d} \tilde{\mathbf{g}}(\boldsymbol{\lambda})^\top \tilde{\mathbf{F}}^{(1)}(\boldsymbol{\lambda}) \mathbf{h}(\boldsymbol{\lambda}) d\boldsymbol{\lambda}$  is continuous and linear on  $\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})$ ). Hence,  $\sum_{k=1}^\infty [(\tilde{\mathbf{g}}_k, (V - I)\tilde{\mathbf{g}}_k)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}]^2 < \infty$ . Recall that  $\tilde{\mathbf{F}}^{(1)} \geq \tilde{\mathbf{F}}^{(0)}$ , so  $(V - I)^{1/2}$  is well-defined and compact, so from the spectral theorem, there exists a sequence of eigenfunctions  $(\mathbf{g}_k)_{k \in \mathbb{N}}$  of  $(V - I)^{1/2}$  with the corresponding eigenvalues  $(\sqrt{\alpha_k})_{k \in \mathbb{N}}$ . Furthermore, we have  $\sum_{k=1}^\infty (\sqrt{\alpha_k})^4 < \infty$ . Hence,  $(\mathbf{g}_k)_{k \in \mathbb{N}}$  is a sequence of eigenfunctions of  $V - I$  with eigenvalues  $(\alpha_k)_{k \in \mathbb{N}}$  with  $\sum_{k=1}^\infty \alpha_k^2 < \infty$  and so  $V - I$  is Hilbert–Schmidt.

We have, for  $r, q \in \mathbb{N}$ , with constants  $c, c'$  and using the equivalence of matrix norms and that  $(\mathbf{g}_k)_{k \in \mathbb{N}}$  is an orthonormal basis of  $\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})$ ,

$$\begin{aligned} &\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\| \sum_{k=r}^{r+q} \alpha_k \mathbf{g}_k(\boldsymbol{\lambda}) \tilde{\mathbf{g}}_k(\boldsymbol{\mu})^\top \right\|^2 \|\tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda})\| \|\tilde{\mathbf{F}}^{(0)}(\boldsymbol{\mu})\| d\boldsymbol{\lambda} d\boldsymbol{\mu} \\ &\leq c \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \sum_{k=r}^{r+q} \|\tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda})^{1/2} (\alpha_k \mathbf{g}_k(\boldsymbol{\lambda}) \tilde{\mathbf{g}}_k(\boldsymbol{\mu})^\top) \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\mu})^{1/2}\|^2 d\boldsymbol{\lambda} d\boldsymbol{\mu} \\ &\leq c' \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \text{tr} \left( \left[ \sum_{k=r}^{r+q} \alpha_k \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda})^{1/2} \mathbf{g}_k(\boldsymbol{\lambda}) \tilde{\mathbf{g}}_k(\boldsymbol{\mu})^\top \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\mu})^{1/2} \right] \right. \\ &\quad \left. \left[ \sum_{\ell=r}^{r+q} \alpha_\ell \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\mu})^{1/2} \mathbf{g}_\ell(\boldsymbol{\mu}) \tilde{\mathbf{g}}_\ell(\boldsymbol{\lambda})^\top \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda})^{1/2} \right] \right) d\boldsymbol{\lambda} d\boldsymbol{\mu} = c' \sum_{k, \ell=r}^{r+q} \alpha_k \alpha_\ell \\ &\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \text{tr} \left( \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda})^{1/2} \mathbf{g}_k(\boldsymbol{\lambda}) \tilde{\mathbf{g}}_k(\boldsymbol{\mu})^\top \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\mu})^{1/2} \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\mu})^{1/2} \mathbf{g}_\ell(\boldsymbol{\mu}) \tilde{\mathbf{g}}_\ell(\boldsymbol{\lambda})^\top \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda})^{1/2} \right) d\boldsymbol{\lambda} d\boldsymbol{\mu} \end{aligned}$$

$$= c' \sum_{k,\ell=r}^{r+q} \alpha_k \alpha_\ell \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \bar{\mathbf{g}}_\ell(\lambda)^\top \tilde{\mathbf{F}}^{(0)}(\lambda) \mathbf{g}_k(\lambda) \bar{\mathbf{g}}_k(\mu)^\top \tilde{\mathbf{F}}^{(0)}(\mu) \mathbf{g}_\ell(\mu) d\lambda d\mu = \sum_{k=r}^{r+q} \alpha_k^2.$$

Since  $(\alpha_k)_{k \in \mathbb{N}}$  is square summable, it follows that

$$\mathbf{B}(\boldsymbol{\mu}, \boldsymbol{\lambda})^\top = \sum_{k=1}^{\infty} \alpha_k \mathbf{g}_k(\boldsymbol{\lambda}) \bar{\mathbf{g}}_k(\boldsymbol{\mu})^\top$$

is well-defined as a limit of Cauchy sequence and we have

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \|\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\mu})\|^2 \|\tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda})\| \|\tilde{\mathbf{F}}^{(0)}(\boldsymbol{\mu})\| d\lambda d\boldsymbol{\mu} < \infty.$$

Let  $\tilde{\mathbf{R}}^{(j)}$  be the matrix covariance function of  $\tilde{\mathbf{Z}}_j$  for  $j = 0, 1$ . For  $i = 1, \dots, p$ , let  $\boldsymbol{\psi}_{i,r}(\boldsymbol{\lambda}) = e^{-i\boldsymbol{\lambda}^\top r} \mathbf{e}_i$  for  $r \in \mathbb{R}^d$ . We have, for  $i, j = 1, \dots, p$ ,

$$\begin{aligned} [\tilde{\mathbf{R}}^{(1)}(\mathbf{t} - \mathbf{s}) - \tilde{\mathbf{R}}^{(0)}(\mathbf{t} - \mathbf{s})]_{i,j} &= \mathbf{e}_i^\top \left( \int_{\mathbb{R}^d} e^{i(\mathbf{t}-\mathbf{s})^\top \boldsymbol{\lambda}} \tilde{\mathbf{F}}^{(1)}(\boldsymbol{\lambda}) d\boldsymbol{\lambda} \right) \mathbf{e}_j - \mathbf{e}_i^\top \left( \int_{\mathbb{R}^d} e^{i(\mathbf{t}-\mathbf{s})^\top \boldsymbol{\lambda}} \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda}) d\boldsymbol{\lambda} \right) \mathbf{e}_j \\ &= \int_{\mathbb{R}^d} \bar{\boldsymbol{\psi}}_{i,t}(\boldsymbol{\lambda})^\top \tilde{\mathbf{F}}^{(1)}(\boldsymbol{\lambda}) \boldsymbol{\psi}_{j,s}(\boldsymbol{\lambda}) d\boldsymbol{\lambda} - \int_{\mathbb{R}^d} \bar{\boldsymbol{\psi}}_{i,t}(\boldsymbol{\lambda})^\top \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda}) \boldsymbol{\psi}_{j,s}(\boldsymbol{\lambda}) d\boldsymbol{\lambda} \\ &= \left( (V - I) \boldsymbol{\psi}_{i,t}, \boldsymbol{\psi}_{j,s} \right)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} = \sum_{k=1}^{\infty} \alpha_k \left( \boldsymbol{\psi}_{i,t}, \mathbf{g}_k \right)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} \left( \mathbf{g}_k, \boldsymbol{\psi}_{j,s} \right)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} \\ &= \sum_{k=1}^{\infty} \alpha_k \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{e}_i^\top e^{i\boldsymbol{\lambda}^\top t} \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda}) \mathbf{g}_k(\boldsymbol{\lambda}) \bar{\mathbf{g}}_k(\boldsymbol{\mu})^\top \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\mu}) e^{-i\boldsymbol{\mu}^\top s} \mathbf{e}_j d\boldsymbol{\lambda} d\boldsymbol{\mu} \\ &= \mathbf{e}_i^\top \left( \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{i(\boldsymbol{\lambda}^\top t - \boldsymbol{\mu}^\top s)} \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda}) \mathbf{B}(\boldsymbol{\mu}, \boldsymbol{\lambda})^\top \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\mu}) d\boldsymbol{\lambda} d\boldsymbol{\mu} \right) \mathbf{e}_j. \end{aligned}$$

Hence using  $\tilde{\mathbf{R}}^{(a)}(-\mathbf{h}) = \tilde{\mathbf{R}}^{(a)}(\mathbf{h})^\top$  for  $a = 0, 1$  and  $\tilde{\mathbf{F}}^{(0)} = (\tilde{\mathbf{F}}^{(0)})^\top$ , we obtain

$$\tilde{\mathbf{R}}^{(1)}(\mathbf{t} - \mathbf{s}) - \tilde{\mathbf{R}}^{(0)}(\mathbf{t} - \mathbf{s}) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-i\boldsymbol{\lambda}^\top t} e^{i\boldsymbol{\mu}^\top s} \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda}) \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\mu}) d\boldsymbol{\lambda} d\boldsymbol{\mu}.$$

Hence, from Theorem 1, the Gaussian measures given by  $\tilde{\mathbf{Z}}^{(0)}$  and  $\tilde{\mathbf{Z}}^{(1)}$  are equivalent, thus so are the Gaussian measures given by  $\mathbf{Z}^{(0)}$  and  $\mathbf{Z}^{(1)}$  as remarked previously.

Let us now drop the assumption that for all  $\boldsymbol{\lambda} \in \mathbb{R}^d$ ,  $\mathbf{F}^{(1)}(\boldsymbol{\lambda}) \geq \mathbf{F}^{(0)}(\boldsymbol{\lambda})$ . Let  $t^+$  be the positive part of  $t \in \mathbb{R}$  and let, for  $\boldsymbol{\lambda} \in \mathbb{R}^d$ ,

$$\begin{aligned} \tilde{\mathbf{F}}^{(0)}(\boldsymbol{\lambda}) &= c_1 \gamma(\boldsymbol{\lambda})^2 \mathbf{I}_p, \\ \tilde{\mathbf{F}}^{(1)}(\boldsymbol{\lambda}) &= c_1 \gamma(\boldsymbol{\lambda})^2 \mathbf{I}_p + \left[ \sup_{\|\mathbf{v}\|=1} (\bar{\mathbf{v}}^\top \mathbf{F}^{(1)}(\boldsymbol{\lambda}) \mathbf{v} - \bar{\mathbf{v}}^\top \mathbf{F}^{(0)}(\boldsymbol{\lambda}) \mathbf{v}) \right]^+ \mathbf{I}_p, \\ \tilde{\mathbf{F}}^{(2)}(\boldsymbol{\lambda}) &= \mathbf{F}^{(0)}(\boldsymbol{\lambda}) - c_1 \gamma(\boldsymbol{\lambda})^2 \mathbf{I}_p \quad \text{and} \\ \mathbf{F}^{(3)}(\boldsymbol{\lambda}) &= \tilde{\mathbf{F}}^{(1)}(\boldsymbol{\lambda}) + \tilde{\mathbf{F}}^{(2)}(\boldsymbol{\lambda}). \end{aligned}$$

Let  $\tilde{\mathbf{Z}}^{(0)}$ ,  $\tilde{\mathbf{Z}}^{(1)}$ ,  $\tilde{\mathbf{Z}}^{(2)}$  and  $\mathbf{Z}^{(3)}$  be four independent  $p$ -variate Gaussian processes with mean function zero and respective matrix spectral densities  $\tilde{\mathbf{F}}^{(0)}$ ,  $\tilde{\mathbf{F}}^{(1)}$ ,  $\tilde{\mathbf{F}}^{(2)}$  and  $\mathbf{F}^{(3)}$ .

Then, in distribution,  $\mathbf{Z}^{(0)} = \tilde{\mathbf{Z}}^{(0)} + \tilde{\mathbf{Z}}^{(2)}$  and  $\mathbf{Z}^{(3)} = \tilde{\mathbf{Z}}^{(1)} + \tilde{\mathbf{Z}}^{(2)}$ . We have

$$\int_{\mathbb{R}^d} \frac{1}{\gamma(\lambda)^4} \|\tilde{\mathbf{F}}^{(0)}(\lambda) - \tilde{\mathbf{F}}^{(1)}(\lambda)\|^2 d\lambda \leq \int_{\mathbb{R}^d} \frac{1}{\gamma(\lambda)^4} \|\mathbf{F}^{(0)}(\lambda) - \mathbf{F}^{(1)}(\lambda)\|^2 d\lambda < +\infty.$$

Hence the Gaussian measures given by  $\mathbf{Z}^{(0)}$  and  $\mathbf{Z}^{(3)}$  are equivalent.

We remark that for any  $\lambda \in \mathbb{R}^d$  and  $\mathbf{v} \in \mathbb{C}^p$  with  $\|\mathbf{v}\| = 1$ , we have

$$\begin{aligned} & \bar{\mathbf{v}}^\top \mathbf{F}^{(3)}(\lambda) \mathbf{v} - \bar{\mathbf{v}}^\top \mathbf{F}^{(1)}(\lambda) \mathbf{v} \\ &= \bar{\mathbf{v}}^\top \mathbf{F}^{(0)}(\lambda) \mathbf{v} - \bar{\mathbf{v}}^\top \mathbf{F}^{(1)}(\lambda) \mathbf{v} + \left[ \sup_{\|\mathbf{v}\|=1} (\bar{\mathbf{v}}^\top \mathbf{F}^{(1)}(\lambda) \mathbf{v} - \bar{\mathbf{v}}^\top \mathbf{F}^{(0)}(\lambda) \mathbf{v}) \right]^+ \geq 0. \end{aligned}$$

Hence, applying the previous step of the proof, the measures given by  $\mathbf{Z}^{(1)}$  and  $\mathbf{Z}^{(3)}$  are equivalent since

$$\int_{\mathbb{R}^d} \frac{1}{\gamma(\lambda)^4} \|\mathbf{F}^{(3)}(\lambda) - \mathbf{F}^{(1)}(\lambda)\|^2 d\lambda \leq 4 \int_{\mathbb{R}^d} \frac{1}{\gamma(\lambda)^4} \|\mathbf{F}^{(0)}(\lambda) - \mathbf{F}^{(1)}(\lambda)\|^2 d\lambda < \infty.$$

Hence the Gaussian measures given by  $\mathbf{Z}^{(0)}$  and  $\mathbf{Z}^{(1)}$  are equivalent. □

The next lemma is immediate.

**Lemma 1.** Assume that Condition 1 holds. With  $c_3 = c_2/c_1$ , for  $\lambda \in \mathbb{R}^d$ ,

$$\mathbf{F}^{(1)}(\lambda) \leq c_3 \mathbf{F}^{(0)}(\lambda).$$

**Lemma 2.** We can find  $\mathbf{g}_1, \dots, \mathbf{g}_n$  in  $\mathcal{W}_D$  as described in (38).

**Proof of Lemma 2.** By density, for  $\epsilon > 0$ , we can find  $\hat{\mathbf{g}}_1, \dots, \hat{\mathbf{g}}_n$  in  $\mathcal{W}_D$  such that for  $k = 1, \dots, n$ ,  $\|\hat{\mathbf{g}}_k - \tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} \leq \epsilon/n^2$ . Let  $\mathbf{M} = [(\hat{\mathbf{g}}_k, \hat{\mathbf{g}}_\ell)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}]_{k, \ell=1, \dots, n}$ . We have, for  $k, \ell = 1, \dots, n$ ,

$$\begin{aligned} \left| (\hat{\mathbf{g}}_k, \hat{\mathbf{g}}_\ell)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} - (\tilde{\mathbf{g}}_k, \tilde{\mathbf{g}}_\ell)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} \right| &\leq \left| (\hat{\mathbf{g}}_k - \tilde{\mathbf{g}}_k, \hat{\mathbf{g}}_\ell)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} \right| + \left| (\tilde{\mathbf{g}}_k, \hat{\mathbf{g}}_\ell - \tilde{\mathbf{g}}_\ell)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} \right| \\ &\leq \frac{3\epsilon}{n^2}, \end{aligned} \tag{39}$$

from Cauchy–Schwarz. Also, from Lemma 1,

$$\begin{aligned} \left| (\hat{\mathbf{g}}_k, \hat{\mathbf{g}}_\ell)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})} - (\tilde{\mathbf{g}}_k, \tilde{\mathbf{g}}_\ell)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})} \right| &\leq \left| (\hat{\mathbf{g}}_k - \tilde{\mathbf{g}}_k, \hat{\mathbf{g}}_\ell)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})} \right| + \left| (\tilde{\mathbf{g}}_k, \hat{\mathbf{g}}_\ell - \tilde{\mathbf{g}}_\ell)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})} \right| \\ &\leq \frac{c_3 3\epsilon}{n^2}. \end{aligned} \tag{40}$$

From Gershogrin’s circle theorem, we thus have  $\|\mathbf{M} - \mathbf{I}_n\| \leq 3\epsilon/n$ . For  $k = 1, \dots, n$ , let  $\mathbf{g}_k = \sum_{\ell=1}^n (\mathbf{M}^{-1/2})_{k, \ell} \hat{\mathbf{g}}_\ell$ . Then,  $\mathbf{g}_1, \dots, \mathbf{g}_n$  satisfy  $(\mathbf{g}_k, \mathbf{g}_l)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} = \delta_{k, l}$  for  $k, l = 1, \dots, n$ . Furthermore, there exists a constant  $c_n < \infty$ , not depending on  $\epsilon$ , such that  $\|\mathbf{g}_k - \tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} \leq c_n \epsilon/n^2$ .



We then have, with a constant  $c'_n$  not depending on  $\epsilon$ ,

$$\begin{aligned} & \left| \sum_{k=1}^n \left[ \|\tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})}^2 - \|\tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}^2 \right]^2 - \sum_{k=1}^n \left[ \|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})}^2 - \|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}^2 \right]^2 \right| \\ &= \left| \sum_{k=1}^n \left( \|\tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})}^2 - \|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})}^2 - \|\tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}^2 + \|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}^2 \right) \right. \\ & \quad \left. \left( \|\tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})}^2 + \|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(1)})}^2 - \|\tilde{\mathbf{g}}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}^2 - \|\mathbf{g}_k\|_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})}^2 \right) \right| \\ & \leq c'_n \frac{\epsilon}{n^2}, \end{aligned}$$

proceeding as for (39) and (40). This concludes the proof since  $\epsilon$  can be chosen arbitrarily small.  $\square$

**Lemma 3.** *In the context of the proof of Theorem 2, for  $n \in \mathbb{N}$ , let  $\mathbf{g}_1, \dots, \mathbf{g}_n$  in  $\mathcal{W}_D$  for which  $(\mathbf{g}_k, \mathbf{g}_l)_{\mathcal{W}_D(\tilde{\mathbf{F}}^{(0)})} = \delta_{k,l}$  for  $k, l = 1, \dots, n$ . Then for any  $\lambda \in \mathbb{R}^d$ , with a finite constant  $T$ ,*

$$\sum_{k=1}^n \|\mathbf{g}_k(\lambda)\|^2 \leq \frac{p2^d T^d}{\gamma^2(\lambda)c_1^2(2\pi)^d}.$$

**Proof of Lemma 3.** For  $\lambda \in \mathbb{R}^d$  and  $k = 1, \dots, n$ , let  $\mathbf{h}_k(\lambda) = (h_{k,1}(\lambda), \dots, h_{k,p}(\lambda))^\top = \gamma(\lambda)\mathbf{g}_k(\lambda)$ . By convolution, with a finite constant  $T$ , there exist  $p$  square summable functions  $\psi_{k,1}, \dots, \psi_{k,p}$  from  $[-T, T]^d$  to  $\mathbb{C}$  such that for  $i = 1, \dots, p$

$$h_{k,i}(\lambda) = \int_{[-T, T]^d} e^{-i\lambda^\top t} \psi_{k,i}(t) dt.$$

We have, for  $k, \ell = 1, \dots, n$

$$c_1 \int_{\mathbb{R}^d} \bar{\mathbf{g}}_k(\lambda)^\top \gamma(\lambda)^2 \mathbf{g}_\ell(\lambda) d\lambda = \delta_{k,\ell}$$

and thus

$$c_1 \int_{\mathbb{R}^d} \bar{\mathbf{h}}_k(\lambda)^\top \mathbf{h}_\ell(\lambda) d\lambda = \delta_{k,\ell}.$$

By Plancherel’s theorem we obtain

$$c_1 \int_{[-T, T]^d} \bar{\psi}_{k,1}(t) \psi_{\ell,1}(t) dt + \dots + c_1 \int_{[-T, T]^d} \bar{\psi}_{k,p}(t) \psi_{\ell,p}(t) dt = \frac{1}{(2\pi)^d} \delta_{k,\ell}.$$

Hence,  $c_1(2\pi)^{d/2}((\psi_{k,1}, \dots, \psi_{k,p})^\top)_{k=1, \dots, n} = c_1(2\pi)^{d/2}(\psi_k)_{k=1, \dots, n}$  is an orthonormal system in  $L^2_{[-T, T]^d}$ .

For  $i = 1, \dots, p$ , let  $\phi_{\lambda,i}(t) = (0, \dots, 0, e^{i\lambda^\top t}, 0, \dots, 0)^\top$  for  $\lambda, t \in \mathbb{R}^d$ , where the non-zero element is at position  $i$ . From Bessel’s inequality we obtain

$$\sum_{k=1}^n |h_{k,i}(\lambda)|^2 = \sum_{k=1}^n \left| \int_{[-T, T]^d} \psi_{k,i}(t) e^{-i\lambda^\top t} dt \right|^2 = \sum_{k=1}^n \left| \int_{[-T, T]^d} \bar{\phi}_{\lambda,i}(t)^\top \psi_k(t) dt \right|^2$$

$$\leq \frac{1}{c_1^2(2\pi)^d} \int_{[-T, T]^d} \|\phi_{\lambda, i}(t)\|^2 dt \leq \frac{1}{c_1^2(2\pi)^d} (2T)^d.$$

Hence

$$\sum_{k=1}^n \gamma(\lambda)^2 \|\mathbf{g}_k(\lambda)\|^2 = \sum_{i=1}^p \sum_{k=1}^n |h_{k, i}(\lambda)|^2 \leq \frac{p(2T)^d}{c_1^2(2\pi)^d}. \quad \square$$

## Acknowledgments

The authors are grateful to the Referees and Associate Editor for their comments, that lead to a significant improvement of the paper, both in terms of content and exposition.

## Funding

The work of Moreno Bevilacqua was partially supported by ANID – Millennium Science Initiative Program – NCN17\_059 from the Chilean government, by FONDECYT grant 1200068 (Chile) and by regional MATH-AmSud program, grant number 20-MATH-03. Reinhard Furrer was supported by the Swiss National Science Foundation (SNSF-175529). This work was supported partially in part by FONDECYT grant 11200749 and in part by grant DIUBB 2020525 IF/R from the university of Bío-Bío. This publication is based upon work supported by the Khalifa University of Science and Technology under Award No. FSU-2021-016 (E. Porcu).

## Supplementary Material

**Remaining proofs** (DOI: [10.3150/21-BEJ1427SUPP](https://doi.org/10.3150/21-BEJ1427SUPP); .pdf). We provide the proofs for Remark 1 and Section 4.

## References

- Abramowitz, M. and Stegun, I.A., eds. (1966). *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*. New York: Dover. [MR0208797](https://doi.org/10.1016/j.jmva.2017.09.001)
- Alegría, A. and Porcu, E. (2017). The dimple problem related to space-time modeling under the Lagrangian framework. *J. Multivariate Anal.* **162** 110–121. [MR3719338](https://doi.org/10.1016/j.jmva.2017.09.001) <https://doi.org/10.1016/j.jmva.2017.09.001>
- Alegría, A., Porcu, E., Furrer, R. and Mateu, J. (2019). Covariance functions for multivariate Gaussian fields. *Stoch. Environ. Res. Risk Assess.* **33** 1593–1608.
- Anderes, E. (2010). On the consistent separation of scale and variance for Gaussian random fields. *Ann. Statist.* **38** 870–893. [MR2604700](https://doi.org/10.1214/09-AOS725) <https://doi.org/10.1214/09-AOS725>
- Apanasovich, T. and Genton, M.G. (2010). Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika* **97** 15–30. [MR2594414](https://doi.org/10.1093/biomet/asp078) <https://doi.org/10.1093/biomet/asp078>
- 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. *J. Amer. Statist. Assoc.* **107** 180–193. [MR2949350](https://doi.org/10.1080/01621459.2011.643197) <https://doi.org/10.1080/01621459.2011.643197>
- Askey, R. (1973). Radial characteristic functions Technical report, Research Center, Univ. Wisconsin.
- Azañs, J.-M., Bachoc, F., Lagnoux, A. and Nguyen, T.M.N. (2020). Semi-parametric estimation of the variogram scale parameter of a Gaussian process with stationary increments. *ESAIM Probab. Stat.* **24** 842–882. [MR4178368](https://doi.org/10.1051/ps/2020021) <https://doi.org/10.1051/ps/2020021>

- Bachoc, F. (2014). Asymptotic analysis of the role of spatial sampling for covariance parameter estimation of Gaussian processes. *J. Multivariate Anal.* **125** 1–35. MR3163828 <https://doi.org/10.1016/j.jmva.2013.11.015>
- Bachoc, F., Porcu, E., Bevilacqua, M., Furrer, R. and Faouzi, T. (2022). Supplement to “Asymptotically equivalent prediction in multivariate geostatistics.” <https://doi.org/10.3150/21-BEJ1427SUPP>
- Bevilacqua, M., Hering, A.S. and Porcu, E. (2015). On the flexibility of multivariate covariance models: Comment on the paper by Genton and Kleiber. *Statist. Sci.* **30** 167–169. MR3353098 <https://doi.org/10.1214/15-STS516>
- Bevilacqua, M., Faouzi, T., Furrer, R. and Porcu, E. (2019). Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics. *Ann. Statist.* **47** 828–856. MR3909952 <https://doi.org/10.1214/17-AOS1652>
- Blackwell, D. and Dubins, L. (1962). Merging of opinions with increasing information. *Ann. Math. Stat.* **33** 882–886. MR0149577 <https://doi.org/10.1214/aoms/1177704456>
- Da Prato, G. and Zabczyk, J. (2014). *Stochastic Equations in Infinite Dimensions*, 2nd ed. *Encyclopedia of Mathematics and Its Applications* **152**. Cambridge: Cambridge Univ. Press. MR3236753 <https://doi.org/10.1017/CBO9781107295513>
- Daley, D., Porcu, E. and Bevilacqua, M. (2015). Classes of compactly supported covariance functions for multivariate random fields. *Stoch. Environ. Res. Risk Assess.* **29** 1249–1263.
- Furrer, R., Genton, M.G. and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets. *J. Comput. Graph. Statist.* **15** 502–523. MR2291261 <https://doi.org/10.1198/106186006X132178>
- Genton, M.G. and Kleiber, W. (2015). Cross-covariance functions for multivariate geostatistics. *Statist. Sci.* **30** 147–163. MR3353096 <https://doi.org/10.1214/14-STS487>
- Gneiting, T. (2002a). Compactly supported correlation functions. *J. Multivariate Anal.* **83** 493–508. MR1945966 <https://doi.org/10.1006/jmva.2001.2056>
- Gneiting, T. (2002b). Nonseparable, stationary covariance functions for space-time data. *J. Amer. Statist. Assoc.* **97** 590–600. MR1941475 <https://doi.org/10.1198/016214502760047113>
- Gneiting, T., Kleiber, W. and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *J. Amer. Statist. Assoc.* **105** 1167–1177. MR2752612 <https://doi.org/10.1198/jasa.2010.tm09420>
- Golubov, B.I. (1981). On Abel-Poisson type and Riesz means. *Anal. Math.* **7** 161–184. MR0635483 <https://doi.org/10.1007/BF01908520>
- Ibragimov, I.D.A. and Rozanov, Y.A. (1978). *Gaussian Random Processes. Applications of Mathematics* **9**. New York-Berlin: Springer. Translated from the Russian by A. B. Aries. MR0543837
- Istas, J. and Lang, G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Henri Poincaré Probab. Stat.* **33** 407–436. MR1465796 [https://doi.org/10.1016/S0246-0203\(97\)80099-4](https://doi.org/10.1016/S0246-0203(97)80099-4)
- Kaufman, C.G., Schervish, M.J. and Nychka, D.W. (2008). Covariance tapering for likelihood-based estimation in large spatial data sets. *J. Amer. Statist. Assoc.* **103** 1545–1555. MR2504203 <https://doi.org/10.1198/016214508000000959>
- Lang, G. and Roueff, F. (2001). Semi-parametric estimation of the Hölder exponent of a stationary Gaussian process with minimax rates. *Stat. Inference Stoch. Process.* **4** 283–306. MR1868724 <https://doi.org/10.1023/A:1012227325436>
- Loh, W.-L. (2015). Estimating the smoothness of a Gaussian random field from irregularly spaced data via higher-order quadratic variations. *Ann. Statist.* **43** 2766–2794. MR3405611 <https://doi.org/10.1214/15-AOS1365>
- Maniglia, S. and Rhandi, A. (2004). Gaussian measures on separable Hilbert spaces and applications. *Quad. Mat.* **2004**.
- Mardia, K.V. and Marshall, R.J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika* **71** 135–146. MR0738334 <https://doi.org/10.1093/biomet/71.1.135>
- Porcu, E., Bevilacqua, M. and Genton, M.G. (2016). Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. *J. Amer. Statist. Assoc.* **111** 888–898. MR3538713 <https://doi.org/10.1080/01621459.2015.1072541>
- Porcu, E., Gregori, P. and Mateu, J. (2006). Nonseparable stationary anisotropic space-time covariance functions. *Stoch. Environ. Res. Risk Assess.* **21** 113–122. MR2307631 <https://doi.org/10.1007/s00477-006-0048-3>
- Porcu, E. and Zastavnyi, V. (2011). Characterization theorems for some classes of covariance functions associated to vector valued random fields. *J. Multivariate Anal.* **102** 1293–1301. MR2811618 <https://doi.org/10.1016/j.jmva.2011.04.013>

- Putter, H. and Young, G.A. (2001). On the effect of covariance function estimation on the accuracy of Kriging predictors. *Bernoulli* **7** 421–438. [MR1836738](#) <https://doi.org/10.2307/3318494>
- Schaback, R. (2011). The missing Wendland functions. *Adv. Comput. Math.* **34** 67–81. [MR2783302](#) <https://doi.org/10.1007/s10444-009-9142-7>
- Skorohod, A.V. and Jadrenko, M.Ī. (1973). Absolute continuity of measures that correspond to homogeneous Gaussian fields. *Theory Probab. Appl.* **18** 27–40.
- Stein, M.L. (1988). Asymptotically efficient prediction of a random field with a misspecified covariance function. *Ann. Statist.* **16** 55–63. [MR0924856](#) <https://doi.org/10.1214/aos/1176350690>
- Stein, M. (1990). Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure. *Ann. Statist.* **18** 850–872. [MR1056340](#) <https://doi.org/10.1214/aos/1176347629>
- Stein, M.L. (1993). A simple condition for asymptotic optimality of linear predictions of random fields. *Statist. Probab. Lett.* **17** 399–404. [MR1237787](#) [https://doi.org/10.1016/0167-7152\(93\)90261-G](https://doi.org/10.1016/0167-7152(93)90261-G)
- Stein, M.L. (1999a). *Interpolation of Spatial Data. Some Theory for Kriging*. Springer Series in Statistics. New York: Springer. [MR1697409](#) <https://doi.org/10.1007/978-1-4612-1494-6>
- Stein, M.L. (1999b). Predicting random fields with increasing dense observations. *Ann. Appl. Probab.* **9** 242–273. [MR1682572](#) <https://doi.org/10.1214/aoap/1029962604>
- Stein, M.L. (2004). Equivalence of Gaussian measures for some nonstationary random fields. *J. Statist. Plann. Inference* **123** 1–11. [MR2058118](#) [https://doi.org/10.1016/S0378-3758\(03\)00144-7](https://doi.org/10.1016/S0378-3758(03)00144-7)
- Vallejos, R., Osorio, F. and Bevilacqua, M. (2020). *Spatial Relationships Between Two Georeferenced Variables—with Applications in R*. Cham: Springer. [MR4248932](#) <https://doi.org/10.1007/978-3-030-56681-4>
- Wackernagel, H. (2003). *Multivariate Geostatistics: An Introduction with Applications*, 3rd ed. New York: Springer.
- Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv. Comput. Math.* **4** 389–396. [MR1366510](#) <https://doi.org/10.1007/BF02123482>
- Zastavnyĭ, V.P. (2006). On some properties of the Buhmann functions. *Ukrain. Mat. Zh.* **58** 1045–1067. [MR2345078](#) <https://doi.org/10.1007/s11253-006-0128-z>
- Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Statist. Assoc.* **99** 250–261. [MR2054303](#) <https://doi.org/10.1198/016214504000000241>
- Zhang, H. and Cai, W. (2015). When doesn't cokriging outperform kriging? *Statist. Sci.* **30** 176–180. [MR3353100](#) <https://doi.org/10.1214/15-ST518>

Received July 2020 and revised September 2021