Estimation and prediction of Gaussian random fields under fixed domain asymptotics using generalized Wendland covariance functions

Universidad de Valparaíso
CHILE
Instituto de Estadística
Facultad de Ciencias

Tarik Faouzi

Supervisor: Prof. Moreno Bevilacqua

Advisor: Prof. Emilio Porcu

Instituto de Estadística
Universidad de Valparaíso

This dissertation is submitted for the degree of
Doctor en Estadística

May 2017

Executive Summary

Covariance functions cover a central aspect of inference and prediction of Gaussian fields defined over some (compact) set of $\mathbb{R}^d$. For instance, the best linear unbiased prediction at an unobserved site depends on the knowledge of the covariance function. Since a covariance function must be positive definite, practical estimation generally requires the selection of some parametric classes of covariances and the corresponding estimation of these parameters.

The maximum likelihood (ML) estimation method is generally considered best for estimating the parameters of covariance models. The study of asymptotic properties of ML estimators is complicated by the fact that more than one asymptotic frameworks can be considered when observing a single realization from a Gaussian field. In particular, under fixed domain asymptotics, one supposes that the sampling domain is bounded and that the sampling set becomes increasingly dense. Under increasing domain, the sampling domain increases with the number of observed data, and the distance between any two sampling locations is bounded away from zero.

The asymptotic behavior of ML estimators of the covariance parameters can be quite different under these two frameworks (Zhang and Zimmerman, 2005). A general result under increasing domain asymptotic framework and some mild regularity conditions is given in Mardia and Marshall (1984). Specifically, they show that ML estimators are consistent and asymptotically Gaussian, with asymptotic covariance matrix equal to the inverse of the Fisher information matrix.

Equivalence of Gaussian measures (Skorokhod and Yadrenko, 1973; Ibragimov and Rozanov, 1978) represents an essential tool to establish the asymptotic properties of both prediction and estimation of Gaussian fields under fixed domain asymptotics. In his *tour de force*, Stein (1988, 1990, 1993, 1999b, 2004) provides conditions under which predictions under a misspecified covariance function are asymptotically efficient, and mean square errors converge almost surely to their targets. Since Gaussian measures depend exclusively on their mean and covariance functions, practical evaluation of Stein's conditions translates into the fact that the true and the misspecified covariances must be compatible, i.e., the induced Gaussian measures are equivalent.

Under fixed domain asymptotics, no general results are available for the asymptotic properties of ML estimators. Yet, some results have been obtained when assuming that the covariance belongs to the parametric family of Matérn covariance functions (Matérn, 1960) that has been very popular in spatial statistics for its flexibility with respect to continuous parameterization of smoothness, in the mean square sense, of the underlying Gaussian field. For a Gaussian field defined over a bounded and infinite set of $\mathbb{R}^d$, Zhang (2004) shows that when the smoothness parameter is known and fixed, not all parameters can be estimated consistently when $d = 1,2,3$. Instead, the ratio of variance and scale (to the power of the smoothing parameter), sometimes called microergodic parameter (Zhang and Zimmerman, 2005; Stein, 1999a), is consistently estimable. In contrast for $d \geq 5$, Anderes (2010) proved the orthogonality of two Gaussian measures with different Matérn covariance functions and hence, in this case, all the parameters can be consistently estimated under fixed-domain asymptotics. The case $d = 4$ is still open.

Microergodic parameters are important since they affect the prediction much more than the non-microergodic parameters and they are generally consistently estimable under the fixed domain asymptotic framework.

Asymptotic results for ML estimator of the microergodic parameter of the Matérn model can be found in Zhang (2004), Du et al. (2009), Wang and Loh (2011) and Kaufman and Shaby (2013). In particular, Kaufman and Shaby (2013) give strong consistency and asymptotic distribution of the microergodic parameter when estimating jointly the scale and variance parameters, generalizing previous results in Zhang (2004) and Wang and Loh (2011) where the scale parameter is assumed known and fixed. Kaufman and Shaby (2013) show by means of a simulation study that asymptotic approximation using a fixed scale parameter can be problematic when applied to finite samples, even for large sample sizes. In contrast, they show that performance is improved and asymptotic approximations are applicable for smaller sample sizes when the parameters are jointly estimated.

Under the Matérn family, similar results have been obtained for the covariance tapering (CT) method of estimation, as originally proposed in Kaufman et al. (2008) and consisting of setting to zero the dependence after a given distance. This is in turn achieved by multiplying the Matérn covariance with a taper function, that is, a correlation function being additionally compactly supported over a ball with given radius. Thus, the resulting covariance tapered matrix is sparse, with the level of sparseness depending on the radius of compact support. Sparse matrix algorithms can then be used to evaluate efficiently an approximate likelihood where the original covariance matrix is replaced by the tapered matrix. The results proposed in Kaufman et al. (2008) have then inspired the works in Du et al. (2009), Wang and Loh (2011) and Kaufman and Shaby (2013), where asymptotic properties of the CT estimator of the Matérn microergodic parameter are given.

Using the Matérn family, Furrer et al. (2006) study CT when applied to the best unbiased linear predictor and show that under fixed domain asymptotics and some specific conditions of the taper function, asymptotic efficiency prediction and asymptotically correct estimation of mean square error can be achieved using a tapered Matérn covariance function. Extensions have been discussed by, e.g., Stein (2013) and Hirano and Yajima (2013). The basic message of CT method is that large data sets can be handled both for estimation and prediction exploiting sparse matrix algorithms when using the Matérn model.

Inspired by this idea, we focus on a covariance model that offers the strength of the Matérn family and allows the use of sparse matrices. Specifically, we study estimation and prediction of Gaussian fields under fixed domain asymptotics, using the generalized Wendland (GW) class of covariance functions (Gneiting, 2002a; Zastavnyi, 2006), the members of which are compactly supported over balls of $\mathbb{R}^d$ with arbitrary radii, and additionally allow for a continuous parameterization of differentiability at the origin, in a similar way to the Matérn family.

In particular, we provide the following results. First, we characterize the equivalence of two Gaussian measures with covariance functions belonging to the GW class and sharing the same smoothness parameter. A consequence of this result is that, as in the Matérn case (Zhang, 2004), when the smoothness parameter is known and fixed, not all parameters can be estimated consistently under fixed domain asymptotics. Then we give sufficient conditions for the equivalence of two Gaussian measures where the state of truth is represented by a member of the Matérn family and the other measure has a GW covariance model and vice versa.

We assess the asymptotic properties of the ML estimator of the microergodic parameter associated with the GW class. Specifically, for a fixed smoothness parameter, we establish strong consistency and asymptotic distribution of the microergodic parameter assuming the compact support parameter fixed and known. Then, we generalize these results when jointly estimating with ML the variance and the compact support parameter.

Finally, using results in Stein (1988, 1993), we study the implications of our results on prediction under fixed domain asymptotics. One remarkable implication is that when the true covariance belongs to the Matérn family, asymptotic efficiency prediction and asymptotically correct estimation of mean square error can be achieved using a compatible GW covariance model.

Possible asked the impossible: Where are you staying? He answered: in the dreams of impotent. Life is full of tests and trials that an individual must face with patience. Thanks you for your patience.

to

My big & small family

Acknowledgements

I would like to express my sincere gratitude to my advisor, Dr. Moreno Bevilacqua. This research would not have been possible without his talented insightful guidance and continuing motivation. I am thankful to my second advisor, Dr. Emilio Porcu, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. Also, I thank Dr. Furrer Reinhard for his fundamental contribution to the studied theme.

Abstract

Spanish Version

En esta tesis, estudiamos la estimación y predicción de campos aleatorios Gaussianos con modelos de covarianza pertenecientes a la clase generalizada de Wendland (GW), bajo asintóticos de dominio fijo. Como el caso Matérn, esta clase permite una parametrización continua de la suavidad del campo aleatorio Gaussiano a soporte compacto.

Específicamente, caracterizamos primero la equivalencia de dos medidas Gaussianas con función de covarianza GW, y proporcionamos condiciones suficientes para la equivalencia de dos medidas Gaussianas de funciones de covarianza Matérn y GW. Elucidar las consecuencias de estos hechos en términos de (mal especificados) mejores predictores lineales no sesgados. En la segunda parte, establecemos una consistencia y distribución asintótica del estimador de máxima verosimilitud del parámetro microergódico asociado al modelo de covarianza GW, bajo dominio fijo asintótico.

Nuestros hallazgos se ilustran a través de un estudio de simulación: el primero compara el comportamiento de la muestra finita de la estimación de máxima verosimilitud del parámetro microergódico con la distribución asintótica dada. El último compara el comportamiento de la muestra finita de la predicción y su error cuadrático medio asociado cuando se usan dos medidas Gaussianas equivalentes con el modelo de covarianza Matérn y GW, utilizando la covarianza tapering como punto de referencia.

English Version

In this thesis, we study estimation and prediction of Gaussian random fields with covariance models belonging to the generalized Wendland (GW) class, under fixed domain asymptotics. As the Matérn case, this class allows a continuous parameterization of smoothness of the underlying Gaussian random field, being additionally compactly supported. Specifically, we first characterize the equivalence of two Gaussian measures with GW covariance function, and we provide sufficient conditions for the equivalence of two Gaussian measures with

Matérn and GW covariance functions. We elucidate the consequences of these facts in terms of (misspecified) best linear unbiased predictors. In the second part, we establish strong consistency and asymptotic distribution of the maximum likelihood estimator of the microergodic parameter associated with GW covariance model, under fixed domain asymptotics. Our findings are illustrated through a simulation study: the former compares the finite sample behavior of the maximum likelihood estimation of the microergodic parameter with the given asymptotic distribution. The latter compares the finite-sample behavior of the prediction and its associated mean square error when using two equivalent Gaussian measures with Matérn and GW covariance model, using covariance tapering as a benchmark.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

For spatial Gaussian random fields, the inverse of the covariance matrix is a fundamental tool when estimating with maximum likelihood and computing prediction at some unknown location sites.

Since the spatial sample size can be quite large, computing the inverse can be numerical challenging if not impossible. If the spatial data does not posses a particular structure (i.e. regular lattice) the most widely used algorithms such as Cholesky decomposition require up to $O(n^3)$ steps if n is the number of location sites. This can be prohibitive if n is large.

A natural idea is to use covariance models with compact support so that the resulting matrix has a high proportion of zero entries. This idea has been exploited, for instance, in the covariance tapering approach as described in Furrer et al. (2006). In this case, algorithm for sparse matrices can be used in order to speed up the computation of the inverse of the covariance matrix. We focus here on the Wendland class which is positive definite and with compact support (Wendland, 1995). the use of Wendland covariance function is justified by the fact that this covariance has the same properties as the Matérn covariance function. Nevertheless the associated covariance matrix is sparse while it is full in the case of the Matérn case.

For parameter estimation, different asymptotic frameworks have been studied. When the sampling domain is bounded and more data are observed from this bounded region, this is the fixed domain or infill asymptotic framework. When the distance between any two sampling locations is bounded away from zero and more data are observed, this is the increasing domain asymptotic framework.

Under increasing domain asymptotic framework Mardia and Marshall (1984) give general results about consistency and asymptotic Gaussianity of the maximum likelihood estimates. Under infill asymptotic results are established only for the Matérn model (Zhang, 2004).

Infill asymptotics has been also considered in Stein (1999a) when developing asymptotic theory for kriging. All the known results in literature for the estimation and prediction of Gaussian random fields, under infill asymptotics, are based on the equivalence of Gaussian measure.

We consider these two problems together, the equivalence of two Gaussian measures with covariance functions belonging to the Generalized Wendland (GW) class and sharing the same smoothness parameter, and the equivalence of two Gaussian measures where the state of truth is represented by a member of the Matérn family and the other measure has a GW covariance model and vice versa.

The goal of this thesis is to offer an alternative covariance functions family that have the same properties as the Matérn covariance functions, and both present the same prediction with the same asymptotic mean square error under some conditions when two Gaussian measures are equivalent where the state of truth is represented by a member of the Matérn family and the other measure has a GW covariance model.

Our thesis is organized in the following form: The second chapter gives basic tools on spatial Gaussian random fields. we present a set of definitions and properties of spatial Gaussian random fields, stationarity, differentiability that are necessary for the development of the proposed theme. Furthermore, we define a relevant method of interpolation.

Chapter 3 describes two covariance models used in this Thesis: Matérn and Generalized Wendland models. Chapter 4 gives some inequalities to demonstrate theorem 13. Also, this chapter offers tools on Fourier theory analysis and completely monotonic functions necessary for the results in Chapter 5. All the results of this Thesis are given in Chapter 5 where we first characterize the equivalence of Gaussian measure under the GW covariance model and we find a sufficient condition for the equivalence of two Gaussian measures with Matérn and GW covariance models. Then we establish strong consistency and asymptotic distribution of the ML estimator of the microergodic parameter of the GW models, under fixed domain asymptotics.

We discuss the consequences of the previous results on prediction under fixed domain asymptotics. Moreover we provide two simulation studies: The first shows how well the given asymptotic distribution of the microergodic parameter applies to finite sample cases when estimating with ML a GW covariance model under fixed domain asymptotics. The second compares the finite-sample behavior of the prediction when using two compatible

Matérn and GW models, using CT as a benchmark. The final section f chapter 5 provides discussion on the consequence of our results and identifies problems for future research. We finish with a small chapter which we offer some open questions and future works.

Chapter 2

Background of spatial statistics

We now review a few important basics of spatial processes.

2.1 Spatial Random Fields

We denote as $Z(\boldsymbol{s})$, where $\boldsymbol{s}$ indexes location and $\boldsymbol{s} \in \mathscr{D} \subset \mathbb{R}^d$, the quantity we are studying. For each $\boldsymbol{s}$, $Z(\boldsymbol{s})$ is a random variable. The collection, $Z(\boldsymbol{s})$, when $\boldsymbol{s}$ varies over all its possible values, is called a spatial process or random field i.e. $Z = \{Z(\boldsymbol{s}), \boldsymbol{s} \in \mathscr{D} \subset \mathbb{R}^d\}$. In practice, $Z(\boldsymbol{s})$ is a random function indexed by the symbol $\boldsymbol{s}$ which belongs to some index set $\mathscr{D}$. When $d = 1$ it is usually called random stochastic or random process, while when $d \geq 2$ it is defined as random field.

Definition 1. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $\mathscr{D} \in \mathbb{R}^d$ be an arbitrary set. For each $\boldsymbol{s} \in \mathscr{D}$ the function $Z(\boldsymbol{s}, .) : \Omega \to \mathbb{R},\ \omega \to Z(\boldsymbol{s}, \omega)$ is a random variable, and any collection of random variables $Z = \{Z(\boldsymbol{s}, .) : \boldsymbol{s} \in \mathscr{D}\}$ defined on $(\Omega, \mathscr{F}, \mathbb{P})$ is a stochastic process with index set $\mathscr{D}$.*

Definition 2. *A sample path of spatial stochastic process is a mapping $Z : \mathscr{D} \to \mathbb{R}, \boldsymbol{s} \to Z(\boldsymbol{s}, \omega)$ which to every event $\omega \in \Omega$ corresponds a sample path or trajectory of random field process Z.*

$Z(\boldsymbol{s})$ is simply a random variable for each $\boldsymbol{s}$ and its properties (e.g. mean and variance) can be described by its distribution function. More generally we are interested in studying the whole collection of random variables $Z = \{Z(\boldsymbol{s}), \boldsymbol{s} \in \mathscr{D}\}$ and its joint distribution function.

The finite-dimensional distributions of the random field Z are the distributions of the finite-dimensional vectors

$$\mathbf{Z} = [Z(\boldsymbol{s}_1), \ldots Z(\boldsymbol{s}_n)]^T, \qquad \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\} \in \mathscr{D},$$

for all possible choices of $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ and every n, i.e.

$$F_{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n}(k_1,\ldots,k_n) = Pr(Z(\boldsymbol{s}_1) \leq k_1, \ldots, Z(\boldsymbol{s}_n) \leq k_n).$$

The family of all finite-dimensional distributions of Z is called the spatial distribution of the process. Given a family of finite-dimensional distributions, Kolmogorov's existence theorem (Abrahamsen, 1997), establishes the existence of a Gaussian stochastic process associated with such family.

2.2 Stationarity

When making inference on the probability structure of the random field based on what we observe, a simplifying assumption often made is the stationarity assumption. Stationarity in simple terms means that the random field looks similar in different parts of the domain. There are different kinds of stationarity.

A random field Z is said strictly stationary if for all $\boldsymbol{s}_1, ..., \boldsymbol{s}_n$ and any $\boldsymbol{h} \in \mathbb{R}^d$, the joint distribution of $Z(\boldsymbol{s}_1), ..., Z(\boldsymbol{s}_n)$ is identical with the joint distribution of $Z(\boldsymbol{s}_1 + \boldsymbol{h}), ..., Z(\boldsymbol{s}_n + \boldsymbol{h})$, i.e.,

$$Pr(Z(\boldsymbol{s}_1) \leq z_1, ..., Z(\boldsymbol{s}_n) \leq z_n) = Pr(Z(\boldsymbol{s}_1 + \boldsymbol{h}) \leq z_1, ..., Z(\boldsymbol{s}_n + \boldsymbol{h}) \leq z_n), \tag{2.1}$$

where $z_1, ..., z_n \in \mathbb{R}$.

That is the probability law of a strictly stationary process is invariant under a shift in space. A lighter type of stationarity is the weak stationarity. A random field Z is weakly stationary (WS) if: $E(Z(\boldsymbol{s})) = \mu$ and $Cov(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)) = \mathrm{C}(\boldsymbol{h})$ with $\boldsymbol{h} = \boldsymbol{s}_2 - \boldsymbol{s}_1$.

Thus a random field whose mean does not depend on the spatial location and whose covariance is a function of the separation lag $\boldsymbol{h}$, is a WS process. $\mathrm{C}(\boldsymbol{h})$ is called the covariance function of $Z(\boldsymbol{s})$. For a WS random field Z, the correlation between $Z(\boldsymbol{s}_1)$ and $Z(\boldsymbol{s}_2)$ is defined as:

$$Corr(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)) = \frac{\mathrm{C}(\boldsymbol{h})}{\mathrm{C}(\boldsymbol{0})} = \rho(\boldsymbol{h}).$$

Strict stationarity (if moments of second order exist) implies WS while the reverse is not true. A common hypothesis regarding the finite-dimensional distribution of the random field Z, is Gaussianity.

Definition 3. *$Z(\boldsymbol{s})$ is called Gaussian process if for all n and admissible $\boldsymbol{s}_1, ..., \boldsymbol{s}_n$, the joint distribution of $Z(\boldsymbol{s}_1), ..., Z(\boldsymbol{s}_n)$ is multivariate normal.*

A multivariate normal distribution is characterized by its mean and covariance matrix, so the first two moments of a Gaussian process completely specify its probability structure. Thus for Gaussian processes, WS implies strict stationarity. Last kind of stationarity regards the increments of the process.

A random field Z is intrinsic stationary (IS) if: $E(Z(\boldsymbol{s}_1)) = E(Z(\boldsymbol{s}_2))$ and $Var(Z(\boldsymbol{s}_1) - Z(\boldsymbol{s}_2)) = 2\gamma(\boldsymbol{h})$. The function $2\gamma(\boldsymbol{h})$ is called variogram. IS is a weaker property than WS. If the process is WS, it is easy to verify that:

$$Var(Z(\boldsymbol{s}_1) - Z(\boldsymbol{s}_2)) = 2\mathrm{C}(\mathbf{0}) - 2\mathrm{C}(\boldsymbol{h})$$

and so $\gamma(\boldsymbol{h}) = \mathrm{C}(\mathbf{0}) - \mathrm{C}(\boldsymbol{h})$. Conversely, in general IS does not imply weak stationarity. For instance if Z is the standard Brownian motion in one dimension, the variogram function is $Var(Z(\boldsymbol{s}_1) - Z(\boldsymbol{s}_2)) = |\boldsymbol{h}|$. However, we can not recover the covariance function since the variogram is unbounded as $\boldsymbol{h}$ tends to infinity.

Gaussian processes play a central role in modeling spatial data. The advantages of the Gaussian process assumption are obvious: it allows convenient distribution theory (for instance, conditional distributions are easily obtained from the joint distributions). Gaussian processes have a rich, detailed and very well understood general theory. Furthermore, in most applications, we observe a single realization of the process at a finite set of locations. It is not easy to criticize a Gaussian assumption since we only have a sample size of one from a finite dimensional distribution. Nevertheless, there are situations in which it is more appropriate to use other processes to model spatial data.

2.3 Covariance functions and variograms properties

For a real-valued mapping $\mathrm{C} : \mathbb{R}^d \to \mathbb{R}$, necessary conditions for C to be a covariance function are:

(i) $\mathrm{C}(\mathbf{0}) \geq 0$;

(ii) $\mathrm{C}(\boldsymbol{h}) = \mathrm{C}(-\boldsymbol{h})$, i.e., C is an even function;

(iii) $\mathrm{C}(\mathbf{0}) \geq |\mathrm{C}(\boldsymbol{h})|$;

(iv) If $\mathrm{C}_j(\boldsymbol{h})$ are valid covariance function for $j = 1 \dots k$ then $\sum_{j=1}^{k} b_j \mathrm{C}_j(\boldsymbol{h})$ is a valid covariance function, if $b_j \geq 0, \quad \forall j$;

(v) If $\mathrm{C}_j(\boldsymbol{h})$ are valid covariance function for $j = 1 \dots k$ then $\prod_{j=1}^{k} \mathrm{C}_j(\boldsymbol{h})$ is a valid covariance function;

(vi) If $\mathrm{C}(\boldsymbol{h})$ is a valid covariance function in $\mathbb{R}^d$, then it is also a valid covariance function in $\mathbb{R}^p$, $p \leq d$.

Analogous properties of the variogram are:

(i) $\gamma(\mathbf{0}) = 0$;

(ii) $\gamma(\boldsymbol{h}) = \gamma(-\boldsymbol{h})$, i.e., γ is an even function;

(iii) $\gamma(\boldsymbol{h}) \geq 0$;

(iv) If $\gamma_j(\boldsymbol{h})$ are valid variograms for $j = 1 \ldots k$ then $\sum_{j=1}^{k} b_j \gamma_j(\boldsymbol{h})$ is a valid variogram,, if $b_j \geq 0, \quad \forall j$;

Valid (or permissible) covariance functions means that they must respect some mathematical constraints.

Indeed, one cannot define a spatial covariance or variogram function in a totally arbitrary way. The key property which has to satisfy is semi-positive definiteness. For a WS random field Z with covariance function $C(.)$, it means that:

$$\sum_{i}^{n} \sum_{j}^{n} a_i a_j \mathrm{C}(\boldsymbol{s}_i - \boldsymbol{s}_j) \geq 0 \tag{2.2}$$

for any set of $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ and all real $a_1, \ldots, a_n$.

The positive semi-definite condition is necessary for the existence of a random field with finite second moments. This condition guarantees that the variance of spatial predictions is non-negative. This simply follows noting that (2.2) is $Var(\sum_i^n a_i Z(\boldsymbol{s}_i))$.

On the other hand, if C is positive semi-definite, there exists a Gaussian random field with covariance matrix K and mean $E(Z(\boldsymbol{s})) = m < \infty$. Thus, positive definiteness is a necessary and sufficient condition for a covariance function. Bochner's theorem (Bochner, 1933) provides necessary and sufficient conditions for a covariance function $\mathrm{C}(\boldsymbol{h})$ of a WS process to be positive semi-definite.

Theorem 1. *(Bochner's Theorem). For a real-valued WS process on* $\mathbb{R}^d$, $\mathrm{C}(\boldsymbol{h})$ *is positive semidefinite if and only if it can be represented as:*

$$\mathrm{C}(\boldsymbol{h}) = \int e^{i\boldsymbol{\omega}^T \boldsymbol{h}} dF(\boldsymbol{\omega}) \tag{2.3}$$

where F *is a positive, symmetric, and finite measure and is called the spectral measure of* $\mathrm{C}(\boldsymbol{h})$*. If* F *is absolutely continuous with respect to Lebesgue measure, i.e.,* $dF(\boldsymbol{\omega}) = f(\boldsymbol{\omega})d\boldsymbol{\omega}$*,* $f(\boldsymbol{\omega})$ *is called the spectral density.*

Analogously to the covariance function, the variogram must respect some conditions to be permissible. Specifically, for any set of $\boldsymbol{s}_1,\dots,\boldsymbol{s}_n$ and any set of real $a_1,\dots,a_n$ such that $\sum_i^n a_i = 0$,

$$\sum_i^n \sum_j^n a_i a_j \gamma(\boldsymbol{s}_i - \boldsymbol{s}_j) \leq 0. \tag{2.4}$$

This follows by noting:

$$\sum_i^n \sum_j^n a_i a_j \gamma(\boldsymbol{s}_i - \boldsymbol{s}_j) = -E(\sum_i^n (a_i Z(\boldsymbol{s}_i))^2).$$

2.4 Isotropy

A random field is said isotropic if its covariance function $\mathrm{C}(\boldsymbol{h})$ only depends on $r = ||\boldsymbol{h}|| \geq 0$, where $||.||$ indicates the Euclidean distance. The class of all valid continuous covariance functions on $\mathbb{R}^d$ can be characterized by the Fourier transforms of all finite positive measures on $\mathbb{R}^d$ (Bochner's Theorem). There is an analogous characterization for isotropic covariance functions. Specifically, Schoenberg (1938) characterized the class Φ_d of isotropic covariance models on $\mathbb{R}^d$ as scale mixtures of the characteristic functions of random vectors uniformly distributed on the spherical shell of $\mathbb{R}^d$, with any *positive measure*, G:

$$\mathrm{C}(r) = 2^{\frac{d-2}{2}} \Gamma(\frac{d}{2}) \int_0^\infty (\boldsymbol{\omega} r)^{-\frac{d-2}{2}} J_{\frac{d-2}{2}}(\boldsymbol{\omega} r) dG(\boldsymbol{\omega}) \tag{2.5}$$

Here J_k is the Bessel function of the first kind of order k and the measure $G(.)$ is nondecreasing bounded in $\mathbb{R}^+$ and $G(\mathbf{0}) = 0$.

A general form for an isotropic covariance function is:

$$c(r, \boldsymbol{\theta}) = \begin{cases} \sigma^2 \rho(r, \boldsymbol{\theta}), & r > 0 \\ \tau + \sigma^2, & r = 0 \end{cases} \tag{2.6}$$

where τ is the nugget parameter. This parameter describes the behavior of the covariance at the origin. A phenomenon quite common in applications is that the variogram is discontinuous at the origin. This is due to microscale variability (variability of a spatial process operating at lag distances shorter than the smallest lag observed in the data) or/and measurement error. In geostatistical literature τ is the nugget, $\sigma^2 + \tau$ is the sill and σ^2 is the partial sill or variance.

In (2.6) $\rho(r, \boldsymbol{\theta})$ is a parametric correlation function which depends on $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$. Typically parametric correlation models depend on few parameters. In Chapter 2 we review the covariance models that we study in this Thesis. Specifically Matern and Generalized Wendland covariance models.

2.5 Spatial continuity and differentiability

Continuity and differentiability of a random fields are important since they are informative about the structure and the smoothness properties of the random field.

Assume the process have zero mean and finite second-order moments.

Definition 4. *A process $Z(\boldsymbol{s})$ is L_2 continuous at $\boldsymbol{s}_0$ if and only if* $\lim_{\boldsymbol{s}\longrightarrow\boldsymbol{s}_0} E[Z(\boldsymbol{s})-Z(\boldsymbol{s}_0)]^2 = 0$

Continuity in L_2 sense is also referred to as mean square continuity and will be denoted by $Z(\boldsymbol{s}) \overset{L_2}{\rightarrow} Z(\boldsymbol{s}_0)$

Definition 5. *A process $Z(\boldsymbol{s})$ is almost surely continuous at $\boldsymbol{s}_0$ if $Z(\boldsymbol{s}) \longrightarrow Z(\boldsymbol{s}_0)$ a.s. as $\boldsymbol{s} \longrightarrow \boldsymbol{s}_0$. If the process is almost surely continuous for every $\boldsymbol{s}_0 \in \mathbb{R}^d$ then the process is said to have continuous realizations.*

In general, one form of continuity does not imply the other since one form of convergence does not imply the other. However, if $Z(\boldsymbol{s})$ is a bounded process then *a.s.* continuity implies L_2 continuity. Of course, each implies $Z(\boldsymbol{s}) \overset{P}{\rightarrow} Z(\boldsymbol{s}_0)$. It is easy to show that for a WS random field, mean square continuity at $\boldsymbol{s}$ implies that:

$$\lim_{\boldsymbol{h}\longrightarrow\boldsymbol{0}} E[Z(\boldsymbol{s})-Z(\boldsymbol{s}+\boldsymbol{h})]^2 = 0$$

Thus, it is easily shown that for a WS random field mean square continuity is equivalent to the covariance function $\mathrm{C}(\boldsymbol{h})$ being continuous at $\boldsymbol{0}$. That is mean square continuity can be verified through the behavior of the covariance function near 0. As explained in section 2.4 some processes appear to have a variogram for which $\gamma(\boldsymbol{h}) \longrightarrow c > 0$ as $\boldsymbol{h} \longrightarrow \boldsymbol{0}$, *i.e* the nugget effect.

Means square continuity by itself does not convey much about the smoothness of the process and how it is related to the covariance function. The smoothness concept is brought into focus by studying the partial derivatives of the random field and introducing the mean square differentiability.

In $\mathbb{R}$ we can define the process:

$$Z_\delta(t) = \frac{Z(t+\delta)-Z(t)}{\delta}$$

with $t,\delta \in \mathbb{R}$ and we say that the process $Z(t)$ is mean square differentiable if the process $Z_\delta(t)$ converges in L_2. More formally, we have the following definition.

Definition 6. *A process $Z(t)$ on R is (mean square) differentiable if there exists a process $Z'(t)$ such that the following holds:* $\lim_{h\longrightarrow 0} E[Z_h(t) - Z'(t)]^2 = 0$

There exist another form of differentiability for a sample path. Differentiability sample path means that the partial derivatives of the sample paths are continuous.

Definition 7. *A process $Z(t)$ has differentiable sample path with probability one in $D \subset \mathbb{R}^d$ if for every sequence $(t_n)_{n\in\mathbb{N}}$ for such that $||t_n - t|| \to 0$ as $n \to \infty$, we have*

$$P\left(\omega : |\dot{Z}_j(t_n,\omega) - \dot{Z}_j(t,\omega)| \to 0 \quad as \quad n \to \infty \quad \forall j = 1,\ldots,n, \forall t \in D\right) = 1 \tag{2.7}$$

with $\dot{Z}_j(t,\omega) = \frac{\partial Z(t)}{\partial t_j}$.

Generalizing the definition 6, we obtain

Definition 8. *A process $Z(t)$ is mean square differentiable in $D \subset \mathbb{R}^d$ if for every sequence $(t_n)_{n\in\mathbb{N}}$ for such that $||t_n - t|| \to 0$ as $n \to \infty$, then*

$$E\left[|\dot{Z}_j(t_n) - \dot{Z}_j(t)|^2\right] \to 0 \quad as \quad n \to \infty \quad \forall j = 1,\ldots,n, \forall t \in D. \tag{2.8}$$

with $\dot{Z}_j(t,\omega) = \frac{\partial Z(t)}{\partial t_j}$.

It can be shown that the stationary process $Z_\delta(h)$ has the covariance function $C_\delta(h)$ such that $\lim_{h\longrightarrow 0} C_\delta(h) = -C''(h)$ provided $C(h)$ is twice differentiable. This also shows that $-C''(h)$ is positive definite. Stein (1999a) proves that $Z(t)$ is m-times mean square differentiable if and only if $[\frac{d^{2m}C(h)}{dh^{2m}}]_0$ exists and is finite. That is there is a strong relation between the mean square differentiability of a process and the derivative of its covariance function.

From this point of view it can be possible to make the selection of a particular correlation function based upon theoretical considerations. This possibility arises from the powerful fact that the choice of a correlation function determines the smoothness of realizations from the spatial process. In this sense Stein (1999a) recommends the Matérn class as a general tool for building spatial models since a parameter can control the degree of smoothness.

2.6 Kriging

The main goal in spatial statistics is often interpolation. There exists different kinds of interpolators but kriging presents relevant advantages. For instance it provides some measure of the accuracy of the prediction with respect to deterministic interpolator such as splines or inverse distance method.

Kriging is a geostatistical interpolation technique that considers both the distance and the degree of variation between known data points when predicting values in unknown location points. A kriged predictin is a weighted linear combination of the known sample values around the point to be predicted. Applied properly, kriging allows the user to derive weights that result in optimal and unbiased preditions. Let us consider a random field $Z(\boldsymbol{s})$, $\boldsymbol{s} \in D \subset \mathbb{R}^d$ and the linear model

$$Z(\boldsymbol{s}) = \mu(\boldsymbol{s}) + \varepsilon(\boldsymbol{s}) \quad \boldsymbol{s} \in D, \tag{2.9}$$

where $\mu(\boldsymbol{s})$ is a deterministic function and $\varepsilon(\boldsymbol{s})$ is a WS random process. We observe the process at n different locations, Z= $(Z(\boldsymbol{s}_1)\dots Z(\boldsymbol{s}_n))$, and wish to predict the process Z at an unobserved location $\boldsymbol{s}_0$. Let us denote with $p(\boldsymbol{s}_0, Z)$ the kriging interpolator at $\boldsymbol{s}_0$. The main properties of this object are:

- $p(\boldsymbol{s}_0, Z) = \lambda_0 + \sum_{i=1}^{n} \lambda_i Z(\boldsymbol{s}_i)$.
- $E(p(\boldsymbol{s}_0, Z)) = \mu(\boldsymbol{s}_0)$.
- $E((Z(\boldsymbol{s}_0) - p(\boldsymbol{s}_0, Z)^2)$ is minimum.

That is, kriging interpolator is optimal in the class of linear interpolator, *i.e* it is unbiased and with minimum variance error. Relaxing the first condition, it is easy to show that the best interpolator is the conditional expectation of $Z(\boldsymbol{s}_0)$ given the observed data:

$$p(\boldsymbol{s}_0, Z) = E(Z(\boldsymbol{s}_0)|Z) \tag{2.10}$$

However, in general $p(\boldsymbol{s}_0, Z)$ is not a linear function of the data and establishing the statistical properties of the best predictor under squared-error loss can be difficult. Fortunately, if the random field is Gaussian the best linear interpolator is also the best interpolator. Kriging appears in many forms and flavors, distinguished by whether the mean is known or not, what the distribution of it, whether predictions are made for points or areas and so forth. Here we describe classical ordinary kriging.

Let be $\mu(\boldsymbol{s}) = \mu$ and $\varepsilon(\boldsymbol{s})$ a zero mean random field with associated covariance matrix C that we assume known and let c the covariance between $Z(\boldsymbol{s}_0)$ and Z. We consider linear predictor of the form $p(\boldsymbol{s}_0, Z) = \lambda_0 + \lambda^T Z$, where $\lambda = (\lambda_1 \dots \lambda_N)$. Since we are looking for unbiased interpolator it is easy to show that it is equivalent to set $\lambda_0 = 0$ and $\sum_i^N \lambda_i = 1$.

Thus the problem now is to choose the λ weights that minimize:

$$E((\lambda^T Z - Z(\boldsymbol{s}_0))^2) \qquad \text{subject to} \quad \sum_i^N \lambda_i = 1.$$

This can be accomplished as an unconstrained minimization problem introducing Lagrange multiplier m:

$$\underset{\lambda}{\operatorname{argmin}} E((\lambda^T Z - Z(\boldsymbol{s}_0))^2) - 2m(\sum_i^N \lambda_i - 1). \tag{2.11}$$

It can be shown that (Cressie, 1993) the solution to this problem is:

$$\lambda^T = c + \mathbf{1}(\frac{1 - \mathbf{1}^T C^{-1} c}{\mathbf{1}^T C^{-1} \mathbf{1}})^T \mathrm{C}^{-1} \quad \text{and } m = \frac{1 - \mathbf{1}^T C^{-1} c}{\mathbf{1}^T C^{-1} \mathbf{1}}.$$

Thus, the optimal linear predictor is

$$p(\boldsymbol{s}_0, Z)_{OK} = \hat{\mu} + c^T C^{-1}(Z - \mathbf{1}\hat{\mu}),$$

with $\hat{\mu} = \frac{C^{-1}Z}{\mathbf{1}^T C^{-1} \mathbf{1}}$, and the minimized mean-square prediction error, *i.e*. the ordinary kriging variance is:

$$\sigma^2(\boldsymbol{s}_0)_{OK} = \mathrm{C}(0) - c^T C^{-1} c + \frac{1 - \mathbf{1}^T C^{-1} c}{\mathbf{1}^T C^{-1} \mathbf{1}}.$$

When μ is known ordinary kriging is called simple kriging. In this case the the optimal linear predictor simplifies to

$$p(\boldsymbol{s}_0, Z)_{SK} = \mu + c^T C^{-1}(Z - \mathbf{1}\mu), \tag{2.12}$$

with associated mean square error

$$\sigma^2(\boldsymbol{s}_0)_{SK} = \mathrm{C}(0) - c^T C^{-1} c. \tag{2.13}$$

Cressie (1993) discusses more complicated versions, such as lognormal and trans-Gaussian kriging, and universal kriging, used in a presence of non-stationary mean field model.

2.7 Equivalence of Gaussian measures

In this section, we present some theory of equivalence of Gaussian measures. Equivalence of two measures is defined as:

Definition 9. *Let $(\Omega, \mathscr{F})$ be a measurable space and P_i, i=1,2, be two probability measures on $\mathscr{F}$. Then,*

- *two probability measures P_1 and P_2 are equivalent ($P_1 \equiv P_2$) means for all $A \in \mathscr{F}$, $P_1(A) = 0$ if and only if $P_2(A) = 0$.*
- *two probability measures P_1 and P_2 are orthogonal ($P_1 \perp P_2$) if there exists $A \in \mathscr{F}$ such that $P_1(A) = 0$ and $P_2(A) = 1$.*

Different conditions have been established in the literature in order to verify the equivalence of two Gaussian measures (Skorokhod and Yadrenko (1973), Ibragimov and Rozanov (1978)). Some of these conditions are based on likelihood ratios or on spectral densities. Using both criteria some equivalence of Gaussian measures have been established when focusing on a specific class of covariance functions. Specifically, let $P_i = P_i^{(m_i,K_i)}$, $i = 1,2$, two probability measures such that under P_i, the process $Z(\boldsymbol{s})$, $\boldsymbol{s} \in \mathbb{R}^d$, is stationary Gaussian with a second-order structure $(0, K_i)$, where

$$K_i(||\boldsymbol{h}||;\sigma_i,\theta_i) = \sigma_i^2 e^{(-||\boldsymbol{h}||\theta_i)} \quad \boldsymbol{h} \in \mathbb{R}^d, d \geq 1.$$

Then, for any bounded infinite set $\boldsymbol{T} \subset \mathbb{R}^d$, $P_1^{(0,K_1)} \equiv P_2^{(0,K2)}$ on the paths of $Z(\boldsymbol{s})$, $\boldsymbol{s} \in \boldsymbol{T}$, if $\sigma_1^2\theta_1 = \sigma_2^2\theta_2$ (Stein, 1999a, 2004).

Zhang (2004) extended the previous result to the Matérn class, defined as

$$K(||\boldsymbol{h}||;\boldsymbol{\theta}) = \frac{\sigma^2(\alpha||\boldsymbol{h}||)^{\upsilon}}{\Gamma(\upsilon)2^{\upsilon-1}} K_{\upsilon}(\alpha||\boldsymbol{h}||)$$

,

with $\boldsymbol{\theta} = (\sigma^2, \alpha, \upsilon)^T$ and where σ^2, α >0, and $\upsilon > 0$ are parameters and K_υ is the modified Bessel function of order υ. In particular, the parameter υ allows to parameterize the behaviour of the correlation at the origin.

Specifically Zhang (2004) shows that given two Gaussian measures P_i, $i = 1,2$, with second-order structures $(0, K_i(||\boldsymbol{h}||, \boldsymbol{\theta}_i))$, where $K_i(||\boldsymbol{h}||, \boldsymbol{\theta}_i)$ is a Matérn covariance, are equivalent for a stationary and isotropic process $Z(\boldsymbol{s})$, $\boldsymbol{s} \in \boldsymbol{T} \subset \mathbb{R}^d$, $d = 1,2,3$ observed in a bounded set $\boldsymbol{T}$, if and only if

$$\sigma_2^2\alpha_2^{2\upsilon} = \sigma_1^2\alpha_1^{2\upsilon}. \tag{2.14}$$

This result has an important consequence from estimation point of view. In fact, Zhang (2004) shows that for a fixed υ, σ^2 and α cannot be consistently estimated, under infill asymptotics.

In order to describe the link between the asymptotic behavior of the kriging and the equivalence of measures, let us consider two possible measures P_j, $j = 1,2$, for a process $Z(t)$ with a second-order structure (m_j, K_j), j=1,2, and suppose to predict $Z(t_0)$, $t_0 \in \mathbb{R} \setminus Q$, where $Q = \{t_k, k = 1,2,...\}$. Define $e_j(t_0,Q) = Z(t_0) - E_j(Z(t_0)|Z(t), t \in Q)$, $j = 1,2$, to be the error of the best linear predictor of $Z(t_0)$ under P_j. To prove how good predictions based on K_2 are when K_1 is the correct covariance function, it is necessary to compare the two mean squared errors by evaluating the ratio $\frac{E_1[(e_2(t_0,Q))^2]}{E_1[(e_1(t_0,Q))^2]} = 1 + \frac{E_1[(e_2(t_0,Q)-e_1(t_0,Q))^2]}{E_1[(e_1(t_0,Q))^2]}$. There is another measure for how good predictions given K_2 which is the ratio $\frac{E_2[(e_2(t_0,Q))^2]}{E_1[(e_1(t_0,Q))^2]}$. The following theorem establishes a relationship between two equivalent probability measures and the ratio mentioned above (Stein, 1999a),(Porcu et al., 2012, Chap 10, page 242).

Theorem 2. *Let $K_i = K(x;\alpha_i,\sigma_i^2)$, i=1,2, be covariance function associated with probability measures P_i. Let Z(s), $s \in D$, and the set of sampling sites $Q = \{s_k, k = 1,2,...\}$ is dense in D, be a Gaussian process with mean m_i and covariance function K_i under P_i, i=1,2. Then, for a bounded set $D \subset \mathbb{R}^d$, if $P_1 \equiv P_2$ on the paths of $Z(s)$, $s \in D$, then uniformly in $s \in D$ such that $E_1[(e_1(t_0,Q))^2] > 0$,*

$$\lim_{n\to\infty} \frac{E_1[(e_2(t_0,Q))^2]}{E_1[(e_1(t_0,Q))^2]} = 1 \textit{ and } \lim_{n\to\infty} \frac{E_2[(e_2(t_0,Q))^2]}{E_1[(e_1(t_0,Q))^2]} = 1.$$

Theorem 2 implies that when the spatial location sites are observed on a compact set, the asymptotic mean squared error of the prediction is the same if the two Gaussian measure are equivalent. Thus finding conditions of equivalence for Gaussian measures is important from the prediction point of view.

Chapter 3

Classes of covariance functions

In this section we review three important covariance models. Specifically the Matérn, the Generalized Wendland and the Generalized Cauchy model. Then we elucidate the connections of these three covariance models with the fractal dimension and Hurst coefficient of a Gaussian process.

3.1 Matérn family

The Matérn family has been very popular in spatial statistics for its flexibility with respect to continuous parameterization of smoothness, in the mean square sense, of the underlying Gaussian field.

The Matérn function, defined as

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

is a member of the class Φ_∞ for any positive values of α and ν. Here, $\mathcal{K}_\nu$ is a modified Bessel function of the second kind of order ν, σ^2 is the variance and α a positive scaling parameter (we use the abuse of notation $\mathcal{M}_\nu$ for $\mathcal{M}_{\nu,1,1}$). The parameter ν 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 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 ν tending to infinity, a rescaled version of the Matérn converges to a Gaussian 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.

3.2 Generalized Wendland (GW) family

A family of piecewise polynomial functions with compact support provide another interesting class of covariance functions. We also define Φ_d^b as the class that consists of members of Φ_d being additionally compactly supported on a given interval, $[0,b]$, $b>0$. Clearly, their radial versions are compactly supported over balls of $\mathbb{R}^d$ with radius b.

A celebrated family from the class Φ_d^1 is the Askey (Askey, 1973) family of functions, $\mathscr{A}_\mu : [0,\infty) \to \mathbb{R}$, defined by

$$\mathscr{A}_\mu(r) = (1-r)_+^\mu, \qquad \mu > 0,$$

where $(\cdot)_+$ denotes the positive part. Arguments in Golubov (1981) show that $\mathscr{A}_\mu \in \Phi_d^1$ if and only if $\mu \geq (d+1)/2$.

We now have the ingredients to define GW functions $\varphi_{\mu,\kappa}$ as introduced by Gneiting (2002b) and Zastavnyi (2006). For $\kappa > 0$, we define

$$\varphi_{\mu,\kappa}(r) := \frac{1}{B(2\kappa,\mu+1)} \int_r^\infty u(u^2-r^2)^{\kappa-1} \mathscr{A}_\mu(u)\,\mathrm{d}u, \quad 0 \leq r \leq 1, \tag{3.1}$$

with B denoting the beta function. Arguments in Gneiting (2002b) and Zastavnyi (2006) show that, for a given $\kappa > 0$, $\varphi_{\mu,\kappa} \in \Phi_d^1$ if and only if

$$\mu \geq \lambda(d,\kappa) := (d+1)/2 + \kappa. \tag{3.2}$$

Throughout, we use λ instead of $\lambda(d,\kappa)$ whenever no confusion arises. Integration by parts shows that (3.1) can be written as

$$\varphi_{\mu,\kappa}(r) = \frac{1}{B(1+2\kappa,\mu)} \int_r^\infty (u^2-r^2)^\kappa \mathscr{A}_{\mu-1}(u)\,\mathrm{d}u, \qquad 0 \leq r \leq 1.$$

Note that $\varphi_{\mu,0}$ is not defined because κ must be strictly positive, so that in this special case we define $\varphi_{\mu,0} := \mathscr{A}_\mu$.

Finally, we define the GW covariance function, with compact support $\beta > 0$, variance σ^2 and smoothness parameter $\kappa \geq 0$ as

$$\varphi_{\mu,\kappa,\beta,\sigma^2}(r) := \sigma^2 \varphi_{\mu,\kappa}(r/\beta), \quad 0 \leq r \leq \beta. \tag{3.3}$$

Table 3.1 GW correlation $\varphi_{\mu,\kappa}(r)$ and Matérn correlation $\mathcal{M}_\nu(r)$ with increasing smoothness parameters κ and ν. $SP(k)$ means that the sample paths of the associated Gaussian field are k times differentiable.

κ	$\varphi_{\mu,\kappa}(r)$	ν	$\mathcal{M}_\nu(r)$	$SP(k)$
0	$(1-r)_+^{\mu}$	0.5	e^{-r}	0
1	$(1-r)_+^{\mu+1}(1+r(\mu+1))$	1.5	$e^{-r}(1+r)$	1
2	$(1-r)_+^{\mu+2}(1+r(\mu+2)+r^2(\mu^2+4\mu+3)\frac{1}{3})$	2.5	$e^{-r}(1+r+\frac{r^2}{3})$	2
3	$(1-r)_+^{\mu+3}(1+r(\mu+3)+r^2(2\mu^2+12\mu+15)\frac{1}{5}$ $+r^3(\mu^3+9\mu^2+23\mu+15)\frac{1}{15})$	3.5	$e^{-r}(1+\frac{r}{2}+r^2\frac{6}{15}+\frac{r^3}{15})$	3

and $\varphi_{\mu,\kappa,\beta,\sigma^2} \in \Phi_d^{\beta}$ for $\mu \geq \lambda$. Accordingly, we define

$$\varphi_{\mu,0,\beta,\sigma^2}(r) := \sigma^2 \mathscr{A}_\mu(r/\beta)(r), \quad 0 \leq r \leq \beta, \tag{3.4}$$

When computing (3.3), numerical integration is obviously feasible, but could be cumbersome to (spatial) statisticians used to handle closed form parametric covariance models. Nevertheless, closed form solutions of the integral in Equation (3.1) can be obtained when $\kappa = k$, a positive integer. In this case, $\varphi_{\mu,k}(r) = \mathscr{A}_{\mu+k}(r)P_k(r)$, 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 (3.1) can be obtained when $\kappa = k + 0.5$, using some results in Schaback (2011). Such solutions are called *missing* Wendland functions.

Recently, Porcu et al. (2016) have shown that the GW class includes almost all celebrated classes of covariance functions with compact supports known to the geostatistical and numerical analysis communities. Not only original and Wendland functions, but also Wu's functions (Wu, 1995), which in turn include the spherical model (Wackernagel, 2003), as well as the celebrated Trigub splines (Zastavnyi, 2006). Finally, Chernih et al. (2014) show that, for κ tending to infinity, a rescaled version of the GW model converges to a Gaussian model.

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

3.3 Generalized Cauchy covariance

A celebrated class of members of Φ_∞ is the generalized Cauchy class (Gneiting and Schlather, 2004) $\mathscr{C}_{\delta,\lambda,\gamma,\sigma^2} : [0,\infty) \to \mathbb{R}$, defined as

$$\mathscr{C}_{\delta,\lambda,\gamma,\sigma^2}(r) = \sigma^2 \left(1 + (r/\gamma)^\delta\right)^{-\lambda/\delta} \qquad r \geq 0, \tag{3.5}$$

where the conditions $\delta \in (0,2]$ and $\lambda > 0, \gamma > 0, \sigma^2 > 0$ are necessary and sufficient for $\mathscr{C}_{\delta,\lambda,\gamma,\sigma^2}$ belonging to the class Φ_∞. The parameter δ is crucial for the differentiability at the origin and, as a consequence, for the differentiability of the associated sample paths. Specifically, for $\delta = 2$, they are infinitely times differentiable and they are not differentiable for $\delta \in (0,2)$.

3.4 Fractal dimension and Hurst coefficient

The smoothness of a Gaussian fields can also be described via the Hausdorff or fractal dimension of a sample path. The fractal dimension $D \in [d, d+1)$ is a measure of the roughness for non-differentiable Gaussian fields and higher values indicating rougher surfaces. For a given covariance function $\phi \in \Phi_d$ if $1 - \phi(r) \sim r^\chi$ as $r \to 0$ for some $\chi \in (0,2]$ then the sample paths of the associated random field have fractal dimension $D = d + 1 - \chi/2$. Here χ is the so called fractal index that governs the roughness of sample paths of a random field.

In the case of a Matérn model $\chi = 2\nu$ so $D = d + 1 - \nu$ if $0 < \nu < 1$ and d otherwise (Adler, 1981). Thus the Matérn model permit the full range of allowable values for the fractal dimension. In the case of Generalized Wendland class the fractal index χ is equal to $2\kappa + 1$, so that in this case $D = d + 0.5 - \kappa$ if $0 \leq \kappa < 0.5$ and d otherwise. Thus the GW model does not allow to cover the full range of allowable values for the fractal dimension.

Long-memory dependence can be defined through the asymptotic behavior of the correlation function at infinity. Specifically, for a given covariance function $\phi \in \Phi_d$, if the power-law $\phi(r) \sim r^{-\varepsilon}$ as $r \to \infty$ holds for some $\varepsilon \in (0,1)$ the random field is said to have long memory with Hurst coefficient $H = 1 - \varepsilon/2$.

The generalized Cauchy class represents a breaking point with respect to earlier literature based on the dogmatic assumption of self similarity, since it decouples the fractal dimension and the Hurst effect. Specifically, the sample paths of the associated random field have fractal dimension $D = n + 1 - \delta/2$ and if $\lambda \in (0,1)$ the random field has long memory with Hurst coefficient $H = 1 - \lambda/2$. Thus, D and H may vary independently of each other.

Chapter 4

Fourier transform theory and complete monoton

In this section we establish some inequalities for estimation for Gaussian random field utilizing Bessel's inequality and Young's inequality.

4.1 Some inequalities

Let x and y be two positives numbers and integers $p,q > 1$ with $\frac{1}{p}+\frac{1}{q}=1$, then

$$\frac{1}{q}x^q + \frac{1}{p}y^p \geq xy. \tag{4.1}$$

The next result follows from Bessel's inequality.

Definition 10. *Let* $\mathbf{H}$ *be a Hilbert space and let* $\{e_1, e_2, \cdots\}$ *be an orthonormal sequence in* $\mathbf{H}$. *Then for any* $\boldsymbol{x} \in \mathbf{H}$ *we have*

$$\sum_{k=1}^{\infty} |\langle \boldsymbol{x}, e_k \rangle|^2 \leq ||\boldsymbol{x}||^2 \tag{4.2}$$

with $\boldsymbol{x} = \sum_{k=1}^{\infty} |\langle \boldsymbol{x}, e_k \rangle| e_k$.

4.2 Fourier transform

Spectral methods present an important tool for studying the spatial-temporal structure of random fields. In this work, the Fourier transform is a powerful technique to find the conditions of equivalence of two Gaussian random fields with different covariance families. Before, we introduce some definitions concerning the Fourier transform.

In his memory Fourier (1807) applied that all periodic functions can be decomposes into infinite sum of sinusoidal signals. For example, let f be a T-periodic function, then it becomes

$$f(x) = \sum_{n=-\infty}^{+\infty} a_n(f) e^{2in\pi x/T}, \tag{4.3}$$

where a_n is a complex Fourier series.

The following definition presents the Fourier transform of f for sufficiently large T.

Definition 11. *Let f be a square integrable function in $\mathbb{R}$. We define its Fourier transform as*

$$\hat{f}(z) = \int_{-\infty}^{+\infty} e^{-izx} f(x) \mathrm{d}x. \tag{4.4}$$

Note that if f is square integrable in $\mathbb{R}$ then so is $\hat{f}(z)$.

Definition 12. *Let $\hat{f}(z) \in \mathrm{L}^2(\mathbb{R})$. Then the inverse Fourier transform of $\hat{f}(z)$ can be defined as follows*

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{izx} \hat{f}(z) \mathrm{d}x. \tag{4.5}$$

Fourier transforms of radial versions of members of Φ_d, for a given d, have a simple expression, as reported in Yaglom (1987) or Stein (1999a). For a member ϕ of the class Φ_d, we define its isotropic spectral density as

$$\widehat{\phi}(z) = \frac{z^{1-d/2}}{(2\pi)^d} \int_0^{\infty} u^{d/2} J_{d/2-1}(uz) \phi(u) \mathrm{d}u \qquad z \geq 0, \tag{4.6}$$

and throughout the thesis, we use the notations $\widehat{\mathscr{M}}_{\nu,\alpha,\sigma^2}$ and $\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2}$ for the Fourier transforms of $\mathscr{M}_{\nu,\alpha,\sigma^2}$ and $\varphi_{\mu,\kappa,\beta,\sigma^2}$, respectively.

A well-known result about the spectral density of the Matérn model is the following:

$$\widehat{\mathscr{M}}_{\nu,\alpha,\sigma^2}(z) = \frac{\Gamma(\nu + d/2)}{\pi^{d/2}\Gamma(\nu)} \frac{\sigma^2 \alpha^d}{(1+\alpha^2 z^2)^{\nu+d/2}} \qquad z \geq 0. \tag{4.7}$$

For two given functions $g_1(x)$ and $g_2(x)$, with $g_1(x) \asymp g_2(x)$ we mean that there exist two constants c and C such that $0 < c < C < \infty$ and $c|g_2(x)| \leq |g_1(x)| \leq C|g_2(x)|$ for each x. The next result follows from Zastavnyi (2006), Chernih and Hubbert (2014), and from standard properties of Fourier transforms. Their proofs are thus omitted. Let us first define the function

${}_1F_2$ as

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

This function is a special case of the generalized hypergeometric functions ${}_qF_p$ (Abramowitz and Stegun, 1970), with

$$ (q)_k = \begin{cases} \frac{\Gamma(k+q)}{\Gamma(q)} & \text{for } k \geq 1 \\ 1 & \text{for } k = 0, \end{cases} $$

being the Pochhammer symbol.

Theorem 3. *Let $\varphi_{\mu,\kappa,\beta,\sigma^2}$ be the function defined at Equation* (3.3). *Then, for $\kappa, \sigma^2, \beta > 0$ and $\mu \geq \lambda$:*

1. $\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2}(z) = \sigma^2 L^{\boldsymbol{\varsigma}} \beta^d {}_1F_2\Big(\lambda; \lambda + \dfrac{\mu}{2}, \lambda + \dfrac{\mu}{2} + \dfrac{1}{2}; -\dfrac{(z\beta)^2}{4}\Big) \quad z > 0;$

2. $$\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2}(z) = \sigma^2 L^{\boldsymbol{\varsigma}} \beta^d \Big[c_3^{\boldsymbol{\varsigma}} (z\beta)^{-2\lambda} \{1 + \mathscr{O}(z^{-2})\} \\ + c_4^{\boldsymbol{\varsigma}} (z\beta)^{-(\mu+\lambda)} \big\{ \cos(z\beta - c_5^{\boldsymbol{\varsigma}}) + \mathscr{O}(z^{-1}) \big\} \Big] \text{ for } z \to \infty;$$

3. $\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2}(z) \asymp z^{-2\lambda}$ *for* $z \to \infty$,

where $c_3^{\boldsymbol{\varsigma}} = \frac{\Gamma(\mu+2\lambda)}{\Gamma(\mu)}$, $c_4^{\boldsymbol{\varsigma}} = \frac{\Gamma(\mu+2\lambda)}{\Gamma(\lambda)2^{\lambda-1}}$, $c_5^{\boldsymbol{\varsigma}} = \frac{\pi}{2}(\mu + \lambda)$, $L^{\boldsymbol{\varsigma}} = \frac{K^{\boldsymbol{\varsigma}}\Gamma(\kappa)}{2^{1-\kappa}B(2\kappa,\mu+1)}$ *and*

$$ K^{\boldsymbol{\varsigma}} = \frac{2^{-\kappa-d+1} \pi^{-\frac{d}{2}} \Gamma(\mu+1)\Gamma(2\kappa+d)}{\Gamma(\kappa + \frac{d}{2})\Gamma(\mu+2\lambda)}, $$

where $\boldsymbol{\varsigma} := (\mu, \kappa, d)'$.

Point 1 has been shown by Zastavnyi (2006). Points 2 and 3 can be found in Chernih and Hubbert (2014).

Note that the case $\kappa = 0$ is not included in Theorem 3. We consider it in the following result, whose proof follows the lines of Zastavnyi (2006) and Chernih and Hubbert (2014) for the case $\kappa > 0$.

Theorem 4. *Let $\mathscr{A}_{\mu,\beta,\sigma^2} = \varphi_{\mu,0,\beta,\mu,\sigma^2}$ as being defined at Equation* (3.4). *Then, for $\sigma^2, \beta > 0$, $\mu \geq (d+1)/2$:*

1. $$\widehat{\mathscr{A}}_{\mu,\beta,\sigma^2}(z) = \sigma^2 K^{\boldsymbol{\varsigma}} \beta^d \\ \times {}_1F_2\left(\frac{d+1}{2}; \frac{d+1}{2} + \frac{\mu}{2}, \frac{d+1}{2} + \frac{\mu}{2} + \frac{1}{2}; -\frac{(z\beta)^2}{4} \right) \quad z > 0;$$

2. $\widehat{\mathscr{A}_{\mu,\beta,\sigma^2}}(z) = \sigma^2 K^{\varsigma} \beta^d \big[c_3^{\varsigma}(z\beta)^{-(d+1)}\{1+\mathscr{O}(z^{-2})\}$
$+c_4^{\varsigma}(z\beta)^{-(\mu+(d+1)/2)}\{\cos(z\beta - c_5^{\varsigma}) + \mathscr{O}(z^{-1})\}\big]$ *for* $z \to \infty$;

3. $\widehat{\mathscr{A}_{\mu,\beta,\sigma^2}}(z) \asymp z^{-(d+1)}$ *for* $z \to \infty$,

with c_3^{ς}, c_4^{ς}, c_5^{ς} *and* K^{ς} *defined as in Theorem 3 but with* $\boldsymbol{\varsigma} := (\mu, 0, d)'$.

The spectral density and its decay for $z \to \infty$ in Theorems 3 and 4 are useful when studying some geometrical properties of a Gaussian field or its associated sample paths (Adler, 1981). For instance, using Theorem 3.1 or 4.1, it is easy to prove that for a positive integer k, the sample paths of a Gaussian field with GW function are k times differentiable if and only if $\kappa > k - 1/2$.

Table 3.1 compares the GW $\varphi_{\mu,\kappa}(r)$ for $\kappa = 0, 1, 2, 3$ with the Matérn correlation model for $\nu = 0.5, 1.5, 2.5, 3.5$ with the associated number of sample paths differentiability.

4.3 Completely monotonic function

The notion of complete monotonicity has been used one hundred years ago. It finds wide applications in areas ranging from probability and spatio-temporal theory. In this section we present some properties of completely monotonic functions and give an example that we need later.

4.3.1 Definitions and properties

Definition 13. *Let f be a non-negative function on $\mathbb{R}_+$. f is called completely monotonic if it is infinitely differentiable and*

$$(-1)^k f^{(n)}(x) \geq 0 \qquad \forall x > 0, \quad k \in \mathbb{N} \tag{4.8}$$

The following theorem gives an integral characterization of completely monotonic functions.

Theorem 5. *(Bernstein)*
A non negative function $f(x)$ on $(0,\infty)$ is completely monotone if and only if it is the Laplace transform of a finite non-negative Radon measure λ on $(0,\infty)$. We write

$$f(z) = \mathscr{L}(\lambda(t))(z) = \int_0^{\infty} e^{-zt} d\lambda(t). \tag{4.9}$$

Properties 1. *Let $f(x)$ and $g(x)$ be a completely monotonic functions on $(0,\infty)$. Then*

1. $f(z)+g(z)$ *is completely monotonic.*

2. $f(z)g(z)$ *is completely monotonic.*

3. $f(z)-f(z+h)$ *is completely monotonic for any* $h \geq 0$.

In 1938, Schoenberg established a relationship between positive definite radial and completely monotonic functions. The following theorem presents this connection.

Theorem 6. *A radial function* $f(\|\cdot\|^2)$ *is positive definite on* $\mathbb{R}^d$ *for any integer d if and only if f is completely monotone on* $(0,\infty)$

Note that if f is not constant, then $f(\|\cdot\|^2)$ is strictly positive definite and radial on $\mathbb{R}^d$ for any integer d.

4.3.2 Application

We show that some parameterized differences of Buhmann functions preserve positive definiteness in d-dimensional Euclidean spaces, and then we determine the exact level of smoothness induced by such an operation. The proof can be find in Porcu et al. (2016). A more detailed proof can be found in theorem (7).

Let $D \subset \mathbb{R}^d$ be a bounded subset of $\mathbb{R}^d$ and let $\boldsymbol{Z}_n = (Z(\boldsymbol{s}_1),\ldots,Z(\boldsymbol{s}_n))'$ be a finite realization of $Z(\boldsymbol{s})$, $\boldsymbol{s} \in D$, a zero mean stationary Gaussian field with a given parametric generalized Wendland covariance function $\sigma^2\phi(\|\cdot\|;\beta)$, with $\sigma^2 > 0$, β a scale parameter and ϕ a member of the Wendland class.

We then write $R_n(\beta) = [\phi(\|\boldsymbol{s}_i - \boldsymbol{s}_j\|;\beta)]_{i,j=1}^n$ for the associated correlation matrix.

From Horn and Johnson (1991) we have $R_n(\beta_1)^{-1}\beta_1^{-(1+2\kappa)} - R_n(\beta_2)^{-1}\beta_2^{-(1+2\kappa)}$ is positive semidefinite, is equivalent to tell that $\beta_2^{1+2\kappa}\widehat{\varphi}_{\mu,\kappa,\beta,1}(z)$ is an increasing function with respect to β, for all $z \in \mathbb{R}_+$.

From theorem 3 it is enough to show that $\beta^{2\lambda}\mathrm{H}_{\lambda,\mu,z}(\beta) = \beta^{2\lambda}{}_1F_2\Big(\lambda;\lambda+\frac{\mu}{2},\lambda+\frac{\mu}{2}+\frac{1}{2};-\frac{(z\beta)^2}{4}\Big)$ is an increasing function respect to β, for all $z \in \mathbb{R}_+$. From lemma 12 of Zastavnyi and Trigub (2002) we deduce the following theorem.

Theorem 7. *The following assertions are equivalent:*

1.
$$x^{2\lambda}\mathrm{H}_{\lambda,\mu,z}(x) \quad \textit{is an increasing function for} \quad x \in (0,\infty), \tag{4.10}$$

2.

$$x^{2\lambda}\mathrm{H}_{\lambda,\mu,1}(x) \quad \textit{is an increasing function for} \quad x \in (0,\infty), \tag{4.11}$$

3.

$$\Theta(x) = 2\lambda\mathrm{H}_{\lambda,\mu,1}(x) - \frac{\lambda x^2}{4(\lambda+\mu/2)(\lambda+\mu/2+1/2)}\mathrm{H}_{\lambda+1,\mu,1}(x) \quad \textit{is positive}, \tag{4.12}$$

4. The Laplace transform of $x^{2\lambda+\mu-1}\Theta(x)$ *is completely monotonic.*

It is sufficient to show the third point of theorem 7. In fact, let $\Psi_{\lambda,\mu} = \frac{2^{-\lambda-3/2}\pi^{1/2}\lambda\Gamma(2\lambda+\mu)}{\Gamma(\lambda)\Gamma(\mu)}$, then we write

$$2\lambda\mathrm{H}_{\lambda,\mu,1}(x) = \Psi_{\lambda,\mu}x^{-(2\lambda+\mu-1)}\int_0^x u^{\lambda-1/2}J_{\lambda-1/2}(u)(x-u)^{\mu-1}\mathrm{d}u,$$

and

$$\mathrm{H}_{\lambda+1,\mu,1}(x) = (2\lambda+2)^{-1}x^{-(2\lambda+\mu+1)}\Psi_{\lambda+1,\mu}\int_0^x u^{\lambda+1/2}J_{\lambda+1/2}(u)(x-u)^{\mu-1}\mathrm{d}u.$$

Applying the Laplace transform on $\Theta(x)$, we obtain,

$$\begin{aligned}
\mathscr{L}\left(x^{2\lambda+\mu-1}\Theta(x)\right)(s) &= (2\lambda)^{-1}\Psi_{\lambda,\mu}\mathscr{L}\left(x^{\mu-1}\right)(s)\mathscr{L}\left(x^{\lambda-1/2}J_{\lambda-1/2}(x)\right)(s) \\
&\quad - \frac{\lambda(2\lambda+2)^{-1}}{4(\lambda+\mu/2)(\lambda+\mu/2+1/2)}\Psi_{\lambda+1,\mu}\mathscr{L}\left(x^{\mu-1}\right)(s)\mathscr{L}\left(x^{\lambda+1/2}J_{\lambda+1/2}(x)\right)(s) \\
&= \lambda^{-1}\Psi_{\lambda,\mu}\left[\mathscr{L}\left(x^{\mu-1}\right)(s)\mathscr{L}\left(x^{\lambda-1/2}J_{\lambda-1/2}(x)\right)(s) - \frac{\lambda+1}{4\lambda}\mathscr{L}\left(x^{\mu-1}\right)(s)\right. \\
&\quad \left.\times\mathscr{L}\left(x^{\lambda+1/2}J_{\lambda+1/2}(x)\right)(s)\right].
\end{aligned}$$

Next, from Erdelyi et al. (1954) we have $\mathscr{L}\left(x^{2\lambda+\mu-1}\Theta(x)\right)(s)$ is completely monotonic if and only if $\mu > \lambda - 1$. The proof is completed.

Chapter 5

Estimation and prediction for Gaussian random fields under infill asymptotics using generalized Wendland covariance functions

In this chapter we characterize the equivalence of two Gaussian measures with GW covariance function, and to provide sufficient conditions for the equivalence of two Gaussian measures with Matérn and GW covariance functions. Then we establish strong consistency and asymptotic distribution of the maximum likelihood estimator of the microergodic parameter associated to GW covariance model, under fixed domain asymptotics. We compare the finite-sample behavior of the maximum likelihood estimator with the theoretical results. Furthermore, we establish asymptotic efficiency predition and asymptotically correct estimation of prediction variance when using two GW models and when using a GW and Matérn model. Finally, we compare the finite-sample behavior of the prediction and its associated mean square error when using two equivalent Gaussian measures with Matérn and GW covariance model, using covariance tapering as benchmark.

5.1 Equivalence of Gaussian measures with GW models

Equivalence and orthogonality of probability measures are useful tools when assessing the asymptotic properties of both prediction and estimation for Gaussian fields. Denote with P_i, $i = 0, 1$, two probability measures defined on the same measurable space $\{\Omega, \mathscr{F}\}$. P_0 and P_1 are called equivalent (denoted $P_0 \equiv P_1$) if $P_1(A) = 1$ for any $A \in \mathscr{F}$ implies $P_0(A) = 1$ and vice versa. On the other hand, P_0 and P_1 are orthogonal (denoted $P_0 \perp P_1$) if there exists an event A such that $P_1(A) = 1$ but $P_0(A) = 0$. For a stochastic process $\{Z(\boldsymbol{s}), \boldsymbol{s} \in \mathbb{R}^d\}$, to define previous concepts, we restrict the event A to the σ-algebra generated by $\{Z(\boldsymbol{s}), \boldsymbol{s} \in D\}$ where $D \subset \mathbb{R}^d$. We emphasize this restriction by saying that the two measures are equivalent on the paths of $\{Z(\boldsymbol{s}), \boldsymbol{s} \in D\}$.

Gaussian measures are completely characterized by their mean and covariance function. We write $P(\rho)$ for a Gaussian measure with zero mean and covariance function ρ. It is well known that two Gaussian measures are either equivalent or orthogonal on the paths of $\{Z(\boldsymbol{s}), \boldsymbol{s} \in D\}$ (Ibragimov and Rozanov, 1978).

Let $P(\rho_i)$, $i = 0,1$ be two zero mean Gaussian measures with isotropic covariance function ρ_i and associated spectral density $\widehat{\rho}_i$, $i = 0,1$, as defined through (4.6). Using results in Skorokhod and Yadrenko (1973) and Ibragimov and Rozanov (1978), Stein (2004) has shown that, if for some $a > 0$, $\widehat{\rho}_0(z)z^a$ is bounded away from 0 and ∞ as $z \to \infty$, and for some finite and positive c,

$$\int_c^\infty z^{d-1} \left\{ \frac{\widehat{\rho}_1(z) - \widehat{\rho}_0(z)}{\widehat{\rho}_0(z)} \right\}^2 \mathrm{d}z < \infty, \tag{5.1}$$

then for any bounded subset $D \subset \mathbb{R}^d$, $P(\rho_0) \equiv P(\rho_1)$ on the paths of $Z(\boldsymbol{s}), \boldsymbol{s} \in D$.

For the remainder of the paper, we denote with $P(\mathscr{M}_{\nu,\alpha,\sigma^2})$ a zero mean Gaussian measure induced by a Matérn covariance function with associated spectral density $\widehat{\mathscr{M}}_{\nu,\alpha,\sigma^2}$, and with $P(\varphi_{\mu,\kappa,\beta,\sigma^2})$ a zero mean Gaussian measure induced by a GW covariance function with associated spectral density $\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2}$.

Using (5.1) and (4.7), Zhang (2004) established the following characterization concerning the equivalence conditions of two Gaussian measures with Matérn covariance models.

Theorem 8 (Zhang, 2004)**.** *Let $P(\mathscr{M}_{\nu,\alpha_i,\sigma_i^2})$, $i = 0,1$, be two zero mean Gaussian measures. For any bounded infinite set $D \subset \mathbb{R}^d$, $d = 1,2,3$, $P(\mathscr{M}_{\nu,\alpha_0,\sigma_0^2}) \equiv P(\mathscr{M}_{\nu,\alpha_1,\sigma_1^2})$ on the paths of $Z(\boldsymbol{s}), \boldsymbol{s} \in D$, if and only if*

$$\sigma_0^2/\alpha_0^{2\nu} = \sigma_1^2/\alpha_1^{2\nu}. \tag{5.2}$$

The first relevant result of this paper concerns the characterization of the equivalence of two zero mean Gaussian measures under GW functions. The crux of the proof is the arguments in Equation (5.1), coupled with the asymptotic expansion of the spectral density as in Theorems 3 and 4.

Theorem 9. *Let $P(\varphi_{\mu,\kappa,\beta_i,\sigma_i^2})$, $i = 0,1$, be two zero mean Gaussian measures. For a given $\kappa \geq 0$, let $\mu > \lambda + d/2$, with λ as defined through Equation* (3.2)*. For any bounded infinite set $D \subset \mathbb{R}^d$, $d = 1,2,3$, $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2}) \equiv P(\varphi_{\mu,\kappa,\beta_1,\sigma_1^2})$ on the paths of $Z(\boldsymbol{s}), \boldsymbol{s} \in D$ if and only* *if*

$$\sigma_0^2/\beta_0^{2\kappa+1} = \sigma_1^2/\beta_1^{2\kappa+1}. \tag{5.3}$$

Proof. We first consider the case $\kappa > 0$. Let us start with the sufficient part of the assertion. From Theorem 3 (Point 3), there exist two constants c_i and C_i such that

$$c_i \le z^{2\lambda}\widehat{\varphi}_{\mu,\kappa,\beta_i,\sigma_i^2}(z) \le C_i \quad i = 0,1.$$

In order to prove the sufficient part of Theorem 4, we need to find conditions such that for some positive and finite c,

$$\int_c^\infty z^{d-1}\left(\frac{\widehat{\varphi}_{\mu,\kappa,\beta_1,\sigma_1^2}(z) - \widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(z)}{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(z)}\right)^2 \mathrm{d}z < \infty, \tag{5.4}$$

We proceed by direct construction, and, using Theorem 3 (Points 1 and 2), we find that, as $z \to \infty$,

$$\begin{aligned}
\left|\frac{\widehat{\varphi}_{\mu,\kappa,\beta_1,\sigma_1^2}(z) - \widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(z)}{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(z)}\right| &\le L^{\varsigma} c_0^{-1} z^{2\lambda}\Big|\sigma_1^2\beta_1^d\Big[c_3^{\varsigma}(\beta_1 z)^{-2\lambda}\big\{1+\mathscr{O}(z^{-2})\big\} \\
&\quad + c_4^{\varsigma}(z\beta_1)^{-(\mu+\lambda)}\big\{\cos(\beta_1 z - c_5^{\varsigma}) + \mathscr{O}(z^{-1})\big\}\Big] \\
&\quad - \sigma_0^2\beta_0^d\Big[c_3^{\varsigma}(\beta_0 z)^{-2\lambda}\big\{1+\mathscr{O}(z^{-2})\big\} \\
&\quad + c_4^{\varsigma}(z\beta_0)^{-(\mu+\lambda)}\big\{\cos(\beta_0 z - c_5^{\varsigma}) + \mathscr{O}(z^{-1})\big\}\Big]\Big| \\
&\le L^{\varsigma} c_0^{-1}\Big|c_3^{\varsigma}\Big[\sigma_1^2\beta_1^{-(1+2\kappa)} - \sigma_0^2\beta_0^{-(1+2\kappa)}\Big] + \mathscr{O}(z^{-2}) \\
&\quad + c_4^{\varsigma} z^{\lambda-\mu}\Big[\sigma_1^2\beta_1^{\tilde{\lambda}}\cos(\beta_1 z - c_5^{\varsigma}) - \sigma_0^2\beta_0^{\tilde{\lambda}}\cos(\beta_0 z - c_5^{\varsigma})\Big] \\
&\quad + c_4^{\varsigma} z^{\lambda-\mu}\mathscr{O}(z^{-1})\big\{\sigma_1^2\beta_1^{\tilde{\lambda}} - \sigma_0^2\beta_0^{\tilde{\lambda}}\big\}\Big|.
\end{aligned}$$

where $\tilde{\lambda} = d - (\mu+\lambda)$. Let us now write

$$\begin{aligned}
A(z) &= c_3^{\varsigma}[\sigma_1^2\beta_1^{-(1+2\kappa)} - \sigma_0^2\beta_0^{-(1+2\kappa)}] + \mathscr{O}(z^{-2}), \\
B(z) &= c_4^{\varsigma} z^{\lambda-\mu}[\sigma_1^2\beta_1^{\tilde{\lambda}}\cos(\beta_1 z - c_5^{\varsigma}) - \sigma_0^2\beta_0^{\tilde{\lambda}}\cos(\beta_0 z - c_5^{\varsigma})], \text{ and} \\
D(z) &= c_4^{\varsigma} z^{\lambda-\mu}\mathscr{O}(z^{-1})\big\{\sigma_1^2\beta_1^{\tilde{\lambda}} - \sigma_0^2\beta_0^{\tilde{\lambda}}\big\}.
\end{aligned}$$

Then a sufficient condition for (5.4) is the following condition:

$$(L^{\varsigma}/c_0)^2 \int_c^{\infty} z^{d-1}\big(A(z)+B(z)+D(z)\big)^2 \mathrm{d}z < \infty. \tag{5.5}$$

Note that $A(z)$ is of order $\mathcal{O}(z^{-2})$ under condition (5.3). We claim that (5.5) is satisfied if $\sigma_1^2\beta_1^{-(1+2\kappa)} = \sigma_0^2\beta_0^{-(1+2\kappa)}$ for $\mu > \lambda + d/2$, $d = 1,2,3$.

In fact, we have, for $z \to \infty$,

$$|B(z)| \leq c_4^{\varsigma} z^{\lambda-\mu}[\sigma_1^2\beta_1^{\tilde{\lambda}} + \sigma_0^2\beta_0^{\tilde{\lambda}}] \leq c_6 z^{\lambda-\mu},$$

and

$$\begin{aligned}|D(z)| &\leq c_4^{\varsigma} z^{\lambda-\mu}\mathcal{O}(z^{-1})\{\sigma_1^2\beta_1^{\tilde{\lambda}} + \sigma_0^2\beta_0^{\tilde{\lambda}}\} \\ &\leq c_7 c_4^{\varsigma} z^{\lambda-\mu-1}\{\sigma_1^2\beta_1^{\tilde{\lambda}} + \sigma_0^2\beta_0^{\tilde{\lambda}}\} \leq c_8 z^{\lambda-\mu-1}\end{aligned}$$

with c_6, c_7 and c_8 being positive and finite constants. Expanding (5.5) we notice that the dominant terms are A^2 and B^2, independently on the cross products. These are respectively of the order $\mathcal{O}(z^{-4})$ and $\mathcal{O}(z^{2(\lambda-\mu)})$. This in turn implies that the integral (5.5) is finite if $\sigma_1^2\beta_1^{-(1+2\kappa)} = \sigma_0^2\beta_0^{-(1+2\kappa)}$, for $\mu > \lambda + d/2$ and $d = 1,2,3$ and this implies that (5.4) is satisfied under the same conditions. The sufficient part of our claim is thus proved.

The necessary part follows the proof of Zhang (2004). For $\mu > \lambda + d/2$ and $d = 1,2,3$, we suppose $\sigma_1^2\beta_1^{-(1+2\kappa)} \neq \sigma_0^2\beta_0^{-(1+2\kappa)}$ and let $\sigma_2^2 = \sigma_0^2(\beta_0/\beta_1)^{-(1+2\kappa)}$. Then $\varphi_{\mu,\kappa,\beta_0,\sigma_0^2}$ and $\varphi_{\mu,\kappa,\beta_1,\sigma_2^2}$ define two equivalent Gaussian measures. We need to show that $\varphi_{\mu,\kappa,\beta_1,\sigma_2^2}$ and $\varphi_{\mu,\kappa,\beta_1,\sigma_1^2}$ define two orthogonal Gaussian measures. The rest of the proof follows the same arguments in Zhang (2004).

We omit the proof of the special case $\kappa = 0$, since is similar to the case $\kappa > 0$, but using the arguments in Theorem 4. □

An immediate consequence of Theorem 9 is that for fixed κ and μ, the β and σ^2 parameters cannot be estimated consistently (Zhang, 2004). Instead, the microergodic parameter $\sigma^2\beta^{-(1+2\kappa)}$ is consistently estimable. In Section 5.2, we establish the asymptotic properties of ML estimation associated with the microergodic parameter.

The next result depicts an interesting scenario in which a GW and Matérn model are considered and gives sufficient conditions for the compatibility of these two covariance models. We offer a constructive proof, the crux of the argument again being Equation (5.1). We treat the cases $\kappa > 0$ and $\kappa = 0$ separately.

Theorem 10. *Let $P(\mathscr{M}_{\nu,\alpha,\sigma_0^2})$ and $P(\varphi_{\mu,\kappa,\beta,\sigma_1^2})$ be two zero mean Gaussian measures with $\kappa > 0$. If $\nu = \kappa + 1/2$, $\mu > \lambda + d/2$, with λ as defined through Equation* (3.2)*, and*

$$\sigma_0^2 \alpha^{-2\nu} = C_{\nu,\kappa,\mu} \sigma_1^2 \beta^{-(1+2\kappa)}, \tag{5.6}$$

where $C_{\nu,\kappa,\mu} = \frac{\mu 2^{-d}\Gamma(\nu)\Gamma(\kappa)\Gamma(2\kappa+d)}{\Gamma(\nu+d/2)\Gamma(\kappa+d/2)B(2\kappa,\mu+1)}$, then for any bounded infinite set $D \subset \mathbb{R}^d$, $d = 1,2,3$, $P(\mathscr{M}_{\nu,\alpha,\sigma_0^2}) \equiv P(\varphi_{\mu,\kappa,\beta,\sigma_1^2})$ on the paths of $Z(\boldsymbol{s}), \boldsymbol{s} \in D$.

Proof. In order to prove Theorem 5, we need to find conditions such that for some positive and finite c,

$$\int_c^\infty z^{d-1} \left(\frac{\widehat{\varphi}_{\nu,\kappa,\beta,\sigma_1^2}(z) - \widehat{\mathscr{M}}_{\nu,\alpha,\sigma_0^2}(z)}{\widehat{\mathscr{M}}_{\nu,\alpha,\sigma_0^2}(z)} \right)^2 \mathrm{d}z < \infty. \tag{5.7}$$

It is known that $\widehat{\mathscr{M}}_{\nu,\alpha,\sigma_0^2}(z) z^a$ is bounded away from 0 and ∞ as $z \to \infty$ for some $a > 0$. (Zhang, 2004). Using (4.7) and Theorem 3 (Points 1 and 2), we have, as $z \to \infty$,

$$\begin{aligned}
&\left| \frac{\widehat{\varphi}_{\mu,\kappa,\beta,\sigma_1^2}(z) - \widehat{\mathscr{M}}_{\nu,\alpha,\sigma_0^2}(z)}{\widehat{\mathscr{M}}_{\nu,\alpha,\sigma_0^2}(z)} \right| \\
&= \left| \frac{\sigma_1^2 \beta^d \Gamma(\nu) L^{\varsigma}}{\Gamma(\nu + d/2) \sigma_0^2 \alpha^{-2\nu} \pi^{-\frac{d}{2}}} \Big[c_3^{\varsigma} (\beta z)^{-2\lambda} \{ 1 + \mathscr{O}(z^{-2}) \} \right. \\
&\quad \left. + c_4^{\varsigma} (z\beta)^{-(\mu+\lambda))} \{ \cos(\beta z - c_5^{\varsigma}) + \mathscr{O}(z^{-1}) \} \Big] (\alpha^{-2} + z^2)^{\nu + \frac{d}{2}} - 1 \right| \\
&= \left| \frac{\sigma_1^2 \beta^d \Gamma(\nu) L^{\varsigma}}{\Gamma(\nu + d/2) \sigma_0^2 \alpha^{-2\nu} \pi^{-\frac{d}{2}}} \Big[c_3^{\varsigma} (\beta z)^{-2\lambda} \{ 1 + \mathscr{O}(z^{-2}) \} \right. \\
&\quad \left. + c_4^{\varsigma} (z\beta)^{-(\mu+\lambda)} \{ \cos(\beta z - c_5^{\varsigma}) + \mathscr{O}(z^{-1}) \} \Big] z^{2\nu+d} ((\alpha z)^{-2} + 1)^{\nu + \frac{d}{2}} - 1 \right| \\
&= \left| w_1 z^{-2\lambda} \{ 1 + \mathscr{O}(z^{-2}) \} z^{2\nu+d} \Big[1 + (\nu + d/2)(\alpha z)^{-2} + \mathscr{O}(z^{-2}) \Big] \right. \\
&\quad \left. + w_2 z^{-(\mu+\lambda)} z^{2\nu+d} \Big[1 + (\nu + d/2)(\alpha z)^{-2} + \mathscr{O}(z^{-2}) \Big] \{ \cos(\beta z - c_5^{\varsigma}) + \mathscr{O}(z^{-1}) \} - 1 \right|,
\end{aligned}$$

where $w_1 = \frac{L^{\varsigma}\sigma_1^2\beta^{-(1+2\kappa)}\Gamma(\nu)c_3^{\varsigma}}{\Gamma(\nu+d/2)\sigma_0^2\alpha^{-2\nu}\pi^{-d/2}}$, $w_2 = w_1 c_4^{\varsigma}\beta^{\lambda-\mu}/c_3^{\varsigma}$. Since $z^{2\nu+d}\left[(\nu+d/2)(\alpha z)^{-2} + \mathscr{O}(z^{-2})\right] = \mathscr{O}(z^{2\nu+d-2})$, we have

$$\int_c^{\infty} z^{d-1}\left|\frac{\widehat{\varphi}_{\mu,\kappa,\beta,\sigma_1^2}(z) - \widehat{\mathscr{M}}_{\nu,\alpha,\sigma_0^2}(z)}{\widehat{\mathscr{M}}_{\nu,\alpha,\sigma_0^2}(z)}\right|^2 \mathrm{d}z$$

$$= \int_c^{\infty} z^{d-1}\Big| w_1 z^{-2\lambda}\mathscr{O}(z^{2\nu+d-2}) + \left\{w_1 z^{2\nu-(1+2\kappa)} - 1\right\} + w_1 z^{-2\lambda} \times \left\{\mathscr{O}(z^{2\nu+d-2}) + \mathscr{O}(z^{2\nu+d-4})\right\} + w_2 z^{-(\mu+\lambda)}\left\{\mathscr{O}(z^{2\nu+d-2}) + z^{2\nu+d}\right\}\left\{\cos(\beta z - c_5^{\varsigma}) + \mathscr{O}(z^{-1})\right\}\Big|^2 \mathrm{d}z.$$

For assessing the last integral, the following is relevant:

(i) $w_1 z^{2\nu-(1+2\kappa)} - 1 = 0$ if $\nu = \kappa + 1/2$ and $w_1 = 1$.

(ii) $\int_c^{\infty} z^{d-1}\left(w_1 z^{-2\lambda}\mathscr{O}(z^{2\nu+d-2})\right)^2 \mathrm{d}z < \infty$ if $d = 1,2,3$ and $\nu = \kappa + 1/2$.

(iii) $\int_c^{\infty} z^{d-1}\left(w_1 z^{-2\lambda}\left\{\mathscr{O}(z^{2\nu+d-2}) + \mathscr{O}(z^{2\nu+d-4})\right\}\right)^2 \mathrm{d}z < \infty$ if $d = 1,2,3$ and $\nu = \kappa + 1/2$.

(iv) $\int_c^{\infty} z^{d-1}\left(w_2 z^{-(\mu+\lambda)}\left\{\mathscr{O}(z^{2\nu+d-2}) + z^{2\nu+d}\right\}\left\{\cos(\beta z - c_5^{\varsigma}) + \mathscr{O}(z^{-1})\right\}\right)^2 \mathrm{d}z < \infty$ if $\mu > \lambda + d/2$ and $\nu = \kappa + 1/2$.

(v) $\int_c^{\infty} z^{d-1}\left(w_1 z^{-2\lambda}\mathscr{O}(z^{2\nu+d-2})\right)\left(w_1 z^{-2\lambda}\mathscr{O}(z^{-2})(\mathscr{O}(z^{2\nu+d-2}) + z^{2\nu+d})\right)\mathrm{d}z < \infty$ if $d = 1,2,3$ and $\nu = \kappa + 1/2$.

(vi) $\int_c^{\infty} z^{d-1}\left(w_1 z^{-2\lambda}\mathscr{O}(z^{2\nu+d-2})\right)\left(w_2 z^{-(\mu+\lambda)}\left\{\mathscr{O}(z^{2\nu+d-2}) + z^{2\nu+d}\right\}\left\{\cos(\beta z - c_5^{\varsigma}) + \mathscr{O}(z^{-1})\right\}\right)\mathrm{d}z < \infty$ if $\mu > \lambda + d - 2$ and $\nu = \kappa + 1/2$.

(vii) $\int_c^{\infty} z^{d-1}\left(w_1 z^{-2\lambda}\mathscr{O}(z^{-2})\left\{\mathscr{O}(z^{2\nu+d-2}) + z^{2\nu+d}\right\}\right)\left(w_2 z^{-(\mu+\lambda)}\left\{\mathscr{O}(z^{2\nu+d-2}) + z^{2\nu+d}\right\} \times \left\{\cos(\beta z - c_5^{\varsigma}) + \mathscr{O}(z^{-1})\right\}\right)\mathrm{d}z < \infty$ if $\mu > \lambda + d - 2$ and $\nu = \kappa + 1/2$.

This allows us to conclude that, for a given $\kappa > 0$, if $w_1 = 1$, $\nu = \kappa + 1/2$, $\mu > \lambda + d/2$ and $d = 1,2,3$ then (5.7) holds and thus $P(\mathscr{M}_{\nu,\alpha,\sigma_0^2}) \equiv P(\varphi_{\mu,\kappa,\beta,\sigma_1^2})$.

Condition $w_1 = 1$ is equivalent to

$$L^{\varsigma}c_3^{\varsigma}\sigma_1^2\beta^{-(1+2\kappa)} = \pi^{-d/2}\Gamma(\nu+d/2)\Gamma(\nu)^{-1}\sigma_0^2\alpha^{-2\nu},$$

and from the definition of c_3^{ς} and L^{ς}, the previous condition can be rewritten as $\sigma_0^2\alpha^{-2\nu} = C_{\nu,\kappa,\mu}\sigma_1^2\beta^{-(1+2\kappa)}$. □

Theorem 11. *Let $P(\mathscr{M}_{\nu,\alpha,\sigma_0^2})$ and $P(\varphi_{\mu,0,\beta,\sigma_1^2})$ be two zero mean Gaussian measures. If $\nu = 1/2$, $\mu > d+1/2$ and*

$$\sigma_0^2 \alpha^{-2\nu} = R_{\nu,\mu} \sigma_1^2 \beta^{-1}, \tag{5.8}$$

where $R_{\nu,\mu} = \mu \left(\frac{2^{1-d}\Gamma(\nu)\Gamma(d)}{\Gamma(\nu+d/2)\Gamma(d/2)} \right)$, then for any bounded infinite set $D \subset \mathbb{R}^d$, $d = 1,2,3$, $P(\mathscr{M}_{\nu,\alpha,\sigma_0^2}) \equiv P(\varphi_{\mu,0,\beta,\sigma_1^2})$ on the paths of $Z(\boldsymbol{s}), \boldsymbol{s} \in D$.

Proof. The proof follows the same arguments exposed for the case $\kappa > 0$ in Theorem 5 , but using (4.7) and Theorem 4 (Points 1 and 2). In this case, it can be shown that if $\nu = 0.5$, $\mu > d+1/2$, $d = 1,2,3$ and $\mu \left(\frac{2^{1-d}\Gamma(\nu)\Gamma(d)}{\Gamma(\nu+d/2)\Gamma(d/2)} \right) \sigma_1^2 \beta^{-1} = \sigma_0^2 \alpha^{-2\nu}$ then (5.7) holds. □

Remark: In Theorems 5 and 6 since $\nu = \kappa + 1/2$ for $\kappa \geq 0$, using the duplication formula of the gamma function, we easily obtain $C_{\kappa+1/2,\kappa,\mu} = \mu\Gamma(2\kappa+\mu+1)/\Gamma(\mu+1)$, and $R_{1/2,\mu} = 1$ in (5.6) and (5.8) respectively. Then the sufficient condition for $P(\mathscr{M}_{\nu,\alpha,\sigma_0^2}) \equiv P(\varphi_{\mu,\kappa,\beta,\sigma_1^2})$ can be simplified as

$$\sigma_0^2 \alpha^{-2\nu} = \left(\frac{\mu\Gamma(2\kappa+\mu+1)}{\Gamma(\mu+1)} \right) \sigma_1^2 \beta^{-(1+2\kappa)}, \tag{5.9}$$

$\nu = \kappa + 1/2$, $\mu > \lambda + d/2$ and $d = 1,2,3$ for $\kappa \geq 0$.

5.2 Asymptotic properties of the ML estimation for the GW model

We now focus on the microergodic parameter $\sigma^2\beta^{-(1+2\kappa)}$ associated with the GW family. The following results fix the asymptotic properties of its ML estimator. In particular, we will show that the microergodic parameter can be estimated consistently, and then we will assess the asymptotic distribution of the ML estimator.

Let $D \subset \mathbb{R}^d$ be a bounded subset of $\mathbb{R}^d$ and let $\boldsymbol{Z}_n = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n))'$ be a finite realization of $Z(\boldsymbol{s})$, $\boldsymbol{s} \in D$, a zero mean stationary Gaussian field with a given parametric covariance function $\sigma^2\phi(\|\cdot\|;\boldsymbol{\tau})$, with $\sigma^2 > 0$, $\boldsymbol{\tau}$ a parameter vector and ϕ a member of the class Φ_d, with $\phi(0;\boldsymbol{\tau}) = 1$.

We then write $R_n(\boldsymbol{\tau}) = [\phi(\|\boldsymbol{s}_i - \boldsymbol{s}_j\|;\boldsymbol{\tau})]_{i,j=1}^n$ for the associated correlation matrix. The Gaussian log-likelihood function is defined as

$$\mathscr{L}_n(\sigma^2,\boldsymbol{\tau}) = -\frac{1}{2}\left(n\log(2\pi\sigma^2) + \log(|R_n(\boldsymbol{\tau})|) + \frac{1}{\sigma^2}\boldsymbol{Z}_n' R_n(\boldsymbol{\tau})^{-1}\boldsymbol{Z}_n \right). \tag{5.10}$$

Under the Matérn model, the Gaussian log-likelihood is obtained with $\phi(\cdot;\boldsymbol{\tau}) \equiv \mathcal{M}_{\nu,\alpha,1}$ and $\boldsymbol{\tau} = (\nu,\alpha)'$. Since in what follows ν is assumed known and fixed, for notation convenience, we write $\tau = \alpha$. Let $\hat{\sigma}_n^2$ and $\hat{\alpha}_n$ be the maximum likelihood estimator obtained maximizing $\mathcal{L}_n(\sigma^2,\alpha)$ for a fixed ν.

Below, we report a result that establishes strong consistency and asymptotic distribution of the ML estimation of the microergodic parameter of the Matérn model, that is $\sigma^2/\alpha^{2\nu}$.

Theorem 12 (Kaufman and Shaby, 2013)**.** *Let $Z(\boldsymbol{s})$, $\boldsymbol{s} \in D \subset \mathbb{R}^d$, $d = 1,2,3$, be a zero mean Gaussian field with a Matérn covariance model $\mathcal{M}_{\nu,\alpha_0,\sigma_0^2}$. Suppose $(\sigma_0^2,\alpha_0)' \in (0,\infty)\times[\alpha_L,\alpha_U]$, for any $0<\alpha_L<\alpha_U<\infty$. Let $(\hat{\sigma}_n^2,\hat{\alpha}_n)'$ maximize* (5.10) *over* $(0,\infty)\times[\alpha_L,\alpha_U]$. *Then as $n \to \infty$,*

1. $\hat{\sigma}_n^2/\hat{\alpha}_n^{2\nu} \xrightarrow{a.s.} \sigma_0^2/\alpha_0^{2\nu}$, *and*
2. $\sqrt{n}(\hat{\sigma}_n^2/\hat{\alpha}_n^{2\nu} - \sigma_0^2/\alpha_0^{2\nu}) \xrightarrow{\mathcal{D}} \mathcal{N}(0,2(\sigma_0^2/\alpha_0^{2\nu})^2)$.

Analogous results can be found in Zhang (2004); Wang and Loh (2011), when $\hat{\alpha}_n$ is replaced by α, an arbitrary positive fixed constant. Kaufman and Shaby (2013) show, through simulation study, that asymptotic approximation using a fixed scale parameter can be problematic when applied to finite samples, even for large sample sizes. In contrast, they show that performance is improved and asymptotic approximations are applicable for smaller sample sizes, when the parameters are jointly estimated.

Now, let us consider the Gaussian log-likelihood under the GW model, so that $\boldsymbol{\tau} = (\mu,\kappa,\beta)'$ and $\phi(\cdot;\boldsymbol{\tau}) = \varphi_{\mu,\kappa,\beta,1}(\cdot)$ according to the previous notation. Since in what follows κ and μ are assumed known and fixed, for notation convenience we write $\tau = \beta$. To prove the analogue of Theorem 12 for the GW case, we consider two types of estimators. The first maximizes (5.10) with respect to σ^2 for a fixed arbitrary compact support $\beta > 0$, obtaining the following estimator

$$\hat{\sigma}_n^2(\beta) = arg\,max_{\sigma^2}\mathcal{L}_n(\sigma^2,\beta) = \mathbf{Z}_n' R_n(\beta)^{-1}\mathbf{Z}_n/n. \tag{5.11}$$

Here $R_n(\beta)$ is the correlation matrix coming from the GW family $\varphi_{\mu,\kappa\beta,1}$. The following result offers some asymptotic properties of the sequence of random variables $\hat{\sigma}_n^2(\beta)/\beta^{(1+2\kappa)}$. The proof is quite technical and has been deferred to the Appendix.

Theorem 13. *Let $Z(\boldsymbol{s})$, $\boldsymbol{s} \in D \subset \mathbb{R}^d$, $d = 1,2,3$, be a zero mean Gaussian field with GW covariance model $\varphi_{\mu,\kappa,\beta_0,\sigma_0^2}$, with $\mu > \lambda + d/2$. Suppose $(\sigma_0^2,\beta_0) \in (0,\infty)\times(0,\infty)$. For a fixed $\beta > 0$, let $\hat{\sigma}_n^2(\beta)$ as defined through Equation* (5.11). *Then, as $n \to \infty$,*

1. $\hat{\sigma}_n^2(\beta)/\beta^{1+2\kappa} \xrightarrow{a.s} \sigma_0^2(\beta_0)/\beta_0^{1+2\kappa}$ *and*

2. $\sqrt{n}(\hat{\sigma}_n^2(\beta)/\beta^{1+2\kappa} - \sigma_0^2(\beta_0)/\beta_0^{1+2\kappa}) \xrightarrow{\mathscr{D}} \mathscr{N}(0, 2(\sigma_0^2(\beta_0)/\beta_0^{1+2\kappa})^2)$.

The second type of estimation technique considers the joint maximization of (5.10) with respect to $(\sigma^2, \beta) \in (0, \infty) \times I$ where $I = [\beta_L, \beta_U]$ and $0 < \beta_L < \beta_U < \infty$. The solution of this optimization problem is given by $(\hat{\sigma}_n^2(\hat{\beta}_n), \hat{\beta}_n)$ where

$$\hat{\sigma}_n^2(\hat{\beta}_n) = \mathbf{Z}_n' R_n(\hat{\beta}_n)^{-1} \mathbf{Z}_n / n$$

and $\hat{\beta}_n = arg\,max_{\beta \in I} \mathscr{PL}_n(\beta)$. Here $\mathscr{PL}_n(\beta)$ is the profile log-likelihood:

$$\mathscr{PL}_n(\beta) = -\frac{1}{2}\left(\log(2\pi) + n\log(\hat{\sigma}_n^2(\beta)) + \log|R_n(\beta)| + n\right). \tag{5.12}$$

In order to establish strong consistency and asymptotic distribution of the sequence of random variables $\hat{\sigma}_n^2(\hat{\beta}_n)/\hat{\beta}_n^{1+2\kappa}$, we use the following Lemma that establishes the monotone behaviour of $\hat{\sigma}_n^2(\beta)/\beta^{1+2\kappa}$ when viewed as a function of $\beta \in I$ under specific condition on the μ parameter.

Lemma 1. *Let $S_n = \{\mathbf{s}_1, \ldots, \mathbf{s}_n \in D \subset \mathbb{R}^d\}$ denote any set of distinct locations. For any $\beta_1 < \beta_2 \in I$ and for each n, $\hat{\sigma}_n^2(\beta_1)/\beta_1^{1+2\kappa} \leq \hat{\sigma}_n^2(\beta_2)/\beta_2^{1+2\kappa}$ if and only if $\mu \geq \lambda + 3$.*

Proof. The proof follows Kaufman and Shaby (2013). Let $0 < \beta_1 < \beta_2$, with $\beta_1, \beta_2 \in I$. Then, for any $\mathbf{Z}_n$,

$$\hat{\sigma}_n^2(\beta_1)/\beta_1^{1+2\kappa} - \hat{\sigma}_n^2(\beta_2)/\beta_2^{1+2\kappa} = \frac{1}{n}\mathbf{Z}_n'(R_n(\beta_1)^{-1}\beta_1^{-(1+2\kappa)} - R_n(\beta_2)^{-1}\beta_2^{-(1+2\kappa)})\mathbf{Z}_n$$

is nonnegative if the matrix $R_n(\beta_1)^{-1}\beta_1^{-(1+2\kappa)} - R_n(\beta_2)^{-1}\beta_2^{-(1+2\kappa)}$ is positive semidefinite and this happens if and only if the matrix $B = R_n(\beta_2)\beta_2^{1+2\kappa} - R_n(\beta_1)\beta_1^{1+2\kappa}$ with generic element

$$B_{ij} = \beta_2^{1+2\kappa}\varphi_{\mu,\kappa,\beta_2,1}(||\mathbf{s}_i - \mathbf{s}_j||) - \beta_1^{1+2\kappa}\varphi_{\mu,\kappa,\beta_1,1}(||\mathbf{s}_i - \mathbf{s}_j||)$$

is positive semidefinite. From Theorem 2 in Porcu et al. (2016), this happens if and only if $\mu \geq \lambda + 3$. □

Theorem 14. *Let $Z(\mathbf{s})$, $\mathbf{s} \in D \subset \mathbb{R}^d$, $d = 1, 2, 3$, be a zero mean Gaussian field with a GW covariance model $\varphi_{\mu,\kappa,\beta_0,\sigma_0^2}$ with $\mu \geq \lambda + 3$. Suppose $(\sigma_0^2, \beta_0) \in (0, \infty) \times I$ where $I = [\beta_L, \beta_U]$ with $0 < \beta_L < \beta_U < \infty$. Let $(\hat{\sigma}_n^2, \hat{\beta}_n)'$ maximize (5.10) over $(0, \infty) \times I$. Then as $n \to \infty$,*

1. $\hat{\sigma}_n^2(\hat{\beta}_n)/\hat{\beta}_n^{1+2\kappa} \xrightarrow{a.s} \sigma_0^2(\beta_0)/\beta_0^{1+2\kappa}$ *and*

2. $\sqrt{n}(\hat{\sigma}_n^2(\hat{\beta}_n)/\hat{\beta}_n^{1+2\kappa} - \sigma_0^2(\beta_0)/\beta_0^{1+2\kappa}) \xrightarrow{\mathscr{D}} \mathscr{N}(0, 2(\sigma_0^2(\beta_0)/\beta_0^{1+2\kappa})^2)$.

Proof. The proof follows Kaufman and Shaby (2013) who use the same arguments in the Matérn case. Let $\mathscr{G}_n(x) = \hat{\sigma}_n^2(x)/x^{1+2\kappa}$ and define the sequences $\mathscr{G}_n(\beta_L)$ and $\mathscr{G}_n(\beta_U)$. Since $\beta_L \leq \hat{\beta}_n \leq \beta_U$ for every n, then, using Lemma 1, $\mathscr{G}_n(\beta_U) \leq \mathscr{G}_n(\hat{\beta}_n) \leq \mathscr{G}_n(\beta_L)$ for all n with probability one. Combining this with Theorem 13 implies the result. □

5.3 Prediction using GW model

We now consider prediction of a Gaussian field at a new location $\boldsymbol{s}_0$, using the GW model, under fixed domain asymptotics. Specifically, we focus on two properties: asymptotic efficiency prediction and asymptotically correct estimation of prediction variance. Stein (1988) shows that both asymptotic properties hold when the Gaussian measures are equivalent. Let $P(\varphi_{\mu,\kappa,\beta_i,\sigma_i^2})$, $i = 1,2$, be two probability zero mean Gaussian measures. Under $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2})$, and using Theorem 9, both properties hold when $\sigma_0^2\beta_0^{-(1+2\kappa)} = \sigma_1^2\beta_1^{-(1+2\kappa)}$, $\mu > \lambda + d/2$ and $d = 1,2,3$.

Similarly, let $P(\mathscr{M}_{\nu,\alpha,\sigma_2^2})$ and $P(\varphi_{\mu,\kappa,\beta_1,\sigma_1^2})$ be two Gaussian measures with Matérn and GW model. Under $P(\mathscr{M}_{\nu,\alpha,\sigma_2^2})$ both properties hold when (5.9) is true, $\mu > \lambda + d/2$, $d = 1,2,3$. Actually, Stein (1993) gives a substantially weaker condition for asymptotic efficiency prediction based on the asymptotic behaviour of the ratio of the isotropic spectral densities. Now, let

$$\widehat{Z}_n(\mu,\kappa,\beta) = \boldsymbol{c}_n(\mu,\kappa,\beta)' R_n(\mu,\kappa,\beta)^{-1} \boldsymbol{Z}_n \tag{5.13}$$

be the best linear unbiased predictor at an unknown location $\boldsymbol{s}_0 \in D \subset \mathbb{R}^d$, under the misspecified model $P(\varphi_{\mu,\kappa,\beta,\sigma^2})$, where $\boldsymbol{c}_n(\mu,\kappa,\beta) = [\varphi_{\mu,\kappa,\beta,1}(\boldsymbol{s}_0 - \boldsymbol{s}_i)]_{i=1}^n$ and $R_n(\mu,\kappa,\beta) = [\varphi_{\mu,\kappa,\beta,1}(\boldsymbol{s}_i - \boldsymbol{s}_j)]_{i,j=1}^n$ is the correlation matrix.

If the correct model is $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2})$, then the mean squared error of the predictor is given by

$$\begin{aligned} Var_{\mu,\kappa,\beta_0,\sigma_0^2}\left[\widehat{Z}_n(\mu,\kappa,\beta) - Z(\boldsymbol{s}_0)\right] &= \sigma_0^2\Big(1 - 2\boldsymbol{c}_n(\mu,\kappa,\beta)'R_n(\mu,\kappa,\beta)^{-1}\boldsymbol{c}_n(\mu,\kappa,\beta_0) \\ &+ \boldsymbol{c}_n(\mu,\kappa,\beta)'R_n(\mu,\kappa,\beta)^{-1}R_n(\mu,\kappa,\beta_0)R_n(\mu,\kappa,\beta)^{-1}\boldsymbol{c}_n(\mu,\kappa,\beta)\Big). \end{aligned} \tag{5.14}$$

In the case that $\beta_0 = \beta$, i.e., true and wrong models coincide, this expression simplifies to

$$\begin{aligned} Var_{\mu,\kappa,\beta_0,\sigma_0^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_0) - Z(\boldsymbol{s}_0)\right] \\ = \sigma_0^2\big(1 - \boldsymbol{c}_n(\mu,\kappa,\beta_0)'R_n(\mu,\kappa,\beta_0)^{-1}\boldsymbol{c}_n(\mu,\kappa,\beta_0)\big). \end{aligned} \tag{5.15}$$

Similarly $Var_{\nu,\alpha,\sigma_2^2}\left[\widehat{Z}_n(\mu,\kappa,\beta)-Z(\boldsymbol{s}_0)\right]$ and $Var_{\nu,\alpha,\sigma_2^2}\left[\widehat{Z}_n(\nu,\alpha)-Z(\boldsymbol{s}_0)\right]$ can be defined under $P(\mathscr{M}_{\nu,\alpha,\sigma_2^2})$, where $\widehat{Z}_n(\nu,\alpha)$ is the best linear unbiased predictor using the Matérn model. The following results are an application of Theorems 1 and 2 of Stein (1993).

Theorem 15. *Let $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2})$, $P(\varphi_{\mu,\kappa,\beta_1,\sigma_1^2})$, $P(\mathscr{M}_{\nu,\alpha,\sigma_2^2})$ be three Gaussian probability measures on $D\subset\mathbb{R}^d$ and let $\mu>\lambda$. Then, for all $\boldsymbol{s}_0\in D$:*

1. *Under $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2})$, as $n\to\infty$,*

$$\frac{Var_{\mu,\kappa,\beta_0,\sigma_0^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_1)-Z(\boldsymbol{s}_0)\right]}{Var_{\mu,\kappa,\beta_0,\sigma_0^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_0)-Z(\boldsymbol{s}_0)\right]}\longrightarrow 1, \tag{5.16}$$

for any fixed $\beta_1>0$.

2. *Under $P(\mathscr{M}_{\nu,\alpha,\sigma_2^2})$, if $\nu=\kappa+\frac{1}{2}$ as $n\to\infty$,*

$$\frac{Var_{\nu,\alpha,\sigma_2^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_1)-Z(\boldsymbol{s}_0)\right]}{Var_{\nu,\alpha,\sigma_2^2}\left[\widehat{Z}_n(\nu,\alpha)-Z(\boldsymbol{s}_0)\right]}\longrightarrow 1, \tag{5.17}$$

for any fixed $\beta_1>0$.

3. *Under $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2})$, if $\sigma_0^2\beta_0^{-(1+2\kappa)}=\sigma_1^2\beta_1^{-(1+2\kappa)}$, then as $n\to\infty$,*

$$\frac{Var_{\mu,\kappa,\beta_1,\sigma_1^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_1)-Z(\boldsymbol{s}_0)\right]}{Var_{\mu,\kappa,\beta_0,\sigma_0^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_1)-Z(\boldsymbol{s}_0)\right]}\longrightarrow 1. \tag{5.18}$$

4. *Under $P(\mathscr{M}_{\nu,\alpha,\sigma_2^2})$, if $\mu\Gamma(2\kappa+\mu+1)/\Gamma(\mu+1)\times\sigma_1^2\beta_1^{-(1+2\kappa)}=\sigma_2^2\alpha^{-2\nu}$, $\nu=\kappa+\frac{1}{2}$, then as $n\to\infty$,*

$$\frac{Var_{\mu,\kappa,\beta_1,\sigma_1^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_1)-Z(\boldsymbol{s}_0)\right]}{Var_{\nu,\alpha,\sigma_2^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_1)-Z(\boldsymbol{s}_0)\right]}\longrightarrow 1. \tag{5.19}$$

Proof. Since $\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(z)$ is bounded away from zero and infinity and as $z\to\infty$,

$$\frac{\widehat{\varphi}_{\mu,\kappa,\beta_1,\sigma_1^2}(z)}{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(z)}=\frac{\sigma_1^2\beta_1^d\left[c_3^{\varsigma}\beta_1^{-2\lambda}\left\{1+\mathscr{O}(z^{-2})\right\}+c_4^{\varsigma}\beta_1^{-(\mu+\lambda)}z^{\lambda-\mu}\left\{\cos(z\beta_1-c_5^{\varsigma})+\mathscr{O}(z^{-1})\right\}\right]}{\sigma_0^2\beta_0^d\left[c_3^{\varsigma}\beta_0^{-2\lambda}\left\{1+\mathscr{O}(z^{-2})\right\}+c_4^{\varsigma}\beta_0^{-(\mu+\lambda)}z^{\lambda-\mu}\left\{\cos(z\beta_0-c_5^{\varsigma})+\mathscr{O}(z^{-1})\right\}\right]}$$

then, for $\mu > \lambda$, we have

$$\lim_{z\to\infty} \frac{\widehat{\varphi}_{\mu,\kappa,\beta_1,\sigma_1^2}(z)}{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(z)} = \frac{\sigma_1^2\beta_1^{-(1+2\kappa)}}{\sigma_0^2\beta_0^{-(1+2\kappa)}}, \tag{5.20}$$

and using Theorem 1 of Stein (1993), we obtain (5.16). If $\sigma_1^2\beta_1^{-(1+2\kappa)} = \sigma_0^2\beta_0^{-(1+2\kappa)}$ and using Theorem 2 of Stein (1993), we obtain (5.18).

Similarly, since $\widehat{\mathscr{M}}_{\nu,\alpha,\sigma_2^2}(z)$ is bounded away from zero and infinity and as $z\to\infty$,

$$\begin{aligned}
&\frac{\widehat{\varphi}_{\mu,\kappa,\beta_1,\sigma_1^2}(z)}{\widehat{\mathscr{M}}_{\nu,\alpha,\sigma_2^2}(z)}\\
&= \frac{\sigma_1^2\beta^d\Gamma(\nu)L^{\varsigma}}{\Gamma(\nu+d/2)\sigma_0^2\alpha^{-2\nu}\pi^{-\frac{d}{2}}}\Big[c_3^{\varsigma}(\beta z)^{-2\lambda}\{1+\mathscr{O}(z^{-2})\}+c_4^{\varsigma}(z\beta)^{-(\mu+\lambda))}\\
&\quad\times\{\cos(\beta z-c_5^{\varsigma})+\mathscr{O}(z^{-1})\}\Big](\alpha^{-2}+z^2)^{\nu+\frac{d}{2}}\\
&= \frac{\sigma_1^2\beta^d\Gamma(\nu)L^{\varsigma}}{\Gamma(\nu+d/2)\sigma_0^2\alpha^{-2\nu}\pi^{-\frac{d}{2}}}\Big[c_3^{\varsigma}(\beta z)^{-2\lambda}\{1+\mathscr{O}(z^{-2})\}+c_4^{\varsigma}(z\beta)^{-(\mu+\lambda))}\\
&\quad\times\{\cos(\beta z-c_5^{\varsigma})+\mathscr{O}(z^{-1})\}\Big]z^{2\nu+d}\Big[1+(\nu+d/2)(\alpha z)^{-2}+\mathscr{O}(z^{-2})\Big]\\
&= \frac{\sigma_1^2\beta^d\Gamma(\nu)L^{\varsigma}}{\Gamma(\nu+d/2)\sigma_0^2\alpha^{-2\nu}\pi^{-\frac{d}{2}}}\Big[c_3^{\varsigma}\beta^{-2\lambda}z^{2\nu-2\lambda+d}\{1+\mathscr{O}(z^{-2})\}+c_4^{\varsigma}\beta^{-(\mu+\lambda)}\\
&\quad\times z^{2\nu-(\mu+\lambda)+d}\{\cos(\beta z-c_5^{\varsigma})+\mathscr{O}(z^{-1})\}\Big]\Big[1+(\nu+d/2)(\alpha z)^{-2}+\mathscr{O}(z^{-2})\Big]
\end{aligned}$$

then, if $2\nu+d=2\lambda$, that is $\kappa+1/2=\nu$, $\mu>\lambda$ and considering Remark 1 then:

$$\lim_{z\to\infty} \frac{\widehat{\varphi}_{\mu,\kappa,\beta_1,\sigma_1^2}(z)}{\widehat{\mathscr{M}}_{\nu,\alpha,\sigma_2^2}(z)} = \frac{\sigma_1^2\beta_1^{-(1+2\kappa)}}{\sigma_2^2\alpha^{-2\nu}}\left(\mu\frac{\Gamma(2\kappa+\mu+1)}{\Gamma(\mu+1)}\right). \tag{5.21}$$

Using Theorem 1 of Stein (1993), we obtain (5.17). If $\sigma_1^2\beta_1^{-(1+2\kappa)}\left(\mu\frac{\Gamma(2\kappa+\mu+1)}{\Gamma(\mu+1)}\right)=\sigma_2^2\alpha^{-2\nu}$ and using Theorem 2 of Stein (1993), we obtain (5.19). □

The implication of point 1 is that under $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2})$, prediction with $\varphi_{\mu,\kappa,\beta_1,\sigma_0^2}$ with an arbitrary $\beta_1>0$ gives asymptotic prediction efficiency, if the correct value of κ and μ are used and $\mu>\lambda$. By virtue of point 2, under $P(\mathscr{M}_{\nu,\alpha,\sigma_2^2})$, prediction with $\varphi_{\mu,\kappa,\beta_1,\sigma_0^2}$, with an arbitrary $\beta_1>0$, gives asymptotic prediction efficiency, if $\nu=\kappa+1/2$, $\mu>\lambda$. For instance, if $\sigma_2^2e^{-r/\alpha}$ is the true covariance, asymptotic prediction efficiency can be achieved with $\sigma_0^2(1-r/\beta_1)_+^{\mu}$, using an arbitrary β_1, and $\mu>1.5$ when $d=2$.

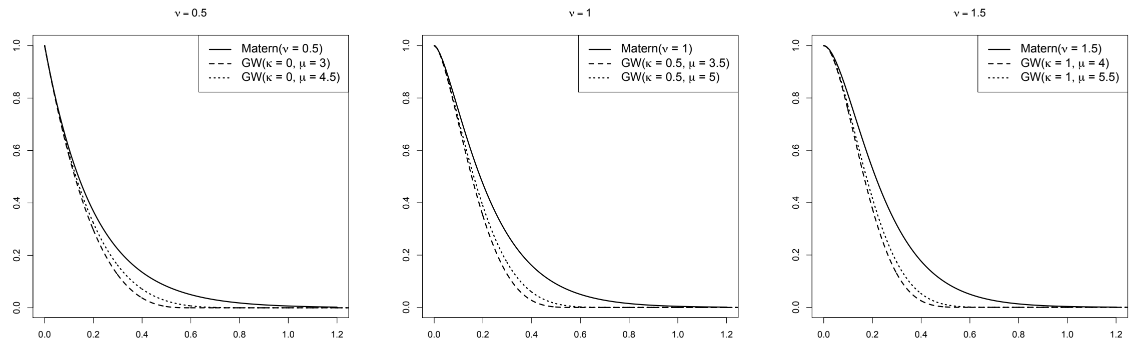

Fig. 5.1 Compatible correlation models for the case $d = 2$: The Matern model when $\nu = 0.5, 1, 1.5$ (from left to right) and the practical range is 0.6 and two compatibles GW models. For the GW models $\kappa = \nu - 0.5$, $\mu = \lambda + 1 + x$, with $x = 0.5, 2$ and the compact support is fixed using the equivalence condition.

In view of point 3, under $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2})$, prediction with $\varphi_{\mu,\kappa,\beta_1,\sigma_1^2}$, when $\sigma_0^2\beta_0^{-(1+2\kappa)} = \sigma_1^2\beta_1^{-(1+2\kappa)}$ provides asymptotic prediction efficiency and asymptotically correct estimates of error variance, if $\mu > \lambda$. Finally, point 4 implies that under $P(\mathscr{M}_{\nu,\alpha,\sigma_2^2})$, prediction using $\varphi_{\mu,\kappa,\beta_1,\sigma_1^2}$, under the conditions $\mu\Gamma(2\kappa+\mu+1)/\Gamma(\mu+1)\sigma_1^2\beta_1^{-(1+2\kappa)} = \sigma_2^2\alpha^{-2\nu}$, $\nu = \kappa + 1/2$ and $\mu > \lambda$, provides asymptotic prediction efficiency and asymptotically correct estimates of error variance.

For instance, if $\sigma_2^2 e^{-r/\alpha}$ is the true covariance and $d = 2$, asymptotic prediction efficiency and asymptotically correct estimates of variance error can be achieved with $\sigma_1^2(1 - r/\beta_1)_+^{\mu}$ setting $\beta_1 = \mu\alpha\sigma_1^2\sigma_2^{-2}$, and $\mu > 1.5$. Setting $\sigma_2^2 = \sigma_1^2 = 1$, $\mu = 3$, $\alpha = x/3$ (x in this case is the so-called practical range, i.e., the correlation is lower than 0.05 when $r > x$), the *equivalent* compact support is $\beta_1 = x$. Note that in this special case, the practical range of the exponential model and the compact support of the Askey function coincide. Figure 5.1 shows the Matérn correlation model with $\nu = 0.5, 1, 1.5$ and practical range equal to 0.6, and two compatible GW correlation models when $d = 2$ with $\kappa = \nu - 0.5$, $\mu = \lambda + 1 + x$, with $x = 0.5, 2$ and the associated compact supports are obtained using the equivalence condition. They are 0.601, 0.595, 0.624 for $\kappa = 0, 0.5, 1$ respectively when $x = 0.5$ and 0.901, 0.821, 0.815 for $\kappa = 0, 0.5, 1$, respectively, when $x = 2$.

In practice, covariance parameters are unknown, so it is common to estimate them and then plug into (5.13) and (5.15). Nevertheless, the asymptotic properties of this procedure are quite difficult to obtain (Putter and Young, 2001). Instead, most theoretical results have been given under a framework in which plug-in parameters are fixed, rather than being estimated from observations.

As in Theorem 4 of Kaufman and Shaby (2013), our Points 3 and 4 may be extended to include estimation of the variance parameter. Specifically let $\widehat{\sigma}_n^2 = \boldsymbol{Z}_n' R_n(\mu,\kappa,\beta_1)^{-1}\boldsymbol{Z}_n/n$. Then as $n \to \infty$,

$$\frac{Var_{\mu,\kappa,\beta_1,\widehat{\sigma}_n^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_1) - Z(\boldsymbol{s}_0)\right]}{Var_{\mu,\kappa,\beta_0,\sigma_0^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_1) - Z(\boldsymbol{s}_0)\right]} \to 1, \tag{5.22}$$

$$\frac{Var_{\mu,\kappa,\beta_1,\widehat{\sigma}_n^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_1) - Z(\boldsymbol{s}_0)\right]}{Var_{\nu,\alpha,\sigma_2^2}\left[\widehat{Z}_n(\mu,\kappa,\beta_1) - Z(\boldsymbol{s}_0)\right]} \to 1. \tag{5.23}$$

The proof follows the lines of Kaufman and Shaby (2013), and we omit it. As outlined in Kaufman and Shaby (2013), we also conjecture that (5.22) and (5.23) hold if β_1 is replaced by its maximum likelihood estimator.

5.4 Simulations and illustrations

The main goals of this section are twofold: on the one hand, we compare the finite sample behavior of the ML estimation of the microergodic parameter of the GW model with the asymptotic distributions given in Theorems 13 and 14. On the other hand, we compare the finite sample behavior of MSE prediction of a zero mean Gaussian field with Matérn covariance model, using both a Matérn and a compatible GW covariance model, using CT applied to a Matérn model as a benchmark.

Regarding the first goal, we simulate, using Cholesky decomposition, and then we estimate with ML, 1000 realizations from a zero mean Gaussian field with GW model. Sampling locations are constructed as in Kaufman et al. (2008), using a perturbated regular grid, to avoid numerical issues. Specifically, we have considered a regular grid with increments 0.03 over $[0,1]^d$, $d=2$. Then the grid points have been perturbated, adding a uniform random value on $[-0.01, 0.01]$ to each coordinate. Figure 5.2 shows the perturbated grid considered, from which we randomly choose $n = 50, 100, 250, 500, 1000$ locations without replacement.

For the GW covariance model $\varphi_{\mu,\kappa,\beta_0,\sigma_0^2}$, we use different values of the compact support and smoothness parameters, that is $\beta_0 = 0.2, 0.4, 0.6$, $\kappa = 0, 0.5, 1$, and fix $\sigma_0^2 = 1$ and, in view of Theorem 14, $\mu = \lambda(2,\kappa)+3$. For each simulation, we consider κ and μ as known and fixed, and we estimate with ML the variance and compact support parameters, obtaining $\hat{\sigma}_i^2$ and $\hat{\beta}_i$, $i = 1,\dots,1000$. To estimate, we first maximize the profile log-likelihood (5.12) to get $\hat{\beta}_i$. Then, we obtain $\hat{\sigma}_i^2(\hat{\beta}_i) = \boldsymbol{z}_i' R(\hat{\beta}_i)^{-1}\boldsymbol{z}_i/n$, where $\boldsymbol{z}_i$ is the data vector of simulation i.

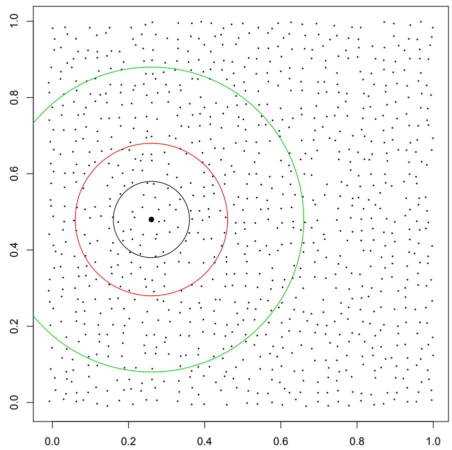

Fig. 5.2 Perturbated grid consisting of $n = 1156$ considered in the simulation study. The black dot has coordinates $(0.26, 0.48)$. In the circles (from smaller to larger) the location sites involved in prediction with GW with compact support equal to $0.1, 0.2, 0.4$.

Optimization was carried out using the R (R Development Core Team, 2007) function *optimize* where, following Kaufman and Shaby (2013), the compact support parameter was restricted to the interval $[\varepsilon, 15\beta]$ and ε is slightly larger than machine precision, about 10^{-15} here.

Using the asymptotic distributions stated in Theorems 13 and 14, Table 5.1 compares the sample quantiles of order $0.05, 0.25, 0.5, 0.75, 0.95$, mean and variance of $\sqrt{n/2}\big(\widehat{\sigma}_i^2(x)\beta_0^{1+2\kappa} / (\sigma_0^2 x^{1+2\kappa}) - 1\big)$ for $x = \widehat{\beta}_i, \beta_0, 0.5\beta_0, 2\beta_0$ with the associated theoretical values of the standard Gaussian distribution, for $\beta_0 = 0.4$, $\kappa = 0, 0.5, 1$ and $n = 250, 500, 1000$.

As expected, the best approximation is achieved overall when using the true compact support, i.e., $x = \beta_0$, with little difference between the different values of β and κ. In the case of $x = \widehat{\beta}_i$, the asymptotic distribution given in Theorem 14 is a satisfactory approximation of the sample distribution, visually improving when increasing n. Using some simulation results, not reported here, it can be highlighted as the asymptotic approximations given in Theorems 13 and 14 improve when increasing β_0. The value of κ has less impact compared to β_0. In general, smaller values lead to better results.

When using compact supports that are too small or too large with respect to the true compact support ($x = 0.5\beta_0, 2\beta_0$), the convergence of the asymptotic distribution given in Theorem 13 is very slow. In particular, when $x = 0.5\beta_0$, the asymptotic approximation is not satisfactory even for $n = 1000$. In other words, confidence intervals for the microergodic parameter, based on Theorem 13, i.e., fixing an arbitrary compact support, can be problematic when applied to finite samples, even for large sample sizes. We strongly recommend jointly

Table 5.1 Sample quantiles, mean and variance of $\sqrt{n/2}(\widehat{\sigma}_i^2(x)\beta_0^{1+2\kappa}/(\sigma_0^2 x^{1+2\kappa})-1)$, $i=1,\dots,1000$, for $x=\widehat{\beta}_i,\beta_0,0.5\beta_0,2\beta_0$ for different values of κ, when $\beta_0=0.4$ and $n=250,500,1000$, compared with the associated theoretical values of the standard Gaussian distribution.

κ	x	n	5%	25%	50%	75%	95%	Mean	Var
0	$\widehat{\beta}$	250	-1.699	-0.721	-0.020	0.798	2.084	0.072	1.375
		500	-1.680	-0.677	0.027	0.758	1.966	0.071	1.212
		1000	-1.614	-0.666	0.062	0.767	1.788	0.057	1.104
	β_0	250	-1.548	-0.670	-0.039	0.675	1.833	0.025	1.058
		500	-1.632	-0.665	0.001	0.661	1.754	0.027	1.047
		1000	-1.629	-0.690	0.020	0.698	1.627	0.011	1.009
	$0.5\beta_0$	250	3.224	4.953	6.163	7.471	9.370	6.234	3.493
		500	3.399	4.762	5.948	7.018	8.879	5.979	2.840
		1000	2.792	4.063	5.059	5.984	7.516	5.088	2.088
	$2\beta_0$	250	-2.443	-1.698	-1.128	-0.490	0.610	-1.065	0.898
		500	-2.485	-1.576	-0.941	-0.313	0.718	-0.904	0.947
		1000	-2.324	-1.438	-0.759	-0.107	0.819	-0.757	0.949
0.5	$\widehat{\beta}$	250	-1.761	-0.786	0.019	0.807	2.271	0.072	1.506
		500	-1.774	-0.714	0.027	0.822	1.978	0.063	1.309
		1000	-1.609	-0.700	0.047	0.761	1.840	0.051	1.152
	β_0	250	-1.548	-0.670	-0.039	0.675	1.833	0.025	1.058
		500	-1.632	-0.665	0.001	0.661	1.754	0.027	1.047
		1000	-1.629	-0.690	0.020	0.698	1.627	0.011	1.009
	$0.5\beta_0$	250	11.462	14.603	16.995	19.573	23.414	17.155	12.818
		500	11.133	13.624	15.459	17.592	21.090	15.697	9.060
		1000	9.192	11.051	12.578	14.187	16.904	12.733	5.560
	$2\beta_0$	250	-3.166	-2.469	-1.914	-1.315	-0.260	-1.860	0.784
		500	-3.136	-2.258	-1.628	-1.037	-0.029	-1.604	0.883
		1000	-2.851	-1.999	-1.353	-0.707	0.207	-1.342	0.907
1	$\widehat{\beta}$	250	-1.825	-0.868	0.042	0.836	2.389	0.078	1.661
		500	-1.869	-0.770	0.027	0.820	2.092	0.059	1.412
		1000	-1.679	-0.719	0.058	0.762	1.836	0.045	1.199
	β_0	250	-1.548	-0.670	-0.039	0.675	1.833	0.025	1.058
		500	-1.632	-0.665	0.001	0.661	1.754	0.027	1.047
		1000	-1.629	-0.690	0.020	0.698	1.627	0.011	1.009
	$0.5\beta_0$	250	28.654	34.704	39.574	44.651	52.477	39.856	51.483
		500	27.166	31.848	35.553	39.808	46.519	35.992	34.995
		1000	22.055	25.398	28.218	31.256	36.451	28.565	19.929
	$2\beta_0$	250	-3.949	-3.312	-2.806	-2.262	-1.288	-2.750	0.666
		500	-3.876	-3.050	-2.445	-1.862	-0.925	-2.427	0.809
		1000	-3.524	-2.675	-2.065	-1.419	-0.532	-2.047	0.856
$N(0,1)$			-1.645	-0.674	0	0.674	1.645	0	1

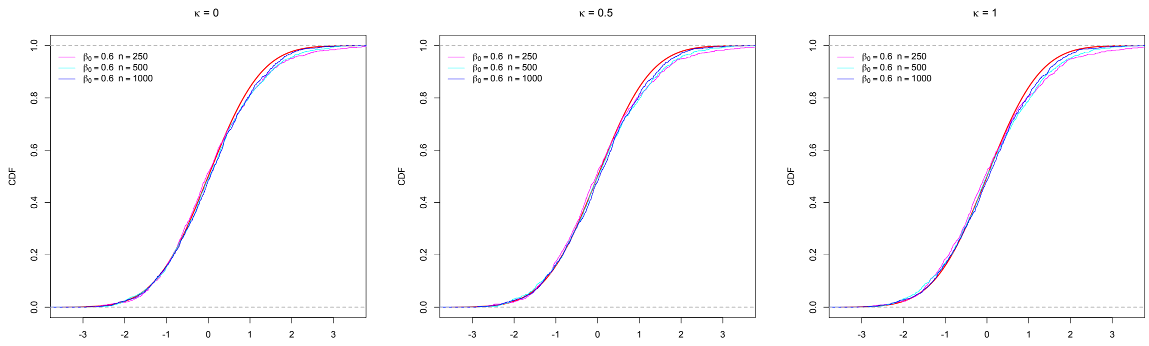

Fig. 5.3 Empirical CDF of the simulated ML estimation of the standardized microergodic parameter vs CDF of a standard Gaussian distribution (red line) when $\sigma_0^2 = 1$, $\kappa = 0, 0.5, 1$ (from left to right), $\beta_0 = 0.6$ and $n = 250, 500, 1000$.

estimating variance and compact support and using the asymptotic distribution give in Theorem 14 or, alternatively, choosing β conservatively.

As a graphical example, Figure 5.3 compares the empirical CDF of the ML estimates of the standardized microergodic parameter with the CDF of the standard Gaussian distribution when $\sigma_0^2 = 1$, $\kappa = 0, 0.5, 1$, $\beta_0 = 0.6$ and $n = 250, 500, 1000$. Finally, our numerical results are consistent with the results in Kaufman and Shaby (2013), in the Matérn case.

As for the second goal, using the results given in Theorem 15 Points 2 and 4, we now specifically compare asymptotic efficiency prediction and asymptotically correct estimation of prediction variance using ratios (5.17) and (5.19) respectively. As a benchmark, we also consider the same ratios using a tapered Matérn model.

More precisely, we consider a Matérn model $\mathcal{M}_{\nu,\alpha,\sigma_2^2}$ setting $\sigma_2^2 = 1$, $\nu = 0.5, 1, 1.5$ and $\alpha = y/c_\nu$ with $y = 0.1, 0.2, 0.4$ if $\nu = 0.5$, $y = 0.101, 0.202, 0.404$ if $\nu = 1$ and $y = 0.097, 0.193, 0.385$ if $\nu = 1.5$. Here c_ν is a scalar depending on ν such that $\mathcal{M}_{\nu,1,1}(r)$ is lower than 0.05 when $r > c_\nu$ that is y is the practical range.

Let us define the ratios (5.17) and (5.19) as $U_1(\beta_1)$ and U_2, respectively. For each ν and α, we randomly select $n_j = 50, 100, 250, 500, 1000$, $j = 1, \ldots, 500$ location sites without replacement from the perturbated grid in Figure 5.2.

For each j, we compute the ratio $U_{1j}(\beta_1)$ and the ratio U_{2j}, $j = 1, \ldots, 500$, using closed form expressions in Equation (5.14) and (5.15) when predicting the location site $(0.26, 0.48)'$ (black dot in Figure 5.2).

Specifically for each U_{2j}, following the conditions in Theorem 15 Point 4, we set $\sigma_1^2 = 1$, $\kappa = \nu - 1/2$, $\mu = \lambda + 1.5$. The "equivalent" compact support is obtained as

$$\beta_1^* = \left[\left(\mu \frac{\Gamma(2\kappa + \mu + 1)}{\Gamma(\mu + 1)} \right) \frac{\sigma_1^2 \alpha^{-2\nu}}{\sigma_2^2} \right]^{1/(1+2\kappa)}.$$

Under this specific setting the "equivalent" compact support associated with the (varying with ν) practical range is approximately $\beta_1^* = 0.1, 0.2, 0.4$, irrespectively of ν. Figure 5.2 shows the location sites involved in the prediction using GW functions with $\beta_1^* = 0.1, 0.2, 0.4$.

For each $U_{1j}(\beta)$, following Theorem 15 point 2, we fix $\kappa = \nu - 1/2$, $\mu = \lambda + 1.5$ and $\beta = \beta_1^*$. Then, to investigate the effect of considering an arbitrary compact support on the convergence of ratio (5.17), we also consider, $U_{1j}(0.2\beta_1^*)$ and $U_{1j}(5\beta_1^*)$.

For each combination of ν, α, Table 5.2 shows the empirical means $\bar{U}_1(x\beta_1^*) = \sum_{j=1}^{500} U_{1j}(x\beta_1^*)/500$ for $x = 1, 0.5, 2$, and $\bar{U}_2 = \sum_{j=1}^{500} U_{2j}/500$ when increasing n.

As a benchmark, we also compute the empirical means replacing the GW model with a tapered Matérn covariance model, that is, considering the model $\mathscr{M}_{\nu,\alpha,\sigma_2^2} K_{x\beta_1^*}$, and we denote these means by $\bar{U}_1^T(x\beta_1^*)$, $x = 1, 0.5, 2$ and $\bar{U}_2^T$. Here, $K_{x\beta_1^*}$ is a known compactly supported correlation function called taper function. Following Furrer et al. (2006), as taper function, we use $K_{x\beta_1^*} = \varphi_{2,0,x\beta_1^*,1}$ if $\nu = 0.5$, $K_{x\beta_1^*} = \varphi_{3,1,x\beta_1^*,1}$ if $\nu = 1$ and $K_{x\beta_1^*} = \varphi_{4,2,x\beta_1^*,1}$ if $\nu = 1.5$. for $x = 1, 0.5, 2$.

These specific choices of taper functions guarantee the convergence of ratios (5.17) and (5.19), using a tapered Matérn model instead of the GW model (see Theorem 2 in Furrer et al., 2006). In Table 5.2, the percentages of nonzero elements in the covariance matrices are also reported in all scenarios and for each n when using the compact support β_1^*.

Table 5.2 shows that $\bar{U}_2$ clearly overall outperforms $\bar{U}_2^T$ in terms of speed of convergence in particular when increasing β_1^*. This implies that in terms of finite sample, if the Matern model is the state of nature, prediction efficiency and correct estimation of prediction variance are better achieved when predicting with the (compatible) GW model with respect to the so-called naive CT predictor (Furrer et al., 2006), sharing the same compact support.

Comparing $\bar{U}_1(x\beta_1^*)$ with $\bar{U}_1^T(x\beta_1^*)$ for $x = 1, 0.5, 2$ note that when $x = 1$, $\bar{U}_1(\beta_1^*)$ overall slightly outperforms $\bar{U}_1^T(\beta_1^*)$ and when $x = 0.5$, the convergence of both ratios seems to be very slow, in particular for larger ν. This suggests that taking an arbitrary compact support too small with respect to the "equivalent" compact support β_1^* can seriously affect the prediction efficiency both for tapered Matérn and GW models. This kind of problem disappears when $x = 2$, as expected. By the tapering effect, i.e., inducing a covariance with an apparent shorter range, $\bar{U}_1^T(2\beta_1^*)$ slightly outperforms $\bar{U}_1(2\beta_1^*)$.

Simulation results have been obtained using an upcoming version of the R package CompRandFld (Padoan and Bevilacqua, 2015a,b).

5.5 Concluding Remarks

Parameter estimation for interpolation of spatially or spatio-temporally correlated random processes is used in many areas and often requires particular models or careful implementation. In recent years the dataset sizes have steadily increased such that straightforward statistical tools are computationally too expensive to use. The use of covariance functions with an (inherent or induced) compact support, leading to sparse matrices, is a very accessible and scalable approach. In this paper we studied estimation and prediction of Gaussian fields with covariance models belonging to the GW class, under fixed domain asymptotics.

Specifically, we first characterize the equivalence of two Gaussian measures with GW models, and then we establish strong consistency and asymptotic Gaussianity of the ML estimator of the associated microergodic parameter when considering both an arbitrary and an estimated compact support. Simulation results show that for a finite sample, the choice of an arbitrary compact support can result in a very poor approximation of the asymptotic distribution. These results are consistent with those in Kaufman and Shaby (2013) in the Matérn case.

In a second aspect, we give a sufficient condition for the equivalence of two Gaussian measures with Matérn and GW model, and we study the effect on prediction when using these two covariance models under fixed domain asymptotics. A first consequence of our results is that GW model is more than a valid competitor of the Matérn model. It allows, as in the Matérn case, a continuous parameterization of smoothness of the underlying Gaussian field and, under fixed domain asymptotics, prediction and mean square error prediction obtained with a Matérn model can be achieved using a GW model inducing an equivalent Gaussian measure, using our condition. For this reason, we advocate the GW class when working with large or huge spatial datasets since well established and implemented algorithms for sparse matrices can be used when estimating the covariance parameters and/or predicting at unknown locations (e.g., Padoan and Bevilacqua, 2015a; Furrer and Sain, 2010). However, the approach does not impose any algorithmic constraints and iterative or hierarchical factorization schemes are possible as well, see, e.g., **??** and references therein.

As the theoretical and numerical results illustrate, CT for prediction is essentially an obsolete approach. When comparing both approaches with the same sensible compact support, the tapered CT is less efficient. For estimation, one has to distinguish between a so-called one-taper or two-taper approach, i.e., a proper likelihood or an estimating function approach, (Kaufman et al., 2008). Fixing again the support, a GW model can approximate a Matérn covariance function much better than a tapered one. Thus, the GW is in an estimation setting superior to a one-taper CT. In both approaches, one needs to be aware of the resulting biases, which can be substantial. In the case of (kriging) predictions based on plug-in estimates, the biases are largely canceled (Furrer et al., 2016). Finally, the two-taper approach is conceptually a different approach and, as it is computationally very expensive, it would not be fair to compare it with the GW model.

Similarly to the Matérn model with smoothness parameter different to $p+1/2$, $p \in \mathbb{N}$, the GW does not have a closed form expression when its smoothness parameter is different from p, and low level software implementations are needed for a computationally efficient use.

Table 5.2 $\bar{U}_1(x)$, $\bar{U}_1^T(x)$, $x = 0.5\beta_1^*, 2\beta_1^*, \beta_1^*$ and $\bar{U}_2$, $\bar{U}_2^T$, as defined in Section 5.4, when considering a Matérn model with increasing practical range y, smoothness parameter ν and n. Here β_1^* is the compact support parameter of the GW model computed using the equivalence condition. The column % reports the mean of percentages of non-zero elements in the covariance matrices involved when considering β_1^*.

		$\nu = 0.5$								$\nu = 1$								$\nu = 1.5$								
	n	$\bar{U}_1(0.5\beta_1^*)$	$\bar{U}_1(2\beta_1^*)$	$\bar{U}_1(\beta_1^*)$	$\bar{U}_1^T(0.5\beta_1^*)$	$\bar{U}_1^T(2\beta_1^*)$	$\bar{U}_1^T(\beta_1^*)$	$\bar{U}_2$	$\bar{U}_2^T$	$\bar{U}_1(0.5\beta_1^*)$	$\bar{U}_1(2\beta_1^*)$	$\bar{U}_1(\beta_1^*)$	$\bar{U}_1^T(0.5\beta_1^*)$	$\bar{U}_1^T(2\beta_1^*)$	$\bar{U}_1^T(\beta_1^*)$	$\bar{U}_2$	$\bar{U}_2^T$	$\bar{U}_1(0.5\beta_1^*)$	$\bar{U}_1(2\beta_1^*)$	$\bar{U}_1(\beta_1^*)$	$\bar{U}_1^T(0.5\beta_1^*)$	$\bar{U}_1^T(2\beta_1^*)$	$\bar{U}_1^T(\beta_1^*)$	$\bar{U}_2$	$\bar{U}_2^T$	%
	$\alpha = \frac{y}{c_\nu}$	$y = 0.1$								$y = 0.101$								$y = 0.097$								
$\beta_1^* = 0.1$	50	1.051	1.029	1.009	1.051	1.008	1.025	1.019	1.029	1.098	1.047	1.018	1.101	1.009	1.041	1.038	1.048	1.124	1.054	1.024	1.134	1.012	1.057	1.050	1.056	4.67
	100	1.096	1.043	1.014	1.096	1.013	1.043	1.035	1.056	1.189	1.076	1.028	1.195	1.013	1.072	1.073	1.097	1.246	1.095	1.039	1.266	1.019	1.105	1.098	1.112	3.70
	250	1.182	1.046	1.019	1.183	1.018	1.069	1.064	1.118	1.379	1.097	1.038	1.393	1.016	1.121	1.138	1.204	1.521	1.156	1.059	1.567	1.025	1.197	1.197	1.241	3.12
	500	1.267	1.030	1.015	1.268	1.016	1.077	1.081	1.211	1.608	1.065	1.030	1.639	1.011	1.132	1.187	1.372	1.928	1.116	1.051	2.039	1.020	1.253	1.300	1.481	2.92
	1000	1.325	1.015	1.009	1.330	1.010	1.061	1.073	1.332	1.858	1.032	1.016	1.923	1.005	1.088	1.168	1.586	2.549	1.054	1.025	2.820	1.008	1.209	1.300	1.877	2.82
	$\alpha = \frac{y}{c_\nu}$	$y = 0.2$								$y = 0.202$								$y = 0.193$								
$\beta_1^* = 0.2$	50	1.151	1.044	1.016	1.157	1.016	1.058	1.053	1.134	1.316	1.090	1.032	1.326	1.014	1.094	1.119	1.217	1.448	1.139	1.048	1.505	1.021	1.149	1.177	1.288	11.95
	100	1.209	1.032	1.013	1.221	1.015	1.066	1.068	1.235	1.471	1.068	1.027	1.491	1.011	1.103	1.162	1.377	1.730	1.120	1.045	1.848	1.018	1.186	1.266	1.534	11.04
	250	1.227	1.012	1.007	1.247	1.008	1.046	1.060	1.397	1.578	1.026	1.013	1.614	1.004	1.061	1.146	1.590	2.085	1.049	1.022	2.363	1.007	1.137	1.271	1.954	10.48
	500	1.152	1.005	1.003	1.178	1.003	1.021	1.040	1.513	1.415	1.009	1.005	1.447	1.002	1.022	1.092	1.625	1.945	1.017	1.009	2.321	1.004	1.051	1.174	2.069	10.27
	1000	1.061	1.002	1.001	1.083	1.001	1.007	1.024	1.586	1.145	1.003	1.002	1.152	1.002	1.014	1.052	1.497	1.358	1.005	1.003	1.554	1.003	1.029	1.093	1.728	10.18
	$\alpha = \frac{y}{c_\nu}$	$y = 0.4$								$y = 0.404$								$y = 0.385$								
$\beta_1^* = 0.4$	50	1.208	1.016	1.008	1.226	1.010	1.050	1.060	1.372	1.507	1.035	1.016	1.530	1.005	1.072	1.148	1.519	1.900	1.066	1.027	2.088	1.010	1.152	1.271	1.823	34.97
	100	1.151	1.006	1.004	1.174	1.004	1.026	1.041	1.499	1.399	1.013	1.007	1.421	1.002	1.030	1.100	1.583	1.846	1.024	1.011	2.106	1.004	1.071	1.196	2.006	34.18
	250	1.050	1.001	1.001	1.066	1.001	1.006	1.020	1.598	1.128	1.003	1.002	1.141	1.001	1.011	1.048	1.454	1.328	1.006	1.003	1.491	1.003	1.028	1.091	1.691	33.90
	500	1.013	1.000	1.000	1.019	1.000	1.002	1.011	1.633	1.024	1.001	1.001	1.044	1.001	1.009	1.025	1.314	1.053	1.002	1.001	1.121	1.001	1.019	1.047	1.373	33.66
	1000	1.003	1.000	1.000	1.005	1.000	1.000	1.006	1.649	1.003	1.000	1.000	1.034	1.000	1.006	1.014	1.208	1.005	1.000	1.000	1.077	1.000	1.009	1.024	1.184	33.59

Open questions and future works

5.6 Covariance family

There exist an interesting point of equivalence of two Gaussian measures in 4-dimensional space. Anderes (2010) proved the orthogonality of two Gaussian measures, for $d \geq 5$, with different Matérn covariance functions. However, the case $d = 4$ is still open. So, all the work of this thesis can reproduce for this case if it is possible to prove the equivalence of two Gaussian measures with Matérn, Wendland, Cauchy, and Dagum family.

5.7 Generalized Cauchy covariance model

A future investigation work is estimation and prediction of Gaussian random fields with covariance models belonging to the generalized Cauchy (GC) class, under fixed domain asymptotics. Gaussian random field with this kind of covariance provide separate characterization of fractal dimension and long range dependence, an appealing feature in many physical, biological or geological systems.

The results of the future work can be classified into three parts: first, we characterize the equivalence of two Gaussian measures with GC covariance function and we provide sufficient conditions for the equivalence of two Gaussian measures with Matérn (MT) and GC covariance functions and two Gaussian Measures with Generalized Wendland (GW) and GC covariance functions.

In the second part, we establish strong consistency and asymptotic distribution of the maximum likelihood estimator of the microergodic parameter associated with GC covariance model, under fixed domain asymptotics. The third part elucidates the consequences of our results in terms of prediction under fixed domain asymptotics.

5.7.1 Technical point

Another important point is to establish strong consistency and asymptotic distribution of the sequence of random variables $\hat{\sigma}_n^2(\hat{\gamma}_n)\lambda/\hat{\gamma}_n^{\delta}$. In fact, computationally we can prove the following conjecture result.

Lemma 2. *The combination of spectral density $\widehat{\mathscr{C}}_{\sigma^2,\lambda,\delta,\gamma}$ and power, $\gamma^{\delta+2}\widehat{\mathscr{C}}_{\sigma^2,\lambda,\delta,\gamma}(z)$, is not decreasing with respect to γ, for $\delta \in (1,2)$, $d = 1,2,3$ and $\lambda > \frac{d}{2}+2$.*

We can illustrate this lemma with the following figure for a specifical values of λ, d, and δ.

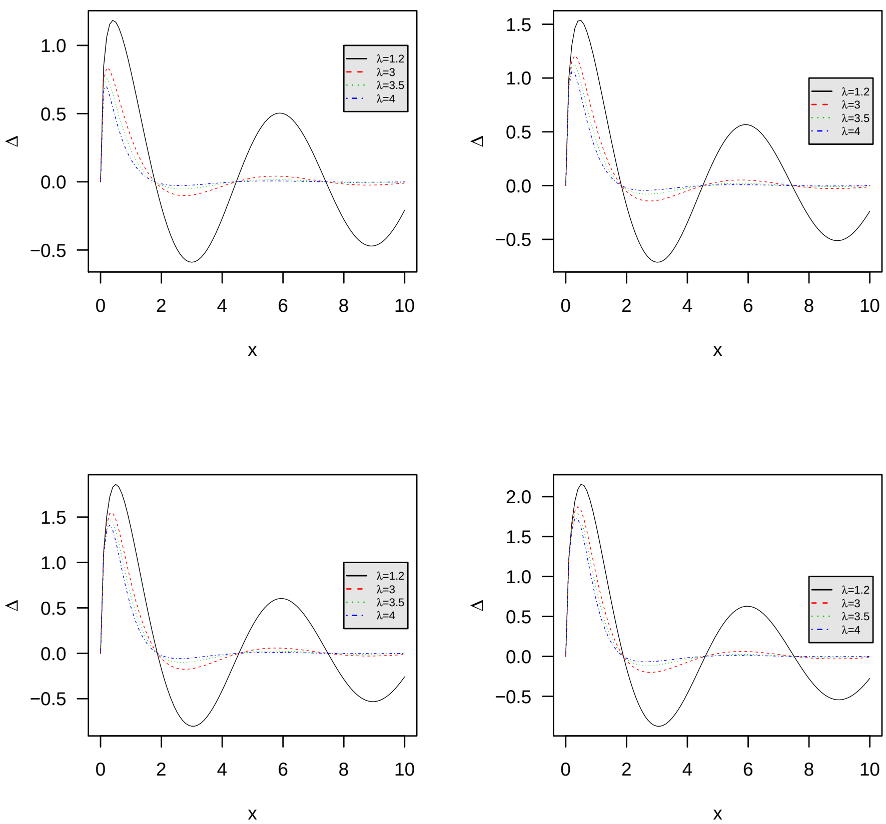

Fig. 5.4 $\Delta' = \frac{x^{d/2+1}}{(1+x^{\delta})^{\lambda\delta}}J_{\nu-1}(x) + \frac{(2+\delta)x^{d/2}}{(1+x^{\delta})^{\lambda/\delta}}J_{\nu}(x)$'s Derivative for the case $d = 1$ and $\delta = 1, 1.3, 1.6, 1.9$.

We observe that the integral of Δ' is positive when $\lambda \geq \frac{d}{2}+2$.

Note that the integral of Δ' for each case is positive when $\lambda \geq \frac{d}{2}+2$ and $d = 2$.

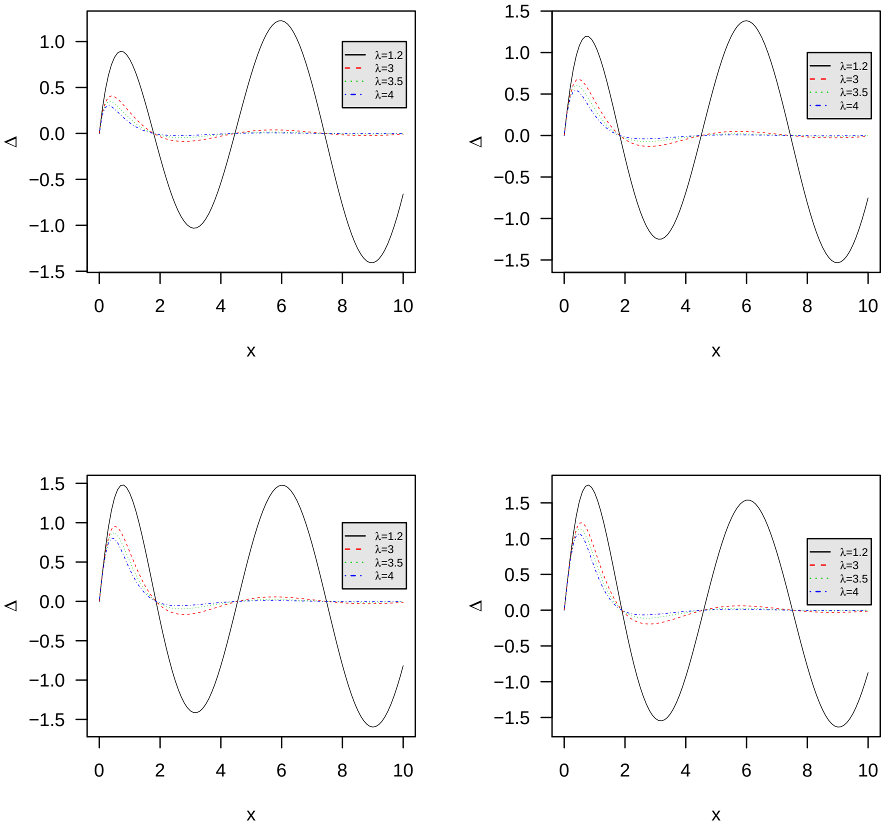

Fig. 5.5 $\Delta' = \frac{x^{d/2+1}}{(1+x^\delta)^{\lambda\delta}} J_{\nu-1}(x) + \frac{(2+\delta)x^{d/2}}{(1+x^\delta)^{\lambda/\delta}} J_\nu(x)$'s Derivative for the case $d = 2$ and $\delta = 1, 1.3, 1.6, 1.9$.

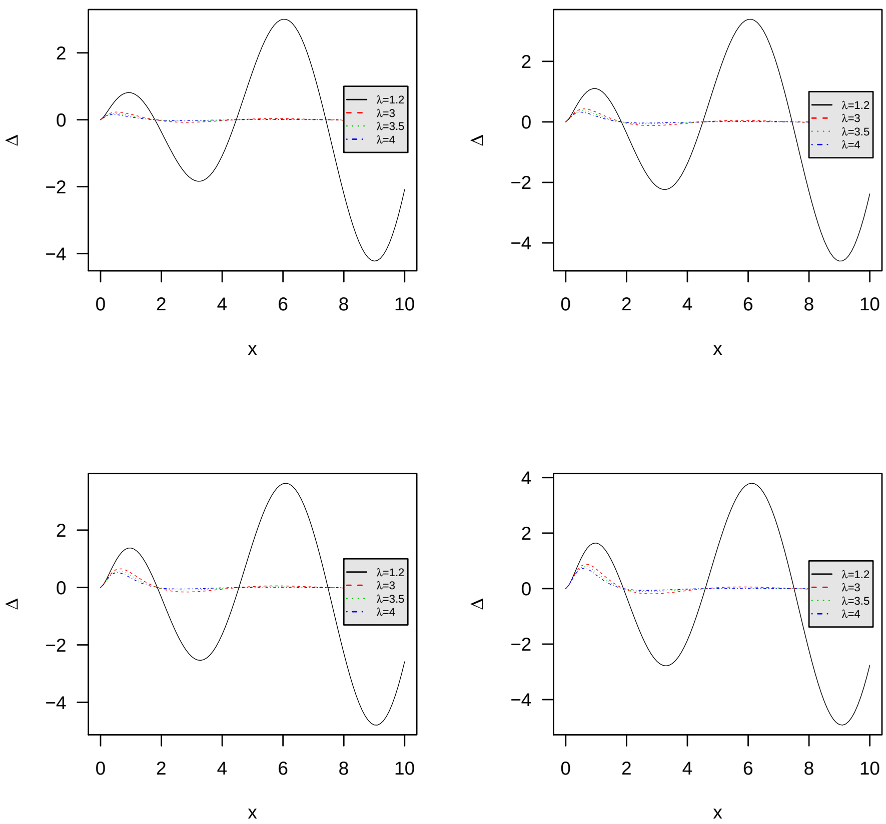

Fig. 5.6 $\Delta' = \frac{x^{d/2+1}}{(1+x^{\delta})^{\lambda\delta}} J_{\nu-1}(x) + \frac{(2+\delta)x^{d/2}}{(1+x^{\delta})^{\lambda/\delta}} J_{\nu}(x)$'s Derivative for the case $d = 3$ and $\delta = 1, 1.3, 1.6, 1.9$.

References

Abrahamsen, P. (1997). *A Review of Gaussian Random Fields and Correlation Functions*. Norsk Regnesentral/Norwegian Computing Center, second edition.

Abramowitz, M. and Stegun, I. A., editors (1970). *Handbook of Mathematical Functions*. Dover, New York.

Adler, R. J. (1981). *The Geometry of Random Fields*. Wiley, New York.

Anderes, E. (2010). On the consistent separation of scale and variance for gaussian random fields. *The Annals of Statistics*, 38(2):870–893.

Askey, R. (1973). Radial characteristic functions. *Technical report, Research Center, University of Wisconsin*.

Bochner, S. (1933). Monotone funktionen, stieltjessche integrale und harmonische analyse. *Mathematische Annalen*, 40:378–410.

Chernih, A. and Hubbert, S. (2014). Closed form representations and properties of the generalised wendland functions. *Journal of Approximation Theory*, 40:17–33.

Chernih, A., Sloan, I. H., and Womersley, R. S. (2014). Wendland functions with increasing smoothness converge to a gaussian. *Advances in Computational Mathematics*, 40:17–33.

Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York, revised edition.

Du, J., Zhang, H., and Mandrekar, V. S. (2009). Fixed-domain asymptotic properties of tapered maximum likelihood estimators. *The Annals of Statistics*, 37:3330–3361.

Erdelyi, A., Magnus, W., Oberhettinger, F., and Tricomi, F., editors (1954). *Tables of integral transforms, vol. I, II*. McGraw-Hill, New York.

Fields, J. L. and Ismail, M. E. H. (1975). On the positivity of some ${}_1f_2$'s. *SIAM Journal on Mathematical Analysis*, 6:551–559.

Fourier, M. (1807). Mémoire sur la propagation de la chaleur dans les corps solides. 6:112–116.

Furrer, R., Bachoc, F., and Du, J. (2016). Asymptotic properties of multivariate tapering for estimation and prediction. *Journal of Multivariate Analysis*, 149:177–191.

Furrer, R., Genton, M. G., and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 15:502–523.

Furrer, R. and Sain, S. R. (2010). spam: a sparse matrix R package with emphasis on mcmc methods for gaussian markov random rields. *Journal of Statistical Software*, 36:1–25.

Gneiting, T. (2002a). Compactly supported correlation functions. *Journal of Multivariate Analysis*, 83:493–508.

Gneiting, T. (2002b). Stationary covariance functions for space-time data. *Journal of the American Statistical Association*, 97:590–600.

Gneiting, T. and Schlather, M. (2004). Stochastic models that separate fractal dimension and the hurst effect. *SIAM review*, 46(2):269–282.

Golubov, B. I. (1981). On abel-poisson type and riesz means. *Anal. Mathematica*, 7:161–184.

Hirano, T. and Yajima, Y. (2013). Covariance tapering for prediction of large spatial data sets in transformed random fields. *Annals of the Institute of Statistical Mathematics*, 65(5):913–939.

Horn, R. A. and Johnson, C. R. (1991). *Topics in Matrix Analysis*. Cambridge University Press, Cambridge.

Ibragimov, I. A. and Rozanov, Y. A. (1978). *Gaussian Random Processes*. Springer, New York.

Kaufman, C. G., Schervish, M. J., and Nychka, D. W. (2008). Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103:1545–1555.

Kaufman, C. G. and Shaby, B. A. (2013). The role of the range parameter for estimation and prediction in geostatistics. *Biometrika*, 100:473–484.

Luke, Y., editor (1969). *The Special Functions and Their Approximations. Vol. 1*. Academic Press, New York.

Mardia, K. V. and Marshall, J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, 71:135–146.

Matérn, B. (1960). *Spatial Variation*. Meddelanden Från Statens Skogsforskningsinstitut, Band 49, Nr 5., Stockholm.

Padoan, S. and Bevilacqua, M. (2015a). Analysis of random fields using comprandfld. *Journal of the Statistical Software*, 63:1–27.

Padoan, S. and Bevilacqua, M. (2015b). *CompRandFld: Composite-Likelihood Based Analysis of Random Fields*. R package version 2.0.

Porcu, E., Montero, J., and Schlather, M. (2012). *Advances and Challenges in Space-time Modelling of Natural Events*. Lecture Notes in Statistics. Springer Berlin Heidelberg.

Porcu, E., Zastavnyi, V., and Bevilacqua, M. (2016). Buhmann covariance functions, their compact supports, and their smoothness. *Submitted*.

Porcu, E., Zastavyi, P., and Bevilacqua, M. (2016). Buhmann covariance functions, their compact supports, and their smoothness. *ArXiv e-prints*.

Putter, H. and Young, G. (2001). On the effect of covariance function estimation on the accuracy of kriging predictors. *Biometrika*, 7:421–438.

R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schaback, R. (2011). The missing wendland functions. *Advances in Computational Mathematics*, 34(1):67–81.

Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. *Ann. Math.*, 39:811–841.

Skorokhod, A. and Yadrenko, M. (1973). On absolute continuity of measures corresponding to homogeneous gaussian fields. *Theory of Probability and Its Applications*, 18:27–40.

Stein, M. (1988). Asymptotically efficient prediction of a random field with a misspecified covariance function. *Annals of Statistics*, 16:55–63.

Stein, M. L. (1990). Uniform asymptotic optimality of linear predictions of a random field using an incorrect second order structure. *Annals of Statistics*, 19:850–872.

Stein, M. L. (1993). A simple condition for asymptotic optimality of linear predictions of random fields. *Statistic and Probability Letters*, 17:399–404.

Stein, M. L. (1999a). *Interpolation of Spatial Data. Some Theory of Kriging*. Springer-Verlag, New York.

Stein, M. L. (1999b). Predicting random frields with increasing dense observations. *Annals of Applied Probability*, 9:242–273.

Stein, M. L. (2004). Equivalence of gaussian measures for some nonstationary random fields. *Journal of Statistical Planning and Inference*, 123:1–11.

Stein, M. L. (2013). Statistical properties of covariance tapers. *J. Comput. Graph. Stat.*, 22(4):866–885.

Wackernagel, H. (2003). *Multivariate Geostatistics: An Introduction with Applications*. Springer, New York, 3rd edition.

Wang, D. (2010). *Fixed domain asymptotics and consistent estimation for gaussian random field models in spatial statistics and computer experiments*. PhD thesis, University of Science and Technology of China, Department of statistics and applied probability national university of Singapore.

Wang, D. and Loh, W.-L. (2011). On fixed-domain asymptotics and covariance tapering in gaussian random field models. *Electronic Journal of Statistics*, 5:238–269.

Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv. Comput. Math.*, 4:389–396.

Wu, Z. (1995). Compactly supported positive definite radial functions. *Advances in Computational Mathematics*, 4(1):283–292.

Yaglom, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions. Volume I: Basic Results*. Springer, New York.

Zastavnyi, V. and Trigub, R. (2002). Positive definite splines of special form. *English transl. in Sb. Math.*, 193:1771–1800.

Zastavnyi, V. P. (2006). On some properties of buhmann functions. *Ukrainian Mathematical Journal*, 58(8):1184–1208.

Zhang, H. (2004). Inconsistent estimation and asymptotically equivalent interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99:250–261.

Zhang, H. and Zimmerman, D. (2005). Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika*, 92:921–936.

Appendix A

Chapter I

For a neater exposition, the lemmas are presented inside the proof.

Proof of Theorem 13. The proof of the first assertion follows the same arguments of the proof of Theorem 3 in Zhang (2004), and we omit it.

For the proof of the second assertion, we follow the arguments in Wang and Loh (2011) and Wang (2010), applied to the GW case. As in Wang and Loh (2011), without loss of generality, we assume $D = [0,T]^d$, $0 < T < \infty$ is a bounded subset of $\mathbb{R}^d$, $d = 1,2,3$. Let $R_\beta = R_n(\beta)$ and $\hat{\sigma}_n^2 = \hat{\sigma}_n^2(\beta)$ for notation convenience and let σ^2 and β two positive constants such that $\sigma^2\beta^{-(1+2\kappa)} = \sigma_0^2\beta_0^{-(1+2\kappa)}$. Then we have

$$\sqrt{n}\left(\hat{\sigma}_n^2\beta^{-(1+2\kappa)} - \sigma_0^2\beta_0^{-(1+2\kappa)}\right) = \frac{\sigma_0^2\beta_0^{-(1+2\kappa)}}{\sqrt{n}}\left(\frac{1}{\sigma^2}\mathbf{Z}_n'R_\beta^{-1}\mathbf{Z}_n - \frac{1}{\sigma_0^2}\mathbf{Z}_n'R_{\beta_0}^{-1}\mathbf{Z}_n\right) + \frac{\sigma_0^2\beta_0^{-(1+2\kappa)}}{\sqrt{n}}\left(\frac{1}{\sigma_0^2}\mathbf{Z}_n'R_{\beta_0}^{-1}\mathbf{Z}_n - n\right).$$

Under the measure $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2})$, we have $\sigma_0^{-2}\mathbf{Z}_n'R_{\beta_0}^{-1}\mathbf{Z}_n \sim \chi_n^2$ and

$$\frac{\sigma_0^2\beta_0^{-(1+2\kappa)}}{\sqrt{n}}\left(\frac{1}{\sigma_0^2}\mathbf{Z}_n'R_{\beta_0}^{-1}\mathbf{Z}_n - n\right) \xrightarrow{\mathscr{D}} \mathscr{N}(0, 2\sigma_0^2\beta_0^{-(1+2\kappa)})$$

as $n \to \infty$. Thus to prove the result, it is sufficient to show that

$$\frac{1}{\sqrt{n}}\left(\frac{1}{\sigma^2}\mathbf{Z}_n'R_\beta^{-1}\mathbf{Z}_n - \frac{1}{\sigma_0^2}\mathbf{Z}_n'R_{\beta_0}^{-1}\mathbf{Z}_n\right) \xrightarrow{p} 0, \text{ as } n \to \infty, \tag{A.1}$$

under $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2})$.

Specifically we need to show that for any $\vartheta > 0$,

$$\begin{aligned} &P_{\sigma_0^2,\beta_0}\left(\frac{1}{\sqrt{n}}\left|\frac{1}{\sigma^2}\mathbf{Z}_n' R_\beta^{-1}\mathbf{Z}_n - \frac{1}{\sigma_0^2}\mathbf{Z}_n' R_{\beta_0}^{-1}\mathbf{Z}_n\right| > \vartheta\right) \\ =&P_{\sigma_0^2,\beta_0}\left(\frac{1}{\sqrt{n}}\left|\sum_{k=1}^{n}(\lambda_{k,n}^{-1}-1)Y_k^2\right| > \vartheta\right) \longrightarrow 0, \text{ as } n \to \infty, \end{aligned} \tag{A.2}$$

where Y_k and $\lambda_{k,n}$ are defined below.

Following Wang and Loh (2011), the quantity in (A.1) can be written as

$$\frac{1}{\sqrt{n}}\sum_{k=0}^{n}(\lambda_{k,n}^{-1}-1)Y_k^2, \tag{A.3}$$

where $(Y_1,\dots,Y_n)' \sim \mathscr{N}_n(0,I_n)$ under $P(\varphi_{\mu,\kappa,\beta_0,\sigma_0^2})$ and $\lambda_{k,n}$, $k=1,\dots,n$, satisfy

$$\sigma^2\left[\sigma_0^{-1}R_{\beta_0}^{-1/2}\right]' R_\beta\left[\sigma_0^{-1}R_{\beta_0}^{-1/2}\right] = \operatorname{diag}(\lambda_{k,n})_{k\in\{1,\dots,n\}}.$$

For the rest of the proof $|\cdot|$ denotes the Euclidean norm, and $|\mathbf{x}|_{\max} = \max\{|x_1|,\dots,|x_d|\}$ with $\mathbf{x} = (x_1,\dots,x_d)' \in \mathbb{R}^d$.

Let $\xi_0 : \mathbb{R}^d \to \mathbb{R}$ be defined as $\xi_0(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} \mathrm{e}^{-i\mathbf{x}'\boldsymbol{\omega}} c_0(\mathbf{x})\mathrm{d}\mathbf{x}$, where $c_0(\mathbf{x}) = |\mathbf{x}|^{\zeta-d} I\{|\mathbf{x}| \leq 1\}$ and $\zeta = \frac{d+1+2\kappa}{2m}$, with $m = \lfloor d+1+2\kappa\rfloor + 1$. Here, $\lfloor x\rfloor$ is the largest integer less than or equal to x and $I\{\cdot\}$ denotes the indicator function.

By Theorem 2 of Fields and Ismail (1975), ξ_0 is a positive function for $d \geq 2$. To cover the case $d = 1$, we need the following lemma.

Lemma 3. *For* $0 < \upsilon < 1/2$*, we have*

$${}_1F_2\left(\upsilon/2;\upsilon/2+1,1/2;-(r/2)^2\right) > 0, \qquad \forall r > 0.$$

Proof of Lemma 3. For all $r > 0$ we have

$${}_1F_2(\upsilon/2;\upsilon/2+1,1/2;-(r/2)^2) = \upsilon r^{-1}\int_0^r \cos(x)x^{\upsilon-1}\mathrm{d}x.$$

From the arguments in Wang (2010, page 32), we have that the integral, is a positive, which concludes the proof. □

Lemma 4. *The function* $\xi_0 : \mathbb{R}^d \to \mathbb{R}$ *is a continuous, isotropic strictly positive function and* $\xi_0(\boldsymbol{\omega}) \asymp |\boldsymbol{\omega}|^{-\zeta}$ *as* $|\boldsymbol{\omega}| \to \infty$.

Proof of Lemma 4. Let U_d be the uniform probability measure on $S^{d-1} = \{\mathbf{u} \in \mathbb{R}^d : |\mathbf{u}| = 1\}$. By isotropy, we have for all $\boldsymbol{\omega} \in \mathbb{R}^d$

$$\begin{aligned}
\xi_0(\boldsymbol{\omega}) &= \int_{|\mathbf{x}|\le 1} \int_{S^{d-1}} e^{-i|\boldsymbol{\omega}|\mathbf{u}'\mathbf{x}} |\mathbf{x}|^{-\zeta-d} U_d(\mathrm{d}\mathbf{u})\mathrm{d}\mathbf{x} \\
&= (2\pi)^{\frac{d}{2}} |\boldsymbol{\omega}|^{\frac{2-d}{2}} \int_0^1 r^{\zeta-\frac{d}{2}} J_{\frac{d-2}{2}}(|\boldsymbol{\omega}|r)\mathrm{d}r \\
&= (2\pi)^{\frac{d}{2}} |\boldsymbol{\omega}|^{-\zeta} \int_0^{|\boldsymbol{\omega}|} r^{\zeta-\frac{d}{2}} J_{\frac{d-2}{2}}(r)\mathrm{d}r \\
&= 2\zeta^{-1}\pi^{d/2}\Gamma(d/2)^{-1} {}_1F_2(\zeta/2;\zeta/2+1,d/2;-(|\boldsymbol{\omega}|/2)^2).
\end{aligned} \tag{A.4}$$

From Theorem 2 in Fields and Ismail (1975) we get that $\xi_0 > 0$ for $d \ge 2$, and using Lemma 3 we have $\xi_0 > 0$ for $d = 1$. Then ξ_0 is a continuous, isotropic and strictly positive on $\mathbb{R}^d$.

Moreover, from Luke (1969, p. 203 (4)) we have, as $|\boldsymbol{\omega}| \to \infty$,

$$\begin{aligned}
{}_1F_2(\zeta/2;\zeta/2+1,d/2;-(|\boldsymbol{\omega}|/2)^2) &= \frac{2^\zeta\Gamma(d/2)}{\Gamma(d/2-\zeta/2)}|\boldsymbol{\omega}|^{-\zeta} \\
&+ \frac{\Gamma(d/2)}{\pi^{1/2}\Gamma(\zeta/2)}|\boldsymbol{\omega}|^{-(1+d)/2}\exp\big(4w_3|\boldsymbol{\omega}|^{-2} + \mathscr{O}(|\boldsymbol{\omega}|^{-4})\big) \\
&\times \cos\Big(|\boldsymbol{\omega}| - \frac{\pi(d+1)}{2} - 2w_4|\boldsymbol{\omega}|^{-1} - 8w_5|\boldsymbol{\omega}|^{-3} + \mathscr{O}(|\boldsymbol{\omega}|^{-5})\Big),
\end{aligned}$$

where $\{w_k\}_{k=3,4,5}$ are constants not depending on $\boldsymbol{\omega} \in \mathbb{R}^d$. Thus

$${}_1F_2(\zeta/2;\zeta/2+1,d/2;-(|\boldsymbol{\omega}|/2)^2) \asymp \frac{2^\zeta\Gamma(d/2)}{\Gamma(d/2-\zeta/2)}|\boldsymbol{\omega}|^{-\zeta},$$

which implies, in concert with Equation (A.4), that $\xi_0(\boldsymbol{\omega}) \asymp |\boldsymbol{\omega}|^{-\zeta}$, as $|\boldsymbol{\omega}| \to \infty$. $\square$

Let $\xi_1(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} e^{-i\boldsymbol{\omega}'\mathbf{x}} c_1(\mathbf{x})\mathrm{dx} = \xi_0(\boldsymbol{\omega})^{2m}$, for all $\boldsymbol{\omega} \in \mathbb{R}^d$, where $c_1 = c_0 * \ldots * c_0$ denote the $2m$-fold convolution of the function c_0 with itself. We define,

$$\eta(\boldsymbol{\omega}) = \frac{\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2}(|\boldsymbol{\omega}|) - \widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(|\boldsymbol{\omega}|)}{\xi_1(\boldsymbol{\omega})}, \qquad \forall \boldsymbol{\omega} \in \mathbb{R}^d.$$

From Theorem 3 Point 3 and Lemma 4, we have

$$\frac{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(|\boldsymbol{\omega}|)}{\xi_1(\boldsymbol{\omega})} \asymp 1, \qquad \text{as } |\boldsymbol{\omega}| \to \infty.$$

Furthermore, this ratio is well definite and continuous on arbitrary compact interval belong to $\mathbb{R}_+$ with $\xi_1 > 0$, so there exist two constants c_{ξ_1} and C_{ξ_1} not depend on $|\boldsymbol{\omega}|$ such that

$$c_{\xi_1} \leq \frac{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(|\boldsymbol{\omega}|)}{\xi_1(\boldsymbol{\omega})} \leq C_{\xi_1}, \qquad \text{as } |\boldsymbol{\omega}| \to \infty. \tag{A.5}$$

Thus, for a constant $C_\eta > 0$, we have

$$\begin{aligned}\int_{\mathbb{R}^d} \eta(\boldsymbol{\omega})^2 \mathrm{d}\boldsymbol{\omega} = \frac{2\pi^{d/2}}{\Gamma(d/2)} \Bigg[\int_0^{C_\eta} r^{d-1} \Big(\frac{\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2}(r) - \widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(r)}{\xi_1(\boldsymbol{r})} \Big)^2 \mathrm{d}r \\ + \int_{C_\eta}^{\infty} r^{d-1} \Big(\frac{\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2}(r) - \widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(r)}{\xi_1(\boldsymbol{r})} \Big)^2 \mathrm{d}r \Bigg].\end{aligned} \tag{A.6}$$

where $\boldsymbol{r} \in \mathbb{R}^d$, with $|\boldsymbol{r}| = r$.
Since $d = 1,2,3$, $\mu > \lambda + d/2$, $\sigma^2 \beta^{-(1+2\kappa)} = \sigma_0^2 \beta_0^{-(1+2\kappa)}$, both terms of Equation (A.6) are finite. Thus, η is square integrable. From the theory of Fourier transforms of $\mathrm{L}^2(\mathbb{R}^d)$ functions, there exists a square integrable function $g : \mathbb{R}^d \to \mathbb{R}$ such that

$$\int_{\mathbb{R}^d} \big(\eta(\boldsymbol{\omega}) - \hat{g}_k(\boldsymbol{\omega})\big)^2 \mathrm{d}\boldsymbol{\omega} \to 0, \quad \text{as } k \to \infty,$$

where

$$\hat{g}_k(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} \mathrm{e}^{-i\boldsymbol{\omega}'\mathbf{x}} g(\mathbf{x}) I\{|\mathbf{x}|_{\max} \leq k\} \mathrm{d}\mathbf{x}, \qquad \forall \boldsymbol{\omega} \in \mathbb{R}^d, \quad k > 0. \tag{A.7}$$

In order to illustrate the following Lemma, some notation is needed. According to Equation (2.44) of Wang (2010), define

$$e_n(\mathbf{x}) = \frac{1}{C_e \varepsilon_n^d} \tilde{c}_1 \left(\frac{\mathbf{x}}{\varepsilon_n} \right), \qquad \forall \mathbf{x} \in \mathbb{R}^d, \tag{A.8}$$

and

$$\tilde{\xi}_1(\boldsymbol{\omega}) = \int_{\mathbb{R}} \mathrm{e}^{-i\boldsymbol{\omega}'\mathbf{x}} \tilde{c}_1(\mathbf{x}) \mathrm{d}\mathbf{x},$$

where $C_e = \int_{\mathbb{R}^d} \tilde{c}_1(\mathbf{x}) \mathrm{d}\mathbf{x}$ and $\tilde{c}_1 = \tilde{c}_0 * \ldots * \tilde{c}_0$ with $\tilde{c}_0(\mathbf{x}) = |\mathbf{x}|^{\frac{a+d+1}{2m_a} - d} I\{|\mathbf{x}| \leq 1\}$ and $m_a = \lfloor a + d + 1 \rfloor + 1$. Here a is an arbitrary positive constant. Write

$$\hat{e}_n(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} \mathrm{e}^{-i\boldsymbol{\omega}'\mathbf{x}} \mathrm{e}_n(\mathbf{x}) \mathrm{d}\mathbf{x} = \frac{\tilde{\xi}_1(\varepsilon_n \boldsymbol{\omega})}{C_e}$$

for the Fourier transform of e_n. Note that there exists a constant $C_{\hat{e}}$ not depending on $\boldsymbol{\omega}$ and n such that

$$|\hat{e}_n(\boldsymbol{\omega})| \leq \frac{C_{\hat{e}}}{(1+\varepsilon_n|\boldsymbol{\omega}|)^{a+d+1}}, \quad \forall \boldsymbol{\omega} \in \mathbb{R}^d. \tag{A.9}$$

Lemma 5. *Let* $(\varepsilon_n)_n : \varepsilon_n \in (0,1]$, $\forall n \in \mathbb{N}$, *and additionally,* $\varepsilon_n \to 0$, *when* $n \to \infty$. *Let* g *as in Equation* (A.7), e_n *as in Equation* (A.8), *and* ι_0 *a constant satisfying* $0 < \iota_0 < \min\{2(\mu-\lambda-d/2), 4-d\}$. *Then, there exists a constant* C_{ι_0} *such that*

$$\int_{\mathbb{R}^d} \left| e_n * g(\mathbf{x}) - g(\mathbf{x}) \right|^2 \mathrm{d}\mathbf{x} \leq C_{\iota_0} \varepsilon_n^{\iota_0}. \tag{A.10}$$

Proof. Lemma (5) can be proved by noting that

$$\begin{aligned}\int_{\mathbb{R}^d} \left| g(\mathbf{x}-\mathbf{y}) - g(\mathbf{x}) \right|^2 \mathrm{d}\mathbf{x} &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |(e^{-i\boldsymbol{w}'\mathbf{y}} - 1)\eta(\boldsymbol{w})|^2 \mathrm{d}\boldsymbol{w} \\ &\leq \frac{2^{2-\iota_0}|\mathbf{y}|^{\iota_0}}{(2\pi)^d} \int_{\mathbb{R}^d} |\boldsymbol{w}|^{\iota_0} |\eta(\boldsymbol{w})|^2 \mathrm{d}\boldsymbol{w}\end{aligned}$$

and

$$\begin{aligned}\left[\int_{\mathbb{R}^d} \left| e_n * g(\mathbf{x}) - g(\mathbf{x}) \right|^2 \mathrm{d}\mathbf{x} \right]^{1/2} &= \left[\int_{\mathbb{R}^d} \int_{|\mathbf{y}| \leq 2m_a\varepsilon_n} |(g(\mathbf{x}-\mathbf{y}) - g(\mathbf{x}))e_n(\mathbf{y})\mathrm{d}\mathbf{y}|^2 \mathrm{d}\mathbf{x} \right]^{1/2} \\ &\leq \frac{2^{(2-\iota_0)/2}(2m_a\varepsilon_n)^{\iota_0/2}}{(2\pi)^d} \left[\int_{\mathbb{R}^d} |\boldsymbol{w}|^{\iota_0} |\eta(\boldsymbol{w})|^2 \mathrm{d}\boldsymbol{w} \right]^{1/2}.\end{aligned}$$

We know that $\int_{\mathbb{R}^d} |\boldsymbol{w}|^{\iota_0} |\eta(\boldsymbol{w})|^2 \mathrm{d}\boldsymbol{w}$ is finite, so the proof is completed. □

Let $b(\mathbf{x},\mathbf{y}) = E_{\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2},n}[Z(\mathbf{x})Z(\mathbf{y})] - E_{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2},n}[Z(\mathbf{x})Z(\mathbf{y})]$, $\forall \mathbf{x},\mathbf{y} \in D = [0,T]^d$. From Wang and Loh (2011, (2.24)), and observing that $supp(c_1) \subseteq [-2m, 2m]^d$, we obtain for $\mathbf{x}$, $\mathbf{y} \in D$,

$$\begin{aligned} b(\mathbf{x},\mathbf{y}) &= (2\pi)^d \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(\mathbf{s}-\mathbf{t}) c_1(\mathbf{x}-\mathbf{s}) c_1(\mathbf{y}-\mathbf{t}) \mathrm{d}\mathbf{s}\mathrm{d}\mathbf{t} \\ &= (2\pi)^d \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e_n * g(\mathbf{s}-\mathbf{t}) c_1(\mathbf{x}-\mathbf{s}) c_1(\mathbf{y}-\mathbf{t}) \mathrm{d}\mathbf{s}\mathrm{d}\mathbf{t} \\ &\quad + (2\pi)^d \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h_n^*(\mathbf{s},\mathbf{t}) c_1(\mathbf{x}-\mathbf{s}) c_1(\mathbf{y}-\mathbf{t}) \mathrm{d}\mathbf{s}\mathrm{d}\mathbf{t}, \end{aligned}$$

where $h_n^*(\mathbf{s},\mathbf{t}) = \left[g(\mathbf{s}-\mathbf{t}) - e_n * g(\mathbf{s}-\mathbf{t}) \right] \mathbf{I}\{|\mathbf{s}+\mathbf{t}|_{\max} \leq 4m + 2T\}, \forall \mathbf{s}, \mathbf{t} \in \mathbb{R}^d$.

Let $\eta_n^* : \mathbb{R}^d \to \mathbb{C}$ denote the Fourier transform of $g - e_n * g$, this implies that

$$\int_{\mathbb{R}^d} \left| \eta_n^*(\boldsymbol{w}) - \hat{g}_{n,k}^*(\boldsymbol{w}) \right|^2 \mathrm{d}\boldsymbol{w} \to 0, \quad \text{as } k \to \infty, \tag{A.11}$$

where $\hat{g}^*_{n,k}(\boldsymbol{w}) = \int_{\mathbb{R}^d} e^{-i\boldsymbol{w}'\mathbf{x}}\big[g(\mathbf{x}) - e_n * g(\mathbf{x})\big]\mathbf{I}\{|\mathbf{x}|_{\max} \le k\}\mathrm{d}\mathbf{x}$.

Thus as in Wang (2010, (2.27)) we have

$$\begin{aligned}
&(2\pi)^d \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h_n^*(\mathbf{s},\mathbf{t})c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})\mathrm{d}\mathbf{s}\mathrm{d}\mathbf{t} \\
&\quad = (2\pi)^{-d}\int_{\mathbb{R}^{2d}} e^{i(\boldsymbol{\omega}'\mathbf{x}-\mathbf{v}'\mathbf{y})}\eta_n^*(\frac{\boldsymbol{\omega}+\mathbf{v}}{2})\theta(\frac{\boldsymbol{\omega}-\mathbf{v}}{2})\xi_1(\boldsymbol{\omega})\xi_1(\mathbf{v})\mathrm{d}\boldsymbol{\omega}\mathrm{d}\mathbf{v},
\end{aligned} \tag{A.12}$$

where $\theta(\mathbf{x}) = 2^{-d}\int_{\mathbb{R}^d} \mathrm{e}^{-i\mathbf{t}^T\mathbf{x}} I\{|\mathbf{t}|_{\max} \le 4m+2T\}\mathrm{d}\mathbf{t}$, $\mathbf{x} \in \mathbb{R}^d$.

We observe that θ is continuous and

$$\int_{\mathbb{R}^d} \theta(\boldsymbol{\omega})^2 \mathrm{d}\boldsymbol{\omega} < \infty. \tag{A.13}$$

Now we define

$$h_n^{**}(\mathbf{s},\mathbf{t}) = \int_{|u|_{\max}\le 2m+2m_a+T} e_n(\mathbf{s}-\mathbf{u})g(\mathbf{u}-\mathbf{t})\mathrm{d}\mathbf{u}, \quad \forall \mathbf{s}, \mathbf{t} \in \mathbb{R}^d.$$

The function $h_n^{**} : \mathbb{R}^{2d} \to \mathbb{C}$ is square-integrable, then

$$\begin{aligned}
&(2\pi)^d \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e_n * g(\mathbf{s}-\mathbf{t})c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})\mathrm{d}\mathbf{s}\mathrm{d}\mathbf{t} \\
&= (2\pi)^d \int_{\mathbb{R}^{2d}} h_n^{**}(\mathbf{s},\mathbf{t})c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})\mathrm{d}\mathbf{s}\mathrm{d}\mathbf{t} \\
&= (2\pi)^{-d}\int_{\mathbb{R}^{2d}} e^{i(\boldsymbol{\omega}'\mathbf{u}-\mathbf{v}'\mathbf{u})}\xi_1(\boldsymbol{\omega})\xi_1(\mathbf{v}) \\
&\quad \times \Big(\int_{|\mathbf{u}|_{\max}\le 2m+2m_a+T} e^{-i(\boldsymbol{\omega}'\mathbf{u}-\mathbf{v}'\mathbf{u})}\hat{e}_n(\boldsymbol{\omega})\eta(\mathbf{v})\mathrm{d}\mathbf{u}\Big)\mathrm{d}\mathbf{v}\mathrm{d}\boldsymbol{\omega}.
\end{aligned} \tag{A.14}$$

It follows, from equations (A.12) and (A.14), that for $\mathbf{x},\mathbf{y} \in D = [0,T]^d$,

$$\begin{aligned}
b(\mathbf{x},\mathbf{y}) &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} \mathrm{e}^{i(\boldsymbol{\omega}'\mathbf{x}-\mathbf{v}'\mathbf{y})}\eta_n^*\Big(\frac{\boldsymbol{\omega}+\mathbf{v}}{2}\Big)\theta\Big(\frac{\boldsymbol{\omega}-\mathbf{v}}{2}\Big)\xi_1(\boldsymbol{\omega})\xi_1(\mathbf{v})\mathrm{d}\boldsymbol{\omega}\mathrm{d}\mathbf{v} \\
&\quad + \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} \mathrm{e}^{i(\boldsymbol{\omega}'\mathbf{x}-\mathbf{v}'\mathbf{y})}\xi_1(\boldsymbol{\omega})\xi_1(\mathbf{v}) \\
&\quad \times \Big(\int_{|\mathbf{u}|_{\max}\le 2m+2m_a+T} e^{-i(\boldsymbol{\omega}'\mathbf{u}-\mathbf{v}'\mathbf{u})}\hat{e}_n(\boldsymbol{\omega})\eta(\mathbf{v})\mathrm{d}\mathbf{u}\Big)\mathrm{d}\mathbf{v}\mathrm{d}\boldsymbol{\omega}.
\end{aligned}$$

Let $\{\psi_1,\dots,\psi_n\}$ be as in (2.15) of Wang (2010). Then using (2.16) and (2.60) in Wang (2010), we have

$$< \psi_k, \psi_k >_{\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2}} - < \psi_k, \psi_k >_{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}} = \lambda_{k,n} - 1 = \varpi^*_{k,n} + \varpi^{**}_{k,n},$$

where

$$\varpi_{k,n}^{*} = \frac{1}{(2\pi)^2}\int_{\mathbb{R}^{2d}} \psi_k(\boldsymbol{\omega})\overline{\psi_k(\mathbf{v})}\eta_n^{*}\Big(\frac{\boldsymbol{\omega}+\mathbf{v}}{2}\Big)\theta\Big(\frac{\boldsymbol{\omega}-\mathbf{v}}{2}\Big)\xi_1(\boldsymbol{\omega})\xi_1(\mathbf{v})\mathrm{d}\boldsymbol{\omega}\mathrm{d}\mathbf{v}, \text{ and}$$

$$\varpi_{k,n}^{**} = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} \psi_k(\boldsymbol{\omega})\overline{\psi_k(\mathbf{v})}\xi_1(\boldsymbol{\omega})\xi_1(\mathbf{v}) \times \Big(\int_{|\mathbf{u}|_{\max}\le 2m+2m_a+T} e^{-i(\boldsymbol{\omega}'\mathbf{u}-\mathbf{v}'\mathbf{u})}\hat{e}_n(\boldsymbol{\omega})\eta(\mathbf{v})\mathrm{d}\mathbf{u}\Big)\mathrm{d}\mathbf{v}\mathrm{d}\boldsymbol{\omega}.$$

Using Bessel's inequality, we have

$$\sum_{k=1}^{n}|\varpi_{k,n}^{*}|^2 \le 2^{-d-1}\pi^{-d}\sup_{\mathbf{s}\in\mathbb{R}^d}\frac{\xi_1(\mathbf{s})^2}{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(\mathbf{s})}\int_{\mathbb{R}^d}|\eta_n^{*}(\boldsymbol{\omega})|^2\mathrm{d}\boldsymbol{\omega}\int_{\mathbb{R}^d}|\theta(\mathbf{v})|^2\mathrm{d}\mathbf{v},$$

and

$$\sum_{k=1}^{n}|\varpi_{k,n}^{**}| \le 2^{-d-1}\pi^{-d}\sup_{\mathbf{s}\in\mathbb{R}^d}\frac{\xi_1(\mathbf{s})^2}{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}(\mathbf{s})}\int_{|\mathbf{u}|_{\max}\le 2m+2m_a+T}\mathrm{d}\mathbf{u} \times \Big(\int_{\mathbb{R}^d}|\hat{e}_n(\boldsymbol{\omega})|^2\mathrm{d}\boldsymbol{\omega}+\int_{\mathbb{R}^d}\eta(\mathbf{v})^2\mathrm{d}\mathbf{v}\Big).$$

From Equations (A.5), (A.6), (A.9), (A.10), (A.11), (A.13), there exists constants C, C_1, C_2 not depending on n such that

$$\sum_{k=1}^{n}|\varpi_{k,n}^{*}|^2 \le C\varepsilon_n^{\iota_0}, \qquad \sum_{k=1}^{n}|\varpi_{k,n}^{*}| \le \sqrt{Cn\varepsilon_n^{\iota_0}} \quad \text{and}$$

$$\sum_{k=1}^{n}|\varpi_{k,n}^{**}| \le \big(C_1/\varepsilon_n^d + C_2\Upsilon\big)$$

with $\Upsilon = \int_{\mathbb{R}^d}\eta(\mathbf{v})^2\mathrm{d}\mathbf{v}$ being finite.

So we conclude that

$$\sum_{k=1}^{n}|\lambda_{k,n}-1| \le \sqrt{Cn\varepsilon_n^{\iota_0}} + \frac{C_1}{\varepsilon_n^d} + C_2\Upsilon. \tag{A.15}$$

We further observe that there exists constants $c^* > 0$ and C^* such that

$$c^* \le \frac{\widehat{\varphi}_{\mu,\kappa,\beta,\sigma^2}}{\widehat{\varphi}_{\mu,\kappa,\beta_0,\sigma_0^2}} \le C^*, \qquad \forall \boldsymbol{\omega}\in\mathbb{R}^d.$$

This implies that $c^* \le \lambda_{k,n} \le C^*$ $\forall k \in \{1,2,\ldots,n\}$.

Finally for any $\vartheta > 0$, using Markov's inequality, (A.15), and using (A.3) we obtain

$$\begin{aligned} P_{\sigma_0^2,\beta_0} & \left(\frac{1}{\sqrt{n}} \left| \frac{1}{\sigma^2} \mathbf{Z}_n' R_\beta^{-1} \mathbf{Z}_n - \frac{1}{\sigma_0^2} \mathbf{Z}_n' R_{\beta_0}^{-1} \mathbf{Z}_n \right| > \vartheta \right) \qquad (A.16) \\ &= P_{\sigma_0^2,\beta_0} \left(\frac{1}{\sqrt{n}} \left| \sum_{k=1}^{n} (\lambda_{k,n}^{-1} - 1) Y_k^2 \right| > \vartheta \right) \\ &\leq P_{\sigma_0^2,\beta_0} \left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} |\lambda_{k,n}^{-1} - 1| Y_k^2 > \vartheta \right) \\ &\leq \frac{1}{\vartheta\sqrt{n}} \sum_{k=1}^{n} |\lambda_{k,n}^{-1} - 1| \\ &\leq \frac{1}{\vartheta\sqrt{n}} \underset{i \in [1,n]}{max} \left\{ \lambda_{i,n}^{-1} \right\} \sum_{k=1}^{n} |\lambda_{k,n} - 1| \\ &\leq \frac{C^{1/2} \varepsilon_n^{\iota_0/2}}{c^* \vartheta} + \frac{1}{c^* \vartheta n^{1/2}} \left(C_1 / \varepsilon_n^d + C_2 \Upsilon \right). \end{aligned}$$

Choose ε_n such that $\varepsilon_n \to 0$ and $n^{1/2} \varepsilon_n^d \to \infty$ as $n \to \infty$. It follows that (A.16) tends to 0 as $n \to \infty$. □

Appendix B

All the estimates and prediction have been carried out using un upcoming version of the R package `CompRandFld` (**?**). Here we present the code for simulating, estimating with maximum likelihood and predicting spatial Gaussian random fields with Matern and Generalized covariance models.

B.1 Code example

B.1.1 Simulation and estimation with ML of GRF with Matern and Generalized Wendland covariance models

```
library(CompRandFld)
library(spam)
# Define the spatial-coordinates of the points:
x <- runif(500)
y <- runif(500)
coords=cbind(x,y)
set.seed(261)
# Simulation of a spatial Gaussian RF with Matern correlation function
smooth_mat=0.5
mean=0
sill=1
scale=0.2/3
nugget=0
param=list(smooth=smooth_mat,mean=mean,sill=sill,
           scale=scale,nugget=nugget)
data1=RFsim(coordx=coords, corrmodel="Matern",
```

```
          param=param)$data

# Estimation of a spatial Gaussian RF with Matern correlation function
fixed=list(mean=mean,smooth=smooth_mat,nugget=nugget)
start=list(sill=sill,scale=scale)
fit_MAT <- FitComposite(data=data1,coordx=coords,corrmodel="Matern",
                    likelihood="Full",type="Standard",
                    start=start,fixed=fixed)

print(fit_MAT)

# Simulation of a spatial Gaussian RF with GW model
c_supp=0.2
smooth_GW=0
power2=4
nugget=0
mean=0
param=list(smooth=smooth_mat,power2=power2,mean=mean,
           sill=sill,scale=c_supp, nugget=nugget)
data2 = RFsim(coordx=coords,  corrmodel="GenWend", param=param)$data
# Estimation of a spatial Gaussian RF with GW model
fixed=list(mean=mean,smooth=smooth_GW,nugget=nugget,power2=power2)
start=list(sill=sill,scale=c_supp)
fit_GW <- FitComposite(data=data2,coordx=coords,corrmodel="GenWend",
                    likelihood="Full",type="Standard",
                    start=start,fixed=fixed)
print(fit_GW)
```

B.1.2 Prediction GRF with Matern, tapered Matern and Generalized Wendland covariance models

```
library(CompRandFld)
library(fields)

################################################################
```

```
############### Examples of Spatial kriging  ###################
################################################################

# Define the spatial-coordinates of the points:
set.seed(79)
x <- runif(200, 0, 1)
y <- runif(200, 0, 1)
coords<-cbind(x,y)
# Set the exponential cov parameters:
corrmodel_1 <- "exponential"
mean<-0
sill<-1
nugget<-0
scale<-0.3/3
param<-list(mean=mean,sill=sill,nugget=nugget,scale=scale)

# Set the wendland parameters (two compatible correlations):
corrmodel_2 <- "Wend0"
mean<-0
sill<-1
nugget<-0
power2=3
c_supp<-0.3
param_wen<-list(mean=mean,sill=sill,nugget=nugget,scale=c_supp,power2=power2)



# Simulation of the spatial Gaussian random field:
data <- RFsim(coordx=coords, corrmodel=corrmodel_1,
              param=param)$data

# locations to predict
xx<-seq(0,1,0.03)
loc_to_pred<-as.matrix(expand.grid(xx,xx))

################################################################
```

```
###
### Example 1. Spatial simple kriging of n sites of a
### Gaussian random fields with exponential correlation.
###
###############################################################
pr<-Kri(loc=loc_to_pred,coordx=coords,corrmodel=corrmodel_1,
       param= param, data=data,mse=TRUE)

################################################################
###
### Example 2. Spatial tapered simple kriging of n sites of a
### Gaussian random fields with exponential correlation.
###
###############################################################
pr_tap=Kri(loc=loc_to_pred,coordx=coords,corrmodel=corrmodel_1,data=data,
       param= param,type="Tapering",maxdist=0.2,taper="Wendland1",mse=TRUE)


################################################################
###
### Example 3. Spatial  simple kriging of n sites of a
### Gaussian random fields using a compatible Wendland model
###
###############################################################
param<-list(mean=mean,sill=sill,nugget=nugget,scale=scale,power2=power2)
pr_wen=Kri(loc=loc_to_pred,coordx=coords,corrmodel=corrmodel_2,data=data,
       param=param_wen,sparse=TRUE,mse=TRUE)

colour <- rainbow(100)
par(mfrow=c(3,2))
zlim=c(-2.6,2.6)
# simple kriging map prediction
image.plot(xx, xx, matrix(pr$pred,ncol=length(xx)),col=colour,
           zlim=zlim,xlab="",ylab="",
           main="Simple Kriging with  exponential model ")
```

```
# simple kriging map prediction variance
image.plot(xx, xx, matrix(pr$mse,ncol=length(xx)),col=colour,
           xlab="",ylab="",main="Std error")

# simple tapered kriging map prediction
image.plot(xx, xx, matrix(pr_tap$pred,ncol=length(xx)),col=colour,
           zlim=zlim,xlab="",ylab="",main="Simple Tapered Kriging")

# simple taperd kriging map prediction variance
image.plot(xx, xx, matrix(pr_tap$mse,ncol=length(xx)),col=colour,
           xlab="",ylab="",main="Std error")

# simple tapered kriging map prediction
image.plot(xx, xx, matrix(pr_wen$pred,ncol=length(xx)),col=colour,
           zlim=zlim,xlab="",ylab="",main="Simple Kriging with Wendland model")

# simple kriging map prediction variance
image.plot(xx, xx, matrix(pr_wen$mse,ncol=length(xx)),col=colour,
           xlab="",ylab="",main="Std error")
```