HAL
open science

On the smallest eigenvalues of covariance matrices of multivariate spatial processes

François Bachoc, Reinhard Furrer

To cite this version:

François Bachoc, Reinhard Furrer. On the smallest eigenvalues of covariance matrices of multivariate spatial processes. Stat, 2016, 5 (1), pp.102-107. 10.1002/sta4.107 . hal-01267398

HAL Id: hal-01267398

https://hal.science/hal-01267398v1

Submitted on 5 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the smallest eigenvalues of covariance matrices of multivariate spatial processes

François Bachoc, Reinhard Furrer
Toulouse Mathematics Institute, University Paul Sabatier, France
Institute of Mathematics and Institute of Computational Science,
University of Zurich, Switzerland

February 5, 2016

Abstract: There has been a growing interest in providing models for multivariate spatial processes. A majority of these models specify a parametric matrix covariance function. Based on observations, the parameters are estimated by maximum likelihood or variants thereof. While the asymptotic properties of maximum likelihood estimators for univariate spatial processes have been analyzed in detail, maximum likelihood estimators for multivariate spatial processes have not received their deserved attention yet. In this article we consider the classical increasing-domain asymptotic setting restricting the minimum distance between the locations. Then, one of the main components to be studied from a theoretical point of view is the asymptotic positive definiteness of the underlying covariance matrix. Based on very weak assumptions on the matrix covariance function we show that the smallest eigenvalue of the covariance matrix is asymptotically bounded away from zero. Several practical implications are discussed as well.

Keywords:increasing-domain asymptotics; matrix covariance function; maximum likelihood; multivariate process; spectral representation; spectrum

1 Introduction

The motivation of this work is the question of what simple conditions on the sampling design and covariance functions have to be imposed for a particular likelihood estimator to be consistent or, say, asymptotically normal, in the setting of multivariate spatial processes. More precisely, let

$$\left\{Z_k(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{D} \subset \mathbb{R}^d, 1 \leq k \leq p\right\} \tag{1}$$

be a multivariate stationary random process, for some fixed $d \in \mathbb{N}^+$ and $p \in \mathbb{N}^+$. Without loss of generality, we assume that the process (1) has zero mean and a matrix covariance

function of the form $\mathbf{C}(\boldsymbol{h}) = \{c_{k\ell}(\boldsymbol{h})\}_{1\leq k,\ell\leq p}$. The diagonal elements c_{kk} are called direct covariance (functions) and the off-diagonal entries $c_{k\ell}$, $k \neq l$, are called cross covariance (functions).

Consider p sets $(\boldsymbol{x}_i^{(1)})_{1\leq i\leq n_1}, \ldots, (\boldsymbol{x}_i^{(p)})_{1\leq i\leq n_p}$ of points in $\mathcal{D}$. In the geostatistical literature, we often denote these points as locations and assume that each $Z_k(\boldsymbol{s})$ is observed at $(\boldsymbol{x}_i^{(k)})_{1\leq i\leq n_k}$. One important problem is to estimate the matrix covariance function from these observations. One common practice is to assume that this function belongs to a given parametric set, and to estimate the corresponding covariance parameters by maximum likelihood approaches. These approaches are now largely implementable in practice, with the availability of massive computing power. In contrast, finite sample and asymptotic properties of these approaches for multivariate processes are still on the research agendas.

Currently, the majority of the asymptotic results are specific to the univariate case, where $d = 1$ and $n_1 = n \to \infty$. While two types of asymptotic settings exist, fixed-domain and increasing-domain [Ste99], we shall focus throughout this article on the increasing-domain setting. In a seminal article, [MM84] give conditions which warrant consistency and asymptotic normality of the maximum likelihood estimator. Additional results and conditions are provided in [Bac14] for the maximum likelihood estimator and in [SR12] for the tapered maximum likelihood estimator. A central quantity in these three references is the covariance matrix of the observation vector. In particular, a necessary condition for the results above is that the smallest eigenvalue of this covariance matrix be bounded away from zero as $n \to \infty$. This seemingly difficult-to-check condition is guaranteed by much simpler and non-restrictive conditions applying only on the covariance function, provided that there exists a fixed minimal distance between any two different observation points, as shown by [Bac14].

The picture is similar in the multivariate setting ($d > 1$), but currently incomplete. The recent references [BFG+15] and [FBD15] provide asymptotic results for maximum likelihood approaches, and require that the smallest eigenvalue of the covariance matrix of the observations be bounded away from zero as $n \to \infty$. However, this condition has currently only been shown to hold for some specific covariance functions and regular grids, see [BFG+15].

In this article, we show that this condition holds in general, provided that the minimal distance between two different observation points of the same process - $\boldsymbol{x}_i^{(k)}$ and $\boldsymbol{x}_j^{(k)}$ for $i \neq j$ - is bounded away from zero as $n_1, ..., n_p \to \infty$, and provided that the matrix covariance function satisfies certain non-restrictive conditions generalizing those of [Bac14]. While the starting argument is similar as in Proposition D.4 of [Bac14] the proof here requires different and more elaborate tools based on complex matrices and Fourier functions. Our approach allows a proof in both cases, collocated and non-collocated observations. We also show that the lower bound we provide can be made uniform over parametric families of matrix covariance functions. We perceive this article as the first step towards a rigorous

analysis of the asymptotic properties of various maximum likelihood type estimators for multivariate spatial processes in the framework of increasing-domain asymptotics.

The next section introduces the notation and states the main result. Due to the interesting tools used in the proof, it is of interest on its own and we have kept it in the main part of the article. Section 3 concludes with some remarks.

2 Notation and main result

For $\boldsymbol{x} \in \mathbb{C}^m$, we let $|\boldsymbol{x}| = \max_{1\leq i\leq m} |x_i|$, where $|z|$ is the modulus of a complex number z. Recall that any Hermitian complex matrix $\mathbf{M}$ of size $m \times m$ has real eigenvalues that we write $\lambda_1(\mathbf{M}) \leq \cdots \leq \lambda_m(\mathbf{M})$. For a complex vector $\boldsymbol{v}$ of size $m \times 1$, we let $\boldsymbol{v}^* = \bar{\boldsymbol{v}}^t$ be the transpose of its conjugate vector and we let $||\boldsymbol{v}||^2 = \boldsymbol{v}^*\boldsymbol{v}$. The next assumption is satisfied for most standard covariance and cross covariance functions.

Assumption 1. *There exists a finite fixed constant $A > 0$ and a fixed constant $\tau > 0$ so that the functions $c_{k\ell}$ satisfy, for all $\boldsymbol{x} \in \mathbb{R}^d$,*

$$|c_{k\ell}(\boldsymbol{x})| \leq \frac{A}{1 + |\boldsymbol{x}|^{d+\tau}}. \tag{2}$$

We define the Fourier transform of a function $g : \mathbb{R}^d \to \mathbb{R}$ by $\hat{g}(\boldsymbol{f}) = (2\pi)^{-d} \int_{\mathbb{R}^d} g(\boldsymbol{x}) e^{-\imath \boldsymbol{f}\cdot\boldsymbol{x}} d\boldsymbol{x}$, where $\imath^2 = -1$. Then, from (2), the covariance functions $c_{k\ell}$ have Fourier transforms $\hat{c}_{k\ell}$ that are continuous and bounded. Also, note that, for any $\boldsymbol{f} \in \mathbb{R}^d$, $\hat{\mathbf{C}}(\boldsymbol{f}) = \{\hat{c}_{k\ell}(\boldsymbol{f})\}_{1\leq k,\ell\leq p}$ is a Hermitian complex matrix, that have real non-negative eigenvalues $0 \leq \lambda_1\{\hat{\mathbf{C}}(\boldsymbol{f})\} \leq \cdots \leq \lambda_p\{\hat{\mathbf{C}}(\boldsymbol{f})\}$. We further assume the following.

Assumption 2. *We have*

$$c_{k\ell}(\boldsymbol{x}) = \int_{\mathbb{R}^d} \hat{c}_{k\ell}(\boldsymbol{f}) e^{\imath \boldsymbol{f}\cdot\boldsymbol{x}} d\boldsymbol{f}. \tag{3}$$

Also, we have $0 < \lambda_1\{\hat{\mathbf{C}}(\boldsymbol{f})\}$ for all $\boldsymbol{f} \in \mathbb{R}^d$.

The condition (3) is very weak and satisfied by most standard covariance and cross covariance functions. On the other hand, the condition $0 < \lambda_1\{\hat{\mathbf{C}}(\boldsymbol{f})\}$ for all $\boldsymbol{f} \in \mathbb{R}^d$ is less innocuous, and is further discussed in Section 3.

Let $d \in \mathbb{N}^+$ and $p \in \mathbb{N}^+$ be fixed. Consider p sequences $(\boldsymbol{x}_i^{(1)})_{i\in\mathbb{N}^+}, \ldots, (\boldsymbol{x}_i^{(p)})_{i\in\mathbb{N}^+}$ of points in $\mathbb{R}^d$, for which we assume the following.

Assumption 3. *There exists a fixed $\Delta > 0$ so that for all k, $\inf_{i,j\in\mathbb{N}^+;i\neq j} |\boldsymbol{x}_i^{(k)} - \boldsymbol{x}_j^{(k)}| \geq \Delta$.*

For all $n_1, \ldots, n_p \in \mathbb{N}^+$, let, for $0 \leq k \leq p$, $N_k = n_1 + \cdots + n_k$, with the convention that $N_0 = 0$. Let also $N = N_p$. Note that $N_1, \ldots, N_p$ depend on $n_1, \ldots, n_p$ but that we do not explicitly write this dependence for concision. Then, let $\boldsymbol{\Sigma}$ (also depending on $n_1, \ldots, n_p$) be the $N \times N$ covariance matrix, filled as follows: For $a = N_{k-1} + i$ and $b = N_{\ell-1} + j$, with $1 \leq k, \ell \leq p$, $1 \leq i \leq n_k$ and $1 \leq j \leq n_l$, $\sigma_{ab} = c_{k\ell}(\boldsymbol{x}_i^{(k)} - \boldsymbol{x}_j^{(\ell)})$.

Our main result is the following.

Theorem 4. *Assume that Assumptions 1, 2 and 3 are satisfied. Then, we have*

$$\inf_{n_1,\ldots,n_p\in\mathbb{N}^+} \lambda_1(\mathbf{\Sigma}) > 0.$$

Proof of Theorem 4. From the proof of Proposition D.4 in [Bac14], there exists a function $\hat{h} : \mathbb{R}^d \to \mathbb{R}^+$ that is C^∞ with compact support $[0,1]^d$ and so that there exists a function $h : \mathbb{R}^d \to \mathbb{R}$ which satisfies $h(0) > 0$,

$$|h(\boldsymbol{x})| \leq \frac{A}{1+|\boldsymbol{x}|^{d+\tau}} \qquad \tau > 0,$$

with the notation of (2), and

$$h(\boldsymbol{x}) = \int_{\mathbb{R}^d} \hat{h}(\boldsymbol{f}) e^{\imath \boldsymbol{f}\cdot\boldsymbol{x}} d\boldsymbol{f}.$$

It can be shown, similarly as in the proof of Lemma D.2 in [Bac14], that there exists a fixed $0 < \delta < \infty$ so that

$$\sup_{1\leq k\leq p, n_k\in\mathbb{N}^+, 1\leq i\leq n_k} \sum_{1\leq j\leq n_k; j\neq i} \left| h\left\{\delta\big(\boldsymbol{x}_i^{(k)} - \boldsymbol{x}_j^{(k)}\big)\right\}\right| \leq \frac{1}{2} h\left(0\right). \tag{4}$$

For all $1 \leq k \leq p$, using Gershgorin circle theorem, the eigenvalues of the $n_k \times n_k$ symmetric matrix $\left[h\{\delta\big(\boldsymbol{x}_i^{(k)} - \boldsymbol{x}_j^{(k)}\big)\}\right]_{1\leq i,j\leq n_k}$ belong to the balls with center $h(0)$ and radius $\sum_{1\leq j\leq n_k, j\neq i} \left|h\{\delta\big(\boldsymbol{x}_i^{(k)} - \boldsymbol{x}_j^{(k)}\big)\}\right|$. Thus, because of (4), these eigenvalues belong to the segment $[h(0)-(1/2)h(0), h(0)+(1/2)h(0)]$ and are larger than $(1/2)h(0)$. Hence, the $N \times N$ matrix $\mathbf{T}$ defined as being block diagonal, with p blocks and with block k equal to $\left[h\{\delta\big(\boldsymbol{x}_i^{(k)} - \boldsymbol{x}_j^{(k)}\big)\}\right]_{1\leq i,j\leq n_k}$, also has eigenvalues larger than $(1/2)h(0)$.

Consider now a real vector $\boldsymbol{v}$ of size $N \times 1$. Then, because $\mathbf{T}$ has eigenvalues larger than $(1/2)h(0)$,

$$\frac{1}{2} h\left(0\right) \sum_{k=1}^{p} \sum_{i=1}^{n_k} |v_{N_{k-1}+i}|^2 \leq \sum_{k=1}^{p} \sum_{i,j=1}^{n_k} v_{N_{k-1}+i} v_{N_{k-1}+j} h\left\{\delta\big(\boldsymbol{x}_i^{(k)} - \boldsymbol{x}_j^{(k)}\big)\right\}. \tag{5}$$

Now, define the $Np \times 1$ real vector $\boldsymbol{w}$ as follows: The N first components are

$$v_1, \ldots, v_{n_1}, \quad \underbrace{0, \ldots, 0,}_{N-n_1 \text{ times}}$$

the components $N+1$ to $2N$ are

$$\underbrace{0, \ldots, 0,}_{n_1 \text{ times}} \; v_{n_1+1}, \ldots, v_{n_1+n_2}, \qquad \underbrace{0, \ldots, 0,}_{N-n_1-n_2 \text{ times}}$$

and so on, until the N last components are

$$\underbrace{0, \ldots, 0,}_{N_{p-1} \text{ times}} \; v_{N_{p-1}+1}, \ldots, v_{N_{p-1}+n_p}.$$

Let also $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\}$ be defined as $\{\boldsymbol{x}_1^{(1)}, \ldots, \boldsymbol{x}_{n_1}^{(1)}, \boldsymbol{x}_1^{(2)}, \ldots, \boldsymbol{x}_{n_2}^{(2)}, \ldots, \boldsymbol{x}_1^{(p)}, \ldots, \boldsymbol{x}_{n_p}^{(p)}\}$. Then we have, from (5),

$$\frac{1}{2}h(0)\sum_{k=1}^{p}\sum_{i=1}^{n_k}|v_{N_{k-1}+i}|^2 \leq \sum_{k=1}^{p}\sum_{i,j=1}^{N} w_{(k-1)N+i}w_{(k-1)N+j}h\left\{\delta(\boldsymbol{s}_i-\boldsymbol{s}_j)\right\}.$$

Let, for $1 \leq i \leq N$, $\boldsymbol{w}_{i\bullet}$ be the $p \times 1$ vector $(w_{(k-1)N+i})_{1\leq k\leq p}$. Then, we can write

$$\begin{aligned}
\frac{1}{2}h(0)\sum_{k=1}^{p}\sum_{i=1}^{n_k}|v_{N_{k-1}+i}|^2 &\leq \sum_{i,j=1}^{N}\sum_{k=1}^{p} w_{(k-1)N+i}w_{(k-1)N+j}\frac{1}{\delta^d}\int_{\mathbb{R}^d}\hat{h}\left(\frac{\boldsymbol{f}}{\delta}\right)e^{\imath\boldsymbol{f}\cdot(\boldsymbol{s}_i-\boldsymbol{s}_j)}d\boldsymbol{f} \\
&= \frac{1}{\delta^d}\int_{\mathbb{R}^d}\hat{h}\left(\frac{\boldsymbol{f}}{\delta}\right)\sum_{i,j=1}^{N}(e^{-\imath\boldsymbol{f}\cdot\boldsymbol{s}_i}\boldsymbol{w}_{i\bullet})^*(e^{-\imath\boldsymbol{f}\cdot\boldsymbol{s}_j}\boldsymbol{w}_{j\bullet})d\boldsymbol{f} \\
&= \frac{1}{\delta^d}\int_{\mathbb{R}^d}\hat{h}\left(\frac{\boldsymbol{f}}{\delta}\right)\left|\left|\sum_{i=1}^{N}e^{-\imath\boldsymbol{f}\cdot\boldsymbol{s}_i}\boldsymbol{w}_{i\bullet}\right|\right|^2 d\boldsymbol{f}. \qquad (6)
\end{aligned}$$

Let E be the set of $\boldsymbol{f}$ so that $\hat{h}(\boldsymbol{f}/\delta)$ is non-zero. Then E is bounded since $\hat{h}$ has a compact support. Hence, because the Hermitian matrix $\hat{\mathbf{C}}(\boldsymbol{f})$ has strictly positive eigenvalues for all $\boldsymbol{f} \in \mathbb{R}^d$ and is continuous, we have $\inf_{\boldsymbol{f}\in E}\lambda_1\{\hat{\mathbf{C}}(\boldsymbol{f})\} > 0$. Also, because $\hat{h}$ is continuous, $\sup_{\boldsymbol{f}\in E}\hat{h}(\boldsymbol{f}/\delta) < +\infty$. Hence, there exists a fixed $\delta_2 > 0$ so that for any $\boldsymbol{z} \in \mathbb{C}^p$, $\boldsymbol{f} \in \mathbb{R}^d$, $\boldsymbol{z}^*\hat{\mathbf{C}}(\boldsymbol{f})\boldsymbol{z} \geq \delta_2\hat{h}(\boldsymbol{f}/\delta)\,||\boldsymbol{z}||^2$. Hence, from (6),

$$\begin{aligned}
\frac{1}{2}h(0)\sum_{k=1}^{p}\sum_{i=1}^{n_k}|v_{N_{k-1}+i}|^2 &\leq \frac{1}{\delta^d\delta_2}\int_{\mathbb{R}^d}\left(\sum_{i=1}^{N}e^{-\imath\boldsymbol{f}\cdot\boldsymbol{s}_i}\boldsymbol{w}_{i\bullet}\right)^*\hat{\mathbf{C}}(\boldsymbol{f})\left(\sum_{i=1}^{N}e^{-\imath\boldsymbol{f}\cdot\boldsymbol{s}_i}\boldsymbol{w}_{i\bullet}\right)d\boldsymbol{f} \\
&= \frac{1}{\delta^d\delta_2}\sum_{i,j=1}^{N}\sum_{k,\ell=1}^{p} w_{(k-1)N+i}w_{(\ell-1)N+j}\int_{\mathbb{R}^d}\hat{c}_{k\ell}(\boldsymbol{f})e^{\imath\boldsymbol{f}\cdot(\boldsymbol{s}_i-\boldsymbol{s}_j)}d\boldsymbol{f} \\
&= \frac{1}{\delta^d\delta_2}\sum_{k,\ell=1}^{p}\sum_{i,j=1}^{N} w_{(k-1)N+i}w_{(\ell-1)N+j}c_{k\ell}(\boldsymbol{s}_i-\boldsymbol{s}_j) \\
&= \frac{1}{\delta^d\delta_2}\boldsymbol{v}^t\boldsymbol{\Sigma}\boldsymbol{v}.
\end{aligned}$$

This concludes the proof. □

We now extend Theorem 4 to the case of a parametric family of covariance and cross covariance functions $\{c_{\boldsymbol{\theta}k\ell}(\boldsymbol{h})\}_{1\leq k,\ell\leq p}$, indexed by a parameter $\boldsymbol{\theta}$ in a compact set Θ of $\mathbb{R}^q$. Let $\hat{\mathbf{C}}_{\boldsymbol{\theta}}(\boldsymbol{f})$ be as $\hat{\mathbf{C}}(\boldsymbol{f})$ with $\{c_{k\ell}(\boldsymbol{h})\}_{1\leq k,\ell\leq p}$ replaced by $\{c_{\boldsymbol{\theta}k\ell}(\boldsymbol{h})\}_{1\leq k,\ell\leq p}$. The proof of the next theorem is identical to that of Theorem 4, up to more cumbersome notations, and is omitted.

Theorem 5. *Assume that, for all $\boldsymbol{\theta} \in \Theta$, the functions $\{c_{\boldsymbol{\theta}k\ell}(\boldsymbol{h})\}_{1\leq k,\ell\leq p}$ satisfy Assumption 1, where A and τ can be chosen independently of $\boldsymbol{\theta}$. Assume that (3) holds with $\{c_{k\ell}(\boldsymbol{h})\}_{1\leq k,\ell\leq p}$ replaced by $\{c_{\boldsymbol{\theta}k\ell}(\boldsymbol{h})\}_{1\leq k,\ell\leq p}$, for all $\boldsymbol{\theta} \in \Theta$. Assume that $\hat{\mathbf{C}}_{\boldsymbol{\theta}}(\boldsymbol{f})$ is jointly*

continuous in $\boldsymbol{f}$ and $\boldsymbol{\theta}$ and that $\lambda_1\{\hat{\mathbf{C}}_{\boldsymbol{\theta}}(\boldsymbol{f})\} > 0$ for all $\boldsymbol{f}$ and $\boldsymbol{\theta}$. Assume finally that Assumption 3 is satisfied.

Then, with $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ being as $\boldsymbol{\Sigma}$ with $\{c_{k\ell}(\boldsymbol{h})\}_{1\leq k,\ell\leq p}$ replaced by $\{c_{\boldsymbol{\theta}k\ell}(\boldsymbol{h})\}_{1\leq k,\ell\leq p}$, we have

$$\inf_{\boldsymbol{\theta}\in\Theta, n_1,\ldots,n_p\in\mathbb{N}^+} \lambda_1(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}) > 0.$$

The lower bound provided by Theorem 5 is typically assumed in the references [SR12], [BFG+15] and [FBD15].

3 Concluding remarks

Note that Theorem 4 is typically applicable for models of stochastic processes with discrete definition spaces, such as time series or Gauss-Markov random fields. Indeed, these models incorporate a fixed minimal distance between any two different observation points.

It is well-known (e.g. Wac03) that covariance and cross covariance functions $\{c_{k\ell}(\boldsymbol{h})\}_{1\leq k,\ell\leq p}$ satisfy $\lambda_1\{\hat{\mathbf{C}}(\boldsymbol{f})\} \geq 0$ for all $\boldsymbol{f}$. It is also known that if $\lambda_1\{\hat{\mathbf{C}}(\boldsymbol{f})\} > 0$ for almost all $\boldsymbol{f} \in \mathbb{R}^d$, then $\lambda_1(\boldsymbol{\Sigma}) > 0$ whenever the points $(\boldsymbol{x}_i^{(k)})_{1\leq i\leq n_k}$ are two-by-two distinct for all k. We believe that this latter assumption on $\lambda_1\{\hat{\mathbf{C}}(\boldsymbol{f})\}$ is nevertheless generally insufficient for Theorem 4 to hold. Indeed, consider for illustration the univariate and unidimensional case with the triangular covariance function, that is $d = 1$, $p = 1$ and $c(h) = c_{11}(h) = (1 - |h|)\mathbf{1}_{\{|h|\leq 1\}}$. Then (see, e.g., Ste99), $\hat{c}(f) = \mathrm{sinc}^2(f/2)/(2\pi)$ has a countable number of zeros. Consider now the sequence of observation points $(x_i^{(1)})_{i\in\mathbb{N}^+} = (x_i)_{i\in\mathbb{N}^+} = (i/2)_{i\in\mathbb{N}^+}$. Then, $\boldsymbol{\Sigma}$ is a tridiagonal Toeplitz matrix with $\sigma_{11} = 1$ and $\sigma_{12} = \sigma_{21} = 1/2$. Hence, see e.g., [NPR13], for any value of $n_1 = n$, the eigenvalues of $\boldsymbol{\Sigma}$ are $[1 + \cos\{(i\pi)/(n+1)\}]_{1\leq i\leq n}$. Thus, although $\lambda_1(\boldsymbol{\Sigma}) = 1 + cos\{(n\pi)/(n+1)\}$ is strictly positive for any n, it goes to 0 as $n \to \infty$. Hence, Assumption 2, stating that $\lambda_1\{\hat{\mathbf{C}}(\boldsymbol{f})\}$ is strictly positive for all $\boldsymbol{f}$, appears to be generally necessary for Theorem 4 to hold.

In order to derive increasing-domain asymptotic properties for covariance tapering (e.g. in SR12 or BFG+15), one typically assumes that the smallest eigenvalues of tapered covariance matrices are lower bounded, uniformly in $n_1, \ldots, n_p$. A tapered covariance matrix is of the form $\boldsymbol{\Sigma} \circ \mathbf{A}$, where $\circ$ is the Schur (component by component) product and where $\mathbf{A}$ is symmetric positive semi-definite with diagonal components 1. Because of the relation $\lambda_1(\boldsymbol{\Sigma} \circ \mathbf{A}) \geq (\min_i A_{ii})\lambda_1(\boldsymbol{\Sigma})$ (see HJ91, Theorem 5.3.4), Theorem 4 directly provides a uniform lower bound for the smallest eigenvalues of tapered covariance matrices.

Acknowledgment: Reinhard Furrer acknowledges support of the UZH Research Priority Program (URPP) on "Global Change and Biodiversity" and the Swiss National Science Foundation SNSF-143282.

References

[Bac14] F. Bachoc. Asymptotic analysis of the role of spatial sampling for covariance parameter estimation of Gaussian processes. *Journal of Multivariate Analysis*, 125(1):1–35, 2014. 2, 4

[BFG+15] M. Bevilacqua, A. Fassò, C. Gaetan, E. Porcu, and D. Velandia. Covariance tapering for multivariate Gaussian random fields estimation. *Statistical Methods & Applications*, pages 1–17, 2015. 2, 6

[FBD15] R. Furrer, F. Bachoc, and J. Du. Asymptotic properties of multivariate tapering for estimation and prediction. *http://arxiv.org/abs/1506.01833*, 2015+. 2, 6

[HJ91] R. Horn and C. Johnson. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1991. 6

[MM84] K.V. Mardia and R.J. Marshall. Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, 71:135–146, 1984. 2

[NPR13] S. Noschese, L. Pasquini, and L. Reichel. Tridiagonal Toeplitz matrices: properties and novel applications. *Numerical Linear Algebra with Applications*, 20(2):302–326, 2013. 6

[SR12] B.A. Shaby and D Ruppert. Tapered covariance: Bayesian estimation and asymptotics. *Journal of Computational and Graphical Statistics*, 21(2):433–452, 2012. 2, 6

[Ste99] M.L Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, New York, 1999. 2, 6

[Wac03] H. Wackernagel. *Multivariate Geostatistics*. Springer, Berlin, 2003. 6