Submitted to Bernoulli
arXiv: 1609.04967

Semiparametric estimation for isotropic max-stable space-time processes

SVEN BUHL[1,*] RICHARD A. DAVIS[2,**] CLAUDIA KLÜPPELBERG[1,†] and CHRISTINA STEINKOHL[1,‡]

[1] *Center for Mathematical Sciences and TUM Institute of Advanced Study, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany.*
E-mail: [*]sven.buhl@tum.de; [†]cklu@tum.de; [‡]christina.steinkohl@gmail.com, *url:* http://www.statistics.ma.tum.de

[2] *Department of Statistics, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, USA. E-mail:* [**]rdavis@stat.columbia.edu, *url:* http://www.stat.columbia.edu/~rdavis

Abstract
Max-stable space-time processes have been developed to study extremal dependence in space-time data. We propose a semiparametric estimation procedure based on a closed form expression of the extremogram to estimate the parameters in a max-stable space-time process. We establish the asymptotic properties of the resulting parameter estimates and propose subsampling procedures to obtain asymptotically correct confidence intervals. A simulation study shows that the proposed procedure works well for moderate sample sizes. Finally, we apply this estimation procedure to fitting a max-stable model to radar rainfall measurements in a region in Florida.

Keywords: Brown-Resnick process, extremogram, max-stable process, semiparametric estimation, space-time process, subsampling, mixing.

1. Introduction

Max-stable processes are a natural extension of the generalized extreme value distributions to infinite dimensions and provide a useful framework for modeling extremal dependence in continuous time or space. In this paper we focus on the max-stable Brown-Resnick process, which was introduced in a time series framework in Brown and Resnick [3], in a spatial setting in Kabluchko et al. [21], and extended to a space-time setting in Davis et al. [8].

In the literature, various max-stable models and estimation procedures have been proposed for extremal data. For the Brown-Resnick process with parametrized dependence structure, inference has been based on composite likelihood methods. In particular, pairwise likelihood estimation has been found useful to estimate parameters in a max-stable process. A description of this method can be found in Padoan et al. [23] for the spatial setting, and Huser and Davison [20] in a space-time setting. Asymptotic results for pairwise likelihood estimates and detailed analyses in the space-time setting for the model analysed in this paper are given in Davis et al. [9]. Unfortunately, parameter estimation using composite likelihood methods can be laborious, since the computation and subsequent optimization of the objective function is time-consuming. Also the choice of good initial values for the optimization of the composite likelihood is essential.

In this paper we introduce a new semiparametric estimation procedure as an alternative to or as a prerequisite for composite likelihood methods. It is based on the extremogram as a natural extremal analog of the correlation function for stationary processes. It was introduced in Davis and Mikosch [7] for time series (also in Fasen et al. [16]), and they show consistency and asymptotic normality of an empirical extremogram estimate under weak mixing conditions. The empirical extremogram and its asymptotic properties in a spatial setting have been investigated in Buhl and Klüppelberg [5] and Cho et al. [6].

Assuming the same dependence structure for the Brown-Resnick space-time process as in [8, 9], we obtain a closed form expression of the extremogram containing the parameters of interest. We first estimate the extremogram nonparametrically by its empirical version, where we separate space and time. Weighted linear regression is then applied in order to produce parameter estimates.

Asymptotic normality of these semiparametric estimates requires asymptotic normality of the extremogram. For the spatial estimate we apply the CLT with mixing conditions as provided in [5], and for the timewise estimate that of [7]. The rate of convergence can be improved by a bias correction term, which we explain in detail for space and time. In a second step we prove then asymptotic normality of the weighted least squares parameter estimates, where constrained optimization has to be applied, since one of the space and one of the time parameters has bounded support. Also the limit laws differ depending whether the parameter lies on the boundary or not. Since the asymptotic covariance matrices in the normal limits are difficult to access, we apply subsampling procedures to obtain pointwise confidence intervals for the parameters, also taking care of the different normal limits.

Our paper is organized as follows. Section 2 defines the isotropic Brown-Resnick process with its choice of dependence function used throughout for modelling extremes observed in space and time. The extremogram is introduced and its parametric form for our model is given. Based on gridded data, the nonparametric extremogram estimation is derived. Asymptotic normality of the parameter estimates is established in Section 3. Section 3.1 is dedicated to the asymptotic normality of the empirical spatial extremogram and its bias correction; Section 3.2 deals with the asymptotic properties of the spatial parameter estimates. Sections 3.3 and 3.4 present the analogues for the time parameters. In Section 4 we explain the subsampling procedure. We test our new semiparametric estimation procedure in a simulation study in Section 5. The paper concludes with an analysis of daily rainfall maxima in a region in Florida in Section 6, where we also compare the semiparametric estimates with the previously obtained pairwise likelihood estimates. Some auxiliary results are summarized in an appendix.

2. Model description and semiparametric estimates

Throughout the paper we consider a strictly stationary Brown-Resnick process in space and time with representation

$$\eta(\boldsymbol{s},t)=\bigvee_{j=1}^{\infty}\left\{\xi_j\, e^{W_j(\boldsymbol{s},t)-\delta(\|\boldsymbol{s}\|,t)}\right\},\quad (\boldsymbol{s},t)\in\mathbb{R}^2\times[0,\infty), \tag{2.1}$$

where $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^2$, $\{\xi_j : j\in\mathbb{N}\}$ are points of a Poisson process on $[0,\infty)$ with intensity $\xi^{-2}d\xi$ and the dependence function δ is *nonnegative and conditionally negative definite*; i.e., for every $m\in\mathbb{N}$ and every $(\boldsymbol{s}^{(1)},t^{(1)}),\dots,(\boldsymbol{s}^{(m)},t^{(m)})\in\mathbb{R}^2\times[0,\infty)$, it holds that

$$\sum_{i=1}^{m}\sum_{j=1}^{m}a_ia_j\delta(\|\boldsymbol{s}^{(i)}-\boldsymbol{s}^{(j)}\|,|t^{(i)}-t^{(j)}|)\le 0$$

for all $a_1,\dots,a_m\in\mathbb{R}$ summing up to 0. The processes $\{W_j(\boldsymbol{s},t):\boldsymbol{s}\in\mathbb{R}^2,t\in[0,\infty)\}$ are independent replicates of a Gaussian process $\{W(\boldsymbol{s},t):\boldsymbol{s}\in\mathbb{R}^2,t\in[0,\infty)\}$ with stationary increments, $W(\mathbf{0},0)=0$, $\mathbb{E}[W(\boldsymbol{s},t)]=0$ and covariance function

$$\begin{aligned}\mathbb{C}\mathrm{ov}[W(\boldsymbol{s}^{(1)},t^{(1)}),&W(\boldsymbol{s}^{(2)},t^{(2)})]\\&=\delta(\|\boldsymbol{s}^{(1)}\|,t^{(1)})+\delta(\|\boldsymbol{s}^{(2)}\|,t^{(2)})-\delta(\|\boldsymbol{s}^{(1)}-\boldsymbol{s}^{(2)}\|,|t^{(1)}-t^{(2)}|).\end{aligned}$$

Representation (2.1) goes back to de Haan [11], Giné et al. [18] and Kabluchko et al. [21]. All finite-dimensional distributions are multivariate extreme value distributions with standard unit Fréchet margins, hence they are in particular multivariate regularly varying. Furthermore, they are perfectly characterised by the dependence function δ, which is termed the *semivariogram* of the process $\{W(\boldsymbol{s},t)\}$ in geostatistics: For $(\boldsymbol{s}^{(1)},t^{(1)}),(\boldsymbol{s}^{(2)},t^{(2)})\in\mathbb{R}^2\times[0,\infty)$, it is given by

$$\mathbb{V}\mathrm{ar}[W(\boldsymbol{s}^{(1)},t^{(1)})-W(\boldsymbol{s}^{(2)},t^{(2)})]=2\delta(\|\boldsymbol{s}^{(1)}-\boldsymbol{s}^{(2)}\|,|t^{(1)}-t^{(2)}|).$$

Since we assume δ to depend only on the norm of $\boldsymbol{s}^{(1)}-\boldsymbol{s}^{(2)}$, the associated process is *(spatially) isotropic*.

In this paper we assume the dependence function δ to be given for $v,u\ge 0$ by

$$\delta(v,u)=2\theta_1v^{\alpha_1}+2\theta_2u^{\alpha_2}, \tag{2.2}$$

where $0<\alpha_1,\alpha_2\le 2$ and $\theta_1,\theta_2>0$. This is the fractional class frequently used for dependence modelling, and here defined with respect to space and time.

The bivariate distribution function of $(\eta(\mathbf{0},0),\eta(\boldsymbol{h},u))$ is given for $x_1,x_2>0$ by

$$F(x_1,x_2)=\exp\Bigg\{-\frac{1}{x_1}\Phi\left(\frac{\log(x_2/x_1)}{\sqrt{2\delta(\|\boldsymbol{h}\|,|u|)}}+\sqrt{\frac{\delta(\|\boldsymbol{h}\|,|u|)}{2}}\right)$$

$$- \frac{1}{x_2} \Phi \left(\frac{\log(x_1/x_2)}{\sqrt{2\delta(\|\boldsymbol{h}\|, |u|)}} + \sqrt{\frac{\delta(\|\boldsymbol{h}\|, |u|)}{2}} \right) \Bigg\}, \tag{2.3}$$

where Φ denotes the standard normal distribution function (cf. Davis et al. [8]).

The parameters of interest are contained in the dependence function δ. We refer to (θ_1, α_1) as the *spatial parameter* and to (θ_2, α_2) as the *temporal parameter*. From the bivariate distribution function in (2.3), the pairwise density can be derived and pairwise likelihood methods can be used to estimate the parameters; cf. Davis et al. [9], Huser and Davison [20] and Padoan et al. [23]. Full likelihood inference is typically hardly tractable in a general multidimensional setting, as the number of terms occurring in the likelihood explode. More recently, however, parametric inference methods based on higher-dimensional margins have been proposed that work in specific scenarios, see for instance Genton et al. [17], who use triplewise instead of pairwise likelihood, Engelke et al. [15], who propose a threshold-based approach, or Thibaud and Opitz [28] and Wadsworth and Tawn [29], who use a censoring scheme for bias reduction.

In the following we introduce an alternative estimation approach, which is based on a closed form expression of the *extremogram*. The latter was introduced for time series by Davis and Mikosch [7] and for spatial and space-time processes by Cho et al. [6] and Steinkohl [27], respectively, and can be regarded as a correlogram for extreme events.

In this paper we consider an isotropic Brown-Resnick process as a regularly varying stochastic processes $\{\eta(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$ where $d = 1$ corresponds to a time series and $d = 2$ to a spatial process, such that $d = 3$ holds for the space-time process.

More precisely, we consider strictly stationary regularly varying processes $\{\eta(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$ for $d \in \mathbb{N}$, where all finite-dimensional distributions are regularly varying (cf. Hult and Lindskog [19] for definitions and results in a general framework and Resnick [25] for details about multivariate regular variation). As a prerequisite, we define for every finite set $\mathcal{I} \subset \mathbb{R}^d$ with cardinality $|\mathcal{I}|$ the vector

$$\eta_{\mathcal{I}} := (\eta(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{I})^\intercal.$$

Throughout, we abbreviate $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$.

Definition* 2.1** (Regularly varying stochastic process).*** A strictly stationary stochastic process $\{\eta(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$ is called *regularly varying*, if there exists some normalizing sequence $0 < a_n \to \infty$ such that $\mathbb{P}(|\eta(\mathbf{0})| > a_n) \sim n^{-d}$ as $n \to \infty$, and if for every finite set $\mathcal{I} \subset \mathbb{R}^d$,

$$n^d \mathbb{P}\Big(\frac{\eta_{\mathcal{I}}}{a_n} \in \cdot\Big) \xrightarrow{v} \mu_{\mathcal{I}}(\cdot), \quad n \to \infty, \tag{2.4}$$

for some non-null Radon measure $\mu_{\mathcal{I}}$ on the Borel sets in $\overline{\mathbb{R}}^{|\mathcal{I}|} \backslash \{\mathbf{0}\}$. In that case,

$$\mu_{\mathcal{I}}(xC) = x^{-\beta} \mu_{\mathcal{I}}(C), \quad x > 0,$$

for every Borel set C in $\overline{\mathbb{R}}^{|\mathcal{I}|} \backslash \{\mathbf{0}\}$. The notation $\xrightarrow{v}$ stands for vague convergence, and $\beta > 0$ is called the *index of regular variation.*

For every $\boldsymbol{s} \in \mathbb{R}^d$ and $\mathcal{I} = \{\boldsymbol{s}\}$ we set $\mu_{\{\boldsymbol{s}\}}(\cdot) = \mu_{\{\boldsymbol{0}\}}(\cdot) =: \mu(\cdot)$, which is justified by stationarity.

Assuming strict stationarity and spatial isotropy of a regularly varying space-time process $\{\eta(\boldsymbol{s},t) : \boldsymbol{s} \in \mathbb{R}^2, t \in [0,\infty)\}$ we can define its extremogram at two points $(\boldsymbol{s}_1, t_1)$ and $(\boldsymbol{s}_2, t_2)$ only in terms of the spatial and temporal lags $v := \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|$ and $u := |t_1 - t_2|$.

***Definition* 2.2** (The extremogram)**.** For a regularly varying strictly stationary isotropic space-time process $\{\eta(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^2 \times [0,\infty)\}$ we define the *space-time extremogram* for two μ-continuous Borel sets A and B in $\overline{\mathbb{R}}\backslash\{0\}$ (i.e. $\mu(\partial A) = \mu(\partial B) = 0$) such that $\mu(A) > 0$ by

$$\rho_{AB}(v,u) = \lim_{n\to\infty} \frac{P\left(\eta(\boldsymbol{s}_1,t_1)/a_n \in A, \eta(\boldsymbol{s}_2,t_2)/a_n \in B\right)}{P\left(\eta(\boldsymbol{s}_1,t_1)/a_n \in A\right)}, \tag{2.5}$$

where $v = \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|$ and $u = |t_1 - t_2|$.

Setting $A = B = (1,\infty)$, we rediscover the *tail dependence coefficient* $\chi(v,u) = \rho_{(1,\infty)(1,\infty)}(v,u)$. For the isotropic Brown-Resnick process there is a closed form expression for $\chi(v,u)$, which is the basis for our estimation procedure.

***Lemma* 2.3** (Davis et al. [8], equation (3.1))**.** Let η be the strictly stationary isotropic Brown-Resnick process in $\mathbb{R}^2 \times [0,\infty)$ as defined in (2.1) with dependence function given in (2.2). Then for $A = B = (1,\infty)$ the extremogram of η is given by

$$\chi(v,u) = 2\Big(1 - \Phi\Big(\sqrt{\frac{1}{2}\delta(v,u)}\Big)\Big) = 2\big(1 - \Phi(\sqrt{\theta_1 v^{\alpha_1} + \theta_2 u^{\alpha_2}})\big), \quad v,u \geq 0. \tag{2.6}$$

Solving equation (2.6) for $\delta(v,u)$ leads to

$$\frac{\delta(v,u)}{2} = \theta_1 v^{\alpha_1} + \theta_2 u^{\alpha_2} = \left(\Phi^{-1}\big(1 - \frac{1}{2}\chi(v,u)\big)\right)^2. \tag{2.7}$$

For temporal lag 0 and taking the logarithm on both sides we have

$$2\log\left(\Phi^{-1}\big(1 - \frac{1}{2}\chi(v,0)\big)\right) = \log(\theta_1) + \alpha_1 \log v.$$

In the same way, we obtain

$$2\log\left(\Phi^{-1}\big(1 - \frac{1}{2}\chi(0,u)\big)\right) = \log(\theta_2) + \alpha_2 \log u.$$

These equations are the basis for parameter estimates. We replace the extremogram on the left hand side in both of these equations by nonparametric estimates at different lags. Then we use constrained weighted least squares estimation in a linear regression framework to obtain parameter estimates.

The estimation procedure is based on the following observation scheme for the space-time data.

***Condition* 2.4.** (1) The locations lie on a regular 2-dimensional grid

$$\mathcal{S}_n = \big\{(i_1, i_2) : i_1, i_2 \in \{1, \dots, n\}\big\} = \big\{\boldsymbol{s}_i : i = 1, \dots, n^2\big\}.$$

(2) The time points are equidistant, given by the set $\{t_1, \dots, t_T\}$.

***Remark* 2.1.** The assumption of a regular grid can be relaxed in various ways. A simple, but notationally more involved extension is the generalisation to rectangular grids, cf. Buhl and Klüppelberg [5], Section 3. Furthermore, it is possible to assume that the observation area consists of random locations given by points of a Poisson process, see for instance Cho et al. [6], Section 2.3, or Steinkohl [27], Section 4.5.2. Also deterministic, but irregularly spaced locations, could be considered as treated in [27] in Section 4.5.1 in the context of pairwise likelihood estimation. In order to make our method transparent we focus on observations on a regular grid. □

The following scheme provides the semiparametric estimation procedure in detail. Denote by $\mathcal{V}$ and $\mathcal{U}$ the sets of spatial and temporal lags, on which the estimation is based. As a general rule, $\mathcal{V}$ and $\mathcal{U}$ should be chosen such that the whole range of extremal dependence is covered. However, beyond that, no lags in space or time should be included for the estimation, since independence effects can introduce a bias in the least squares estimates, similarly as in pairwise likelihood estimation; cf. Buhl and Klüppelberg [4], Section 5.3. One way to determine the range of extremal dependence are permutation tests, which we describe at the end of Section 6.

(1) Nonparametric estimates for the extremogram:
Summarize all pairs of $\mathcal{S}_n$ which give rise to the same spatial lag $v \in \mathcal{V}$ into

$$N(v) = \{(i,j) \in \{1, \dots, n^2\}^2 : \ \|\boldsymbol{s}_i - \boldsymbol{s}_j\| = v\}.$$

For all $t \in \{t_1, \dots, t_T\}$ estimate the *spatial extremogram* by

$$\widehat{\chi}^{(t)}(v, 0) = \frac{\dfrac{1}{|N(v)|} \displaystyle\sum_{i=1}^{n^2} \sum_{\substack{j=1 \\ \|\boldsymbol{s}_i - \boldsymbol{s}_j\| = v}}^{n^2} \mathbb{1}_{\{\eta(\boldsymbol{s}_i, t) > q, \eta(\boldsymbol{s}_j, t) > q\}}}{\dfrac{1}{n^2} \displaystyle\sum_{i=1}^{n^2} \mathbb{1}_{\{\eta(\boldsymbol{s}_i, t) > q\}}}, \quad v \in \mathcal{V}, \tag{2.8}$$

where q is a large quantile (to be specified) of the standard unit Frechét distribution.

For all $\boldsymbol{s} \in \mathcal{S}_n$ estimate the *temporal extremogram* by

$$\widehat{\chi}^{(\boldsymbol{s})}(0, u) = \frac{\frac{1}{T-u} \displaystyle\sum_{k=1}^{T-u} \mathbb{1}_{\{\eta(\boldsymbol{s}, t_k) > q, \eta(\boldsymbol{s}, t_k + u) > q\}}}{\frac{1}{T} \displaystyle\sum_{k=1}^{T} \mathbb{1}_{\{\eta(\boldsymbol{s}, t_k) > q\}}}, \quad u \in \mathcal{U}, \tag{2.9}$$

where again q is a large (possibly different) quantile of the standard unit Frechét distribution
(2) The overall "spatial" and "temporal" extremogram estimates are defined as averages over the temporal and spatial locations, respectively; i.e.,

$$\widehat{\chi}(v,0) = \frac{1}{T}\sum_{k=1}^{T}\widehat{\chi}^{(t_k)}(v,0), \quad v \in \mathcal{V}, \tag{2.10}$$

$$\widehat{\chi}(0,u) = \frac{1}{n^2}\sum_{i=1}^{n^2}\widehat{\chi}^{(\boldsymbol{s}_i)}(0,u), \quad u \in \mathcal{U}. \tag{2.11}$$

Parameter estimates for $\theta_1, \alpha_1, \theta_2$ and α_2 are found by using constrained weighted least squares estimation:

$$\binom{\widehat{\theta}_1}{\widehat{\alpha}_1} = \underset{\substack{\theta_1,\alpha_1>0\\ \alpha_1\in(0,2]}}{\arg\min}\sum_{v\in\mathcal{V}} w_v\Big(2\log\Big(\Phi^{-1}\big(1-\frac{1}{2}\widehat{\chi}(v,0)\big)\Big) - \big(\log(\theta_1)+\alpha_1\log(v)\big)\Big)^2, \tag{2.12}$$

$$\binom{\widehat{\theta}_2}{\widehat{\alpha}_2} = \underset{\substack{\theta_2,\alpha_2>0\\ \alpha_2\in(0,2]}}{\arg\min}\sum_{u\in\mathcal{U}} w_u\Big(2\log\Big(\Phi^{-1}\big(1-\frac{1}{2}\widehat{\chi}(0,u)\big)\Big) - \big(\log(\theta_2)+\alpha_2\log(u)\big)\Big)^2, \tag{2.13}$$

with weights $w_u > 0$ and $w_v > 0$.
We call the estimates $(\widehat{\theta}_1, \widehat{\alpha}_1)$ and $(\widehat{\theta}_2, \widehat{\alpha}_2)$ *weighted least squares estimates* (WLSE).

3. Estimation of the isotropic Brown-Resnick process

In this section we investigate aymptotic properties of the WLSE $(\widehat{\theta}_1, \widehat{\alpha}_1)$ and $(\widehat{\theta}_2, \widehat{\alpha}_2)$.
For a central limit theorem of the extremogram we need a sufficiently precise estimate for the extremogram (2.6), which we give now.

***Lemma* 3.1.** Let $\boldsymbol{s}, \boldsymbol{h} \in \mathbb{R}^2$ and $t \in [0,\infty)$. For every sequence $a_n \to \infty$ we have

$$\begin{aligned}&\frac{\mathbb{P}(\eta(\boldsymbol{s},t) > a_n, \eta(\boldsymbol{s}+\boldsymbol{h},t) > a_n)}{\mathbb{P}(\eta(\boldsymbol{s},t) > a_n)}\\ =&\Big[\chi(\|\boldsymbol{h}\|,0) + \frac{1}{2a_n}\big(\chi(\|\boldsymbol{h}\|,0)-2\big)\big(\chi(\|\boldsymbol{h}\|,0)-1\big)\Big](1+o(1)).\end{aligned}$$

Lemma 3.1 is a direct application of Lemma A.1(b) of Buhl and Klüppelberg [5] for $A = B = (1,\infty)$ and their equation (A.4). This applies since $\{\eta(\boldsymbol{s},t) : \boldsymbol{s} \in \mathbb{R}^2\}$ has finite-dimensional standard unit Fréchet marginal distributions. Also note that $a_n \sim n^2$ as $n \to \infty$ according to Definition 2.1.

Since the WLSE are functions of the spatial and temporal extremograms, we first derive the asymptotic properties of $\widehat{\chi}^{(t)}$ and $\widehat{\chi}^{(\boldsymbol{s})}$ for a fixed time point t and a fixed location $\boldsymbol{s}$, respectively. Sections 3.1 and 3.2 focus on the spatial parameters, whereas Sections 3.3 and 3.4 handle the temporal parameters. We use several results for the extremogram provided in Appendix A and in Buhl and Klüppelberg [5].

3.1. Asymptotics of the empirical spatial extremogram

We prove a central limit theorem for the empirical spatial extremogram of the Brown-Resnick process (2.1) sampled at a finite set of spatial lags, which we summarize in

$$\mathcal{V} = \{v_1, \dots, v_p\}.$$

First we show that the empirical extremogram centred by the pre-asymptotic version is asymptotically normal.

Theorem 3.2. *For a fixed time point $t \in \{t_1, \dots, t_T\}$, consider the spatial Brown-Resnick process $\left\{\eta(\boldsymbol{s}, t) : \boldsymbol{s} \in \mathbb{R}^2\right\}$ as defined in (2.1) with dependence function given in (2.2). Set $m_n = n^{\beta_1}$ for $\beta_1 \in (0, 1/2)$. Then the empirical spatial extremogram $\widehat{\chi}^{(t)}(v, 0)$ defined in (2.8) with the quantile $q = m_n^2$ satisfies*

$$\frac{n}{m_n}\left(\widehat{\chi}^{(t)}(v,0) - \chi_n(v,0)\right)_{v \in \mathcal{V}} \xrightarrow{d} \mathcal{N}(0, \Pi_1^{(\mathrm{iso})}), \quad n \to \infty,$$

where the covariance matrix $\Pi_1^{(\mathrm{iso})}$ is specified in equation (3.6) below, and χ_n is the pre-asymptotic spatial extremogram,

$$\chi_n(v, 0) = \frac{\mathbb{P}(\eta(\mathbf{0}, 0) > m_n^2, \eta(\boldsymbol{h}, 0) > m_n^2)}{\mathbb{P}(\eta(\mathbf{0}, 0) > m_n^2)}, \quad v = \|\boldsymbol{h}\| \in \mathcal{V}\,. \tag{3.1}$$

Proof. As $\eta(\mathbf{0}, t)$ has standard unit Fréchet marginal distributions, we can choose $a_{m_n} = m_n^2$ by Definition 2.1 of regular variation.

We apply Theorem 4.2 of Buhl and Klüppelberg [5] by verifying conditions (M1)-(M4) of that theorem for $\{\eta(\boldsymbol{s}, t) : \boldsymbol{s} \in \mathbb{R}^2\}$, $d = 2$, and $A = B = (1, \infty)$. Condition (M1) is satisfied by equation (A.2).

To show conditions (M2)-(M4) we choose sequences $m_n = n^{\beta_1}$ and $r_n = n^{\beta_2}$ for $0 < \beta_1 < 1/2$ and $0 < \beta_2 < \beta_1$. For this choice m_n and r_n increase to infinity with $m_n = o(n)$ and $r_n = o(m_n)$ as required.

Condition (M2); i.e., $m_n^2 r_n^2 / n = n^{2(\beta_1 + \beta_2) - 1} \to 0$ holds if and only if $\beta_2 \in (0, \min\{\beta_1, (1/2 - \beta_1)\})$.

We now show condition (M3). Choose $\gamma > 0$, such that all lags in $\mathcal{V}$ lie in $B(\mathbf{0}, \gamma) := \{\boldsymbol{s} \in \mathbb{Z}^2 : \|\boldsymbol{s}\| \leq \gamma\}$. Denote by $B(\boldsymbol{h}, \gamma) := \{\boldsymbol{s} \in \mathbb{Z}^2 : \|\boldsymbol{s} - \boldsymbol{h}\| \leq \gamma\} = \boldsymbol{h} + B(\mathbf{0}, \gamma)$ for $\boldsymbol{h} \in \mathbb{R}^2$.

For $\varepsilon > 0$, like in Example 4.6 of Buhl and Klüppelberg [5], we have for $\boldsymbol{s}, \boldsymbol{s}' \in \mathbb{R}^2$ by a Taylor expansion,

$$\begin{aligned}
&\mathbb{P}(\eta(\boldsymbol{s},t) > \varepsilon m_n^2, \eta(\boldsymbol{s}',t) > \varepsilon m_n^2) \\
&\qquad = 1 - 2\mathbb{P}(\eta(\boldsymbol{0},0) \le \varepsilon m_n^2) + \mathbb{P}(\eta(\boldsymbol{s},t) \le \varepsilon m_n^2, \eta(\boldsymbol{s}',t) \le \varepsilon m_n^2) \\
&\qquad = 1 - 2\exp\Big\{ -\frac{1}{x}\Big\} + \exp\Big\{ -\frac{2-\chi(\|\boldsymbol{s}-\boldsymbol{s}'\|,0)}{\varepsilon m_n^2}\Big\} \\
&\qquad = \frac{1}{\varepsilon m_n^2}\chi(\|\boldsymbol{s}-\boldsymbol{s}'\|,0) + \mathcal{O}\Big(\frac{1}{m_n^4}\Big), \quad n \to \infty.
\end{aligned}$$

Therefore, for $\|\boldsymbol{h}\| \ge 2\gamma$,

$$\begin{aligned}
&\mathbb{P}(\max_{\boldsymbol{s}\in B(\boldsymbol{0},\gamma)} \eta(\boldsymbol{s},t) > \varepsilon m_n^2, \max_{\boldsymbol{s}'\in B(\boldsymbol{h},\gamma)} \eta(\boldsymbol{s}',t) > \varepsilon m_n^2) \\
\le& \sum_{\boldsymbol{s}\in B(\boldsymbol{0},\gamma)} \sum_{\boldsymbol{s}'\in B(\boldsymbol{h},\gamma)} \mathbb{P}(\eta(\boldsymbol{s},t) > \varepsilon m_n^2, \eta(\boldsymbol{s}',t) > \varepsilon m_n^2) \\
=& \sum_{\boldsymbol{s}\in B(\boldsymbol{0},\gamma)} \sum_{\boldsymbol{s}'\in B(\boldsymbol{h},\gamma)} \Big\{\frac{1}{\varepsilon m_n^2}\chi(\|\boldsymbol{s}-\boldsymbol{s}'\|,0) + \mathcal{O}\Big(\frac{1}{m_n^4}\Big)\Big\} \\
\le& \frac{2|B(\boldsymbol{0},\gamma)|^2}{\varepsilon m_n^2}\big(1-\Phi(\sqrt{\theta_1(\|\boldsymbol{h}\|-2\gamma)^{\alpha_1}}\big) + \mathcal{O}\Big(\frac{1}{m_n^4}\Big),
\end{aligned} \tag{3.2}$$

as $n \to \infty$, where we have used (2.6). Summarize $V := \{v = \|\boldsymbol{h}\| : \boldsymbol{h} \in \mathbb{Z}^2\}$ and note that $|\{\boldsymbol{h} \in \mathbb{Z}^2 : \|\boldsymbol{h}\| = v\}| = \mathcal{O}(v)$. Therefore, for $k \ge 2\gamma$,

$$\begin{aligned}
L_{m_n} :=& \limsup_{n\to\infty} m_n^2 \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^2 \\ k<\|\boldsymbol{h}\|\le r_n}} \mathbb{P}\Big(\max_{\boldsymbol{s}\in B(\boldsymbol{0},\gamma)} \eta(\boldsymbol{s},t) > \varepsilon m_n^2, \max_{\boldsymbol{s}'\in B(\boldsymbol{h},\gamma)} \eta(\boldsymbol{s}',t) > \varepsilon m_n^2\Big) \\
\le& 2|B(\boldsymbol{0},\gamma)|^2 \limsup_{n\to\infty} \Big\{ \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^2 \\ k<\|\boldsymbol{h}\|\le r_n}} \Big\{\frac{1}{\varepsilon}(1-\Phi(\sqrt{\theta_1(\|\boldsymbol{h}\|-2\gamma)^{\alpha_1}}))\Big\} + \mathcal{O}\Big(\Big(\frac{r_n}{m_n}\Big)^2\Big)\Big\} \\
\le& C_1 \limsup_{n\to\infty} \sum_{\substack{v\in V: \\ k<v\le r_n}} \Big\{\frac{v}{\varepsilon}2\big(1-\Phi(\sqrt{\theta_1(v-2\gamma)^{\alpha_1}})\big)\Big\},
\end{aligned}$$

for some constant $C_1 > 0$. For the term $\mathcal{O}((r_n/m_n)^2)$ we use that $r_n/m_n \to 0$. From Lemma A.3 and the fact that $1-\Phi(x) \le \exp\{-x^2/2\}$ for $x > 0$, we find for $C_2 > 0$,

$$L_{m_n} \le C_2 k^2 \exp\Big\{ -\frac{1}{2}\theta_1(k-2\gamma)^{\alpha_1}\Big\}.$$

Since $\alpha_1 > 0$, the right hand side converges to 0 as $k \to \infty$ ensuring condition (M3).

Now we turn to the mixing conditions (M4).
We start with (M4i). With V as before, and with equation (A.2), we estimate, recalling

from above that the number of lags $\|\boldsymbol{h}\| = v$ is of oder $\mathcal{O}(v)$,

$$m_n^2 \sum_{\boldsymbol{h}\in\mathbb{Z}^2:\|\boldsymbol{h}\|>r_n} \alpha_{1,1}(\|\boldsymbol{h}\|) \le C_1 m_n^2 \sum_{v\in V:v>r_n} v\,\alpha_{1,1}(v) \le 4C_1 m_n^2 \sum_{v\in V:v>r_n} v\,e^{-\theta_1 v^{\alpha_1}/2}.$$

By Lemma A.3 we find

$$m_n^2 \sum_{v\in V:v>r_n} v\,e^{-\theta_1 v^{\alpha_1}/2} \le c m_n^2 r_n^2\ e^{-\theta_1 r_n^{\alpha_1}/2} = c m_n^2 r_n^2\ e^{-\theta_1 n^{\alpha_1\beta_2}/2} \to 0, \quad n\to\infty.$$

By the same arguments condition (M4ii) is satisfied.
Condition (M4iii) holds by equation (A.2), since

$$m_n\, n\,\alpha_{1,n^2}(r_n) \le 4n^3 m_n\, e^{-\theta_1 r_n^{\alpha_1}/2} \to 0, \quad n\to\infty.$$

For the specification of the asymptotic covariance matrix we apply Theorem 4.2 of [5] for the isotropic case, where each spatial lag v_i arises from a set of different vectors $\boldsymbol{h}$, all with same Euclidean norm v_i. For $i\in\{1,\dots,p\}$ such that $v_i\in\mathcal{V}$, we summarize these into

$$L(v_i) := \{\boldsymbol{h}\in\mathbb{Z}^2 : \|\boldsymbol{h}\| = v_i\} = \{\boldsymbol{h}_1^{(i)},\dots,\boldsymbol{h}_{\ell_i}^{(i)}\},$$

where $\ell_i := |L(v_i)|$. Based on the preceding steps of the proof, we conclude from that theorem that

$$\frac{n}{m_n}\big(\widehat{\chi}^{(t)}(\boldsymbol{h}_1^{(i)},0) - \chi_n(\boldsymbol{h}_1^{(i)},0),\dots,\widehat{\chi}^{(t)}(\boldsymbol{h}_{\ell_i}^{(i)},0) - \chi_n(\boldsymbol{h}_{\ell_i}^{(i)},0)\big)^\intercal_{i=1,\dots,p} \xrightarrow{d} \mathcal{N}(\mathbf{0},\Pi_1^{(\mathrm{space})}),$$

where $\Pi_1^{(\mathrm{space})}$ is specified in equation (4.3)-(4.6) of [5]. Furthermore, $\widehat{\chi}^{(t)}(\boldsymbol{h},0)$ is the empirical extremogram for each vector $\boldsymbol{h}$ as specified above.

Define $N(\boldsymbol{h}) := \{(i,j)\in\{1,\dots,n^2\} : \boldsymbol{s}_i - \boldsymbol{s}_j = \boldsymbol{h}\}$, then the numerator in (2.8) normalizes by $|N(\boldsymbol{h})|$ (instead of $|N(v)|$) and the sum runs over $\boldsymbol{s}_i - \boldsymbol{s}_j = \boldsymbol{h}$ vector-wise (instead of equality in norm). Hence, $N(v_i) = \sum_{\boldsymbol{h}\in L(v_i)} N(\boldsymbol{h})$. Isotropy implies for the pre-asymptotic extremogram that $\chi_n(v_i,0) = \chi_n(\boldsymbol{h},0)$ for all $\boldsymbol{h}\in L(v_i)$, such that

$$\chi_n(v_i,0) = \sum_{\boldsymbol{h}\in L(v_i)} \frac{N(\boldsymbol{h})}{N(v_i)}\chi_n(v_i,0) = \sum_{\boldsymbol{h}\in L(v_i)} \frac{N(\boldsymbol{h})}{N(v_i)}\chi_n(\boldsymbol{h},0) \tag{3.3}$$

as well as, by the definition of the estimator in (2.8),

$$\widehat{\chi}^{(t)}(v_i,0) = \sum_{\boldsymbol{h}\in L(v_i)} \frac{N(\boldsymbol{h})}{N(v_i)}\widehat{\chi}^{(t)}(v_i,0) = \sum_{\boldsymbol{h}\in L(v_i)} \frac{N(\boldsymbol{h})}{N(v_i)}\widehat{\chi}^{(t)}(\boldsymbol{h},0). \tag{3.4}$$

We conclude by (3.3) and (3.4) that

$$\widehat{\chi}^{(t)}(v_i,0) - \chi_n(v_i,0) = \sum_{\boldsymbol{h}\in L(v_i)} \frac{N(\boldsymbol{h})}{N(v_i)}\big(\widehat{\chi}^{(t)}(\boldsymbol{h},0) - \chi_n(\boldsymbol{h},0)\big).$$

To obtain a concise representation of the asymptotic normal law for the isotropic extremogram, we define row vectors $(N(\boldsymbol{h})/N(v_i) : \boldsymbol{h} \in L(v_i))$ for $i = 1, \ldots, p$. Set $L := \sum_{i=1}^{p} \ell_i$ and define the $p \times L-$matrix

$$N := \begin{pmatrix} \left(\frac{N(\boldsymbol{h})}{N(v_1)} : \boldsymbol{h} \in L(v_1)\right) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \left(\frac{N(\boldsymbol{h})}{N(v_2)} : \boldsymbol{h} \in L(v_2)\right) & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \left(\frac{N(\boldsymbol{h})}{N(v_p)} : \boldsymbol{h} \in L(v_p)\right) \end{pmatrix}. \tag{3.5}$$

Then we find

$$\begin{aligned} & \frac{n}{m_n}\left(\widehat{\chi}^{(t)}(v_i,0) - \chi_n(v_i,0)\right)^{\intercal}_{i=1,\ldots,p} \\ = \ & \frac{n}{m_n} N \left(\widehat{\chi}^{(t)}(\boldsymbol{h}_1^{(i)},0) - \chi_n(\boldsymbol{h}_1^{(i)},0), \ldots, \widehat{\chi}^{(t)}(\boldsymbol{h}_{\ell_i}^{(i)},0) - \chi_n(\boldsymbol{h}_{\ell_i}^{(i)},0)\right)^{\intercal}_{i=1,\ldots,p} \\ \stackrel{d}{\rightarrow} \ & \mathcal{N}(\mathbf{0}, N\Pi_1^{(\text{space})}N^{\intercal}), \quad n \to \infty, \end{aligned}$$

such that

$$\Pi_1^{(\text{iso})} := N\Pi_1^{(\text{space})}N^{\intercal}. \tag{3.6}$$

□

***Corollary* 3.3.** Under the conditions of Theorem 3.2 the averaged spatial extremogram in (2.10) satisfies

$$\frac{n}{m_n}\left(\frac{1}{T}\sum_{k=1}^{T}\widehat{\chi}^{(t_k)}(v,0) - \chi_n(v,0)\right)_{v\in\mathcal{V}} \stackrel{d}{\rightarrow} \mathcal{N}(0,\Pi_2^{(\text{iso})}), \quad n \to \infty,$$

with covariance matrix $\Pi_2^{(\text{iso})}$ specified in (3.10) below.

Proof. For the first part of the proof, we neglect spatial isotropy. This part is similar to the proof of Theorem 4.2 in Buhl and Klüppelberg [5] and Corollary 3.4 of Davis and Mikosch [7]. We use the notation of the proof of Theorem 3.2. Enumerate the set of spatial lag vectors inherent in the estimation of the extremogram as $\{\boldsymbol{h}_1^{(i)}, \ldots, \boldsymbol{h}_{\ell_i}^{(i)} : i = 1, \ldots, p\}$ and let $\gamma \geq \max\{v_1, \ldots, v_p\}$. Define the vector process

$$\{\boldsymbol{Y}(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^2\} = \{(\eta(\boldsymbol{s}+\boldsymbol{h}, t_k) : \boldsymbol{h} \in B(\mathbf{0},\gamma))^{\intercal}_{k=1,\ldots,T} : \boldsymbol{s} \in \mathbb{R}^2\}.$$

Let $A = B = (1,\infty)$. Consider $i = 1, \ldots, p$, $j = 1, \ldots, \ell_i$, and $k = 1, \ldots, T$. Define sets $D_{j,k}^{(i)}$ by

$$\{\boldsymbol{Y}(\boldsymbol{s}) \in D_{j,k}^{(i)}\} = \{\eta(\boldsymbol{s}, t_k) \in A, \eta(\boldsymbol{s}', t_k) \in B : \boldsymbol{s} - \boldsymbol{s}' = \boldsymbol{h}_j^{(i)}\},$$

and the sets D_k by
$$\{\boldsymbol{Y}(\boldsymbol{s}) \in D_k\} = \{\eta(\boldsymbol{s}, t_k) \in A\}.$$
For $\boldsymbol{h} \in \mathbb{R}^2$ let $B_T(\boldsymbol{h}, \gamma) := B(\boldsymbol{h}, \gamma) \times \{t_1, \dots, t_T\}$. For $\mu_{B_T(\boldsymbol{0},\gamma)}$-continuous Borel sets C and D in $\overline{\mathbb{R}}^{T|B(\boldsymbol{0},\gamma)|}\backslash\{\boldsymbol{0}\}$, regular variation yields the existence of the limit measures

$$\mu_{B_T(\boldsymbol{0},\gamma)}(C) := \lim_{n\to\infty} m_n^2 \mathbb{P}\Big(\frac{\boldsymbol{Y}(\boldsymbol{0})}{m_n^2} \in C\Big)$$
$$\tau_{B_T(\boldsymbol{0},\gamma)\times B_T(\boldsymbol{h},\gamma)}(C \times D) := \lim_{n\to\infty} m_n^2 \mathbb{P}\Big(\frac{\boldsymbol{Y}(\boldsymbol{0})}{m_n^2} \in C, \frac{\boldsymbol{Y}(\boldsymbol{h})}{m_n^2} \in D\Big).$$

By time stationarity we have $\mu_{B_T(\boldsymbol{0},\gamma)}(D_k) = \mu(A)$,

$$\widehat{\chi}^{(t_k)}(\boldsymbol{h}_j^{(i)}, 0) \sim \widehat{R}_{m_n}(D_{j,k}^{(i)}, D_k) := \widehat{\mu}_{B_T(\boldsymbol{0},\gamma),m_n}(D_{j,k}^{(i)})/\widehat{\mu}_{B_T(\boldsymbol{0},\gamma),m_n}(D_k), \quad n \to \infty, \tag{3.7}$$

where the $\widehat{\mu}_{B_T(\boldsymbol{0},\gamma),m_n}(\cdot)$ are empirical estimators of $\mu_{B_T(\boldsymbol{0},\gamma)}(\cdot)$ defined as

$$\widehat{\mu}_{B_T(\boldsymbol{0},\gamma),m_n}(\cdot) := \Big(\frac{m_n}{n}\Big)^2 \sum_{\boldsymbol{s}\in\mathcal{S}_n} \mathbb{1}_{\{\frac{\boldsymbol{Y}(\boldsymbol{s})}{a_n} \in \cdot\}}. \tag{3.8}$$

Likewise we have for the pre-asymptotic quantities

$$\chi_n(\boldsymbol{h}_j^{(i)}, 0) = R_{m_n}(D_{j,k}^{(i)}, D_k) := \frac{\mathbb{P}(\boldsymbol{Y}(\boldsymbol{0})/m_n^2 \in D_{j,k}^{(i)})}{\mathbb{P}(\boldsymbol{Y}(\boldsymbol{0})/m_n^2 \in D_k)} =: \frac{\mu_{B_T(\boldsymbol{0},\gamma),m_n}(D_{j,k}^{(i)})}{\mu_{B_T(\boldsymbol{0},\gamma),m_n}(D_k)}, \tag{3.9}$$

which are independent of time t_k by stationarity. For notational ease we abbreviate in the following

$$\mu_{B_T(\boldsymbol{0},\gamma)}(\cdot) = \mu_\gamma(\cdot), \quad \mu_{B_T(\boldsymbol{0},\gamma),m_n}(\cdot) = \mu_{\gamma,m_n}(\cdot), \quad \text{and} \quad \widehat{\mu}_{B_T(\boldsymbol{0},\gamma),m_n}(\cdot) = \widehat{\mu}_{\gamma,m_n}(\cdot)$$

For each $k \in \{1, \dots, T\}$ we now define the matrices

$$F^{(k)} = [F_1, F_2^{(k)}]$$

with $F_1 \in \mathbb{R}^{L\times L}$ and $F_2^{(k)} \in \mathbb{R}^L$ given by

$$F_1 = \operatorname{diag}(\mu(A)) \quad \text{and} \quad F_2^{(k)} := (-\mu_\gamma(D_{1,k}^{(1)}), \dots, -\mu_\gamma(D_{\ell_1,k}^{(1)}), \dots, -\mu_\gamma(D_{\ell_p,k}^{(p)}))^\top.$$

Although $F_2^{(k)}$ is constant over $k \in \{1, \dots, T\}$ by time stationarity, we keep the index to clarify the notation. Define the $TL \times T(L+1)$-matrix $\boldsymbol{F}$ and the column vector $\widehat{\boldsymbol{\chi}} - \boldsymbol{\chi}_n$

with TL components as

$$\boldsymbol{F} := \begin{pmatrix} F^{(1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & F^{(2)} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & F^{(T)} \end{pmatrix} \quad \text{and} \quad \widehat{\boldsymbol{\chi}} - \boldsymbol{\chi}_n := \begin{pmatrix} \widehat{\chi}^{(t_1)}(h_1^{(1)},0) - \chi_n(h_1^{(1)},0) \\ \vdots \\ \widehat{\chi}^{(t_1)}(h_{\ell_1}^{(1)},0) - \chi_n(h_{\ell_1}^{(1)},0) \\ \vdots \\ \widehat{\chi}^{(t_1)}(h_{\ell_p}^{(p)},0) - \chi_n(h_{\ell_p}^{(p)},0) \\ \vdots \\ \widehat{\chi}^{(t_T)}(h_{\ell_p}^{(p)},0) - \chi_n(h_{\ell_p}^{(p)},0) \end{pmatrix}.$$

Define the vector $(\widehat{\boldsymbol{R}}_{m_n} - \boldsymbol{R}_{m_n})$ with the quantities from (3.7) and the corresponding pre-asymptotic quantities from (3.9) exactly in the same way. Furthermore, define for $k = 1,\dots,T$ the vectors in $\mathbb{R}^{L+1}$

$$\boldsymbol{\mu}_{\gamma,m_n}^{(k)} = \\ \big(\mu_{\gamma,m_n}(D_{1,k}^{(1)}),\dots,\mu_{\gamma,m_n}(D_{\ell_1,k}^{(1)}),\dots\dots,\mu_{\gamma,m_n}(D_{1,k}^{(p)}),\dots,\mu_{\gamma,m_n}(D_{\ell_p,k}^{(p)}),\mu_{\gamma,m_n}(D_k)\big)^\intercal,$$

which we stack one on top of the other giving a vector $\boldsymbol{\mu}_{\gamma,m_n}$ in $\mathbb{R}^{T(L+1)}$, and $\widehat{\boldsymbol{\mu}}_{\gamma,m_n}$ analogously. Then we obtain

$$\widehat{\boldsymbol{\chi}} - \boldsymbol{\chi}_n = (1+o(1))(\widehat{\boldsymbol{R}}_{m_n} - \boldsymbol{R}_{m_n}) = \frac{1+o_p(1)}{\mu(A)^2}\boldsymbol{F}\,(\widehat{\boldsymbol{\mu}}_{\gamma,m_n} - \boldsymbol{\mu}_{\gamma,m_n}),$$

where the last step follows as in the proof of Theorem 4.2 of [5] and involves Slutzky's theorem. Using ideas of the proof of their Lemma 5.1, we observe that as $n \to \infty$,

$$\mathbb{C}\mathrm{ov}\Big[\widehat{\mu}_{B_T(\mathbf{0},\gamma),m_n}(C),\widehat{\mu}_{B_T(\mathbf{0},\gamma),m_n}(D)\Big] \\ \sim \Big(\frac{m_n}{n}\Big)^2\Big(\mu_{B_T(\mathbf{0},\gamma)}(C\cap D) + \sum_{\mathbf{0}\neq \boldsymbol{h}\in\mathbb{Z}^2}\tau_{B_T(\mathbf{0},\gamma)\times B_T(\boldsymbol{h},\gamma)}(C\times D)\Big) =: \Big(\frac{m_n}{n}\Big)^2 c_{C,D}.$$

With $\Sigma \in \mathbb{R}^{T(L+1)\times T(L+1)}$ defined as

$$\Sigma = \begin{pmatrix} c_{D_{1,1}^{(1)},D_{1,1}^{(1)}} & \cdots & c_{D_{1,1}^{(1)},D_1} & \cdots & c_{D_{1,1}^{(1)},D_{1,T}^{(p)}} & \cdots & c_{D_{1,1}^{(1)},D_T} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{D_T,D_{1,1}^{(1)}} & \cdots & c_{D_T,D_1} & \cdots & c_{D_T,D_{1,T}^{(p)}} & \cdots & c_{D_T,D_T} \end{pmatrix},$$

we thus conclude that

$$\frac{n}{m_n}\begin{pmatrix} \widehat{\chi}^{(t_1)}(\boldsymbol{h}_1^{(1)},0) - \chi_n(\boldsymbol{h}_1^{(1)},0) \\ \vdots \\ \widehat{\chi}^{(t_T)}(\boldsymbol{h}_{\ell_p}^{(p)},0) - \chi_n(\boldsymbol{h}_{\ell_p}^{(p)},0) \end{pmatrix} \xrightarrow{d} \mathcal{N}(\mathbf{0},\mu(A)^{-4}\boldsymbol{F}\Sigma(\boldsymbol{F})^\top).$$

To obtain the asymptotic covariance matrix in the spatially isotropic case, we proceed as in the proof of Theorem 3.2. We define the $Tp \times TL$-matrix

$$\boldsymbol{N} := \begin{pmatrix} N & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & N & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & N \end{pmatrix}$$

with N given in equation (3.5). Then we have

$$\begin{aligned} \frac{n}{m_n} \begin{pmatrix} \widehat{\chi}^{(t_1)}(v_1,0) - \chi_n(v_1,0) \\ \vdots \\ \widehat{\chi}^{(t_T)}(v_p,0) - \chi_n(v_p,0) \end{pmatrix} &= \frac{n}{m_n}\boldsymbol{N} \begin{pmatrix} \widehat{\chi}^{(t_1)}(\boldsymbol{h}_1^{(1)},0) - \chi_n(\boldsymbol{h}_1^{(1)},0) \\ \vdots \\ \widehat{\chi}^{(t_T)}(\boldsymbol{h}_{\ell_p}^{(p)},0) - \chi_n(\boldsymbol{h}_{\ell_p}^{(p)},0) \end{pmatrix} \\ \stackrel{d}{\to} \quad & \mathcal{N}(\mathbf{0}, \mu(A)^{-4}\boldsymbol{N}\boldsymbol{F}\Sigma(\boldsymbol{N}\boldsymbol{F})^\top), \quad n \to \infty, \end{aligned}$$

and we conclude that for the averaged spatial extremogram the statement holds with

$$\begin{aligned} \Pi_2^{(\mathrm{iso})} =& \mu(A)^{-4}T^{-2} \begin{pmatrix} 1 & 0\cdots 0 & 1 & 0\cdots 0 & \cdots & 1 & 0\cdots 0 \\ 0 & 1\cdots 0 & 0 & 1\cdots 0 & \cdots & 0 & 1\cdots 0 \\ & & & \ddots & & & \\ 0 & 0\cdots 1 & 0 & 0\cdots 1 & \cdots & 0 & 0\cdots 1 \end{pmatrix} \boldsymbol{N}\boldsymbol{F}\Sigma(\boldsymbol{N}\boldsymbol{F})^\top \\ & \begin{pmatrix} 1 & 0\cdots 0 & 1 & 0\cdots 0 & \cdots & 1 & 0\cdots 0 \\ 0 & 1\cdots 0 & 0 & 1\cdots 0 & \cdots & 0 & 1\cdots 0 \\ & & & \ddots & & & \\ 0 & 0\cdots 1 & 0 & 0\cdots 1 & \cdots & 0 & 0\cdots 1 \end{pmatrix}^\top . \end{aligned} \tag{3.10}$$

□

***Remark* 3.1.** In the central limit theorem the pre-asymptotic extremogram (3.1) can be replaced by the theoretical one (2.6), provided that

$$\frac{n}{m_n}\left(\chi_n(v,0) - \chi(v,0)\right) \to 0, \quad n \to \infty, \tag{3.11}$$

is satisfied for all spatial lags $v \in \mathcal{V}$. For the Brown-Resnick process (2.1) we obtain from Lemma 3.1,

$$\begin{aligned} & \frac{n}{m_n}\left(\chi_n(v,0) - \chi(v,0)\right) \\ &= \frac{n}{m_n}\left(\frac{P(\eta(\boldsymbol{s},t) > m_n^2, \eta(\boldsymbol{s}+\boldsymbol{h},t) > m_n^2)}{\mathbb{P}(\eta(\boldsymbol{s},t) > m_n^2)} - \chi(v,0)\right) \\ &\sim \frac{n}{2m_n^3}\big(\chi(v,0) - 2\big)\big(\chi(v,0) - 1\big) \end{aligned}$$

$$= n^{1-3\beta_1}\frac{1}{2}\big(\chi(v,0)-2\big)\big(\chi(v,0)-1\big) \quad \to \quad 0 \quad \text{if and only if } \beta_1 > 1/3;$$

cf. Theorem 4.4 of Buhl and Klüppelberg [5]. Thus we have to distinguish two cases:

(I) For $\beta_1 \leq 1/3$ we cannot replace the pre-asymptotic extremogram by the theoretical version, but can resort to a bias correction, which is decribed in (3.14) below.

(II) For $1/3 < \beta_1 < 1/2$ we obtain indeed

$$n^{1-\beta_1}\big(\widehat{\chi}^{(t)}(v,0)-\chi(v,0)\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(0,\Pi_1^{(\mathrm{iso})}), \quad n\to\infty. \tag{3.12}$$

□

We now turn to the bias correction needed in case (I). By Lemma 3.1 the pre-asymptotic extremogram has representation

$$\begin{aligned}\chi_n(v,0) &= \Big[\chi(v,0)+\frac{1}{2m_n^2}\big(\chi(v,0)-2\big)\big(\chi(v,0)-1\big)\Big](1+o(1))\\ &= \Big[\chi(v,0)+\frac{1}{2m_n^2}\nu(v,0)\Big](1+o(1)), \quad n\to\infty,\end{aligned} \tag{3.13}$$

where $\nu(v,0) := \big(\chi(v,0)-2\big)\big(\chi(v,0)-1\big)$.

Consequently, we propose for fixed $t\in\{t_1,\dots,t_T\}$ and all $v\in\mathcal{V}$ the *bias corrected empirical spatial extremogram*

$$\widehat{\chi}^{(t)}(v,0)-\frac{1}{2m_n^2}\big(\widehat{\chi}^{(t)}(v,0)-2\big)\big(\widehat{\chi}^{(t)}(v,0)-1\big) =: \widehat{\chi}^{(t)}(v,0)-\frac{1}{2m_n^2}\widehat{\nu}^{(t)}(v,0),$$

and set

$$\widetilde{\chi}^{(t)}(v,0) := \begin{cases}\widehat{\chi}^{(t)}(v,0)-\dfrac{1}{2m_n^2}\widehat{\nu}^{(t)}(v,0) & \text{if } m_n = n^{\beta_1} \text{ with } \beta_1\in(\frac{1}{5},\frac{1}{3}],\\ \widehat{\chi}^{(t)}(v,0) & \text{if } m_n = n^{\beta_1} \text{ with } \beta_1\in(\frac{1}{3},\frac{1}{2}).\end{cases} \tag{3.14}$$

Theorem 3.4 below shows asymptotic normality of the bias corrected extremogram centred by the true one and, in particular, why β_1 has to be larger than $1/5$.

Theorem 3.4. *For a fixed time point $t\in\{t_1,\dots,t_T\}$ consider the spatial Brown-Resnick process $\{\eta(\boldsymbol{s},t),\boldsymbol{s}\in\mathbb{R}^2\}$ defined in (2.1) with dependence function given in (2.2). Set $m_n = n^{\beta_1}$ for $\beta_1\in\big(\frac{1}{5},\frac{1}{3}\big]$. Then the bias corrected empirical spatial extremogram (3.14) satisfies*

$$\frac{n}{m_n}\big(\widetilde{\chi}^{(t)}(v,0)-\chi(v,0)\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(0,\Pi_1^{(\mathrm{iso})}), \quad n\to\infty, \tag{3.15}$$

where $\Pi_1^{(\mathrm{iso})}$ is the covariance matrix as given in equation (3.6). Furthermore, the corresponding bias corrected version of (2.10) satisfies

$$\frac{n}{m_n}\Big(\frac{1}{T}\sum_{k=1}^{T}\widetilde{\chi}^{(t_k)}(v,0)-\chi(v,0)\Big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(0,\Pi_2^{(\mathrm{iso})}), \quad n\to\infty,$$

with covariance matrix $\Pi_2^{(\mathrm{iso})}$ *specified in* (3.10).

Proof. For simplicity we suppress the time point t in the notation. By (3.13) and (3.14) we have as $n \to \infty$,

$$\frac{n}{m_n}(\widetilde{\chi}(v,0) - \chi(v,0)) \sim \frac{n}{m_n}(\widehat{\chi}(v,0) - \chi_n(v,0)) - \frac{n}{2m_n^3}(\widehat{\nu}(v,0) - \nu(v,0)).$$

By Theorem 3.2 it suffices to show that $(n/(2m_n^3))(\widehat{\nu}(v,0) - \nu(v,0)) \stackrel{P}{\to} 0$. Setting $\nu_n(v,0) := \big(\chi_n(v,0) - 2\big)\big(\chi_n(v,0) - 1\big)$ we have

$$\frac{n}{2m_n^3}(\widehat{\nu}(v,0) - \nu(v,0)) = \frac{n}{2m_n^3}(\widehat{\nu}(v,0) - \nu_n(v,0)) + \frac{n}{2m_n^3}(\nu_n(v,0) - \nu(v,0)) =: A_1 + A_2.$$

We calculate

$$\begin{aligned}
&\frac{n}{m_n(2\chi(v,0)-3)}\Big(\widehat{\nu}(v,0) - \nu_n(v,0)\Big)\\
&= \frac{n}{m_n(2\chi(v,0)-3)}\Big(\widehat{\chi}^2(v,0) - 3\widehat{\chi}(v,0) - (\chi_n^2(v,0) - 3\chi_n(v,0))\Big)\\
&= \frac{n}{m_n(2\chi(v,0)-3)}\Big((\widehat{\chi}(v,0) - \chi_n(v,0))(\widehat{\chi}(v,0) + \chi_n(v,0)) - 3(\widehat{\chi}(v,0) - \chi_n(v,0))\Big)\\
&= \frac{n}{m_n}\Big(\widehat{\chi}(v,0) - \chi_n(v,0)\Big)\frac{\widehat{\chi}(v,0) + \chi_n(v,0) - 3}{2\chi(v,0) - 3}.
\end{aligned}$$

The first term converges by Theorem 3.2 weakly to a normal distribution, and the second term, together with the fact that $\widehat{\chi}(v,0) \stackrel{P}{\to} \chi(v,0)$ and $\chi_n(v,0) \stackrel{P}{\to} \chi(v,0)$, converges to 1 in probability. Hence, it follows from Slutzky's theorem that $A_1 \stackrel{P}{\to} 0$. Now we turn to A_2 and calculate

$$\begin{aligned}
\nu_n(v,0) &= \chi_n^2(v,0) - 3\chi_n(v,0) + 2\\
&\sim \Big(\chi(v,0) + \frac{1}{2m_n^2}\nu(v,0)\Big)^2 - 3\Big(\chi(v,0) + \frac{1}{2m_n^2}\nu(v,0)\Big) + 2\\
&= \chi^2(v,0) - 3\chi(v,0) + 2 + \frac{1}{m_n^2}\chi(v,0)\nu(v,0) + \frac{1}{4m_n^4}\nu(v,0)^2 - \frac{3}{2m_n^2}\nu(v,0)\\
&= \big(\chi(v,0) - 2\big)\big(\chi(v,0) - 1\big) + \frac{1}{m_n^2}\chi(v,0)\nu(v,0) + \frac{1}{4m_n^4}\nu(v,0)^2 - \frac{3}{2m_n^2}\nu(v,0)\\
&= \nu(v,0) + \frac{\nu(v,0)}{m_n^2}\Big(\chi(v,0) + \frac{1}{4m_n^2}\nu(v,0) - \frac{3}{2}\Big),
\end{aligned}$$

where we have used (3.13). Therefore, A_2 converges to 0, if $n/m_n^5 \to 0$ as $n \to \infty$. With $m_n = n^{\beta_1}$ it follows that $\beta_1 > \frac{1}{5}$. Finally, the last statement follows from Corollary 3.3. □

Remark 3.2. Note that in (3.12) the rate of convergence is of the order n^a for $a \in (1/2, 2/3)$. On the other hand, after bias correction in (3.15) we obtain convergence of the order n^a for $a \in [2/3, 4/5)$; i.e. a better rate. □

***Example* 3.5.** We generate 100 realisations of the Brown-Resnick process in (2.1) using the `R`-package `RandomFields` [26] and the exact method via extremal functions proposed in Dombry et al. [13], Section 2. We then compare the empirical estimates of the spatial extremogram $\widehat{\chi}(v,0)$ in (2.8) and the bias corrected ones $\widetilde{\chi}(v,0)$ in (3.14) with the true theoretical extremogram $\chi(v,0)$ for lags $v \in \{1, \sqrt{2}, 2, \sqrt{5}, \sqrt{8}, 3, \sqrt{10}, \sqrt{13}, 4, \sqrt{17}\}$. We choose the parameters $\theta_1 = 0.4$ and $\alpha_1 = 1.5$. The grid size and the number of time points are given by $n = 70$ and $T = 10$. The results are summarised in Figure 3.1. We see that the bias corrected extremogram is closer to the true one. $\square$

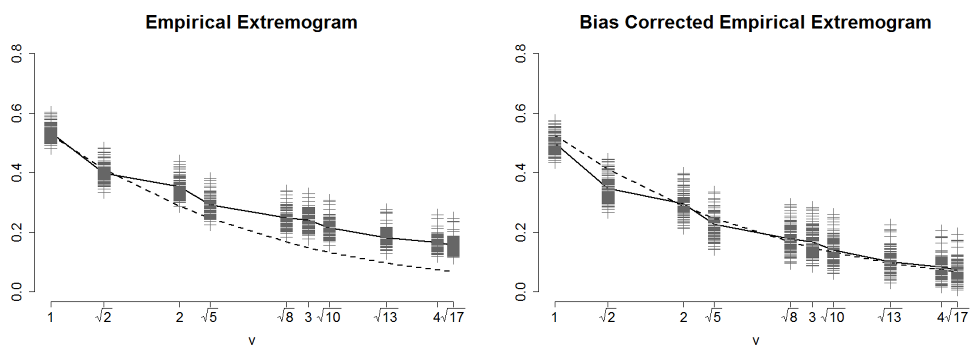

Figure 3.1: Empirical spatial extremogram (left) and its bias corrected version (right) for one hundred simulated max-stable random fields in (2.1) with $\delta(v,0) = 2 \cdot 0.4v^{1.5}$. The dotted line represents the theoretical spatial extremogram and the solid line is the mean over all estimates.

3.2. Asymptotic properties of spatial parameter estimates

In this section we prove asymptotic normality of the WLSE $(\widehat{\theta}_1, \widehat{\alpha}_1)$ of Section 2. We use the following notation:

$$y_v := 2\log\Big(\Phi^{-1}\big(1 - \frac{1}{2}\widetilde{\chi}(v,0)\big)\Big) \quad \text{and} \quad x_v := \log(v), \quad v \in \mathcal{V},$$

with $\widetilde{\chi}(v,0) = \frac{1}{T}\sum_{k=1}^{T} \widetilde{\chi}^{(t_k)}(v,0)$ as in (2.10), possibly after a bias correction, which depends on the two cases described in Remark 3.1. Then (2.12) reads as

$$\begin{pmatrix} \widehat{\theta}_1 \\ \widehat{\alpha}_1 \end{pmatrix} = \underset{\substack{\theta_1, \alpha_1 > 0 \\ \alpha_1 \in (0,2]}}{\arg\min} \sum_{v \in \mathcal{V}} w_v \big(y_v - \big(\log(\theta_1) + \alpha_1 x_v\big)\big)^2 \tag{3.16}$$

and we are in the setting of weighted linear regression. To show asymptotic normality of the WLSE as in (3.16), we define the design matrix X and weight matrix W as

$$X = [\mathbf{1}, (x_v)_{v\in\mathcal{V}}^\intercal] \in \mathbb{R}^{p\times 2} \quad \text{and} \quad W = \operatorname{diag}\{w_v : v \in \mathcal{V}\} \in \mathbb{R}^{p\times p},$$

respectively, where $\mathbf{1} = (1,\dots,1)^\intercal \in \mathbb{R}^p$. Let $\boldsymbol{\psi}_1 = (\log(\theta_1), \alpha_1)^\intercal$ be the parameter vector with parameter space $\Psi = \mathbb{R}\times(0,2]$. Then the WLSE; i.e., the solution to (3.16) is given by

$$\widehat{\boldsymbol{\psi}}_1 := \begin{pmatrix}\log(\widehat{\theta}_1)\\ \widehat{\alpha}_1\end{pmatrix} = (X^\intercal W X)^{-1} X^\intercal W (y_v)_{v\in\mathcal{V}}^\intercal.$$

Without any constraints $\widehat{\boldsymbol{\psi}}_1$ may produce estimates of α_1 outside its parameter space $(0,2]$. In such cases we set the parameter estimate equal to 2, and we denote the resulting estimate by $\widehat{\boldsymbol{\psi}}_1^c = (\log(\widehat{\theta}_1)^c, \widehat{\alpha}_1^c)^\intercal$.

Theorem 3.6. *Let $\widehat{\boldsymbol{\psi}}_1^c = (\log(\widehat{\theta}_1^c), \widehat{\alpha}_1^c)^\intercal$ denote the WLSE resulting from the constrained minimization problem (3.16) and $\boldsymbol{\psi}_1^* = (\log(\theta_1^*), \alpha_1^*)^\intercal \in \Psi$ the true parameter vector. Set $m_n = n^{\beta_1}$ for $\beta_1 \in (1/5, 1/2)$. Then as $n\to\infty$,*

$$\frac{n}{m_n}\left(\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^*\right) \xrightarrow{d} \begin{cases}\boldsymbol{Z}_1 & \text{if } \alpha_1^* < 2,\\ \boldsymbol{Z}_2 & \text{if } \alpha_1^* = 2,\end{cases} \tag{3.17}$$

where $\boldsymbol{Z}_1 \sim \mathcal{N}(0, \Pi_3^{(\mathrm{iso})})$, and the distribution of $\boldsymbol{Z}_2$ is given by

$$\begin{aligned}\mathbb{P}(\boldsymbol{Z}_2 \in B) = &\int_{B\cap\{(b_1,b_2)\in\mathbb{R}^2: b_2<0\}} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(z_1,z_2)\,dz_1dz_2\\ &+ \int_0^\infty \int_{\{b_1\in\mathbb{R}:(b_1,0)\in B\}} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}\Big(z_1 - \frac{1}{\sum_{v\in\mathcal{V}} w_v}\sum_{v\in\mathcal{V}}(w_v x_v)\ z_2,\ z_2\Big)dz_1dz_2\end{aligned} \tag{3.18}$$

for every Borel set B in $\mathbb{R}^2$, and $\varphi_{\mathbf{0},\Sigma}$ denotes the bivariate normal density with mean vector $\mathbf{0}$ and covariance matrix Σ. In particular, the joint distribution function of $\boldsymbol{Z}_2$ is given for $(p_1,p_2)^\intercal \in \mathbb{R}^2$ by

$$\begin{aligned}\mathbb{P}\left(\boldsymbol{Z}_2 \le (p_1,p_2)^\intercal\right) = &\int_{-\infty}^{\min\{0,p_2\}}\int_{-\infty}^{p_1} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(z_1,z_2)\,dz_1dz_2\\ &+ \mathbb{1}_{\{p_2\ge 0\}}\int_0^\infty\int_{-\infty}^{p_1} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}\Big(z_1 - \frac{1}{\sum_{v\in\mathcal{V}} w_v}\sum_{v\in\mathcal{V}}(w_v x_v)\ z_2,\ z_2\Big)dz_1dz_2.\end{aligned} \tag{3.19}$$

The covariance matrix of $\boldsymbol{Z}_1$ has representation

$$\Pi_3^{(\mathrm{iso})} = Q_x^{(w)} G \Pi_2^{(\mathrm{iso})} G Q_x^{(w)\intercal}, \tag{3.20}$$

where $\Pi_2^{(\mathrm{iso})}$ *is the covariance matrix given in* (3.10),

$$Q_x^{(w)} = (X^\intercal W X)^{-1} X^\intercal W \quad \text{ and } \tag{3.21}$$

$$G = \operatorname{diag}\left\{ \sqrt{\frac{2\pi}{\theta_1^* v^{\alpha_1^*}}} \exp\left\{\frac{1}{2}\theta_1^* v^{\alpha_1^*}\right\} : \; v \in \mathcal{V} \right\}. \tag{3.22}$$

Proof. For $v \in \mathcal{V}$ we have $y_v = g(\widetilde{\chi}(v,0))$ with $g(x) = 2\log(\Phi^{-1}(1-x/2))$. The derivative of g is given by

$$g'(x) = -\Big(\Phi^{-1}(1-\frac{x}{2})\,\varphi(\Phi^{-1}(1-\frac{x}{2}))\Big)^{-1},$$

where φ is the univariate standard normal density. Thus,

$$g'(\chi(v,0)) = -\Big(\sqrt{\theta_1^* v^{\alpha_1^*}}\,\varphi\big(\sqrt{\theta_1^* v^{\alpha_1^*}}\big)\Big)^{-1} = -\sqrt{\frac{2\pi}{\theta_1^* v^{\alpha_1^*}}} \exp\left\{\frac{1}{2}\theta_1^* v^{\alpha_1^*}\right\}.$$

Using the multivariate delta method together with Theorems 3.2 and 3.4 it follows that

$$\frac{n}{m_n}\big(y_v - g(\chi(v,0))\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(\mathbf{0}, G\Pi_2^{(\mathrm{iso})}G), \quad n \to \infty,$$

where G is defined in (3.22). Since

$$\min_{\substack{\theta_1,\alpha_1\\ \alpha_1\in(0,2]}} \sum_{v\in\mathcal{V}} w_v\big(g(\chi(v,0)) - \big(\log(\theta_1) + \alpha_1 x_v\big)\big)^2 = \sum_{v\in\mathcal{V}} w_v\big(g(\chi(v,0)) - \big(\log(\theta_1^*) + \alpha_1^* x_v\big)\big)^2,$$

we find the well-known property of unbiasedness of the WLSE,

$$Q_x^{(w)}(g(\chi(v,0)))_{v\in\mathcal{V}}^\intercal = \underset{\substack{\theta_1,\alpha_1\\ \alpha_1\in(0,2]}}{\arg\min} \sum_{v\in\mathcal{V}} w_v\big(g(\chi(v,0)) - \big(\log(\theta_1) + \alpha_1 x_v\big)\big)^2 = \boldsymbol{\psi}_1^*.$$

It follows that, as $n \to \infty$,

$$\frac{n}{m_n}\left(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*\right) = \frac{n}{m_n} Q_x^{(w)}\big(y_v - g(\chi(v,0))\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, Q_x^{(w)} G\Pi_2^{(\mathrm{iso})} G Q_x^{(w)\intercal}\right).$$

We now turn to the constraints on α_1. Since the objective function is quadratic, if the unconstrained estimate exceeds two, the constraint $\alpha_1 \in (0,2]$ results in an estimate $\widehat{\alpha}_1^c = 2$. We consider separately the cases $\alpha_1^* < 2$ and $\alpha_1^* = 2$; i.e., the true parameter lies either in the interior or on the boundary of the parameter space. The constrained estimator $\widehat{\boldsymbol{\psi}}_1^c$ can be written as

$$\widehat{\boldsymbol{\psi}}_1^c = \widehat{\boldsymbol{\psi}}_1 \mathbb{1}_{\{\widehat{\alpha}_1 \le 2\}} + (\widehat{\theta}_1, 2)^\intercal \mathbb{1}_{\{\widehat{\alpha}_1 > 2\}}.$$

We calculate the asymptotic probabilities for the events $\{\widehat{\alpha}_1 \le 2\}$ and $\{\widehat{\alpha}_1 > 2\}$,

$$\mathbb{P}(\widehat{\alpha}_1 \le 2) = P\Big(\frac{n}{m_n}(\widehat{\alpha}_1 - \alpha_1^*) \le \frac{n}{m_n}(2 - \alpha_1^*)\Big).$$

Since for $\alpha_1^* < 2$ as $n \to \infty$

$$\frac{n}{m_n}\left(\widehat{\alpha}_1 - \alpha_1^*\right) \xrightarrow{d} \mathcal{N}\left(0, (0,1)\Pi_3^{(\mathrm{iso})}(0,1)^\intercal\right) \quad \text{and} \quad \frac{n}{m_n}(2-\alpha_1^*) \to \infty,$$

it follows that

$$\mathbb{P}(\widehat{\alpha}_1 \leq 2) \to 1 \quad \text{and} \quad \mathbb{P}(\widehat{\alpha}_1 > 2) \to 0, \quad n \to \infty. \tag{3.23}$$

Therefore, for $\alpha_1^* < 2$,

$$\frac{n}{m_n}\left(\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^*\right) \xrightarrow{d} \mathcal{N}(0, \Pi_3^{(\mathrm{iso})}), \quad n \to \infty.$$

We now consider the case $\alpha_1^* = 2$ and $\widehat{\alpha}_1 > 2$ (the unconstrained estimate exceeds 2). In this case (3.16) leads to the constrained optimization problem

$$\begin{aligned}&\min_{\boldsymbol{\psi}_1}\{[W^{1/2}((y_v)_{v\in\mathcal{V}}^\intercal - X\boldsymbol{\psi}_1)]^\intercal [W^{1/2}((y_v)_{v\in\mathcal{V}}^\intercal - X\boldsymbol{\psi}_1]\},\\ &\text{s.t.} \quad (0,1)\boldsymbol{\psi}_1 = 2.\end{aligned}$$

To obtain asymptotic results for $\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^*$, the vector $\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*$ is projected onto the line $\Lambda = \{\boldsymbol{\psi} \in \mathbb{R}^2, (0,1)\boldsymbol{\psi} = 0\}$, i.e., denoting by I_2 the 2×2-identity matrix, the projection matrix is given by

$$P_\Lambda = I_2 - (X^\intercal W X)^{-1}(0,1)^\intercal((0,1)(X^\intercal WX)^{-1}(0,1)^\intercal)^{-1}(0,1).$$

For simplicity we use the abbreviation $p_{wx} = \sum_{v\in\mathcal{V}} w_v x_v / \sum_{v\in\mathcal{V}} w_v$. We calculate

$$\begin{aligned}&(\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^*)\mathbb{1}_{\{\widehat{\alpha}_1>2\}} = P_\Lambda(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*)\mathbb{1}_{\{\widehat{\alpha}_1>2\}}\\ &= (\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*)\mathbb{1}_{\{\widehat{\alpha}_1>2\}} - (X^\intercal WX)^{-1}(0,1)^\intercal\left((0,1)(X^\intercal WX)^{-1}(0,1)^\intercal\right)^{-1}(\widehat{\alpha}_1-2)\mathbb{1}_{\{\widehat{\alpha}_1>2\}}\\ &= (\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*)\mathbb{1}_{\{\widehat{\alpha}_1>2\}} + \begin{pmatrix} p_{wx} \\ -1 \end{pmatrix}(\widehat{\alpha}_1-2)\mathbb{1}_{\{\widehat{\alpha}_1>2\}}.\end{aligned}$$

For the joint constrained estimator $\boldsymbol{\psi}_1^c$ we obtain

$$\begin{aligned}\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^* &= (\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^*)\mathbb{1}_{\{\widehat{\alpha}_1\leq 2\}} + (\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^*)\mathbb{1}_{\{\widehat{\alpha}_1>2\}}\\ &= (\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*)\mathbb{1}_{\{\widehat{\alpha}_1\leq 2\}} + (\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*)\mathbb{1}_{\{\widehat{\alpha}_1>2\}} + \begin{pmatrix} p_{wx} \\ -1 \end{pmatrix}(\widehat{\alpha}_1-2)\mathbb{1}_{\{\widehat{\alpha}_1>2\}}\\ &= (\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*) + \begin{pmatrix} p_{wx} \\ -1 \end{pmatrix}(\widehat{\alpha}_1-2)\mathbb{1}_{\{\widehat{\alpha}_1>2\}}.\end{aligned}$$

This implies

$$\frac{n}{m_n}(\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^*) = \frac{n}{m_n}\begin{pmatrix} (\log(\widehat{\theta}_1) - \log(\theta_1^*)) + p_{wx}(\widehat{\alpha}_1-2)\mathbb{1}_{\{\widehat{\alpha}_1>2\}} \\ (\widehat{\alpha}_1-2) - (\widehat{\alpha}_1-2)\mathbb{1}_{\{\widehat{\alpha}_1>2\}} \end{pmatrix}.$$

Let $f(x_1,x_2) = (x_1 + p_{wx}x_2\mathbb{1}_{\{x_2>0\}}, x_2 - x_2\mathbb{1}_{\{x_2>0\}})^\intercal$ and observe that $f(c(x_1,x_2)) = cf(x_1,x_2)$ for $c \geq 0$. For the asymptotic distribution we calculate

$$\begin{aligned}
&\mathbb{P}\Big(\frac{n}{m_n}(\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^*) \in B\Big) \\
&= \mathbb{P}\Big(\frac{n}{m_n}f(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*) \in B\Big) = \mathbb{P}\Big(f\big(\frac{n}{m_n}(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*)\big) \in B\Big) \\
&= \mathbb{P}\Big(\frac{n}{m_n}(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*) \in f^{-1}(B \cap \{(b_1,b_2) \in \mathbb{R}^2 : b_2 < 0\}) \cup f^{-1}(B \cap \{(b_1,0) : b_1 \in \mathbb{R}\})\Big) \\
&= \mathbb{P}\Big(\frac{n}{m_n}(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*) \in (B \cap \{(b_1,b_2) \in \mathbb{R}^2 : b_2 < 0\}) \\
&\quad \cup (\{(b_1 - p_{wx}b_2, b_2), b_2 \geq 0, (b_1,0) \in B)\Big) \\
&\to \int_{B\cap\{(b_1,b_2)\in\mathbb{R}^2, b_2<0\}} \varphi_{\mathbf{0},\Pi_3^{(\text{iso})}}(z_1,z_2)dz_1dz_2 \\
&\qquad\qquad + \int_0^\infty \int_{\{b_1\in\mathbb{R},(b_1,0)\in B\}} \varphi_{\mathbf{0},\Pi_3^{(\text{iso})}}(z_1 - p_{wx}z_2, z_2)dz_1dz_2, \quad n \to \infty.
\end{aligned}$$

Plugging in $B = (-\infty, p_1] \times (-\infty, p_2]$ and using the Fubini-Tonelli theorem yields (3.19). □

***Remark* 3.3.** The derivation of the asymptotic properties for the constrained estimate is in fact a special case of Corollary 1 in Andrews [1], who shows asymptotic properties of parameter estimates in a very general setting, when the true parameter is on the boundary of the parameter space. The asymptotic distribution of the estimates in the case $\alpha_1^* = 2$ results from the fact that approximately half of the estimates lie above the true value and are therefore equal to two, which is reflected by the second term in the asymptotic distribution of the estimates. □

3.3. Asymptotic properties of the empirical temporal extremogram

The results for the temporal parameter (θ_2, α_2) are analogous to those for the spatial parameter as presented in Sections 3.1 and 3.2. The set of temporal lags used in the estimation is given by

$$\mathcal{U} = \{1, \dots, \overline{p}\}.$$

Theorem 3.7. *For fixed location $\boldsymbol{s} \in \mathcal{S}_n$, consider the Brown-Resnick time series $\{\eta(\boldsymbol{s},t) : t \in [0,\infty)\}$ as defined in (2.1) with dependence function given in (2.2). Set $m_T = T^{\beta_1}$ for $\beta_1 \in (0,1)$. Then the empirical temporal extremogram $\widehat{\chi}^{(\boldsymbol{s})}(0,u)$ defined in (2.9) with the quantile $q = m_T$ satisfies*

$$\left(\frac{T}{m_T}\right)^{1/2} \left(\widehat{\chi}^{(\boldsymbol{s})}(0,u) - \chi_T(0,u)\right)_{u\in\mathcal{U}} \xrightarrow{d} \mathcal{N}(0, \Pi_1^{(\text{time})}), \quad T \to \infty,$$

where the covariance matrix $\Pi_1^{(\mathrm{time})}$ is specified in Corollary 3.4 of Davis and Mikosch [7], and χ_T is the pre-asymptotic extremogram

$$\chi_T(0,u) \quad = \quad \frac{\mathbb{P}(\eta(\mathbf{0},0)>m_T,\eta(\mathbf{0},u)>m_T)}{\mathbb{P}(\eta(\boldsymbol{s},0)>m_T)}, \quad u\in\mathcal{U}. \tag{3.24}$$

Proof. We verify the mixing conditions for the central limit theorem of the temporal extremogram in Davis and Mikosch [7], Corollary 3.4.

Define sequences $m_T=T^{\beta_1}$ for $\beta_1\in(0,1)$ and $r_T=T^{\beta_2}$ for $0<\beta_2<\beta_1$, which both tend to infinity as $T\to\infty$ as well as $m_T/T\to 0$ and $r_T/m_T\to 0$. From equation (A.3) and Lemma A.3 the time series $\{\eta(\boldsymbol{s},t),t\in[0,\infty)\}$ is α-mixing with mixing coefficients $\alpha(u)\leq Cu\exp\{-\theta_2u^{\alpha}/2\}$ for some positive constant C. Hence, by Lemma A.3, and temporal parameters (θ_2,α_2),

$$\begin{aligned} m_T\sum_{u=r_T}^{\infty}\alpha(u) \quad &\leq \quad Cm_T\sum_{u=r_T}^{\infty}ue^{-\theta_2u^{\alpha_2}/2}\leq Cm_T\,r_T^2\,e^{-\theta_2r_T^{\alpha_2}/2} \\ &= \quad CT^{\beta_1+2\beta_2}\exp\Big\{-\frac{1}{2}\theta_2T^{\alpha_2\beta_2}\Big\}\to 0, \quad T\to\infty. \end{aligned}\tag{3.25}$$

In addition, for arbitrary $\varepsilon>0$,

$$\begin{aligned} &m_T\sum_{u=k}^{r_T}P\left(\|(\eta(\boldsymbol{s},t_{1+u}),\ldots,\eta(\boldsymbol{s},t_{1+u+p}))\|>\varepsilon m_T,\|(\eta(\boldsymbol{s},t_1),\ldots,\eta(\boldsymbol{s},t_{1+p}))\|>\varepsilon m_T\right) \\ &\leq m_T\sum_{u=k}^{\infty}\sum_{i=u}^{u+p}\sum_{j=0}^{p}P\left(\eta(\boldsymbol{s},t_{1+i})>\varepsilon m_T,\eta(\boldsymbol{s},t_{1+j})>\varepsilon m_T\right). \end{aligned}$$

By a time-wise version of (3.2) and the fact that $\mathbb{P}(\eta(\boldsymbol{s},t_k)>m_T)\sim m_T^{-1}$, it suffices to show that the following sum is finite, which we estimate by

$$\begin{aligned} &\sum_{u=k}^{\infty}\sum_{i=u}^{u+p}\sum_{j=0}^{p}2\Big(1-\Phi\big(\sqrt{\theta_2|t_{i+1}-t_{j+1}|^{\alpha_2}}\big)\Big) \\ \leq &2\sum_{u=k}^{\infty}\sum_{i=0}^{p}\sum_{j=0}^{p}\exp\left\{-\frac{\theta_2}{2}|t_{i+1}-t_{j+1}|^{\alpha_2}\right\}<\infty, \end{aligned}$$

where we use that $1-\Phi(x)\leq\exp\{-x^2/2\}$ for $x>0$. This establishes Condition (M) in Davis and Mikosch [7]. As in equation (3.25), we get for $m_T=T^{\beta_1}$,

$$\frac{T}{m_T}\alpha(r_T)\leq CT^{1+\beta_2-\beta_1}\exp\Big\{-\frac{1}{2}\theta_2T^{\alpha_2\beta_2}\Big\}\to 0, \quad T\to\infty.$$

Furthermore, we need one of the following conditions:

(I) $m_T=o(T^{1/3})$, which is satisfied if and only if $\beta_1<1/3$; or

(II) $m_T r_T^3/T = T^{\beta_1+3\beta_2-1} \to 0$ as $T \to \infty$,
which in particular holds for $\beta_1 \in [\frac{1}{3}, 1)$ and $\beta_2 \in (0, \min\{\beta_1, \frac{1}{3}(1-\beta_1)\})$, and
$m_T^4 T^{-1} \sum_{u=r_T}^{m_T} \alpha(u) \le C T^{4\beta_1+2\beta_2-1} \exp\big\{-\theta_2 T^{\alpha_2\beta_2}/2\big\} \to 0$ as $T \to \infty$,
which is satisfied for $\beta_1 \in [1/3, 1)$ and $\beta_2 < \beta_1$.

□

***Remark* 3.4.** The following is the time-wise analogue of Remark 3.1, and gives the convergence rate of the pre-asymptotic extremogram (3.24) to the extremogram. First note that the pre-asymptotic extremogram has for $u \in \mathcal{U}$ the representation

$$\chi_T(0,u) = \Big[\chi(0,u) + \frac{1}{2m_T}\big(\chi(0,u)-2\big)\big(\chi(0,u)-1\big)\Big](1+o(1)), \quad T \to \infty, \tag{3.26}$$

which can, similarly as Lemma 3.1, be deduced from equation (A.4) of Buhl and Klüppelberg [5]. Hence we have for $u \in \mathcal{U}$,

$$\begin{aligned}&\Big(\frac{T}{m_T}\Big)^{1/2}\Big(\chi_T(0,u)-\chi(0,u)\Big)\\ &\sim \Big(\frac{T}{m_T}\Big)^{1/2}\frac{1}{2m_T}\Big[\big(\chi(0,u)-2\big)\big(\chi(0,u)-1\big)\Big] \to 0, \quad T \to \infty\end{aligned}$$

for $m_T = T^{\beta_1}$ if and only if $\beta_1 > 1/3$. Thus we have the two cases:

(I) For $\beta_1 \le 1/3$ we cannot replace the pre-asymptotic extremogram with the theoretical version, but can resort to a bias correction, which is described in (3.28) below.

(II) For $1/3 < \beta_1 < 1$ we obtain indeed

$$\Big(\frac{T}{m_T}\Big)^{1/2}\big(\widehat{\chi}^{(\boldsymbol{s})}(0,u)-\chi(0,u)\big)_{u\in\mathcal{U}} \xrightarrow{d} \mathcal{N}(0,\Pi_1^{(\mathrm{time})}), \quad T \to \infty. \tag{3.27}$$

□

We now turn to the bias correction needed in case (I) for asymptotic normality. Motivated by equation (3.26), we propose for fixed $\boldsymbol{s} \in \mathcal{S}_n$ and all $u \in \mathcal{U}$ the *bias corrected empirical temporal extremogram*

$$\widehat{\chi}^{(\boldsymbol{s})}(0,u) - \frac{1}{2m_T}\big(\widehat{\chi}^{(\boldsymbol{s})}(0,u)-2\big)\big(\widehat{\chi}^{(\boldsymbol{s})}(0,u)-1\big), \quad \boldsymbol{s} \in \mathcal{S}_n,$$

and set

$$\widetilde{\chi}^{(\boldsymbol{s})}(0,u) := \tag{3.28}$$

$$\begin{cases}\widehat{\chi}^{(\boldsymbol{s})}(0,u) - \frac{1}{2m_T}\big(\widehat{\chi}^{(\boldsymbol{s})}(0,u)-2\big)\big(\widehat{\chi}^{(\boldsymbol{s})}(0,u)-1\big) & \text{if } m_T = T^{\beta_1} \text{ with } \beta_1 \in (\frac{1}{5},\frac{1}{3}],\\ \widehat{\chi}^{(\boldsymbol{s})}(0,u) & \text{if } m_T = T^{\beta_1} \text{ with } \beta_1 \in (\frac{1}{3},1).\end{cases} \tag{3.29}$$

We conclude this section by proving asymptotic normality of the bias corrected temporal extremogram centred by the true one. The proof is analogous to that of Theorem 3.4 and shows in particular why β_1 needs to be larger than $1/5$. The extension of the statement to spatial means of extremograms follows in the same way as in Corollary 3.3 by using the vectorized process

$$\{\boldsymbol{Y}(t) : t \in [0,\infty)\} = \{(\eta(\boldsymbol{s},t),\ldots,\eta(\boldsymbol{s},t+\overline{p}))^{\intercal}_{\boldsymbol{s}\in\mathcal{S}_n} : t \in [0,\infty)\}$$

and defining sets $D_{u,k}$ and D_k for $u = 1,\ldots,\overline{p}$ and $k = 1,\ldots,n^2$ properly to extend the covariance matrix. This leads to the statement in (3.31), where

$$\begin{aligned}\Pi_2^{(\text{time})} =&(n\,\mu((1,\infty)))^{-4}\begin{pmatrix}1 & 0\cdots0 & 1 & 0\cdots0 & \cdots & 1 & 0\cdots0\\ 0 & 1\cdots0 & 0 & 1\cdots0 & \cdots & 0 & 1\cdots0\\ & & & \ddots & & & \\ 0 & 0\cdots1 & 0 & 0\cdots1 & \cdots & 0 & 0\cdots1\end{pmatrix}\boldsymbol{F}'\Sigma'(\boldsymbol{F}')^{\intercal}\\ &\begin{pmatrix}1 & 0\cdots0 & 1 & 0\cdots0 & \cdots & 1 & 0\cdots0\\ 0 & 1\cdots0 & 0 & 1\cdots0 & \cdots & 0 & 1\cdots0\\ & & & \ddots & & & \\ 0 & 0\cdots1 & 0 & 0\cdots1 & \cdots & 0 & 0\cdots1\end{pmatrix}^{\intercal},\end{aligned}\tag{3.30}$$

and $\boldsymbol{F}'$ and Σ' are defined in a similar fashion as the matrices $\boldsymbol{F}$ and Σ in Corollary 3.3.

Theorem 3.8. *For a fixed location $\boldsymbol{s} \in \mathcal{S}_n$ consider the Brown-Resnick time series $\{\eta(\boldsymbol{s},t), t \in [0,\infty)\}$ as defined in (2.1) with dependence function given in (2.2). Set $m_T = T_1^{\beta}$ for $\beta_1 \in \big(1/5, 1/3\big]$. Then the bias corrected empirical temporal extremogram (3.28) satisfies*

$$\left(\frac{T}{m_T}\right)^{1/2}\Big(\widetilde{\chi}^{(\boldsymbol{s})}(0,u)-\chi(0,u)\Big)_{u\in\mathcal{U}} \xrightarrow{d} \mathcal{N}(0,\Pi_1^{(\text{time})}),\quad T\to\infty,$$

with covariance matrix $\Pi_1^{(\text{time})}$ as in Theorem 3.7. Furthermore, the corresponding bias corrected version $\widetilde{\chi}(0,u) = n^{-2}\sum_{i=1}^{n^2}\widetilde{\chi}^{(\boldsymbol{s}_i)}(0,u)$ of (2.11) satisfies

$$\left(\frac{T}{m_T}\right)^{1/2}\Big(\frac{1}{n^2}\sum_{i=1}^{n^2}\widetilde{\chi}^{(\boldsymbol{s}_i)}(0,u)-\chi(0,u)\Big)_{u\in\mathcal{U}} \xrightarrow{d} \mathcal{N}(0,\Pi_2^{(\text{time})}),\quad T\to\infty,\tag{3.31}$$

with covariance matrix $\Pi_2^{(\text{time})}$ specified in equation (3.30).

***Remark* 3.5.** Note that in (3.27) the rate of convergence is of the order n^a for $a \in (0,1/3)$. On the other hand, after bias correction in (3.31) we obtain convergence of the order n^a for $a \in [1/3, 2/5)$. Thus, the bias correction leads to better rates compared to those in Davis and Mikosch [7], where no bias correction is applied. □

3.4. Asymptotic properties of temporal parameter estimates

The asymptotic normality of the WLSE $(\widehat{\theta}_2, \widehat{\alpha}_2)$ of Section 2 can be derived in exactly the same way as for the spatial parameter estimates $(\widehat{\theta}_1, \widehat{\alpha}_1)$. Accordingly, we define

$$y_u := 2\log\Big(\Phi^{-1}\big(1-\frac{1}{2}\widetilde{\chi}(0,u)\big)\Big) \quad \text{and} \quad x_u := \log(u), \quad u \in \mathcal{U},$$

where $\widetilde{\chi}(0,u) = \frac{1}{n^2}\sum_{i=1}^{n^2} \widetilde{\chi}^{(\boldsymbol{s}_i)}(0,u)$ as in (2.11), possibly after a bias correction, which depends on the two cases described in Remark 3.4. Then (2.13) reads as

$$\begin{pmatrix} \widehat{\theta}_2 \\ \widehat{\alpha}_2 \end{pmatrix} = \mathop{\arg\min}_{\substack{\theta_2,\alpha_2>0\\ \alpha_1\in(0,2]}} \sum_{v\in\mathcal{V}} w_v\big(y_v - \big(\log(\theta_2)+\alpha_2 x_v\big)\big)^2 \tag{3.32}$$

We also define the design matrix X and weight matrix W as

$$X = [\mathbf{1}, (x_u)_{u\in\mathcal{U}}^{\intercal}] \in \mathbb{R}^{\overline{p}\times 2} \quad \text{and} \quad W = \operatorname{diag}\{w_u : u\in\mathcal{U}\} \in \mathbb{R}^{\overline{p}\times\overline{p}},$$

where $\mathbf{1} = (1,\dots,1)^{\intercal} \in \mathbb{R}^{\overline{p}}$. We state asymptotic normality of the WLSE of the time parameters.

Theorem 3.9. *Let $\widehat{\boldsymbol{\psi}}_2^c = (\log(\widehat{\theta}_2^c), \widehat{\alpha}_2^c)^{\intercal}$ denote the WLSE resulting from the constrained minimization problem in (3.32) and $\boldsymbol{\psi}_2^* = (\log(\theta_2^*), \alpha_2^*)^{\intercal} \in \Psi$ the true parameter vector. Set $m_T = T^{\beta_1}$ for $\beta_1 \in (1/5, 1)$. Then as $T\to\infty$,*

$$\left(\frac{T}{m_T}\right)^{1/2}\left(\widehat{\boldsymbol{\psi}}_2^c - \boldsymbol{\psi}_2^*\right) \xrightarrow{d} \begin{cases} \boldsymbol{Z}_1 & \text{if } \alpha_2^* < 2, \\ \boldsymbol{Z}_2 & \text{if } \alpha_2^* = 2, \end{cases} \tag{3.33}$$

where $\boldsymbol{Z}_1 \sim \mathcal{N}(\mathbf{0}, \Pi_3^{(\text{time})})$, and the distribution of $\boldsymbol{Z}_2$ is given by

$$\begin{aligned} P\left(\boldsymbol{Z}_2 \in B\right) = & \int_{B\cap\{(b_1,b_2)\in\mathbb{R}^2: b_2<0\}} \varphi_{\mathbf{0},\Pi_3^{(\text{time})}}(z_1,z_2)dz_1dz_2 \\ & + \int_0^\infty \int_{\{b_1\in\mathbb{R}:(b_1,0)\in B\}} \varphi_{\mathbf{0},\Pi_3^{(\text{time})}}\Big(z_1 - \frac{1}{\sum_{u\in\mathcal{U}} w_u}\sum_{u\in\mathcal{U}} w_u x_u\ z_2,\ z_2\Big)dz_1dz_2 \end{aligned} \tag{3.34}$$

for every Borel set $B\subset\mathbb{R}$, and $\varphi_{\mathbf{0},\Sigma}$ denotes the bivariate normal density with mean vector $\mathbf{0}$ and covariance matrix Σ. In particular, the joint distribution function of $\boldsymbol{Z}_2$ is given for $(p_1,p_2)^{\intercal}\in\mathbb{R}^2$ by

$$\mathbb{P}\left(\boldsymbol{Z}_2 \le (p_1,p_2)^{\intercal}\right) = \int_{-\infty}^{\min\{0,p_2\}} \int_{-\infty}^{p_1} \varphi_{\mathbf{0},\Pi_3^{(\text{time})}}(z_1,z_2)dz_1dz_2$$

$$+ \mathbb{1}_{\{p_2 \geq 0\}} \int_0^{\infty} \int_{-\infty}^{p_1} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{time})}} \Big(z_1 - \frac{1}{\sum_{u\in\mathcal{U}} w_u} \sum_{u\in\mathcal{U}} (w_u x_u)\ z_2,\ z_2 \Big) dz_1 dz_2. \quad (3.35)$$

The covariance matrix of $\boldsymbol{Z}_1$ *has representation*

$$\Pi_3^{(\mathrm{time})} = Q_x^{(w)} G \Pi_2^{(\mathrm{time})} G Q_x^{(w)\intercal}, \quad (3.36)$$

where $\Pi_2^{(\mathrm{time})}$ *is the covariance matrix given in* (3.30),

$$Q_x^{(w)} = (X^\intercal W X)^{-1} X^\intercal W \quad \textit{and} \quad (3.37)$$

$$G = \mathrm{diag}\left\{ \sqrt{\frac{2\pi}{\theta_2^* u^{\alpha_2^*}}} \exp\left\{ \theta_2^* u^{\alpha_2^*}/2 \right\},\ u \in \mathcal{U} \right\}. \quad (3.38)$$

4. Subsampling for confidence regions

Let $\{\eta(\boldsymbol{s},t) : \boldsymbol{s} \in \mathbb{R}^2, t \in [0,\infty)\}$ be the isotropic Brown-Resnick process as in (2.1) with dependence function δ given in (2.2); i.e.,

$$\delta(v,u) = 2\theta_1 v^{\alpha_1} + 2\theta_2 u^{\alpha_2}, \quad v,u \geq 0,$$

for $\theta_1, \theta_2 > 0$ and $\alpha_1, \alpha_2 \in (0,2]$. We assume to observe the process on a regular grid $\mathcal{S}_n = \{(i_1,i_2) : i_1, i_2 \in \{1,\ldots,n\}\}$ and at time points $\{t_1,\ldots,t_T\} = \{1,\ldots,T\}$ as specified in Condition 2.4. The estimation methods based on the spatial and temporal extremograms described in Sections 2 and 3 yield consistent estimators $\widehat{\boldsymbol{\psi}}_1^c = (\log(\widehat{\theta}_1^c), \widehat{\alpha}_1^c)^\intercal$ and $\widehat{\boldsymbol{\psi}}_2^c = (\log(\widehat{\theta}_2^c), \widehat{\alpha}_2^c)^\intercal$ of the true spatial and temporal parameters $\boldsymbol{\psi}_1^\star = (\log(\theta_1^\star), \alpha_1^\star)^\intercal$ and $\boldsymbol{\psi}_2^\star = (\log(\theta_2^\star), \alpha_2^\star)^\intercal$, respectively. Furthermore, we have the limit theorems

$$\tau_n\Big(\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^\star\Big) \stackrel{d}{\rightarrow} \boldsymbol{Z}^{(1)}, \quad n\to\infty, \quad \text{and} \quad \tau_T\Big(\widehat{\boldsymbol{\psi}}_2^c - \boldsymbol{\psi}_2^\star\Big) \stackrel{d}{\rightarrow} \boldsymbol{Z}^{(2)}, \quad T\to\infty,$$

where the bivariate distributions of $\boldsymbol{Z}^{(1)}$ and $\boldsymbol{Z}^{(2)}$ depend on the true parameter values $\alpha_1^\star$ and $\alpha_2^\star$, respectively. The rates of convergence are given by $\tau_n := \frac{n}{m_n}$ and $\tau_T := \sqrt{\frac{T}{m_T}}$, where m_n and m_T are appropriately chosen scaling sequences.

Due to the complicated forms of the covariance matrices of $\boldsymbol{Z}^{(1)}$ and $\boldsymbol{Z}^{(2)}$ (cf. Theorems 3.6 and 3.9) we use resampling methods to construct asymptotic confidence regions of the parameter vectors $\boldsymbol{\psi}_1^\star$ and $\boldsymbol{\psi}_2^\star$. One appealing method is subsampling (see Politis et al. [24], Chapter 5), since it works under weak regularity conditions asymptotically correct. The central assumption is the existence of weak limit laws, which is guaranteed by Theorems 3.6 and 3.9. In Section 4.1 we consider the spatial case, whereas Section 4.2 deals with the temporal case.

We have applied subsampling successfully already for confidence bounds of pairwise likelihood estimates of the space-time Brown-Resnick process in Buhl and Klüppelberg

[4], Section 4. The procedure is as follows: understanding inequalities between vectors componentwise, we choose block lengths $(1,1,1) \leq \boldsymbol{b} = (b_s, b_s, b_t) \leq (n, n, T)$ and the degree of overlap $(1,1,1) \leq \boldsymbol{e} = (e_s, e_s, e_t) \leq (b_s, b_s, b_t)$, where $\boldsymbol{e} = (1,1,1)$ corresponds to maximum overlap and $\boldsymbol{e} = \boldsymbol{b}$ to no overlap. The blocks are indexed by $\boldsymbol{i} = (i_1, i_2, i_3) \in \mathbb{N}^3$ with $i_j \leq q_s$ for $q_s := \lfloor \frac{n-b_s}{e_s} \rfloor + 1$ and $j = 1, 2$ and $i_3 \leq q_t := \lfloor \frac{T-b_t}{e_t} \rfloor + 1$. This results in a total number of $q = q_s^2 q_t$ blocks, which we summarise in the sets

$$E_{\boldsymbol{i},\boldsymbol{b},\boldsymbol{e}} = \big\{(s_1, s_2, t) \in \mathcal{S}_n \times \{1, \dots, T\} : (i_j - 1)e_s + 1 \leq s_j \leq (i_j - 1)e_s + b_s \text{ for } j = 1, 2; \\ (i_3 - 1)e_t + 1 \leq t \leq (i_3 - 1)e_t + b_t\big\}.$$

We estimate $\theta_1, \alpha_1, \theta_2, \alpha_2$ based on the observations in each block as described in the previous sections. This yields different estimates for each spatial and temporal parameter, which we denote by $\widehat{\boldsymbol{\psi}}^c_{1,\boldsymbol{i}}$ and $\widehat{\boldsymbol{\psi}}^c_{2,\boldsymbol{i}}$, respectively.

4.1. Subsampling: the spatial parameters

Our first theorem below provides a basis for constructing asymptotically valid confidence intervals for the true spatial parameters $\theta_1^\star$ and $\alpha_1^\star$. We define τ_{b_s} as the analogue of $\tau_n = n/m_n = n^{1-\beta}$ where $\beta \in (1/5, 1/2)$; i.e., $\tau_{b_s} := b_s^{1-\beta}$ (cf. Remark 3.1 and Theorem 3.4)

Theorem 4.1. *Assume that the conditions of Theorem 3.6 hold and, as $n \to \infty$,*

(i) $b_s \to \infty$ such that $b_s = o(n)$ (hence, $\tau_{b_s}/\tau_n = (b_s/n)^{1-\beta} \to 0$),
(ii) $\boldsymbol{e}$ does not depend on n.

Define the empirical distribution function $L_{b_s,\boldsymbol{s}}$

$$L_{b_s,s}(x) := \frac{1}{q} \sum_{i_1=1}^{q_s} \sum_{i_2=1}^{q_s} \sum_{i_3=1}^{q_t} \mathbf{1}_{\left\{\tau_{b_s} \|\widehat{\boldsymbol{\psi}}^c_{1,\boldsymbol{i}} - \widehat{\boldsymbol{\psi}}^c_1\| \leq x\right\}}, \quad x \in \mathbb{R},$$

and the empirical quantile function

$$c_{b_s,s}(1-\alpha) := \inf\left\{x \in \mathbb{R} : L_{b_s,s}(x) \geq 1 - \alpha\right\}, \quad \alpha \in (0,1).$$

Then

$$\mathbb{P}\left(\tau_n \|\widehat{\boldsymbol{\psi}}^c_1 - \boldsymbol{\psi}_1^\star\| \leq c_{b_s,s}(1-\alpha)\right) \to 1 - \alpha, \quad n \to \infty. \tag{4.1}$$

Proof. We apply Corollary 5.3.4 of Politis et al. [24]. Their main Assumption 5.3.4 is the existence of a weak limit distribution of $\tau_n \|\widehat{\boldsymbol{\psi}}^c_1 - \boldsymbol{\psi}_1^\star\|$. By Theorem 3.6, the continuous mapping theorem and the Fubini-Tonelli theorem we have for $\gamma \geq 0$, as $n \to \infty$,

$$\mathbb{P}(\tau_n \|\widehat{\boldsymbol{\psi}}^c_1 - \boldsymbol{\psi}_1^\star\| \leq \gamma) \to \mathbb{P}(\|\boldsymbol{Z}_1\| \leq \gamma) = \mathbb{P}(\boldsymbol{Z}_1 \in B(\mathbf{0}, \gamma)) = 2 \int_{-\gamma}^{\gamma} \int_0^{\sqrt{\gamma^2 - r^2}} \varphi_{\mathbf{0}, \Pi_3^{(\mathrm{iso})}}(r, s) ds dr$$

if $\alpha_1^\star < 2$. For $\alpha_1^\star = 2$ we obtain

$$
\begin{aligned}
&\mathbb{P}(\tau_n\|\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^\star\| \le \gamma) \to \mathbb{P}(\|\boldsymbol{Z}_2\| \le \gamma) = \mathbb{P}(\boldsymbol{Z}_2 \in B(\mathbf{0}, \gamma)) \\
&= \int_{-\gamma}^{\gamma} \int_{-\sqrt{\gamma^2 - r^2}}^{0} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(r,s)\,ds\,dr \\
&\quad + \int_{-\gamma}^{\gamma}\int_{0}^{\infty} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(r - \frac{1}{\sum_{v\in\mathcal{V}} w_v}\sum_{v\in\mathcal{V}}(w_v x_v)s, s)\,ds\,dr \\
&= \int_{-\gamma}^{\gamma}\Big\{ \int_{-\sqrt{\gamma^2-r^2}}^{0} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(r,s)\,ds + \int_0^\infty \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}\Big(r - \frac{1}{\sum_{v\in\mathcal{V}} w_v}\sum_{v\in\mathcal{V}}(w_v x_v)s, s\Big)ds\Big\}dr.
\end{aligned}
$$

In particular, the limiting distribution function of the scaled norm $\tau_n\|\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^\star\|$ is continuous in γ both for $\alpha_1^\star < 2$ and $\alpha_1^\star = 2$. Assumptions (i) and (ii) are also presumed in Politis et al. [24]. The required condition on the α-mixing coefficients is satisfied, similarly as in the proof of Theorem 3.2, by equation (A.2) and Lemma A.3, and the result follows. □

As a consequence of equation (4.1), for n large enough, an approximate $(1-\alpha)$-confidence region for the true parameter vector $\boldsymbol{\psi}_1^\star = (\log(\theta_1^\star), \alpha_1^\star)$ is given by

$$\{\boldsymbol{\psi} \in \mathbb{R}\times(0,2] : \|\boldsymbol{\psi} - \widehat{\boldsymbol{\psi}}_1^c\| \le c_{b_s,s}(1-\alpha)/\tau_n\}.$$

The one-dimensional approximate $(1-\alpha)$-confidence intervals for the parameters $\theta_1^\star$ and $\alpha_1^\star$ can be read off from this as

$$
\begin{aligned}
\Big[\widehat{\theta}_1^c \exp\Big\{-\frac{c_{b_s,s}(1-\alpha)}{\tau_n}\Big\}, \widehat{\theta}_1^c \exp\Big\{\frac{c_{b_s,s}(1-\alpha)}{\tau_n}\Big\}\Big] &\text{ and} \\
\Big[\widehat{\alpha}_1^c - \frac{c_{b_s,s}(1-\alpha)}{\tau_n}, \widehat{\alpha}_1^c + \frac{c_{b_s,s}(1-\alpha)}{\tau_n}\Big] \cap (0,2]&.
\end{aligned}
$$

4.2. Subsampling: the temporal parameters

The theorem below provides a basis for constructing asymptotically valid confidence intervals for the true temporal parameters $\theta_2^\star$ and $\alpha_2^\star$. We define τ_{b_t} as the analogue of $\tau_T = \sqrt{\frac{T}{m_T}} = T^{(1-\beta)/2}$ where $\beta \in (1/5, 1)$; i.e., $\tau_{b_t} := b_t^{(1-\beta)/2}$ (cf. Remark 3.4 and Theorem 3.8). The proof is completely analogous to that of Theorem 4.1 for the spatial parameters.

Theorem 4.2. *Assume that the conditions of Theorem 3.9 hold and, as $T \to \infty$,*

(i) $b_t \to \infty$ *such that* $b_t = o(T)$ *(hence,* $\tau_{b_t}/\tau_T = (b_t/T)^{(1-\beta)/2} \to 0$*),*
(ii) $\boldsymbol{e}$ *does not depend on* T.

Define the empirical distribution function $L_{b_t,t}$

$$L_{b_t,t}(x) := \frac{1}{q} \sum_{i_1=1}^{q_s} \sum_{i_2=1}^{q_s} \sum_{i_3=1}^{q_t} \mathbf{1}_{\left\{\tau_{b_t} \|\widehat{\boldsymbol{\psi}}_{2,i}^c - \widehat{\boldsymbol{\psi}}_2^c\| \leq x\right\}}, \quad x \in \mathbb{R},$$

and the empirical quantile function

$$c_{b_t,t}(1-\alpha) := \inf\left\{x \in \mathbb{R} : L_{b_t,t}(x) \geq 1-\alpha\right\}, \quad \alpha \in (0,1).$$

Then

$$\mathbb{P}\left(\tau_T \|\widehat{\boldsymbol{\psi}}_2^c - \boldsymbol{\psi}_2^\star\| \leq c_{b_t,t}(1-\alpha)\right) \to 1-\alpha, \quad n \to \infty. \tag{4.2}$$

As a consequence of equation (4.2), for T large enough, an approximate $(1-\alpha)$-confidence region for the true parameter vector $\boldsymbol{\psi}_2^\star = (\log(\theta_2^\star), \alpha_2^\star)$ is given by

$$\{\boldsymbol{\psi} \in \mathbb{R} \times (0,2] : \|\boldsymbol{\psi} - \widehat{\boldsymbol{\psi}}_2^c\| \leq c_{b_t,t}(1-\alpha)/\tau_T\}.$$

The one-dimensional approximate $(1-\alpha)$-confidence intervals for the parameters $\theta_2^\star$ and $\alpha_2^\star$ can be read off from this as

$$\left[\widehat{\theta}_2^c \exp\left\{-\frac{c_{b_t,t}(1-\alpha)}{\tau_T}\right\}, \widehat{\theta}_2^c \exp\left\{\frac{c_{b_t,t}(1-\alpha)}{\tau_T}\right\}\right] \text{ and}$$
$$\left[\widehat{\alpha}_2^c - \frac{c_{b_t,t}(1-\alpha)}{\tau_T}, \widehat{\alpha}_2^c + \frac{c_{b_t,t}(1-\alpha)}{\tau_T}\right] \cap (0,2].$$

5. Simulation study

We examine the performance of the WLSEs by a simulation study. The estimation of the spatial parameters relies on a rather large number of spatial observations and the estimation of the temporal parameters on a rather large number of observed time points. However, simulation of Brown-Resnick space-time processes based on the exact method proposed by Dombry et al. [13] can be time consuming, if both a large number of spatial locations and of time points is taken. For a time-saving method we generate the process on two different space-time observation areas, one for examining the performance of the spatial estimates and one for the temporal ones, which we call $\mathcal{S}^{(1)} \times \mathcal{T}^{(1)}$ and $\mathcal{S}^{(2)} \times \mathcal{T}^{(2)}$, respectively. The design for the simulation experiment is given in more details as follows:

1. We choose two space-time observation areas

$$\mathcal{S}^{(1)} \times \mathcal{T}^{(1)} = \{(i_1,i_2) : i_1,i_2 \in \{1,\ldots,70\}\} \times \{1,\ldots,10\}$$
$$\mathcal{S}^{(2)} \times \mathcal{T}^{(2)} = \{(i_1,i_2) : i_1,i_2 \in \{1,\ldots,5\}\} \times \{1,\ldots,300\}$$

and the sets $\mathcal{V} = \{1, \sqrt{2}, 2, \sqrt{5}, \sqrt{8}, 3, \sqrt{10}, \sqrt{13}, 4, \sqrt{17}\}$ and $\mathcal{U} = \{1,\ldots,10\}$.

2. We simulate the Brown-Resnick space-time process (2.1) based on the exact method proposed in Dombry et al. [13], using the `R`-package `RandomFields` [26]. The dependence function δ is modelled as in (2.2); i.e.,

$$\delta(v,u) = 2\theta_1 v^{\alpha_1} + 2\theta_2 u^{\alpha_2}, \quad v,u \geq 0,$$

with parameters

$$\theta_1 = 0.4,\ \alpha_1 = 1.5, \quad \theta_2 = 0.2,\ \alpha_2 = 1.$$

3. The parameters $\theta_1, \alpha_1, \theta_2$ and α_2 are estimated.
   - For the estimation of the empirical extremograms (cf. equations 2.8-2.11) we have to choose high empirical quantiles q. In practice, q is chosen from an interval of high quantiles for which the empirical extremogram is robust, see the remarks of Davis et al. [10] after Theorem 2.1. We choose the $90\%-$empirical quantile for the estimation of the spatial parameters and the $70\%-$quantile for the temporal part. The quantile for the temporal part is lower to ensure reliable estimation of the extremogram, because the number of time points (300) used for the estimation of the time parameters is much smaller than the number of spatial locations ($70 \cdot 70 = 4900$) used for the estimation of the spatial parameters.
   - The weights in the constrained weighted linear regression problem (see 3.16 and 3.32) are chosen such that locations and time points which are further apart of each other have less influence on the estimation. More precisely, we choose

$$w_u = \exp\{-u^2\} \text{ for } u \in \mathcal{U} \quad \text{ and } \quad w_v = \exp\{-v^2\} \text{ for } v \in \mathcal{V}.$$

   This choice of weights reflects the exponential decay of $\chi(v,0)$ and $\chi(0,u)$ defined in (2.6), which are tail probabilities of the standard normal distribution Φ.
4. Pointwise confidence bounds are computed by subsampling as described in Sections 4.1 and 4.2. We choose block lengths $\boldsymbol{b} = (60,60,10)$ and overlap $\boldsymbol{e} = (2,2,2)$ for the space-time process with observation area $\mathcal{S}^{(1)} \times \mathcal{T}^{(1)}$ and $\boldsymbol{b} = (5,5,200)$, $\boldsymbol{e} = (1,1,1)$ for the process with observation area $\mathcal{S}^{(2)} \times \mathcal{T}^{(2)}$.
5. Steps 1 - 5 are repeated 100 times.

Figures 5.1 and 5.2 show the estimates of the spatial parameters (θ_1, α_1) and temporal parameters (θ_2, α_2) for each of the 100 realizations of the Brown-Resnick space-time process. The dotted lines above and below the dots are pointwise confidence intervals based on subsampling. Table 1 shows the mean, RMSE and MAE of the 100 simulations. Altogether, we observe that the estimates are close to the true values. Moreover, the spatial confidence intervals are more accurate than the temporal ones, which is due to the larger number of observations in space than in time.

	MEAN	RMSE	MAE
θ_1	0.4033	0.0678	0.0559
α_1	1.4984	0.0521	0.0400
θ_2	0.2249	0.0649	0.0526
α_2	0.9563	0.0939	0.0767

Table 1.: Mean, root mean squared error and mean absolute error of the WLSE.

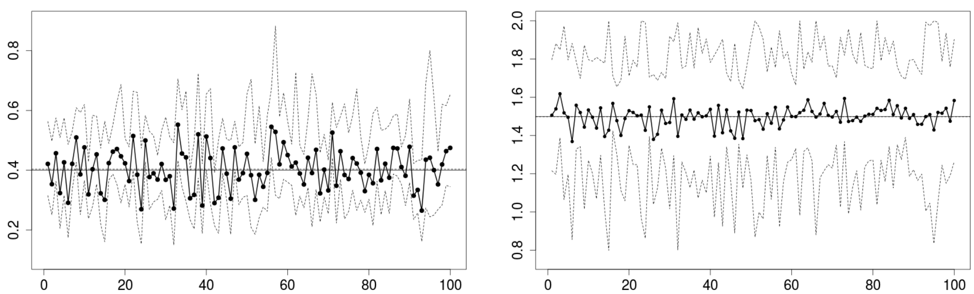

Figure 5.1: WLSEs of θ_1 (left) and α_1 (right) for 100 simulated Brown-Resnick space-time processes together with pointwise 95%−subsampling confidence intervals (dotted). The middle solid line is the true value and the middle dotted line represents the mean over all estimates.

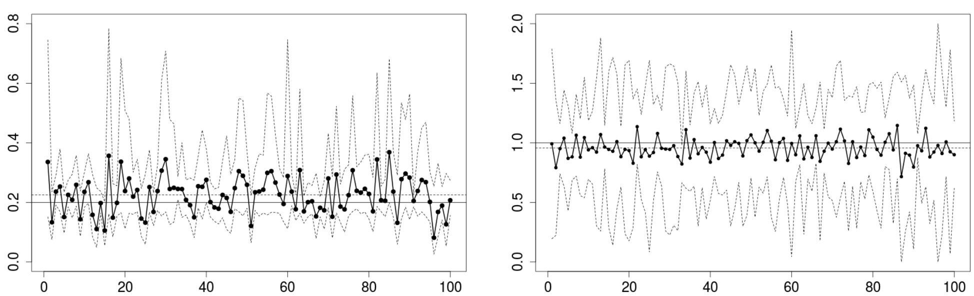

Figure 5.2: WLSEs of θ_2 (left) and α_2 (right) for 100 simulated Brown-Resnick space-time processes together with pointwise 95%−subsampling confidence intervals (dotted). The middle solid line is the true value and the middle dotted line represents the mean over all estimates.

In Section 6, Steinkohl [27] carries out a detailed simulation study with the goal to compare the performance of the semiparametric estimation method with the pairwise likelihood approach in finite samples. To summarise her findings, the WLSE is slightly more biased than the pairwise likelihood estimator. This is due to the fact that the bias correction in the semiparametric estimation depends crucially on the chosen threshold as it applies only asymptotically. The main advantage of the semiparametric method is computation time, which is about 15 times lower in the setting considered by [27].

6. Analysis of radar rainfall measurements

Finally, we apply the Brown-Resnick space-time model in (2.1) and the WLSE to radar rainfall data. The data were collected by the Southwest Florida Water Management District (SWFWMD)[1]. Our objective is to quantify the extremal behaviour of radar rainfall data in a region in Florida by using spatial and temporal block maxima and fitting a Brown-Resnick space-time model to the block maxima.

The data base consists of radar values in inches measured on a 120×120km region containing 3600 grid locations. We calculate the spatial and temporal maxima over subregions of size 10×10km and over 24 subsequent measurements of the corresponding hourly accumulated time series in the wet season (June to September) from the years 1999-2004 for further analysis. In this way we obtain 12×12 locations during 732 days containing space-time block maxima of rainfall observations.

We denote the set of locations by $\mathcal{S} = \{(i_1, i_2), i_1, i_2 \in \{1, \ldots, 12\}\}$ and the space-time observations by $\{\eta(\boldsymbol{s}, t), \boldsymbol{s} \in \mathcal{S}, t \in \{t_1, \ldots, t_{732}\}\}$. This setup is also considered in Buhl and Klüppelberg [4], Section 5 and Steinkohl [27], Chapter 7. To make the results obtained there comparable with the results here, we use the data preprocessed as there and after the same marginal modelling steps; for precise description cf. [4], Section 5.1. Since the data do not fail the max-stability check described in Section 5.2 of that paper, we assume that $\{\eta(\boldsymbol{s}, t), \boldsymbol{s} \in \mathcal{S}, t \in \{t_1, \ldots, t_{732}\}\}$ are realisations of a max-stable space-time process with standard unit Fréchet margins.

We then fit the Brown-Resnick model (2.1) by estimating the extremal dependence structure (2.2) as follows:

1. We estimate the parameters θ_1, α_1, θ_2 and α_2 by WLSE as described in Section 2 based on the sets $\mathcal{V} = \{1, \sqrt{2}, 2, \sqrt{5}, \sqrt{8}, 3, \sqrt{10}, \sqrt{13}, 4, \sqrt{17}\}$ and $\mathcal{U} = \{1, \ldots, 10\}$ for the linear regression. Permutation tests as described below and visualised in Figure 6.3 indicate that these lags are sufficient to cover the relevant extremal dependence structure. Since the true extremogram χ is unknown, we choose as weights for the different spatial and temporal lags $v \in \mathcal{V}$ and $u \in \mathcal{U}$ the corresponding estimated averaged extremogram values; i.e., $w_v = T^{-1} \sum_{k=1}^{T} \widetilde{\chi}^{(t_k)}(v, 0)$ and $w_u = n^{-2} \sum_{i=1}^{n^2} \widetilde{\chi}^{(\boldsymbol{s}_i)}(0, u)$, respectively. Since the so defined weights are random, what follows is conditional on the realisations of these weights.
   As the number of spatial points in the analysis is rather small, we cannot choose a very high empirical quantile q, since this would in turn result in a too small number of exceedances to get a reliable estimate of the extremogram. Hence, we choose q as the empirical $60\%-$quantile, relying on the fact that the block maxima generate a max-stable process.
   For the temporal estimation, we choose the empirical $90\%-$quantile for q.
2. We perform subsampling (see Section 4) to construct 95%-confidence intervals for each parameter estimate. As subsample block sizes we choose $b_s = 12$ (due to the small number of spatial locations) and $b_t = 300$. We further choose $e_s = e_t = 1$, which corresponds to the maximum degree of overlap.

[1] http://www.swfwmd.state.fl.us/

The results are shown in Figures 6.2, 6.3 and Table 2. Figure 6.1 visualizes the daily rainfall maxima for the two grid locations $(1, 1)$ and $(5, 6)$. The semiparametric estimates together with subsampling confidence intervals are given in Table 2.

For comparison we present the parameter estimates from the pairwise likelihood estimation (for details for the isotropic Brown-Resnick model see Davis et al. [8] and [27], Chapter 7), where we obtained $\widetilde{\theta}_1 = 0.3485$, $\widetilde{\alpha}_1 = 0.8858$, $\widetilde{\theta}_2 = 2.4190$ and $\widetilde{\alpha}_2 = 0.1973$. From Table 2 we recognize that these estimates are close to the semiparametric estimates and even lie in most cases in the 95%-subsampling confidence intervals.

Figure 6.2 shows the spatial and temporal mean of empirical temporal (left) and spatial (right) extremograms as described in (2.10) and (2.11) together with 95% subsampling confidence intervals. We perform a permutation test to test the presence of extremal independence. To this end we randomly permute the space-time data and calculate empirical extremograms as before. More precisely, we compute the empirical temporal extremogram as before and repeat the procedure 1000 times. From the resulting temporal extremogram sample we determine nonparametric 97.5% and 2.5% empirical quantiles, which gives a 95%−confidence region for temporal extremal independence. The analogue procedure is performed for the spatial extremogram.

The results are shown in Figure 6.3 together with the extremogram fit based on the WLSE. The plots indicate that for time lags larger than 3 there is no temporal extremal dependence, and for spatial lags larger than 4 no spatial extremal dependence.

Estimate Subsampling-CI	$\widehat{\theta}_1$	0.3611 [0.3472,0.3755]	$\widehat{\alpha}_1$	0.9876 [0.9482,1.0267]
Estimate Subsampling-CI	$\widehat{\theta}_2$	2.3650 [1.9110,2.7381]	$\widehat{\alpha}_2$	0.0818 [0.0000,0.2680]

Table 2.: Semiparametric estimates for the spatial parameters θ_1 and α_1 and the temporal parameters θ_2 and α_2 of the Brown-Resnick process in (2.1) together with 95% subsampling confidence intervals.

7. Conclusions and Outlook

For the isotropic Brown-Resnick space-time process with flexible dependence structure we have suggested a new semiparametric estimation method, which works remarkably well in an extreme value setting. The method results in quite reliable estimates, much faster than the composite likelihood methods used so far. These estimates can also be used as initial values for a composite likelihood optimization routine to obtain more accurate estimates.

Future work will be dedicated to generalisations of the semiparametric method based on extremogram estimation. At present we work on three topics:

1. Generalize the dependence function (2.2) to anisotropic and appropriate mixed models.

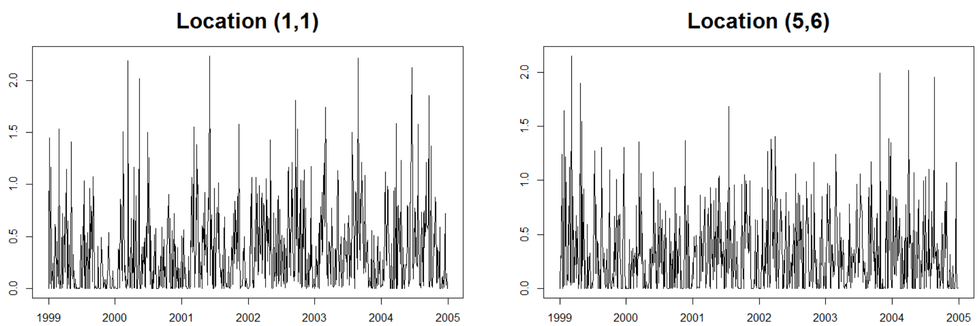

Figure 6.1: Daily rainfall maxima over hourly accumulated measurements from 1999-2004 in inches for two grid locations.

2. Generalize the sampling scheme to allow for a fixed (small) number of spatial observations and consider limit results for the number of temporal observations to tend to infinity.
3. Generalize the least squares estimation to estimate spatial and temporal parameters simultaneously, also in the situation described in topic 2.

Another interesting question concerns the optimal choice of the weight matrix W, such that the asymptotic variance of the WLSE is minimal. Some ideas can be found in the geostatistics literature in the context of least squares estimation of the variogram parameters; see e.g. Lahiri et al. [22], Section 4. They describe the situation, where the optimal choice of the weight matrix is given by the inverse of the asymptotic covariance matrix of the nonparametric estimates; i.e., of $\big(T^{-1}\sum_{k=1}^{T}\widetilde{\chi}^{(t_k)}(v,0)\big)^{\intercal}_{v\in\mathcal{V}}$ in the spatial case and of $\big(n^{-2}\sum_{i=1}^{n^2}\widetilde{\chi}^{(\boldsymbol{s}_i)}(0,u)-\chi(0,u)\big)^{\intercal}_{u\in\mathcal{U}}$ in the temporal case. In our case, however, this involves the matrices $\Pi_2^{(\text{iso})}$ and $\Pi_2^{(\text{time})}$ (given in (4.3)-(4.6) of Buhl and Klüppelberg [5]), whose components are infinite sums.

Appendix A: α-mixing of the Brown-Resnick space-time process

In the following we define α-mixing for spatial processes; see e.g. Doukhan [14] or Bolthausen [2].

***Definition* A.1.** For $d\in\mathbb{N}$, consider a strictly stationary random field $\{X(\boldsymbol{s}):\boldsymbol{s}\in\mathbb{R}^d\}$ and let $d(\cdot,\cdot)$ be some metric induced by a norm on $\mathbb{R}^d$. For $\Lambda_1,\Lambda_2\subset\mathbb{Z}^d$ set

$$d(\Lambda_1,\Lambda_2):=\inf\{d(\boldsymbol{s}_1,\boldsymbol{s}_2):\ \boldsymbol{s}_1\in\Lambda_1,\boldsymbol{s}_2\in\Lambda_2\}.$$

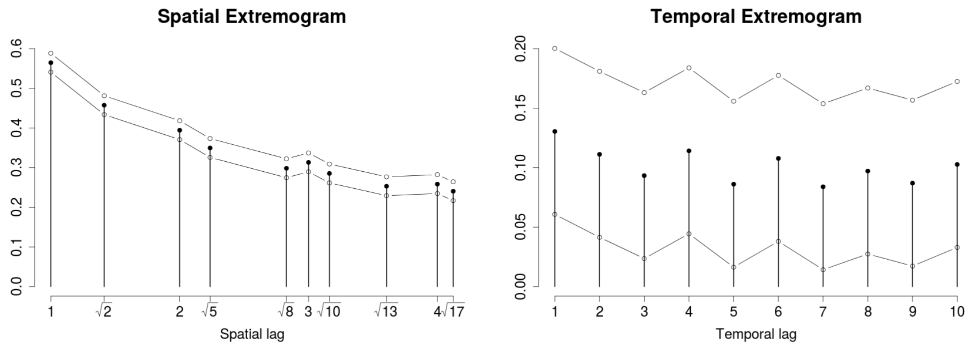

Figure 6.2: Empirical spatial (left) and temporal (right) extremogram based on spatial and temporal means for the space-time observations as given in (2.10) and (2.11) together with 95%−subsampling confidence intervals.

Further, for $i = 1,2$ denote by $\mathcal{F}_{\Lambda_i} = \sigma\{X(\boldsymbol{s}), \boldsymbol{s} \in \Lambda_i\}$ the σ-algebra generated by $\{X(\boldsymbol{s}) : \ \boldsymbol{s} \in \Lambda_i\}$.

(i) The α-*mixing coefficients* are defined for $k, l \in \mathbb{N} \cup \{\infty\}$ and $r \geq 0$ by

$$\alpha_{k,l}(r) = \sup\left\{|\mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1)\mathbb{P}(A_2)| : \ A_i \in \mathcal{F}_{\Lambda_i}, |\Lambda_1| \leq k, |\Lambda_2| \leq l, d(\Lambda_1, \Lambda_2) \geq r\right\}, \tag{A.1}$$

where $|\Lambda_i|$ is the cardinality of the set Λ_i for $i = 1, 2$.

(ii) The random field is called α-*mixing*, if $\alpha_{k,l}(r) \to 0$ as $r \to \infty$ for all $k, l \in \mathbb{N}$.

For a strictly stationary max-stable processes Corollary 2.2 of Dombry and Eyi-Minko [12] shows that the α-mixing coefficients can be related to the extremogram of the max-stable process. Equations (A.2) and (A.3) follow as in the proofs of Proposition 1 and 2 of Buhl and Klüppelberg [4].

***Proposition* A.2.** For all fixed time points $t \in \mathbb{N}$ the random field $\{\eta(\boldsymbol{s}, t), \boldsymbol{s} \in \mathbb{Z}^2\}$ (2.1) is α-mixing with mixing coefficients satisfying

$$\alpha_{k,l}(r) \leq 2kl \sup_{s \geq r} \chi(s, 0) \leq 4kle^{-\theta_1 r^{\alpha_1}/2}, \quad k, l \in \mathbb{N}, \ r \geq 0. \tag{A.2}$$

For all fixed locations $\boldsymbol{s} \in \mathbb{R}^2$ the time series $\{\eta(\boldsymbol{s}, t) : t \in [0, \infty)\}$ in (2.1) is α-mixing with mixing coefficients satisfying for some constant $c > 0$

$$\alpha(r) := \alpha_{\infty,\infty}(r) \leq c \sum_{u=r}^{\infty} e^{-\theta_2 u^{\alpha_2}/2}, \quad r \geq 0. \tag{A.3}$$

We will make frequent use of the following simple result.

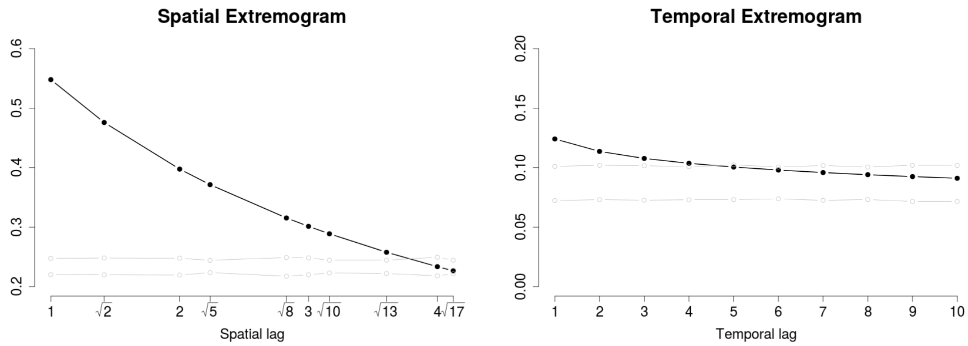

Figure 6.3: Permutation test for extremal independence: The gray lines show the 97.5%− and 2.5%−quantiles of the extremogram estimates for 1000 random space-time permutations for the empirical spatial (left) and the temporal (right) extremogram estimates.

***Lemma* A.3.** Let $z \in \mathbb{N}$. For $(\theta, \alpha) \in \{(\theta_1, \alpha_1), (\theta_2, \alpha_2)\}$ we have

$$g_z(r) = \sum_{u=r}^{\infty} u^z e^{-\theta u^\alpha/2} \leq c e^{-\theta r^\alpha/2} r^{z+1}, \quad r \in \mathbb{N}.$$

for some constant $c = c(z) > 0$.

Proof. An integral bound together with a change of variables yields

$$\begin{aligned}
g_z(r) = \sum_{u=r}^{\infty} u^z e^{-\theta u^\alpha/2} &\leq \int_r^\infty u^z e^{-\theta u^\alpha/2} du \\
&= \left(\frac{2}{\theta}\right)^{(z+1)/\alpha} \frac{1}{\alpha} \int_{\theta r^\alpha/2}^\infty t^{(z+1)/\alpha - 1} e^{-t} dt \\
&\leq c_1 \Gamma\left(\lceil (z+1)/\alpha \rceil, \theta r^\alpha/2\right) \\
&= c_1 \left(\lceil (z+1)/\alpha \rceil - 1\right)!\; e^{-\theta r^\alpha/2} \sum_{k=0}^{\lceil (z+1)/\alpha \rceil - 1} \frac{\theta^k r^{\alpha k}}{2^k k!} \\
&\leq c e^{-\theta r^\alpha/2} r^{\alpha(\lceil (z+1)/\alpha \rceil - 1)} \\
&\leq c e^{-\theta r^\alpha/2} r^{z+1},
\end{aligned}$$

where $\Gamma(s, r) = \int_r^\infty t^{s-1} e^{-t} dt = (s-1)! e^{-r} \sum_{k=0}^{s-1} r^k/k!$, $s \in \mathbb{N}$, is the incomplete gamma function and $c_1, c > 0$ are constants depending on z. $\square$

Acknowledgements

The three last authors gratefully acknowledge support by the TUM Institute for Advanced Study. Furthermore, all authors would like to thank Chin Man Bill Mok and Daniel Straub for their help in finding the data and discussions regarding the results. We further acknowledge the Southwest Florida Water Management District (SWFWMD) for providing the data. We also thank Dennis Leber for comparing simulation methods of the Brown-Resnick process, and we thank Ton Dieker and Marius Hofert for providing and improving the simulation programme. SB and CS also thank the International Graduate School of Science and Engineering (IGSSE) of the Technische Universität München for support. The research of RD was supported in part by the National Science Foundation grant DMS-1107031, and ARO MURI grant W11NF-12-1-0385.

References

[1] D. W. K. Andrews. Estimation when a parameter is on a boundary. *Econometrica*, 67(6):1341–1383, 1999.

[2] E. Bolthausen. On the central limit theorem for stationary mixing random fields. *Annals of Probability*, 10(4):1047–1050, 1982.

[3] B.M. Brown and S.I. Resnick. Extreme values of independent stochastic processes. *Journal of Applied Probability*, 14(4):732–739, 1977.

[4] S. Buhl and C. Klüppelberg. Anisotropic Brown-Resnick space-time processes: estimation and model assessment. *Extremes*, 19(4):627–660, 2016.

[5] S. Buhl and C. Klüppelberg. Limit theory for the empirical extremogram of random fields. Submitted. arXiv 1609.04961, 2016.

[6] Y. Cho, R.A. Davis, and S. Ghosh. Asymptotic properties of the spatial empirical extremogram. *Scandinavian Journal of Statistics*, 43(3):757–773, 2016.

[7] R.A. Davis and T. Mikosch. The extremogram: A correlogram for extreme events. *Bernoulli*, 15(4):977–1009, 2009.

[8] R.A. Davis, C. Klüppelberg, and C. Steinkohl. Max-stable processes for extremes of processes observed in space and time. *Journal of the Korean Statistical Society*, 42(3):399–414, 2013.

[9] R.A. Davis, C. Klüppelberg, and C. Steinkohl. Statistical inference for max-stable processes in space and time. *Journal of the Royal Statistical Society B*, 75(5):791–819, 2013.

[10] R.A. Davis, T. Mikosch, and Y. Zhao. Measures of serial extremal dependence and their estimation. *Stochastic Processes and Their Applications*, 123(7):2575–2602, 2013.

[11] L. de Haan. A spectral representation for max-stable processes. *Annals of Probability*, 12(4):1194–1204, 1984.

[12] C. Dombry and F. Eyi-Minko. Strong mixing properties of max-infinitely divisible random fields. *Stochastic Processes and their Applications*, 122(11):3790–3811, 2012.

[13] C. Dombry, S. Engelke, and M. Oesting. Exact simulation of max-stable processes. *Biometrika*, 103, 2016.

[14] P. Doukhan. *Mixing: Properties and Examples.* Springer, New York, 1994.

[15] S. Engelke, A. Malinowski, Z. Kabluchko, and M. Schlather. Estimation of Hüsler-Reiss distributions and Brown-Resnick processes. *Journal of the Royal Statistical Society B*, 77(1):239–265, 2015.

[16] V. Fasen, C. Klüppelberg, and M. Schlather. High-level dependence in time series models. *Extremes*, 13(1):1–33, 2010.

[17] M.G. Genton, Y. Ma, and H. Sang. On the likelihood function of Gaussian max-stable processes. *Biometrika*, 98(2):481–488, 2011.

[18] E. Giné, M. G. Hahn, and P. Vatan. Max-infinitely divisible and max-stable sample continuous processes. *Probab. Theory Rel.*, 87:139–165, 1990.

[19] H. Hult and F. Lindskog. Regular variation for measures on metric spaces. *Publications de l'Institut Mathématique (Beograd)*, 80:121–140, 2006.

[20] R. Huser and A. Davison. Space-time modelling of extreme events. *Journal of the Royal Statistical Society B*, 76(2):439–461, 2014.

[21] Z. Kabluchko, M. Schlather, and L. de Haan. Stationary max-stable fields associated to negative definite functions. *Annals of Probability*, 37(5):2042–2065, 2009.

[22] S. N. Lahiri, Y. Lee, and N. Cressie. On asymptotic distribution and asymptotic efficiency of least squares estimators of spatial variogram parameters. *J. Stat. Plan. Inf.*, 103(1):65–85, 2002.

[23] S.A. Padoan, M. Ribatet, and S.A. Sisson. Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, 105(489):263–277, 2010.

[24] D.N. Politis, J.P. Romano, and M. Wolf. *Subsampling.* Springer, New York, 1999.

[25] S.I. Resnick. *Heavy-Tail Phenomena, Probabilistic and Statistical Modeling.* Springer Series in Operations Research and Financial Engineering, 2007.

[26] M. Schlather. Randomfields, contributed package on random field simulation for R. http://cran.r-project.org/web/packages/RandomFields/.

[27] C. Steinkohl. *Statistical Modelling of Extremes in Space and Time using Max-Stable Processes.* Dissertation, Technische Universität München, München, 2013.

[28] E. Thibaud and T. Opitz. Efficient inference and simulation for elliptical Pareto processes. *Biometrika*, 102(4):855–870, 2015.

[29] J.L. Wadsworth and J.A. Tawn. Efficient inference for spatial extreme value processes associated to log-Gaussian random functions. *Biometrika*, 101(1):1–15, 2014.