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

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

Regularly varying space-time processes have proved useful 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 parametric models of extremal dependence functions. 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 and is robust to small departures from the underlying model. Finally, we apply this estimation procedure to fitting a max-stable process to radar rainfall measurements in a region in Florida. Complementary results and some proofs of key results are presented together with the simulation study in the supplement Buhl et al. [7].

1. Introduction

Regularly varying processes provide a useful framework for modeling extremal dependence in continuous time or space. They have been investigated in Hult and Lindskog [19, 20]. A prominent class of examples consists of max-stable processes. A key example in this paper is the max-stable Brown-Resnick process which was introduced in a time series framework in Brown and Resnick [2], in a spatial setting in Kabluchko et al. [22], and extended to a space-time setting in Davis et al. [10].

In the literature, various dependence 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. [24] for the spatial setting, and Huser and Davison [21] 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. [11]. 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 for regularly varying processes which is based on the extremogram as a natural extremal analog of the correlation function for stationary processes. The extremogram was introduced in Davis and Mikosch [9] 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. [8]. It can serve as a useful graphical tool for assessing extremal dependence structures in spatial and space-time processes that provides clues about potential parametric models, a critical step in the model building paradigm. For example, compatibility with various assumptions such as isotropy and stationarity (see Buhl and Klüppelberg [4] and Davis et al. [11] for some examples), can be assessed by examining invariance of the empirical extremogram when computed over specially chosen subsets of the data. Ultimately, a number of families of proposed parametric models are often fitted before deciding on a particular class of models. Therefore it is of interest to be able to not only have a procedure that can compute estimates rapidly, but also to serve as a check on the efficacy of model choice. Additionally, the new estimation procedure allows one to provide parameter estimates that can be used as initial values in more refined procedures, such as composite likelihood.

Our semiparametric estimation method assumes a spatially isotropic and additively separable dependence structure for regularly varying space-time processes. We first estimate the extremogram nonparametrically by its empirical version, where we can hence 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 empirical extremogram, and we apply the CLT with mixing conditions as provided in [5]. The rate of convergence can be improved by a bias correction term, a fact which we explain in detail. The proofs of the asymptotic properties of semiparametric spatial and temporal parameter estimates are analogous, and we present the details on the spatial parameters only, referring to Buhl [3], Chapter 3, for details about the asymptotic properties of the semiparametric temporal parameter.

In a second step we establish asymptotic normality of the weighted least squares parameter estimates. When the dependence parameters have bounded support, as for the Brown-Resnick process in Section 4, constrained optimization has to be applied. Then also the limit law differs depending whether the true parameters lie on the boundary or not. Since the asymptotic covariance matrix in the normal limit is difficult to access, we apply subsampling procedures to obtain pointwise confidence intervals for the parameters.

The semiparametric estimates converge at a slower rate than the square root rate of a fully parametric procedure such as pairwise likelihood estimation. However, it is known that likelihood-based estimates may be inefficient and even not consistent if the model is slightly misspecified. The semiparametric estimates, however, are often unaffected by slight deviations in the model. This is proved in Section 9 and illustrated in Section 10 of the supplement [7], where data are generated from a Brown-Resnick process, but with observational noise. The semiparametric estimates clearly outperform pairwise likelihood estimates in this case. On the other hand, the semiparametric estimates perform admirably well relative to the pairwise likelihood estimates when the underlying process is in fact a Brown-Resnick process.

Our paper is organized as follows. Section 2 defines regularly varying processes in space and time and their extremogram. Based on gridded data, the nonparametric extremogram estimation is derived and used for parametric model fitting. Asymptotic normality of the parameter estimates is established in Section 3. Section 3.1 is dedicated to the asymptotic normality of the empirical extremogram; and Section 3.2 deals with the asymptotic properties of the parameter estimates. The subsampling procedure – as well as results and proofs for our setting – is given in Section 7 of the supplement [7]. In Section 4 we apply the semiparametric method to the Brown-Resnick process and verify the required conditions. Here we also calculate the bias corrected estimator. We test our new semiparametric estimation procedure in a simulation study presented in the supplement [7] and compare it to pairwise likelihood estimation, both when applied to data generated by a Brown-Resnick process and when the data are affected by observational noise. In the latter, our procedure produces estimates with less bias than those based on pairwise likelihood (see Section 10 of the supplement [7]). The paper concludes with an analysis of daily rainfall maxima in a region in Florida in Section 5, where we also compare the semiparametric estimates with previously obtained pairwise likelihood estimates. The supplement [7] contains four sections, on subsampling, on α-mixing of the Brown-Resnick process, a robustness result for the bias corrected estimator, and a simulation study.

2. Model description and semiparametric estimates

In this paper we consider strictly stationary *regularly varying processes* in space and time $\{\eta(\boldsymbol{s},t) : \boldsymbol{s} \in \mathbb{R}^{d-1}, t \in [0,\infty)\}$ for $d \in \mathbb{N}$, where all finite-dimensional distributions are regularly varying (cf. Hult and Lindskog [20] for definitions and results in a general framework and Resnick [25] for details about multivariate regular variation). Throughout, $f(n) \sim g(n)$ means that $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$. As a prerequisite, we define for every finite set $\mathcal{I} \subset \mathbb{R}^{d-1} \times [0,\infty)$ with cardinality $|\mathcal{I}|$ the vector

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

Let furthermore $\|\cdot\|$ be the Euclidean norm on $\mathbb{R}^{d-1}$.

***Definition* 2.1** (Regularly varying stochastic process)**.** A strictly stationary stochastic space-time process $\{\eta(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^{d-1} \times [0,\infty)\}$ is called *regularly varying*, if there exists some normalizing sequence $0 < a_n \to \infty$ such that $\mathbb{P}(|\eta(\mathbf{0},0)| > a_n) \sim n^{-d}$ as $n \to \infty$, and if for every finite set $\mathcal{I} \subset \mathbb{R}^{d-1} \times [0,\infty)$,

$$n^d \mathbb{P}\Big(\frac{\eta_{\mathcal{I}}}{a_n} \in \cdot\Big) \xrightarrow{v} \mu_{\mathcal{I}}(\cdot), \quad n \to \infty, \tag{2.1}$$

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 $(\boldsymbol{s},t) \in \mathbb{R}^{d-1} \times [0,\infty)$ and $\mathcal{I} = \{(\boldsymbol{s},t)\}$ we set $\mu_{\{(\boldsymbol{s},t)\}}(\cdot) = \mu_{\{(\mathbf{0},0)\}}(\cdot) =: \mu(\cdot)$, which is justified by stationarity. Throughout we furthermore consider the space-time process $\{\eta(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^{d-1} \times [0,\infty)\}$ to be spatially isotropic. Together with the assumption of strict stationarity, this means that extremal dependence between two space-time points $(\boldsymbol{s}_1,t_1)$ and $(\boldsymbol{s}_2,t_2)$ is only driven by the spatial and temporal lags $v := \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|$ and $u := |t_1 - t_2|$, respectively, and we can define the extremogram only as a function of v and u. The extremogram was introduced for spatial and space-time processes by Buhl and Klüppelberg [5] and Cho et al. [8], based on Steinkohl [27], and can be regarded as a correlogram for extreme events.

***Definition* 2.2** (The extremogram)**.** For a regularly varying strictly stationary isotropic space-time process $\{\eta(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^{d-1} \times [0,\infty)\}$ 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

$$\rho_{AB}(v,u) = \lim_{n\to\infty} \frac{\mathbb{P}\left(\eta(\boldsymbol{s}_1,t_1)/a_n \in A, \eta(\boldsymbol{s}_2,t_2)/a_n \in B\right)}{\mathbb{P}\left(\eta(\boldsymbol{s}_1,t_1)/a_n \in A\right)}, \tag{2.2}$$

where $v = \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|$ and $u = |t_1 - t_2|$. Setting $A = B = (1,\infty)$, this reduces to the *tail dependence coefficient* $\chi(v,u) = \rho_{(1,\infty)(1,\infty)}(v,u)$.

In what follows we propose a two-step semiparametric estimation procedure of a parametric model of the extremogram. In particular, we assume that the model is additively separable such that setting either the temporal lag u or the spatial lag v equal to 0, it can be linearly parametrized as

$$T_1(\chi(v,0)) = T_1(\chi(v,0;C_1,\alpha_1)) = C_1 + \alpha_1 v, \quad (C_1,\alpha_1) \in \Theta_{\mathcal{S}}, \quad v \geq 0, \tag{2.3}$$

and

$$T_2(\chi(0,u)) = T_2(\chi(0,u;C_2,\alpha_2)) = C_2 + \alpha_2 u, \quad (C_2,\alpha_2) \in \Theta_{\mathcal{T}} \quad u \geq 0, \tag{2.4}$$

where T_1 and T_2 are known suitable strictly monotonous continuously differentiable transformations and the parameters (C_1, α_1) and (C_2, α_2) lie in appropriate parameter spaces $\Theta_{\mathcal{S}}$ and $\Theta_{\mathcal{T}}$. We refer to (C_1, α_1) as the *spatial parameter* and to (C_2, α_2) as the *temporal parameter*. Equations (2.3) and (2.4) are the basis for parameter estimates. We replace the extremogram on the left hand side in both of these equations by nonparametric estimates sampled at different lags. Then we use constrained weighted least squares estimation in a linear regression framework to obtain parameter estimates.

For better understanding, we stick to the 2-dimensional spatial case $d-1=2$; however, the method can directly be generalized and applied to higher dimensions. The estimation procedure is based on the following observation scheme for the space-time data.

***Condition* 2.3.** (1) The locations lie on a regular 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 generalization 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. [8], 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 $\mathcal{U}$ finite sets of spatial and temporal lags, on which the estimation is based. Concerning their choice, we generally include those lags which show clear extremal dependence between locations or time points. Larger lags should not be considered, since they may 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 clear extremal dependence are permutation tests, which we describe at the end of Section 5.

(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

$$\widehat{\chi}^{(t)}(v,0) = \frac{\dfrac{1}{|N(v)|}\displaystyle\sum_{i=1}^{n^2}\sum_{\substack{j=1 \\ \|\boldsymbol{s}_i-\boldsymbol{s}_j\|=v}}^{n^2} \mathbb{1}_{\{\eta(\boldsymbol{s}_i,t)>q,\eta(\boldsymbol{s}_j,t)>q\}}}{\dfrac{1}{n^2}\displaystyle\sum_{i=1}^{n^2}\mathbb{1}_{\{\eta(\boldsymbol{s}_i,t)>q\}}}, \quad v \in \mathcal{V}, \tag{2.5}$$

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

$$\widehat{\chi}^{(\boldsymbol{s})}(0,u) = \frac{\frac{1}{T-u}\sum_{k=1}^{T-u} \mathbb{1}_{\{\eta(\boldsymbol{s},t_k)>q,\eta(\boldsymbol{s},t_k+u)>q\}}}{\frac{1}{T}\sum_{k=1}^{T} \mathbb{1}_{\{\eta(\boldsymbol{s},t_k)>q\}}}, \quad u \in \mathcal{U}, \tag{2.6}$$

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.,

$$\widehat{\chi}(v,0) = \frac{1}{T}\sum_{k=1}^{T} \widehat{\chi}^{(t_k)}(v,0), \quad v \in \mathcal{V}, \tag{2.7}$$

$$\widehat{\chi}(0,u) = \frac{1}{n^2}\sum_{i=1}^{n^2} \widehat{\chi}^{(\boldsymbol{s}_i)}(0,u), \quad u \in \mathcal{U}. \tag{2.8}$$

(3) Parameter estimates for C_1, α_1, C_2 and α_2 are found by using weighted least squares estimation:

$$\begin{pmatrix}\widehat{C}_1\\ \widehat{\alpha}_1\end{pmatrix} = \underset{(C_1,\alpha_1)\in\Theta_{\mathcal{S}}}{\arg\min} \sum_{v\in\mathcal{V}} w_v \Big(T_1(\widehat{\chi}(v,0)) - \big(C_1+\alpha_1 v\big)\Big)^2, \tag{2.9}$$

$$\begin{pmatrix}\widehat{C}_2\\ \widehat{\alpha}_2\end{pmatrix} = \underset{(C_2,\alpha_2)\in\Theta_{\mathcal{T}}}{\arg\min} \sum_{u\in\mathcal{U}} w_u \Big(T_2(\widehat{\chi}(0,u)) - \big(C_2+\alpha_2 u\big)\Big)^2, \tag{2.10}$$

with weights $w_u > 0$ and $w_v > 0$.

We call the estimates $(\widehat{C}_1, \widehat{\alpha}_1)$ and $(\widehat{C}_2, \widehat{\alpha}_2)$ *weighted least squares estimates* (WLSE). This approach bears similarity with that proposed by Einmahl et al. [14], who suggest semiparametric weighted least squares estimation of the parameters of parametric models of the stable tail dependence function based on iid random vector observations.

3. Asymptotic properties of the WLSE

In this section we investigate aymptotic properties of the WLSE $(\widehat{C}_1, \widehat{\alpha}_1)$ and $(\widehat{C}_2, \widehat{\alpha}_2)$. Recall from (2.9) and (2.10) that they are functions of the averaged empirical extremogram $\widehat{\chi}(\cdot,\cdot)$. Its definition is given in (2.7) and (2.8) and implies that we first need CLTs of the pointwise empirical extremograms $\widehat{\chi}^{(t)}$ 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. The corresponding results for the temporal case can be derived similarly by replacing n with $\sqrt{T}$ and can be found with full details in Buhl [3], Chapter 3 for the Brown-Resnick space-time process. We use several results for the extremogram provided in Section 8 of the supplement [7] and in Buhl and Klüppelberg [5].

3.1. Asymptotics of the empirical spatial extremogram

We show a CLT for the empirical spatial extremogram of regularly varying space-time processes, which is defined in (2.1) and based on a finite set of observed spatial lags

$$\mathcal{V} = \{v_1, \ldots, v_p\},$$

which show clear extremal dependence as explained in Section 2. First we state conditions under which the empirical extremogram centred by the pre-asymptotic version is asymptotically normal.

Theorem 3.1. *For a fixed time point $t \in \{t_1, \ldots, t_T\}$, consider a regularly varying spatial process $\left\{\eta(\boldsymbol{s}, t) : \boldsymbol{s} \in \mathbb{R}^2\right\}$ as defined in Definition 2.1. Let a_n be a sequence as in (2.1). Assume that there exists $\gamma > 0$ that satisfies $\max\{v_1, \ldots, v_p\} \leq \gamma$, such that the following conditions are satisfied:*

(M1) $\{\eta(\boldsymbol{s}, t) : \boldsymbol{s} \in \mathbb{R}^2\}$ *is α-mixing with α-mixing coefficients $\alpha_{k,\ell}(\cdot)$.*

There exist sequences $m = m_n, r = r_n \to \infty$ with $m_n/n \to 0$ and $r_n/m_n \to 0$ as $n \to \infty$ such that the following hold:

(M2) $m_n^2 r_n^2 / n \to 0$.

(M3) For all $\epsilon > 0$:

$$\lim_{k\to\infty} \limsup_{n\to\infty} \sum_{\boldsymbol{h}\in\mathbb{Z}^2 : k < \|\boldsymbol{h}\| \leq r_n} m_n^2 \\ \mathbb{P}\Big(\max_{\boldsymbol{s}\in B(\boldsymbol{0},\gamma)} |\eta(\boldsymbol{s}, t)| > \epsilon a_m, \max_{\boldsymbol{s}'\in B(\boldsymbol{h},\gamma)} |\eta(\boldsymbol{s}', t)| > \epsilon a_m \Big) = 0,$$

where $B(\boldsymbol{h}, \gamma) := \{\boldsymbol{s} \in \mathbb{Z}^2 : \|\boldsymbol{s} - \boldsymbol{h}\| \leq \gamma\}$ for $\boldsymbol{h} \in \mathbb{R}^2$.

(M4) *(i)* $\lim_{n\to\infty} m_n^2 \sum_{\boldsymbol{h}\in\mathbb{Z}^2 : \|\boldsymbol{h}\| > r_n} \alpha_{1,1}(\|\boldsymbol{h}\|) = 0,$

(ii) $\sum_{\boldsymbol{h}\in\mathbb{Z}^2} \alpha_{p,q}(\|\boldsymbol{h}\|) < \infty$ *for* $2 \leq p + q \leq 4,$

(iii) $\lim_{n\to\infty} m_n n \; \alpha_{1,n^2}(r_n) = 0,$

Then the empirical spatial extremogram $\widehat{\chi}^{(t)}(v, 0)$ defined in (2.5) with the quantile $q = a_m$ satisfies

$$\frac{n}{m_n} \big(\widehat{\chi}^{(t)}(v, 0) - \chi_n(v, 0)\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \Pi_1^{(\mathrm{iso})}), \quad n \to \infty, \tag{3.1}$$

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

$$\chi_n(v,0) = \frac{\mathbb{P}(\eta(\mathbf{0},0) > a_m, \eta(\boldsymbol{h},0) > a_m)}{\mathbb{P}(\eta(\mathbf{0},0) > a_m)}, \quad v = \|\boldsymbol{h}\| \in \mathcal{V}. \tag{3.2}$$

Proof. Theorem 3.1 is a direct application of Theorem 4.2 of Buhl and Klüppelberg [5] to the process $\{\eta(\boldsymbol{s},t) : \boldsymbol{s} \in \mathbb{R}^2\}$ for $d = 2$ and $A = B = (1, \infty)$. For the specification of the asymptotic covariance matrix we need to adapt that theorem to 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 $i \in \{1, \ldots, p\}$ such that $v_i \in \mathcal{V}$, we summarize these into

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

where $\ell_i := |L(v_i)|$. We conclude that

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

where $\Pi_1^{(\mathrm{space})}$ is specified in equation (4.3)-(4.6) of [5]. Note the slight misuse of notation committed here for the sake of simplicity: by $\widehat{\chi}^{(t)}(\boldsymbol{h},0)$ (instead of $\widehat{\chi}^{(t)}(v,0)$) we denote the empirical extremogram for each single vector $\boldsymbol{h} \in L(v_i)$ specified above; i.e.,

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

where $N(\boldsymbol{h}) := \{(i,j) \in \{1, \ldots, n^2\} : \boldsymbol{s}_i - \boldsymbol{s}_j = \boldsymbol{h}\}$ (instead of $N(v)$). Analogously we define the pre-asymptotic extremogram $\chi_n(\boldsymbol{h},0)$ w.r.t. a vector $\boldsymbol{h}$.

It holds that $|N(v_i)| = \sum_{\boldsymbol{h}\in L(v_i)} |N(\boldsymbol{h})|$. Isotropy implies furthermore 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 (2.5),

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

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

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

such that

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

□

***Corollary* 3.2.** Under the conditions of Theorem 3.1 the averaged spatial extremogram in (2.7) satisfies (with covariance matrix $\Pi_2^{(\mathrm{iso})}$ specified in (3.11) below)

$$\frac{n}{m_n}\left(\frac{1}{T}\sum_{k=1}^{T} \widehat{\chi}^{(t_k)}(v,0) - \chi_n(v,0)\right)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Pi_2^{(\mathrm{iso})}), \quad n \to \infty. \tag{3.7}$$

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 [9]. We use the notation of the proof of Theorem 3.1. 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 $\mu_{B_T(\mathbf{0},\gamma)}$-continuous Borel sets C and D in $\overline{\mathbb{R}}^{T|B(\mathbf{0},\gamma)|}\backslash\{\mathbf{0}\}$, regular variation yields the existence of the limit measures

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

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

$$\widehat{\chi}^{(t_k)}(\boldsymbol{h}_j^{(i)},0) \sim \widehat{R}_{m_n}(D_{j,k}^{(i)},D_k) := \widehat{\mu}_{B_T(\mathbf{0},\gamma),m_n}(D_{j,k}^{(i)})/\widehat{\mu}_{B_T(\mathbf{0},\gamma),m_n}(D_k), \quad n\to\infty, \tag{3.8}$$

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

$$\widehat{\mu}_{B_T(\mathbf{0},\gamma),m_n}(\cdot) := \Big(\frac{m_n}{n}\Big)^2 \sum_{\boldsymbol{s}\in\mathcal{S}_n} \mathbb{1}_{\{\frac{\boldsymbol{Y}(\boldsymbol{s})}{m_n^2}\in\cdot\}}. \tag{3.9}$$

Likewise we have for the pre-asymptotic quantities

$$\chi_n(\boldsymbol{h}_j^{(i)},0) = R_{m_n}(D_{j,k}^{(i)},D_k) := \frac{\mathbb{P}(\boldsymbol{Y}(\mathbf{0})/m_n^2 \in D_{j,k}^{(i)})}{\mathbb{P}(\boldsymbol{Y}(\mathbf{0})/m_n^2 \in D_k)} =: \frac{\mu_{B_T(\mathbf{0},\gamma),m_n}(D_{j,k}^{(i)})}{\mu_{B_T(\mathbf{0},\gamma),m_n}(D_k)}, \tag{3.10}$$

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

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

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 (3.8) and the corresponding pre-asymptotic quantities from (3.10) exactly in the same way. Furthermore, define for $k = 1, \ldots, T$ the vectors in $\mathbb{R}^{L+1}$

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

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

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

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 $n \to \infty$,

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

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

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

we thus conclude that

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

To obtain the asymptotic covariance matrix in the spatially isotropic case, we proceed as in the proof of Theorem 3.1. 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

$$\frac{n}{m_n}\begin{pmatrix} \widehat{\chi}^{(t_1)}(v_1, 0) - \chi_n(v_1, 0) \\ \vdots \\ \widehat{\chi}^{(t_T)}(v_p, 0) - \chi_n(v_p, 0) \end{pmatrix} = \frac{n}{m_n}\boldsymbol{N}\begin{pmatrix} \widehat{\chi}^{(t_1)}(\boldsymbol{h}^{(1)}_1, 0) - \chi_n(\boldsymbol{h}^{(1)}_1, 0) \\ \vdots \\ \widehat{\chi}^{(t_T)}(\boldsymbol{h}^{(p)}_{\ell_p}, 0) - \chi_n(\boldsymbol{h}^{(p)}_{\ell_p}, 0) \end{pmatrix}$$

$$\xrightarrow{d} \quad \mathcal{N}(\mathbf{0}, \mu(A)^{-4} \boldsymbol{NF} \Sigma (\boldsymbol{NF})^\top), \quad n \to \infty,$$

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

$$\begin{aligned}\Pi_2^{(\mathrm{iso})} =& \mu(A)^{-4} T^{-2} \begin{pmatrix} 1 & 0\cdots 0 & 1 & 0\cdots 0 & \cdots & 1 & 0\cdots 0 \\ 0 & 1\cdots 0 & 0 & 1\cdots 0 & \cdots & 0 & 1\cdots 0 \\ & & & \ddots & & & \\ 0 & 0\cdots 1 & 0 & 0\cdots 1 & \cdots & 0 & 0\cdots 1 \end{pmatrix} \boldsymbol{NF}\Sigma(\boldsymbol{NF})^\top \\ & \begin{pmatrix} 1 & 0\cdots 0 & 1 & 0\cdots 0 & \cdots & 1 & 0\cdots 0 \\ 0 & 1\cdots 0 & 0 & 1\cdots 0 & \cdots & 0 & 1\cdots 0 \\ & & & \ddots & & & \\ 0 & 0\cdots 1 & 0 & 0\cdots 1 & \cdots & 0 & 0\cdots 1 \end{pmatrix}^\top . \end{aligned} \tag{3.11}$$

□

***Condition* 3.3.** In the CLTs (3.1) and (3.7), the pre-asymptotic extremogram (3.2) can be replaced by the theoretical one (eq. (2.2) with $A = B = (1, \infty)$), provided that

$$\frac{n}{m_n} \left(\chi_n(v,0) - \chi(v,0) \right) \to 0, \quad n \to \infty, \tag{3.12}$$

is satisfied for all spatial lags $v \in \mathcal{V}$. In particular, we then obtain

$$\frac{n}{m_n} \big(\widehat{\chi}(v,0) - \chi(v,0) \big)_{v \in \mathcal{V}} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Pi_2^{(\mathrm{iso})}), \quad n \to \infty. \tag{3.13}$$

This bias condition turns out to be central in order to obtain a CLT for the WLSE $(\widehat{C}_1, \widehat{\alpha}_1)$ in Section 3.2 below. However, even if it is not satisfied, the empirical extremogram keeps its important asymptotic interpretation as a conditional probability of extremal events. Furthermore there are cases where we can resort to a bias correction, ensuring again a CLT for $(\widehat{C}_1, \widehat{\alpha}_1)$. For examples we refer to Section 4 below.

3.2. Asymptotic properties of spatial parameter estimates

In this section we state conditions that yield asymptotic normality of the WLSE $(\widehat{C}_1, \widehat{\alpha}_1)$ of Section 2. Recall the weighted least squares optimization problem (2.9); i.e.,

$$\begin{pmatrix} \widehat{C}_1 \\ \widehat{\alpha}_1 \end{pmatrix} = \underset{(C_1, \alpha_1) \in \Theta_{\mathcal{S}}}{\arg\min} \sum_{v \in \mathcal{V}} w_v \Big(T_1(\widehat{\chi}(v,0)) - \big(C_1 + \alpha_1 v \big) \Big)^2 .$$

To show asymptotic normality of the WLSE, we define the design matrix X and weight matrix W as

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

respectively, where $\mathbf{1} = (1,\dots,1)^\intercal \in \mathbb{R}^p$. If neither C_1 nor α_1 have bounded support, then the WLSE; i.e., the solution to (2.9), is given by

$$\widehat{\psi}_1 := \begin{pmatrix} \widehat{C}_1 \\ \widehat{\alpha}_1 \end{pmatrix} = (X^\intercal W X)^{-1} X^\intercal W (T_1(\widehat{\chi}(v,0)))^\intercal_{v\in\mathcal{V}}.$$

If one of the parameters C_1 or α_1 does have bounded support, we need to constrain $\widehat{\psi}_1$ properly, obtaining a CLT that might differ considerably from that given in Theorem 3.4 below. An important example of this is treated in Section 4.

Theorem 3.4. *For a fixed time point $t \in \{t_1,\dots,t_T\}$, consider a regularly varying spatial process $\left\{\eta(\boldsymbol{s},t) : \boldsymbol{s} \in \mathbb{R}^2\right\}$ as defined in Definition 2.1. Assume that it satisfies the conditions of Theorem 3.1. Let $\widehat{\boldsymbol{\psi}}_1 = (\widehat{C}_1, \widehat{\alpha}_1)^\intercal$ denote the WLSE resulting from the minimization problem (2.9) and $\boldsymbol{\psi}_1^* = (C_1^*, \alpha_1^*)^\intercal \in \Theta_{\mathcal{S}}$ the true parameter vector. Assume that the CLT (3.13) holds, possibly after a bias correction of the empirical extremogram $(\widehat{\chi}_v : v \in \mathcal{V})$. Then for a suitably chosen scaling sequence m_n, we obtain, as $n \to \infty$,*

$$\frac{n}{m_n}\left(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, Q_x^{(w)} G \Pi_2^{(\mathrm{iso})} G Q_x^{(w)\intercal}). \tag{3.14}$$

Here $\Pi_2^{(\mathrm{iso})}$ is the covariance matrix given in (3.11),

$$Q_x^{(w)} = (X^\intercal W X)^{-1} X^\intercal W \quad \text{and} \quad G = \operatorname{diag}\left\{T_1'(\chi(v,0)) :\ v \in \mathcal{V}\right\}, \tag{3.15}$$

where $T_1'(x)$ denotes the derivative of $T_1(x)$ with respect to x for $0 < x < 1$.

Proof. Using the multivariate delta method together with the CLT (3.13) it directly follows that

$$\frac{n}{m_n}\big(T_1(\widehat{\chi}(v,0)) - T_1(\chi(v,0))\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(\mathbf{0}, G\Pi_2^{(\mathrm{iso})}G), \quad n \to \infty,$$

where G is defined in (3.15). Since

$$\min_{(C_1,\alpha_1)\in\Theta_{\mathcal{S}}} \sum_{v\in\mathcal{V}} w_v\big(T_1(\chi(v,0)) - (C_1 + \alpha_1 v)\big)^2 = \sum_{v\in\mathcal{V}} w_v\big(T_1(\chi(v,0)) - (C_1^* + \alpha_1^* v)\big)^2,$$

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

$$Q_x^{(w)}(T_1(\chi(v,0)))^\intercal_{v\in\mathcal{V}} = \operatorname*{arg\,min}_{(C_1,\alpha_1)\in\Theta_{\mathcal{S}}} \sum_{v\in\mathcal{V}} w_v\big(T_1(\chi(v,0)) - \big(\log(\theta_1) + \alpha_1 x_v\big)\big)^2 = \boldsymbol{\psi}_1^*.$$

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

$$\frac{n}{m_n}\left(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*\right) = \frac{n}{m_n} Q_x^{(w)}\big(T_1(\widehat{\chi}(v,0)) - T_1(\chi(v,0))\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, Q_x^{(w)} G \Pi_2^{(\mathrm{iso})} G Q_x^{(w)\intercal}\right).$$

□

4. Example: the Brown-Resnick process

We illustrate the results of the previous sections by applying them to a max-stable strictly stationary and isotropic Brown-Resnick space-time process with representation

$$\eta(\boldsymbol{s},t)=\bigvee_{j=1}^{\infty}\left\{\xi_j\, e^{W_j(\boldsymbol{s},t)-\delta(\|\boldsymbol{s}\|,t)}\right\},\quad (\boldsymbol{s},t)\in\mathbb{R}^2\times[0,\infty),\tag{4.1}$$

where $\{\xi_j : j\in\mathbb{N}\}$ 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 (4.1) goes back to de Haan [12], Giné et al. [18] and Kabluchko et al. [22]. 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 characterized 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.

We assume the dependence function δ to be given for $v,u\ge 0$ by

$$\delta(v,u)=2\theta_1v^{\alpha_1}+2\theta_2u^{\alpha_2},\tag{4.2}$$

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

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

$$- \frac{1}{x_2}\Phi\left(\frac{\log(x_1/x_2)}{\sqrt{2\delta(\|\boldsymbol{h}\|,|u|)}} + \sqrt{\frac{\delta(\|\boldsymbol{h}\|,|u|)}{2}}\right)\Bigg\}, \tag{4.3}$$

where Φ denotes the standard normal distribution function (cf. Davis et al. [10]).

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 (4.3), the pairwise density can be derived and pairwise likelihood methods can be used to estimate the parameters; cf. Davis et al. [11], Huser and Davison [21] and Padoan et al. [24]. Full likelihood inference is virtually intractable 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 apply the estimation method introduced in Section 2 based on the extremogram of more general regularly varying processes to the special case of the Brown-Resnick process (4.1). We make use of the fact that its extremogram possesses a closed-form expression which is characterized by the dependence function δ.

***Lemma* 4.1** (Davis et al. [10], equation (3.1))**.** Let $\{\eta(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^2 \times [0,\infty)\}$ be the strictly stationary isotropic Brown-Resnick process in $\mathbb{R}^2 \times [0,\infty)$ as defined in (4.1) with dependence function given in (4.2). Then the extremogram of η is given by

$$\chi(v,u) = 2\Big(1 - \Phi\Big(\sqrt{\frac{1}{2}\delta(v,u)}\Big)\Big) = 2\big(1 - \Phi(\sqrt{\theta_1 v^{\alpha_1} + \theta_2 u^{\alpha_2}})\big), \quad v,u \geq 0. \tag{4.4}$$

Solving equation (4.4) for $\delta(v,u)$ leads to

$$\frac{\delta(v,u)}{2} = \theta_1 v^{\alpha_1} + \theta_2 u^{\alpha_2} = \Big(\Phi^{-1}\big(1 - \frac{1}{2}\chi(v,u)\big)\Big)^2. \tag{4.5}$$

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

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

In the same way, we obtain

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

To put this in the context of equations (2.3) and (2.4), first note that in the weighted linear regression, instead of working with the "original" lags v and u, we consider their log transformations $x_v = \log(v)$ and $x_u = \log(u)$; hence in particular, we need to exclude the lags $v = 0$ and $u = 0$. The observation scheme described in Condition 2.3 then yields that $u, v \geq 1$ and thus $x_v, x_u \geq 0$. We furthermore set $C_1 = \log(\theta_1)$, $C_2 = \log(\theta_2)$ and choose the transformations T_1 and T_2 defined by $T_1(\chi(v,0)) = 2\log\big(\Phi^{-1}\big(1 - \frac{1}{2}\chi(v,0)\big)\big)$ and $T_2(\chi(0,u)) = 2\log\big(\Phi^{-1}\big(1 - \frac{1}{2}\chi(0,u)\big)\big)$. The parameter spaces are given by $\Theta_{\mathcal{S}} = \Theta_{\mathcal{T}} = \mathbb{R} \times (0,2]$.

In the following we work out necessary and sufficient conditions for the Brown-Resnick process (4.1) with dependence function (4.2) to satisfy the conditions of Theorem 3.4, focusing again on the spatial case; i.e., on the processes $\{\eta(\boldsymbol{s},t) : \boldsymbol{s} \in \mathbb{R}^2\}$ for fixed observed $t \in \{t_1, \dots, t_T\}$. Furthermore we show how the fact that the model parameter $\alpha_1 \in (0,2]$ has bounded support influences the asymptotics of the WLSE $(\widehat{\theta}_1, \widehat{\alpha}_1)$.

4.1. Asymptotics of the empirical spatial extremogram of the Brown-Resnick process

For a start, we need a sufficiently precise estimate for the extremogram (4.4) of the Brown-Resnick process, which we give now.

***Lemma* 4.2.** Let $\boldsymbol{s}, \boldsymbol{h} \in \mathbb{R}^2$.For every sequence $a_n \to \infty$ we have for fixed $t \in [0,\infty)$,

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

Lemma 4.2 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. We can choose in the following $a_n = n^2$ in order to satisfy the condition $\mathbb{P}(|\eta(\boldsymbol{0},0)| > a_n) \sim n^{-2}$ as $n \to \infty$ from Definition 2.1. Recall furthermore that we have to choose a finite set $\mathcal{V} = \{v_1, \dots, v_p\}$ of observed lags, which show clear extremal dependence as explained in Section 2.

Theorem 4.3. *Consider the spatial Brown-Resnick process $\{\eta(\boldsymbol{s},t) : \boldsymbol{s} \in \mathbb{R}^2\}$ as defined in (4.1) with dependence function given in (4.2). 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 (2.5) with the quantile $q = a_{m_n} = m_n^2$ satisfies*

$$\frac{n}{m_n}\big(\widehat{\chi}^{(t)}(v,0) - \chi_n(v,0)\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \Pi_1^{(\mathrm{iso})}), \quad n \to \infty, \tag{4.6}$$

where the covariance matrix $\Pi_1^{(\mathrm{iso})}$ is specified in equation (3.6), and χ_n is the pre-asymptotic spatial extremogram as in (3.2).

Furthermore, for the averaged empirical extremogram $\widetilde{\chi}(v,0) = T^{-1}\sum_{k=1}^{T}\widetilde{\chi}^{(t_k)}(v,0)$ defined in (2.7), we obtain (with covariance matrix $\Pi_2^{(\mathrm{iso})}$ given in equation (3.11))

$$\frac{n}{m_n}\big(\widehat{\chi}(v,0) - \chi_n(v,0)\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \Pi_2^{(\mathrm{iso})}), \quad n \to \infty. \tag{4.7}$$

Proof. We need to verify the conditions of Corollary 3.2; i.e., conditions (M1)-(M4) of Theorem 3.1 for $a_{m_n} = m_n^2$, and apply results of Section 8 of the supplement [7].

Condition (M1) is satisfied by equation (8.2).

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 $B(\mathbf{0},\gamma) = \{\boldsymbol{s} \in \mathbb{Z}^2 : \|\boldsymbol{s}\| \le \gamma\}$. For $\epsilon > 0$, 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,

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

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

$$
\begin{aligned}
&\mathbb{P}\big(\max_{\boldsymbol{s}\in B(\mathbf{0},\gamma)} \eta(\boldsymbol{s},t) > \epsilon m_n^2, \max_{\boldsymbol{s}'\in B(\boldsymbol{h},\gamma)} \eta(\boldsymbol{s}',t) > \epsilon m_n^2\big)\\
&\le \sum_{\boldsymbol{s}\in B(\mathbf{0},\gamma)} \sum_{\boldsymbol{s}'\in B(\boldsymbol{h},\gamma)} \mathbb{P}(\eta(\boldsymbol{s},t) > \epsilon m_n^2, \eta(\boldsymbol{s}',t) > \epsilon m_n^2)\\
&= \sum_{\boldsymbol{s}\in B(\mathbf{0},\gamma)} \sum_{\boldsymbol{s}'\in B(\boldsymbol{h},\gamma)} \Big\{\frac{1}{\epsilon m_n^2}\chi(\|\boldsymbol{s}-\boldsymbol{s}'\|,0) + \mathcal{O}\Big(\frac{1}{m_n^4}\Big)\Big\}\\
&\le \frac{2|B(\mathbf{0},\gamma)|^2}{\epsilon m_n^2}\big(1 - \Phi(\sqrt{\theta_1(\|\boldsymbol{h}\| - 2\gamma)^{\alpha_1}}\big) + \mathcal{O}\Big(\frac{1}{m_n^4}\Big),
\end{aligned} \tag{4.8}
$$

as $n \to \infty$, where we have used (4.4). 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 $k \ge 2\gamma$,

$$
\begin{aligned}
L_{m_n} &:= \limsup_{n\to\infty} m_n^2 \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^2\\ k<\|\boldsymbol{h}\|\le r_n}} \mathbb{P}\Big(\max_{\boldsymbol{s}\in B(\mathbf{0},\gamma)} \eta(\boldsymbol{s},t) > \epsilon m_n^2, \max_{\boldsymbol{s}'\in B(\boldsymbol{h},\gamma)} \eta(\boldsymbol{s}',t) > \epsilon m_n^2\Big)\\
&\le 2|B(\mathbf{0},\gamma)|^2 \limsup_{n\to\infty} \Big\{ \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^2\\ k<\|\boldsymbol{h}\|\le r_n}} \Big\{\frac{1}{\epsilon}(1 - \Phi(\sqrt{\theta_1(\|\boldsymbol{h}\| - 2\gamma)^{\alpha_1}}))\Big\} + \mathcal{O}\Big(\Big(\frac{r_n}{m_n}\Big)^2\Big)\Big\}\\
&\le K_1 \limsup_{n\to\infty} \sum_{\substack{v\in V:\\ k<v\le r_n}} \Big\{\frac{v}{\epsilon}2\big(1 - \Phi(\sqrt{\theta_1(v - 2\gamma)^{\alpha_1}})\big)\Big\},
\end{aligned}
$$

for some constant $K_1 > 0$. For the term $\mathcal{O}((r_n/m_n)^2)$ we use that $r_n/m_n \to 0$. From Lemma 8.3 and the fact that $1-\Phi(x) \le \exp\{-x^2/2\}$ for $x > 0$, we find for $K_2 > 0$,

$$L_{m_n} \le K_2 k^2 \exp\Big\{-\frac{1}{2}\theta_1 (k-2\gamma)^{\alpha_1}\Big\}.$$

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

$$m_n^2 \sum_{\boldsymbol{h}\in\mathbb{Z}^2:\|\boldsymbol{h}\|>r_n} \alpha_{1,1}(\|\boldsymbol{h}\|) \le K_1 m_n^2 \sum_{v\in V: v>r_n} v\,\alpha_{1,1}(v) \le 4K_1 m_n^2 \sum_{v\in V: v>r_n} v\, e^{-\theta_1 v^{\alpha_1}/2}.$$

By Lemma 8.3 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 (8.2), since

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

□

***Remark* 4.1.** We want to examine for which choices of β_1, introduced with the sequence $m_n = n^{\beta_1}$ in Theorem 4.3, we can replace the pre-asymptotic extremogram by the theoretical one in the CLTs (4.6) and (4.7); that is, the bias condition (3.12),

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

is satisfied for all spatial lags $v \in \mathcal{V}$. For the Brown-Resnick process (4.1) we obtain from Lemma 4.2,

$$\begin{aligned}
&\frac{n}{m_n}\,(\chi_n(v,0) - \chi(v,0)) \\
&= \frac{n}{m_n}\left(\frac{\mathbb{P}(\eta(\boldsymbol{s},t) > m_n^2, \eta(\boldsymbol{s}+\boldsymbol{h},t) > m_n^2)}{\mathbb{P}(\eta(\boldsymbol{s},t) > m_n^2)} - \chi(v,0)\right) \\
&\sim \frac{n}{2m_n^3}\big(\chi(v,0)-2\big)\big(\chi(v,0)-1\big) \\
&= n^{1-3\beta_1}\frac{1}{2}\big(\chi(v,0)-2\big)\big(\chi(v,0)-1\big) \quad \to \quad 0 \quad \text{ if and only if } \beta_1 > 1/3;
\end{aligned}$$

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 (4.12) below.

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

$$n^{1-\beta_1}\big(\widehat{\chi}^{(t)}(v,0)-\chi(v,0)\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(\mathbf{0},\Pi_1^{(\mathrm{iso})}), \quad n\to\infty, \tag{4.9}$$

and likewise for the averaged empirical extremogram,

$$n^{1-\beta_1}\big(\widehat{\chi}(v,0)-\chi(v,0)\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(\mathbf{0},\Pi_2^{(\mathrm{iso})}), \quad n\to\infty. \tag{4.10}$$

□

We now turn to the bias correction needed in case (I). By Lemma 4.2 the pre-asymptotic extremogram has representation

$$\begin{aligned}\chi_n(v,0) &= \chi(v,0) + \Big[\frac{1}{2m_n^2}\big(\chi(v,0)-2\big)\big(\chi(v,0)-1\big)\Big](1+o(1))\\ &= \chi(v,0)+\frac{1}{2m_n^2}\nu(v,0)(1+o(1)), \quad n\to\infty,\end{aligned} \tag{4.11}$$

where $\nu(v,0) := \big(\chi(v,0)-2\big)\big(\chi(v,0)-1\big)$. Consequently, we propose for fixed $t\in\{t_1,\ldots,t_T\}$ 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

$$\widetilde{\chi}^{(t)}(v,0) := \begin{cases}\widehat{\chi}^{(t)}(v,0)-\dfrac{1}{2m_n^2}\widehat{\nu}^{(t)}(v,0) & \text{if } m_n=n^{\beta_1} \text{ with } \beta_1\in(\frac{1}{5},\frac{1}{3}],\\ \widehat{\chi}^{(t)}(v,0) & \text{if } m_n=n^{\beta_1} \text{ with } \beta_1\in(\frac{1}{3},\frac{1}{2}).\end{cases} \tag{4.12}$$

Theorem 4.4 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 4.4. *For a fixed time point $t\in\{t_1,\ldots,t_T\}$ consider the spatial Brown-Resnick process $\big\{\eta(\boldsymbol{s},t),\boldsymbol{s}\in\mathbb{R}^2\big\}$ defined in (4.1) with dependence function given in (4.2). 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 (4.12) satisfies*

$$\frac{n}{m_n}\big(\widetilde{\chi}^{(t)}(v,0)-\chi(v,0)\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}(\mathbf{0},\Pi_1^{(\mathrm{iso})}), \quad n\to\infty, \tag{4.13}$$

where $\Pi_1^{(\mathrm{iso})}$ is the covariance matrix as given in equation (3.6). Furthermore, the corresponding bias corrected averaged version $\widetilde{\chi}(v,0)=T^{-1}\sum_{k=1}^{T}\widetilde{\chi}^{(t_k)}(v,0)$ satisfies

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

with covariance matrix $\Pi_2^{(\mathrm{iso})}$ specified in (3.11).

Proof. For simplicity we suppress the time point t in the notation. By (4.11) and (4.12) 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 4.3 it suffices to show that $(n/(2m_n^3))(\widehat{\nu}(v,0) - \nu(v,0)) \xrightarrow{P} 0$. 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 4.3 weakly to a normal distribution, and the second term, together with the fact that $\widehat{\chi}(v,0) \xrightarrow{P} \chi(v,0)$ and $\chi_n(v,0) \xrightarrow{P} \chi(v,0)$, converges to 1 in probability. Hence, it follows from Slutzky's theorem that $A_1 \xrightarrow{P} 0$. 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 (4.11). 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 as Corollary 3.2. □

Remark 4.2. Note that in (4.9) and (4.10) 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 (4.13) we obtain convergence of the order n^a for $a \in [2/3, 4/5)$; i.e. a better rate. □

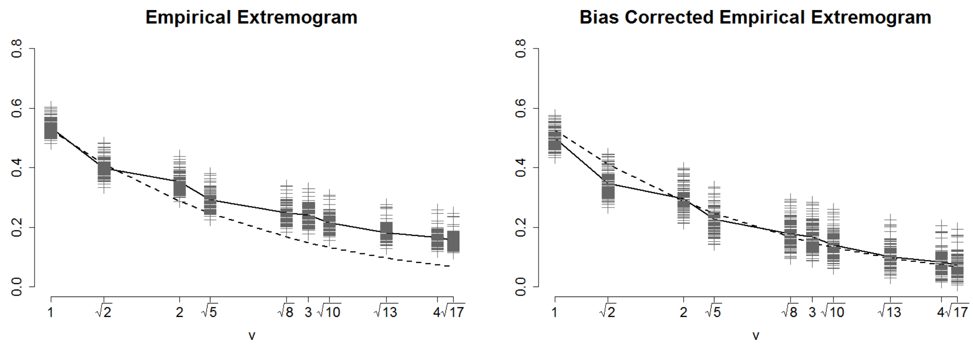

Figure 4.1: Empirical spatial extremogram (left) and its bias corrected version (right) for 100 simulated max-stable random fields in (4.1) with $\delta(v,0) = 2\cdot 0.4v^{1.5}$. The dashed line represents the theoretical spatial extremogram and the solid line is the mean over all 100 replicates.

***Example* 4.5.** We generate 100 realizations of the Brown-Resnick process in (4.1) 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 (2.5) and the bias corrected ones $\widetilde{\chi}(v,0)$ in (4.12) 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 summarized in Figure 4.1. We see that the bias corrected extremogram is closer to the true one. $\square$

4.2. Asymptotic properties of spatial parameter estimates of the Brown-Resnick process

In this section we prove asymptotic normality of the WLSE $(\widehat{\theta}_1,\widehat{\alpha}_1)$. We proceed as in the more general setting in Section 3.2. Recall that in the more specific situation here we have $C_1 = \log(\theta_1)$ and choose the transformation $T_1(\chi(v,0)) = 2\log\big(\Phi^{-1}\big(1-\frac{1}{2}\chi(v,0)\big)\big)$, where the log transformed version of the spatial lag satisfies $x_v = \log(v) \geq 0$ for $v \in \mathcal{V}$. We set $\widetilde{\chi}(v,0) = \frac{1}{T}\sum_{k=1}^{T}\widetilde{\chi}^{(t_k)}(v,0)$ as in (2.7), possibly after a bias correction, which depends on the two cases described in Remark 4.1. The analogue of the weighted least squares optimization problem (2.9) then reads as

$$\begin{pmatrix}\widehat{\theta}_1\\ \widehat{\alpha}_1\end{pmatrix} = \underset{\substack{\theta_1,\alpha_1>0\\ \alpha_1\in(0,2]}}{\arg\min}\sum_{v\in\mathcal{V}} w_v\big(T_1(\widetilde{\chi}(v,0)) - \big(\log(\theta_1)+\alpha_1 x_v\big)\big)^2. \tag{4.14}$$

Note in particular that α_1 has bounded support; this has to be treated as a special case in what follows. To show asymptotic normality of the WLSE in (4.14), we define as before design and weight matrices X and W as

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

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 $\Theta_{\mathcal{S}} = \mathbb{R}\times(0,2]$. Then the WLSE; i.e., the solution to (4.14) is given by

$$\widehat{\boldsymbol{\psi}}_1 = \begin{pmatrix} \log(\widehat{\theta}_1) \\ \widehat{\alpha}_1 \end{pmatrix} = (X^\intercal W X)^{-1} X^\intercal W (T_1(\widetilde{\chi}(v,0)))^\intercal_{v\in\mathcal{V}}. \tag{4.15}$$

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 4.6. *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 (4.14) and $\boldsymbol{\psi}_1^* = (\log(\theta_1^*), \alpha_1^*)^\intercal \in \Theta_{\mathcal{S}}$ the true parameter vector. Set $m_n = n^{\beta_1}$ for $\beta_1 \in (1/5, 1/2)$. Then as $n\to\infty$,*

$$\frac{n}{m_n}\Big(\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^*\Big) \xrightarrow{d} \begin{cases} \boldsymbol{Z}_1 & \text{if } \alpha_1^* < 2, \\ \boldsymbol{Z}_2 & \text{if } \alpha_1^* = 2, \end{cases} \tag{4.16}$$

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

$$\begin{aligned} \mathbb{P}(\boldsymbol{Z}_2 \in B) &= \int_{B\cap\{(b_1,b_2)\in\mathbb{R}^2: b_2<0\}} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(z_1,z_2)dz_1dz_2 \\ &\quad + \int_0^\infty \int_{\{b_1\in\mathbb{R}:(b_1,0)\in B\}} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}\Big(z_1 - \frac{1}{\sum_{v\in\mathcal{V}} w_v}\sum_{v\in\mathcal{V}}(w_v x_v)\ z_2,\ z_2\Big)dz_1dz_2 \end{aligned} \tag{4.17}$$

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

$$\begin{aligned} \mathbb{P}\Big(\boldsymbol{Z}_2 \le (p_1,p_2)^\intercal\Big) &= \int_{-\infty}^{\min\{0,p_2\}} \int_{-\infty}^{p_1} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(z_1,z_2)dz_1dz_2 \\ &\quad + \mathbb{1}_{\{p_2\ge 0\}} \int_0^\infty \int_{-\infty}^{p_1} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}\Big(z_1 - \frac{1}{\sum_{v\in\mathcal{V}} w_v}\sum_{v\in\mathcal{V}}(w_v x_v)\ z_2,\ z_2\Big)dz_1dz_2. \end{aligned} \tag{4.18}$$

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

$$\Pi_3^{(\mathrm{iso})} = Q_x^{(w)} G \Pi_2^{(\mathrm{iso})} G Q_x^{(w)\intercal}, \tag{4.19}$$

where $\Pi_2^{(\mathrm{iso})}$ *is the covariance matrix given in* (3.11),

$$Q_x^{(w)} = (X^\intercal W X)^{-1} X^\intercal W \quad \text{and} \quad G = \operatorname{diag}\left\{ \sqrt{\frac{2\pi}{\theta_1^* v^{\alpha_1^*}}} \exp\left\{\frac{1}{2}\theta_1^* v^{\alpha_1^*}\right\} : \ v \in \mathcal{V} \right\}.$$

Proof. For the first part of the proof, we neglect the constraints on α_1. Then we can directly use Theorem 3.4, observing that the derivative of T_1 is given by

$$T_1'(x) = -\Big(\Phi^{-1}(1-\frac{x}{2})\,\varphi(\Phi^{-1}(1-\frac{x}{2}))\Big)^{-1}, \quad 0 < x < 1,$$

where φ is the univariate standard normal density. Thus,

$$T_1'(\chi(v,0)) = -\Big(\sqrt{\theta_1^* v^{\alpha_1^*}}\,\varphi\big(\sqrt{\theta_1^* v^{\alpha_1^*}}\big)\Big)^{-1} = -\sqrt{\frac{2\pi}{\theta_1^* v^{\alpha_1^*}}} \exp\left\{\frac{1}{2}\theta_1^* v^{\alpha_1^*}\right\}.$$

Hence, as $n \to \infty$,

$$\frac{n}{m_n}\left(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*\right) = \frac{n}{m_n} Q_x^{(w)} \big(T_1(\widetilde{\chi}(v,0)) - T_1(\chi(v,0))\big)_{v\in\mathcal{V}} \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, Q_x^{(w)} G \Pi_2^{(\mathrm{iso})} G Q_x^{(w)\intercal}\right).$$

Note that we can define the diagonal matrix G unsigned, since signs cancel out. 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 $\{\widehat{\alpha}_1 > 2\}$,

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

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

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

it follows that

$$\mathbb{P}(\widehat{\alpha}_1 \le 2) \to 1 \quad \text{and} \quad \mathbb{P}(\widehat{\alpha}_1 > 2) \to 0, \quad n \to \infty. \tag{4.20}$$

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

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

We now consider the case $\alpha_1^* = 2$ and $\widehat{\alpha}_1 > 2$ (the unconstrained estimate exceeds 2). In this case (4.14) leads to the constrained optimization problem

$$\min_{\boldsymbol{\psi}_1}\{[W^{1/2}((T_1(\widetilde{\chi}(v,0)))_{v\in\mathcal{V}}^\intercal - X\boldsymbol{\psi}_1)]^\intercal [W^{1/2}((T_1(\widetilde{\chi}(v,0)))_{v\in\mathcal{V}}^\intercal - X\boldsymbol{\psi}_1]\},$$

$$\text{s.t.} \quad (0,1)\boldsymbol{\psi}_1 = 2.$$

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 with respect to the induced norm $\boldsymbol{\psi} \mapsto (\boldsymbol{\psi}^\intercal X^\intercal W X \boldsymbol{\psi})^{1/2}$ is given by (cf. Andrews [1], page 1365)

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

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

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

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

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

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 $c \ge 0$. For the asymptotic distribution we calculate, denoting by f^{-1} the inverse image of f,

$$\begin{aligned}
&\mathbb{P}\Big(\frac{n}{m_n}(\widehat{\boldsymbol{\psi}}_1^c - \boldsymbol{\psi}_1^*) \in B\Big) \\
&= \mathbb{P}\Big(\frac{n}{m_n}f(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*) \in B\Big) = \mathbb{P}\Big(f\big(\frac{n}{m_n}(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*)\big) \in B\Big) \\
&= \mathbb{P}\Big(\frac{n}{m_n}(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*) \in f^{-1}(B \cap \{(b_1,b_2)\in\mathbb{R}^2 : b_2 < 0\}) \cup f^{-1}(B \cap \{(b_1,0) : b_1 \in \mathbb{R}\})\Big) \\
&= \mathbb{P}\Big(\frac{n}{m_n}(\widehat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*) \in [B \cap \{(b_1,b_2)\in\mathbb{R}^2 : b_2 < 0\}] \\
&\qquad \cup [\{(b_1 - p_{wx}b_2, b_2), b_2 \ge 0, (b_1,0) \in B\}]\Big)
\end{aligned}$$

$$\to \int_{B\cap\{(b_1,b_2)\in\mathbb{R}^2, b_2<0\}} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(z_1,z_2)dz_1dz_2 \\ + \int_0^\infty \int_{\{b_1\in\mathbb{R},(b_1,0)\in B\}} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(z_1 - p_{wx}z_2, z_2)dz_1dz_2, \quad n\to\infty.$$

Plugging in $B = (-\infty, p_1] \times (-\infty, p_2]$ and using the Fubini-Tonelli theorem yields (4.18). □

***Remark* 4.3.** The asymptotic properties for the constrained estimate are derived as 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 for $\alpha_1^* = 2$ results from the fact that approximately half of the estimates lie above the true value and are therefore equal to two. □

5. Analysis of radar rainfall measurements

Finally, we apply the Brown-Resnick space-time process in (4.1) and the WLSE to radar rainfall data provided by the Southwest Florida Water Management District (SWFWMD)[1]. Our objective is to quantify their extremal behaviour by using spatial and temporal block maxima and fitting a Brown-Resnick space-time process 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. In this way we obtain 12×12 locations on 732 days of space-time block maxima of rainfall observations. Taking block maxima yields a process consistent with the assumption of a max-stable process, or at least to lie in the domain of attraction of a max-stable process. Taking daily data, we can furthermore ignore diurnal patterns.

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 to ours, we use the the same preprocessing steps; for a precise description cf. [4], Section 5.1.

The data do not fail the max-stability check described in Section 5.2 of [4], such that we assume that $\{\eta(\boldsymbol{s}, t), \boldsymbol{s} \in \mathcal{S}, t \in \{t_1, \ldots, t_{732}\}\}$ are realizations of a max-stable space-time process with standard unit Fréchet margins. Nevertheless, the assumption that the data are in fact an exact realization from a max-stable process is only approximate. Hence there is no guarantee that composite likelihood estimation applied to these transformed data outperforms the semiparametric estimation introduced in Section 2; cf. the results obtained in Section 10 of the supplement [7] when data have observational noise. Here we use this data example to illustrate our new semiparametric methodology.

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

We fit the Brown-Resnick process (4.1) by estimating (4.2) 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 $\mathcal{U} = \{1, \ldots, 10\}$. Permutation tests as described below and visualized in Figure 5.3 indicate that these lags are sufficient to cover the relevant extremal dependence structure. 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 realizations 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 at least approximately a max-stable process and on the robustness of the estimates derived in Section 9 of the supplement [7].
For the temporal estimation, we choose the empirical 90%−quantile for q.

(2) We perform subsampling by constructing subsets of the observations and estimating on the subsets (see Section 7 of the supplement [7]) 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) for the spatial dimensions and $b_t = 300$ for the temporal one. As overlap parameters we take $e_s = e_t = 1$, which corresponds to the maximum degree of overlap.

The results are shown in Figures 5.2, 5.3 and Table 1. Figure 5.1 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 1.

For comparison we present the parameter estimates from the pairwise likelihood estimation (for details see Davis et al. [10] 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 1 we recognize that these estimates are close to the semiparametric estimates and even lie in most cases in the 95%-subsampling confidence intervals.

Figure 5.2 shows the temporal and spatial mean of empirical temporal (left) and spatial (right) extremograms as described in (2.7) and (2.8) 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 5.3 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	$\widehat{\theta}_1$	0.3611	$\widehat{\alpha}_1$	0.9876
Subsampling-CI		[0.3472,0.3755]		[0.9482,1.0267]
Estimate	$\widehat{\theta}_2$	2.3650	$\widehat{\alpha}_2$	0.0818
Subsampling-CI		[1.9110,2.7381]		[0.0000,0.2680]

Table 1.: 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 (4.1) together with 95% subsampling confidence intervals.

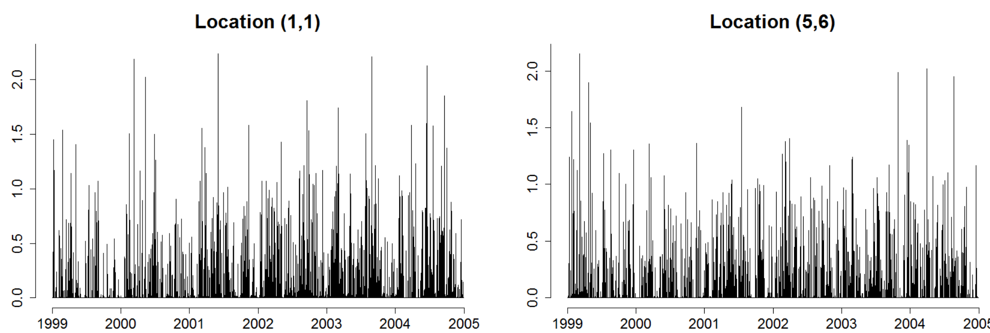

Figure 5.1: Daily rainfall maxima over hourly accumulated measurements from 1999-2004 in inches for two grid locations.

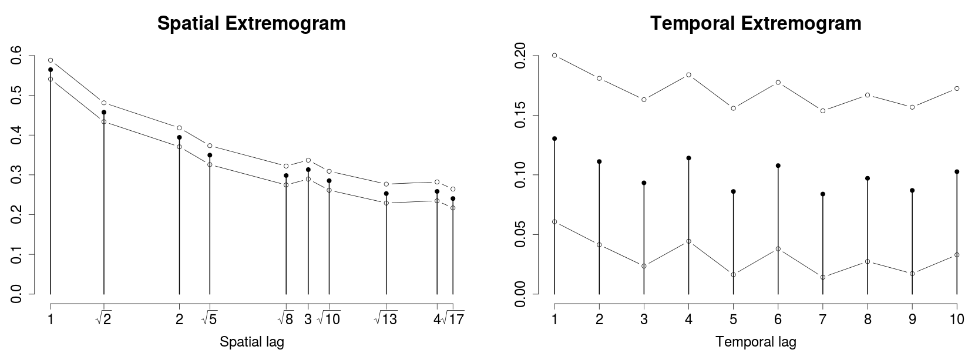

Figure 5.2: Empirical spatial (left) and temporal (right) extremogram based on spatial and temporal means for the space-time observations as given in (2.7) and (2.8) together with 95%−subsampling confidence intervals.

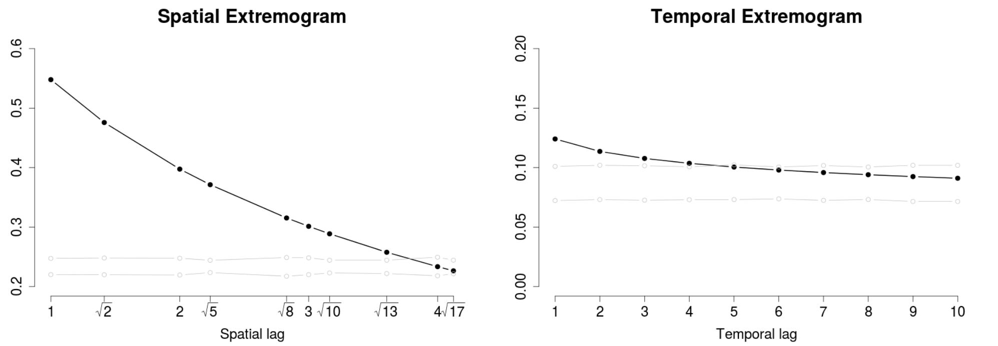

Figure 5.3: Permutation test for extremal independence: The gray lines show the 97.5%− and 2.5%−quantiles of the extremogram estimates for 1000 random space-time permutations for the empirical spatial (left) and the temporal (right) extremogram estimates.

6. Conclusions and Outlook

For isotropic strictly stationary regularly-varying space-time processes with additively separable dependence structure we have suggested a new semiparametric estimation method. The method works remarkably well and produces reliable estimates that are much faster to compute than composite likelihood estimates. These estimates can also be useful as initial values for a composite likelihood optimization.

Meanwhile, we have generalized the semiparametric method based on extremogram estimation. The paper Buhl and Klüppelberg [6] is dedicated to the three topics:

1. Generalize the dependence function (4.2) to anisotropic and appropriate mixed models and get rid of the assumption of separability.
2. Generalize the sampling scheme to a fixed (small) number of spatial observations and 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 2.

Another 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 LSE of the variogram parameters; e.g. in Lahiri et al. [23], Section 4. Here 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 $\Pi_2^{(\mathrm{iso})}$ and $\Pi_2^{(\mathrm{time})}$ (given in equations (4.3)-(4.6) of Buhl and Klüppelberg [5]), whose components are infinite sums.

Acknowledgements

The three last authors gratefully acknowledge support by the TUM Institute for Advanced Study. Furthermore, all authors would like to thank Chin Man 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 for providing the data. We thank Dennis Leber for comparing simulation methods of the BR process, and Ton Dieker and Marius Hofert for improving the simulation code. SB and CS thank the International Graduate School of Science and Engineering (IGSSE) of the Technical University of Munich 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.

Supplementary Material

Supplement to "Semiparametric estimation for isotropic max-stable space-time processes"
(doi: COMPLETED BY THE TYPESETTER; BDKSsupp.pdf). We provide additional results on $\alpha-$mixing, subsampling for confidence regions, and a simulation study supporting the theoretical results. Our method is extended to max-stable date with observational noise and applied to both exact realizations of the Brown-Resnick process and to realizations with observational noise, thus verifying the robustness of our approach.

References

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

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

[3] S. Buhl. *Statistical Modelling and Estimation of Space-Time Extremes*. Dissertation, Technische Universität München, München, 2017.

[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. Stochastic Processes and their Applications, in Press. arXiv 1609.04961, 2017.

[6] S. Buhl and C. Klüppelberg. Generalised least squares estimation of regularly varying space-time processes based on flexible observation schemes. arXiv:1862319[math.ST], 2017.

[7] S. Buhl, R.A. Davis, C. Klüppelberg, and C. Steinkohl. Supplement to "Semiparametric estimation for isotropic max-stable space-time processes". 2018.

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

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

[10] 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.

[11] 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.

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

[13] C. Dombry, S. Engelke, and M. Oesting. Exact simulation of max-stable processes. *Biometrika*, 103:303–317, 2016.

[14] J.H.J. Einmahl, A. Kiriliouk, and J. Segers. A continuous updating weighted least squares estimator of tail dependence in high dimension. arXiv:1601.04826vl[stat.ME], 2016.

[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. Extremal behavior of regularly varying stochastic processes. *Stochastic Processes and their Applications*, 115:249–274, 2005.

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

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

[22] 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.

[23] 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.

[24] 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.

[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.