Bernoulli **25**(4A), 2019, 2508–2537
https://doi.org/10.3150/18-BEJ1061

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*

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. (2018)].

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

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. [9].

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. [10]. 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.

1350-7265

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 [11] 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 [7] 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 [5] and Davis et al. [10] 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 [7]. 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 [4], 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 [4]. 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 [4] 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 [4]). 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 [4] 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

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

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 $\mathbb{P}(|\eta(\boldsymbol{0},0)| > a_n) \sim n^{-d}$ as $n \to \infty$, and if for every finite set $\mathcal{I} \subset \mathbb{R}^{d-1} \times [0,\infty)$,

$$n^d \mathbb{P}\left(\frac{\eta_{\mathcal{I}}}{a_n} \in \cdot\right) \stackrel{v}{\to} \mu_{\mathcal{I}}(\cdot), \qquad n \to \infty, \tag{2.1}$$

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

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

for every Borel set C in $\overline{\mathbb{R}}^{|\mathcal{I}|}\backslash\{\boldsymbol{0}\}$. The notation $\stackrel{v}{\to}$ 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 $\mu_{\{(\boldsymbol{s},t)\}}(\cdot) = \mu_{\{(\boldsymbol{0},0)\}}(\cdot) =: \mu(\cdot)$, 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 [7] 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

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

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

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

and

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

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 [7], 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 [5], 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

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

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

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

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

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

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; that is,

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

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

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

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

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

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 i.i.d. 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 [4] and in Buhl and Klüppelberg [7].

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

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

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* $t \in \{t_1, \dots, t_T\}$, *consider a regularly varying spatial process* $\{\eta(\boldsymbol{s}, t) : \boldsymbol{s} \in \mathbb{R}^2\}$ *as defined in Definition 2.1. Let* a_n *be a sequence as in* (2.1). *Assume that there exists* $\gamma > 0$ *that satisfies* $\max\{v_1, \dots, v_p\} \le \gamma$, *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* $\varepsilon > 0$:

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

where $B(\boldsymbol{h},\gamma) := \{\boldsymbol{s}\in\mathbb{Z}^2 : \|\boldsymbol{s}-\boldsymbol{h}\|\le\gamma\}$ *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* $2\le p+q\le 4$,
(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*

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

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

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

Proof. Theorem 3.1 is a direct application of Theorem 4.2 of Buhl and Klüppelberg [7] 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 $i\in\{1,\dots,p\}$ such that $v_i\in\mathcal{V}$, we summarize these into

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

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

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

where $\Pi_1^{(\mathrm{space})}$ is specified in equation (4.3)–(4.6) of [7]. 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; that is,

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

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

$$\begin{aligned} &\frac{n}{m_n}\big(\widehat{\chi}^{(t)}(v_i,0) - \chi_n(v_i,0)\big)^{\mathsf{T}}_{i=1,\dots,p} \\ &= \frac{n}{m_n} N\big(\widehat{\chi}^{(t)}\big(\boldsymbol{h}_1^{(i)},0\big) - \chi_n\big(\boldsymbol{h}_1^{(i)},0\big),\dots,\widehat{\chi}^{(t)}\big(\boldsymbol{h}_{\ell_i}^{(i)},0\big) - \chi_n\big(\boldsymbol{h}_{\ell_i}^{(i)},0\big)\big)^{\mathsf{T}}_{i=1,\dots,p} \\ &\xrightarrow{d} \mathcal{N}\big(\mathbf{0}, N\Pi_1^{(\text{space})}N^{\mathsf{T}}\big), \qquad n \to \infty, \end{aligned}$$

such that

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

□

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*)

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

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 [7] and Corollary 3.4 of Davis and Mikosch [11]. 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 $\{\boldsymbol{h}_1^{(i)},\dots,\boldsymbol{h}_{\ell_i}^{(i)}: i=1,\dots,p\}$ and let $\gamma\ge\max\{v_1,\dots,v_p\}$. Define the vector process

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

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

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

and the sets D_k by

$$\big\{\boldsymbol{Y}(\boldsymbol{s})\in D_k\big\}=\big\{\eta(\boldsymbol{s},t_k)\in A\big\}.$$

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

$$\mu_{B_T(\mathbf{0},\gamma)}(C):=\lim_{n\to\infty}m_n^2\mathbb{P}\left(\frac{\boldsymbol{Y}(\mathbf{0})}{m_n^2}\in C\right),$$

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

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

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

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

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

Likewise we have for the pre-asymptotic quantities

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

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

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

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

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

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

$$F_1=\operatorname{diag}\big(\mu(A)\big)$$

and

$$F_2^{(k)}:=\big(-\mu_\gamma\big(D_{1,k}^{(1)}\big),\dots,-\mu_\gamma\big(D_{\ell_1,k}^{(1)}\big),\dots,-\mu_\gamma\big(D_{\ell_p,k}^{(p)}\big)\big)^\top.$$

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

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

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 $k=1,\dots,T$ the vectors in $\mathbb{R}^{L+1}$

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

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

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

where the last step follows as in the proof of Theorem 4.2 of [7] and involves Slutzky's theorem. Using ideas of the proof of their Lemma 5.1, we observe that as $n \to \infty$,

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

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

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

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

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

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

$$\Pi_2^{(\mathrm{iso})} = \mu(A)^{-4} T^{-2} \begin{pmatrix} 1 & 0\cdots0 & 1 & 0\cdots0 & \cdots & 1 & 0\cdots0 \\ 0 & 1\cdots0 & 0 & 1\cdots0 & \cdots & 0 & 1\cdots0 \\ & & & \ddots & & & \\ 0 & 0\cdots1 & 0 & 0\cdots1 & \cdots & 0 & 0\cdots1 \end{pmatrix} \boldsymbol{NF}\Sigma(\boldsymbol{NF})^\top \times \begin{pmatrix} 1 & 0\cdots0 & 1 & 0\cdots0 & \cdots & 1 & 0\cdots0 \\ 0 & 1\cdots0 & 0 & 1\cdots0 & \cdots & 0 & 1\cdots0 \\ & & & \ddots & & & \\ 0 & 0\cdots1 & 0 & 0\cdots1 & \cdots & 0 & 0\cdots1 \end{pmatrix}^\top . \tag{3.11}$$

□

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

$$\frac{n}{m_n}\big(\chi_n(v, 0) - \chi(v, 0)\big) \to 0, \qquad n \to \infty, \tag{3.12}$$

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

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

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 example 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); that is,

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

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

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

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

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

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 $\{\eta(\boldsymbol{s}, t) : \boldsymbol{s} \in \mathbb{R}^2\}$ 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 $n \to \infty$,*

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

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

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

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

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

where G is defined in (3.15). Since

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

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

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

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

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

□

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

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

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*; that is, for every $m \in \mathbb{N}$ and every $(\boldsymbol{s}^{(1)}, t^{(1)}), \ldots, (\boldsymbol{s}^{(m)}, t^{(m)}) \in \mathbb{R}^2 \times [0, \infty)$, it holds that

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

for all $a_1, \ldots, a_m \in \mathbb{R}$ 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, $W(\boldsymbol{0}, 0) = 0$, $\mathbb{E}[W(\boldsymbol{s}, t)] = 0$ and covariance function

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

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

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

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

$$\delta(v, u) = 2\theta_1 v^{\alpha_1} + 2\theta_2 u^{\alpha_2}, \tag{4.2}$$

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 $(\eta(\boldsymbol{0}, 0), \eta(\boldsymbol{h}, u))$ is given for $x_1, x_2 > 0$ by

$$\begin{aligned}F(x_1, x_2) = \exp\bigg\{&-\frac{1}{x_1}\Phi\bigg(\frac{\log(x_2/x_1)}{\sqrt{2\delta(\|\boldsymbol{h}\|, |u|)}} + \sqrt{\frac{\delta(\|\boldsymbol{h}\|, |u|)}{2}}\bigg)\\ &- \frac{1}{x_2}\Phi\bigg(\frac{\log(x_1/x_2)}{\sqrt{2\delta(\|\boldsymbol{h}\|, |u|)}} + \sqrt{\frac{\delta(\|\boldsymbol{h}\|, |u|)}{2}}\bigg)\bigg\},\end{aligned} \tag{4.3}$$

where Φ denotes the standard normal distribution function (cf. Davis et al. [9]).

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. [10], 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. [9], equation (3.1)). *Let $\{\eta(s,t) : (s,t) \in \mathbb{R}^2 \times [0,\infty)\}$ 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*

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

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

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

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

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

In the same way, we obtain

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

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 $T_1(\chi(v,0)) = 2\log(\Phi^{-1}(1 - \frac{1}{2}\chi(v,0)))$ and $T_2(\chi(0,u)) = 2\log(\Phi^{-1}(1 - \frac{1}{2}\chi(0,u)))$. 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; that is, 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 $t\in[0,\infty)$,*

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

Lemma 4.2 is a direct application of Lemma A.1(b) of Buhl and Klüppelberg [7] 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 $\mathbb{P}(|\eta(\mathbf{0},0)|>a_n)\sim n^{-2}$ 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*

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

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))

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

Proof. We need to verify the conditions of Corollary 3.2; that is, conditions (M1)–(M4) of Theorem 3.1 for $a_{m_n}=m_n^2$, and apply results of Section 8 of the supplement [4].

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 of m_n and r_n, we have $m_n = o(n)$ 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 $\varepsilon > 0$, like in Example 4.6 of Buhl and Klüppelberg [7], we have for $\boldsymbol{s}, \boldsymbol{s}' \in \mathbb{R}^2$ by a Taylor expansion,

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

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

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

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 $k \ge 2\gamma$,

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

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 $K_2 > 0$,

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

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 (M4)(i). With V as before, and with equation (8.2), we estimate, recalling from above that the number of lags $\|\boldsymbol{h}\| = v$ is of order $\mathcal{O}(v)$,

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

By Lemma 8.3, we find

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

By the same arguments condition (M4)(ii) is satisfied.

Condition (M4)(iii) holds by equation (8.2), since

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

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),

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

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

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

cf. Theorem 4.4 of Buhl and Klüppelberg [7]. Thus, we have to distinguish two cases:

(I) For $\beta_1 \le 1/3$ we cannot replace the pre-asymptotic extremogram by the theoretical version, but can resort to a bias correction, which is described in (4.12) below.

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

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

and likewise for the averaged empirical extremogram,

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

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

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

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

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

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 $t\in\{t_1,\dots,t_T\}$ consider the spatial Brown–Resnick process $\{\eta(\boldsymbol{s},t), \boldsymbol{s}\in\mathbb{R}^2\}$ defined in (4.1) with dependence function given in (4.2). Set $m_n = n^{\beta_1}$ for $\beta_1 \in (\frac{1}{5},\frac{1}{3}]$. Then the bias corrected empirical spatial extremogram (4.12) satisfies*

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

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

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

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 $n \to \infty$,

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

By Theorem 4.3 it suffices to show that $(n/(2m_n^3))(\widehat{\nu}(v,0) - \nu(v,0)) \xrightarrow{P} 0$. Setting $\nu_n(v,0) := (\chi_n(v,0) - 2)(\chi_n(v,0) - 1)$ we have

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

We calculate

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

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

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

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 $a \in [2/3, 4/5)$; that is, a better rate.

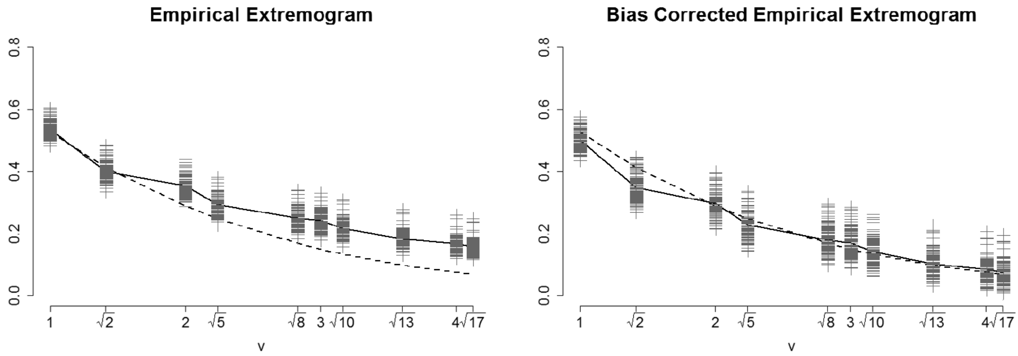

Figure 1. Empirical spatial extremogram (left) and its bias corrected version (right) for 100 simulated max-stable random fields in (4.1) with $\delta(v,0) = 2 \cdot 0.4v^{1.5}$. The dashed line represents the theoretical spatial extremogram and the solid line is the mean over all 100 replicates.

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 1. We see that the bias corrected extremogram is closer to the true one.

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 $T_1(\chi(v,0)) = 2\log(\Phi^{-1}(1 - \frac{1}{2}\chi(v,0)))$, 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

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

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

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

respectively, where $\mathbf{1} = (1, \dots, 1)^\top \in \mathbb{R}^p$. Let $\boldsymbol{\psi}_1 = (\log(\theta_1), \alpha_1)^\top$ be the parameter vector with parameter space $\Theta_{\mathcal{S}} = \mathbb{R} \times (0, 2]$. Then the WLSE; that is, the solution to (4.14) is given by

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

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 $\widehat{\boldsymbol{\psi}}_1^c = (\log(\widehat{\theta}_1^c), \widehat{\alpha}_1^c)^\top$.

Theorem 4.6. *Let $\widehat{\boldsymbol{\psi}}_1^c = (\log(\widehat{\theta}_1^c), \widehat{\alpha}_1^c)^\top$ denote the WLSE resulting from the constrained minimization problem (4.14) and $\boldsymbol{\psi}_1^* = (\log(\theta_1^*), \alpha_1^*)^\top \in \Theta_{\mathcal{S}}$ the true parameter vector. Set $m_n = n^{\beta_1}$ for $\beta_1 \in (1/5, 1/2)$. Then as $n \to \infty$,*

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

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

$$\begin{aligned} \mathbb{P}(\mathbf{Z}_2 \in B) = {} & \int_{B \cap \{(b_1, b_2) \in \mathbb{R}^2 : b_2 < 0\}} \varphi_{\mathbf{0}, \Pi_3^{(\text{iso})}}(z_1, z_2)\, dz_1\, dz_2 \\ & + \int_0^\infty \int_{\{b_1 \in \mathbb{R} : (b_1, 0) \in B\}} \varphi_{\mathbf{0}, \Pi_3^{(\text{iso})}}\Big(z_1 - \frac{1}{\sum_{v \in \mathcal{V}} w_v} \sum_{v \in \mathcal{V}} (w_v x_v) z_2, z_2\Big)\, dz_1\, dz_2 \end{aligned} \tag{4.17}$$

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 $\mathbf{Z}_2$ is given for $(p_1, p_2)^\top \in \mathbb{R}^2$ by

$$\begin{aligned} & \mathbb{P}\big(\mathbf{Z}_2 \le (p_1, p_2)^\top\big) \\ & = \int_{-\infty}^{\min\{0, p_2\}} \int_{-\infty}^{p_1} \varphi_{\mathbf{0}, \Pi_3^{(\text{iso})}}(z_1, z_2)\, dz_1\, dz_2 \\ & \quad + \mathbb{1}_{\{p_2 \ge 0\}} \int_0^\infty \int_{-\infty}^{p_1} \varphi_{\mathbf{0}, \Pi_3^{(\text{iso})}}\Big(z_1 - \frac{1}{\sum_{v \in \mathcal{V}} w_v} \sum_{v \in \mathcal{V}} (w_v x_v) z_2, z_2\Big)\, dz_1\, dz_2. \end{aligned} \tag{4.18}$$

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

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

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

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

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

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

where φ is the univariate standard normal density. Thus,

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

Hence, as $n \to \infty$,

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

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 $\alpha_1^* = 2$; that is, 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

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

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

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

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

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

it follows that

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

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

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

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

$$\min_{\boldsymbol{\psi}_1}\big\{\big[W^{1/2}\big(\big(T_1\big(\widetilde{\chi}(v,0)\big)\big)_{v\in\mathcal{V}}^\mathsf{T} - X\boldsymbol{\psi}_1\big)\big]^\mathsf{T}\big[W^{1/2}\big(\big(T_1\big(\widetilde{\chi}(v,0)\big)\big)_{v\in\mathcal{V}}^\mathsf{T} - X\boldsymbol{\psi}_1\big]\big\},$$
$$\text{s.t. } (0,1)\boldsymbol{\psi}_1 = 2.$$

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 $\boldsymbol{\psi} \mapsto (\boldsymbol{\psi}^\top X^\top W X \boldsymbol{\psi})^{1/2}$ is given by (cf. Andrews [1], page 1365)

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

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

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

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

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

This implies

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

Let $f(x_1,x_2) = (x_1 + p_{wx}x_2\mathbb{1}_{\{x_2>0\}}, x_2 - x_2\mathbb{1}_{\{x_2>0\}})^\top$ 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,

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

$$\to \int_{B\cap\{(b_1,b_2)\in\mathbb{R}^2, b_2<0\}} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(z_1,z_2)\,dz_1\,dz_2$$
$$+ \int_0^\infty \int_{\{b_1\in\mathbb{R},(b_1,0)\in B\}} \varphi_{\mathbf{0},\Pi_3^{(\mathrm{iso})}}(z_1 - p_{wx}z_2, z_2)\,dz_1\,dz_2, \qquad n\to\infty.$$

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 120×120 km region containing 3600 grid locations. We calculate the spatial and temporal maxima over subregions of size 10×10 km 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 $\mathcal{S} = \{(i_1, i_2), i_1, i_2 \in \{1, \dots, 12\}\}$ and the space-time observations by $\{\eta(\boldsymbol{s}, t), \boldsymbol{s} \in \mathcal{S}, t \in \{t_1, \dots, t_{732}\}\}$. This setup is also considered in Buhl and Klüppelberg [5], 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. [5], Section 5.1.

The data do not fail the max-stability check described in Section 5.2 of [5], such that we assume that $\{\eta(\boldsymbol{s}, t), \boldsymbol{s} \in \mathcal{S}, t \in \{t_1, \dots, t_{732}\}\}$ 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 [4] 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 $\theta_1, \alpha_1, \theta_2$ 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 $\mathcal{U} = \{1, \dots, 10\}$. Permutation tests as described below and visualized in Figure 4 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; that is, $w_v = T^{-1} \sum_{k=1}^{T} \widetilde{\chi}^{(t_k)}(v, 0)$ and $w_u = n^{-2} \sum_{i=1}^{n^2} \widetilde{\chi}^{(s_i)}(0, u)$, 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 [4].

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 [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) 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 3, 4 and Table 1. Figure 2 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. [9] 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 3 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

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

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]

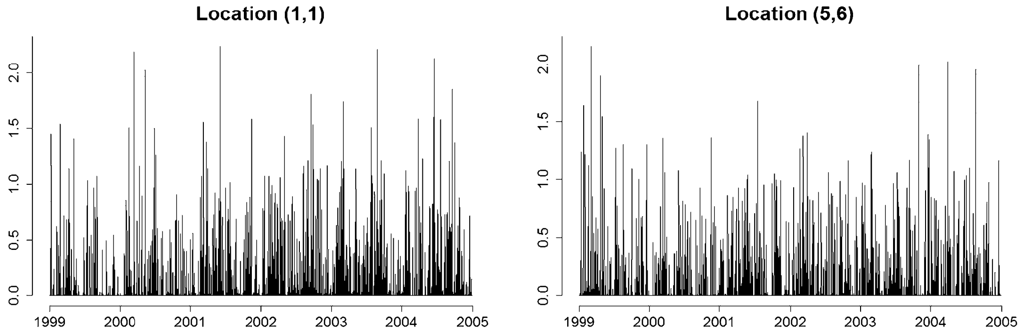

Figure 2. Daily rainfall maxima over hourly accumulated measurements from 1999–2004 in inches for two grid locations.

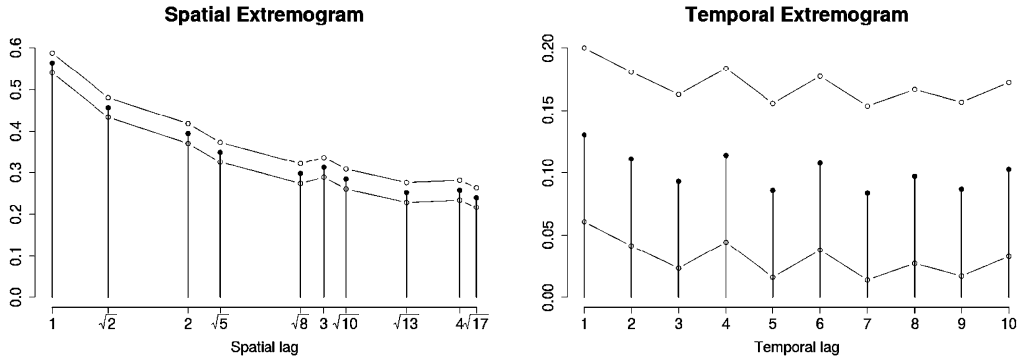

Figure 3. Empirical spatial (left) and temporal (right) extremogram based on spatial and temporal means for the space-time observations as given in (2.7) and (2.8) together with 95%-subsampling confidence intervals.

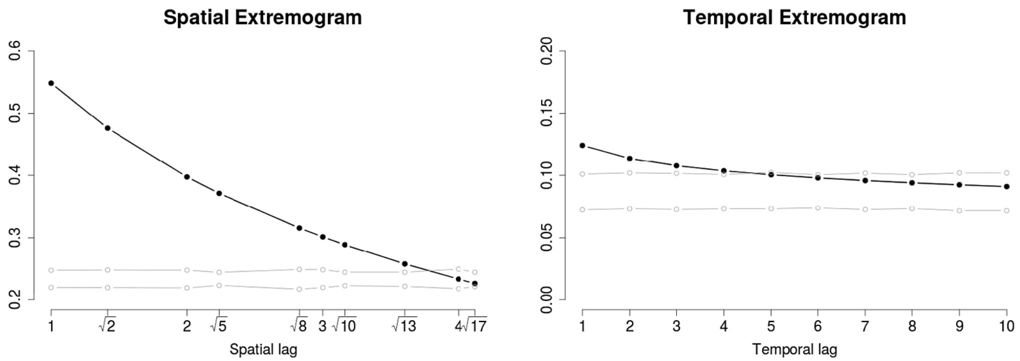

Figure 4. Permutation test for extremal independence: The gray lines show the 97.5%- and 2.5%-quantiles of the extremogram estimates for 1000 random space-time permutations for the empirical spatial (left) and the temporal (right) extremogram estimates.

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 4 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.

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; for example, 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; that is, of $(T^{-1}\sum_{k=1}^{T}\widetilde{\chi}^{(t_k)}(v,0))^{\mathsf{T}}_{v\in\mathcal{V}}$ in the spatial case and of $(n^{-2}\sum_{i=1}^{n^2}\widetilde{\chi}^{(s_i)}(0,u)-\chi(0,u))^{\mathsf{T}}_{u\in\mathcal{U}}$ in the temporal case. In our case, however, this involves the matrices $\Pi_2^{(\text{iso})}$ and $\Pi_2^{(\text{time})}$ (given in equations (4.3)–(4.6) of Buhl and Klüppelberg [7]), 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: 10.3150/18-BEJ1061SUPP; .pdf). We provide additional results on α-mixing, subsampling for confidence regions, and a simulation study supporting the theoretical results. Our method is extended to max-stable data 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] Andrews, D.W.K. (1999). Estimation when a parameter is on a boundary. *Econometrica* **67** 1341–1383. MR1720781

[2] Brown, B.M. and Resnick, S.I. (1977). Extreme values of independent stochastic processes. *J. Appl. Probab.* **14** 732–739. MR0517438

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

[4] Buhl, S., Davis, R.A., Klüppelberg, C. and Steinkohl, C. (2018). Supplement to "Semiparametric estimation for isotropic max-stable space-time processes." DOI:10.3150/18-BEJ1061SUPP.

[5] Buhl, S. and Klüppelberg, C. (2016). Anisotropic Brown–Resnick space-time processes: Estimation and model assessment. *Extremes* **19** 627–660. MR3558348

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

[7] Buhl, S. and Klüppelberg, C. (2018). Limit theory for the empirical extremogram of random fields. *Stochastic Process. Appl.* **128** 2060–2082. MR3797653

[8] Cho, Y.B., Davis, R.A. and Ghosh, S. (2016). Asymptotic properties of the empirical spatial extremogram. *Scand. J. Stat.* **43** 757–773. MR3543321

[9] Davis, R.A., Klüppelberg, C. and Steinkohl, C. (2013). Max-stable processes for modeling extremes observed in space and time. *J. Korean Statist. Soc.* **42** 399–414. MR3255398

[10] Davis, R.A., Klüppelberg, C. and Steinkohl, C. (2013). Statistical inference for max-stable processes in space and time. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 791–819. MR3124792

[11] Davis, R.A. and Mikosch, T. (2009). The extremogram: A correlogram for extreme events. *Bernoulli* **15** 977–1009. MR2597580

[12] de Haan, L. (1984). A spectral representation for max-stable processes. *Ann. Probab.* **12** 1194–1204. MR0757776

[13] Dombry, C., Engelke, S. and Oesting, M. (2016). Exact simulation of max-stable processes. *Biometrika* **103** 303–317. MR3509888

[14] Einmahl, J.H.J., Kiriliouk, A. and Segers, J. (2018). A continuous updating weighted least squares estimator of tail dependence in high dimensions. *Extremes* **21** 205–233. MR3800297

[15] Engelke, S., Malinowski, A., Kabluchko, Z. and Schlather, M. (2015). Estimation of Hüsler–Reiss distributions and Brown–Resnick processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 239–265. MR3299407

[16] Fasen, V., Klüppelberg, C. and Schlather, M. (2010). High-level dependence in time series models. *Extremes* **13** 1–33. MR2593949

[17] Genton, M.G., Ma, Y. and Sang, H. (2011). On the likelihood function of Gaussian max-stable processes. *Biometrika* **98** 481–488. MR2806443

[18] Giné, E., Hahn, M.G. and Vatan, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. *Probab. Theory Related Fields* **87** 139–165. MR1080487
[19] Hult, H. and Lindskog, F. (2005). Extremal behavior of regularly varying stochastic processes. *Stochastic Process. Appl.* **115** 249–274. MR2111194
[20] Hult, H. and Lindskog, F. (2006). Regular variation for measures on metric spaces. *Publ. Inst. Math.* (*Beograd*) (*N.S.*) **80(94)** 121–140. MR2281910
[21] Huser, R. and Davison, A.C. (2014). Space-time modelling of extreme events. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 439–461. MR3164873
[22] Kabluchko, Z., Schlather, M. and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37** 2042–2065. MR2561440
[23] Lahiri, S.N., Lee, Y. and Cressie, N. (2002). On asymptotic distribution and asymptotic efficiency of least squares estimators of spatial variogram parameters. *J. Statist. Plann. Inference* **103** 65–85. MR1896984
[24] Padoan, S.A., Ribatet, M. and Sisson, S.A. (2010). Likelihood-based inference for max-stable processes. *J. Amer. Statist. Assoc.* **105** 263–277. MR2757202
[25] Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2271424
[26] Schlather, M. RandomFields, contributed package on random field simulation for R. http://cran.r-project.org/web/packages/RandomFields/.
[27] Steinkohl, C. (2013). Statistical Modelling of Extremes in Space and Time using Max-Stable Processes. Dissertation, Technische Universität München, München.
[28] Thibaud, E. and Opitz, T. (2015). Efficient inference and simulation for elliptical Pareto processes. *Biometrika* **102** 855–870. MR3431558
[29] Wadsworth, J.L. and Tawn, J.A. (2014). Efficient inference for spatial extreme value processes associated to log-Gaussian random functions. *Biometrika* **101** 1–15. MR3180654

Received September 2016 and revised July 2018

arXiv: 1609.04967

Supplement to the paper "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, semi-parametric estimation, space-time process, subsampling, mixing.

This supplementary material provides additional definitions and results to the paper [4], where the setting, notation, equation reference numbers are retained from that paper. Section 8 defines α-mixing and states results for Brown-Resnick space-time processes used in the proof of Theorem 4.3 and throughout this supplement. Within the particular space-time setting considered in the paper, we provide insight into subsampling to obtain adequate confidence regions for the true parameters in Section 7. Section 9 states and proves an important result related to the extremogram for the Brown-Resnick process observed with noise. This result provides the theoretical justification for the robustness of WLSE for space-time data based on small departures from the Brown-Resnick model. The simulation study presented in Section 10 confirms these results and other findings of the paper.

7. Subsampling for confidence regions

As in Sections 2 and 3 of the paper [4] we consider a strictly stationary regularly varying process in space and time $\{\eta(\boldsymbol{s},t) : \boldsymbol{s} \in \mathbb{R}^{d-1}, t \in [0,\infty)\}$ for $d \in \mathbb{N}$. We assume additively separable parametric models for its extremogram $\{\chi(v,u), v, u \geq 0\}$, such that setting either the temporal lag u or the spatial lag v equal to 0, it can be linearly parametrized as

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

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

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 $\Theta_{\mathcal{T}}$.

The estimation method described in Sections 2 and 3, which is based on the (averaged) empirical extremogram computed by means of space-time observations on the grid $\mathcal{S}_n \times \{t_1, \dots, t_T\}$ defined in Condition 2.3, yields a consistent and asymptotically normal estimator $\widehat{\boldsymbol{\psi}}_1 = (\widehat{C}_1, \widehat{\alpha}_1)^\intercal$ of the true parameter vector $\boldsymbol{\psi}_1^\star = (C_1^\star, \alpha_1^\star)^\intercal$. The rate of convergence is given by $\tau_n := n/m_n$, where m_n is an appropriately chosen scaling sequence.

Due to the complicated forms of the covariance matrix of the normal limit distribution (cf. Theorem 3.1 and Theorem 3.19 of [2]) we use resampling methods to construct asymptotic confidence regions for $\boldsymbol{\psi}_1^\star$. One appealing method is subsampling (see Politis et al. [10], Chapter 5), since it works under weak regularity conditions and produces asymptotically correct coverage. The central assumption is the existence of a continuous weak limit law, which is guaranteed by Theorem 3.1. Again we only consider the spatial case, the temporal case is described (again for the example of the Brown-Resnick process) in Section 3.4.2 of [2].

We have applied subsampling successfully already for confidence bounds of pairwise likelihood estimates of the max-stable space-time Brown-Resnick process in Buhl and Klüppelberg [3], Section 4. The procedure is as follows: understanding inequalities between vectors componentwise, we choose block lengths $\boldsymbol{b} = (b_s, b_s, T)$ with $(1,1) \leq (b_s, b_s) \leq (n,n)$ and the degree of overlap $\boldsymbol{e} = (e_s, e_s, T)$ with $(1,1) \leq (e_s, e_s) \leq (b_s, b_s)$, where $\boldsymbol{e} = (1,1,T)$ corresponds to maximum overlap and $\boldsymbol{e} = \boldsymbol{b}$ to no overlap. The blocks are indexed by $\boldsymbol{i} = (i_1, i_2) \in \mathbb{N}^2$ with $i_j \leq q_s$ for $q_s := \lfloor \frac{n-b_s}{e_s} \rfloor + 1$ and $j = 1,2$. This results in a total number of $q = q_s^2$ blocks, which we summarize in the sets

$$E_{\boldsymbol{i},\boldsymbol{b},\boldsymbol{e}} = \big\{(s_1,s_2) \in \mathcal{S}_n : (i_j - 1)e_s + 1 \leq s_j \leq (i_j - 1)e_s + b_s \text{ for } j = 1,2\big\} \times \{t_1, \dots, t_T\}.$$

We estimate the parameters based on the observations in each block as described in the previous sections. This yields different estimates, which we denote by $\widehat{\boldsymbol{\psi}}_{1,\boldsymbol{i}}$.

Theorem 7.1 below provides a basis for constructing asymptotically valid confidence intervals for the true parameters $C_1^\star$ and $\alpha_1^\star$. We define $\tau_{b_s} = b_s/m_{b_s}$ as the analogue of $\tau_n = n/m_n$.

Theorem 7.1. *Assume that the conditions of Theorem 3.4 hold, and*

(i) $b_s \to \infty$ *such that* $b_s = o(n)$ *and* $\tau_{b_s}/\tau_n \to 0$ *as* $n \to \infty$
(ii) $\boldsymbol{e}$ *does not depend on* n,
(iii) the α*-mixing coefficients* $\alpha_{k,\ell}(\cdot)$ *defined in* (8.1) *satisfy*

$$\frac{1}{n^2} \sum_{r=1}^{n} r \alpha_{b,b}(r) \to 0, \quad n \to \infty,$$

where $b := b_s^2$.

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

$$L_{b_s,s}(x) := \frac{1}{q}\sum_{i_1=1}^{q_s}\sum_{i_2=1}^{q_s} \mathbb{1}_{\{\tau_{b_s}\|\widehat{\psi}_{1,i}-\widehat{\psi}_1\|\leq x\}}, \quad x\in\mathbb{R},$$

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

$$\mathbb{P}\Big(\tau_n\|\widehat{\boldsymbol{\psi}}_1-\boldsymbol{\psi}_1^\star\|\leq c_{b_s,s}(1-\alpha)\Big)\to 1-\alpha, \quad n\to\infty. \tag{7.1}$$

Proof. We apply Corollary 5.3.3 of Politis et al. [10]. Their main Assumption 5.3.3 is the existence of a continuous weak limit distribution of $\tau_n\|\widehat{\boldsymbol{\psi}}_1-\boldsymbol{\psi}_1^\star\|$, which holds by Theorem 3.1. The remaining assumptions (i)-(iii) are also presumed in Politis et al. [10]. □

As a consequence of equation (7.1), for n large enough, an approximate $(1-\alpha)$-confidence region for the true parameter vector $\boldsymbol{\psi}_1^\star=(C_1^\star,\alpha_1^\star)$ is given by

$$\{\boldsymbol{\psi}\in\Theta_{\mathcal{S}} : \|\boldsymbol{\psi}-\widehat{\boldsymbol{\psi}}_1\|\leq c_{b_s,s}(1-\alpha)/\tau_n\}, \tag{7.2}$$

where $\Theta_{\mathcal{S}}$ denotes as before the parameter space.

Remark 7.1. Consider the special case of the Brown-Resnick process (4.1) with dependence function δ given in (4.2) as

$$\delta(v,u)=2\theta_1 v^{\alpha_1}+2\theta_2 u^{\alpha_2}, \quad v,u\geq 0, \quad \theta_1,\theta_2>0, \quad 0<\alpha_1,\alpha_2\leq 2, \tag{7.3}$$

whose parameters α_1 and α_2 have bounded support. Recall from Section 4 that to put this in the context of this section, we set

$$C_1=\log(\theta_1) \quad \text{and} \quad C_2=\log(\theta_2)$$

and choose the transformations T_1 and T_2 defined by

$$T_1(\chi(v,0))=2\log\big(\Phi^{-1}\big(1-\frac{1}{2}\chi(v,0)\big)\big) \quad \text{and} \quad T_2(\chi(0,u))=2\log\big(\Phi^{-1}\big(1-\frac{1}{2}\chi(0,u)\big)\big).$$

In the following, we focus on the spatial parameters. The parameter space is given by $\Theta_{\mathcal{S}}=\mathbb{R}\times(0,2]$. Since the parameter space for α is bounded, we use the constrained estimate $\widehat{\boldsymbol{\psi}}_1^c$ defined before Theorem 4.6. We denote the true parameter by $\boldsymbol{\psi}_1^\star=(\log(\theta_1^\star),\alpha_1^\star)$. The assumptions of Theorem 7.1 for subsampling are satisfied in this setting. Particularly important is the existence of a continuous weak limit distribution of $\tau_n\|\widehat{\boldsymbol{\psi}}_1^c-\boldsymbol{\psi}_1^\star\|$, where the scaling sequence is given by $\tau_n=n/m_n$. By Theorem 4.6, the continuous mapping theorem and the Fubini-Tonelli theorem we have for $\gamma \geq 0$, as $n \to \infty$,

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

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

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

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 $\alpha_1^\star = 2$.

The required condition (iii) on the α-mixing coefficients is satisfied, similarly as in the proof of Theorem 3.1, by equation (8.2) below.

As in (7.2), 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

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

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

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

□

8. α-mixing of the Brown-Resnick space-time process

We define α-mixing for spatial processes; see e.g. Doukhan [9] or Bolthausen [1].

***Definition* 8.1.** For $d \in \mathbb{N}$, consider a strictly stationary process $\left\{X(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\right\}$ 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

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

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

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

$$\alpha_{k,l}(r) = \sup \left\{|\mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1)\mathbb{P}(A_2)| : \ A_i \in \mathcal{F}_{\Lambda_i}, |\Lambda_1| \leq k, |\Lambda_2| \leq l, d(\Lambda_1, \Lambda_2) \geq r\right\}, \tag{8.1}$$

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 process Corollary 2.2 of Dombry and Eyi-Minko [7] shows that the α-mixing coefficients can be related to the extremogram of the max-stable process. Equations (8.2) and (8.3) follow as in the proofs of Proposition 1 and 2 of Buhl and Klüppelberg [3].

***Proposition* 8.2.** For all fixed time points $t \in \mathbb{N}$ the random field $\left\{\eta(\boldsymbol{s}, t), \boldsymbol{s} \in \mathbb{Z}^2\right\}$ defined by (4.1) is α-mixing with mixing coefficients satisfying

$$\alpha_{k,l}(r) \leq 2kl \sup_{s \geq r} \chi(s, 0) \leq 4kl e^{-\theta_1 r^{\alpha_1}/2}, \quad k, l \in \mathbb{N}, \ r \geq 0. \tag{8.2}$$

For all fixed locations $\boldsymbol{s} \in \mathbb{R}^2$ the time series $\{\eta(\boldsymbol{s}, t) : t \in [0, \infty)\}$ in (4.1) is α-mixing with mixing coefficients satisfying for some constant $c > 0$

$$\alpha(r) := \alpha_{\infty,\infty}(r) \leq c \sum_{u=r}^{\infty} u e^{-\theta_2 u^{\alpha_2}/2}, \quad r \geq 0. \tag{8.3}$$

We will make frequent use of the following simple result.

***Lemma* 8.3.** Let $z \in \mathbb{N}$. For $(\theta, \alpha) \in \{(\theta_1, \alpha_1), (\theta_2, \alpha_2)\}$ and sufficiently large r such that the sequence $u^z e^{-\theta u^\alpha/2}$ is decreasing for $u \geq r$, 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

$$g_z(r) = r^z e^{-\theta r^\alpha/2} + \sum_{u=r+1}^{\infty} u^z e^{-\theta u^\alpha/2} \leq r^z e^{-\theta r^\alpha/2} + \int_r^\infty u^z e^{-\theta u^\alpha/2} du$$

$$
\begin{aligned}
&= r^z e^{-\theta r^\alpha/2} + \left(\frac{2}{\theta}\right)^{(z+1)/\alpha} \frac{1}{\alpha} \int_{\theta r^\alpha/2}^{\infty} t^{(z+1)/\alpha - 1} e^{-t} dt \\
&\le r^z e^{-\theta r^\alpha/2} + c_1 \Gamma\left(\lceil (z+1)/\alpha \rceil, \theta r^\alpha/2\right) \\
&= r^z e^{-\theta r^\alpha/2} + c_1 \left(\lceil (z+1)/\alpha \rceil - 1\right)!\; e^{-\theta r^\alpha/2} \sum_{k=0}^{\lceil (z+1)/\alpha \rceil - 1} \frac{\theta^k r^{\alpha k}}{2^k k!} \\
&\le r^z e^{-\theta r^\alpha/2} + c_2 e^{-\theta r^\alpha/2} r^{\alpha(\lceil (z+1)/\alpha \rceil - 1)} \\
&\le c e^{-\theta r^\alpha/2} r^{z+1},
\end{aligned}
$$

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. □

9. Robustness of the bias corrected estimator

As shown in the simulation study in Section 10 below, the WLSEs are robust with respect to small deviations from the model assumptions. Specifically, if one adds measurement noise to the underlying Brown-Resnick process, the WLSEs still perform well. This is in contrast to the composite likelihood procedure for which the estimates become biased. The theoretical foundation for the good performance of the WLSEs is given in Lemma 9.1, which is the analogue of Lemma 4.2 for the Brown-Resnick process without noise.

***Lemma* 9.1.** Let $\{Z(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^2 \times [0,\infty)\}$ be i.i.d. random variables which are independent of the space-time process $\{\eta(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^2 \times [0,\infty)\}$. Assume the moment condition $E|Z(\mathbf{0},0)|^{2+\epsilon} < \infty$ for some $\epsilon > 0$. Then for every sequence $a_n \to \infty$ we have for fixed $t \in [0,\infty)$,

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

Proof. For notational simplicity, write $\eta_1 = \eta(\boldsymbol{s},t)$, $\eta_2 = \eta_(\boldsymbol{s}+\boldsymbol{h},t), Z_1 = Z(\boldsymbol{s},t), Z_2 = Z(\boldsymbol{s}+\boldsymbol{h},t)$, and $\chi = \chi(\|\boldsymbol{h}\|)$. Here we assume that $\boldsymbol{h} \neq \mathbf{0}$, since otherwise, $\chi = 1$ and thus

$$
0 = (\chi - 2)(\chi - 1) = \left(\frac{\mathbb{P}(\eta_1 + Z_1 > a_n)}{\mathbb{P}(\eta_1 + Z_1 > a_n)} - \chi\right).
$$

Using (4.3) and the independence of (η_1, η_2) with (Z_1, Z_2), we have

$$
\begin{aligned}
& \mathbb{P}(\eta_1 + Z_1 > a_n, \eta_2 + Z_2 > a_n) \\
= \;& 1 - \mathbb{P}(\eta_1 + Z_1 \le a_n) - \mathbb{P}(\eta_2 + Z_2 \le a_n) + \mathbb{P}(\eta_1 + Z_1 \le a_n, \eta_2 + Z_2 \le a_n) \\
= \;& 2\mathbb{P}(\eta_1 + Z_1 > a_n) - (1 - \mathbb{P}(\eta_1 + Z_1 \le a_n, \eta_2 + Z_2 \le a_n))
\end{aligned}
$$

$$= \quad 2\mathbb{P}(\eta_1 + Z_1 > a_n) - \mathbb{E}\Big[1 - \exp\Big\{-\frac{1}{(a_n - Z_1)_+}\Phi^n_{1,2} - \frac{1}{(a_n - Z_2)_+}\Phi^n_{2,1}\Big\}\Big]\,,$$

where $x_+ = \max\{0, x\}$, $\Phi^n_{i,j} = \Phi(c\log((a_n - Z_j)_+/(a_n - Z_i)_+) + c/2)$, and $c = \sqrt{2\delta(\|h\|)}$. Set $\Phi^* = \lim_{n\to\infty}\Phi^n_{i,j} = \Phi(c/2)$ *a.s.*

Take $b_n = a_n^{1-\epsilon/4}$, where $\epsilon \in (0,1)$ is specified in the statement of the lemma. Then it follows that $b_n/a_n \to 0$, $a_n^2/b_n^{2+\epsilon} = a_n^{-(2-\epsilon)\epsilon/4} \to 0$ and hence

$$\mathbb{P}(|Z| \geq b_n) \leq \frac{\mathbb{E}|Z|^{2+\epsilon}}{b_n^{2+\epsilon}} = o(a_n^{-2})\,.$$

Writing $\mathbb{E}_n$ for expectation relative to the restriction on the event $\{|Z_1| \vee |Z_2| \leq b_n\}$, we have for any bounded sequence of random variables Y_n that $a_n^2(\mathbb{E}Y_n - \mathbb{E}_nY_n) \to 0$. Hence, using a Taylor series approximation, we obtain

$$\begin{aligned}
\mathbb{P}(\eta_1 + Z_1 > a_n) &= \mathbb{E}\Big[1 - \exp\Big\{-\frac{1}{(a_n - Z_1)_+}\Big\}\Big] \\
&= \mathbb{E}_n\frac{1}{(a_n - Z_1)_+} - \mathbb{E}_n\frac{1}{2(a_n - Z_2)^2} + o(a_n^{-2}) \\
&= \mathbb{E}_n\frac{1}{(a_n - Z_1)_+} - a_n^{-2} + o(a_n^{-2}) \qquad (9.1)
\end{aligned}$$

and

$$\begin{aligned}
I_1 &:= \mathbb{E}\Big[1 - \exp\Big\{-\frac{1}{(a_n - Z_1)_+}\Phi^n_{1,2} - \frac{1}{(a_n - Z_2)_+}\Phi^n_{2,1}\Big\}\Big] \\
&= \mathbb{E}_n\Big[\frac{1}{(a_n - Z_1)_+}\Phi^n_{1,2} + \frac{1}{(a_n - Z_2)_+}\Phi^n_{2,1}\Big] \\
&\qquad -\frac{1}{2}\mathbb{E}_n\Big[\frac{1}{(a_n - Z_1)_+}\Phi^n_{1,2} + \frac{1}{(a_n - Z_2)_+}\Phi^n_{2,1}\Big]^2 + o(a_n^{-2}) \\
&= 2\mathbb{E}_n\Big[\frac{1}{(a_n - Z_1)_+}\Phi^n_{1,2}\Big] - \mathbb{E}_n\Big[\frac{1}{(a_n - Z_1)_+}\Phi^n_{1,2}\Big]^2 \\
&\qquad -\mathbb{E}_n\Big[\frac{1}{(a_n - Z_1)_+}\frac{1}{(a_n - Z_2)_+}\Phi^n_{1,2}\Phi^n_{2,1}\Big] + o(a_n^{-2}).
\end{aligned}$$

In order to complete the proof it suffices to show the following two relations:

$$2\mathbb{E}_n\Big[\frac{1}{(a_n - Z_1)_+}\Phi^n_{1,2}\Big] - 2\Phi^*\mathbb{P}(\eta_1 + Z_1 > a_n) + \frac{\Phi^*}{a_n^2} = o(a_n^{-2}), \qquad (9.2)$$

$$\begin{aligned}
\mathbb{E}_n\Big[\frac{1}{(a_n - Z_1)_+}\Phi^n_{1,2}\Big]^2 &+ \mathbb{E}_n\Big[\frac{1}{(a_n - Z_1)_+}\frac{1}{(a_n - Z_2)_+}\Phi^n_{1,2}\Phi^n_{2,1}\Big] \\
&= 2\frac{(\Phi^*)^2}{a_n^2} + o(a_n^{-2}). \qquad (9.3)
\end{aligned}$$

To see that (9.2) and (9.3) yield the result in Lemma 9.1, observe that

$$\begin{aligned}\frac{\mathbb{P}(\eta_1 + Z_1 > a_n, \eta_2 + Z_2 > a_n)}{\mathbb{P}(\eta_1 + Z_1 > a_n)} &= 2 - \frac{I_1}{\mathbb{P}(\eta_1 + Z_1 > a_n)} \\ &= 2 - 2\Phi^* + \frac{\Phi^* - 2(\Phi^*)^2}{a_n^2 \mathbb{P}(\eta_1 + Z_1 > a_n)} + o_(a_n^{-1}) \\ &= \chi + \frac{1}{2} a_n^{-1} (\chi - 2)(\chi - 1) + o(a_n^{-1}),\end{aligned}$$

where we have used the properties $\chi = 2 - 2\Phi^*$ and $a_n \mathbb{P}(\eta_1 + Z_1 > a_n) \to 1$..

Using a Taylor series expansion and the relation

$$\log \frac{(a_n - Z_2)_+}{(a_n - Z_1)_+} = \frac{Z_1 - Z_2}{a_n} + (Z_1 - Z_2)^2 \cdot O(a_n^{-2}),$$

it follows that on the set $\{|Z_1| \vee |Z_2| \leq b_n\}$,

$$\begin{aligned}\Phi_{1,2}^n - \Phi^* &= c \log \frac{(a_n - Z_2)_+}{(a_n - Z_1)_+} \Phi'(c/2) + O_p(a_n^{-2}) \\ &= c \frac{Z_1 - Z_2}{a_n} \Phi'(c/2) + O_p(a_n^{-2}).\end{aligned}$$

Finally turning to (9.2) and applying (9.1), the left-hand side is equal to

$$\begin{aligned}&2\mathbb{E}_n \Big[\frac{1}{(a_n - Z_1)_+} (\Phi_{1,2}^n - \Phi^*) \Big] + \frac{\Phi^*}{a_n^2} - o(a_n^{-2}) \\ = \; &c\mathbb{E}_n \frac{Z_1 - Z_2}{a_n(a_n - Z_1)} + \Phi'(c/2) \frac{\Phi^*}{a_n^2} - o(a_n^{-2}),\end{aligned}$$

which by multiplying by a_n^2 and taking limits gives the desired limit of Φ^*.

Finally, (9.3) is obtained by multiplying both sides of the equation a_n^2 and taking limits, where the interchange of limits and expectation are justified by the dominating convergence theorem. This completes the proof. □

10. Simulation study

We examine the performance of the WLSEs by simulating a large number of Brown-Resnick processes with dependence function (7.3). Many real data may not follow a Brown-Resnick process precisely. For a more realistic setting, we thus do not perform the simulation study for a Brown-Resnick process only, but also for the sum of a Brown-Resnick process and some noise. This sum is regularly varying and possesses the same dependence function and extremogram for which all results of Sections 2 and 3 hold. As we want to use the bias reduction procedure of Section 4, we have to make sure that Lemma 4.2 extends to this setting; all other results of this section follow from that. This is guaranteed by Lemma 9.1.

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. [8] 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 estimates, 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 detail as follows:

1. We choose two space-time observation areas

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

   and the sets $\mathcal{V} = \{1, \sqrt{2}, 2, \sqrt{5}, \sqrt{8}, 3, \sqrt{10}, \sqrt{13}, 4, \sqrt{17}\}$ and $\mathcal{U} = \{1, \dots, 10\}$.
2. We simulate the Brown-Resnick space-time process (4.1) based on the exact method proposed in Dombry et al. [8], using the `R`-package `RandomFields` [11]. The dependence function δ is modelled as in (4.2) (cf. (7.3)); 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 (2.5)-(2.8)) 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. [6] 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 temporal 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 (2.9) and (2.10)) 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 $\chi(0, u)$ given in (4.4), which are tail probabilities of the standard normal distribution Φ.

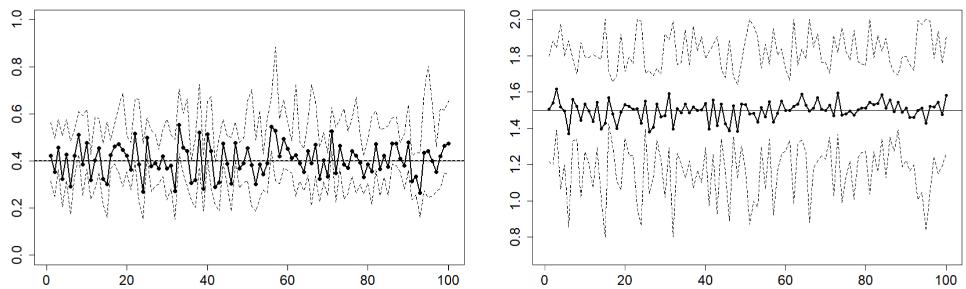

Figure 10.1: WLSEs of θ_1 (left) and α_1 (right) for 100 simulated Brown-Resnick space-time processes together with pointwise 95%−subsampling confidence intervals (dashed). The middle solid line is the true value and the middle dashed line represents the mean over all estimates.

4. Pointwise confidence bounds are computed by subsampling as described in Section 7 for the spatial parameters of general regularly varying processes and in Section 3.4.2 of [2] for the temporal parameters of the Brown-Resnick space-time process considered in this section. We choose block lengths $\boldsymbol{b} = (50, 50, 10)$ and overlap $\boldsymbol{e} = (2, 2, 10)$ for the space-time process with observation area $\mathcal{S}^{(1)} \times \mathcal{T}^{(1)}$ and $\boldsymbol{b} = (5, 5, 200)$, $\boldsymbol{e} = (5, 5, 1)$ for the process with observation area $\mathcal{S}^{(2)} \times \mathcal{T}^{(2)}$.
5. Steps 1 - 5 are repeated 100 times.

Figure 10.1 shows the WLSEs of the spatial parameters θ_1 and α_1 for each of the 100 realizations of the Brown-Resnick space-time process. The dashed lines above and below the dots are pointwise confidence intervals based on subsampling. Panel (a) of Table 1 shows the mean, root mean squared error (RMSE) and mean absolute error (MAE) of both the spatial and the temporal WLSEs based on the 100 simulations. Altogether, we observe that the estimates are close to the true values. The spatial estimates are slightly superior to 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

(a)

	MEAN	RMSE	MAE
θ_1	0.4008	0.0668	0.0552
α_1	1.4946	0.0525	0.0400
θ_2	0.2188	0.0597	0.0489
α_2	0.9275	0.0976	0.0799

(b)

Table 1.: Mean, RMSE and MAE of the WLSEs when applied to exact realizations from the Brown-Resnick process (a) and to realizations with observational noise (b).

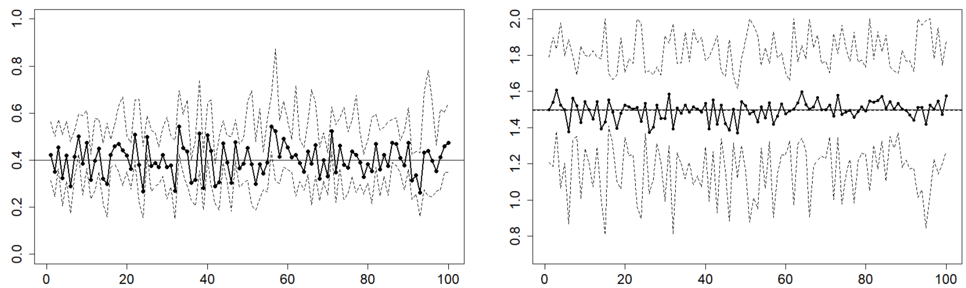

Figure 10.2: WLSEs of θ_1 (left) and α_1 (right) for 100 simulated Brown-Resnick space-time processes with noise together with pointwise 95%−subsampling confidence intervals (dashed). The middle solid line is the true value and the middle dashed line represents the mean over all estimates.

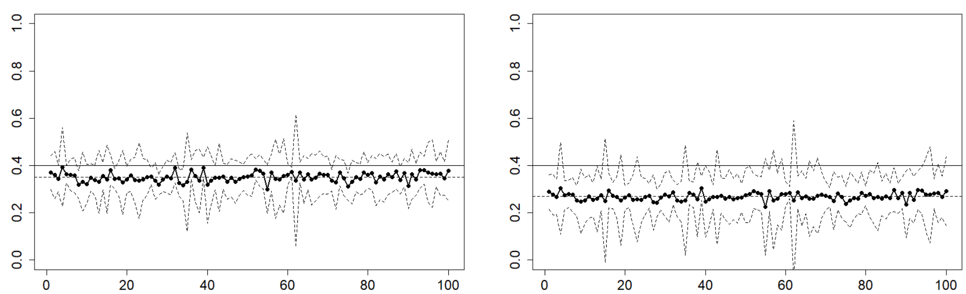

Figure 10.3: Pairwise likelihood estimates of θ_1 for 100 simulated Brown-Resnick space-time processes together with pointwise 95%−subsampling confidence intervals (dashed) in the left-hand plot; the corresponding estimates and pointwise 95%−subsampling confidence intervals (dashed) for the simulated processes with noise are presented in the right-hand plot. The middle solid line is the true value and the middle dashed line represents the mean over all estimates.

In comparison with pairwise likelihood estimation in finite samples (cf. Davis et al. [5]), a big advantage of the semiparametric method is the substantial reduction of computation time by about a factor 15. Moreover, the semiparametric estimation method is much more robust when applied to observations that reveal slight deviations from the model assumptions. To illustrate this point, we repeated the simulation study described above with data obtained from the original ones by adding to each measurement the absolute value of an independent $\mathcal{N}(0, 0.2)$-distributed error. The results of the semiparametric estimation remain practically unaffected; see the summary measures in the panel (b) of Table 1 and Figure 10.2. This result can be explained theoretically by Lemma 9.1. Adding noise to observations of the Brown-Resnick process does not affect the underlying true extremogram nor the rate of convergence of the empirical extremogram to the true one. In contrast, when applying pairwise likelihood estimation to the same simulated data we observe that the estimates are much more sensitive to small disturbations than the semiparametric estimation. Whereas for the original data the estimates are slightly biased, for the corrupted data the bias increases considerably; their variances, however, remain nearly unaffected and are (not surprisingly) smaller than the corresponding variances of the semiparametric estimates. Figure 10.3 illustrates the pairwise likelihood estimates of the spatial parameter θ_1 for the simulated Brown-Resnick space-time processes together with 95%-subsampling confidence intervals and for the data with noise.

References

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

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

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

[4] S. Buhl, R.A. Davis, C. Klüppelberg, and C. Steinkohl. Semiparametric estimation for isotropic max-stable space-time processes. *Bernoulli*, 2018.

[5] R.A. Davis, C. Klüppelberg, and C. Steinkohl. Statistical inference for max-stable processes in space and time. *JRSSB*, 75(5):791–819, 2013.

[6] 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.

[7] 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.

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

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

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

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