Roadmap Max

Max Thannheimer & Marco Oesting

July 2023

1 Abstract

Bayesian inference for functional extreme events defined via partially unobserved processes In order to describe the extremal behaviour of some stochastic process X approaches from univariate extreme value theory are typically generalized to the spacial domain. Besides max-stable processes, that can be used in analogy to the block maxima approach, a generalized peaks-over-threshold approach can be used, allowing us to consider single extreme events. These can be flexibly defined as exceedances of a risk functional ℓ, such as a spatial average, applied to X. Inference for the resulting limit process, the so-called ℓ-Pareto process, requires the evaluation of $\ell(X)$ and thus the knowledge of the whole process X. In practical application we face the challenge that observations of X are only available at single sites. To overcome this issue, we propose a two-step MCMC-algorithm in a Bayesian framework. In a first step, we sample from X conditionally on the observations in order to evaluate which observations lead to ℓ-exceedances. In a second step, we use these exceedances to sample from the posterior distribution of the parameters of the limiting ℓ-Pareto process. Alternating these steps results in a full Bayesian model for the extremes of X. We show that, under appropriate assumptions, the probability of classifying an observation as ℓ-exceedance in the first step converges to the desired probability. Furthermore, given the first step, the distribution of the Markov chain constructed in the second step converges to the posterior distribution of interest. Our procedure is compared to the Bayesian version of the standard procedure in a simulation study.

2 Basics

Let $S \subset \mathbb{R}^d$ be compact and $C(S)^+$ the Banach space of non-negative continuous real-valued functions $f : S \to [0, \infty)$ (without the zero function) equipped with norm $\|f\|_\infty := \sup_{s\in S} |f(S)|$ and the σ-algebra generated by cylinder sets. Consider a sample-continuous process $X = \{X(s) : s \in S\}$ which is in the max-domain of some max-stable process $Z = \{Z(s) : s \in S\}$ of α-Fréchet margins for some $\alpha > 0$, i.e. there exist continuous functions $a_n : S \to (0, \infty)$ and $b_n : S \to \mathbb{R}$ such that

$$\left\{\max_{i=1}^{n} \frac{X_i(s) - b_n(s)}{a_n(s)}, s \in S\right\} \to_d \{Z(s), s \in S\}$$

weakly in $C(S)^+$ for independent stochastic processes $X_1, X_2, \dots$ with the same law as X. Here the choices $b_n \equiv 0$ and

$$a_n(s) := \inf\left\{x \geq 0 : \mathbb{P}(X(s) \leq x) \leq 1 - \frac{1}{n}\right\} \tag{1}$$

can be made.

By de Haan (1984), the process Z allows for the spectral representation

$$Z(s) = \bigvee_{k=1}^{\infty} \zeta_k W_k(s), \quad s \in S,$$

where $\{U_k\}_{k\in\mathbb{N}}$ are the points of a Poisson point process with intensity measure $\alpha\zeta^{-\alpha-1}\mathrm{d}\zeta$ and W_k are independent copies of a non-negative sample-continuous stochastic process $W = \{W(s) : s \in S\}$, the so-called *spectral process*, satisfying $\mathbb{E}(W(s)^\alpha) = 1$ for all $s \in S$. Note that the law of the spectral process W is not unique, but, for instance, can be normalized with respect to different functionals.

ℓ-normalized spectral representation Let $\ell : C(S)^+ \to \mathbb{R}$ be a non-trivial homogeneous continuous functional. Then, there exists a non-negative sample-continuous stochastic process $W^{(\ell)} = \{W^{(\ell)}(s),\, s \in S\}$ satisfying $\ell(W^{(\ell)}) = 1$ almost surely such that

$$Z(s) =_d \sqrt[\alpha]{c_\ell} \bigvee_{k=1}^{\infty} \zeta_k \frac{W_k^{(\ell)}(s)}{\ell(W_k^{(\ell)})}, \quad s \in S, \tag{2}$$

for independent copies $W_1^{(\ell)}, W_2^{(\ell)}, \dots$ of $W^{(\ell)}$ and

$$c_\ell = \int_{C(S)^+} \ell(w)^\alpha \mathrm{d}\mathbb{P}_W(w).$$

The law of $W^{(\ell)}$ is uniquely given by

$$\mathbb{P}(W^{(\ell)} \in A) = \frac{1}{c_\ell} \int_{C(S)^+} \ell(w)^\alpha \mathbb{1}_{\{w\in A\}} \mathrm{d}\mathbb{P}_W(w), \quad A \in \mathcal{B}\left(\mathcal{C}(S)^+\right). \tag{3}$$

From the definition of $W^{(\ell)}$ follows

$$\mathbb{E}\left(g(W^{(\ell)}/\ell(W^{(\ell)}))\right) = \int_{C(S)^+} g(f/\ell(f))\mathrm{d}\mathbb{P}_{W^{(\ell)}}(f) = \frac{1}{c_\ell}\int_{C(S)} \ell(f)^\alpha g\left(\frac{f}{\ell(f)}\right)\mathrm{d}\mathbb{P}_W(f). \tag{4}$$

By decomposing X into a ℓ normalized part $\frac{X}{\ell(X)}$ and an intensity part $\ell(X)$ we have for a_n as in (1) that

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X}{\ell(X)} \in B, \ell(X) > ra_n \mid \ell(X) > a_n\right) = \mathbb{P}\left(\frac{W^{(\ell)}}{\ell(W^{(\ell)})} \in B\right)\mathbb{P}\left(P_\alpha > r\right). \tag{5}$$

3 Note on spectral functions for Brown–Resnick Processes

Standard choice for W:

$$W(t) = \exp\left(G(t) - \frac{1}{2}\mathrm{Var}(G(t))\right), \quad t \in \mathbb{R}^d,$$

where G is a centered Gaussian process. It does not matter which centered Gaussian process you choose as long as it possesses the "right" semi-variogram

$$\gamma(h) = \frac{1}{2}\mathrm{Var}(G(h) - G(0)), \quad h \in \mathbb{R}^d.$$

Starting with any centered Gaussian process $\widetilde{G}$ with semi-variogram γ, consider the Gaussian process G obtained by

$$G(t) = \widetilde{G}(t) - \widetilde{G}(s_0), \quad t \in \mathbb{R}^d.$$

This process is obviously centered with the same semi-variogram γ and satisfies $G(s_0) \equiv 0$. Its variance can be easily calculated by noting that

$$\mathrm{Var}(G(t)) = \mathrm{Var}(\widetilde{G}(t) - \widetilde{G}(s_0)) = 2\gamma(t - s_0), \qquad t \in \mathbb{R}^d.$$

For the semi-variogram γ a convenient choice is

$$\gamma(h) = c||h||^\beta, \ h \in \mathbb{R}^d$$

with $\beta \in (0, 2)$ and $c > 0$. Further some anisotropy matrix can be introduced. The normalization of W at s_0

$$G(s_0) = 0 \Leftrightarrow W(s_0) = 1$$

translates via (3) to $W^{(\ell)}$, namely

$$\begin{aligned}\mathbb{P}(W^{(\ell)} \in \{f : f(s_0) = 1\}) &= \frac{1}{c_\ell}\int_{C(S)^+} \ell(w)^\alpha \mathbb{1}_{\{w \in \{f : f(s_0)=1\}\}} \mathrm{d}\mathbb{P}_W(w) \\ &\overset{W(s_0)=1 \text{ a.s.}}{=} \frac{1}{c_\ell}\int_{C(S)^+} \ell(w)^\alpha \mathrm{d}\mathbb{P}_W(w) = \frac{c_\ell}{c_\ell} = 1.\end{aligned}$$

4 Another note on spectral functions for Brown–Resnick Processes

The circulant embedding approach gives us simulations of a centered Gaussian process $\left\{\tilde{G}(s) : s \in S\right\}$ with the nonstationary covariance function

$$\mathrm{Cov}(\tilde{G}(s), \tilde{G}(t)) = ||s||^\beta + ||t||^\beta - ||s - t||^\beta$$

with $\tilde{G}(0) = 0$ on the fine grid. For the semi-variogram of intrinsic stationary random fields like $\tilde{G}$ holds

$$\mathrm{Cov}(\tilde{G}(s), \tilde{G}(t)) = \gamma(s)^\beta + \gamma(t)^\beta - \gamma(s - t)^\beta$$

and for the semi-variogram γ of $\tilde{G}$ holds $\gamma(h) = ||h||^\beta, \ h \in \mathbb{R}^d, \beta \in (0, 2)$. Now we shift the process $\tilde{G}$ in one of our observation sites and also multiply a parameter $\sqrt{c}, c > 0$ to get the process G with the desired semi-variogram $\gamma(h) = c||h||^\beta, \ h \in \mathbb{R}^d$ normalized to $G(s_0) = 0$.
Then we can use that to construct the Brown-Resnick spectral process as

$$W(t) = \exp\left(G(t) - \frac{1}{2}\mathrm{Var}(G(t))\right), \quad t \in \mathbb{R}^d,$$

5 Markov Chain Monte Carlo

Markov Chains Let E be a subspace of $\mathbb{R}^k$ for some $k > 0$. Let $\mathcal{B}(E)$ be the Borel σ-algebra on E. A **Markov chain** with stationary distribution π on E is a sequence of E valued random variables $\{M_n\}_{n\geq 0}$ such that for the Markov transition kernel P defined via

$$P(M_n, A) = P(M_{n+1} \in A \mid M_0, \dots, M_n)$$

holds

$$\pi(A) = \int P(x, A)\pi(dx), \quad A \in \mathcal{B}(E).$$

We assume that π has a density with respect to some σ-finite measure μ, i.e. $\pi(dx) = \pi(x)\mu(dx)$. Note that π denotes the density and distribution at the same time. The distribution of M_0 is called initial distribution. With P^n denoting the n-th iterate of the Markov transition kernel P we can write the distribution of the M_n conditional on M_0 as

$$P(M_n \in A \mid M_0) = P^n(M_0, A).$$

We are now interested under which conditions the convergence

$$\lim_{n\to\infty} ||P^n(x, \cdot) - \pi||_{\mathrm{TV}} = 0,$$

holds for $x \in E$ with $|| \cdot ||_{\mathrm{TV}}$ denoting the total variation norm.
We call a Markov transition kernel π**-irreducible** if $\pi(E) > 0$ and for all $x \in E$ and each $A \in \mathcal{B}(E)$ with $\pi(A) > 0$ exists a $n > 0$ such that $P^n(x, A) > 0$.
A π-irreducible Markov transition kernel P is called **periodic** if there exists a sequence of disjoint sets $\{E_0, \dots, E_{d-1}\}$ in $\mathcal{B}(E)$ such that

$$P(x, E_j) = 1 \text{ for } j = i + 1 (\text{mod } d),$$

for all $i = 0, \dots, d-1, x \in E_i$. Else it is called **aperiodic**.

Markov Chain Monte Carlo The idea now is to construct a Markov Chain which has the right stationary distribution π and guaranteed convergence. Therefore we define a Markov transition kernel

$$Q(x, dy) = q(x, y)\mu(dy)$$

and call $q(\cdot, \cdot)$ our proposal density. We assume that on $E^+ = \{x : \pi(x) > 0\}$ holds $Q(x, E^+) = 1$ for $x \notin E^+$. Then let

$$\alpha(x, y)$$

Bla bla chain via proposal, accept uniformly with rate α. Then right stationary distrib. no problems with multiplicative constant. Bla bla metropolis kernel P
For a Metropolis kernels holds the following theorem (compare Theorem 1 and Corollary 2 in Tierney (1994)).

Theorem 1. *Suppose P is π-irreducible Metropolis kernel with stationary distribution π. Then π is unique. If P is also aperiodic, then*

$$\lim_{n\to\infty} ||P^n(x, \cdot) - \pi||_{TV} = 0,$$

for all $x \in E$.

Metropolis kernel P is π-irreducible, since
Metropolis kernel P is aperiodic, since

6 First setting

In a first step we assume to be able to get an exact evaluation of ℓ. Further let $\mathbf{s} = (s_1, \ldots, s_n)$ be a set of sites where we have observations and define $W(\mathbf{s}) = (W(s_1), \ldots, W(s_n))$. When density of W is known we have from the definition of $W^{(\ell)}$ that

$$f_{W^{(\ell)}(\mathbf{s})}(\mathbf{x}) = \frac{1}{c_\ell} \ell(\mathbf{x})^\alpha f_{W(s)}(\mathbf{x}). \tag{6}$$

Now introduce MCMC setting and parameterization of W as W_θ with $\theta \in \Theta \subset \mathbb{R}^k$ and proposal density $q(\cdot, \theta)$ on parameter space Θ and conditional on θ the likelihood

$$f_{W^{(\ell)}(\mathbf{s})}(\mathbf{x}|\theta) = \frac{1}{c_{\ell,\theta}} \ell(\mathbf{x})^\alpha f_{W(s)}(\mathbf{x}|\theta). \tag{7}$$

Then there holds convergence of typical metropolis hastings MCMC when $q(\cdot, \theta) > 0$ (irreducible guaranteed) and by the construction also aperiodic. Further show that unbiased estimator of $\frac{1}{c_{\ell,\theta}}$ is sufficient to guarantee convergence.

7 Sampling

Now let X be in the domain of attraction of a max-stable process and $X(\mathbf{s}) := (X(s_0), \ldots, X(s_n))$ an observation on sites $\mathbf{s} := (s_0, \ldots, s_n)$, $s_i \in S \subset \mathbb{R}^d$. Further we assume an additional structure for the observed process:

$$\frac{X}{\ell(X)} \stackrel{d}{=} \frac{W^{(\ell)}}{\ell(W^{(\ell)})} \tag{8}$$

where W is the spectral process of the Brown–Resnick process normalized such that $W(s_0) = 1$ and $W^{(\ell)}$ is the process defined by (3). This assumption holds in the limit via (5)

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X}{\ell(X)} \in B, \ell(X) > r a_n \mid \ell(X) > a_n \right) = \mathbb{P}\left(\frac{W^{(\ell)}}{\ell(W^{(\ell)})} \in B \right) \mathbb{P}\left(P_\alpha > r\right).$$

From the assumption (8) we get

$$\frac{X(\tilde{\mathbf{s}})}{X(s_0)} = \frac{X(\tilde{\mathbf{s}})/\ell(X)}{X(s_0)/\ell(X)} \stackrel{d}{=} \frac{\frac{W^{(\ell)}(\tilde{\mathbf{s}})}{\ell(W^{(\ell)})}/\ell(X)}{\frac{W^{(\ell)}(s_0)}{\ell(W^{(\ell)})}/\ell(X)} = \frac{W^{(\ell)}(\tilde{\mathbf{s}})}{W^{(\ell)}(s_0)} = W^{(\ell)}(\tilde{\mathbf{s}}), \tag{9}$$

where $\tilde{\mathbf{s}} := (s_1, \ldots, s_n)$, $s_i \in S \subset \mathbb{R}^d$.
We now want to sample in a two-step-algorithm to estimate the parameters θ of the parametric model of W as well as the Fréchet parameter α. In the first step we assume them to be fixed and want to sample from the process X such that we can estimate which observations of X lead to a large value of $\ell(X)$. Knowing the extreme observations of X, we assume them to be samples of the limit process via

$$\mathbb{P}\left(\frac{X}{\ell(X)} \in B, \ell(X) > r \cdot \text{threshold} \;\middle|\; \ell(X) > \text{threshold} \right) \approx \mathbb{P}\left(W^{(\ell)} \in B \right) \mathbb{P}\left(P_\alpha > r\right).$$

and can estimate the parameters. i.e. sample from the posterior!!!

7.1 Sampling from X

To get a good approximation of $\ell(X)$ we want to sample X on a denser set of sites $\mathbf{t} := (t_1, \ldots, t_N)$, $t_i \in S \subset \mathbb{R}^d$ with $N \geq n$ large enough that an approximation $\ell(X(\mathbf{t})) := \tilde{\ell}(X(t_1), \ldots, X(t_N)) \approx \ell(X)$. For now we assume equality.
To achieve our goal of sampling from X conditioned on the observations $X(s_0), \ldots, X(s_n)$

$$X \mid X(s_0), \ldots, X(s_n)$$

we sample firstly from

$$
\begin{aligned}
& W \mid W(s_0) = \frac{X(s_0)}{X(s_0)}, \dots, W(s_n) = \frac{X(s_n)}{X(s_0)} \\
\overset{W(s_0)=1}{=} & W \mid W(s_1) = \frac{X(s_1)}{X(s_0)}, \dots, W(s_n) = \frac{X(s_n)}{X(s_0)}
\end{aligned}
$$

Therefore we take a look at the conditional distributions of $W^{(\ell)}$. The joint distribution on the observations and the finer grid fulfils

$$
\begin{aligned}
& \mathbb{P}\left(W^{(\ell)}(\mathbf{t}) \in A, W^{(\ell)}(\mathbf{s}) \in B\right) \\
\overset{\ell((X(\mathbf{t}))=\ell(X)\text{a.s.}}{=} & \frac{1}{c_\ell} \int_{C(S)^+} \ell(\mathbf{z})^\alpha \mathbb{1}_{\{\mathbf{z}\in A\}} \mathbb{1}_{\{\mathbf{x}\in B\}} \mathbb{P}_{(W(\mathbf{t}),W(\mathbf{s}))}(\mathrm{d}(\mathbf{z},\mathbf{x})) \\
= & \frac{1}{c_\ell} \int_B \int_A \ell(\mathbf{z})^\alpha \mathbb{P}_{W(\mathbf{t})\mid W(\mathbf{s})=\mathbf{x}}\left(\mathrm{d}(\mathbf{z})\right) \mathbb{P}_{W(\mathbf{s})}\left(\mathrm{d}(\mathbf{x})\right)
\end{aligned}
$$

and therefore the conditional density of $W^{(\ell)}(\mathbf{t}) \mid W^{(\ell)}(\tilde{\mathbf{s}}) = \tilde{\mathbf{x}}$

$$
\begin{aligned}
f_{W^{(\ell)}(\mathbf{t})\mid W^{(\ell)}(\tilde{\mathbf{s}})=\tilde{\mathbf{x}}}(\mathbf{z}) &= \frac{f_{(W^{(\ell)}(\mathbf{t}),W^{(\ell)}(\tilde{\mathbf{s}}))}(\mathbf{z},\tilde{\mathbf{x}})}{\int_{\mathbb{R}^N} f_{(W^{(\ell)}(\mathbf{t}),W^{(\ell)}(\tilde{\mathbf{s}}))}(\tilde{\mathbf{z}},\tilde{\mathbf{x}})\mathrm{d}\tilde{\mathbf{z}}} \\
&= \frac{1}{c_\ell}\ell^\alpha(\mathbf{z}) f_{W(\mathbf{t})\mid W(\tilde{\mathbf{s}})=\tilde{\mathbf{x}}}(\mathbf{z}) \underbrace{\frac{f_{W(\tilde{\mathbf{s}})}(\tilde{\mathbf{x}})}{\int_{\mathbb{R}^N} f_{(W^{(\ell)}(\mathbf{t}),W^{(\ell)}(\tilde{\mathbf{s}}))}(\tilde{\mathbf{z}},\tilde{\mathbf{x}})\mathrm{d}\tilde{\mathbf{z}}}}_{=:C(\tilde{\mathbf{x}}) \text{ depends only on observation } \tilde{\mathbf{x}}} \\
&= \frac{1}{c_\ell} C(\tilde{\mathbf{x}}) \ell^\alpha(\mathbf{z}) f_{W(\mathbf{t})\mid W(\tilde{\mathbf{s}})=\tilde{\mathbf{x}}}(\mathbf{z}), \quad \mathbf{z} \in \mathbb{R}^N.
\end{aligned}
$$

We choose $\tilde{\mathbf{x}}$ as justified by our assumption (8) and (9) as our observations

$$
\tilde{\mathbf{x}} = \left(\frac{X(s_1)}{X(s_0)}, \dots, \frac{X(s_n)}{X(s_0)}\right).
$$

Then MCMC with acceptance rate

$$
\begin{aligned}
\alpha(\mathbf{v},\tilde{\mathbf{v}}) : &= \text{change } v \text{ and } \tilde{v} \min\left\{\frac{f_{W^{(\ell)}(\mathbf{t})\mid W^{(\ell)}(\tilde{\mathbf{s}})=\tilde{\mathbf{x}}}(\mathbf{v}) / f_{W(\mathbf{t})\mid W(\tilde{\mathbf{s}})=\tilde{\mathbf{x}}}(\mathbf{v})}{f_{W^{(\ell)}(\mathbf{t})\mid W^{(\ell)}(\tilde{\mathbf{s}})=\tilde{\mathbf{x}}}(\tilde{\mathbf{v}}) / f_{W(\mathbf{t})\mid W(\tilde{\mathbf{s}})=\tilde{\mathbf{x}}}(\tilde{\mathbf{v}})}, 1\right\} \\
&= \min\left\{\frac{\frac{1}{c_\ell}C(\tilde{\mathbf{x}})\ell^\alpha(\exp(\mathbf{v}))}{\frac{1}{c_\ell}C(\tilde{\mathbf{x}})\ell^\alpha(\exp(\tilde{\mathbf{v}}))}, 1\right\} \\
&= \min\left\{\frac{\ell^\alpha(\exp(\mathbf{v}))}{\ell^\alpha(\exp(\tilde{\mathbf{v}}))}, 1\right\}, \qquad \mathbf{v},\tilde{\mathbf{v}} \in \mathbb{R}^N.
\end{aligned}
$$

Therefore we have a sample from process $W^{(\ell)}$. Via assumption (8) $\frac{X}{\ell(X)} \overset{d}{=} \frac{W^{(\ell)}}{\ell(W^{(\ell)})}$ we get

$$
\ell(X) = X(s_0) \cdot \frac{\ell(W^{(\ell)})}{W^{(\ell)}(s_0)} \overset{W^{(\ell)}(s_0)=1 \text{ a.s.}}{=} X(s_0) \cdot \ell(W^{(\ell)}).
$$

Since W is a log Gaussian process it can be represented as $W = \exp(V)$ where V is Gaussian. The conditioned Gaussian process is Gaussian again. We therefore only have to sample from the Gaussian process

$$
\mathbf{v} \sim V \mid V(s_1) = \log\left(\frac{X(s_1)}{X(s_0)}\right), \dots, V(s_n) = \log\left(\frac{X(s_n)}{X(s_0)}\right).
$$

and take $\mathbf{w} := \exp(\mathbf{v})$.
Now that we know $\ell(X)$ for our sample $\tilde{\mathbf{x}}$ we can determine if it is an extreme observation, i.e.

$$
\ell(X) > \text{threshold},
$$

and finally use the approximation

$$
\begin{aligned}
&\mathbb{P}\left(\frac{X}{X(s_0)} \in B \,\middle|\, \ell(X) > \text{threshold}\right) \\
&= \mathbb{P}\left(\frac{X/\ell(X)}{X(s_0)/\ell(X)} \in B \,\middle|\, \ell(X) > \text{threshold}\right) \\
&= \mathbb{P}\left(\frac{X/\ell(X)}{X(s_0)/\ell(X)} \in B, \ell(X) > 1 \cdot \text{threshold} \,\middle|\, \ell(X) > \text{threshold}\right) \\
&\approx \mathbb{P}\left(\frac{W^{(\ell)}}{W^{(\ell)}(s_0)} \in B\right) \mathbb{P}\left(P_\alpha > 1\right) \\
&= \mathbb{P}\left(\frac{W^{(\ell)}}{W^{(\ell)}(s_0)} \in B\right) \\
&= \mathbb{P}\left(W^{(\ell)} \in B\right).
\end{aligned}
$$

Now we have to introduce the parameter θ for the dependence structure in the log Gaussian process W. The next MCMC step will propose a new θ' with some proposal density $q(\theta, \cdot)$ (symmetric, > 0, I guess) We have to calculate the likelihoods for the acceptance rate. Therefore we again look at the densities. It holds

$$
\begin{aligned}
f_{W^{(\ell)}(\tilde{\mathbf{s}})}(\tilde{\mathbf{x}} \mid \theta) &= \int_{\mathbb{R}^N} f_{(W^{(\ell)}(\mathbf{t}), W^{(\ell)}(\tilde{\mathbf{s}}))}(\mathbf{z}, \tilde{\mathbf{x}} \mid \theta) \mathrm{d}\mathbf{z} \\
&= f_{W(\tilde{\mathbf{s}})}(\tilde{\mathbf{x}} \mid \theta) \frac{1}{c_{\ell,\theta}} \int_{\mathbb{R}^N} \ell(\mathbf{z})^\alpha f_{W(\mathbf{t}) \mid W(\tilde{\mathbf{s}}) = \tilde{\mathbf{x}}}(\mathbf{z} \mid \theta) \mathrm{d}\mathbf{z}.
\end{aligned}
$$

The left factor of the likelihood is just a log Gaussian density and can be directly calculated. The rest is a bit more complicated. We can estimate the integral unbiasedly via samples from the conditional density $f_{W^{(\ell)}(\mathbf{t}) \mid W^{(\ell)}(\tilde{\mathbf{s}}) = \tilde{\mathbf{x}}}$. Let $Y_i \overset{\text{i.i.d.}}{\sim} f_{W^{(\ell)}(\mathbf{t}) \mid W^{(\ell)}(\tilde{\mathbf{s}}) = \tilde{\mathbf{x}}}$, then we estimate the integral $\int_{\mathbb{R}^N} \ell(\mathbf{z})^\alpha f_{W(\mathbf{t}) \mid W(\tilde{\mathbf{s}}) = \tilde{\mathbf{x}}}(\mathbf{z} \mid \theta) \mathrm{d}\mathbf{z}$ unbiasedly via

$$
Y_{\text{int},\theta,\tilde{\mathbf{x}}} := \frac{1}{n_{\text{est}}} \sum_{i=1}^{n_{\text{est}}} \ell(Y_i)^\alpha
$$

The inverse of the constant

$$
c_{\ell,\theta} = \int_{C(S)^+} \ell(w)^\alpha \mathrm{d}\mathbb{P}_{W_\theta}(w) = \mathbb{E}_\theta(\ell^\alpha(W))
$$

can be unbiasedly estimated directly via samples from W, details follow in the next chapter, but there is an unbiased estimator β_θ of $\frac{1}{c_{\ell,\theta}}$. Using these two unbiased estimator, we get an unbiased estimator for the conditional density $f_{W^{(\ell)}(\tilde{\mathbf{s}})}(\tilde{\mathbf{x}} \mid \theta)$, namely

$$
\overline{f_{W^{(\ell)}(\tilde{\mathbf{s}})}(\tilde{\mathbf{x}} \mid \theta)} := f_{W(\tilde{\mathbf{s}})}(\tilde{\mathbf{x}} \mid \theta) \cdot \beta_\theta \cdot Y_{\text{int},\theta,\tilde{\mathbf{x}}}
$$

Given a new $\theta' \sim q(\theta, \cdot)$ we use a MCMC with acceptance rate

$$
\alpha(\theta, \theta') := \min\left\{\frac{q(\theta, \theta') f_{W(\tilde{\mathbf{s}})}(\tilde{\mathbf{x}} \mid \theta') \cdot \beta'_\theta \cdot Y_{\text{int},\theta',\tilde{\mathbf{x}}}}{q(\theta', \theta) f_{W(\tilde{\mathbf{s}})}(\tilde{\mathbf{x}} \mid \theta) \cdot \beta_\theta \cdot Y_{\text{int},\theta,\tilde{\mathbf{x}}}}, 1\right\}
$$

7.2 Unbiased Estimation of the Reciprocal Mean

To estimate $1/c_{\ell,\theta}$ unbiasedly we rely on an estimator introduced in Moka et al. (2019). For a non negative random variable $Z \geq 0$ we want an unbiased estimator for $\beta = \frac{1}{\mathbb{E}(Z)}$. The idea is using a geometric series and rewrite β in

$$
\beta = \frac{1}{\mathbb{E}(Z)} = w \sum_{n=0}^{\infty} (1 - w\mathbb{E}(Z))^n,
$$

for $w < 2\beta$. Let $Z_i \overset{\text{i.i.d.}}{\sim} Z$ be i.i.d. copies of Z. Further let N be a non negative integer-valued random variable with

$$q_n := \mathbb{P}\left(N = n\right) > 0, \quad n \geq 0.$$

Then we have

$$\begin{aligned}
\beta = \frac{1}{\mathbb{E}Z} &= w \sum_{n=0}^{\infty} q_n \frac{(1 - w\mathbb{E}(Z))^n}{q_n} \\
&= w \sum_{n=0}^{\infty} q_n \frac{\mathbb{E}\prod_{i=1}^{n}(1 - wZ_i)}{q_n} \\
&= w\mathbb{E}\left[\sum_{n=0}^{\infty} \mathbb{P}(N = n) \frac{\prod_{i=1}^{n}(1 - wZ_i)}{q_n}\right] \\
&= w\mathbb{E}\left[\frac{1}{q_N} \prod_{i=1}^{N}(1 - wZ_i)\right].
\end{aligned}$$

With samples from Z and one sample of N we now have a unbiased estimator $\hat{\beta}(w)$ depending on the choice of w. The idea now is to chose w and N in such a way that the sample variance of the estimator $\mathrm{Var}(\hat{\beta}(w))$ is minimal. Moka et al. (2019) show that for any $w < 2\frac{\mathbb{E}(Z)}{\mathbb{E}(Z^2)}$ the variance $\mathrm{Var}(\hat{\beta}(w))$ is finite and minimal if N has a geometric distribution with success probability

$$p_w = 1 - \sqrt{\mathbb{E}(1 - wZ)^2},$$

i.e. we have

$$\mathbb{P}(N = n) = (1 - p_w)^n p_w, \quad n \geq 0.$$

Here $p_w > 0$ holds for $w < 2\frac{\mathbb{E}(Z)}{\mathbb{E}(Z^2)}$, since

$$p_w = 1 - \sqrt{\mathbb{E}(1 - wZ)^2} = 1 - \sqrt{1 - 2w\mathbb{E}(Z) + w^2\mathbb{E}(Z^2)}$$

We know that, for $\alpha \geq 1$,

$$\begin{aligned}
c_{\ell,\theta} &= \mathbb{E}_\theta(\ell(W)^\alpha) \\
&= \mathbb{E}_\theta\left(\frac{1}{|S|}\int_S W(s)\mathrm{d}s\right)^\alpha \\
&\overset{\text{Jensen with } \alpha \geq 1}{\leq} \mathbb{E}_\theta\left(\frac{1}{|S|}\int_S W(s)^\alpha \mathrm{d}s\right) \\
&= \frac{1}{|S|}\int_S \mathbb{E}_\theta\left(W(s)^\alpha\right)\mathrm{d}s \\
&= \frac{1}{|S|}\int_S 1\mathrm{d}s = 1.
\end{aligned}$$

Therefore we have

$$\beta(w) = \frac{1}{c_{\ell,\theta}} = \frac{1}{\mathbb{E}_\theta(\ell(W)^\alpha)} \geq 1$$

and we only need to ensure $w < 2 \leq 2\beta$ to guarantee the convergence of the series.

7.3 Estimation of shape parameter α

We want to make use of the limit theorem for a fixed location s_0 and large u, i.e.

$$\mathcal{L}\left(\frac{X(s_0)}{u} \,\middle|\, \ell(X) > u\right) \approx \mathcal{L}\left(P^\alpha \frac{W^{(\ell)}(s_0)}{\ell\left(W^{(\ell)}\right)}\right).$$

Now we look at the distribution of $\frac{W^{(\ell)}(s_0)}{\ell\left(W^{(\ell)}\right)}$. We have

$$\begin{aligned}\mathbb{P}\left(\frac{W^{(\ell)}(s_0)}{\ell\left(W^{(\ell)}\right)} \leq z\right) &= \frac{1}{c_\ell}\int_{C(S)^+} \ell(w)^\alpha \mathbb{1}_{\left\{\frac{w(s_0)}{\ell(w)} \leq z\right\}} \mathrm{d}\mathbb{P}_W(w) \\ &= \frac{1}{c_\ell}\int_{C(S)^+} \ell(w)^\alpha \mathbb{1}_{\left\{\ell(w) \geq \frac{1}{z}\right\}} \mathrm{d}\mathbb{P}_W(w) = \frac{1}{c_\ell}\mathbb{E}\left[\ell(W)^\alpha \mathbb{1}_{\left\{\ell(W)\geq\frac{1}{z}\right\}}\right].\end{aligned}$$

As the density of P_α is known, we can use its independence from $\frac{W^{(\ell)}(s_0)}{\ell\left(W^{(\ell)}\right)} = \frac{1}{\ell\left(W^{(\ell)}\right)}$ to obtain

$$\begin{aligned}\mathbb{P}\left(P_\alpha \cdot \frac{1}{\ell\left(W^{(\ell)}\right)} \leq z\right) &= \mathbb{P}\left(\frac{1}{\ell\left(W^{(\ell)}\right)} \leq \frac{z}{P_\alpha}\right) = \int_1^\infty f_{P_\alpha}(\tilde{x})\mathbb{P}\left(\frac{1}{\ell\left(W^{(\ell)}\right)} \leq \frac{z}{\tilde{x}}\right)\mathrm{d}\tilde{x} \\ &= \int_0^z \alpha\left(\frac{z}{x}\right)^{-\alpha-1}\mathbb{P}\left(\frac{1}{\ell\left(W^{(\ell)}\right)} \leq x\right)\frac{z}{x^2}\mathrm{d}x \\ &= \alpha z^{-\alpha}\int_0^z x^{\alpha-1}\mathbb{P}\left(\frac{1}{\ell\left(W^{(\ell)}\right)} \leq x\right)\mathrm{d}x.\end{aligned}$$

Now we get the density as the derivate via the product rule

$$\begin{aligned}f_{P_\alpha\cdot\frac{1}{\ell\left(W^{(\ell)}\right)}}(z) &= \frac{d}{dz}\mathbb{P}\left(P_\alpha \cdot \frac{1}{\ell\left(W^{(\ell)}\right)} \leq z\right) \\ &= \alpha(-\alpha)z^{-\alpha-1}\int_0^z x^{\alpha-1}\mathbb{P}\left(\frac{1}{\ell\left(W^{(\ell)}\right)} \leq x\right)\mathrm{d}x + \alpha z^{-\alpha}z^{\alpha-1}\mathbb{P}\left(\frac{1}{\ell\left(W^{(\ell)}\right)} \leq z\right) \\ &= \alpha z^{-1}\left(\mathbb{P}\left(\frac{1}{\ell\left(W^{(\ell)}\right)} \leq z\right) - \alpha z^{-\alpha}\int_0^z x^{\alpha-1}\mathbb{P}\left(\frac{1}{\ell\left(W^{(\ell)}\right)} \leq x\right)\mathrm{d}x\right) \\ &= \alpha z^{-1}\left(\frac{1}{c_\ell}\mathbb{E}\left[\ell(W)^\alpha\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{z}\right\}}\right] - \alpha z^{-\alpha}\int_0^z x^{\alpha-1}\frac{1}{c_\ell}\mathbb{E}\left[\ell(W)^\alpha\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{x}\right\}}\right]\mathrm{d}x\right) \\ &\overset{\text{Fubini}}{=} \frac{\alpha}{c_\ell}z^{-1}\mathbb{E}\left[\ell(W)^\alpha\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{z}\right\}} - \alpha z^{-\alpha}\int_0^z x^{\alpha-1}\ell(W)^\alpha\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{x}\right\}}\mathrm{d}x\right] \\ &= \frac{\alpha}{c_\ell}z^{-1}\mathbb{E}\left[\ell(W)^\alpha\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{z}\right\}} - z^{-\alpha}\ell(W)^\alpha\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{z}\right\}}\int_{\frac{1}{\ell(W)}}^z \alpha x^{\alpha-1}\mathrm{d}x\right] \\ &= \frac{\alpha}{c_\ell}z^{-1}\mathbb{E}\left[\ell(W)^\alpha\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{z}\right\}} - z^{-\alpha}\ell(W)^\alpha\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{z}\right\}}\left(z^\alpha - \frac{1}{\ell(W)^\alpha}\right)\right] \\ &= \frac{\alpha}{c_\ell}z^{-1}\mathbb{E}\left[\ell(W)^\alpha\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{z}\right\}} - \ell(W)^\alpha\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{z}\right\}} + z^{-\alpha}\mathbb{1}_{\left\{\ell(W)\geq\frac{1}{z}\right\}}\right] \\ &= \frac{\alpha}{c_\ell}z^{-\alpha-1}\mathbb{P}\left(\frac{1}{\ell(W)} \leq z\right).\end{aligned}$$

and so we get a density via a difference quotient

$$
\begin{aligned}
f_{\frac{W^{(\ell)}(s_0)}{\ell\left(W^{(\ell)}\right)}}(z) &\overset{\varepsilon\to 0}{\leftarrow} \frac{1}{\varepsilon}\mathbb{P}\left(z \le \frac{W^{(\ell)}(s_0)}{\ell\left(W^{(\ell)}\right)} \le z+\varepsilon\right) \\
&= \frac{1}{\varepsilon}\frac{1}{c_\ell}\mathbb{E}\left[\underbrace{\ell(W)}_{\frac{1}{z}\ge\cdots\ge\frac{1}{z+\varepsilon}}{}^{\alpha}\mathbb{1}_{\left\{\frac{1}{z}\ge\ell(W)\ge\frac{1}{z+\varepsilon}\right\}}\right] \\
&\sim \frac{1}{\varepsilon}\frac{1}{c_\ell}\left(\frac{1}{z}\right)^{\alpha}\mathbb{E}\left[\mathbb{1}_{\left\{\frac{1}{z}\ge\ell(W)\ge\frac{1}{z+\varepsilon}\right\}}\right] \\
&= \frac{1}{\varepsilon}\frac{1}{c_\ell}\left(\frac{1}{z}\right)^{\alpha}\mathbb{E}\left[\mathbb{1}_{\left\{z\le\frac{1}{\ell(W)}\le z+\varepsilon\right\}}\right] \\
&= \frac{1}{c_\ell}\left(\frac{1}{z}\right)^{\alpha}\frac{\mathbb{P}\left\{\frac{1}{\ell(W)}\le z+\varepsilon\right\}-\mathbb{P}\left\{\frac{1}{\ell(W)}\le z\right\}}{\varepsilon} \\
&\overset{\varepsilon\to 0}{\to} \frac{1}{c_\ell}\left(\frac{1}{z}\right)^{\alpha} f_{\frac{1}{\ell(W)}}(z).
\end{aligned}
$$

Given that and the product form in the limit, we use a formula for products of independent positive random variables X, Y

$$
f_{X\cdot Y}(z) = \int_0^\infty \frac{1}{t} f_X(t)\cdot f_Y\left(\frac{z}{t}\right) dt. \tag{10}
$$

This gives the density $f_{\frac{X(s_0)}{u}|\ell(X)>u}$ as

$$
\begin{aligned}
f_{\frac{X(s_0)}{u}|\ell(X)>u}(z) &\approx f_{\frac{W^{(\ell)}(s_0)}{\ell\left(W^{(\ell)}\right)}\cdot P^\alpha}(z) \\
&= \int_0^\infty \frac{1}{t} f_{\frac{W^{(\ell)}(s_0)}{\ell\left(W^{(\ell)}\right)}}(t) f_{P^\alpha}\left(\frac{z}{t}\right)\mathrm{d}t \\
&= \int_{\left\{t:\frac{z}{t}>1\right\}} \frac{1}{t}\frac{1}{c_\ell}\left(\frac{1}{t}\right)^{\alpha} f_{\frac{1}{\ell(W)}}(t)\alpha\left(\frac{z}{t}\right)^{-\alpha-1}\mathrm{d}t \\
&= \int_0^z \frac{1}{t}\frac{1}{c_\ell}\left(\frac{1}{t}\right)^{\alpha} f_{\frac{1}{\ell(W)}}(t)\alpha\left(\frac{z}{t}\right)^{-\alpha-1}\mathrm{d}t \\
&= \alpha\frac{z^{-\alpha-1}}{c_\ell}\int_0^z f_{\frac{1}{\ell(W)}}(t)\mathrm{d}t \\
&= \alpha\frac{z^{-\alpha-1}}{c_\ell}\mathbb{P}\left(\ell(W)\ge\frac{1}{z}\right).
\end{aligned}
$$

8 Using Fractional Brownian Motion structure of the Underlying Process for Fast Simulation

The fractional Brownian motion is a Gaussian Process in $\mathbf{R}^d$ with covariance function

$$
c(s,t) := \frac{1}{2}\left(||s||^{2H} + ||t||^{2H} - ||s-t||^{2H}\right),
$$

where $H \in (0,1]$ is the so called Hurst parameter. We use the notation $B_H(p,q)$ for the two dimensional version. We look at a finite $N\times N$ grid where we want to simulate on. Therefore we make use of the independent increments. We define

$$
\begin{aligned}
\Delta B_H^x(p) &= B_H(p,0) - B_H(p+1,0) \\
\Delta B_H^y(q) &= B_H(0,q+1) - B_H(0,q) \\
\Delta B_H^{x,y}(p,q) &= B_H(p+1,q+1) - B_H(p+1,q) - B_H(p,q+1) + B_H(p,q),
\end{aligned}
$$

where $p, q \in \{0, \dots, N-1\}$ and the process for increments I as

$$I(p,q) = \begin{cases} \Delta B_H^x(p) & \text{for } q = 0 \\ \Delta B_H^x(q) & \text{for } p = 0 \\ \Delta B_H^{x,y}(p,q) & \text{for } 1 \leq p, q \leq N-1 \end{cases}$$
$$= \begin{cases} B_H(p,0) - B_H(p+1,0) & \text{for } q = 0 \\ B_H(0,q+1) - B_H(0,q) & \text{for } p = 0 \\ B_H(p+1,q+1) - B_H(p+1,q) - B_H(p,q+1) + B_H(p,q) & \text{for } 1 \leq p, q \leq N-1. \end{cases}$$

Now we want to show stationarity of the increment process. We look at a point on the x-axis and some point in the grid.

$$\begin{aligned}
&2Cov(I(r+p,q), I(r,0)) \\
=&2Cov(B_H(r+p+1,q+1) - B_H(r+p+1,q) - B_H(r+p,q+1) + B_H(r+p,q), B_H(r,0) - B_H(r+1,0)) \\
=&2Cov(B_H(r+p+1,q+1), B_H(r,0)) - 2Cov(B_H(r+p+1,q), B_H(r,0)) \\
&- 2Cov(B_H(r+p,q+1), B_H(r,0)) + 2Cov(B_H(r+p,q), B_H(r,0))) \\
&- 2Cov(B_H(r+p+1,q+1), B_H(r+1,0)) + 2Cov(B_H(r+p+1,q), B_H(r+1,0)) \\
&+ 2Cov(B_H(r+p,q+1), B_H(r+1,0)) - 2Cov(B_H(r+p,q), B_H(r+1,0))) \\
=&||r+p+1,q+1||^{2H} + ||r,0||^{2H} - ||p+1,q+1||^{2H} \\
&- ||r+p+1,q||^{2H} - ||r,0||^{2H} + ||p+1,q||^{2H} \\
&- ||r+p,q+1||^{2H} - ||r,0||^{2H} + ||p,q+1||^{2H} \\
&+ ||r+p,q||^{2H} + ||r,0||^{2H} - ||p,q||^{2H} \\
&- ||r+p+1,q+1||^{2H} - ||r+1,0||^{2H} + ||p,q+1||^{2H} \\
&+ ||r+p+1,q||^{2H} + ||r+1,0||^{2H} - ||p,q||^{2H} \\
&+ ||r+p,q+1||^{2H} + ||r+1,0||^{2H} - ||p-1,q+1||^{2H} \\
&- ||r+p,q||^{2H} - ||r+1,0||^{2H} + ||p-1,q||^{2H} \\
=&- ||p+1,q+1||^{2H} + ||p+1,q||^{2H} + ||p,q+1||^{2H} - ||p,q||^{2H} \\
&+ ||p,q+1||^{2H} - ||p,q||^{2H} - ||p-1,q+1||^{2H} + ||p-1,q||^{2H} \\
=&2Cov(I(p,q), I(0,0)),
\end{aligned}$$

where the last step is easily done by setting $r = 0$. Stationarity for the rest is clear, since its just increments. Now we calculate all the remaining covariance combinations to be able to simulate it. First the case on the edges. We get

$$\begin{aligned}
&2Cov(I(r+p,0), I(r,0)) \\
=&2Cov(B_H(r+p,0) - B_H(r+p+1,0), B_H(r,0) - B_H(r+1,0)) \\
=&2Cov(B_H(r+p,0), B_H(r,0)) - 2Cov(B_H(r+p+1,0), B_H(r,0)) \\
&- 2Cov(B_H(r+p,0), B_H(r+1,0)) + 2Cov(B_H(r+p+1,0), B_H(r+1,0)) \\
=&||r+p,0||^{2H} + ||r,0||^{2H} - ||p,0||^{2H} \\
&- ||r+p+1,0||^{2H} - ||r,0||^{2H} + ||p+1,0||^{2H} \\
&- ||r+p,0||^{2H} - ||r+1,0||^{2H} + ||p-1,0||^{2H} \\
&+ ||r+p+1,0||^{2H} + ||r+1,0||^{2H} - ||p,0||^{2H} \\
&= ||p+1,0||^{2H} + ||p-1,0||^{2H} - 2||p,0||^{2H}
\end{aligned}$$

for the x-axis and

$$\begin{aligned}
&2Cov(I(0,r+q),I(0,r))\\
=&2Cov(B_H(0,r+q+1)-B_H(0,r+q),B_H(0,r+1)-B_H(0,r))\\
=&2Cov(B_H(0,r+q+1),B_H(0,r+1))-2Cov(B_H(0,r+q),B_H(0,r+1))\\
&-2Cov(B_H(0,r+q+1),B_H(0,r))+2Cov(B_H(0,r+q),B_H(0,r))\\
=&||0,r+q+1||^{2H}+||0,r+1||^{2H}-||0,q||^{2H}\\
&-||0,r+q||^{2H}-||0,r+1||^{2H}+||0,q-1||^{2H}\\
&-||0,r+q+1||^{2H}-||0,r||^{2H}+||0,q+1||^{2H}\\
&+||0,r+q||^{2H}+||0,r||^{2H}-||0,q||^{2H}\\
=&||0,q-1||^{2H}+||0,q+1||^{2H}-2||0,q||^{2H}
\end{aligned}$$

for the y-axis. With one point on the y-axis and one within the grid we get

$$\begin{aligned}
&2Cov(I(p,r+q),I(0,r))\\
=&2Cov(B_H(p+1,r+q+1)-B_H(p+1,r+q)-B_H(p,r+q+1)+B_H(p,r+q),B_H(0,r+1)-B_H(0,r))\\
=&2Cov(B_H(p+1,r+q+1),B_H(0,r+1))\\
&-2Cov(B_H(p+1,r+q),B_H(0,r+1))\\
&-2Cov(B_H(p,r+q+1),B_H(0,r+1))\\
&+2Cov(B_H(p,r+q),B_H(0,r+1))\\
&-2Cov(B_H(p+1,r+q+1),B_H(0,r))\\
&+2Cov(B_H(p+1,r+q),B_H(0,r))\\
&+2Cov(B_H(p,r+q+1),B_H(0,r))\\
&-2Cov(B_H(p,r+q),B_H(0,r))\\
=&||p+1,r+q+1||^{2H}+||0,r+1||^{2H}-||p+1,q||^{2H}\\
&-||p+1,r+q||^{2H}-||0,r+1||^{2H}+||p+1,q-1||^{2H}\\
&-||p,r+q+1||^{2H}-||0,r+1||^{2H}+||p,q||^{2H}\\
&+||p,r+q||^{2H}+||0,r+1||^{2H}-||p,q-1||^{2H}\\
&-||p+1,r+q+1||^{2H}-||0,r||^{2H}+||p+1,q+1||^{2H}\\
&+||p+1,r+q||^{2H}+||0,r||^{2H}-||p+1,q||^{2H}\\
&+||p,r+q+1||^{2H}+||0,r||^{2H}-||p,q+1||^{2H}\\
&-||p,r+q||^{2H}-||0,r||^{2H}+||p,q||^{2H}\\
=&||p+1,q-1||^{2H}-||p,q-1||^{2H}\\
&+||p+1,q+1||^{2H}-2||p+1,q||^{2H}-||p,q+1||^{2H}+2||p,q||^{2H}.
\end{aligned}$$

Within the grid we have

$$
\begin{aligned}
&2Cov(I(p,q), I(p+x,q+y)) \\
=&2Cov(B_H(p+1,q+1) - B_H(p+1,q) - B_H(p,q+1) + B_H(p,q), \\
&\qquad B_H(p+x+1,q+y+1) - B_H(p+x+1,q+y) - B_H(p+x,q+y+1) + B_H(p+x,q+y)) \\
=&2Cov(B_H(p+1,q+1), B_H(p+x+1,q+y+1)) - 2Cov(B_H(p+1,q), B_H(p+x+1,q+y+1)) \\
&- 2Cov(B_H(p,q+1), B_H(p+x+1,q+y+1)) + 2Cov(B_H(p,q), B_H(p+x+1,q+y+1)) \\
&- 2Cov(B_H(p+1,q+1), B_H(p+x+1,q+y)) + 2Cov(B_H(p+1,q), B_H(p+x+1,q+y)) \\
&+ 2Cov(B_H(p,q+1), B_H(p+x+1,q+y)) - 2Cov(B_H(p,q), B_H(p+x+1,q+y)) \\
&- 2Cov(B_H(p+1,q+1), B_H(p+x,q+y+1)) + 2Cov(B_H(p+1,q), B_H(p+x,q+y+1)) \\
&+ 2Cov(B_H(p,q+1), B_H(p+x,q+y+1)) - 2Cov(B_H(p,q), B_H(p+x,q+y+1)) \\
&+ 2Cov(B_H(p+1,q+1), B_H(p+x,q+y)) - 2Cov(B_H(p+1,q), B_H(p+x,q+y)) \\
&- 2Cov(B_H(p,q+1), B_H(p+x,q+y)) + 2Cov(B_H(p,q), B_H(p+x,q+y)) \\
=&||p+1,q+1||^{2H} + ||p+x+1,q+y+1||^{2H} - ||x,y||^{2H} \\
&- ||p+1,q||^{2H} - ||p+x+1,q+y+1||^{2H} + ||x,y+1||^{2H} \\
&- ||p,q+1||^{2H} - ||p+x+1,q+y+1||^{2H} + ||x+1,y||^{2H} \\
&+ ||p,q||^{2H} + ||p+x+1,q+y+1||^{2H} - ||x+1,y+1||^{2H} \\
&- ||p+1,q+1||^{2H} - ||p+x+1,q+y||^{2H} + ||x,y-1||^{2H} \\
&+ ||p+1,q||^{2H} + ||p+x+1,q+y||^{2H} - ||x,y||^{2H} \\
&+ ||p,q+1||^{2H} + ||p+x+1,q+y||^{2H} - ||x+1,y-1||^{2H} \\
&- ||p,q||^{2H} - ||p+x+1,q+y||^{2H} + ||x+1,y||^{2H} \\
&- ||p+1,q+1||^{2H} - ||p+x,q+y+1||^{2H} + ||x-1,y||^{2H} \\
&+ ||p+1,q||^{2H} + ||p+x,q+y+1||^{2H} - ||x-1,y+1||^{2H} \\
&+ ||p,q+1||^{2H} + ||p+x,q+y+1||^{2H} - ||x,y||^{2H} \\
&- ||p,q||^{2H} - ||p+x,q+y+1||^{2H} + ||x,y+1||^{2H} \\
&+ ||p+1,q+1||^{2H} + ||p+x,q+y||^{2H} - ||x-1,y-1||^{2H} \\
&- ||p+1,q||^{2H} - ||p+x,q+y||^{2H} + ||x-1,y||^{2H} \\
&- ||p,q+1||^{2H} - ||p+x,q+y||^{2H} + ||x,y-1||^{2H} \\
&+ ||p,q||^{2H} + ||p+x,q+y||^{2H} - ||x,y||^{2H} \\
=& -4||x,y||^{2H} + 2||x,y+1||^{2H} + 2||x+1,y||^{2H} - ||x+1,y+1||^{2H} + 2||x,y-1||^{2H} \\
&- ||x+1,y-1||^{2H} + 2||x-1,y||^{2H} - 2||x-1,y+1||^{2H}
\end{aligned}
$$

Combining the results in one closed expression we can make use of the stationarity and get

$$
\begin{aligned}
Cov(I(p,q), I(p+x,q+y)) &= \\
Cov(I(0,0), I(x,y)) &= \begin{cases} \Delta B_H^x(p) & \text{for } q = 0 \\ \Delta B_H^x(q) & \text{for } p = 0 \\ \Delta B_H^{x,y}(p,q) & \text{for } 1 \le p,q \le N-1 \end{cases}
\end{aligned}
$$

References

L. de Haan. A spectral representation for max-stable processes. *Ann. Probab.*, 12(4):1194–1204, 1984.

S. B. Moka, D. P. Kroese, and S. Juneja. Unbiased estimation of the reciprocal mean for non-negative random variables. In *2019 Winter Simulation Conference (WSC)*, pages 404–415. IEEE, 2019.

L. Tierney. Markov chains for exploring posterior distributions. *the Annals of Statistics*, pages 1701–1728, 1994.

Vorschlag für neue Struktur der Arbeit:

Basics

- Einführung in max-stabile Prozesse und Spektraldarstellung
- Einführung der äquivalenten Darstellung mit Spektralprozessen der Form $\widetilde{W}^{(\ell)}/\ell(\widetilde{W}^{(\ell)})$
  Frage: Brauche wir das ursprüngliche $W^{(\ell)}$ überhaupt noch???

Bayesian Inference for ℓ-Pareto processes if $\ell(X)$ is known

- fehlt noch; meiner Meinung nach aber sehr hilfreich zum Verständnis
- Probleme wie erwartungstreue Schätzung treten bereits hier auf
- Am Ende: Theorem: If X is an ℓ-Pareto process, the distribution of the MCMC algorithm converges to the posterior distribution in total variation norm.

Ab hier: $\ell(X)$ nicht beobachtbar!

Selection of partial observations of X with $\ell(X) > u$

- bisherige Section 4.1, erster Teil
- Am Ende:

Bayesian Inference for ℓ-Pareto processes in case of partial observations of X

- bisherige Section 4.1, zweiter Teil
- Am Ende: Theorem: If X is an ℓ-Pareto process and the selection of exceedances is fixed, then, the distribution of the MCMC algorithm converges to the posterior distribution in total variation norm.