Conditional simulations of max-stable processes

C. Dombry†, F. Éyi-Minko†, M. Ribatet‡

† Laboratoire de Mathématiques et Application, Université de Poitiers
‡ Institut de Mathématiques et de Modélisation, Université Montpellier 2
Spatial extremes

- Max-stable processes
- Max-i.d. processes
- Latent variable sampling (max-stable)
- Brown-Resnick processes
- Conditional sampling (max-linear)
- Conditional sampling (regular)

Timeline:
- 1985
- 1990
- 1995
- 2000
- 2005
- 2010
- 2015
Definition 1. A process $Z$ defined on a compact metric space $\mathcal{X}$ is max-i.d. in $C(\mathcal{X})$ if it is sample continuous and for each $n \in \mathbb{N}$, there exists independent identically distributed sample continuous processes $Z_{i,n}$ such that

$$Z \stackrel{d}{=} \max_{i=1,\ldots,n} Z_{i,n}, \quad n \in \mathbb{N},$$

(1)

where $(\max Z_{i,n})(x) = \max Z_{i,n}(x)$ for all $x \in \mathcal{X}$.

Remark. If (1) holds with

$$Z_{i,n} = \frac{Z_i - b_n}{a_n},$$

for some continuous functions $a_n > 0$ and $b_n \in \mathbb{R}$ and where $Z_i$ are independent copies of $Z$, then $Z$ is said to be max-stable.
**Theorem 1** (de Haan 1984 & Giné, Hahn and Vatan 1990). *Let $Z$ be a max-.i.d. process on $\mathcal{X}$ such that $\text{ess inf } Z(x) \equiv 0$. Then there exists a unique $\sigma$–finite measure $\Lambda$ on $\mathcal{C}_0 = \mathcal{C}\{\mathcal{X}, [0, \infty)\} \setminus \{0\}$ such that*

$$Z \overset{d}{=} \max_{\varphi \in \Phi} \varphi,$$

*where $\Phi$ is a Poisson point process on $\mathcal{C}_0$ with intensity measure $\Lambda$.*

**Remark.** If $Z$ is max-stable with unit Fréchet margins, i.e.,

$$\text{Pr}\{Z(x) \leq z\} = \exp(-1/z), \ z > 0,$$

*then*

$$d\Lambda = \zeta^{-2} d\zeta d\sigma,$$

*where $\sigma$ is a finite measure on $\mathcal{C}_1 = \{f \in \mathcal{C}_0 : \|f\| = 1\}$ such that*

$$\int_{\mathcal{C}_1} f(x) d\sigma(f) = 1, \quad x \in \mathcal{X}.$$
The specific form of the intensity measure $d\Lambda = \zeta^{-2} d\xi \, d\sigma$ is well known in extreme value theory.

It factorizes into a radial part $\zeta^{-2}$ and an angular part $\sigma$ using the bijection

$$C_0 \longrightarrow (0, \infty) \times C_1$$

$$f \longmapsto (\| f \|, \, f / \| f \|).$$

The measure $\sigma$ is called the spectral measure and characterizes the spatial dependence of extremes—indeed independently from the radius.

For statistical purposes, it is often more convenient to “think of” $\sigma$ as the distribution of a non-negative, sample continuous stochastic process $Y$ such that $\mathbb{E}\{Y(x)\} = 1$, $x \in \mathcal{X}$. 
Smith’s model (1990)

\[ \varphi_i(x) = \zeta_i \phi(x - U_i; 0, \Sigma), \quad x \in \mathcal{X}, \]

where \( \{(\zeta_i, U_i)\}_{i \geq 1} \) are the points of a Poisson process on \((0, \infty) \times \mathbb{R}^d\) with intensity measure \(d\Lambda(\zeta, u) = \zeta^{-2} d\zeta \, du\) and \(\phi(\cdot; 0, \Sigma)\) is the centered \(d\)-variate normal density with covariance matrix \(\Sigma\).

**Figure 1:** One realization from a Smith process on \([-10, 10]\) with \(\Sigma = 3\).
Schlahter’s model (2002)

\[ \varphi_i(x) = \sqrt{2\pi} \zeta_i \max\{0, \varepsilon_i(x)\}, \quad x \in \mathcal{X}, \]

where \( \{\zeta_i\}_{i \geq 1} \) are the points of a Poisson process on \((0, \infty)\) with intensity measure \(d\Lambda(\zeta) = \zeta^{-2} d\zeta\) and \( \varepsilon_i \) independent copies of a standard Gaussian process.

**Figure 2:** One realization from a Schlather process on \([-10, 10]\) with correlation function \( \rho(h) = \exp(-h/3) \).
\[ \varphi_i(x) = \zeta_i \exp\{\varepsilon_i(x) - \gamma(x)\}, \quad x \in \mathcal{X}, \]

where \( \{\zeta_i\}_{i \geq 1} \) are the points of a Poisson process on \((0, \infty)\) with intensity measure \( d\Lambda(\zeta) = \zeta^{-2} d\zeta \) and \( \varepsilon_i \) independent copies of a centered Gaussian process with semi variogram \( \gamma \).

**Figure 3:** One realization from a Brown–Resnick process on \([-10, 10]\) with semi variogram \( \gamma(h) = \sqrt{h/3} \).
Let $x \in \mathcal{X}^k$ and $z = (0, \infty)^k$, then

$$
\Pr\{Z(x) \leq z\} = \exp \left[ -\Lambda\{(0, z)^c}\right] = \exp\{-V(z)\}.
$$
Let \( x \in X^k \) and \( z = (0, \infty)^k \), then

\[
\Pr\{Z(x) \leq z\} = \exp \left[ -\Lambda\{(0, z)^c\} \right] = \exp\{-V(z)\}.
\]

In particular when

\( k = 2 \):

\[
f(z) = (V_1 V_2 - V_{12}) \exp\{-V(z)\}
\]

\( k = 3 \):

\[
f(z) = (-V_1 V_2 V_3 + V_{12} V_3 + V_{13} V_2 + V_1 V_{23} - V_{123}) \exp\{-V(z)\}
\]

\( k = n \):

\[
f(z) = (\text{sum of many many terms}) \exp\{-V(z)\}
\]

Use of the maximum pairwise likelihood estimator

\[
\hat{\theta}_p = \arg\max_{\theta \in \Theta} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \omega_{i,j} \ln f(z_i, z_j; \theta).
\]
Let $Z$ be a max-stable process defined on $\mathcal{X}$ with unit Fréchet margins.

We observe $Z$ at some conditioning locations $\mathbf{x} = (x_1, \ldots, x_k) \in \mathcal{X}^k$ giving rise to some (critical) values $\mathbf{z} = (z_1, \ldots, z_k) \in (0, \infty)^k$. 
Let $Z$ be a max-stable process defined on $\mathcal{X}$ with unit Fréchet margins.

We observe $Z$ at some conditioning locations $\mathbf{x} = (x_1, \ldots, x_k) \in \mathcal{X}^k$ giving rise to some (critical) values $\mathbf{z} = (z_1, \ldots, z_k) \in (0, \infty)^k$.

Our goal is to sample from $Z(\cdot) | \{Z(x_1) = z_1, \ldots, Z(x_k) = z_k\}$. 
Outline

1. Conditional distributions
2. MCMC sampler
3. Simulation Study
4. Applications
1. Conditional distributions of max-stable processes
Decomposition of $\Phi$

1. Conditional distributions
   - Decomposition of $\Phi$
   - Sub-extremal functions
   - Random partitions
   - Sampling scheme
   - Examples

2. MCMC sampler

3. Simulation Study

4. Applications

\[ Z(x) = \max_{\varphi \in \Phi} \varphi(x), \quad x \in \mathcal{X} \]

- Consider the two following Poisson point processes

  \[ \Phi^- = \{ \varphi \in \Phi : \varphi(x_i) < z_i, \text{ for all } i \in \{1, \ldots, k\} \} \] \hspace{5mm} (sub-extremal functions)

  \[ \Phi^+ = \{ \varphi \in \Phi : \varphi(x_i) = z_i, \text{ for some } i \in \{1, \ldots, k\} \} \] \hspace{5mm} (extremal functions)

- Clearly $\Phi = \Phi^- \cup \Phi^+$. 

Conditional simulations of max-stable processes

Mathieu Ribatet – 13 / 50
Decomposition of $\Phi$

\[ Z(x) = \max_{\varphi \in \Phi} \varphi(x), \quad x \in \mathcal{X} \]

\[ \square \quad \text{Consider the two following Poisson point processes} \]

\[ \Phi^- = \{ \varphi \in \Phi : \varphi(x_i) < z_i, \text{ for all } i \in \{1, \ldots, k\} \}, \quad \text{(sub-extremal functions)} \]

\[ \Phi^+ = \{ \varphi \in \Phi : \varphi(x_i) = z_i, \text{ for some } i \in \{1, \ldots, k\} \}. \quad \text{(extremal functions)} \]

\[ \square \quad \text{Clearly } \Phi = \Phi^- \cup \Phi^+. \]

Key point #1: Conditionally on $Z(x) = z$, $\Phi^-$ and $\Phi^+$ are independent.
Why should we bother about $\Phi^-$?

1. Conditional distributions
   - Decomposition of $\Phi$
   - Sub-extremal functions
2. Random partitions
3. Sampling scheme
4. Examples
5. MCMC sampler
6. Simulation Study
7. Applications
Why should we bother about $\Phi^-$?

1. Conditional distributions
   - Decomposition of $\Phi$
     - Sub-extremal functions
   - Random partitions
   - Sampling scheme
   - Examples
2. MCMC sampler
3. Simulation Study
4. Applications

Conditional simulations of max-stable processes
Why should we bother about $\Phi^-$?

1. Conditional distributions
   - Decomposition of $\Phi$
     - Sub-extremal functions
   - Random partitions
   - Sampling scheme
   - Examples

2. MCMC sampler

3. Simulation Study

4. Applications

Conditional simulations of max-stable processes
Why should we bother about $\Phi^-$?

1. Conditional distributions
   - Decomposition of $\Phi$
     - Sub-extremal functions
   - Random partitions
   - Sampling scheme
   - Examples

2. MCMC sampler

3. Simulation Study

4. Applications
Why should we bother about $\Phi^-$?

The atoms of $\Phi^+$ are only of interest if we restrict our attention to the conditioning points $x$;

But most often one would like to get realizations at $s \neq x$.

The atoms of $\Phi^-$ are needed since it is likely that $\max \Phi^- (s) > \max \Phi^+ (s)$!
Conditional intensity function

\[ Z(x) = \max_{i \geq 1} \varphi_i(x), \quad x = (x_1, \ldots, x_k). \]

- The Poisson point process \( \{\varphi_i(x)\}_{i \geq 1} \) has intensity measure

\[ \Lambda_x(A) = \int_0^\infty \Pr\{\zeta Y(x) \in A\} \zeta^{-2} d\zeta, \quad \text{Borel set } A \subset \mathbb{R}^k. \]

- We assume that \( \Phi \) is regular, i.e., \( \Lambda_x(dz) = \lambda_x(z) dz \), for all \( x \in \mathcal{X}^k \).
Conditional intensity function

\[ Z(x) = \max_{i \geq 1} \phi_i(x), \quad x = (x_1, \ldots, x_k). \]

- The Poisson point process \( \{\phi_i(x)\}_{i \geq 1} \) has intensity measure
  \[ \Lambda_x(A) = \int_0^\infty \Pr\{\zeta Y(x) \in A\} \zeta^{-2} d\zeta, \quad \text{Borel set } A \subset \mathbb{R}^k. \]

- We assume that \( \Phi \) is regular, i.e., \( \Lambda_x(dz) = \lambda_x(z) dz \), for all \( x \in \mathcal{X}^k \).

Key point #2: The conditional intensity function

\[ \lambda_{x_1|x_2,z_2}(u) = \frac{\lambda_{(x_1,x_2)}(u,z_2)}{\lambda_{x_2}(z_2)}, \quad x = (x_1, x_2), \ z = (z_1, z_2), \]

characterizes (up to a truncation) the distribution of the extremal functions.
Random partitions?

1. Conditional distributions
   Decomposition of $\Phi$
   Sub-extremal functions
   Random partitions
2. MCMC sampler
3. Simulation Study
4. Applications

Conditional simulations of max-stable processes
Random partitions?

1. Conditional distributions
   - Decomposition of $\Phi$
   - Sub-extremal functions
   - Random partitions
   - Sampling scheme
   - Examples

2. MCMC sampler

3. Simulation Study

4. Applications

Conditional simulations of max-stable processes
1. Conditional distributions
   Decomposition of $\Phi$
   Sub-extremal functions
   Random partitions
   Sampling scheme
   Examples
2. MCMC sampler
3. Simulation Study
4. Applications
Random partitions?

1. Conditional distributions
   Decomposition of $\Phi$
   Sub-extremal functions
   Random partitions
   Sampling scheme
   Examples

2. MCMC sampler

3. Simulation Study

4. Applications

Conditional simulations of max-stable processes

Mathieu Ribatet – 16 / 50
1. Conditional distributions
   - Decomposition of $\Phi$
   - Sub-extremal functions
   - Random partitions
   - Sampling scheme
   - Examples

2. MCMC sampler
3. Simulation Study
4. Applications

Conditional simulations of max-stable processes
Here the set \( \{x_1, \ldots, x_5\} \) is partitioned into \( \{(x_1, x_3), \{x_2\}, \{x_4\}, \{x_5\}\)
Here the set \( \{x_1, \ldots, x_5\} \) is partitioned into \( \{\{x_1, x_3\}, \{x_2\}, \{x_4\}, \{x_5\}\} \)

- The hitting bounds \( \{z_i\}_{i=1,\ldots,k} \) might be reached by several extremal functions, i.e., \( \Phi^+ = \{\varphi_1^+, \ldots, \varphi_k^+\} = \{\varphi_1^+, \ldots, \varphi_\ell^+\} \) a.s., \( 1 \leq \ell \leq k \).
- So we need to take into account all possible ways these hitting bounds are reached: the hitting scenarios
□ This suggests a three step sampling scheme:
This suggests a three step sampling scheme:

**Step 1** Draw a random partition $\tau$, i.e., a hitting scenario;

**Step 2** Given $\tau$ of size $\ell$, draw the extremal functions $\varphi_1^+, \ldots, \varphi_\ell^+$ independently;

**Step 3** Independently from Steps 1 & 2, draw the sub-extremal functions $\varphi_i^-$, $i \geq 1$. 
Step 1: The random partitions

- Let $\mathcal{P}_k$ the set of all possible partitions of the set $\{x_1, \ldots, x_k\}$.
- Draw a random partition $\tau \in \mathcal{P}_k$ with distribution

$$\pi_x(z, \tau) = \frac{1}{C(x, z)} \prod_{j=1}^{\mid \tau \mid} \lambda_{x_{\tau_j}}(z_{\tau_j}) \int_{\{u < z_{\tau_j}^c\}} \lambda_{x_{\tau_j}^c|x_{\tau_j}, z_{\tau_j}}(u)\,du,$$

where the normalization constant $C(x, z)$ is given by

$$C(x, z) = \sum_{\theta \in \mathcal{P}_k} \prod_{j=1}^{\mid \theta \mid} \lambda_{x_{\theta_j}}(z_{\theta_j}) \int_{\{u < z_{\theta_j}^c\}} \lambda_{x_{\theta_j}^c|x_{\theta_j}, z_{\theta_j}}(u)\,du,$$

and $\mid \tau \mid$ is the “size” of the partition $\tau$. 

- Conditional simulations of max-stable processes
Step 2: The extremal functions

☐ Given \( \tau = (\tau_1, \ldots, \tau_\ell) \), draw \( \ell \) independent random vectors \( \varphi_1^+(s), \ldots, \varphi_\ell^+(s) \) from the distribution

\[
\Pr \left[ \varphi_j^+(s) \in dv_j \right] = \frac{1}{C_j} \left\{ \int \mathbb{1}_{\{u < z_{c_j}\}} \lambda(s, x_{\tau_j} | x_{\tau_j}, z_{\tau_j} \left( v_j, u \right)) \, du \right\} dv_j,
\]

where \( \mathbb{1}_{\{\cdot\}} \) is the indicator function and

\[
C_j = \int \mathbb{1}_{\{u < z_{c_j}\}} \lambda(s, x_{\tau_j} | x_{\tau_j}, z_{\tau_j} \left( v_j, u \right)) \, du \, dv_j.
\]

☐ Define the random vector

\[
Z^+(s) = \max_{j=1, \ldots, \ell} \varphi_j^+(s), \quad s \in \mathcal{X}^m.
\]
□ Independently

\[ Z^-(s) = \max_{\varphi \in \Phi} \varphi(s) \mathbb{1}_{\{\varphi(s) < z\}}, \quad s \in \mathcal{X}^m. \]
Independent

\[ Z^- (s) = \max_{\varphi \in \Phi} \varphi(s) 1_{\{\varphi(s) < z\}}, \quad s \in \mathcal{X}^m. \]

Then provided \( \Phi \) is regular, the random vector

\[ \tilde{Z}(s) = \max\{Z^+(s), Z^-(s)\} \]

follows the conditional distribution of \( Z(s) \) given \( Z(x) = z \).
The conditional cumulative distribution function is

\[
\Pr\{Z(s) \leq a \mid Z(x) = z\} = \left\{ \sum_{\tau \in \mathcal{P}_k} \pi_{x}(z, \tau) \prod_{j=1}^{|	au|} F_{\tau, j}(a) \right\} \frac{\Pr[Z(s) \leq a, Z(x) \leq z]}{\Pr[Z(x) \leq z]},
\]

where

\[
F_{\tau, j}(a) = \frac{\int_{\{y < z_{\tau_j} \cap u < a\}} \lambda(s, x_{\tau_j} | x_{\tau_j}, z_{\tau_j} (u, y)) dy du}{\int_{\{y < z_{\tau_j}\}} \lambda_{\tau_j} | x_{\tau_j}, z_{\tau_j} (y) dy}.
\]
The conditional cumulative distribution function is

\[
\Pr \{ Z(s) \leq a \mid Z(x) = z \} = \left\{ \sum_{\tau \in \mathcal{D}_k} \pi_x(z, \tau) \prod_{j=1}^{|	au|} F_{\tau, j}(a) \right\} \frac{\Pr[Z(s) \leq a, Z(x) \leq z]}{\Pr[Z(x) \leq z]},
\]

where

\[
F_{\tau, j}(a) = \frac{\int_{\{y < z_{r_j}, u < a\}} \lambda(s, x_{r_j}) | x_{r_j}, z_{r_j} \rangle (u, y) dy du}{\int_{\{y < z_{r_j}\}} \lambda_{r_j} | x_{r_j}, z_{r_j} \rangle (y) dy}.
\]

**Remark.** It is “clear” that \( Z(\cdot) \mid \{Z(x) = z\} \) is not max-stable.
Example 1 (Brown–Resnick process).

\[
Z(x) = \max_{i \geq 1} \zeta_i \exp \{ \epsilon_i (x) - \gamma(x) \}, \quad x \in \mathcal{X}.
\]

The intensity function is

\[
\lambda_x(z) = C_x \exp \left( -\frac{1}{2} \log z^T Q_x \log z + L_x \log z \right) \prod_{i=1}^k z_i^{-1}, \quad z \in (0, \infty)^k,
\]

and the conditional intensity function is

\[
\lambda_{s|x,z}(u) = (2\pi)^{-m/2} |\Sigma_{s|x}|^{-1/2} \exp \left\{ -\frac{1}{2} (\log u - \mu_{s|x,z})^T \Sigma_{s|x}^{-1} (\log u - \mu_{s|x,z}) \right\} \prod_{i=1}^m u_i^{-1},
\]

i.e., the extremal functions are log-Normal processes.
### Example 2 (Schlather process).

$$Z(x) = \sqrt{2\pi} \max_{i \geq 1} \zeta_i \max\{0, \varepsilon_i(x)\}, \quad x \in \mathcal{X}.$$  

The intensity function is

$$\lambda_x(z) = \pi^{-(k-1)/2} |\Sigma_x|^{-1/2} a_x(z)^{(k+1)/2} \Gamma\left(\frac{k+1}{2}\right), \quad z \in \mathbb{R}^k,$$

where $a_x(z) = z^T \Sigma_x^{-1} z$, and the conditional intensity function is

$$\lambda_{s|x,z}(u) = \pi^{-m/2} (k+1)^{-m/2} |\tilde{\Sigma}|^{-1/2} \left\{ 1 + \frac{(u - \mu)^T \tilde{\Sigma}^{-1} (u - \mu)}{k + 1} \right\}^{-(m+k+1)/2} \Gamma\left(\frac{m+k+1}{2}\right) \Gamma\left(\frac{k+1}{2}\right),$$

i.e., the extremal functions are Student processes.
2. Markov chain Monte–Carlo sampler
(for Step 1)
Do you recognize these numbers?

| 1. Conditional distributions | 2 | 5 | 15 |
|------------------------------|---|---|----|
| 1                            | 1 | 2 | 5 |
| 52                           | 203| 877| 4140|
| 115975                       | 678570| 4213597| 27644437|
| 1382958545                   | 10480142147| 82864869804| 682076806159|
| ...                          | ...| ...| ...|

… the state space $P_k$ isn't! (really?)

3. Simulation Study

4. Applications
Do you recognize these numbers?

|   | 1  | 1  | 2  | 5  | 15 |
|---|----|----|----|----|----|
|   | 52 | 203| 877| 4140| 21147|
|   | 115975 | 678570 | 4213597 | 27644437 | 190899322|
|   | 1382958545 | 10480142147 | 82864869804 | 682076806159 | 5832742205057|
|   | ... | ... | ... | ... | ...|

These are the first 20 Bell numbers.

Remark. Recall that Bell($k$) is the number of partitions of a set with $k$ elements.

# hitting scenarios = Card ($\mathcal{P}_k$) = Bell($k$)
In Step 1, we need to sample from a discrete distribution whose state space is $\mathcal{P}_k$, i.e., all possible hitting scenarios.

Combinatorial explosion

Hence we cannot compute $C(x,z)$ in

$$
\pi_x(z, \tau) = \frac{1}{C(x,z)} \prod_{j=1}^{\tau} \lambda_{x_{\tau_j}}(z_{\tau_j}) \int_{\{u < z_{\tau_j} \}} \lambda_{x_{\tau_j}^c|x_{\tau_j},z_{\tau_j}}(u) du.
$$
Computational burden

□ In Step 1, we need to sample from a discrete distribution whose state space is $\mathcal{P}_k$, i.e., all possible hitting scenarios.

Combinatorial explosion

Hence we cannot compute $C(x, z)$ in

$$\pi_x(z, \tau) = \frac{1}{C(x, z)} \prod_{j=1}^{||\tau||} \lambda_{x_{\tau_j}}(z_{\tau_j}) \int_{\{u < z_{\tau_j}^c\}} \lambda_{x_{\tau_j}^c | x_{\tau_j}, z_{\tau_j}}(u) du.$$ 

Use of MCMC samplers to sample from the target $\pi_x(z, \cdot)$.

□ We will use a Gibbs sampler that generates a Markov chain

$$\{\theta_n \in \mathcal{P}_k : n \in \mathbb{N}\}$$

whose invariant distribution is $\pi_x(z, \cdot)$. 

Conditional simulations of max-stable processes
Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions

\[ \Pr(\theta \in \cdot \mid \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi_x(z, \cdot), \quad j = 1, \ldots, k, \]

where \( \tau_{-j} \) drops the \( j \)-th location \( x_j \) in \( \tau \).
Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions

\[
\Pr(\theta \in \cdot \mid \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi_x(z, \cdot), \quad j = 1, \ldots, k,
\]

where \(\tau_{-j}\) drops the \(j\)-th location \(x_j\) in \(\tau\).

\[
\theta_0 : \quad \{x_1, x_3\} \quad \{x_2, x_5\} \quad \{x_4\}
\]
Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions

\[ \Pr(\theta \in \cdot | \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi_{\mathbf{z}}(\mathbf{z}, \cdot), \quad j = 1, \ldots, k, \]

where \(\tau_{-j}\) drops the \(j\)-th location \(x_j\) in \(\tau\).

\[ \theta_0 : \quad \{x_1, x_3\} \quad \{x_2, x_5\} \quad \{x_4\} \]

\[ \{x_1, x_3\} \quad \{x_2, x_5\} \quad \{x_4\} \]
Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions

$$\Pr(\theta \in \cdot | \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi_x(z, \cdot), \quad j = 1, \ldots, k,$$

where $\tau_{-j}$ drops the $j$-th location $x_j$ in $\tau$.
Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions

\[ \Pr(\theta \in \cdot \mid \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi(x(z, \cdot)), \quad j = 1, \ldots, k, \]

where \( \tau_{-j} \) drops the \( j \)-th location \( x_j \) in \( \tau \).

\[ \theta_0: \{x_1, x_3\} \quad \{x_2, x_5\} \quad \{x_4\} \]

\[ \{x_1, x_3\} \quad \{x_2\} \quad \{\{x_5\}\} \quad \{x_4\} \]
Full conditional distributions

Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions

\[ \Pr(\theta \in \cdot \mid \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi_x(z, \cdot), \quad j = 1, \ldots, k, \]

where \( \tau_{-j} \) drops the \( j \)-th location \( x_j \) in \( \tau \).

\[ \theta_0: \quad \{x_1, x_3\} \quad \{x_2, x_5\} \quad \{x_4\} \]

\[ \theta_1: \quad \{x_1, x_3\} \quad \{x_2\} \quad \{x_5\} \quad \{x_4\} \]
Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions:

\[ \Pr(\theta \in \cdot | \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi_x(z, \cdot), \quad j = 1, \ldots, k, \]

where \( \tau_{-j} \) drops the \( j \)-th location \( x_j \) in \( \tau \).

\[ \theta_0 : \{x_1, x_3\} \quad \{x_2, x_5\} \quad \{x_4\} \]

\[ \theta_1 : \{x_1, x_3\} \quad \{x_2\} \quad \{x_5\} \quad \{x_4\} \]

\[ \{x_1, x_3\} \quad \{x_2\} \quad \{x_5\} \quad \{x_4\} \]
Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions

$$\Pr(\theta \in \cdot \mid \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi_x(z, \cdot), \quad j = 1, \ldots, k,$$

where $\tau_{-j}$ drops the $j$-th location $x_j$ in $\tau$.

\[
\begin{align*}
\theta_0 : & \quad \{x_1, x_3\} & \quad \{x_2, x_5\} & \quad \{x_4\} \\
\theta_1 : & \quad \{x_1, x_3\} & \quad \{x_2\} & \quad \{x_5\} & \quad \{x_4\}
\end{align*}
\]
Full conditional distributions

Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions

\[ \Pr(\theta \in \cdot \mid \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi_x(z, \cdot), \quad j = 1, \ldots, k, \]

where \( \tau_{-j} \) drops the \( j \)-th location \( x_j \) in \( \tau \).

\[ \theta_0: \quad \{x_1, x_3\} \quad \{x_2, x_5\} \quad \{x_4\} \]
\[ \theta_1: \quad \{x_1, x_3\} \quad \{x_2\} \quad \{x_5\} \quad \{x_4\} \]

\[ \{x_1, x_3, x_4\} \quad \{x_2\} \quad \{x_5\} \quad \emptyset \]
Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions

\[ \Pr(\theta \in \cdot \mid \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi_x(z, \cdot), \quad j = 1, \ldots, k, \]

where \( \tau_{-j} \) drops the \( j \)-th location \( x_j \) in \( \tau \).

\[ \theta_0: \{x_1, x_3\} \quad \{x_2, x_5\} \quad \{x_4\} \]
\[ \theta_1: \{x_1, x_3\} \quad \{x_2\} \quad \{x_5\} \quad \{x_4\} \]
\[ \theta_2: \{x_1, x_3, x_4\} \quad \{x_2\} \quad \{x_5\} \]
Full conditional distributions

Our (random scan) Gibbs sampler amounts to sample from the full conditional distributions

$$\Pr(\theta \in \cdot \mid \theta_{-j} = \tau_{-j}), \quad \theta \sim \pi_x(z, \cdot), \quad j = 1, \ldots, k,$$

where $\tau_{-j}$ drops the $j$-th location $x_j$ in $\tau$.

$$\begin{align*}
\theta_0 : & \{x_1, x_3\} & \{x_2, x_5\} & \{x_4\} \\
\theta_1 : & \{x_1, x_3\} & \{x_2\} & \{x_5\} & \{x_4\} \\
\theta_2 : & \{x_1, x_3, x_4\} & \{x_2\} & \{x_5\} & \{x_4\} \\
\vdots \\
\theta_N : & \{x_1, x_5\} & \{x_2\} & \{x_3, x_4\}
\end{align*}$$
If the full conditional distributions are nice, . . .

For all $\tau^* \in \mathcal{P}_k$ such that $\tau^*_j = \tau_j$,

$$\Pr[\theta = \tau^* | \theta_{-j} = \tau_{-j}] = \frac{\pi_x(z, \tau^*)}{\sum_{\tilde{\tau} \in \mathcal{P}_k} \pi_x(z, \tilde{\tau}) \mathbf{1}_{\tilde{\tau}_{-j} = \tau_{-j}}} \propto \frac{\prod_{j=1}^{\tau^*} w_{\tau^*, j}}{\prod_{j=1}^{\tau} w_{\tau, j}},$$

where $w_{\tau, j} = \lambda(x_{\tau_j}(z_{\tau_j}) \int_{\{u < z_{\tau_j}\}} \lambda(x_{\tau^*_j} | x_{\tau_j}, z_{\tau_j})(u) du.$

In particular at most 4 weights $w_{\cdot, \cdot}$ need to be evaluated and the Gibbs sampler is especially convenient!
But how do I implement a Gibbs sampler whose states are partitions of a set???
... the state space $\mathcal{P}_k$ isn’t! (really?)

But how do I implement a Gibbs sampler whose states are partitions of a set???

Lemma 1. There is a one-one mapping between $\mathcal{P}_k$ and

$$\mathcal{P}_k^* = \left\{ (a_1, \ldots, a_k), \forall i \in \{2, \ldots, k\}: a_1 \leq a_i \leq \max_{1 \leq j < i} a_j + 1, a_i \in \mathbb{Z} \right\},$$

where $a_1 = 1$ by convention.

Example 3. $([x_1, x_2], [x_3])$ is identified to $(1, 1, 2)$ while $([x_1, x_3], [x_2])$ is identified to $(1, 2, 1)$. 
3. Simulation Study

Checking the Gibbs sampler
What we expect
Test cases
Test case: Schlather
What we get
Spatial dependence
CPU times
Checking Step 1, i.e., the Gibbs sampler (i)

1. Conditional distributions
2. MCMC sampler
3. Simulation Study
   - Checking the Gibbs sampler
   - What we expect
   - Test cases
     - Test case: Schlather
   - What we get
   - Spatial dependence
   - CPU times
4. Applications

---

**Figure 4:** Left: Trace plot of one simulated Markov chain with $k = 5$ conditioning locations. Right: Comparison of the theoretical probabilities $\{\pi_x(z, \tau), \tau \in \mathcal{P}_k\}$ to the empirical ones estimated from the simulated Markov chain.
What we expect

- Less variability in regions close to some conditioning points;
- The coverage is OK, i.e., pointwise confidence intervals have the nominal coverage;
- “Unconditional like behavior” in regions far away from any conditioning point.
Test case: Brown–Resnick

Table 1: Spatial dependence structures of Brown–Resnick processes with (semi) variogram \( \gamma(h) = (h/\lambda)^{\kappa} \). The variogram parameters are set to ensure that the extremal coefficient function satisfies \( \theta(115) = 1.7 \).

| Sample path properties | \( \gamma_1 \): Very wiggly | \( \gamma_2 \): Wiggly | \( \gamma_3 \): Smooth |
|-------------------------|-----------------------------|------------------------|------------------------|
| \( \lambda \)           | 25                          | 54                     | 69                     |
| \( \kappa \)            | 0.5                         | 1.0                    | 1.5                    |

Figure 5: Three realizations of a Brown–Resnick process with standard Gumbel margins and (semi) variograms \( \gamma_1 \), \( \gamma_2 \) and \( \gamma_3 \). The squares correspond to the 15 conditioning values that will be used in the simulation study. The right panel shows the associated extremal coefficient functions.
Table 2: Spatial dependence structures of Schlather processes with correlation function \( \rho(h) = \exp\{- (h/\lambda)^\kappa\} \). The correlation function parameters are set to ensure that the extremal coefficient function satisfies \( \theta(100) = 1.5 \).

| Sample path properties | \( \rho_1 \): Very wiggly | \( \rho_2 \): Wiggly | \( \rho_3 \): Smooth |
|-------------------------|--------------------------|-----------------|------------------|
| \( \lambda \)           | 208                      | 144             | 128              |
| \( \kappa \)             | 0.5                      | 1.0             | 1.5              |

Figure 6: Three realizations of a Schlather process with standard Gumbel margins and correlation functions \( \rho_1 \), \( \rho_2 \) and \( \rho_3 \). The squares correspond to the 15 conditioning values that will be used in the simulation study. The right panel shows the associated extremal coefficient functions.
What we get: Brown–Resnick

1. Conditional distributions
2. MCMC sampler
3. Simulation Study
   Checking the Gibbs sampler
   What we expect
   Test cases
   Test case: Schlather
   ▶ What we get
   Spatial dependence
   CPU times
4. Applications

Figure 7: Pointwise sample quantiles (0.025, 0.5, 0.975) estimated from 1000 conditional simulations of Brown–Resnick processes.
What we get: Schlather

1. Conditional distributions
2. MCMC sampler
3. Simulation Study
   Checking the Gibbs sampler
   What we expect
   Test cases
   Test case: Schlather

What we get: Spatial dependence
CPU times

4. Applications

Figure 8: Pointwise sample quantiles (0.025, 0.5, 0.975) estimated from 1000 conditional simulations of Schlather processes.
One last point ;-) (text continues)

- Is the spatial dependence correct?
- Want to compare the theoretical extremal coefficient function $\theta(\cdot)$ to the pairwise extremal coefficient estimates.

But recall, $Z(\cdot) | \{Z(x) = z\}$ is not max-stable and the extremal coefficient function does not exist!!!
One last point ;-) 

- Is the spatial dependence correct?
- Want to compare the theoretical extremal coefficient function \( \theta(\cdot) \) to the pairwise extremal coefficient estimates.

But recall, \( Z(\cdot) \mid \{ Z(x) = z \} \) is not max-stable and the extremal coefficient function does not exist!!!

Since

\[
f(x) = \int f(x \mid y) f(y) dy,
\]

and to recover the max-stability property, we

1. Generate 1000 independent conditional events;
2. For each such conditional event, one conditional realization.
Checking the spatial dependence structure

1. Conditional distributions
2. MCMC sampler
3. Simulation Study
   Checking the Gibbs sampler
   What we expect
   Test cases
   Test case: Schlather
   What we get
   Spatial dependence
   CPU times
4. Applications

Figure 9: Comparison of the extremal coefficient estimates (using a binned F-madogram with 250 bins) and the theoretical extremal coefficient function for a varying number of conditioning locations and different (semi) variograms. From left to right, $k = 5, 10, 15$. The ‘o’, ‘+’ and ‘x’ symbols correspond respectively to $\gamma_1$, $\gamma_2$ and $\gamma_3$. The solid, dashed and dotted grey lines correspond to the theoretical extremal coefficient functions for $\gamma_1$, $\gamma_2$ and $\gamma_3$. 
Checking the spatial dependence structure

1. Conditional distributions
2. MCMC sampler
3. Simulation Study
   Checking the Gibbs sampler
   What we expect
   Test cases
   Test case: Schlather
   What we get
   Spatial dependence
   CPU times
4. Applications

Figure 10: Comparison of the extremal coefficient estimates (using a binned $F$-madogram with 250 bins) and the theoretical extremal coefficient function for a varying number of conditioning locations and different correlation functions. From left to right, $k = 5, 10, 15$. The 'o', '+' and 'x' symbols correspond respectively to $\rho_1$, $\rho_2$ and $\rho_3$. 
Table 3: Timings† for conditional simulations of max-stable processes on a 50 × 50 grid defined on the square [0, 100 × 2^{1/2}]^2 for a varying number k of conditioning locations uniformly distributed over the region. The times, in seconds, are mean values over 100 conditional simulations; standard deviations are reported in brackets.

|          | Brown–Resnick: γ(h) = (h/25)^{0.5} | Schlather: ρ(h) = exp \{- (h/208)^{0.50}\} |
|----------|-------------------------------------|--------------------------------------|
|          | Step 1     | Step 2    | Step 3   | Overall | Step 1     | Step 2    | Step 3   | Overall   |
| k = 5    | 0.21 (0.01) | 49 (11)   | 1.4 (0.1)| 50 (11) | 1.4 (0.02) | 1.9 (0.7) | 0.9 (0.3)| 4.2 (0.8) |
| k = 10   | 8 (2)      | 76 (18)   | 1.4 (0.1)| 85 (19) | 12 (4)     | 2.4 (0.8) | 1.0 (0.3)| 15 (4)    |
| k = 25   | 95 (38)    | 117 (30)  | 1.4 (0.1)| 214 (61)| 86 (42)    | 4 (1)     | 1.0 (0.3)| 90 (43)   |
| k = 50   | 583 (313)  | 348 (391) | 1.5 (0.1)| 931 (535)| 367 (222)  | 62 (113)  | 1.0 (0.3)| 430 (262) |

†Conditional simulations with k = 5 do not use a Gibbs sampler.
4. Applications

Conditional simulations of max-stable processes
We re-analyze the data of Davison et al. (2012), i.e., summer precipitation around Zurich.

Figure 11: Left: Map of Switzerland showing the stations of the 24 rainfall gauges used for the analysis, with an insert showing the altitude. The station marked with a blue square corresponds to Zurich. Middle: Summer maximum daily rainfall values for 1962–2008 at Zurich. Right: Comparison between the pairwise extremal coefficient estimates for the 51 original weather stations and the extremal coefficient function derived from a fitted Brown–Resnick processes having (semi) variogram $\gamma(h) = (h/\lambda)^K$. The grey points are pairwise estimates; the black ones are binned estimates and the red curve is the fitted extremal coefficient function.
We fit a Brown–Resnick process by maximizing the pairwise likelihood with the following trend surfaces

\[
\eta(x) = \beta_{0,\eta} + \beta_{1,\eta}\text{lon}(x) + \beta_{2,\eta}\text{lat}(x),
\]

\[
\sigma(x) = \beta_{0,\sigma} + \beta_{1,\sigma}\text{lon}(x) + \beta_{2,\sigma}\text{lat}(x),
\]

\[
\xi(x) = \beta_{0,\xi},
\]

where \(\eta(x), \sigma(x), \xi(x)\) are the location, scale and shape parameters of the generalized extreme value distribution and \(\text{lon}(x), \text{lat}(x)\) the longitude and latitude of the stations at which the data are observed.
We fit a Brown–Resnick process by maximizing the **pairwise likelihood** with the following trend surfaces

\[
\eta(x) = \beta_{0,\eta} + \beta_{1,\eta}\text{lon}(x) + \beta_{2,\eta}\text{lat}(x),
\]
\[
\sigma(x) = \beta_{0,\sigma} + \beta_{1,\sigma}\text{lon}(x) + \beta_{2,\sigma}\text{lat}(x),
\]
\[
\xi(x) = \beta_{0,\xi},
\]

where \(\eta(x), \sigma(x), \xi(x)\) are the location, scale and shape parameters of the generalized extreme value distribution and \(\text{lon}(x), \text{lat}(x)\) the longitude and latitude of the stations at which the data are observed.

- **Take as conditional event the values observed during year 2000.**
- **Simulate a Markov chain of length 15000 from** \(\pi_x(z, \cdot)\) **to estimate the distribution of the partition size.**
- **And perform a bunch of conditional simulations from our fitted model to get a nice map!**
Table 4: Empirical distribution of the partition size for the rainfall data estimated from a simulated Markov chain of length 15000.

| Partition size | 1    | 2    | 3    | 4    | 5    | 6    | 7–24 |
|----------------|------|------|------|------|------|------|------|
| Empirical probabilities (%) | 66.2 | 28.0 | 4.8  | 0.5  | 0.2  | 0.2  | <0.05 |

- Around 65% of the time, the maxima at the 24 locations are a consequence of a single extremal function, i.e., only one storm, and around 30% of the time of two extremal functions.
- Focusing only on partitions of size 2, around 65% of the time at least one of the four up-north locations are impacted by a first extremal function while the remaining 20 stations are always influenced by a second extremal function.
Figure 12: From left to right, maps on a $50 \times 50$ grid of the pointwise $0.025$, $0.5$ and $0.975$ sample quantiles for rainfall (mm) obtained from 10000 conditional simulations of Brown–Resnick processes having semi variogram $\gamma(h) = (h/38)^{0.69}$. The rightmost panel plots the ratio of the width of the pointwise confidence intervals with and without taking estimation uncertainties into account. The squares show the conditional locations.
We re-analyze the data of Davison and Gholamrezaee (2012), i.e., annual maxima temperature in Switzerland.

Figure 13: Left: Topographical map of Switzerland showing the sites and altitudes in metres above sea level of 16 weather stations for which annual maxima temperature data are available. Middle: Times series of the daily maxima temperatures at the 16 weather stations for year 2003. The 'o', '+' and 'x' symbols indicate the annual maxima that occurred the 11th, 12th and 13th of August respectively. Right: Comparison between the fitted extremal coefficient function from a Schlather process (solid red line) and the pairwise extremal coefficient estimates (gray circles). The black circles denote binned estimates with 16 bins.
We fit a Schlather process by maximizing the **pairwise likelihood** with the following trend surfaces

\[
\eta(x) = \beta_{0,\eta} + \beta_{1,\eta}\text{alt}(x), \\
\sigma(x) = \beta_{0,\sigma}, \\
\xi(x) = \beta_{0,\xi} + \beta_{1,\xi}\text{alt}(x),
\]

where \(\eta(x), \sigma(x), \xi(x)\) are the location, scale and shape parameters of the generalized extreme value distribution and \(\text{alt}(x)\) the altitude of the stations at which the data are observed.
We fit a Schlather process by maximizing the **pairwise likelihood** with the following trend surfaces

\[
\eta(x) = \beta_{0,\eta} + \beta_{1,\eta}\text{alt}(x), \\
\sigma(x) = \beta_{0,\sigma}, \\
\xi(x) = \beta_{0,\xi} + \beta_{1,\xi}\text{alt}(x),
\]

where \(\eta(x), \sigma(x), \xi(x)\) are the location, scale and shape parameters of the generalized extreme value distribution and \(\text{alt}(x)\) the altitude of the stations at which the data are observed.

Take as conditional event the values observed during the 2003 European heatwave.

Simulate a Markov chain of length 10000 from \(\pi_x(z, \cdot)\) to estimate the distribution of the partition size.

And perform a bunch of conditional simulations from our fitted model to get a nice map!
Distribution of the partition size

| Partition size | 1   | 2   | 3   | 4   | 5–16 |
|---------------|-----|-----|-----|-----|------|
| Empirical probabilities (%) | 2.47 | 21.55 | 64.63 | 10.74 | 0.61 |

- Around **90% of the time**, the conditional simulations are a consequence of **at most 3 extremal functions**;
- Inspecting the data, we found that the annual maxima in 2003 occurred between the 11th and 13rd of August
Figure 14: Left: Topographical map of Switzerland showing the sites and altitudes in metres above sea level of 16 weather stations for which annual maxima temperature data are available. Right: Map of temperature anomalies (°C), i.e., the difference between the pointwise medians obtained from 10000 conditional simulations and unconditional medians estimated from the fitted Schlather process.

- As expected the largest deviations occur in the plateau region of Switzerland
- The differences range between 2.5°C and 4.75°C
## What’s next?

- Conditional distributions
- MCMC sampler
- Simulation Study
- Applications
  - Precipitation
  - Temperature

- Inference for max-stable processes based on the (full) likelihood
- Conditional distributions: grid cell conditioning
- Statistical modeling with Pareto processes
What's next?

- Inference for max-stable processes based on the (full) likelihood
- Conditional distributions: grid cell conditioning
- Statistical modeling with Pareto processes

THANK YOU!

Dombry, C. Řyi-Minko, F. and Ribatet, M. *Conditional simulation of max-stable processes*. Biometrika (in press). (*doi: 10.1093/biomet/ass067*)