Monte Carlo Methods for Spatial Extremes

Jose Blanchet (joint with Liu, Dieker, Mikosch).

Columbia Departments of IEOR and Statistics

What is our Goal?

Our goal is to enable efficient Monte Carlo of extreme events in space...

Our focus here is on Max-stable Fields.

Elements of Extreme Value Analysis

- A little story... Gumbel.

Elements of Extreme Value Analysis

- A little story... Gumbel.

- **Textbook Example:** Given storm surges observed over the last 100 yrs in NYC, what's the maximum height of a storm surge which gets exceeded only about once in 1000 yrs?

Elements of Extreme Value Analysis

- A little story... Gumbel.
- **Textbook Example:** Given storm surges observed over the last 100 yrs in NYC, what's the maximum height of a storm surge which gets exceeded only about once in 1000 yrs?
- **Answer:** One must *extrapolate*...

Extreme Value Theory (EVT): The IID case

Theorem (Fisher-Tippet-Gnedenko)

Suppose $X_1, ..., X_n$ *is an IID sequence and define* $Z_n = \max\{X_1, ..., X_n\}$. **If there exists** $\{(a_n, b_n)\}_{n\geq 1}$ **deterministic numbers such that**

$$Z_n \overset{D}{\approx} b_n M + a_n,$$

then M **must be a max-stable distribution.**

Max-stable Distributions

- A max-stable r.v. M is characterized by the fact that if $M_1, M_2, ..., M_n$ are i.i.d. copies of M then

$$M \overset{D}{=} \max_{i=1}^{n} \left(M_i - a_n \right) / b_n.$$

Max-stable Distributions

- A max-stable r.v. M is characterized by the fact that if $M_1, M_2, ..., M_n$ are i.i.d. copies of M then

$$M \stackrel{D}{=} \max_{i=1}^{n} (M_i - a_n) / b_n.$$

- We also have that

$$\begin{aligned} P(M \le x) &= \exp\left(-(1+\gamma x)^{-1/\gamma}\right) \quad 1+\gamma x > 0. \\ P(M \le x) &= \exp(-\exp(-x)); \gamma = 0 \text{ (the Gumbel case)}. \end{aligned}$$

Max-stable Distributions

- A max-stable r.v. M is characterized by the fact that if $M_1, M_2, ..., M_n$ are i.i.d. copies of M then

$$M \stackrel{D}{=} \max_{i=1}^{n} (M_i - a_n) / b_n.$$

- We also have that

$$\begin{aligned} P(M \leq x) &= \exp\left(-(1+\gamma x)^{-1/\gamma}\right) \quad 1+\gamma x > 0. \\ P(M \leq x) &= \exp(-\exp(-x))\,;\ \gamma = 0 \text{ (the Gumbel case).} \end{aligned}$$

- In practice $Z_m = \max\{X_{(m-1)n+1}, ..., X_{mn}\} \stackrel{D}{=} a + bM_m$, apply MLE to estimate a, b, γ & ready to extrapolate...

Max-stable Distributions

- A max-stable r.v. M is characterized by the fact that if $M_1, M_2, ..., M_n$ are i.i.d. copies of M then

$$M \stackrel{D}{=} \max_{i=1}^{n} \left(M_i - a_n\right) / b_n.$$

- We also have that

$$\begin{aligned} P\left(M \leq x\right) &= \exp\left(-\left(1+\gamma x\right)^{-1/\gamma}\right) \quad 1+\gamma x > 0. \\ P\left(M \leq x\right) &= \exp\left(-\exp\left(-x\right)\right);\ \gamma = 0 \text{ (the Gumbel case).} \end{aligned}$$

- In practice $Z_m = \max\{X_{(m-1)n+1}, ..., X_{mn}\} \stackrel{D}{=} a + bM_m$, apply MLE to estimate a, b, γ & ready to extrapolate...

- **But in many cases it is important to account for spatial dependence...**

What is a Catastrophe Bond?

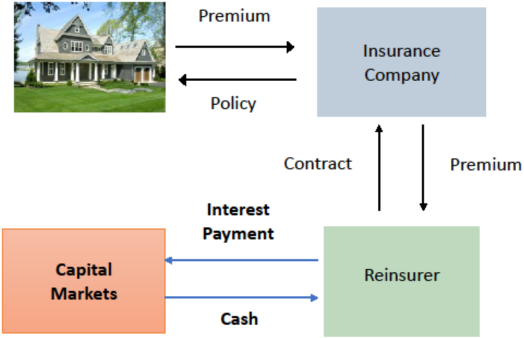
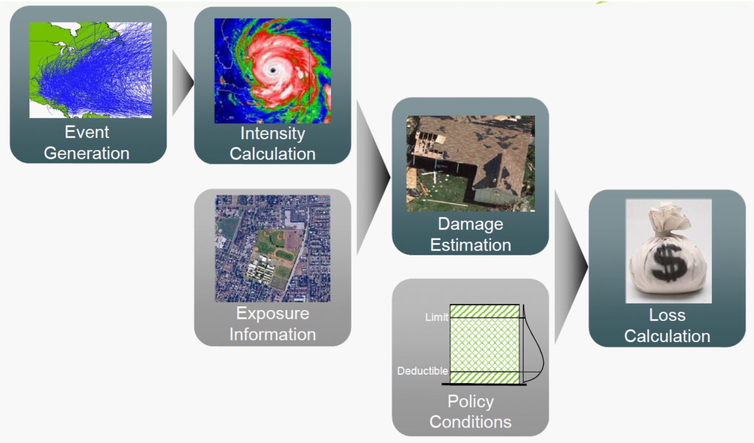


Diagram Illustrating Catastrophe Bonds Risk Calculation



Case Study: MetroCat Re Ltd

- The MTA (New York & New Jersey authorities) obtained parametric triggered storm surge.

Case Study: MetroCat Re Ltd

- The MTA (New York & New Jersey authorities) obtained parametric triggered storm surge.
- Motivated by high insurance costs after Hurricane Sandy in 10/2012.

Case Study: MetroCat Re Ltd

- The MTA (New York & New Jersey authorities) obtained parametric triggered storm surge.
- Motivated by high insurance costs after Hurricane Sandy in 10/2012.
- The parametric trigger is based on weighted average of maximum water levels in several locations.

Max-stable Random Fields

- How does one define natural models of extremes in space?

Max-stable Random Fields

- How does one define natural models of extremes in space?
- **Definition:** $M(\cdot)$ is a max-stable field if

$$M(\cdot) \stackrel{D}{=} \max_{i=1}^{n} (M_i(\cdot) - a_n(\cdot)) / b_n(\cdot),$$

for i.i.d. $M_1(\cdot), ..., M_n(\cdot)$ of $M(\cdot)$ and some $a_n(\cdot), b_n(\cdot)$.

Max-stable Random Fields

- How does one define natural models of extremes in space?

- **Definition:** $M(\cdot)$ is a max-stable field if

$$M(\cdot) \stackrel{D}{=} \max_{i=1}^{n} \left(M_i(\cdot) - a_n(\cdot)\right) / b_n(\cdot),$$

  for i.i.d. $M_1(\cdot), ..., M_n(\cdot)$ of $M(\cdot)$ and some $a_n(\cdot), b_n(\cdot)$.

- How to construct max-stable fields?

Characterizations of Max-stable Fields

- A large class of max-stable fields can be represented as:

$$M(t) = \sup_{n \geq 1} \{-\log(A_n) + X_n(t)\},$$

where $X_n(\cdot)$ is Gaussian.

Characterizations of Max-stable Fields

- A large class of max-stable fields can be represented as:

$$M(t) = \sup_{n \geq 1} \{-\log(A_n) + X_n(t)\},$$

  where $X_n(\cdot)$ is Gaussian.

- $A_n = n$-th arrival of Poisson process with rate 1 (independent of $X_n(\cdot)$).

Characterizations of Max-stable Fields

- A large class of max-stable fields can be represented as:

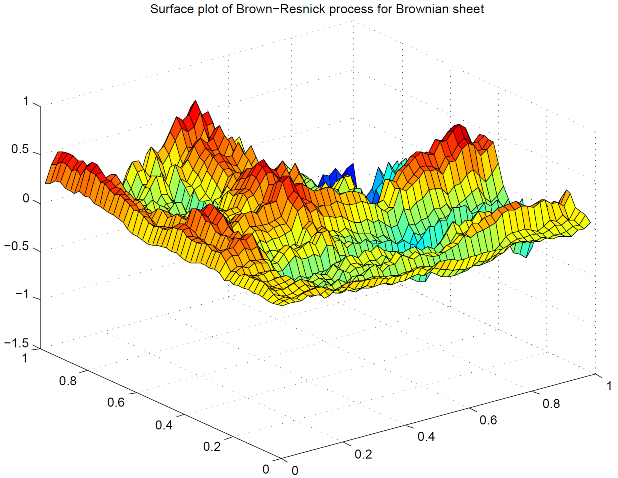
$$M(t) = \sup_{n\geq 1}\{-\log(A_n) + X_n(t)\},$$

  where $X_n(\cdot)$ is Gaussian.

- $A_n = n$-th arrival of Poisson process with rate 1 (independent of $X_n(\cdot)$).

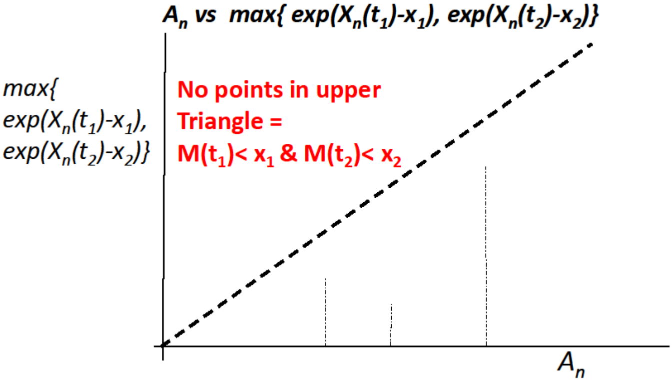
- Brown-Resnick, de Haan, Smith, Schlather, Kabluchko,...

This is How a Max-Stable Field Looks Like



Computing Joint CDFs

$$P\left(e^{M(t_1)} \leq e^{x_1}, e^{M(t_1)} \leq e^{x_2}\right) = P\left(\text{Poisson r.v. } = 0\right)$$
$$= \exp\left(-E\max\left\{\exp\left(X_n\left(t_1\right) - x_1\right), \exp\left(X_n\left(t_2\right) - x_2\right)\right\}\right)$$



Max-Stable Fields: Computational Challenges

- Joint densities for $M(t_1), ..., M(t_d)$ quickly become intractable ($d = 10$ contains more than 10^5 terms).

Max-Stable Fields: Computational Challenges

- Joint densities for $M(t_1), ..., M(t_d)$ quickly become intractable ($d = 10$ contains more than 10^5 terms).

- Likelihood methods applicable in low dimensions only.

Max-Stable Fields: Computational Challenges

- Joint densities for $M(t_1), \ldots, M(t_d)$ quickly become intractable ($d = 10$ contains more than 10^5 terms).
- Likelihood methods applicable in low dimensions only.
- In the end we want to efficiently evaluate quantities such as

$$P\left(\int_T w(t)\, M(t)\, dt > b\right), \quad \text{<–}\textbf{CAT Bond Example}$$

$$f_{M(t_1),\ldots,M(t_k)}(x_1, \ldots, x_k),$$

$$f_{M(t_1),\ldots,M(t_k)|M(s_1),\ldots,M(s_d)}(x_1, \ldots, x_k \mid z_1, \ldots, z_d).$$

Our Contributions

Theorem 1 (B., Dieker, Liu, Mikosch '16): Algorithm for sampling $M(t_1), \ldots, M(t_d)$ with virtually the same asymptotic complexity as sampling $X_1(t_1), \ldots, X_1(t_d)$ as $d \to \infty$.

In other words, $M(t_1), \ldots, M(t_d)$ *is basically as "easy" as* $X_1(t_1), \ldots, X_1(t_d)$...

Our Contributions

- **Theorem 2** (B. and Liu '16): Construction of a finite variance $W(y_1, ..., y_d)$ such that

$$f_{M(t_1),...,M(t_d)}(y_1, ..., y_d) = E(W(y_1, ..., y_d)).$$

The estimator takes $O\left(\varepsilon^{-2}\log^2(1/\varepsilon)\right)$ samples to achieve ε error.

Our Contributions

- **Theorem 2** (B. and Liu '16): Construction of a finite variance $W(y_1, ..., y_d)$ such that

$$f_{M(t_1),...,M(t_d)}(y_1, ..., y_d) = E(W(y_1, ..., y_d)).$$

The estimator takes $O\left(\varepsilon^{-2}\log^2(1/\varepsilon)\right)$ samples to achieve ε error.

- **Theorem 3** (B. and Liu '16): Construction of $W(y_1, ..., y_k, z_1, ..., z_d)$ such that

$$\begin{aligned} & f_{M(t_1),...,M(t_k)|M(s_1),...,M(s_d)}(y_1, ..., y_k \mid z_1, ..., z_d) \\ = \; & E(W(y_1, ..., y_k, z_1, ..., z_d)), \end{aligned}$$

similar complexity as Theorem 2.

A Quick Review of Current Methodology

- State-of-the-art: Exact sampling algorithms due to Dieker and Mikosch (2015), and Dombry, Engelke, Oesting (2016).

A Quick Review of Current Methodology

- State-of-the-art: Exact sampling algorithms due to Dieker and Mikosch (2015), and Dombry, Engelke, Oesting (2016).
- Summary of computational complexity (BDML = Our method):

$$\text{Dieker \& Mikosch} = \text{BDML} \quad \times d \times \log(d).$$
$$\text{DEO} = \text{BDML} \quad \times d.$$

A Quick Review of Current Methodology

- State-of-the-art: Exact sampling algorithms due to Dieker and Mikosch (2015), and Dombry, Engelke, Oesting (2016).

- Summary of computational complexity (BDML = Our method):

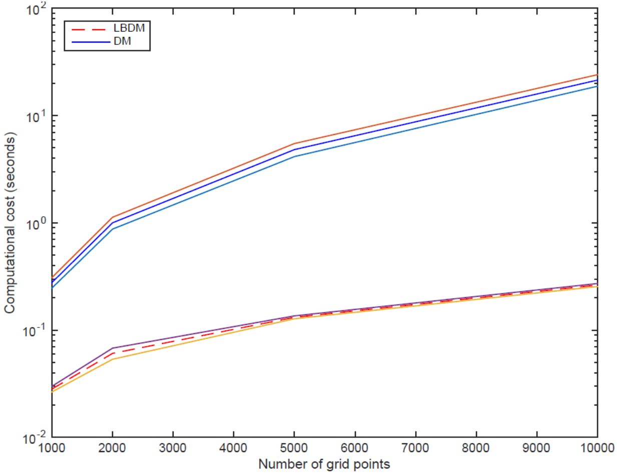
$$\text{Dieker \& Mikosch} = \text{BDML} \qquad \times d \times \log(d).$$
$$\text{DEO} = \text{BDML} \qquad \times d.$$

- For example if $X_n(\cdot)$ is Brownian motion our method takes $O(d)$ complexity to sample $M(t_1), ..., M(t_d)$. Both DM and DEO take at least $O(d^2)$ complexity.

Comparison vs Dieker & Mikosch

Log-log plot: Grid points vs time in seconds

RED = Our Method vs *BLUE = Alternative method.*



Optimal Simulation of Max-stable Processes

- Find two crucial quantities:

Optimal Simulation of Max-stable Processes

- Find two crucial quantities:
  1. N_X such that for $n \geq N_X$:

$$\max_{i=1}^{d} |X_n(t_i)| \leq \frac{1}{2} \log(n+1).$$

Optimal Simulation of Max-stable Processes

- Find two crucial quantities:
  1. N_X such that for $n \geq N_X$:
  $$\max_{i=1}^{d} |X_n(t_i)| \leq \frac{1}{2}\log(n+1).$$
  2. N_A such that for $n \geq N_A$:
  $$A_n > \frac{n+1}{2}.$$

Optimal Simulation of Max-stable Processes

- Find two crucial quantities:
  1. N_X such that for $n \geq N_X$:
  $$\max_{i=1}^{d} |X_n(t_i)| \leq \frac{1}{2} \log(n+1).$$
  2. N_A such that for $n \geq N_A$:
  $$A_n > \frac{n+1}{2}.$$

- Note that for $m \geq \max(N_A, N_X)$
$$\max_{n \geq m} \{X_n(t_i) - \log(A_n)\} \leq -\frac{\log(m+1)}{2} + \log(2).$$

Optimal Simulation of Max-stable Processes

- Find two crucial quantities:
  1. N_X such that for $n \geq N_X$:
  $$\max_{i=1}^{d} |X_n(t_i)| \leq \frac{1}{2}\log(n+1).$$
  2. N_A such that for $n \geq N_A$:
  $$A_n > \frac{n+1}{2}.$$
- Note that for $m \geq \max(N_A, N_X)$
$$\max_{n\geq m}\{X_n(t_i) - \log(A_n)\} \leq -\frac{\log(m+1)}{2} + \log(2).$$
- Need to detect when
$$-\log(m+1)/2 + \log(2) \leq \min_{i=1}^{d} X_1(t_i) - \log(A_1).$$

- If $X_n(\cdot)$ is continuous on the compact set T, then

$$\max_{t \in T} |X_n(t)| = O_p\left((\log(n))^{1/2}\right).$$

- If $X_n(\cdot)$ is continuous on the compact set T, then

$$\max_{t\in T}|X_n(t)| = O_p\left((\log(n))^{1/2}\right).$$

- By the LLNs

$$A_n/n \rightarrow 1.$$

Optimal Simulation of Max-stable Processes

- If $X_n(\cdot)$ is continuous on the compact set T, then

$$\max_{t \in T} |X_n(t)| = O_p\left((\log(n))^{1/2} \right).$$

- By the LLNs

$$A_n / n \longrightarrow 1.$$

- So detection of N_X and N_A should take $O(1)$ time as $d \longrightarrow \infty$.

The Milestone / Record Breaking Events Technique

- Used several times: B. & Sigman (2011), B. & Chen (2012), B. & Dong (2013), B. & Wallwater (2014), B., Chen & Dong (2015).

The Milestone / Record Breaking Events Technique

- Used several times: B. & Sigman (2011), B. & Chen (2012), B. & Dong (2013), B. & Wallwater (2014), B., Chen & Dong (2015).
- Borrow intuition from rare event simulation.

The Milestone / Record Breaking Events Technique

- Used several times: B. & Sigman (2011), B. & Chen (2012), B. & Dong (2013), B. & Wallwater (2014), B., Chen & Dong (2015).
- Borrow intuition from rare event simulation.
- Let us write $X_n = \{X_n(t_i)\}_{i=1}^d$ so $\|X_n\|_\infty = \max_{i=1}^d |X_n(t_i)|$.

The Milestone / Record Breaking Events Technique

- Used several times: B. & Sigman (2011), B. & Chen (2012), B. & Dong (2013), B. & Wallwater (2014), B., Chen & Dong (2015).
- Borrow intuition from rare event simulation.
- Let us write $X_n = \{X_n(t_i)\}_{i=1}^d$ so $\|X_n\|_\infty = \max_{i=1}^d |X_n(t_i)|$.
- Let $\tau_0 = n_0 - 1$ (n_0 chosen later) and let

$$\begin{aligned}
\tau_{k+1} &= \inf\{n > \tau_k : \|X_n\|_\infty > \frac{1}{2}\log(n+1)\}, \\
N_X &= \sup\{\tau_k : \tau_k < \infty\}.
\end{aligned}$$

Can we sample Bernoulli with $P(\tau_1 < \infty)$? Can we simulate $X_1(\cdot), ..., X_{\tau_1}(\cdot)$ given $\tau_1 < \infty$?

Borrow from Rare Event Simulation Intuition

- Say n_0 is large: How does the rare event

$$\{\tau_1 < \infty\} = \bigcup_{n=n_0}^{\infty} \left\{ \|X_n\|_\infty > \frac{1}{2} \log (n+1) \right\} \quad \text{occur?}$$

Borrow from Rare Event Simulation Intuition

- Say n_0 is large: How does the rare event

$$\{\tau_1 < \infty\} = \bigcup_{n=n_0}^{\infty} \{\|X_n\|_\infty > \frac{1}{2}\log(n+1)\} \quad \text{occur?}$$

- **Approximate the optimal importance sampling distribution**

$$P(\tau_1 = n | \tau_1 < \infty) = \frac{P\left(\|X_n\|_\infty > \frac{1}{2}\log(n+1), \tau_1 > n-1\right)}{P(\tau_1 < \infty)}$$

$$\approx \frac{P\left(\|X_n\|_\infty > \frac{1}{2}\log(n+1)\right)}{P(\tau_1 < \infty)} \approx \frac{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)}{P(\tau_1 < \infty)}$$

Borrow from Rare Event Simulation Intuition

- Say n_0 is large: How does the rare event

$$\{\tau_1 < \infty\} = \bigcup_{n=n_0}^{\infty} \{\|X_n\|_\infty > \frac{1}{2}\log(n+1)\} \quad \text{occur?}$$

- **Approximate the optimal importance sampling distribution**

$$P(\tau_1 = n | \tau_1 < \infty) = \frac{P\left(\|X_n\|_\infty > \frac{1}{2}\log(n+1), \tau_1 > n-1\right)}{P(\tau_1 < \infty)}$$

$$\approx \quad \frac{P\left(\|X_n\|_\infty > \frac{1}{2}\log(n+1)\right)}{P(\tau_1 < \infty)} \approx \frac{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)}{P(\tau_1 < \infty)}$$

- This suggests a natural importance sampling strategy...

Will the First Record Breaking Event Occur?

1. Sample K so that for $K \geq n_0$

$$Q\left(K = n\right) = \frac{\sum_{i=1}^{d} P\left(\left|X_n\left(t_i\right)\right| > \frac{1}{2}\log\left(n+1\right)\right)}{\sum_{k=n_0}^{\infty}\sum_{i=1}^{d} P\left(\left|X_k\left(t_i\right)\right| > \frac{1}{2}\log\left(k+1\right)\right)}.$$

Will the First Record Breaking Event Occur?

1. Sample K so that for $K \geq n_0$

$$Q(K=n) = \frac{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)}{\sum_{k=n_0}^{\infty}\sum_{i=1}^{d} P\left(|X_k(t_i)| > \frac{1}{2}\log(k+1)\right)}.$$

2. Define $h(n) = P(\tau_1 = n)/Q(K=n)$ then

$$P(\tau_1 < \infty) = E^{Q}(h(K)) = \sum_{n=n_0}^{\infty} P(\tau_1 = n).$$

Will the First Record Breaking Event Occur?

1. Sample K so that for $K \geq n_0$

$$Q(K = n) = \frac{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)}{\sum_{k=n_0}^{\infty} \sum_{i=1}^{d} P\left(|X_k(t_i)| > \frac{1}{2}\log(k+1)\right)}.$$

2. Define $h(n) = P(\tau_1 = n) / Q(K = n)$ then

$$P(\tau_1 < \infty) = E^Q(h(K)) = \sum_{n=n_0}^{\infty} P(\tau_1 = n).$$

3. **If** $h(k) \leq 1$ **then** $I\{\tau_1 < \infty\}$ **is Bernoulli(**$h(K)$**) with** K **sampled under** Q**.**

Will the First Record Breaking Event Occur?

1. Choose n_0 such that

$$
\begin{aligned}
h(n) \quad = \quad & \sum_{k=n_0}^{\infty} \sum_{i=1}^{d} P\left(|X_k(t_i)| > \frac{1}{2}\log(k+1)\right) \\
& \times P\left(\|X_k\|_\infty \le \frac{1}{2}\log(1+k) \ \forall\ n_0 \le k < n\right) \\
& \times \frac{P\left(\|X_n\|_\infty > \frac{1}{2}\log(n+1)\right)}{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)} < 1.
\end{aligned}
$$

Will the First Record Breaking Event Occur?

1. Choose n_0 such that

$$
\begin{aligned}
h(n) \;=\; & \sum_{k=n_0}^{\infty} \sum_{i=1}^{d} P\left(|X_k(t_i)| > \frac{1}{2}\log(k+1)\right) \\
& \times P\left(\|X_k\|_\infty \le \frac{1}{2}\log(1+k)\ \ \forall\ n_0 \le k < n\right) \\
& \times \frac{P\left(\|X_n\|_\infty > \frac{1}{2}\log(n+1)\right)}{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)} < 1.
\end{aligned}
$$

2. **Sample a Bernoulli with parameter:**

$$
\frac{P\left(\|X_n\|_\infty > \frac{1}{2}\log(n+1)\right)}{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)}.
$$

A Related Acceptance / Rejection Question

- **Sampling from** $P\left(\cdot \mid \left\|X_n\right\|_\infty > \frac{1}{2}\log\left(n+1\right)\right)$ **by acceptance / rejection.**

A Related Acceptance / Rejection Question

- **Sampling from** $P\left(\cdot \mid \|X_n\|_\infty > \frac{1}{2}\log(n+1)\right)$ **by acceptance / rejection.**
- Proposal distribution $Q(\cdot)$ described next:

A Related Acceptance / Rejection Question

- **Sampling from** $P\left(\cdot \mid \|X_n\|_\infty > \frac{1}{2}\log(n+1)\right)$ **by acceptance / rejection.**
- Proposal distribution $Q(\cdot)$ described next:

1. Sample J with probability

$$Q(J=j) = \frac{P\left(|X_n(t_j)| > \frac{1}{2}\log(n+1)\right)}{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)}.$$

A Related Acceptance / Rejection Question

- **Sampling from** $P\left(\cdot \mid \|X_n\|_\infty > \frac{1}{2}\log(n+1)\right)$ **by acceptance / rejection.**
- Proposal distribution $Q(\cdot)$ described next:

1. Sample J with probability

$$Q(J=j) = \frac{P\left(|X_n(t_j)| > \frac{1}{2}\log(n+1)\right)}{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)}.$$

2. Given $J=j$, sample $X_n = \{X_n(t_i)\}_{i=1}^{d}$

$$\begin{aligned}
& Q\left(X_n(t_i) \in dx_i, \ldots, X_n(t_d) \in dx_d \mid J=j\right) \\
= \; & P\left(X_n(t_i) \in dx_i, \ldots, X_n(t_d) \in dx_d \mid |X_n(t_j)| > \frac{1}{2}\log(n+1)\right) \\
= \; & \frac{P\left(X_n(t_i) \in dx_i, \ldots, X_n(t_d) \in dx_d\right) \cdot I\left(|x_n(t_j)| > \frac{1}{2}\log(n+1)\right)}{P\left(|X_n(t_j)| > \frac{1}{2}\log(n+1)\right)}.
\end{aligned}$$

Computing the Likelihood Ratio

- Likelihood of X_n under Q

$$
\begin{aligned}
Q(X_n) &= \sum_j Q(X_n|J=j)\,Q(J=j) \\
&= \sum_j P(X_n)\,\frac{I\left(|x_n(t_j)| > \frac{1}{2}\log(n+1)\right)}{P\left(|X_n(t_j)| > \frac{1}{2}\log(n+1)\right)}\,Q(J=j) \\
&= P(X_n)\sum_j\left(\frac{I\left(|x_n(t_j)| > \frac{1}{2}\log(n+1)\right)}{P\left(|X_n(t_j)| > \frac{1}{2}\log(n+1)\right)}\right. \\
&\quad \left.\times\frac{P\left(|X_n(t_j)| > \frac{1}{2}\log(n+1)\right)}{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)}\right) \\
&= P(X_n)\,\frac{\sum_j I\left(|x_n(t_j)| > \frac{1}{2}\log(n+1)\right)}{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2}\log(n+1)\right)}.
\end{aligned}
$$

Acceptance / Rejection Execution

- Likelihood of X_n under Q

$$\frac{dQ}{dP}\left(X_n\right) = \frac{\sum_{j=1}^{d} I\left(\left|X_n\left(t_j\right)\right| > \log\left(n+1\right)/2\right)}{\sum_{i=1}^{d} P\left(\left|X_n\left(t_i\right)\right| > \log\left(n+1\right)/2\right)}.$$

Acceptance / Rejection Execution

- Likelihood of X_n under Q

$$\frac{dQ}{dP}\left(X_n\right)=\frac{\sum_{j=1}^{d} I\left(\left|X_n\left(t_j\right)\right|>\log\left(n+1\right)/2\right)}{\sum_{i=1}^{d} P\left(\left|X_n\left(t_i\right)\right|>\log\left(n+1\right)/2\right)}.$$

- We conclude

$$\frac{dP\left(X_n \mid \left\|X_n\right\|_\infty > \frac{\log(n+1)}{2}\right)}{dQ}$$

$$=\frac{I\left(\left\|X_n\right\|_\infty > \frac{\log(n+1)}{2}\right)}{P\left(\left\|X_n\right\|_\infty > \frac{\log(n+1)}{2}\right)}\frac{dP}{dQ}\leq\frac{\sum_{i=1}^{d} P\left(\left|X_n\left(t_i\right)\right|>\frac{\log(n+1)}{2}\right)}{P\left(\left\|X_n\right\|_\infty > \frac{\log(n+1)}{2}\right)}.$$

Acceptance / Rejection Execution

- Likelihood of X_n under Q

$$\frac{dQ}{dP}(X_n) = \frac{\sum_{j=1}^{d} I\left(|X_n(t_j)| > \log(n+1)/2\right)}{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \log(n+1)/2\right)}.$$

- We conclude

$$\frac{dP\left(X_n \mid \|X_n\|_\infty > \frac{\log(n+1)}{2}\right)}{dQ}$$
$$= \frac{I\left(\|X_n\|_\infty > \frac{\log(n+1)}{2}\right)}{P\left(\|X_n\|_\infty > \frac{\log(n+1)}{2}\right)} \frac{dP}{dQ} \leq \frac{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{\log(n+1)}{2}\right)}{P\left(\|X_n\|_\infty > \frac{\log(n+1)}{2}\right)}.$$

- By general acceptance / rejection theory we have:

$$\text{Probability of accepting} = \frac{P\left(\|X_n\|_\infty > \log(n+1)/2\right)}{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \log(n+1)/2\right)}.$$

Summary

- We sampled $B = Bernoulli\left(h\left(n\right)\right) = I\left(\tau_1 < \infty\right)$ with $h\left(n\right)$:

$$
\begin{aligned}
h(n) &= \sum_{k=n_0}^{\infty} \sum_{i=1}^{d} P\left(|X_k(t_i)| > \frac{1}{2} \log(k+1) \right) \\
&\times P\left(\|X_k\|_\infty \leq \frac{1}{2} \log(1+k) \ \forall \ n_0 \leq k < n \right) \\
&\times \frac{P\left(\|X_n\|_\infty > \frac{1}{2} \log(n+1) \right)}{\sum_{i=1}^{d} P\left(|X_n(t_i)| > \frac{1}{2} \log(n+1) \right)} < 1.
\end{aligned}
$$

Summary

- We sampled $B = Bernoulli\left(h\left(n\right)\right) = I\left(\tau_1 < \infty\right)$ with $h\left(n\right)$:

$$
\begin{aligned}
h\left(n\right) &= \sum_{k=n_0}^{\infty}\sum_{i=1}^{d} P\left(\left|X_k\left(t_i\right)\right| > \frac{1}{2}\log\left(k+1\right)\right) \\
&\times P\left(\left\|X_k\right\|_\infty \leq \frac{1}{2}\log\left(1+k\right)\ \forall\ n_0 \leq k < n\right) \\
&\times \frac{P\left(\left\|X_n\right\|_\infty > \frac{1}{2}\log\left(n+1\right)\right)}{\sum_{i=1}^{d} P\left(\left|X_n\left(t_i\right)\right| > \frac{1}{2}\log\left(n+1\right)\right)} < 1.
\end{aligned}
$$

- Also sampled X_n given $\left\|X_n\right\|_\infty > \frac{1}{2}\log\left(n+1\right)$ by acceptance rejection:

$$
\text{Probability of accepting} = \frac{P\left(\left\|X_n\right\|_\infty > \log\left(n+1\right)/2\right)}{\sum_{i=1}^{d} P\left(\left|X_n\left(t_i\right)\right| > \log\left(n+1\right)/2\right)}.
$$

Exact Sampling Property

- **Also we actually obtain:**

$$\mathcal{L}aw_Q\left(X_1, ..., X_K, K \middle| B = 1\right) = \mathcal{L}aw_P\left(X_1, ..., X_{\tau_1}, \tau_1 \middle| \tau_1 < \infty\right).$$

Exact Sampling Property

- **Also we actually obtain:**

$$\mathcal{L}aw_Q\left(X_1, ..., X_K, K | B = 1\right) = \mathcal{L}aw_P\left(X_1, ..., X_{\tau_1}, \tau_1 | \tau_1 < \infty\right).$$

- Let's check this identity:

$$\begin{aligned}
& Q(X_1 \in dx_1, ..., X_n \in dx, K = n, B = 1) \\
= \; & Q\left(K = n\right) \sum_{k=n_0}^{\infty} \sum_{i=1}^{d} P\left(\left|X_k\left(t_i\right)\right| > \frac{1}{2}\log\left(k+1\right)\right) \\
& \times P\left(X_{n_0} \in dx, ..., X_{n-1} \in dx_{n-1} \mid \tau_1 > n-1\right) \\
& \times P\left(X_n \in dx | \left\|X_n\right\|_\infty > \frac{\log\left(n+1\right)}{2}\right) \\
= \; & P\left(\tau_1 = n\right) P\left(X_1 \in dx_1, ..., X_n \in dx \mid \tau_1 = n\right) \\
= \; & P\left(X_1 \in dx_1, ..., X_n \in dx, \; \tau_1 = n, \; \tau_1 < \infty\right).
\end{aligned}$$

How to Continue? What Drives the Complexity?

- Subsequently sample $\tau_2, ..., \tau_{L-1} = N_X$ until $\tau_L = \infty$.

How to Continue? What Drives the Complexity?

- Subsequently sample $\tau_2, ..., \tau_{L-1} = N_X$ until $\tau_L = \infty$.

- Note that X_n can be easily simulated even if $n > N_X$.

$$\{X_n \quad : \quad n \geq N_X + 1\} \text{ are still independent but conditioned}$$
$$\text{on } \left\|X_n\right\|_\infty \leq \log\left(n+1\right)/2) \text{ for } n \geq N_X + 1$$

How to Continue? What Drives the Complexity?

- Subsequently sample $\tau_2, ..., \tau_{L-1} = N_X$ until $\tau_L = \infty$.
- Note that X_n can be easily simulated even if $n > N_X$.

$$\{X_n \ : \ n \geq N_X + 1\} \text{ are still independent but conditioned}$$
$$\text{on } \|X_n\|_\infty \leq \log(n+1)/2) \ \text{ for } n \geq N_X + 1$$

- Total complexity turns out to be driven by n_0 such that

$$\sum_{k=n_0}^{\infty} \sum_{i=1}^{d} P\left(|X_k(t_i)| > \frac{1}{2}\log(k+1)\right) < 1$$
$$\Longrightarrow \quad n_0 = \exp\left(c\sqrt{\log(d)}\right) = O\left(d^{\varepsilon}\right) \text{ for any } \varepsilon > 0.$$

- A similar technique used to sample *the last time* N such that

$$A_n < \frac{1}{2}n.$$

Milestone / Record Breaking for Random Walks

- A similar technique used to sample *the last time* N such that

$$A_n < \frac{1}{2}n.$$

- In this case it suffices to sample A_n jointly with

$$\max_{m \geq n}\{A_m - m/2\}$$

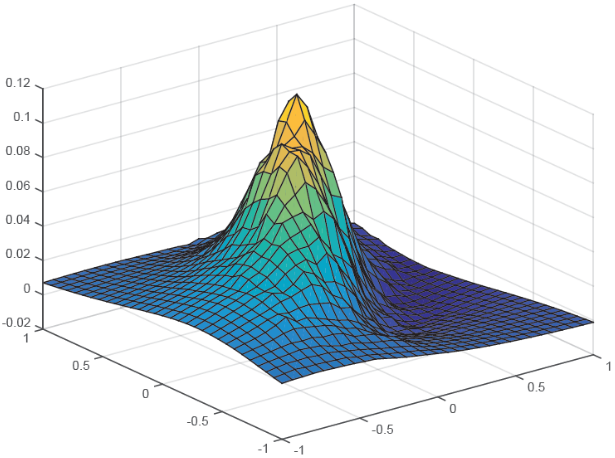
rare event simulation techniques can be applied (see B. & Wallwater (2014)).

Density Estimation

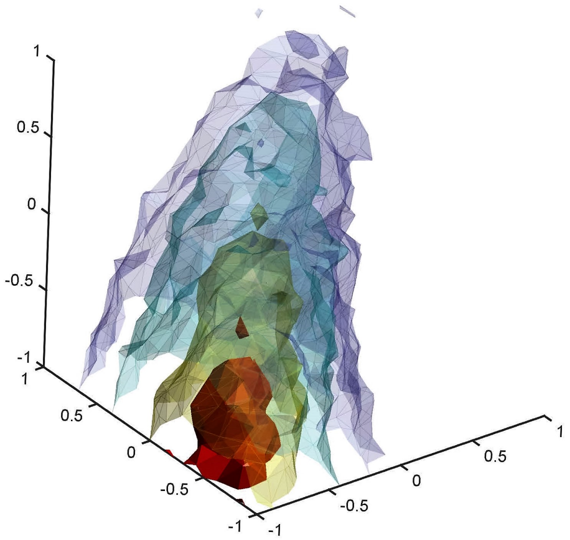
Want to design a finite variance estimator $W(y_1, ..., y_d)$ so that

$$E(W(y_1, ..., y_d)) = f_M(y_1, ..., y_d).$$

An Illustration of the Density Pictures



Some Pictures: 3-d Density



Some Malliavin Calculus

- Follow an idea from Malliavin & Thalmaier.

Some Malliavin Calculus

- Follow an idea from Malliavin & Thalmaier.
- Consider Poisson's equation
$$\Delta u = f$$
formally if $G(\cdot)$ is the Poisson kernel we have
$$u(x) = \int f(y) G(x-y) dy.$$

Some Malliavin Calculus

- Follow an idea from Malliavin & Thalmaier.
- Consider Poisson's equation

$$\Delta u = f$$

formally if $G(\cdot)$ is the Poisson kernel we have

$$u(x) = \int f(y) G(x-y) dy.$$

- So, formally

$$\begin{aligned} \Delta u(x) &= \int f(y) \Delta G(x-y) dy = f(x) \\ \Delta G(x-y) dy &= \delta(x-y) dy \\ f_M(x_1, ..., x_d) &= E(\Delta G(x-M)). \end{aligned}$$

Applying Malliavin Thalmaier Formula

- For $d \geq 3$ we have that

$$\Delta G(x-y) = \sum_{i=1}^{d} \frac{\partial^2 G(x)}{\partial x_i^2}(x-y),$$

$$G_i(x) = \frac{\partial G(x)}{\partial x_i} = \kappa_d \frac{x_i}{\|x\|_2^d}, \quad G(x) = \kappa_d' \frac{1}{\|x\|_2^{d-2}}$$

Applying Malliavin Thalmaier Formula

- For $d \geq 3$ we have that

$$
\begin{aligned}
\Delta G(x-y) &= \sum_{i=1}^{d} \frac{\partial^2 G(x)}{\partial x_i^2}(x-y), \\
G_i(x) &= \frac{\partial G(x)}{\partial x_i} = \kappa_d \frac{x_i}{\|x\|_2^d}, \quad G(x) = \kappa_d' \frac{1}{\|x\|_2^{d-2}}
\end{aligned}
$$

- Higher order derivatives of G implies higher variability...

Applying Malliavin Thalmaier Formula

- For $d \geq 3$ we have that

$$\Delta G(x-y) = \sum_{i=1}^{d} \frac{\partial^2 G(x)}{\partial x_i^2}(x-y),$$
$$G_i(x) = \frac{\partial G(x)}{\partial x_i} = \kappa_d \frac{x_i}{\|x\|_2^d}, \quad G(x) = \kappa'_d \frac{1}{\|x\|_2^{d-2}}$$

- Higher order derivatives of G implies higher variability...

- **Two ways to fix this: Integration by parts & Randomized Multilevel Monte Carlo.**

Improving Variance: Integration by Parts for Finite Maxima

- We claim

$$\frac{\partial G_i\left(x-M_n\right)}{\partial x_i}=-\frac{\partial G_i\left(x-M_n\right)}{\partial M_n\left(t_i\right)}=-\sum_{k=1}^{n}\frac{\partial G_i\left(x-M_n\right)}{\partial X_k\left(t_i\right)}.$$

Improving Variance: Integration by Parts for Finite Maxima

- We claim

$$\frac{\partial G_i\left(x-M_n\right)}{\partial x_i}=-\frac{\partial G_i\left(x-M_n\right)}{\partial M_n\left(t_i\right)}=-\sum_{k=1}^{n}\frac{\partial G_i\left(x-M_n\right)}{\partial X_k\left(t_i\right)}.$$

- Let's check second equality:

$$\begin{aligned}
M_n\left(t_i\right) &= \max_{1\leq k\leq n}\left\{X_k\left(t_i\right)-a_k\right\}. \\
\sum_{k=1}^{n}\frac{\partial M_n\left(t_i\right)}{\partial X_k\left(t_i\right)} &= \sum_{k=1}^{n} I\left(M_n\left(t_i\right)=X_k\left(t_i\right)-a_k\right)=1. \\
\frac{\partial G_i\left(x-M_n\right)}{\partial X_k\left(t_i\right)} &= \frac{\partial G_i\left(x-M_n\right)}{\partial M_n\left(t_i\right)}\frac{\partial M_n\left(t_i\right)}{\partial X_k\left(t_i\right)} \\
\sum_{k=1}^{n}\frac{\partial G_i\left(x-M_n\right)}{\partial X_k\left(t_i\right)} &= \frac{\partial G_i\left(x-M_n\right)}{\partial M_n\left(t_i\right)}.
\end{aligned}$$

Improving Variance: Integration by Parts for Finite Maxima

- Integration by parts yields (with $\Sigma = cov(X_k)$)

$$E\left(\frac{\partial G_i(x - M_n)}{\partial X_k(t_i)}\right) = -E\left(G_i(x - M_n) \cdot e_i^T \Sigma^{-1} X_k\right)$$

Improving Variance: Integration by Parts for Finite Maxima

- Integration by parts yields (with $\Sigma = cov\left(X_k\right)$)

$$E\left(\frac{\partial G_i\left(x-M_n\right)}{\partial X_k\left(t_i\right)}\right) = -E\left(G_i\left(x-M_n\right)\cdot e_i^T\Sigma^{-1}X_k\right)$$

- Therefore

$$\begin{aligned} E\left(\frac{\partial G_i\left(x-M_n\right)}{\partial x_i}\right) &= -\sum_{k=1}^{n} E\left(\frac{\partial G_i\left(x-M_n\right)}{\partial X_k\left(t_i\right)}\right) \\ &= E\left(G_i\left(x-M_n\right)\cdot e_i^T\Sigma^{-1}\sum_{k=1}^{n} X_k\right) \\ f_{M_n}\left(x_1,\ldots,x_d\right) &= E\left(\sum_{k=1}^{n}\sum_{i=1}^{d} G_i\left(x-M_n\right) e_i^T\Sigma^{-1}X_k\right). \end{aligned}$$

Improving Variance: Infinite Horizon Maxima

- Let $\mathcal{F}_N$ be the information generated by the exact sampling procedure so that
$$M_N = M,$$
then we have that for $m \geq 1$
$$E\left(X_{N+m} | \mathcal{F}_N\right) = 0,$$
because given $n > N$, we have that $X_n\left(t_i\right)$ has a symmetric distribution.

Improving Variance: Infinite Horizon Maxima

- Let $\mathcal{F}_N$ be the information generated by the exact sampling procedure so that

$$M_N = M,$$

then we have that for $m \geq 1$

$$E\left(X_{N+m} | \mathcal{F}_N\right) = 0,$$

because given $n > N$, we have that $X_n\left(t_i\right)$ has a symmetric distribution.

- Thus

$$\begin{aligned} \lim_{n\to\infty} f_{M_n}\left(x_1, ..., x_d\right) &= f_{M_n}\left(x_1, ..., x_d\right) \\ &= E\left(\sum_{k=1}^{N}\sum_{i=1}^{d} G_i\left(x - M\right) e_i^T \Sigma^{-1} X_k\right). \end{aligned}$$

Continue Integrating by Parts

- The estimator

$$W\left(x-M\right)=\sum_{k=1}^{N}\sum_{i=1}^{d}G_{i}\left(x-M\right)e_{i}^{T}\Sigma^{-1}X_{k}.$$

Continue Integrating by Parts

- The estimator

$$W\left(x-M\right)=\sum_{k=1}^{N}\sum_{i=1}^{d}G_i\left(x-M\right)e_i^T\Sigma^{-1}X_k.$$

- One can continue obtaining

$$\begin{aligned}\text{quadratic weight} &\Rightarrow \text{finite variance } d\leq 3\\ \text{Cubic weight} &\Rightarrow \text{finite variance } d\leq 5\\ &\quad\ldots\end{aligned}$$

Applying Multilevel Monte Carlo

- Out starting point is estimator

$$W(x) = \sum_{k=1}^{N} \sum_{i=1}^{d} G_i(x - M)\, e_i^T \Sigma^{-1} X_k.$$

Applying Multilevel Monte Carlo

- Out starting point is estimator

$$W(x) = \sum_{k=1}^{N} \sum_{i=1}^{d} G_i(x - M)\, e_i^T \Sigma^{-1} X_k.$$

- Randomized Multilevel Monte Carlo, introducing the differences

$$\begin{aligned} \Delta_n &= G_i^{n+1}(x - M) - G_i^n(x - M); \\ G_i^n(x - M) &= \kappa_d \frac{x_i - M(t_i)}{\|x - M\|_2^d + \|x - M\|_2 / \log(n+1)}. \end{aligned}$$

Conclusions

- Max-stable fields are natural models for extremes, but present computational challenges.

Conclusions

- Max-stable fields are natural models for extremes, but present computational challenges.
- First asymptotically optimal exact simulation algorithms for max-stable fields.

Conclusions

- Max-stable fields are natural models for extremes, but present computational challenges.
- First asymptotically optimal exact simulation algorithms for max-stable fields.
- Ideas based on record-breaking events & rare-event simulation.

Conclusions

- Max-stable fields are natural models for extremes, but present computational challenges.
- First asymptotically optimal exact simulation algorithms for max-stable fields.
- Ideas based on record-breaking events & rare-event simulation.
- Malliavin calculus ideas for first efficient density estimators for max-stable fields.