Non-Parametric Simulation of Multivariate Extreme Events via Spectral Bootstrap

joint with Juliette Legrand (UBO) and Nisrine Madhar (Natixis)

Maud Thomas · ISFA, Université Lyon 1

6th European
Congress of Actuaries
www.eca2026.org

Elements of Extreme Value Theory

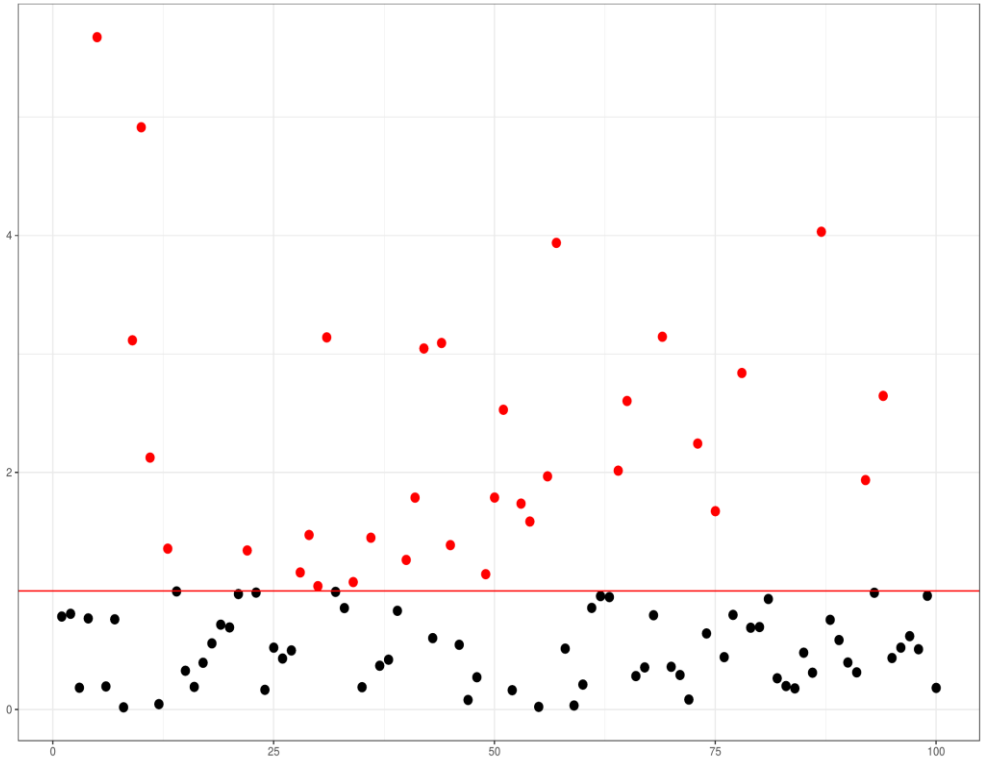
Goals of Extreme Value Theory

Goals of Extreme Value Theory

1. Estimate the probability of occurrence of an event more severe than previously observed
2. Estimate an extreme quantile

⇒ **Inference outside the sample support**

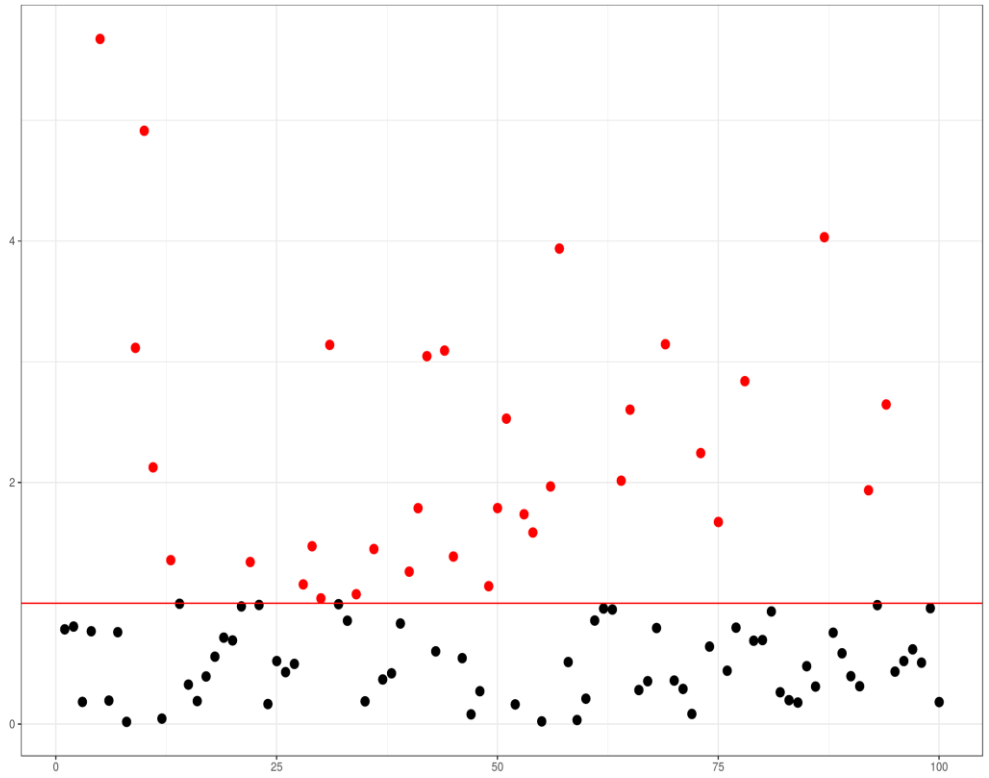
Univariate Peaks-over-Threshold Method



- $Y_1, Y_2, \ldots$ a series of i.i.d. random variables
- Fix a (high) threshold u
- Extreme event = Y_i exceeds u

  → Given that $Y_i > u$, an excess is defined by $Z_i = Y_i - u$

Univariate Peaks-over-Threshold Method



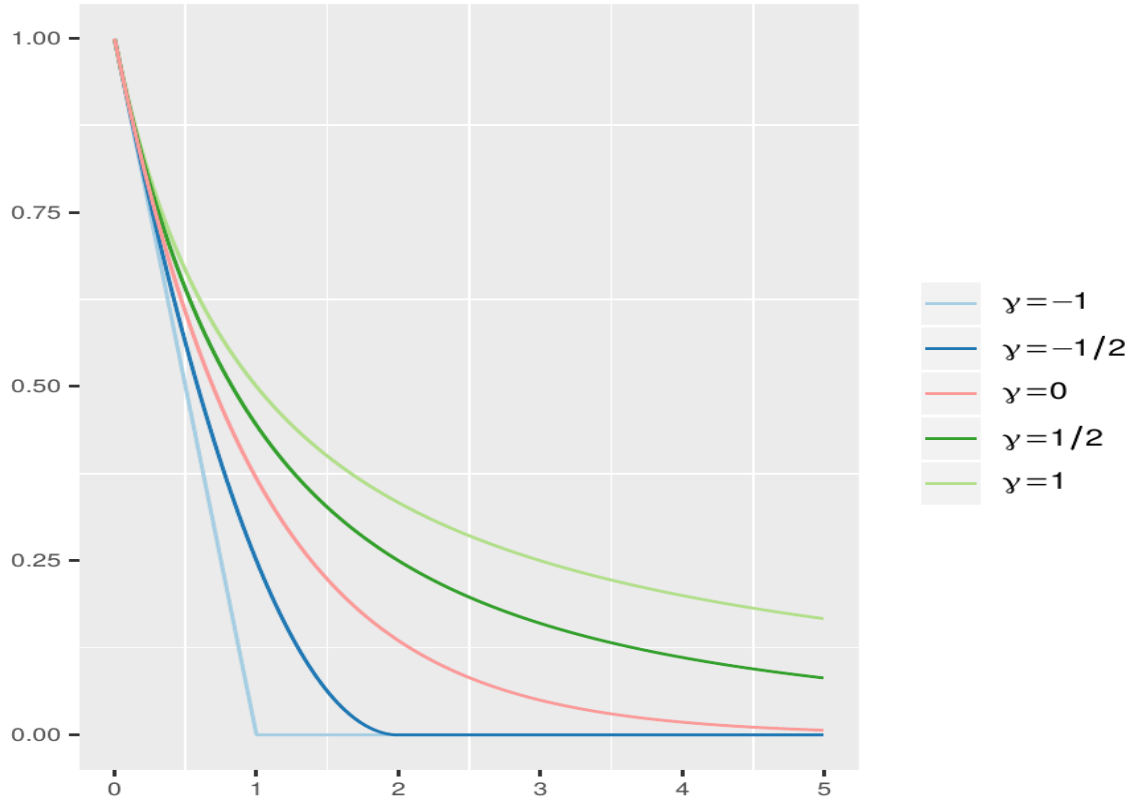
- Excess distribution : $F_u(z) = P[Y_1 - u \leq z \mid Y_1 > u]$
- **Balkema & de Haan (1974), Pickands (1975)**

Under certain conditions, the distribution of excesses F_u converges as u → ∞ to a Generalized Pareto Distribution:

$$H_{\sigma,\gamma}(z) = 1 - (1 + \gamma z/\sigma)^{-1/\gamma} \quad \text{if } \gamma \neq 0$$
$$H_{\sigma,\gamma}(z) = 1 - \exp(-z/\sigma) \quad \text{if } \gamma = 0$$

→ Generalized Pareto distributions (GPD) as limit parametric family

Generalized Pareto Distributions



GPD survival functions

Three Domains of Attraction

1. **Fréchet domain (γ > 0) — Heavy-tailed distributions**

   $1 - H_\gamma(z) \sim \gamma^{-1/\gamma} z^{-1/\gamma}$

   Examples: Cauchy, Log-gamma, Student

2. **Gumbel domain (γ = 0) — Thin tail distributions**

   $1 - H_0(z) \sim \exp(-z)$

   Examples: Gaussian, Gamma, Exponential

3. **Weibull domain (γ < 0) — Finite tail distributions**

   $1 - H_\gamma(z) = 0 \quad \text{for } x \geq -1/\gamma$

   Examples: Uniform, Beta

Non-Parametric Simulation of Multivariate Extremes via Spectral Bootstrap

Motivation and Goal

- EVT focuses on behaviour in the tail of multivariate distributions
- Few extreme observations → difficult inference
  → Natural idea = bootstrap to increase the number of data
- **BUT, standard bootstrap fails for EVT**

Goal

Non-parametric simulation to expand observations above the extreme level
→ Extension of Legrand et al. [2023] (dim 2) to larger dimensions d

- Expands the number of observations above extreme level
- Ensures more reliable estimations
- Enables extrapolation beyond observed data
  → uses spectral representation

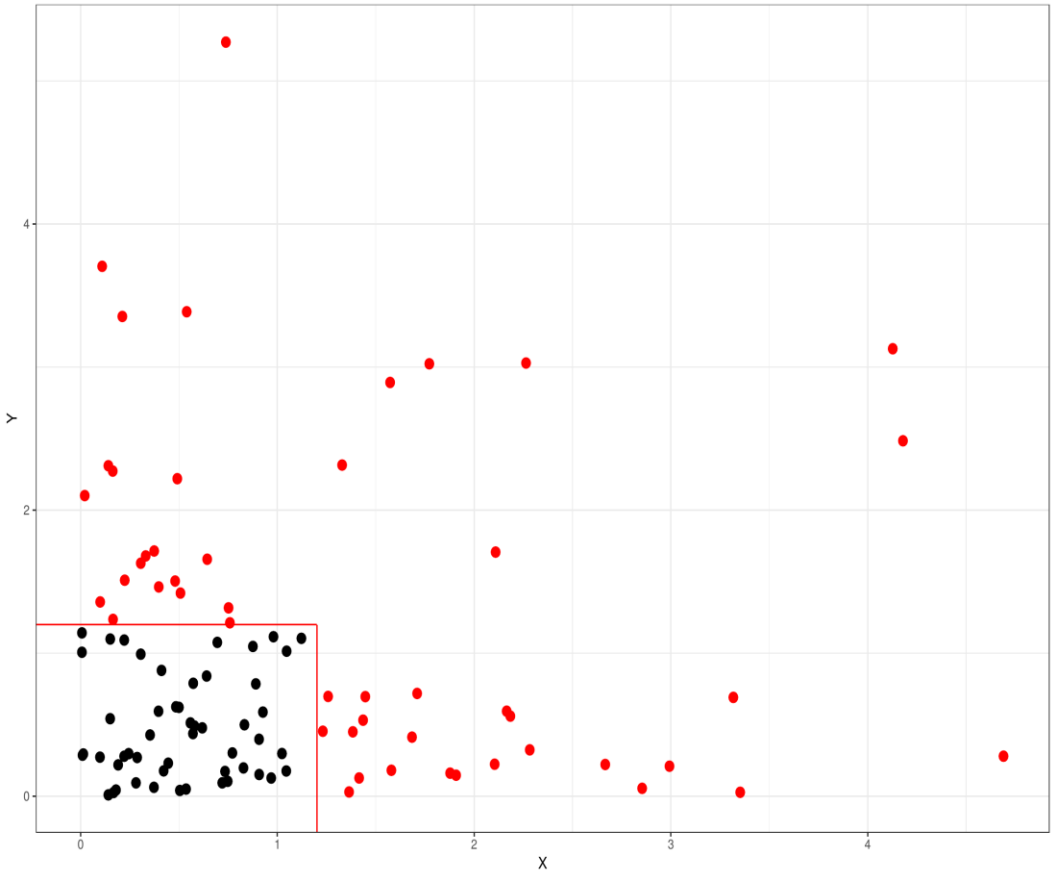
Multivariate Generalized Pareto Distributions

- $\mathbf{X} = (X_1, \ldots, X_d)$ observations
- Choose (high) thresholds $\mathbf{u} = (u_1, \ldots, u_d)$
- Extreme event = AT LEAST one X_j exceeds u_j

⚠ **NO parametric family of limit distributions**

- $\mathbf{Z} = \mathbf{X} - \mathbf{u} \mid \mathbf{X} \not\leq \underline{u}$ follows a MGPD (asymptotically)
- Standard MGPD: $\boldsymbol{\sigma} = \mathbf{1}$, $\boldsymbol{\gamma} = \mathbf{0}$

  → Exponential marginals
- *Rootzén et al. proposed explicit densities for specific models*



Multivariate Generalized Pareto Vectors

The vector of excesses:

$\boldsymbol{Z} = \boldsymbol{X} - \boldsymbol{u} \mid \boldsymbol{X} \not\preceq \boldsymbol{u}$ *(at least one component positive)*

Rootzén et al. [2018] — Spectral decomposition:

$$\mathbf{Z} = E + \mathbf{T} - \max(\mathbf{T})$$

- E ~ Exp(1) — unit exponential variable
- **T** d-dimensional random vector
- **T** and E are independent

Non-Parametric Joint MGP Simulation — Dimension 2

Legrand et al. [2023]

$$Z_1 = E + T_1 - \max(T_1, T_2)$$
$$Z_2 = E + T_2 - \max(T_1, T_2)$$

Noting $\Delta = Z_1 - Z_2 = T_1 - T_2$:

$$Z_1 = E + \Delta \cdot 1\{\Delta<0\}$$
$$Z_2 = E - \Delta \cdot 1\{\Delta\geq 0\}$$

- Simulate E and Δ independently — simulating E is trivial
- Difficulty: simulate Δ

⇒ **Bootstrapping on Δ**

Spectral Representation of MGP in Dimension d

Key Theorem — Rootzén et al. [2018]

A standard MGP vector **Z** admits the representation

$$\mathbf{Z} = E + \mathbf{T} - \max_k T_k$$

where $E \sim \text{Exp}(1)$ and **T** is a d-dim random vector.

- Define spectral differences: $\Delta_j = Z_j - \max_k Z_k$
- Then $\mathbf{Z} = E + \mathbf{\Delta}$
- **Spectral Bootstrap: bootstrap on observed Δ values**

Algorithm 1: Multivariate Extreme Spectral Bootstrap

Algorithm

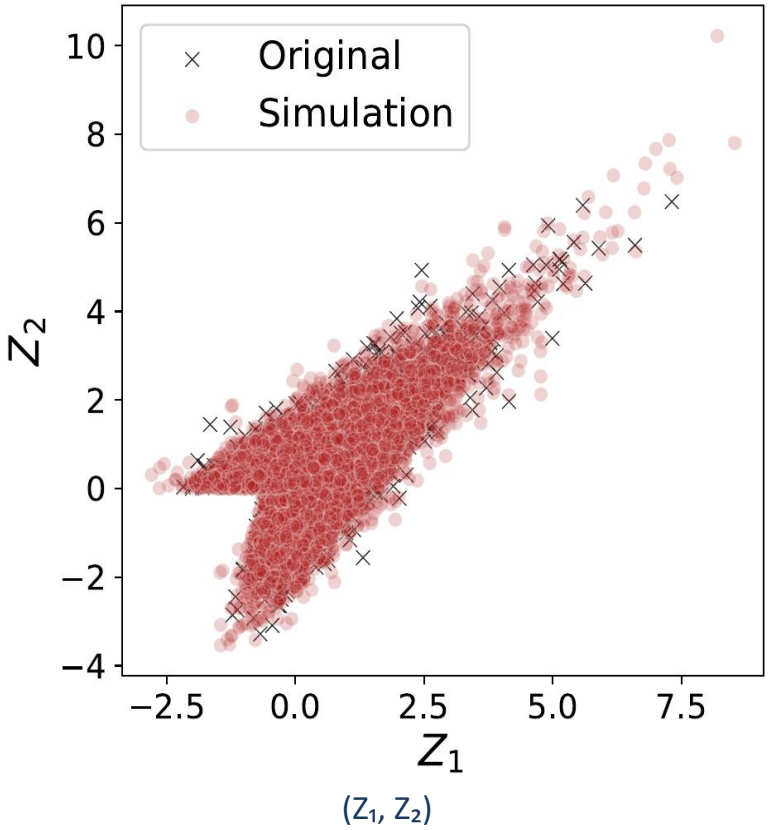
Input: n i.i.d. observations $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ of a standard MGP distribution in $\mathbb{R}^d$
Output: m standard MGP simulated samples $\mathbf{Z}_1^*, \ldots, \mathbf{Z}_m^*$ in $\mathbb{R}^d$

1. Compute $\Delta_{i,j} \leftarrow Z_{i,j} - \max_\{1 \le k \le d\} Z_{i,k}$ for $1 \le i \le n$, $1 \le j \le d$
2. Generate $E_1, \ldots, E_m \sim$ i.i.d. Exp(1), independently of $\mathbf{\Delta}_1, \ldots, \mathbf{\Delta}_n$
3. Generate m bootstrap samples $\mathbf{\Delta}_1^*, \ldots, \mathbf{\Delta}_m^*$ from $(\mathbf{\Delta}_1, \ldots, \mathbf{\Delta}_n)$
4. Return $\mathbf{Z}_\ell^* \leftarrow E_\ell + \mathbf{\Delta}_\ell^*$ for $1 \le \ell \le m$

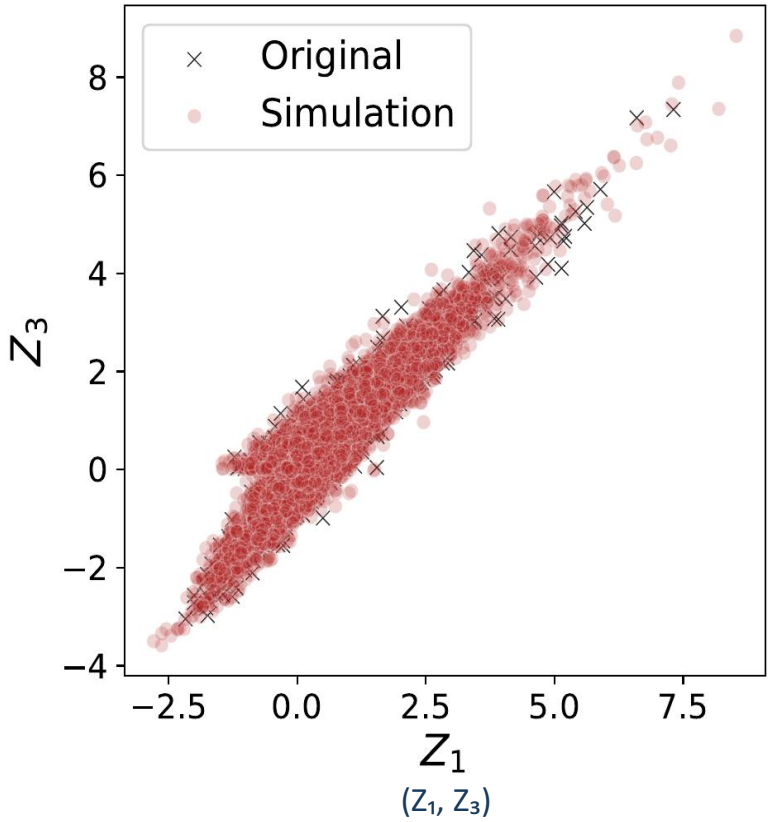
- Preserves tail dependence structure
- Theoretical guarantees: simulated samples follow same distribution

Non-Parametric Joint MGP Simulation — Scatter Plots

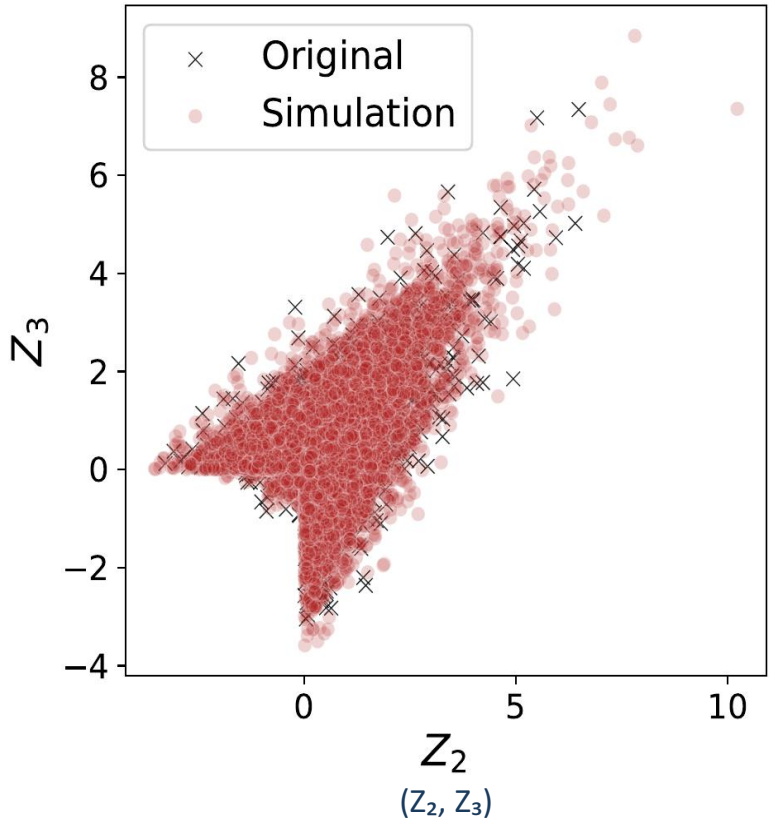
Bivariate representations of a simulated sample (n = 2,000) of a MGP vector $\mathbf{Z} \in \mathbb{R}^3$



(Z_1, Z_2)



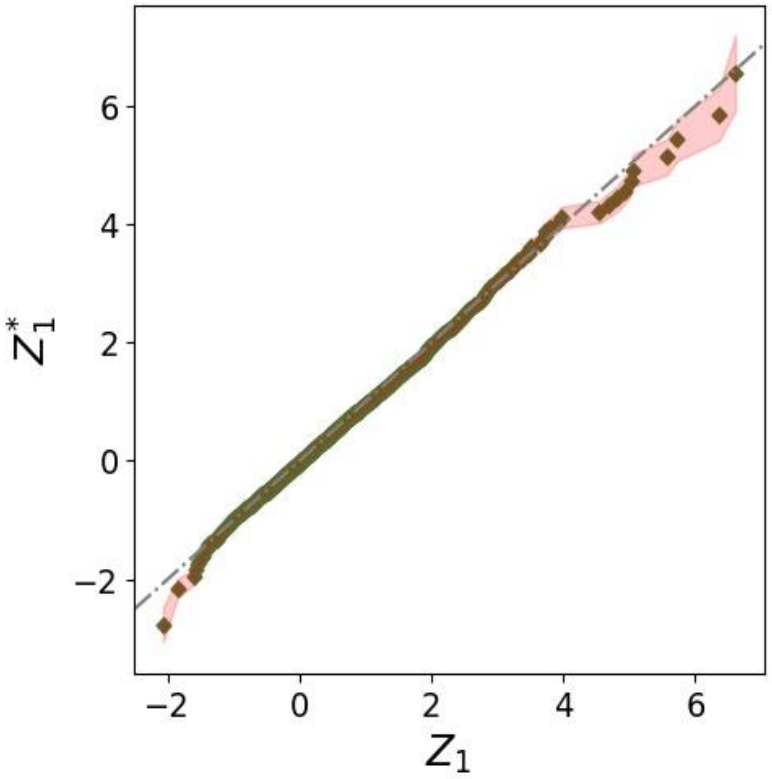
(Z_1, Z_3)



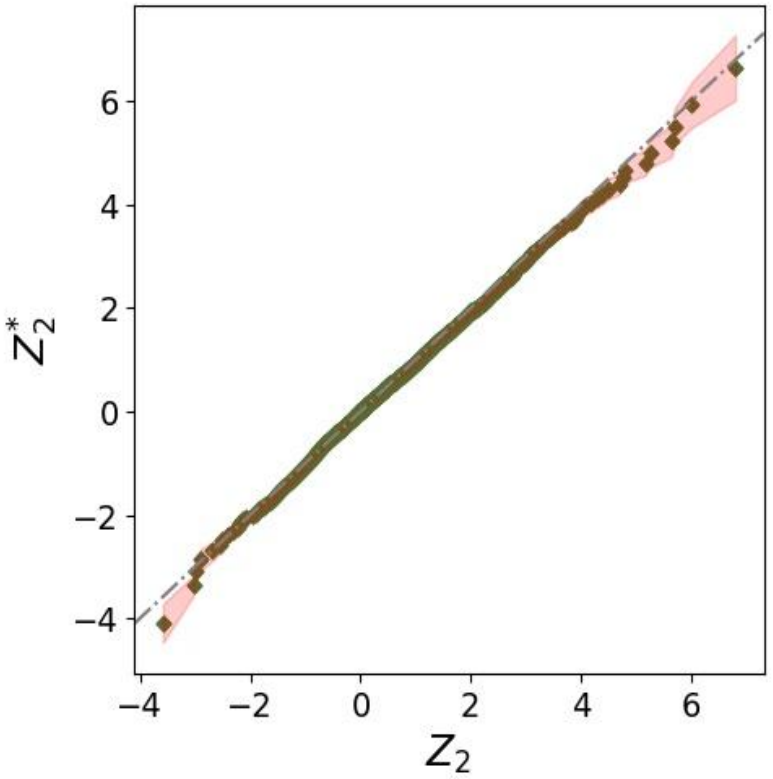
(Z_2, Z_3)

Non-Parametric Joint MGP Simulation — QQ Plots

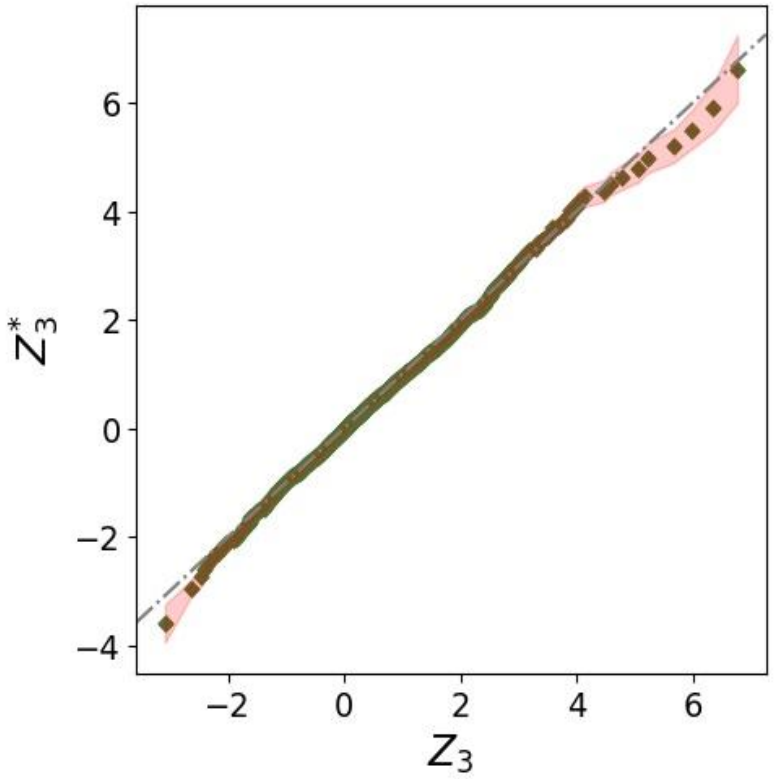
*QQ plots: simulated (**Z***) vs original (**Z**) marginals — each component aligns on the diagonal*



Z_1* vs Z_1



Z_2* vs Z_2



Z_3* vs Z_3

Illustration on Risk Management

Illustration on Risk Management

- Risk management = Crucial in various sectors, specific to each sector
- Risk factor = any variable that could result in a loss or damage

  Climatology: drought, floods

  → wave heights, wind gusts, precipitation

  Finance: market movements

  → market parameters, interest rates

- **Tail risk: very severe magnitude, very low probability**

Aim

For a given target risk factor X_j, accurately quantify its risk through the estimation of Tail Risk Metrics (TRM)

Tail Risk Metrics (TRM)

$\mathbf{X} = (X_1, \ldots, X_d) \in \mathbb{R}^d$ — 3 TRMs:

Expected Shortfall (ES) [Artzner et al., 1999]

$$ES_\alpha(X_j) = E[X_j \mid X_j > VaR_\alpha(X_j)] = v_j^\alpha$$

Multivariate Marginal ES (MMES)

$$MMES_\alpha(X_j; X) = E[X_j \mid \mathbf{X}_{-j} \geq \mathbf{v}_{-j}^\alpha]$$

Dependent Conditional Tail Expectation (DCTE)

$$DCTE\alpha(X_j; \mathbf{X}) = E[X_j \mid \mathbf{X} \geq \mathbf{v}^\alpha]$$

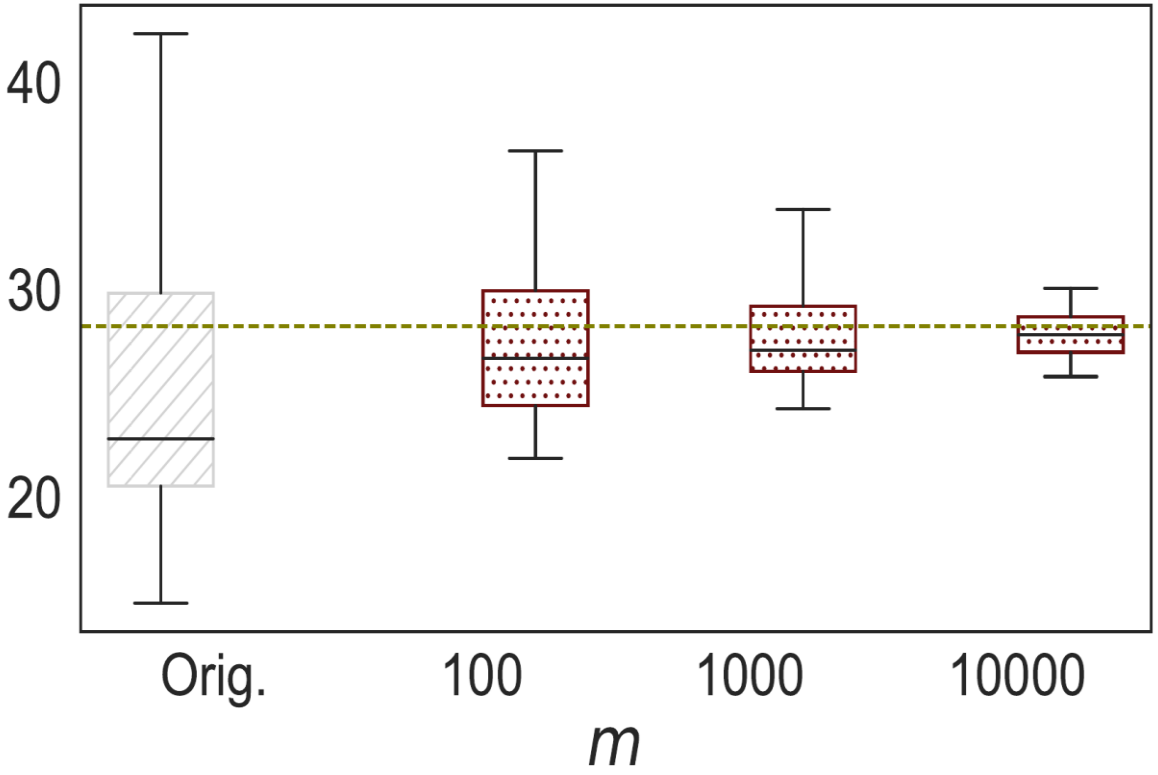
Simulation Framework

- $\mathbf{X} = (X_1, X_2, X_3)$ with Student t-marginals: $\nu_1 = 2$, $\nu_2 = 3$, $\nu_3 = 2.5$
- Components of X are asymptotically dependent
- Gumbel copula to ensure dependent extremes in the upper tail

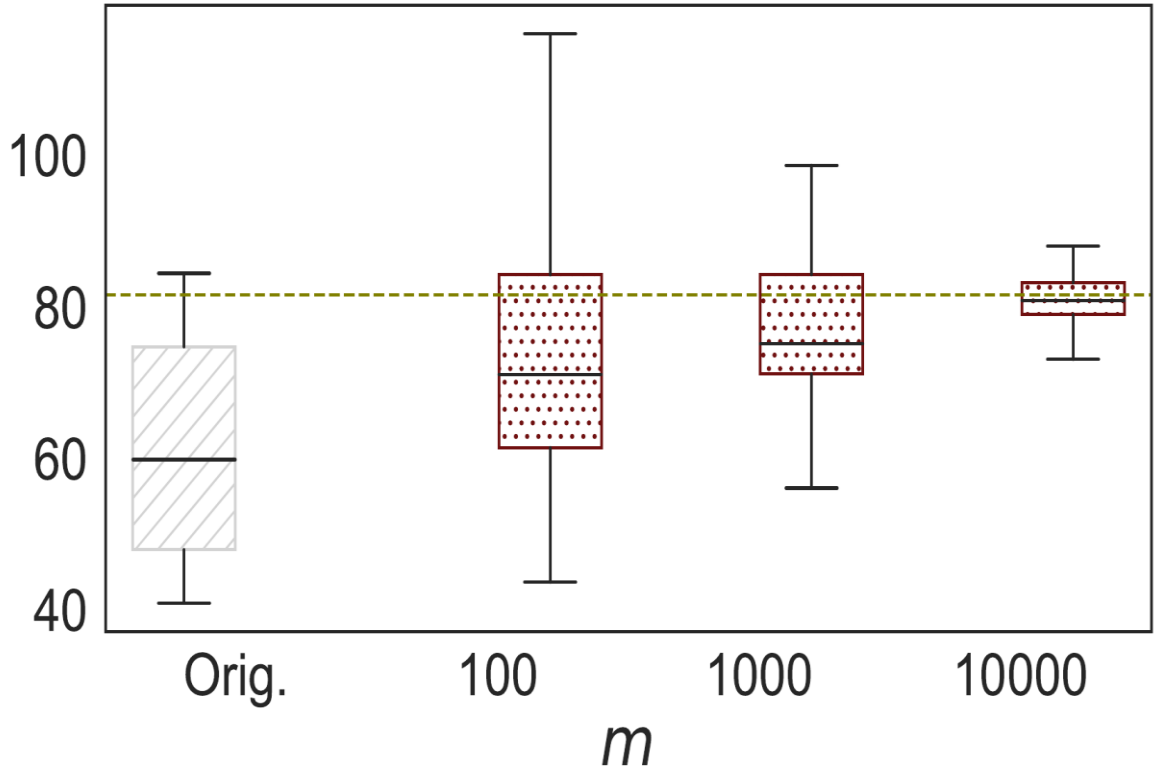
- Numerical experiments on simulated datasets $D \in \mathbb{R}^{1500 \times 3}$

TRM Estimation — Expected Shortfall

Grey hatched = 100 original samples · Red dotted = 100 extended samples · Dashed green = theoretical value



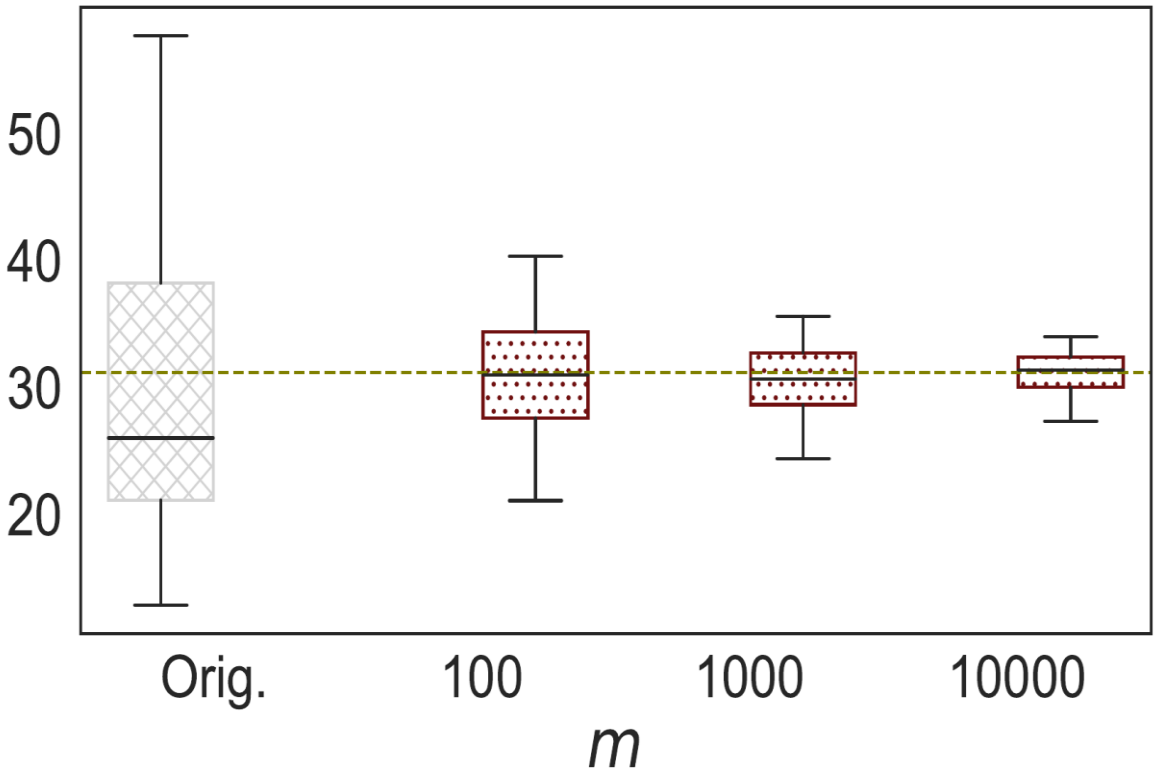
$ES_{0.0025}$



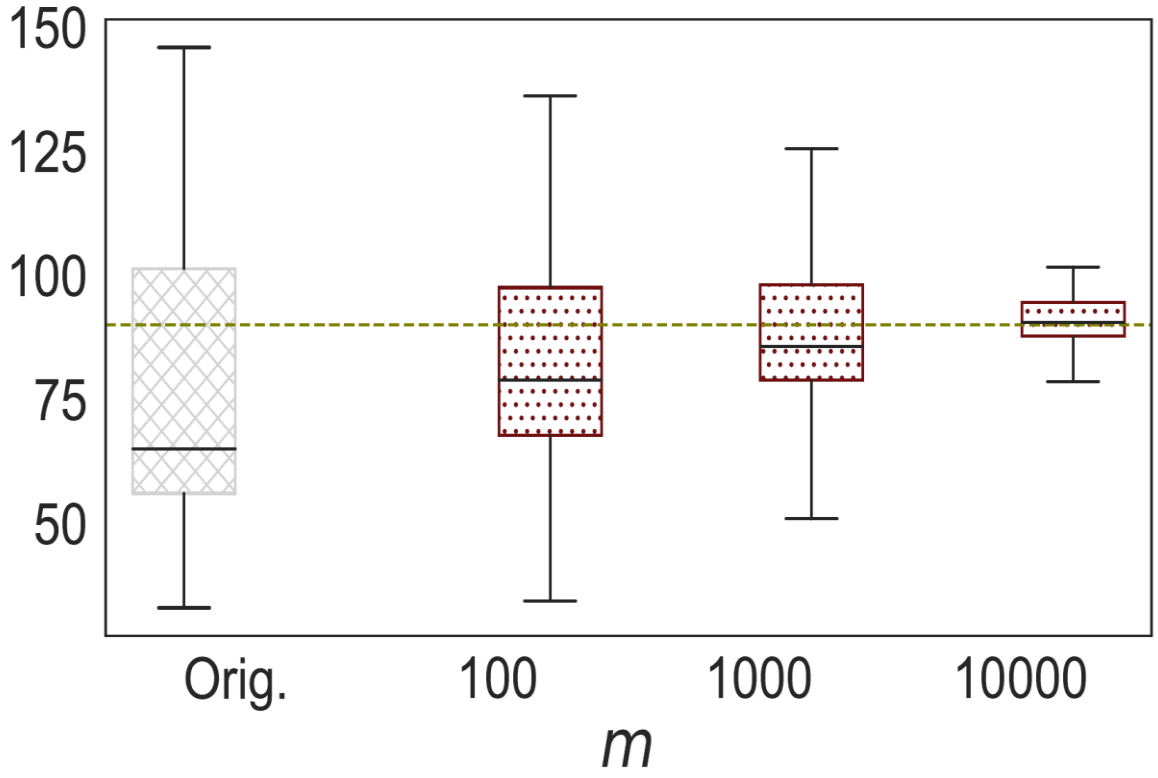
$ES_{0.0003}$

TRM Estimation — Multivariate Marginal ES

Grey hatched = 100 original samples · Red dotted = 100 extended samples · Dashed green = theoretical value



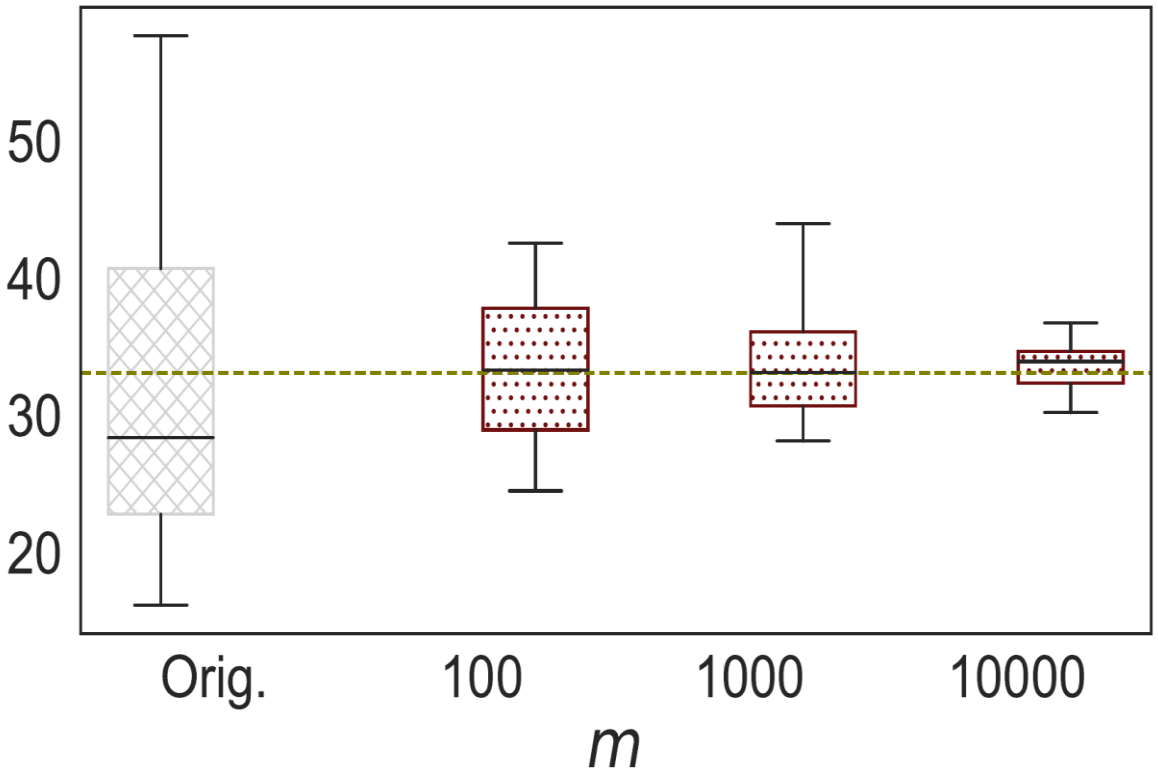
$\mathrm{MMES}_{0.0025}$



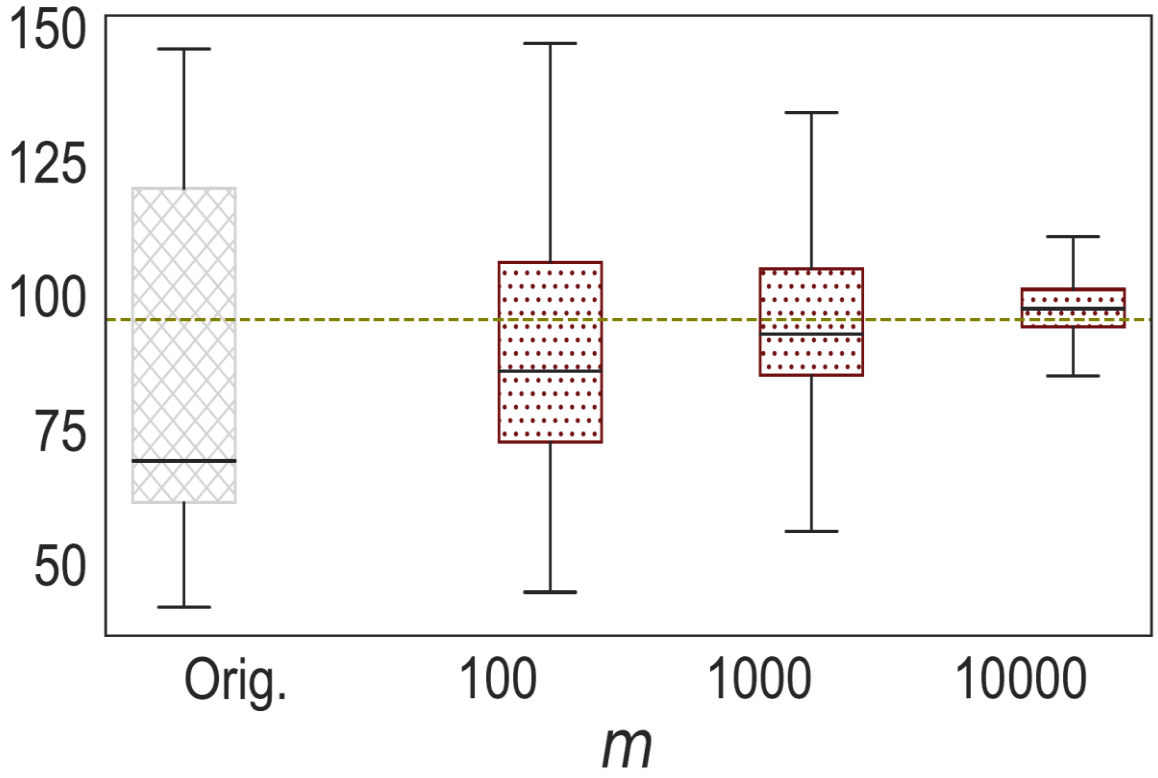
$\mathrm{MMES}_{0.0003}$

TRM Estimation — Dependent Conditional Tail Expectation

Grey hatched = 100 original samples · Red dotted = 100 extended samples · Dashed green = theoretical value



$DCTE_{0.0025}$



$DCTE_{0.0003}$

Thank you for your attention!