Piecewise-linear Modeling of Multivariate Geometric Extremes

Ryan Campbell Jennifer Wadsworth

Lancaster University

26 June 2025

14th International Conference on Extreme Value Analysis
UNC–Chapel Hill

Limit sets and gauge functions

- $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} f_{\boldsymbol{X}}$, d-dimensional, Exponential(1) margins...

Limit sets and gauge functions

- $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} f_{\boldsymbol{X}}$, d-dimensional, Exponential(1) margins...
- *If* $f_{\boldsymbol{X}}$ satisfies

$$\frac{-\log f_{\boldsymbol{X}}(t\boldsymbol{x})}{t} \longrightarrow g(\boldsymbol{x}) \quad , \text{ as } t \to \infty,$$

where the **gauge function**, $g : \mathbb{R}^d \longrightarrow \mathbb{R}_+$, is 1-homogeneous and continuous...

Limit sets and gauge functions

- $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} f_{\boldsymbol{X}}$, d-dimensional, Exponential(1) margins...
- *If* $f_{\boldsymbol{X}}$ satisfies

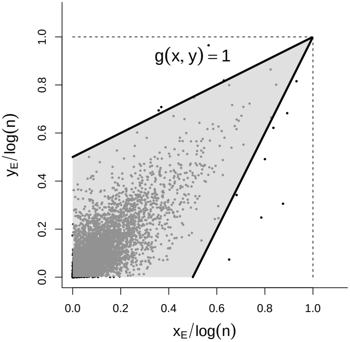
$$\frac{-\log f_{\boldsymbol{X}}(t\boldsymbol{x})}{t} \longrightarrow g(\boldsymbol{x}) \quad , \text{ as } t \to \infty,$$

  where the **gauge function**, $g : \mathbb{R}^d \longrightarrow \mathbb{R}_+$, is 1-homogeneous and continuous...
- ...*then* scaled sample clouds $\left\{\frac{\boldsymbol{X}_1}{\log n}, \ldots, \frac{\boldsymbol{X}_n}{\log n}\right\}$ converge onto a **limit set**,

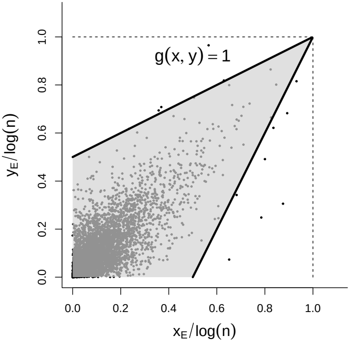
$$G := \left\{\boldsymbol{x} \in \mathbb{R}^d \,\middle|\, g(\boldsymbol{x}) \leq 1\right\}$$

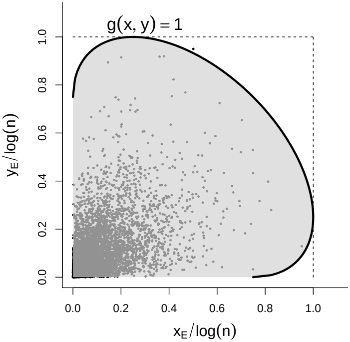
  as $n \to \infty$ (Balkema and Nolde, 2010; Nolde and Wadsworth, 2022).

Limit sets and gauge functions



Limit sets and gauge functions

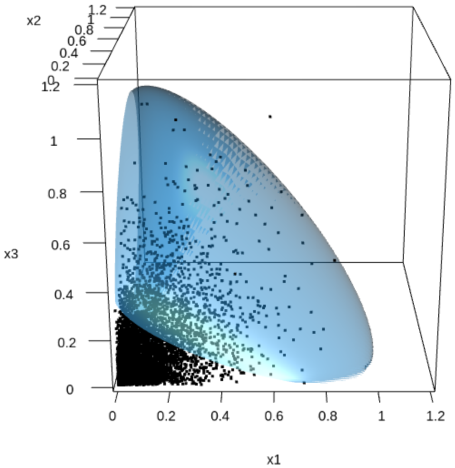




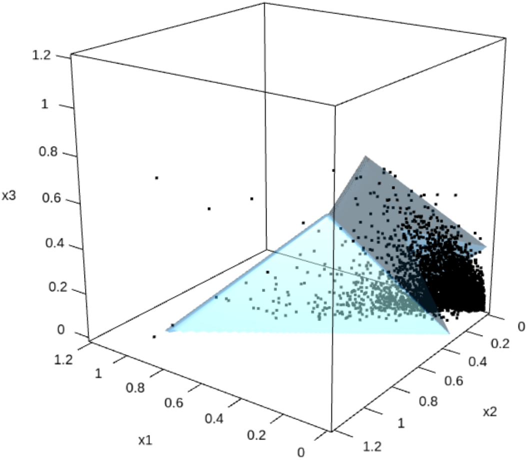
“pointy” $\Longrightarrow$ AD

“not pointy” $\Longrightarrow$ AI

Limit sets and gauge functions



Limit sets and gauge functions



Limit sets and gauge functions

Is this approach useful for extremal statistical inference across the multivariate tail?

If so, how can we estimate g from data and use it for extremal statistical inference?

Extremal inference with gauge functions

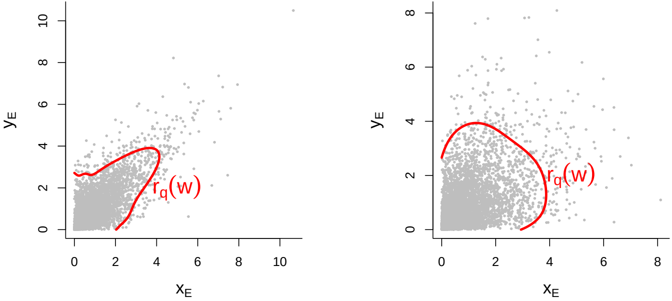
- Define $(R, \boldsymbol{W}) = (\|\boldsymbol{X}\|_1, \boldsymbol{X}/\|\boldsymbol{X}\|_1)$
- $f_{\boldsymbol{X}}$ satisfies $-\log f_{\boldsymbol{X}}(r\boldsymbol{w}) \sim rg(\boldsymbol{w})$ as $r \to \infty$.

Extremal inference with gauge functions

- Define $(R, \boldsymbol{W}) = (\|\boldsymbol{X}\|_1, \boldsymbol{X}/\|\boldsymbol{X}\|_1)$
- $f_{\boldsymbol{X}}$ satisfies $-\log f_{\boldsymbol{X}}(r\boldsymbol{w}) \sim rg(\boldsymbol{w})$ as $r \to \infty$.
- **Goal:** Model for $R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{W})\}$.

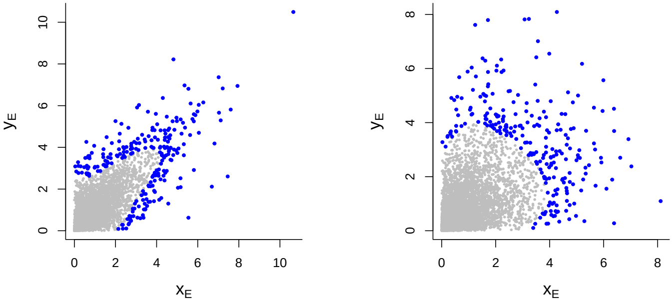
Extremal inference with gauge functions

- Define $(R, \boldsymbol{W}) = (\|\boldsymbol{X}\|_1, \boldsymbol{X}/\|\boldsymbol{X}\|_1)$
- $f_{\boldsymbol{X}}$ satisfies $-\log f_{\boldsymbol{X}}(r\boldsymbol{w}) \sim rg(\boldsymbol{w})$ as $r \to \infty$.
- **Goal:** Model for $R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{W})\}$.



Extremal inference with gauge functions

- Define $(R, \boldsymbol{W}) = (\|\boldsymbol{X}\|_1, \boldsymbol{X}/\|\boldsymbol{X}\|_1)$
- $f_{\boldsymbol{X}}$ satisfies $-\log f_{\boldsymbol{X}}(r\boldsymbol{w}) \sim rg(\boldsymbol{w})$ as $r \to \infty$.
- **Goal:** Model for $R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{W})\}$.



Geometric approach: model fitting

- Wadsworth and Campbell (2024): fit the model

$$R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{w})\} \sim \text{truncGamma}(\alpha, g(\boldsymbol{w}; \boldsymbol{\theta}))$$

by maximizing

$$L(\boldsymbol{\theta}; r_{1:n}, \boldsymbol{w}_{1:n}) = \prod_{i:r_i > r_q(\boldsymbol{w}_i)} \frac{f_{\text{Gamma}}\left(r_i; \alpha, g(\boldsymbol{w}_i; \boldsymbol{\theta})\right)}{\bar{F}_{\text{Gamma}}\left(r_q(\boldsymbol{w}_i); \alpha, g(\boldsymbol{w}_i; \boldsymbol{\theta})\right)}$$

Geometric approach: model fitting

- Wadsworth and Campbell (2024): fit the model

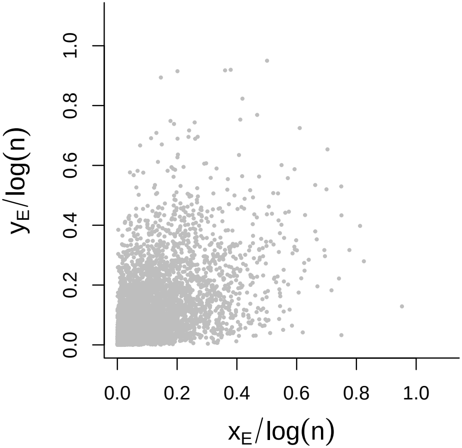
$$R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{w})\} \sim \text{truncGamma}(\alpha, g(\boldsymbol{w}; \boldsymbol{\theta}))$$

by maximizing

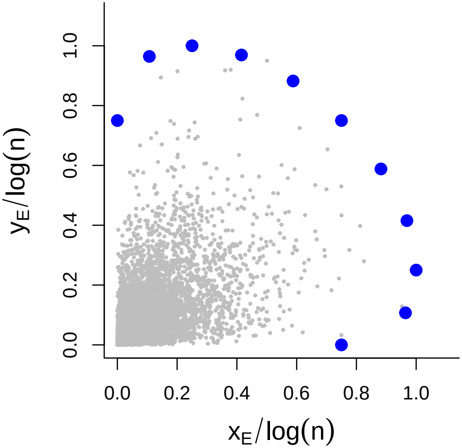
$$L(\boldsymbol{\theta}; r_{1:n}, \boldsymbol{w}_{1:n}) = \prod_{i: r_i > r_q(\boldsymbol{w}_i)} \frac{f_{\text{Gamma}}\left(r_i; \alpha, g(\boldsymbol{w}_i; \boldsymbol{\theta})\right)}{\bar{F}_{\text{Gamma}}\left(r_q(\boldsymbol{w}_i); \alpha, g(\boldsymbol{w}_i; \boldsymbol{\theta})\right)}$$

- **Campbell and Wadsworth (2024): Define $g(\boldsymbol{w}; \boldsymbol{\theta})$ piecewise-linearly.**

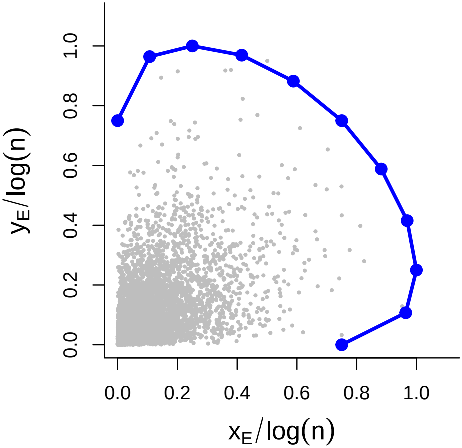
Semiparametric piecewise-linear approach



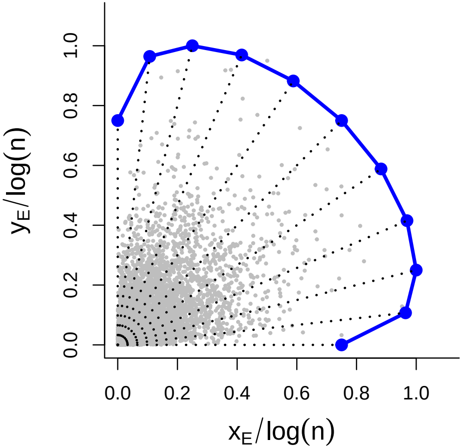
Semiparametric piecewise-linear approach



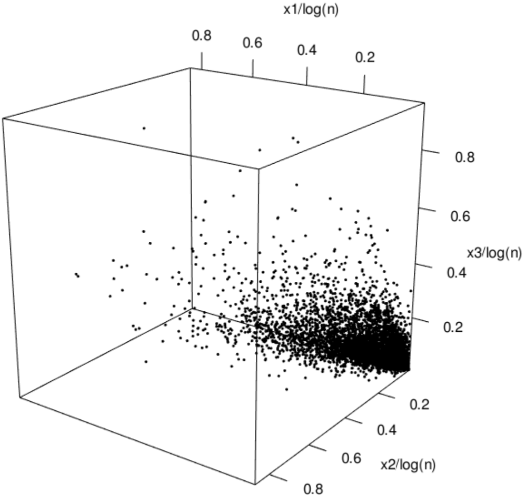
Semiparametric piecewise-linear approach



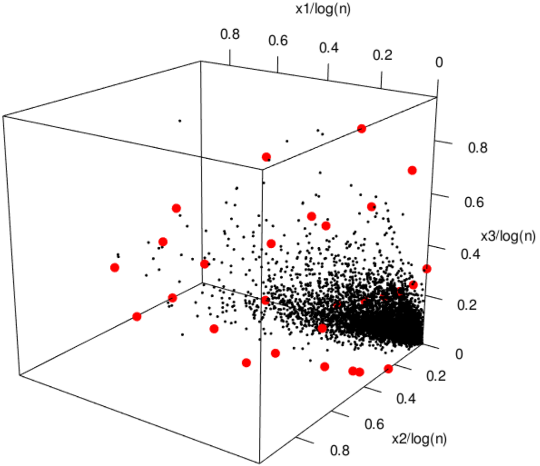
Semiparametric piecewise-linear approach



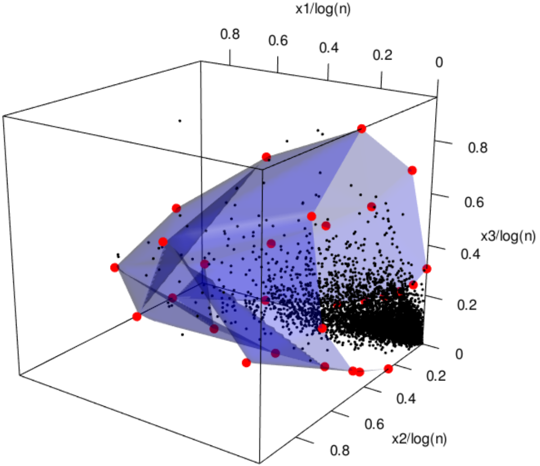
Semiparametric piecewise-linear approach



Semiparametric piecewise-linear approach



Semiparametric piecewise-linear approach



Semiparametric piecewise-linear approach

In d-dimensions...

- Define a set of N reference angles $\boldsymbol{w}^{\star 1}, \ldots, \boldsymbol{w}^{\star N} \in \mathcal{S}_{d-1}$.
- Results in M regions: $\triangle^{(1)}, \triangle^{(2)}, \ldots, \triangle^{(M)}$.

Semiparametric piecewise-linear approach

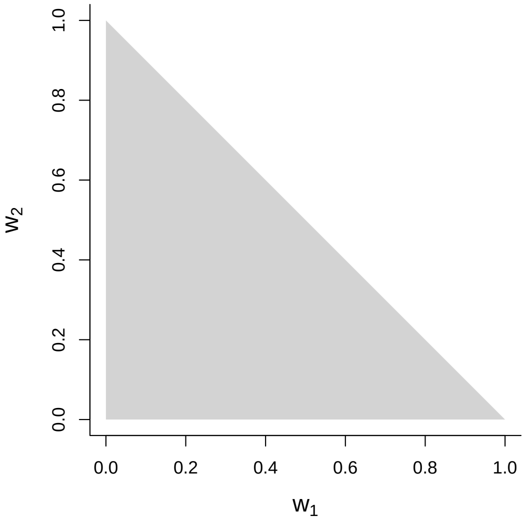
In d-dimensions...

- Define a set of N reference angles $\boldsymbol{w}^{\star 1}, \ldots, \boldsymbol{w}^{\star N} \in \mathcal{S}_{d-1}$.
- Results in M regions: $\triangle^{(1)}, \triangle^{(2)}, \ldots, \triangle^{(M)}$.
- At a point $\boldsymbol{x} \in \mathbb{R}^d$, the gauge function value is given by

$$g_{\mathrm{PWL}}(\boldsymbol{x}; \boldsymbol{\theta}) = \sum_{k=1}^{M} \mathbf{1}_{\triangle^{(k)}}\left(\boldsymbol{x}/\|\boldsymbol{x}\|\right) \frac{\boldsymbol{n}^{(k)\top}\boldsymbol{x}}{\boldsymbol{n}^{(k)\top}\theta_1^{(k)}\boldsymbol{w}^{\star(k),1}}$$

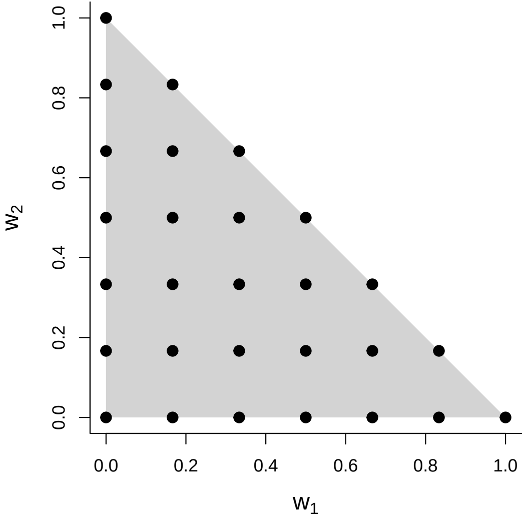
Semiparametric piecewise-linear approach

Partitioning the angular space, $\mathcal{S}_2 = \{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\|_1 = 1\}$



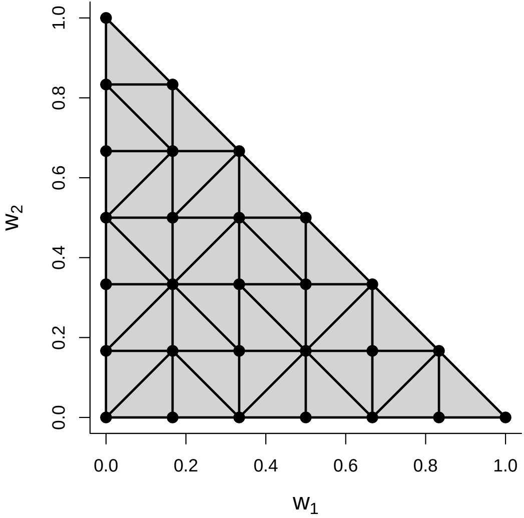
Semiparametric piecewise-linear approach

Partitioning the angular space, $\mathcal{S}_2 = \{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\|_1 = 1\}$

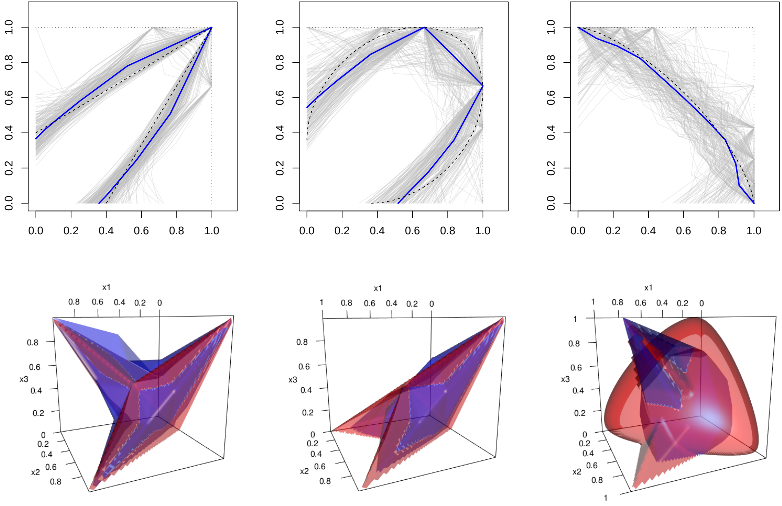


Semiparametric piecewise-linear approach

Partitioning the angular space, $\mathcal{S}_2 = \{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\|_1 = 1\}$



Semiparametric piecewise-linear approach



Geometric approach: extrapolation

- Can estimate extremal probabilities via

$$\Pr\left(\boldsymbol{X} \in B\right) = \Pr\left(\boldsymbol{X} \in B \middle| R > r_q(\boldsymbol{W})\right) \Pr\left(R > r_q(\boldsymbol{W})\right)$$

Geometric approach: extrapolation

- Can estimate extremal probabilities via

$$\Pr\left(\boldsymbol{X} \in B\right) = \Pr\left(\boldsymbol{X} \in B | R > r_q(\boldsymbol{W})\right) \Pr\left(R > r_q(\boldsymbol{W})\right)$$

- Compute $\Pr\left(\boldsymbol{X} \in B | R > r_q(\boldsymbol{W})\right)$ via sampling:
  1. sample $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N$ from $\boldsymbol{W} \mid \{R > r_q(\boldsymbol{W})\}$.
  2. sample r_i from truncGamma$(\widehat{\alpha}, g(\boldsymbol{w}_i; \widehat{\boldsymbol{\theta}}))$ for $i = 1, \ldots, N$
  3. return $\boldsymbol{x}_i = r_i \boldsymbol{w}_i$, $i = 1, \ldots, N$

Geometric approach: extrapolation

- Can estimate extremal probabilities via

$$\Pr\left(\boldsymbol{X} \in B\right) = \Pr\left(\boldsymbol{X} \in B | R > r_q(\boldsymbol{W})\right) \Pr\left(R > r_q(\boldsymbol{W})\right)$$

- Compute $\Pr\left(\boldsymbol{X} \in B | R > r_q(\boldsymbol{W})\right)$ via sampling:
  1. sample $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N$ from $\boldsymbol{W} \mid \{R > r_q(\boldsymbol{W})\}$.
  2. sample r_i from truncGamma$(\widehat{\alpha}, g(\boldsymbol{w}_i; \widehat{\boldsymbol{\theta}}))$ for $i = 1, \ldots, N$
  3. return $\boldsymbol{x}_i = r_i \boldsymbol{w}_i$, $i = 1, \ldots, N$
- Compute $\Pr\left(R > r_q(\boldsymbol{W})\right)$ empirically.

Geometric approach: extrapolation

- Can estimate extremal probabilities via

$$\Pr(\boldsymbol{X} \in B) = \Pr(\boldsymbol{X} \in B | R > r_q(\boldsymbol{W})) \Pr(R > r_q(\boldsymbol{W}))$$

- Compute $\Pr(\boldsymbol{X} \in B | R > r_q(\boldsymbol{W}))$ via sampling:
    1. sample $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N$ from $\boldsymbol{W} \mid \{R > r_q(\boldsymbol{W})\}$.
    2. sample r_i from truncGamma$(\widehat{\alpha}, g(\boldsymbol{w}_i; \widehat{\boldsymbol{\theta}}))$ for $i = 1, \ldots, N$
    3. return $\boldsymbol{x}_i = r_i \boldsymbol{w}_i$, $i = 1, \ldots, N$
- Compute $\Pr(R > r_q(\boldsymbol{W}))$ empirically.
- **Bonus!** Can extrapolate far into the tails using

$$\Pr(\boldsymbol{X} \in B) = \Pr(\boldsymbol{X} \in B | R > k r_q(\boldsymbol{W})) \Pr(R > k r_q(\boldsymbol{W}))$$

for $k > 1$.

Angular model

- $R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{w})\}$ only tells half the story.

Angular model

- $R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{w})\}$ only tells half the story.
- Papastathopoulos et al. (2025): Model $\boldsymbol{W} \mid \{R > r_q(\boldsymbol{W})\}$ via

$$f_{\boldsymbol{W}|\{R>r_q(\boldsymbol{W})\}}(\boldsymbol{w};\boldsymbol{\theta}) = \frac{g(\boldsymbol{w};\boldsymbol{\theta})^{-d}}{d\mathrm{vol}(\{\boldsymbol{x} : g(\boldsymbol{x};\boldsymbol{\theta}) \le 1\})}$$

Angular model

- $R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{w})\}$ only tells half the story.
- Papastathopoulos et al. (2025): Model $\boldsymbol{W} \mid \{R > r_q(\boldsymbol{W})\}$ via

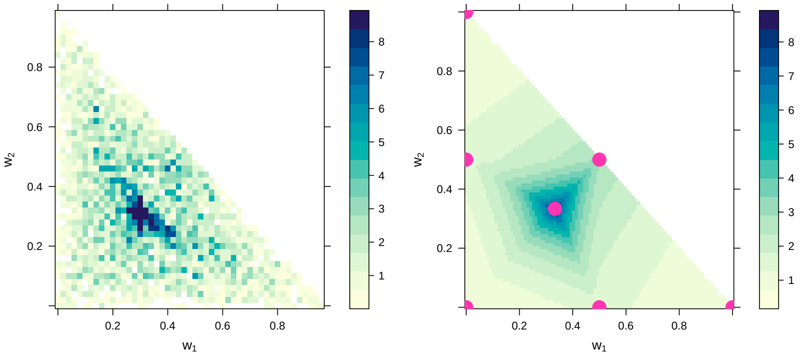
$$f_{\boldsymbol{W}\mid\{R>r_q(\boldsymbol{W})\}}(\boldsymbol{w};\boldsymbol{\theta}) = \frac{g(\boldsymbol{w};\boldsymbol{\theta})^{-d}}{d\mathrm{vol}(\{\boldsymbol{x} : g(\boldsymbol{x};\boldsymbol{\theta}) \leq 1\})}$$

- Campbell and Wadsworth (2024):
  - $\mathrm{vol}(\{\boldsymbol{x} : g_{\mathrm{PWL}}(\boldsymbol{x};\boldsymbol{\theta}) \leq 1\})$ has a closed-form expression.
  - samples are drawn from $f_{\boldsymbol{W}\mid\{R>r_q(\boldsymbol{W})\}}$ using MCMC.

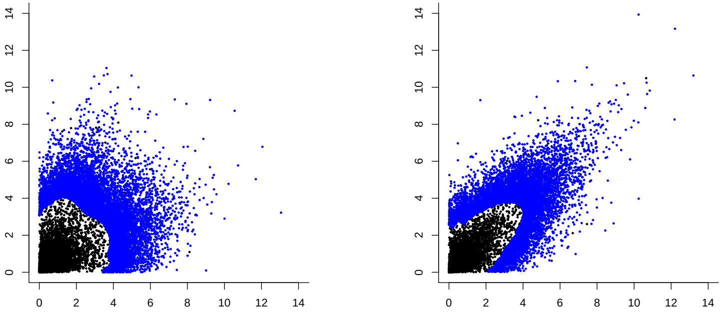
Angular model

Fitting the angular density

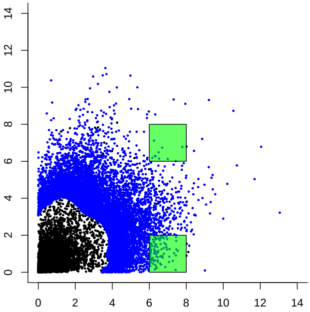
$$f_{\boldsymbol{W}|\{R>r_q(\boldsymbol{W})\}}(\boldsymbol{w};\boldsymbol{\theta}) = \frac{g_{\mathrm{PWL}}(\boldsymbol{w};\boldsymbol{\theta})^{-d}}{d\mathrm{vol}(\{\boldsymbol{x} : g_{\mathrm{PWL}}(\boldsymbol{x};\boldsymbol{\theta}) \leq 1\})}$$

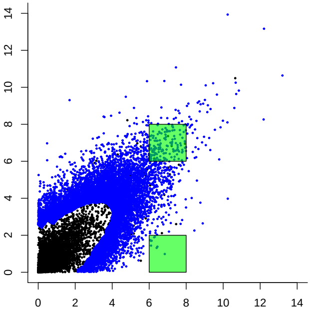


Geometric approach: extrapolation

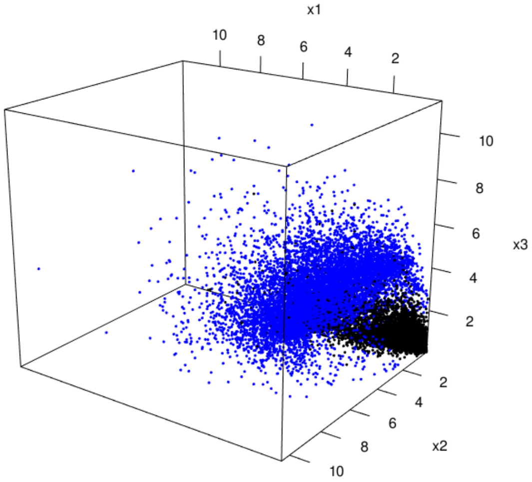


Geometric approach: extrapolation

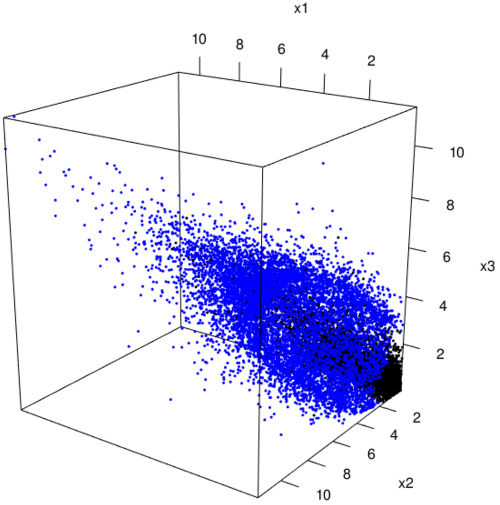




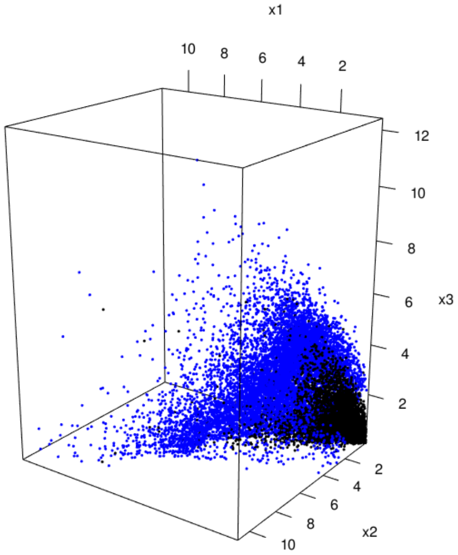
Geometric approach: extrapolation



Geometric approach: extrapolation



Geometric approach: extrapolation



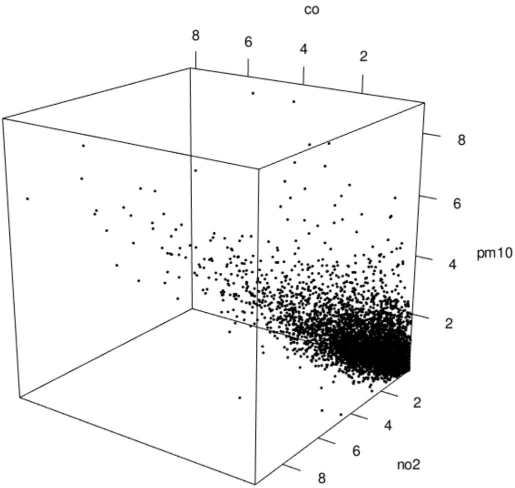
Application to air pollution measurements ($d = 3$)

- North Kensington site, London, UK
- carbon monoxide (CO, mg/m^3), nitrogen dioxide (NO_2, $\mu\text{g/m}^3$), and particles with a diameter of 10 μm or less (PM10, mg/m^3).
- $n = 5,584$ daily maximum measurements, October–April. 1996–2024.

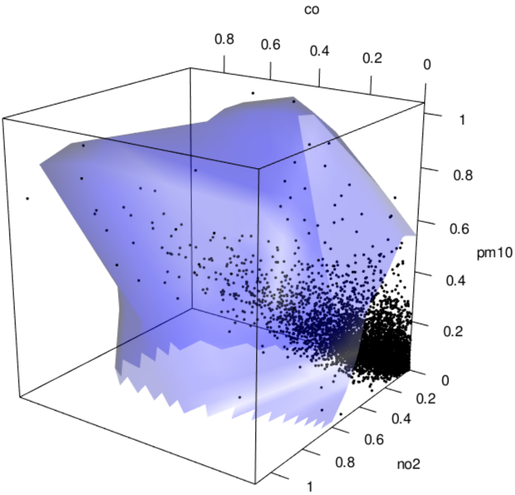
Application to air pollution measurements ($d = 3$)

- North Kensington site, London, UK
- carbon monoxide (CO, mg/m^3), nitrogen dioxide (NO_2, μg/m^3), and particles with a diameter of 10 μm or less (PM10, mg/m^3).
- $n = 5,584$ daily maximum measurements, October–April. 1996–2024.
- Using methods from Simpson et al. (2020), evidence that
  - $\{\text{PM10}\}$ is large when $\{\text{CO}, \text{NO}_2\}$ are small
  - $\{\text{CO}, \text{NO}_2, \text{PM10}\}$ grow large together.

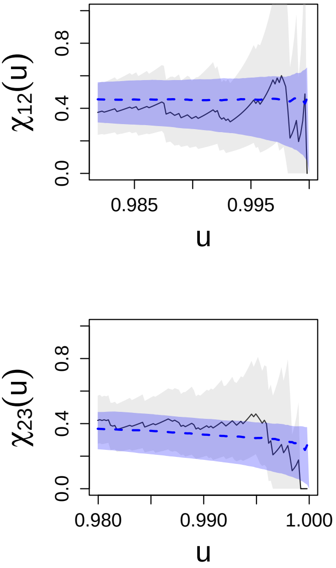
Application to air pollution measurements ($d = 3$)



Application to air pollution measurements ($d = 3$)



$\chi_C(u)$ estimates, $C \subseteq \{1, 2, 3\}$

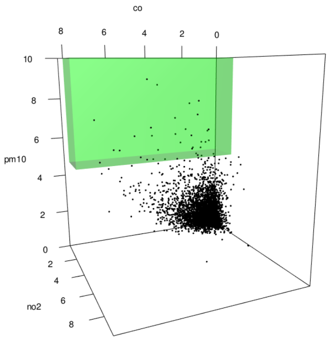


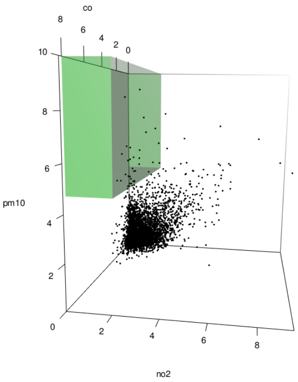
An alternate tail probability

For $d = 3$ pollution data, consider

$$\psi_{\{3\}}(u; \delta_1, \delta_2) = \Pr\left[F_E(X_1) < \delta_1, F_E(X_2) < \delta_2, F_E(X_3) > u\right]$$

for $\delta_1, \delta_2 \in [0, 1)$, $u > u_0$, $u_0 \in [0, 1)$ close to 1.



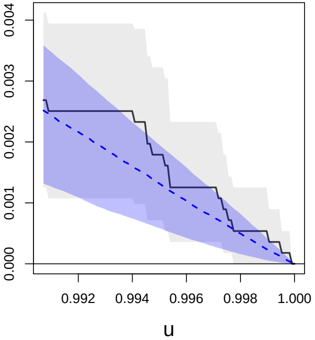


An alternate tail probability

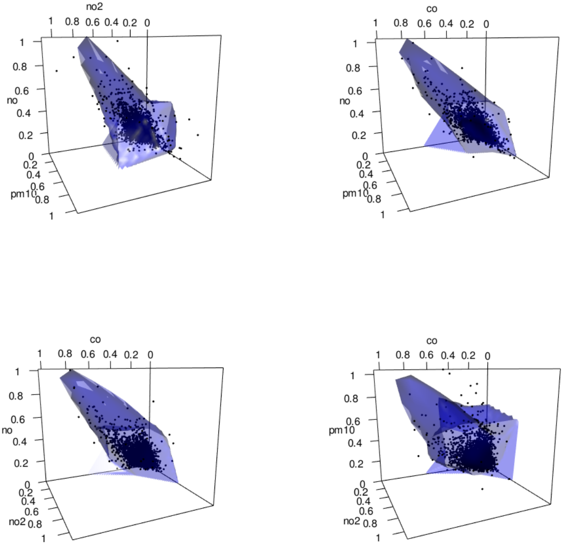
For $d = 3$ pollution data, consider

$$\psi_{\{3\}}(u; \delta_1, \delta_2) = \Pr\left[F_E(X_1) < \delta_1, F_E(X_2) < \delta_2, F_E(X_3) > u\right]$$

for $\delta_1, \delta_2 \in [0, 1)$, $u > u_0$, $u_0 \in [0, 1)$ close to 1.



Application to air pollution measurements ($d = 4$)



Thank you!

Bibliography

A. A. Balkema and N. Nolde. Asymptotic independence for unimodal densities. *Advances in Applied Probability*, 42(2):411–432, 2010.

R. Campbell and J. Wadsworth. Piecewise-linear modeling of multivariate geometric extremes. *arXiv:2412.05195*, 2024.

N. Nolde and J. L. Wadsworth. Linking representations for multivariate extremes via a limit set. *Advances in Applied Probability*, 54(3): 688–717, 2022.

I. Papastathopoulos, L. de Monte, R. Campbell, and H. Rue. Statistical inference for radial generalized pareto distributions and return sets in geometric extremes. *arXiv:2310.06130*, 2025.

E. S. Simpson, J. L. Wadsworth, and J. A. Tawn. Determining the dependence structure of multivariate extremes. *Biometrika*, 107(3): 513–532, 2020.

J. L. Wadsworth and R. Campbell. Statistical inference for multivariate extremes via a geometric approach. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 86(5):1243–1265, 03 2024.