Robust Risk Management

Carole Bernard

Grenoble EM and VU Brussel

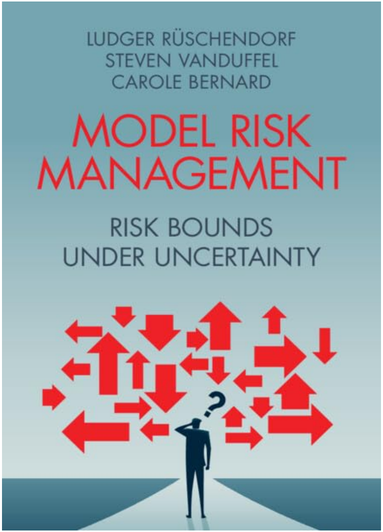
January 2024,
21st Winter School on Mathematical Finance
Part I

First Part: Model Risk on the Dependence: Theory and Computational Approach via The Rearrangement Algorithm

Second Part: Model Risk on the Aggregate Variable

A book to appear in January 2024...

L. Rüschendorf , S. Vanduffel, C. Bernard, *Cambridge Univ. Press*.



Acknowledgment of Collaboration (1/2)

- Bernard, C., X. Jiang, S. Vanduffel, (2012). *Note on Improved Fréchet Bounds*, **Journal of Applied Probability**.
- Bernard, C., X. Jiang, R. Wang, (2013) *Risk Aggregation with Dependence Uncertainty*, **Insurance: Mathematics and Economics**.
- Bernard, C., Vanduffel, S. (2015). *A new approach to assessing model risk in high dimensions.* **Journal of Banking and Finance**.
- Bernard, C. , Rüschendorf, L., Vanduffel, S., Yao, J. (2015). *How robust is the Value-at-Risk of credit risk portfolios?* **European Journal of Finance**.
- Bernard, C. , McLeish D. (2016). *Algorithms for Finding Copulas Minimizing Convex Functions of Sums* **Asia Pacific Journal of Operational Research**.
- Bernard, C., Rüschendorf, L., Vanduffel, S. (2017). *Value-at-Risk bounds with variance constraints*. **Journal of Risk and Insurance**.
- Bernard, C., L. Rüschendorf, S. Vanduffel, R. Wang (2017) *Risk bounds for factor models,* 2017, **Finance and Stochastics**.
- Bernard, C., Denuit, M., Vanduffel, S. (2018). *Measuring Portfolio Risk Under Partial Dependence Information*. **Journal of Risk and Insurance**.
- Bernard, C., Bondarenko, O., Vanduffel, S. (2018). *Rearrangement Algorithm and Maximum Entropy*. **Annals of Operational Research**.

Acknowledgement of Collaboration (2/2)

- Bernard, C., Kazzi R., Vanduffel, S. (2020). *Range Value-at-Risk bounds for unimodal distributions under partial information*. **Insurance: Mathematics and Economics**.
- De Gennaro, L. Bernard, C., (2020). *Bounds on Multi-asset Derivatives via Neural Networks*. **International Journal of Theoretical and Applied Finance**.
- Bernard, C., Müller, A. (2020). *Dependence uncertainty bounds for the energy score and the multivariate Gini mean difference*. **Dependence Modelling**.
- Bernard, C., Bondarenko, O., Vanduffel, S. (2021). *A Model-free Approach to Multivariate Option Pricing*. **Review of Derivatives Research**.
- C. Bernard, C. De Vecchi, S. Vanduffel (2023) *Impact of Correlation on the (Range) Value-at-Risk,* **Scandinavian Actuarial Journal**.
- Bernard, C., Chen, J., Rüschendorf, L., Vanduffel, S. (2023) *Coskewness under dependence uncertainty*, **Statistics & Probability Letters**.
- Bondarenko, O., Bernard, C., (2024). *Option Implied dependence and Correlation Risk Premium*. **Journal of Financial and Quantitative Analysis**, forthcoming.
- Bernard, C., Pesenti, S., Vanduffel, S. (2024) *Robust Distortions Measures*, **Mathematical Finance**, forthcoming.
- Bernard, C., Müller, A. and M. Oesting, (2024), *Lp-norm spherical copulas*, **Journal of Multivariate Analysis**, forthcoming.
- Bernard, C., Chen, J., Rüschendorf, L., Vanduffel, S. (2024) *Improved block rearrangement algorithm*, Working paper.

Outline

Part 1: The Rearrangement Algorithm

- Minimizing variance of a sum with full dependence uncertainty
- Variance bounds

Part 2: Application to Model-Risk Assessment, e.g., Uncertainty on Value-at-Risk

- With 2 risks and full dependence uncertainty
- With d risks and full dependence uncertainty

Part 3: Adding information on dependence

- Moment constraints
- Information on a subset...

Part 4: Using the RA to infer dependence

- Add information about the sum of the risks
- Application to explain the correlation risk premium
- Application to multivariate option pricing

Part 5: Improved Rearrangement Algorithm

Part I

The Rearrangement Algorithm

Portfolio with minimum variance

Background

Assumptions:

- **Marginals known**: $X_i \sim F_i$ for $i = 1, 2, ..., n$
- **Dependence fully unknown** (any dependence structure (copula) is possible)

With f convex,

- In two dimensions $n = 2$, bounds on variance are obtained using **Fréchet-Hoeffding bounds** or "extreme dependence".

$$E[f(F_1^{-1}(U)+F_2^{-1}(1-U))] \leqslant E[f(X_1+X_2)] \leqslant E[f(F_1^{-1}(U)+F_2^{-1}(U))]$$

Background

Assumptions:

- **Marginals known**: $X_i \sim F_i$ for $i = 1, 2, ..., n$
- **Dependence fully unknown** (any dependence structure (copula) is possible)

With f convex,

- In two dimensions $n = 2$, bounds on variance are obtained using **Fréchet-Hoeffding bounds** or "extreme dependence".

$$E[f(F_1^{-1}(U)+F_2^{-1}(1-U))] \leqslant E[f(X_1+X_2)] \leqslant E[f(F_1^{-1}(U)+F_2^{-1}(U))]$$

- When $n \geqslant 2$, the **upper** bound corresponds to the comonotonic scenario,

$$E[f(X_1+X_2+...+X_n)] \leqslant E[f(F_1^{-1}(U)+F_2^{-1}(U)+...+F_n^{-1}(U))]$$

Results on the lower bound in dimensions $n \geqslant 3$

- If $n \geqslant 3$, the Fréchet-Hoeffding lower bound does not exist:

Definition: Complete mixability (Wang and Wang (2011))

$X_1 \sim F_1, ..., X_n \sim F_n$ are **completely mixable** if there exists a dependence structure between X_1,...,X_n such that $X_1 + X_2 + ... + X_n = \sum_i E[X_i]$.

Results on the lower bound in dimensions $n \geqslant 3$

- If $n \geqslant 3$, the Fréchet-Hoeffding lower bound does not exist:

Definition: Complete mixability (Wang and Wang (2011))

$X_1 \sim F_1, ..., X_n \sim F_n$ are **completely mixable** if there exists a dependence structure between X_1,...,X_n such that $X_1 + X_2 + ... + X_n = \sum_i E[X_i]$.

- Puccetti and Rüschendorf (2012): algorithm (RA)
  - Inputs: $X_1 \sim F_1$, ..., $X_n \sim F_n$
  - Goal: look for copulas that solve min $E[f(X_1 + X_2 + ... + X_n)]$ for f convex

Solving for the minimum variance

- **Inputs:** $X_1 \sim F_1$, $X_2 \sim F_2$ ..., $X_d \sim F_d$
- **Goal:** look for a dependence such that

$$\min var(X_1 + X_2 + ... + X_d)$$

- *Algorithm:* Each X_j is sampled into n **equiprobable** values: consider the realizations $x_{ij} := F_j^{-1}(\frac{i-0.5}{n})$:

$$\mathbf{X} = [X_1, X_2, \dots, X_d] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1d} \\ x_{21} & x_{22} & \dots & x_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nd} \end{bmatrix}$$

- Rearrange elements x_{ij} (by columns) such that after the rearrangement variance of sum S is minimized?
- This is an **NP complete** problem (Haus (2014)). Brute force search requires checking $(n!)^{(d-1)}$ rearrangements.

Rearrangement Algorithm

$N = 4$ observations of $d = 3$ variables: X_1, X_2, X_3

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 6 & 3 \\ 4 & 0 & 0 \\ 6 & 3 & 4 \end{bmatrix}$$

Each column: **marginal** distribution.
Interaction among columns: **dependence** among the risks.

Before the Rearrangement Algorithm...

Partition problem

Partition the multiset $\mathcal{S}$ of positive integers into two subsets $\mathcal{S}_1$ and $\mathcal{S}_2$ such that the difference between the sum of elements in $\mathcal{S}_1$ and the sum of elements in $\mathcal{S}_2$ is minimized.

Example: $\mathcal{S} = \{8, 7, 6, 5, 4\}$ would optimally be split as $\mathcal{S}_1 = \{8, 7\}$ and $\mathcal{S}_2 = \{6, 5, 4\}$.

Greedy Algorithm: iterates through the numbers in descending order, assigning each of them to whichever subset has the smaller sum.

$\mathcal{S}_1 = \{8, 5, 4\}$ and $\mathcal{S}_2 = \{7, 6\}$

Numerical example of the Greedy Algorithm

$$\begin{bmatrix} 8 & 7 & 6 & 5 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 8 & 0 & \cdot & \cdot & \cdot \\ 0 & 7 & \cdot & \cdot & \cdot \end{bmatrix} \qquad \begin{bmatrix} 8 \\ 7 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 8 & 0 & 0 & \cdot & \cdot \\ 0 & 7 & 6 & \cdot & \cdot \end{bmatrix} \qquad \begin{bmatrix} 8 \\ 13 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 8 & 0 & 0 & 5 & 4 \\ 0 & 7 & 6 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 17 \\ 13 \end{bmatrix}$$

In the Greedy algorithm, sort the elements of subsequent columns in reverse order of the row sums taken across all **previous** columns

Assembly Line Crew Scheduling

Assembly Line Crew Scheduling problem

How to rearrange elements within columns of a matrix such that variability among the row sums becomes minimum

Greedy algorithm works in higher dimensions

$$\Rightarrow \begin{bmatrix} 5 & 4 & 3 \\ 4 & 0 & 5 \\ 3 & 3 & 0 \end{bmatrix} \qquad \begin{bmatrix} 12 \\ 9 \\ 6 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 5 & \mathbf{0} & 3 \\ 4 & \mathbf{3} & 5 \\ 3 & \mathbf{4} & 0 \end{bmatrix} \qquad \begin{bmatrix} 8 \\ 12 \\ 7 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 5 & 0 & \mathbf{5} \\ 4 & 3 & \mathbf{3} \\ 3 & 4 & \mathbf{0} \end{bmatrix} \qquad \begin{bmatrix} 10 \\ 10 \\ 7 \end{bmatrix}$$

Coffman and Yannakis (MOR, 1984) and Hsu (MOR, 1984)

Assembly Line Crew Scheduling

Assembly Line Crew Scheduling problem

How to rearrange elements within columns of a matrix such that variability among the row sums becomes minimum

Rearrangement Algorithm (Rüschendorf, ZOR, 1983): sort the elements of subsequent columns in reverse order of the row sums taken across **all other** columns

$$\begin{bmatrix} 5 & 4 & 3 \\ 4 & 0 & 5 \\ 3 & 3 & 0 \end{bmatrix} \begin{bmatrix} 12 \\ 9 \\ 6 \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{3} & 4 & 3 \\ \mathbf{4} & 0 & 5 \\ \mathbf{5} & 3 & 0 \end{bmatrix} \begin{bmatrix} 10 \\ 9 \\ 8 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3 & \mathbf{3} & 3 \\ 4 & \mathbf{0} & 5 \\ 5 & \mathbf{4} & 0 \end{bmatrix} \begin{bmatrix} 9 \\ 9 \\ 9 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & 3 & \mathbf{3} \\ 4 & 0 & \mathbf{5} \\ 5 & 4 & \mathbf{0} \end{bmatrix} \begin{bmatrix} 9 \\ 9 \\ 9 \end{bmatrix}$$

Same marginals, different dependence ⇒ Effect on the sum!

$$\begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{2} \\ \mathbf{0} & \mathbf{6} & \mathbf{3} \\ \mathbf{4} & \mathbf{0} & \mathbf{0} \\ \mathbf{6} & \mathbf{3} & \mathbf{4} \end{bmatrix} \qquad S_N = \overset{X_1 + X_2 + X_3}{\begin{bmatrix} 4 \\ 9 \\ 4 \\ 13 \end{bmatrix}}$$

$$\begin{bmatrix} \mathbf{6} & \mathbf{6} & \mathbf{4} \\ \mathbf{4} & \mathbf{3} & \mathbf{3} \\ \mathbf{1} & \mathbf{1} & \mathbf{2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad S_N = \overset{X_1 + X_2 + X_3}{\begin{bmatrix} 16 \\ 10 \\ 3 \\ 0 \end{bmatrix}}$$

Aggregate Risk with Maximum Variance

comonotonic scenario S^c

Rearrangement Algorithm: Sum with Minimum Variance

minimum variance with $d = 2$ risks X_1 and X_2

Antimonotonicity: $var(\mathbf{X_1^a} + X_2) \leqslant var(\mathbf{X_1} + X_2)$.

How about in d dimensions?

Rearrangement Algorithm: Sum with Minimum Variance

minimum variance with $d = 2$ risks X_1 and X_2

Antimonotonicity: $var(\mathbf{X_1^a} + X_2) \leqslant var(\mathbf{X_1} + X_2)$.

How about in d dimensions?
Use of the rearrangement algorithm on the original matrix M.

Aggregate Risk with Minimum Variance

- Columns of M are rearranged such that they become anti-monotonic with the sum of all other columns:

$$\forall k \in \{1, 2, ..., d\}, \mathbf{X_k^a} \textit{ antimonotonic with } \sum_{j \neq k} X_j.$$

- After each step, $var\left(\mathbf{X_k^a} + \sum_{j \neq k} X_j\right) \leqslant var\left(\mathbf{X_k} + \sum_{j \neq k} X_j\right)$ where $\mathbf{X_k^a}$ is antimonotonic with $\sum_{j \neq k} X_j$.

Aggregate risk with minimum variance
Step 1: First column

$$\begin{array}{c} \downarrow \\ \begin{bmatrix} \mathbf{6} & 6 & 4 \\ \mathbf{4} & 3 & 2 \\ \mathbf{1} & 1 & 1 \\ \mathbf{0} & 0 & 0 \end{bmatrix} \end{array} \quad \begin{array}{c} X_2 + X_3 \\ 10 \\ 5 \\ 2 \\ 0 \end{array} \quad \text{becomes} \quad \begin{bmatrix} \mathbf{0} & 6 & 4 \\ \mathbf{1} & 3 & 2 \\ \mathbf{4} & 1 & 1 \\ \mathbf{6} & 0 & 0 \end{bmatrix}$$

Aggregate risk with minimum variance

$$\begin{bmatrix} \downarrow & & \\ \mathbf{6} & 6 & 4 \\ \mathbf{4} & 3 & 2 \\ \mathbf{1} & 1 & 1 \\ \mathbf{0} & 0 & 0 \end{bmatrix} \quad \begin{matrix} X_2 + X_3 \\ 10 \\ 5 \\ 2 \\ 0 \end{matrix} \quad \text{becomes} \quad \begin{bmatrix} \mathbf{0} & 6 & 4 \\ \mathbf{1} & 3 & 2 \\ \mathbf{4} & 1 & 1 \\ \mathbf{6} & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} & \downarrow & \\ 0 & \mathbf{6} & 4 \\ 1 & \mathbf{3} & 2 \\ 4 & \mathbf{1} & 1 \\ 6 & \mathbf{0} & 0 \end{bmatrix} \quad \begin{matrix} X_1 + X_3 \\ 4 \\ 3 \\ 5 \\ 6 \end{matrix} \quad \text{becomes} \quad \begin{bmatrix} 0 & \mathbf{3} & 4 \\ 1 & \mathbf{6} & 2 \\ 4 & \mathbf{1} & 1 \\ 6 & \mathbf{0} & 0 \end{bmatrix}$$

$$\begin{bmatrix} & & \downarrow \\ 0 & 3 & \mathbf{4} \\ 1 & 6 & \mathbf{2} \\ 4 & 1 & \mathbf{1} \\ 6 & 0 & \mathbf{0} \end{bmatrix} \quad \begin{matrix} X_1 + X_2 \\ 3 \\ 7 \\ 5 \\ 6 \end{matrix} \quad \text{becomes} \quad \begin{bmatrix} 0 & 3 & \mathbf{4} \\ 1 & 6 & \mathbf{0} \\ 4 & 1 & \mathbf{2} \\ 6 & 0 & \mathbf{1} \end{bmatrix}$$

Aggregate risk with minimum variance

Each column is antimonotonic with the sum of the others:

$$\begin{array}{c} \downarrow \\ \begin{bmatrix} 0 & 3 & 4 \\ 1 & 6 & 0 \\ 4 & 1 & 2 \\ 6 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} X_2+X_3 \\ 7 \\ 6 \\ 3 \\ 1 \end{array}, \quad \begin{array}{c} \downarrow \\ \begin{bmatrix} 0 & 3 & 4 \\ 1 & 6 & 0 \\ 4 & 1 & 2 \\ 6 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} X_1+X_3 \\ 4 \\ 1 \\ 6 \\ 7 \end{array}, \quad \begin{array}{c} \downarrow \\ \begin{bmatrix} 0 & 3 & 4 \\ 1 & 6 & 0 \\ 4 & 1 & 2 \\ 6 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} X_1+X_2 \\ 3 \\ 7 \\ 5 \\ 6 \end{array}$$

$$\begin{bmatrix} 0 & 3 & 4 \\ 1 & 6 & 0 \\ 4 & 1 & 2 \\ 6 & 0 & 1 \end{bmatrix} \qquad S_N = \overset{X_1+X_2+X_3}{\begin{bmatrix} 7 \\ 7 \\ 7 \\ 7 \end{bmatrix}}$$

The minimum variance of the sum is equal to 0! Ideal case of a constant sum (*complete mixability*, see Wang and Wang (2011)).

Block Rearrangement Algorithm

With more than 3 variables, we can **improve the standard algorithm** (which proceeds column by column) by proceeding by block!

ρ : Pearson correlation

Necessary condition to minimize variance

If $\text{var}\left(\sum_i \mathbf{X}_i\right)$ is minimum then $\rho\left(\sum_{i\in\Pi} \mathbf{X}_i, \sum_{i\in\bar{\Pi}} \mathbf{X}_i\right)$ is minimum for every partition of $\{1, 2, ..., n\}$ into two sets Π and $\bar{\Pi}$. However, the converse does not hold in general.

Block Rearrangement Algorithm

With more than 3 variables, we can **improve the standard algorithm** (which proceeds column by column) by proceeding by block!

ρ : Pearson correlation

Necessary condition to minimize variance

If $\mathrm{var}\left(\sum_i \mathbf{X}_i\right)$ is minimum then $\rho\left(\sum_{i\in\Pi} \mathbf{X}_i, \sum_{i\in\bar{\Pi}} \mathbf{X}_i\right)$ is minimum for every partition of $\{1, 2, ..., n\}$ into two sets Π and $\bar{\Pi}$. However, the converse does not hold in general.

Block Rearrangement Algorithm:

1. Select a random sample of n_{sim} possible partitions of the columns $\{1, 2, ..., n\}$ into two non-empty subsets $\{\Pi, \bar{\Pi}\}$.
2. For each partition, rearrange the second block so that $S_{\bar{\Pi}}$ is antimonotonic to the values of S_{Π}.
3. If there is no improvement in $\mathrm{var}(\sum_i \mathbf{X}_i)$, output the current matrix $\mathbf{X}$, otherwise return to step 1.

A New Multivariate Measure

Definition

Let $\phi(\mathbf{X}_1, \mathbf{X}_2)$ be a measure of dependence between $\mathbf{X}_1$ and $\mathbf{X}_2$. For a matrix of data $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_{n-1}, \mathbf{X}_n]$ with n columns, define

$$\varrho(\mathbf{X}) := \frac{1}{2^{n-1}-1} \sum_{\Pi \in \mathcal{P}} \phi\left(\sum_{i \in \Pi} \mathbf{X}_i, \sum_{i \in \bar{\Pi}} \mathbf{X}_i \right)$$

where the sum is over the set $\mathcal{P}$ consisting of $2^{n-1}-1$ distinct partitions of $\{1, 2, ..., n\}$ into **non-empty** subsets Π and its complement $\bar{\Pi}$.

- Using a bivariate dependence measure that is minimum at -1 (Spearman's rho, Kendall's tau). Then, a necessary condition to be at the minimum variance is that $\varrho(\mathbf{X}) = -1$.
- This condition can also be used as a *stopping rule*.

Some observations on the Block Rearrangement Algorithm

1. In general, many local minima for the variance of the sum:
   - **not at the minimum variance but very close to it**.
2. the BRA outperforms the RA by several order of magnitude (variance is 10 to 100 times smaller, global minimum is reached more often,...)

Information on the RA, R codes available from `https://sites.google.com/site/rearrangementalgorithm/`.
Matlab codes can be obtained from myself.

Bounds on variance (theory)

Analytical Bounds on Standard Deviation

Consider d risks X_i with standard deviation σ_i

$$0 \leqslant std(X_1 + X_2 + ... + X_d) \leqslant \sigma_1 + \sigma_2 + ... + \sigma_d.$$

Example with 20 normal N(0,1)

$$0 \leqslant std(X_1 + X_2 + ... + X_{20}) \leqslant 20,$$

in this case, both bounds are sharp and too wide for practical use!

And the dependence structures that achieve these bounds are relatively easy to guess.

Bounds on variance (theory)

Case of Bernoulli distributions:

- X_i takes value 1 with probability p_i
- Define M such that
  $\sum_i E[X_i] = \mu := p_1 + p_2 + ... + p_N \in [M, M+1[$
- The dependence between X_i such that $var(\sum_i X_i)$ is minimum is such that $\sum_i X_i$ takes exactly two values M with probability $p_M = M + 1 - \mu$, and $M + 1$ with probability $1 - p_M = \mu - M$.

And the dependence structure that achieves this minimum bound is relatively easy to guess.

Bounds on variance (practice)

Case of Arbitrary Distributions

In general the dependence structure that minimizes the variance is not easy to guess:

- X_i has distribution F_i
- Discretize F_i and put the values in a matrix.
- Apply the RA or the BRA
- The dependence between X_i such that $var(\sum_i X_i)$ is minimum is obtained as the output of the algorithm.

Part II-a

Introduction to Model Risk

- Due to Uncertainty on the Dependence
- Why the RA allows to quantify model risk on variance estimation but also on many other risk measures

Motivation on VaR aggregation with dependence uncertainty

Full information on **marginal distributions**:

$$X_j \sim F_j$$

$$+$$

Full Information on **dependence**:

(known copula)

$$\Rightarrow$$

$\mathrm{VaR}_q\left(X_1 + X_2 + ... + X_d\right)$ can be computed!

Motivation on VaR aggregation with dependence uncertainty

Full information on **marginal distributions**:

$$X_j \sim F_j$$

$$+$$

Partial or **no** Information on **dependence**:
(incomplete information on copula)

$$\Rightarrow$$

$\mathrm{VaR}_q\left(X_1 + X_2 + ... + X_d\right)$ **cannot** be computed!

Only a range of possible values for $\mathrm{VaR}_q\left(X_1 + X_2 + ... + X_d\right)$.

Model Risk

1. Goal: Assess the risk of a portfolio sum $S = \sum_{i=1}^{d} X_i$.
2. Choose a risk measure $\rho(\cdot)$: variance, Value-at-Risk...
3. "Fit" a multivariate distribution for $(X_1, X_2, ..., X_d)$ and compute $\rho(S)$
4. How about model risk? How wrong can we be?

Model Risk

1. Goal: Assess the risk of a portfolio sum $S = \sum_{i=1}^{d} X_i$.
2. Choose a risk measure $\rho(\cdot)$: variance, Value-at-Risk...
3. "Fit" a multivariate distribution for $(X_1, X_2, ..., X_d)$ and compute $\rho(S)$
4. How about model risk? How wrong can we be?

Assume $\rho(S) = var(S)$,

$$\rho^+_{\mathcal{F}} := \sup \left\{ var\left(\sum_{i=1}^{d} X_i\right) \right\}, \quad \rho^-_{\mathcal{F}} := \inf \left\{ var\left(\sum_{i=1}^{d} X_i\right) \right\}$$

where the bounds are taken over all other (joint distributions of) random vectors $(X_1, X_2, ..., X_d)$ that "agree" with the available information $\mathcal{F}$

Aggregation with dependence uncertainty: Case of Variance - First Application of the RA

- **Marginals known**
- **Dependence fully unknown**

Minimum variance of the portfolio can be obtained using the RA. Similarly, the uncertainty on any risk measure that is consistent with convex order can be assessed.

Part II-b

Another application of the Rearrangement Algorithm

VaR aggregation with dependence uncertainty

- Maximum Value-at-Risk is not caused by the comonotonic scenario.
- Maximum Value-at-Risk is achieved when the variance is *minimum* in the tail. The RA is then used in the tails only.

Risk Aggregation and full dependence uncertainty
Literature review

- Marginals known
- Dependence fully unknown
- Explicit sharp (attainable) bounds
  - $n = 2$ (Makarov (1981), Rüschendorf (1982))
  - Rüschendorf & Uckelmann (1991), Denuit, Genest & Marceau (1999), Embrechts & Puccetti (2006),
- A challenging problem in $n \geqslant 3$ dimensions
- Approximate sharp bounds
  - Puccetti and Rüschendorf (2012): algorithm (RA) useful to approximate the minimum variance.
  - Embrechts, Puccetti, Rüschendorf (2013): algorithm (RA) to find bounds on VaR

Bounds with full dependence uncertainty

(Unconstrained bounds)

TVaR Bounds with full dependence uncertainty

$$\sum_{i=1}^{d} E[X_i] \leqslant TVaR\left(\sum_{i=1}^{d} X_i\right) \leqslant TVaR\left(\sum_{i=1}^{d} X_i^c\right)$$

in which X_i^c denotes a randiom variable with the same distribution F_i as X_i such that for all i

$$X_i^c = F_i^{-1}(U)$$

for some U uniformly distributed over $(0, 1)$.

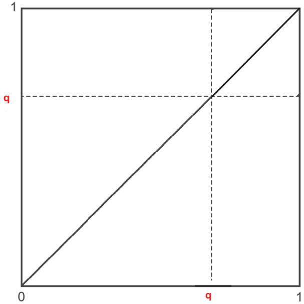
VaR Bounds with full dependence uncertainty

(Unconstrained VaR bounds)

"Riskiest" Dependence: maximum VaR_q in 2 dims?

If X_1 and X_2 are U(0,1) comonotonic, then

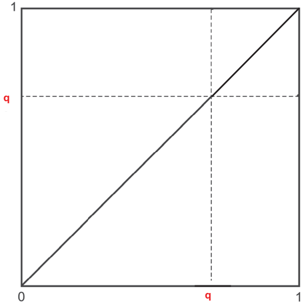
$$VaR_q(S^c) = VaR_q(X_1) + VaR_q(X_2) = 2q.$$



"Riskiest" Dependence: maximum VaR_q in 2 dims?

If X_1 and X_2 are U(0,1) comonotonic, then

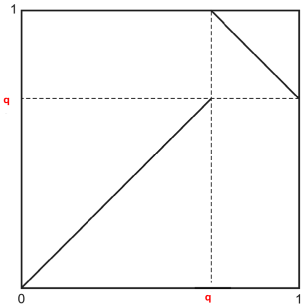
$$VaR_q(S^c) = VaR_q(X_1) + VaR_q(X_2) = 2q.$$



Note that $TVaR_q(S^c) = \frac{\int_q^1 2pdp}{1-q} = 1 + q.$

"Riskiest" Dependence: maximum VaR$_q$ in 2 dims

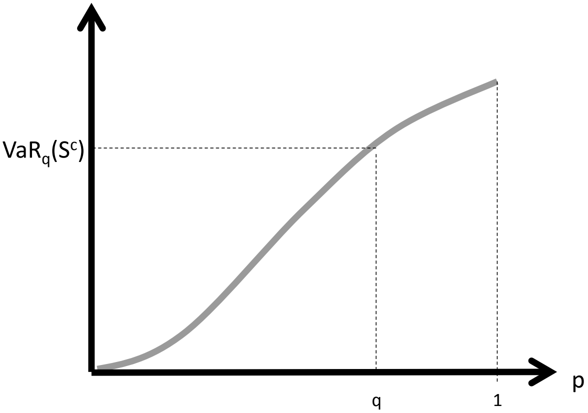
If X_1 and X_2 are U(0,1) and antimonotonic in the tail, then $VaR_q(S^*) = 1 + q$ (which is maximum possible).



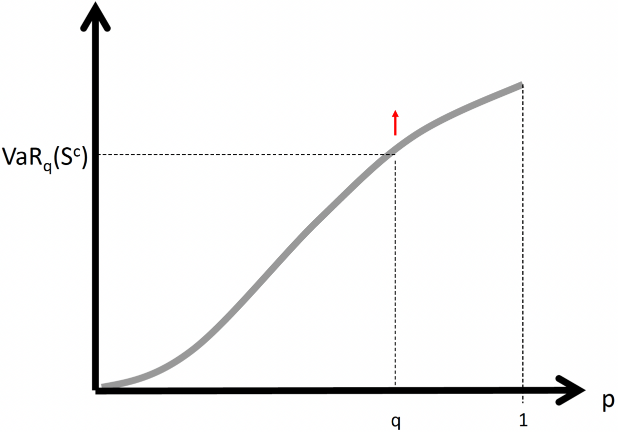
$$VaR_q(S^*) = 1 + q > VaR_q(S^c) = 2q$$

$\Rightarrow$ to maximize VaR$_q$, the idea is to change the comonotonic dependence such that the sum is constant in the tail

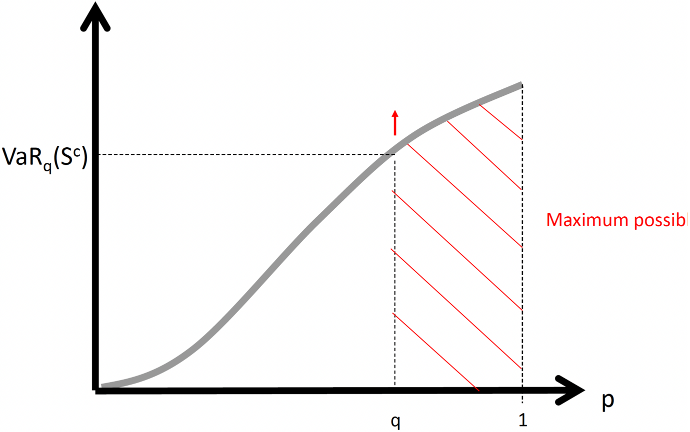
VaR at level q of the comonotonic sum w.r.t. q



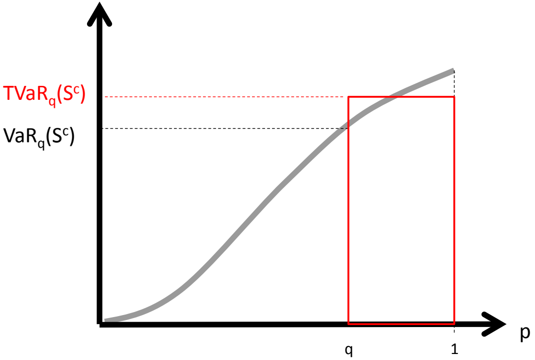
VaR at level q of the comonotonic sum w.r.t. q



VaR at level q of the comonotonic sum w.r.t. q

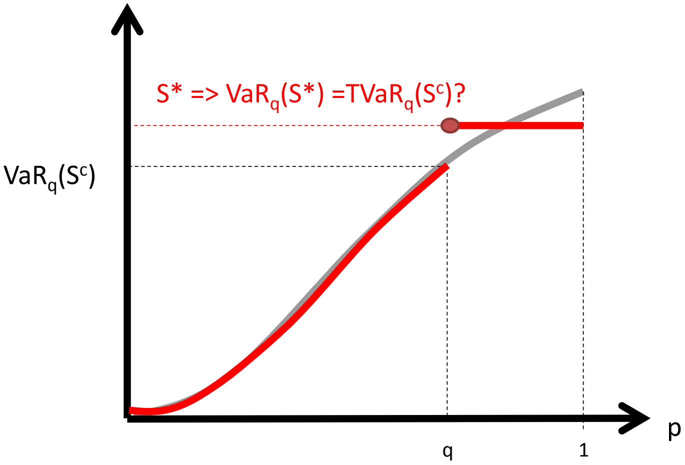


VaR at level q of the comonotonic sum w.r.t. q



where TVaR (Expected shortfall):$\text{TVaR}_q(X) = \dfrac{1}{1-q}\int_q^1 \text{VaR}_u(X)\mathrm{d}u$

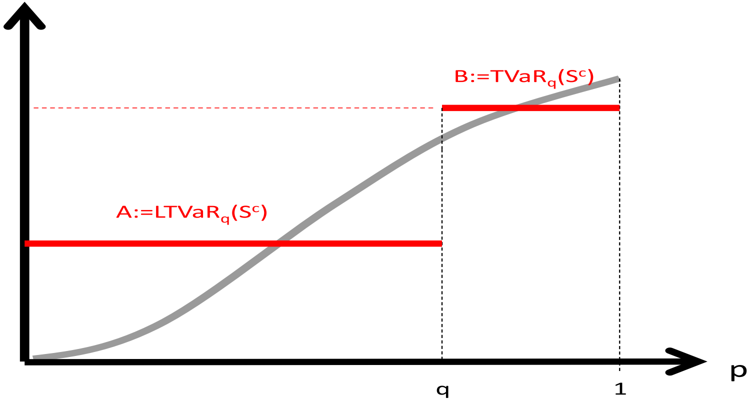
Riskiest Dependence Structure VaR at level q



Analytic expressions (not sharp)

Analytical Unconstrained Bounds with $X_j \sim F_j$

$$A = LTVaR_q(S^c) \leqslant \mathrm{VaR}_q[X_1 + X_2 + ... + X_n] \leqslant B = TVaR_q(S^c)$$



VaR Bounds with full dependence uncertainty

Approximate sharp bounds:

- Puccetti and Rüschendorf (2012): algorithm (RA) useful to approximate the minimum variance.
- Embrechts, Puccetti, Rüschendorf (2013): algorithm (RA) to find bounds on VaR

Illustration for the maximum VaR_q (1/3)

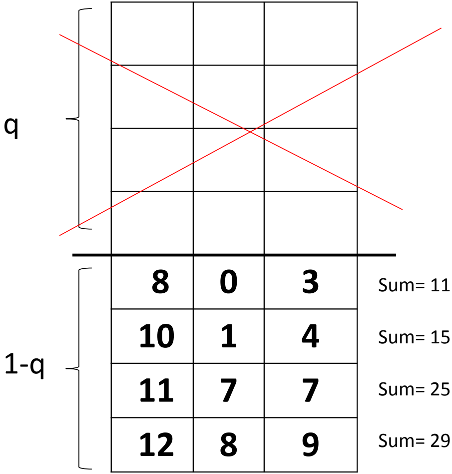


Illustration for the maximum VaR_q (2/3)

q

1-q

8	0	3	Sum= 11
10	1	4	Sum= 15
11	7	7	Sum= 25
12	8	9	Sum= 29

Rearrange **within** columns..to make the sums as constant as possible…

B=(11+15+25+29)/4=20

Illustration for the maximum VaR_q (3/3)

q				
1-q	**8**	**8**	**4**	Sum= 20
	10	**7**	**3**	Sum= 20
	12	**1**	**7**	Sum= 20
	11	**0**	**9**	Sum= 20

=B!

Numerical Results for VaR, 20 risks $N(0,1)$

When marginal distributions are given,

- What is the maximum Value-at-Risk?
- What is the minimum Value-at-Risk?
- A portfolio of 20 risks normally distributed N(0,1). Bounds on VaR_q (by the rearrangement algorithm applied on each tail)

q=95%	(-2.17 , 41.3)
q=99.95%	(-0.035 , 71.1)

- More examples in Embrechts, Puccetti, and Rüschendorf (2013): "Model uncertainty and VaR aggregation," *Journal of Banking and Finance*
- Very wide bounds
- All dependence information ignored

Idea: add information on dependence from a fitted model or from experts' opinions

Regulation challenge

The Basel Committee (2013) insists that **a desired objective of a Solvency framework concerns comparability**:

"*Two banks with portfolios having identical risk profiles apply the frameworks rules and arrive at the same amount of risk-weighted assets, and two banks with different risk profiles should produce risk numbers that are different proportionally to the differences in risk*"

Outline

Part 1: The Rearrangement Algorithm

- Minimizing variance of a sum with full dependence uncertainty
- Variance bounds

Part 2: Application to Model-Risk Assessment, e.g., Uncertainty on Value-at-Risk

- With 2 risks and full dependence uncertainty
- With d risks and full dependence uncertainty

Part 3: Adding information on dependence

- **Moment constraints**
- **Information on a subset...**

Part 4: Using the RA to infer dependence

- Add information about the sum of the risks
- Application to explain the correlation risk premium
- Application to multivariate option pricing

Part 5: Improved Rearrangement Algorithm

Part III

VaR Bounds with partial dependence uncertainty

VaR Bounds with Dependence Information...

Aggregation with dependence uncertainty: Example - Credit Risk

- **Marginals known**
- **Dependence fully unknown**

Consider a portfolio of 10,000 loans all having a default probability $p = 0.049$.

		Min VaR_q	Max VaR_q
q = 0.95		0%	98%
q = 0.995		4.4%	100%

Portfolio models are subject to significant model uncertainty (defaults are rare and correlated events).

Aggregation with dependence uncertainty: Example - Credit Risk

- **Marginals known**
- **Dependence fully unknown**

Consider a portfolio of 10,000 loans all having a default probability $p = 0.049$. *The default correlation is* $\rho = 0.0157$ *(for KMV).*

	KMV VaR_q	Min VaR_q	Max VaR_q
q = 0.95	10.1%	0%	98%
q = 0.995	15.1%	4.4%	100%

Portfolio models are subject to significant model uncertainty (defaults are rare and correlated events).
Using dependence information is crucial to try to get more "reasonable" bounds.

Adding dependence information

Finding minimum and maximum possible values for VaR of the credit portfolio loss, $S = \sum_{i=1}^{n} X_i$, given that

- known marginal distributions of the risks X_i.
- some **dependence information.**

Example 1: Variance constraint - with Rüschendorf and Vanduffel

$$M := \sup \mathrm{VaR}_q \left[X_1 + X_2 + ... + X_n\right],$$
$$\text{subject to } \; X_j \sim F_j, \mathrm{var}(X_1 + X_2 + ... + X_n) \leqslant s^2$$

Journal of Risk and Insurance (2017) and Chapter 6 from the book.

Example 2: Moments constraint - with Denuit, Rüschendorf, Vanduffel, Yao

$$M := \sup \mathrm{VaR}_q \left[X_1 + X_2 + ... + X_n\right],$$
$$\text{subject to } \; X_j \sim F_j, \mathrm{E}(X_1 + X_2 + ... + X_n)^k = c_k$$

European Journal of Finance (2015) and Chapter 6 from the book.

Adding dependence information

Example 3: with Rüschendorf, Vanduffel and Wang

$$M := \sup \mathrm{VaR}_q \left[X_1 + X_2 + ... + X_n\right],$$
$$\text{subject to } \ (X_j, Z) \sim H_j,$$

where Z is a factor.
Finance and Stochastics (2017) and Chapter 9 from the book.
Example 4: with Vanduffel

$$M := \sup \mathrm{VaR}_q \left[X_1 + X_2 + ... + X_n\right],$$

where the joint distribution is known on a subset of $\mathbb{R}^n$.
Journal of Banking and Finance (2015) and Chapter 7 from the book.

Examples

Example 1: variance constraint

$$M := \sup \mathrm{VaR}_q\left[X_1 + X_2 + ... + X_n\right],$$
$$\text{subject to } \; X_j \sim F_j, \mathrm{var}(X_1 + X_2 + ... + X_n) \leqslant s^2$$

Example 2: Moments constraint

$$M := \sup \mathrm{VaR}_q\left[X_1 + X_2 + ... + X_n\right],$$
$$\text{subject to } \; X_j \sim F_j, \mathrm{E}(X_1 + X_2 + ... + X_n)^k \leqslant c_k$$

for all k in 2,...,K

VaR bounds with moment constraints

- Without moment constraints, VaR bounds are attained if there exists a dependence among risks X_i such that

$$S = \begin{cases} A & \text{probability } q \\ B & \text{probability } 1-q \end{cases} \quad \text{a.s.}$$

- If the "distance" between A and B is too wide then improved bounds are obtained with

$$S^* = \begin{cases} a & \text{with probability } q \\ b & \text{with probability } 1-q \end{cases}$$

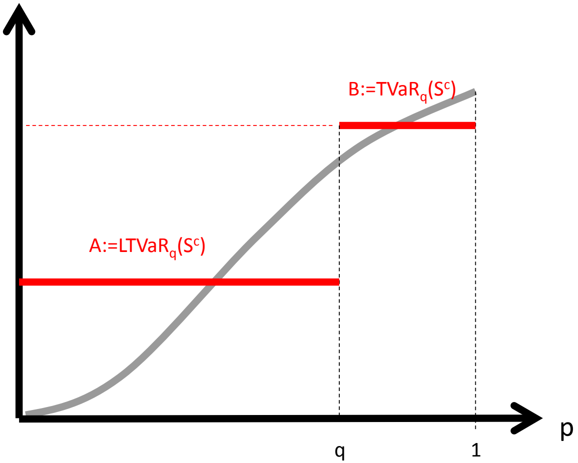
such that

$$\begin{cases} a^k q + b^k(1-q) \leqslant c_k \\ aq + b(1-q) = E[S] \end{cases}$$

in which a and b are "as distant as possible while satisfying all constraints" (for all k)

Unconstrained Bounds with $X_j \sim F_j$

$$A = LTVaR_q(S^c) \leqslant \mathrm{VaR}_q\left[X_1 + X_2 + ... + X_n\right] \leqslant B = TVaR_q(S^c)$$



Analytical result for variance constraint

A and B: unconstrained bounds on Value-at-Risk, $\mu = E[S]$.

Constrained Bounds with $X_j \sim F_j$ and variance $\leqslant s^2$

$$a = \max\left(A, \mu - s\sqrt{\frac{1-q}{q}}\right) \leqslant \mathrm{VaR}_q\left[X_1 + X_2 + \dots + X_n\right]$$

$$\leqslant b = \min\left(B,\ \mu + s\sqrt{\frac{q}{1-q}}\right)$$

• If the variance s^2 is not "too large" (i.e. when $s^2 \leqslant q(A-\mu)^2 + (1-q)(B-\mu)^2)$, then $b < B$.

• The "target" distribution for the sum: a two-point cdf that takes two values a and b. We can write

$$X_1 + X_2 + \dots + X_n - S = 0$$

and apply the standard RA.

Extended RA

q	...	...	...	-a	
	...	...	...	-a	
	...	...	...	-a	
	...	...	...	-a	
1-q	8	8	4	-b	
	10	7	3	-b	
	12	1	7	-b	
	11	0	9	-b	

Rearrange now within all columns such that all sums becomes close to zero

Bounds on VaR of sum of Pareto ($\theta = 3$) with $\rho = 0.15$

Panel A: Approximate sharp bounds obtained by the ERA

(m_d, M_d)		$n = 10$	$n = 100$
$\mathrm{VaR}_{95\%}$	$d = 1,000$	(4.118 ; 19.93)	(42.55 ; 174.0)
$\mathrm{VaR}_{99.5\%}$	$d = 1,000$	(4.868 ; 53.99)	(47.07 ; 457.6)

Bounds on VaR of sum of Pareto ($\theta = 3$) with $\rho = 0.15$

Panel A: Approximate sharp bounds obtained by the ERA

(m_d, M_d)		$n = 10$	$n = 100$
$\mathrm{VaR}_{95\%}$	$d = 1,000$	(4.118 ; 19.93)	(42.55 ; 174.0)
$\mathrm{VaR}_{99.5\%}$	$d = 1,000$	(4.868 ; 53.99)	(47.07 ; 457.6)

Panel B: Variance-constrained VaR bounds (theoretical)

(m_d, M_d)		$n = 10$	$n = 100$
$\mathrm{VaR}_{95\%}$,	$d = 1,000$	(4.100 ; 20.35)	(42.45 ; 175.9)
$\mathrm{VaR}_{99.5\%}$,	$d = 1,000$	(4.662 ; 54.87)	(47.06 ; 459.4)
$\mathrm{VaR}_{95\%}$,	$d = +\infty$	(4.037 ; 23.30)	(42.09 ; 200.3)
$\mathrm{VaR}_{99.5\%}$,	$d = +\infty$	(4.702 ; 64.22)	(47.56 ; 536.4)

Bounds on VaR of sum of Pareto ($\theta = 3$) with $\rho = 0.15$

Panel A: Approximate sharp bounds obtained by the ERA

(m_d, M_d)		$n = 10$	$n = 100$
$\mathrm{VaR}_{95\%}$	$d = 1,000$	(4.118 ; 19.93)	(42.55 ; 174.0)
$\mathrm{VaR}_{99.5\%}$	$d = 1,000$	(4.868 ; 53.99)	(47.07 ; 457.6)

Panel B: Variance-constrained VaR bounds (theoretical)

(m_d, M_d)		$n = 10$	$n = 100$
$\mathrm{VaR}_{95\%}$,	$d = 1,000$	(4.100 ; 20.35)	(42.45 ; 175.9)
$\mathrm{VaR}_{99.5\%}$,	$d = 1,000$	(4.662 ; 54.87)	(47.06 ; 459.4)
$\mathrm{VaR}_{95\%}$,	$d = +\infty$	(4.037 ; 23.30)	(42.09 ; 200.3)
$\mathrm{VaR}_{99.5\%}$,	$d = +\infty$	(4.702 ; 64.22)	(47.56 ; 536.4)

Panel C: Unconstrained VaR bounds (theoretical)

(m_d, M_d)		$n = 10$	$n = 100$
$\mathrm{VaR}_{95\%}$,	$d = 1,000$	(3.642 ; 29.05)	(36.42 ; 290.5)
$\mathrm{VaR}_{99.5\%}$,	$d = 1,000$	(4.615 ; 64.06)	(46.15 ; 640.6)
$\mathrm{VaR}_{95\%}$,	$d = +\infty$	(3.647 ; 30.72)	(36.47 ; 307.2)
$\mathrm{VaR}_{99.5\%}$,	$d = +\infty$	(4.635 ; 77.72)	(46.35 ; 777.2)

Corporate portfolio

- a corporate portfolio of a major European Bank.
- 4495 loans mainly to medium sized and large corporate clients
- total exposure (EAD) is 18642.7 (million Euros), and the top 10% of the portfolio (in terms of EAD) accounts for 70.1% of it.
- portfolio exhibits some heterogeneity.

Summary statistics of a corporate portfolio

	Minimum	Maximum	Average
Default probability	0.0001	0.15	0.0119
EAD	0	750.2	116.7
LGD	0	0.90	0.41

Comparison of Industry Models

VaRs of the corporate portfolio under different industry models

	$q =$	Comon.	KMV	Credit Risk$^+$	Beta
	95%	393.5	340.6	346.2	347.4
$\rho = 0.10$	99%	2374.1	539.4	513.4	520.2
	99.5%	5088.5	631.5	582.9	593.5

VaR bounds with Moments Information

Model risk assessment of the VaR of the corporate portfolio

(we use $\rho = 0.1$ to construct moments constraints)

$q =$	KMV	Comon.	Unconstrained	$K = 2$	$K = 3$
95%	340.6	393.3	(34.0 ; 2083.3)	(97.3 ; 614.8)	(100.9 ; 562.8)
99%	539.4	2374.1	(56.5 ; 6973.1)	(111.8 ; 1245)	(115.0 ; 941.2)
99.5%	631.5	5088.5	(89.4 ; 10120)	(114.9 ; 1709)	(117.6 ; 1177.8)

- Obs 1: Comparison with analytical bounds
- Obs 2: Significant bounds reduction with moments information
- Obs 3: Significant model risk

Objectives and Findings in Example 4:

Example 4: with Vanduffel

$$M := \sup \mathrm{VaR}_q \left[X_1 + X_2 + ... + X_n\right],$$

where the joint distribution is known on a subset of $\mathbb{R}^n$.
Journal of Banking and Finance (2015) and Chapter 7 from the book.

- Model uncertainty on the risk assessment of an aggregate portfolio: the sum of d dependent risks.
  - Given all information available in the market, what can we say about the maximum and minimum possible values of a given risk measure of a portfolio?

Objectives and Findings in Example 4:

Example 4: with Vanduffel

$$M := \sup \mathrm{VaR}_q \left[X_1 + X_2 + ... + X_n\right],$$

where the joint distribution is known on a subset of $\mathbb{R}^n$.
Journal of Banking and Finance (2015) and Chapter 7 from the book.

- Model uncertainty on the risk assessment of an aggregate portfolio: the sum of d dependent risks.
  - Given all information available in the market, what can we say about the maximum and minimum possible values of a given risk measure of a portfolio?
- Findings / Implications:
  - Current VaR based regulation is subject to high model risk, even if one knows the multivariate distribution "almost completely".

Illustration with 2 risks with marginals N(0,1)

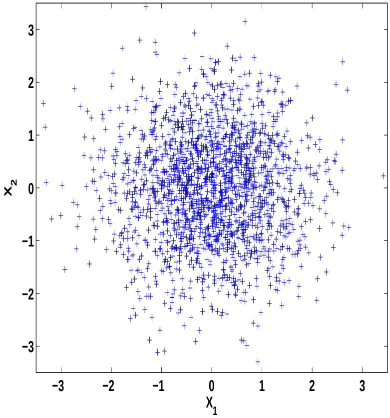
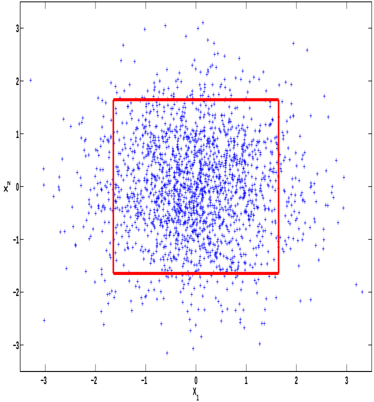


Illustration with 2 risks with marginals N(0,1)



$$\text{Assumption: Independence on } \mathcal{F} = \bigcap_{k=1}^{2} \{q_\beta \leqslant X_k \leqslant q_{1-\beta}\}.$$

Our assumptions on the cdf of $(X_1, X_2, ..., X_d)$

$\mathcal{F} \subset \mathbb{R}^d$ ("trusted" or "fixed" area)
$\mathcal{U} = \mathbb{R}^d \backslash \mathcal{F}$ ("untrusted").
We assume that we know:

(i) the marginal distribution F_i of X_i on $\mathbb{R}$ for $i = 1, 2, ..., d$,

(ii) the distribution of $(X_1, X_2, ..., X_d) \mid \{(X_1, X_2, ..., X_d) \in \mathcal{F}\}$.

(iii) $P\left((X_1, X_2, ..., X_d) \in \mathcal{F}\right)$.

- When only marginals are known: $\mathcal{U} = \mathbb{R}^d$ and $\mathcal{F} = \emptyset$.
- **Our Goal: Find bounds on $\rho(S) := \rho(X_1 + ... + X_d)$ when $(X_1, ..., X_d)$ satisfy (i), (ii) and (iii).**

Example:

$N = 8$ observations, $d = 3$ dimensions
and 3 observations trusted ($p_f = 3/8$).

$$\begin{bmatrix} 3 & 4 & 1 \\ 1 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 2 & 1 \\ 2 & 4 & 2 \\ 3 & 0 & 1 \\ 1 & 1 & 2 \\ 4 & 2 & 3 \end{bmatrix} \qquad S_N = \begin{bmatrix} 8 \\ 3 \\ 5 \\ 3 \\ 8 \\ 4 \\ 4 \\ 9 \end{bmatrix}$$

Example: $N = 8$, $d = 3$ with 3 observations trusted

Maximum variance:

$$M = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 4 & 2 \\ 0 & 2 & 1 \\ 4 & 3 & 3 \\ 3 & 2 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad S_N^f = \begin{bmatrix} 8 \\ 8 \\ 3 \end{bmatrix}, \quad S_N^u = \begin{bmatrix} 10 \\ 7 \\ 4 \\ 3 \\ 1 \end{bmatrix}$$

Minimum variance:

$$M = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 4 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 3 \\ 0 & 3 & 2 \\ 1 & 2 & 2 \\ 3 & 1 & 1 \\ 4 & 0 & 1 \end{bmatrix}, \quad S_N^f = \begin{bmatrix} 8 \\ 8 \\ 3 \end{bmatrix}, \quad S_N^u = \begin{bmatrix} 5 \\ 5 \\ 5 \\ 5 \\ 5 \end{bmatrix}$$

Example $d = 20$ risks N(0,1)

- $(X_1, ..., X_{20})$ independent N(0,1) on

$$\mathcal{F} := [q_\beta, q_{1-\beta}]^d \subset \mathbb{R}^d \qquad p_f = P\left((X_1, ..., X_{20}) \in \mathcal{F}\right)$$

(for some $\beta \leqslant 50\%$) where q_γ: γ-quantile of N(0,1).

- $\beta = 0\%$: no uncertainty (20 independent N(0,1)).
- $\beta = 50\%$: full uncertainty.

$\mathcal{F} = [q_\beta, q_{1-\beta}]^d$	$\mathcal{U} = \emptyset$ $\beta = 0\%$			$\mathcal{U} = \mathbb{R}^d$ $\beta = 50\%$
$\rho = 0$	4.47			(0 , 20)

Example $d = 20$ risks N(0,1)

- $(X_1, ..., X_{20})$ independent N(0,1) on

$$\mathcal{F} := [q_\beta, q_{1-\beta}]^d \subset \mathbb{R}^d \qquad p_f = P\left((X_1, ..., X_{20}) \in \mathcal{F}\right)$$

(for some $\beta \leqslant 50\%$) where q_γ: γ-quantile of N(0,1)

- $\beta = 0\%$: no uncertainty (20 independent N(0,1))
- $\beta = 50\%$: full uncertainty

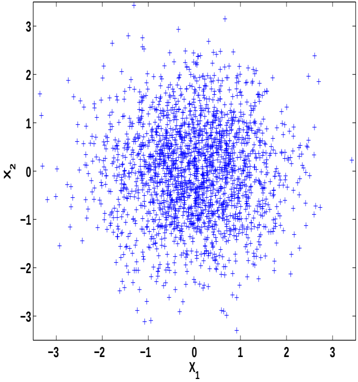
$\mathcal{F} = [q_\beta, q_{1-\beta}]^d$	$\mathcal{U} = \emptyset$ $\beta = 0\%$	$p_f \approx 98\%$ $\beta = 0.05\%$	$p_f \approx 82\%$ $\beta = 0.5\%$	$\mathcal{U} = \mathbb{R}^d$ $\beta = 50\%$
$\rho = 0$	4.47	(4.4 , 5.65)	(3.89 , 10.6)	(0 , 20)

Model risk on the volatility of a portfolio is reduced a lot by incorporating information on dependence!

Information on the joint distribution

- Can come from a fitted model
- Can come from experts' opinions
- Dependence "known" on specific scenarios

Illustration with marginals N(0,1)



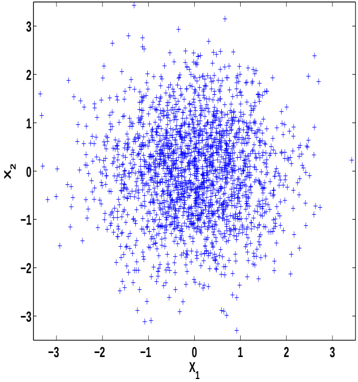
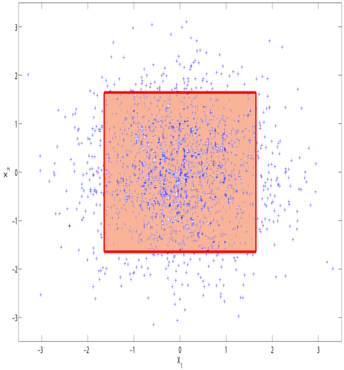
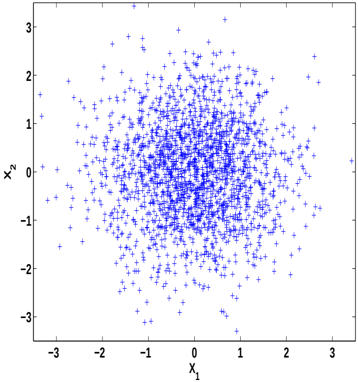


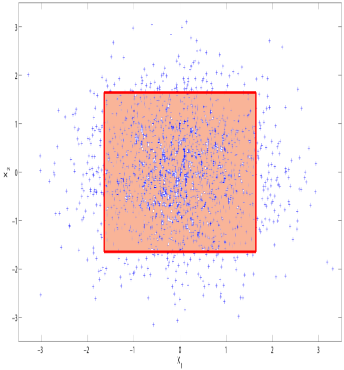
Illustration with marginals N(0,1)

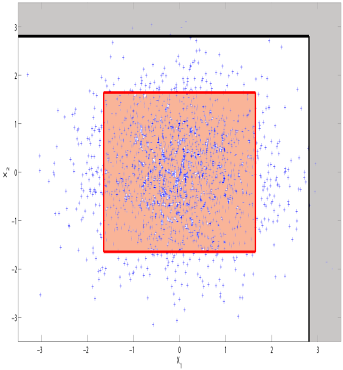




$$\mathcal{F}_1 = \bigcap_{k=1}^{2} \{q_\beta \leqslant X_k \leqslant q_{1-\beta}\}$$

Illustration with marginals N(0,1)

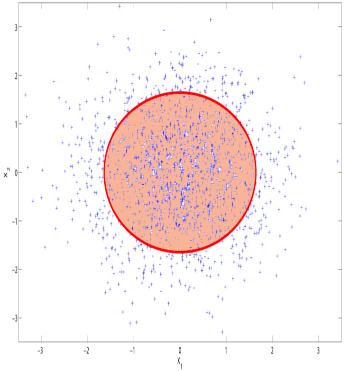




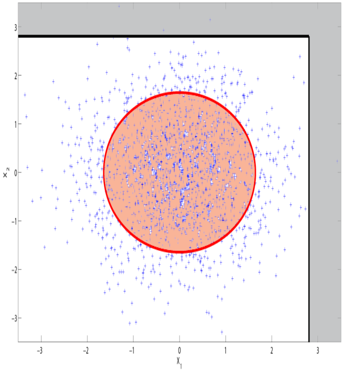
$$\mathcal{F}_1 = \bigcap_{k=1}^{2} \{q_\beta \leqslant X_k \leqslant q_{1-\beta}\}$$

$$\mathcal{F} = \bigcup_{k=1}^{2} \{X_k > q_p\} \bigcup \mathcal{F}_1$$

Illustration with marginals N(0,1)



$\mathcal{F}_1$ =contour of MVN at β



$$\mathcal{F} = \bigcup_{k=1}^{2} \{X_k > q_p\} \bigcup \mathcal{F}_1$$

Comments on bounds on variance with partial information

- Model risk for variance of a portfolio of risks with given marginals but partially known dependence.
- Same method applies to TVaR (expected Shortfall) or any risk measure that satisfies convex order (but not for Value-at-Risk).

Adding information for VaR bounds

Information on a subset

VaR bounds when the joint distribution of $(X_1, X_2, ..., X_n)$ is known on a subset of the sample space.

Our assumptions on the cdf of $(X_1, X_2, ..., X_n)$

$\mathcal{F} \subset \mathbb{R}^n$ ("trusted" or "fixed" area)

$\mathcal{U} = \mathbb{R}^n \backslash \mathcal{F}$ ("untrusted").

We assume that we know:

(i) the marginal distribution F_i of X_i on $\mathbb{R}$ for $i = 1, 2, ..., n$,

(ii) the distribution of $(X_1, X_2, ..., X_n) \mid \{(X_1, X_2, ..., X_n) \in \mathcal{F}\}$.

(iii) $P((X_1, X_2, ..., X_n) \in \mathcal{F})$

- **Goal: Find bounds on $\mathbf{VaR}_q(S) := \mathbf{VaR}_q(X_1 + ... + X_n)$ when $(X_1, ..., X_n)$ satisfy (i), (ii) and (iii).**

Numerical Results, 20 correlated $N(0,1)$ on $\mathcal{F} = [q_\beta, q_{1-\beta}]^n$

$\mathcal{F}$	$\mathcal{U} = \emptyset$ $\beta = 0\%$			$\mathcal{U} = \mathbb{R}^n$ $\beta = 50\%$
q=95%	12.5			(-2.17 , 41.3)
q=99.5%	19.6			(-0.29 , 57.8)
q=99.95%	25.1			(-0.035 , 71.1)

- $\mathcal{U} = \emptyset$: 20 correlated standard normal variables ($\rho = 0.1$).

$$\text{VaR}_{95\%} = 12.5 \quad \text{VaR}_{99.5\%} = 19.6 \quad \text{VaR}_{99.95\%} = 25.1$$

Numerical Results, 20 correlated $N(0,1)$ on $\mathcal{F} = [q_\beta, q_{1-\beta}]^n$

	$\mathcal{U} = \emptyset$ $\beta = 0\%$	$p_f \approx 98\%$ $\beta = 0.05\%$	$p_f \approx 82\%$ $\beta = 0.5\%$	$\mathcal{U} = \mathbb{R}^n$ $\beta = 50\%$
q=95%	12.5	(12.2 , 13.3)	(10.7 , 27.7)	(-2.17 , 41.3)
q=99.5%	19.6	(19.1 , **31.4**)	(16.9 , **57.8**)	(-0.29 , **57.8**)
q=99.95%	25.1	(24.2 , **71.1**)	(21.5 , **71.1**)	(-0.035 , **71.1**)

- $\mathcal{U} = \emptyset$: 20 correlated standard normal variables ($\rho = 0.1$).

$$\text{VaR}_{95\%} = 12.5 \quad \text{VaR}_{99.5\%} = 19.6 \quad \text{VaR}_{99.95\%} = 25.1$$

- **The risk for an underestimation of VaR is increasing in the probability level used to assess the VaR.**
- **For VaR at high probability levels ($q = 99.95\%$), despite all the added information on dependence, the bounds are still wide!**

With Pareto risks

Consider $d = 20$ risks distributed as Pareto with parameter $\theta = 3$.

- Assume we trust the independence conditional on being in $\mathcal{F}_1$

$$\mathcal{F}_1 = \bigcap_{k=1}^{d} \{q_\beta \leqslant X_k \leqslant q_{1-\beta}\}$$

where $q_\beta = (1-\beta)^{-1/\theta} - 1$.

$\mathcal{F}_1$	$\mathcal{U} = \emptyset$ $\beta = 0\%$	$\beta = 0.05\%$	$\beta = 0.5\%$	$\mathcal{U} = \mathbb{R}^d$ $\beta = 0.5$
α=95%	16.6	(16 , 18.4)	(13.8 , 37.4)	(7.29 , 61.4)
α=99.95%	43.5	(26.5 , 359)	(20.5 , 359)	(9.83 , 359)

Incorporating Expert's Judgements

Consider $d = 20$ risks distributed as Pareto $\theta = 3$.

- Assume comonotonicity conditional on being in $\mathcal{F}_2$

$$\mathcal{F}_2 = \bigcup_{k=1}^{d} \{X_k > q_q\}$$

Comonotonic estimates of Value-at-Risk

$VaR_{95\%}(S^c) = 34.29$, $VaR_{99.95\%}(S^c) = 231.98$

$\mathcal{F}_2$	$\mathcal{U} = \emptyset$ (Model)	$q = 99.5\%$	$q = 99.9\%$	$q = 99.95\%$
α=95%	16.6	(8.35 , 50)	(7.47 , 56.7)	(7.38 , 58.3)
α=99.95%	43.5	(232 , 232)	(232 , 232)	(180 , 232)

Comparison

Independence within a rectangle $\mathcal{F}_1 = \bigcap_{k=1}^{d} \{q_\beta \leqslant X_k \leqslant q_{1-\beta}\}$

$\mathcal{F}_1$	$\mathcal{U} = \emptyset$ $\beta = 0\%$	$\beta = 0.05\%$	$\beta = 0.5\%$	$\mathcal{U} = \mathbb{R}^d$ $\beta = 0.5$
α=95%	16.6	(16 , 18.4)	(13.8 , 37.4)	(7.29 , 61.4)
α=99.95%	43.5	(26.5 , 359)	(20.5 , 359)	(9.83 , 359)

Comonotonicity when one of the risks is large
$\mathcal{F}_2 = \bigcup_{k=1}^{d} \{X_k > q\}$

$\mathcal{F}_2$	$\mathcal{U} = \emptyset$ (Model)	$q = 99.5\%$	$q = 99.9\%$	$p = 99.95\%$
α=95%	16.6	(8.35 , 50)	(7.47 , 56.7)	(7.38 , 58.3)
α=99.95%	43.5	(232 , 232)	(232 , 232)	(180 , 232)

With Pareto risks

Consider $d = 20$ risks distributed as Pareto with parameter $\theta = 3$.

- Assume we trust the independence conditional on being in $\mathcal{F}_1$

$$\mathcal{F}_1 = \bigcap_{k=1}^{d} \{q_\beta \leqslant X_k \leqslant q_{1-\beta}\}$$

where $q_\beta = (1-\beta)^{-1/\theta} - 1$.

Comonotonic estimates of Value-at-Risk

$VaR_{95\%}(S^c) \approx 34.3$, $VaR_{99.95\%}(S^c) \approx 232$

$\mathcal{F}_1$	$\mathcal{U} = \emptyset$ $\beta = 0\%$	$\beta = 0.05\%$	$\beta = 0.5\%$	$\mathcal{U} = \mathbb{R}^d$ $\beta = 0.5$
α=95%	16.6	(16 , 18.4)	(13.8 , 37.4)	(7.29 , 61.4)
α=99.95%	43.5	(26.5 , 359)	(20.5 , 359)	(9.83 , 359)

Incorporating Expert's Judgements

Consider $d = 20$ risks distributed as Pareto $\theta = 3$.

- Assume comonotonicity conditional on being in $\mathcal{F}_2$

$$\mathcal{F}_2 = \bigcup_{k=1}^{d} \{X_k > q_p\}$$

Comonotonic estimates of Value-at-Risk
$VaR_{95\%}(S^c) \approx 34.3$, $VaR_{99.95\%}(S^c) \approx 232$

$\mathcal{F}_2$	$\mathcal{U} = \emptyset$ (Model)	$p = 99.5\%$	$p = 99.9\%$	$p = 99.95\%$
α=95%	16.6	(8.35 , 50)	(7.47 , 56.7)	(7.38 , 58.3)
α=99.95%	43.5	(232 , 232)	(232 , 232)	(180 , 232)

Comparison

Independence within a rectangle $\mathcal{F}_1 = \bigcap_{k=1}^{d} \{q_\beta \leqslant X_k \leqslant q_{1-\beta}\}$

$\mathcal{F}_1$	$\mathcal{U} = \emptyset$ $\beta = 0\%$	$\beta = 0.05\%$	$\beta = 0.5\%$	$\mathcal{U} = \mathbb{R}^d$ $\beta = 0.5$
α=95%	16.6	(16 , 18.4)	(13.8 , 37.4)	(7.29 , 61.4)
α=99.95%	43.5	(26.5 , 359)	(20.5 , 359)	(9.83 , 359)

Comonotonicity when one of the risks is large $\mathcal{F}_2 = \bigcup_{k=1}^{d} \{X_k > q_p\}$

$\mathcal{F}_2$	$\mathcal{U} = \emptyset$ (Model)	$p = 99.5\%$	$p = 99.9\%$	$p = 99.95\%$
α=95%	16.6	(8.35 , 50)	(7.47 , 56.7)	(7.38 , 58.3)
α=99.95%	43.5	(232 , 232)	(232 , 232)	(180 , 232)

Some Remaining Challenges

Challenges:

- Choosing the trusted area $\mathcal{F}$
- N too small: possible to improve the efficiency of the algorithm by re-discretizing using the fitted marginal $\hat{f}_i$.
- Possible to amplify the tails of the marginals

Conclusions

- Maximum Value-at-Risk is not caused by comonotonicity.
- Maximum Value-at-Risk is achieved when the variance is *minimum* in the tail. The RA is then used in the tails only.
- Bounds on Value-at-Risk at high confidence level stay wide even if the multivariate dependence is known in 98% of the space!

- Assess model risk with partial information and given marginals
- Design algorithms for bounds on variance, TVaR and VaR and many more risk measures.
- A regulation challenge...

Outline

Part 1: The Rearrangement Algorithm

- Minimizing variance of a sum with full dependence uncertainty
- Variance bounds

Part 2: Application to Model-Risk Assessment, e.g., Uncertainty on Value-at-Risk

- With 2 risks and full dependence uncertainty
- With *d* risks and full dependence uncertainty

Part 3: Adding information on dependence

- Moment constraints
- Information on a subset...

Part 4: Using the RA to infer dependence

- **Add information about the sum of the risks**
- **Application to explain the correlation risk premium**
- **Application to multivariate option pricing**

Part 5: Improved Rearrangement Algorithm

Part IV-A

Use of the Rearrangement Algorithm when one knows marginals and information on the sum to find a possible dependence...

Method: Block RA to infer the dependence

- Inputs:
  - $X_1 \sim F_1$, ... $X_d \sim F_d$
  - $X_1 + ... + X_d \sim G$
- Method (use the fact that $X_1 + X_2 + .. + X_n - Sum = 0$):
  - Matrix m rows (discretization step) by $n = d + 1$ columns.
  - In each of the first d columns

$$F_j^{-1}\left(\frac{i}{m+1}\right), \qquad i = 1, 2, ..., m$$

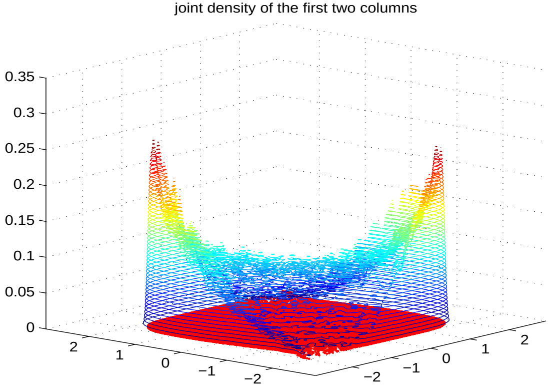
  - In the last column

$$-G^{-1}\left(\frac{i}{m+1}\right), \qquad i = 1, 2, ..., m$$

  - Apply the Block RA on the full matrix
- Output: Extract the d first columns, and they describe a discrete copula that is consistent with the cdfs of the risks and of their sum.

Using the Block RA to infer the dependence

- find the dependence between two uniformly distributed variables that makes the distribution of the sum of two uniform statistically indistinguishable from a normal distribution



How can it be useful?

- When we have information on the distribution of the sum, of linear combinations and of the marginal distributions?
- Infer the dependence between business lines assuming that you have access to individual performance of business lines and of the aggregate performance of the company. In this case you typically are unable to observe the joint distribution.
- When you have information on options on an index and options on its components:
  - Study the properties of the dependence in the risk neutral world of the 9 sectors comprising the SP 500 index
  - Infer a possible model to price basket options when you know a few basket option prices and you want to give a quote of a basket option on an underlying that is a basket with different weights

Rearrangement Algorithm and Maximum Entropy, **Annals of Operational Research**, 2018 with Oleg Bondarenko and Steven Vanduffel.

A Model-free Approach to Multivariate Option Pricing , **Review of Derivatives Research**, 2021 with Oleg Bondarenko and Steven Vanduffel.

Option Implied Dependence and Correlation Risk Premium , **Journal of Financial and Quantitative Analysis**, 2023 with Oleg Bondarenko.

Algorithm to infer dependence

Inputs

- Option prices written on X_i for $i = 1, 2, ..., d$
- Basket option prices on the index S

Output

- A joint distribution of $(X_1, X_2, ..., X_d)$
  - compatible with inputs
  - that maximizes "entropy"

How?

Algorithm to infer dependence

Inputs

- Option prices written on X_i for $i = 1, 2, ..., d$
- Basket option prices on the index S

Output

- A joint distribution of $(X_1, X_2, ..., X_d)$
  - compatible with inputs
  - that maximizes "entropy"

How? Using the Rearrangement Algorithm...

Inferring Dependence

- **Inputs:** d r.v. $X_1 \sim F_1$, ..., $X_d \sim F_d$ and their sum $S \sim F_S$.
- Sample X_j and S into n equiprobable values, arranged in an $n \times (d+1)$ matrix ($s_i = F_S^{-1}((i-0.5)/n)$):

$$[X_1, \ldots, X_d, -S] = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1d} & -s_1 \\ x_{21} & x_{22} & \ldots & x_{2d} & -s_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{nd} & -s_n \end{bmatrix}.$$

- Apply BRA on $[X_1, \ldots, X_d, -S]$.
- Row sums of the rearranged matrix are close to zero, i.e. a **compatible dependence** has been found.

Properties of the output dependence?

- We run BRA K times to obtain different solutions $\mathbf{X}^{(k)}$ $(k = 1, ..., K)$. Let $R^{(k)}$ denote the correlation matrix of $\mathbf{X}^{(k)}$:

$$R^{(k)} := \begin{bmatrix} 1 & \rho_{12}^{(k)} & \dots & \rho_{1d}^{(k)} \\ \rho_{21}^{(k)} & 1 & \dots & \rho_{2d}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{d1}^{(k)} & \rho_{d2}^{(k)} & \dots & 1 \end{bmatrix}.$$

- We compute $\Delta^{(k)} := \det[R^{(k)}]$

Possible correlation matrices

- Standard deviations $\sigma_1, \ldots, \sigma_d$ and σ_S are fixed (since F_1, ..., F_d and F_S are given) and related by

$$\sigma_S^2 = \sum_{i=1}^{d} \sigma_i^2 + 2 \sum_{i=1}^{d-1} \sum_{j>i} \sigma_i \sigma_j \rho_{ij},$$

- Hence for all possible dependences, the average (implied) correlation ρ^{imp} is constant,

$$\rho^{imp} = \frac{\sum_{i=1}^{d-1} \sum_{j>i} \sigma_i \sigma_j \rho_{ij}}{\sum_{i=1}^{d-1} \sum_{j>i} \sigma_i \sigma_j}.$$

- Let $\mathcal{C}(\rho^{imp})$ denote the set of correlation matrices R with average correlation ρ^{imp}.

Constrained set $\mathcal{C}(\rho^{imp})$, $d = 3$

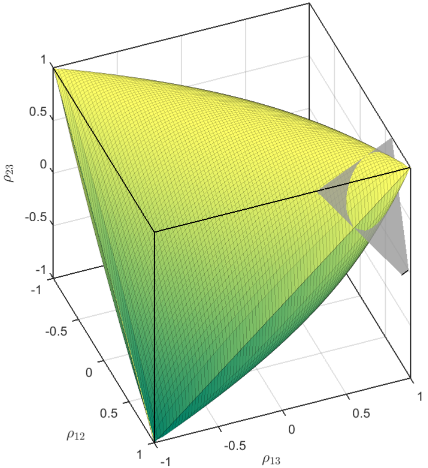


Figure: The set of correlation matrices $(\rho_{12}, \rho_{12}, \rho_{23})$ is intersected by the plane $\sigma_1\sigma_2(\rho_{12} - \rho^{imp}) + \sigma_1\sigma_3(\rho_{13} - \rho^{imp}) + \sigma_2\sigma_3(\rho_{23} - \rho^{imp}) = 0$.

Maximum Determinant and Maximum Entropy

- Entropy refers to disorder of a system, Shannon (1948).
- Let f be the density of a multivariate distribution of $(X_1, ..., X_d)$, then the entropy is defined as

$$H(X_1, ..., X_d) = -\mathrm{E}(\log(f(X_1, ..., X_d))).$$

Proposition: Maximum entropy for a given correlation matrix

The entropy of the multivariate distribution of a random vector $(X_1, ..., X_d)$ and invertible correlation matrix R satisfies

$$H(X_1, .., X_d) \leqslant \frac{d}{2}(1 + \ln(2\pi)) + \frac{1}{2}\sum_{i=1}^{d} \ln\left(\sigma_i^2\right) + \frac{1}{2}\ln\left(\det(R)\right)$$

where the equality holds *iff* $(X_1, ..., X_d)$ is multivariate Gaussian.

Maximum Determinant and Maximum Entropy

- Entropy refers to disorder of a system, Shannon (1948).
- Let f be the density of a multivariate distribution of $(X_1, ..., X_d)$, then the entropy is defined as

$$H(X_1, ..., X_d) = -\mathrm{E}(\log(f(X_1, ..., X_d))).$$

Proposition: Maximum entropy for a given correlation matrix

The entropy of the multivariate distribution of a random vector $(X_1, ..., X_d)$ and invertible correlation matrix R satisfies

$$H(X_1, .., X_d) \leqslant \frac{d}{2}(1 + \ln(2\pi)) + \frac{1}{2}\sum_{i=1}^{d} \ln\left(\sigma_i^2\right) + \frac{1}{2}\ln\left(\det(R)\right)$$

where the equality holds *iff* $(X_1, ..., X_d)$ is multivariate Gaussian.

- **We are interested in** $\Delta_M := \max_{R \in \mathcal{C}(r)} \det[R]$ **and the correlation matrix** R_M **that achieves it.**

Gaussian Case

- Gaussian margins $X_i \sim N[0, \sigma_i^2]$, $i = 1, ..., d$, and Gaussian sum $S \sim N[0, \sigma_S^2]$.
- Standard deviations σ_i are linearly decreasing from 1 to $1/d$.
- Set σ_S such that $\rho_{imp} = 0.8$.
- Number of components d ranges from 3 to 10.
- Discretization level n from 1,000 to 10,000.
- Run BRA $K = 500$ times.
- For each run k, correlation matrix $R^{(k)}$ and its determinant $\Delta^{(k)}$
- Compare with correlation matrix R_M and its maximum determinant $\Delta_M(\rho^{imp})$

Stability of BRA

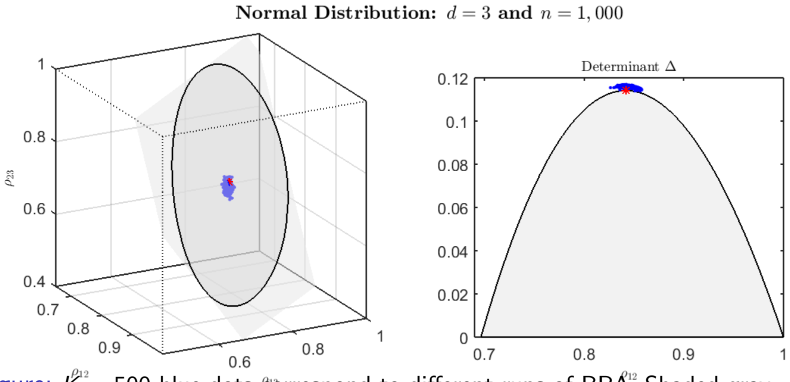


Figure: $K = 500$ blue dots correspond to different runs of BRA. Shaded gray area is constrained set $\mathcal{C}(\rho^{imp})$; red star is maximal correlation matrix R_M. Left panel shows realizations of correlations ρ_{12}, ρ_{13}, and ρ_{23}. Right panel shows the relation of Δ versus ρ_{12}.

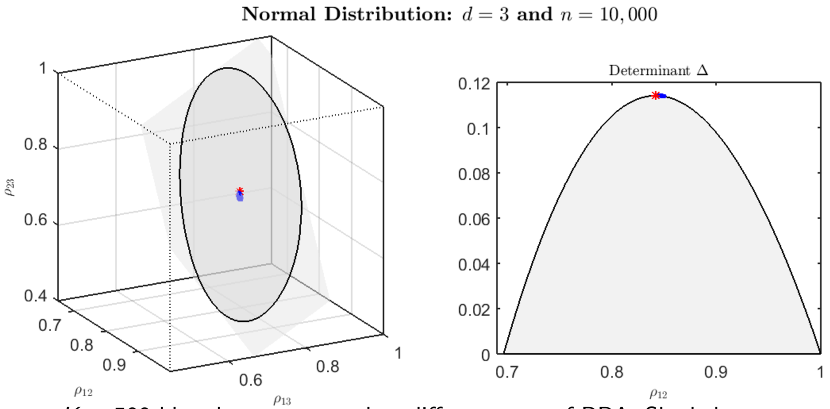


Figure: $K = 500$ blue dots correspond to different runs of BRA. Shaded gray area is constrained set $\mathcal{C}(\rho^{imp})$; red star is maximal correlation matrix R_M. Left panel shows realizations of correlations ρ_{12}, ρ_{13}, and ρ_{23}. Right panel shows the relation of Δ versus ρ_{12}.

Robustness Check

1. Robustness to Initial Conditions (supplement)
    - Start from a particular candidate solution
    - Introduce small noise, by randomly swapping 0.2% of rows:
    - Check where $K = 500$ runs of BRA converge.
2. Robustness to Distributional assumptions - Skewed distributions? (supplement)

Part IV-B

Inferring Dependence: Applications to Options

Application to Implied Correlation Premium

1. Example in 2 dimensions with specified distributions for two variables and for their sum
2. Study of the dependence among the 9 sectors of the SP 500 index
   - Extracting a compatible risk neutral 10-dimensional distribution among the 9 sectors and the SP 500 that is consistent with all option prices written on these 10 underlying variables
   - Study some of its properties
   - New insights about the correlation risk premium

Illustration when X_1, X_2 are $N(0, \sigma_i)$ and S is $N(0, \sigma_S)$ such that implied correlation is 0.

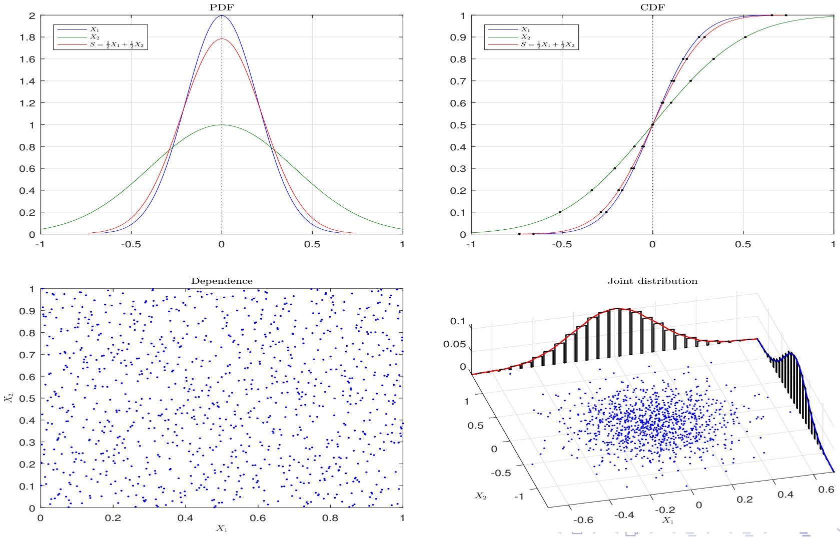


Illustration when X_1, X_2 are $N(0, \sigma_i)$ and S is $N(0, \sigma_S)$ such that implied correlation is 0.97.

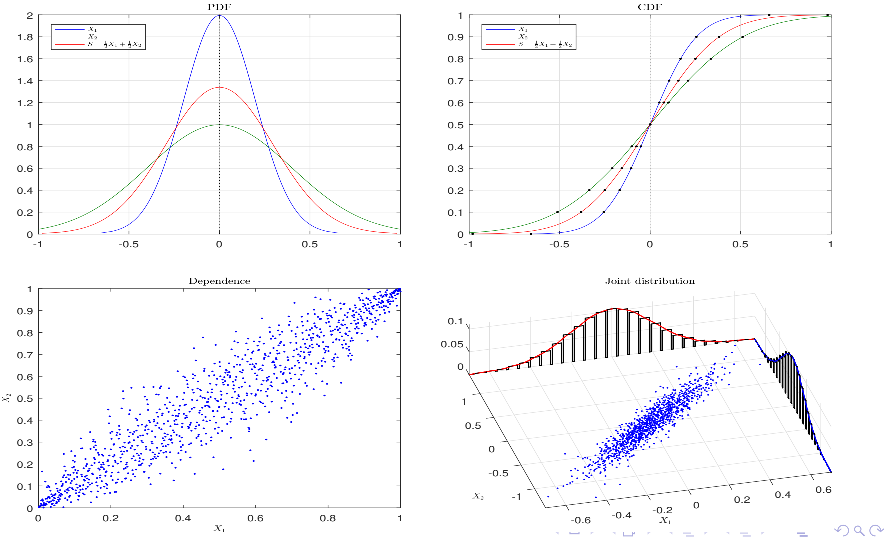
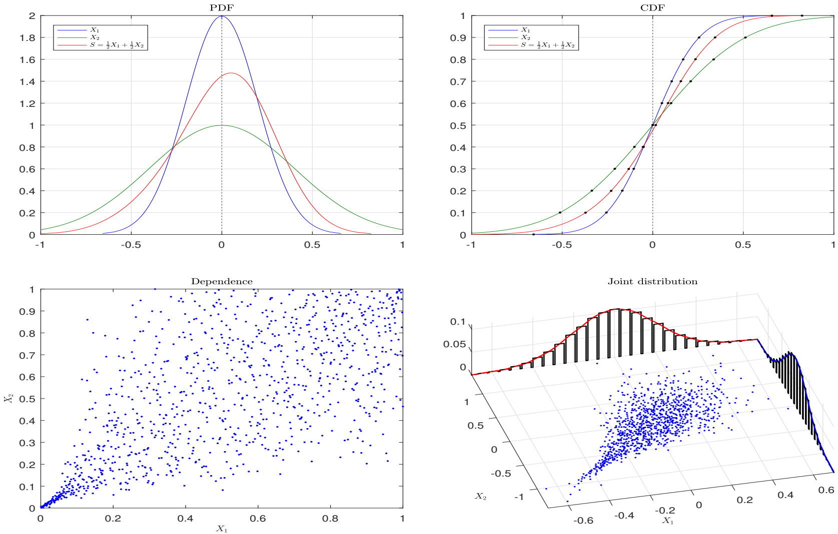


Illustration when X_1, X_2 are $N(0, \sigma_i)$ and S is skewed.



Empirical Application – S&P 500 Sectors

- SPDR ETFs, S&P 500 Index and its 9 sectors:

Description	Ticker	Abbreviation
SPDR S&P 500 ETF Trust	SPY	spx
Consumer Discretionary Sector SPDR Fund	XLY	cdi
Consumer Staples Sector SPDR Fund	XLP	cst
Energy Sector SPDR Fund	XLE	ene
Financial Sector SPDR Fund	XLF	fin
Health Care Sector SPDR Fund	XLV	hea
Industrial Sector SPDR Fund	XLI	ind
Materials Sector SPDR Fund	XLB	mat
Technology Sector SPDR Fund	XLK	tec
Utilities Sector SPDR Fund	XLU	uti

- **9 sectors** that do not overlap and that cover entire S&P 500
- Daily option data from CBOE
- Sample: 04/2007 - 09/2017

S&P 500 Sectors

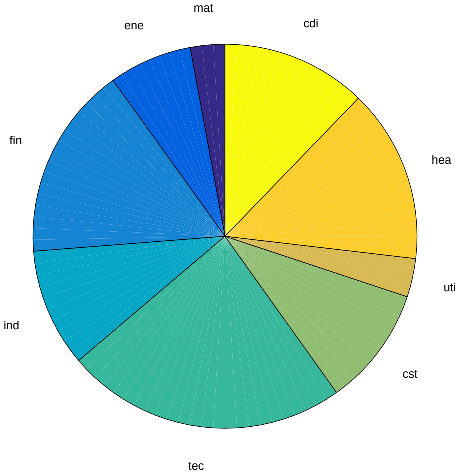


Figure: Sector weights in September 2016.

S&P 500 Sectors

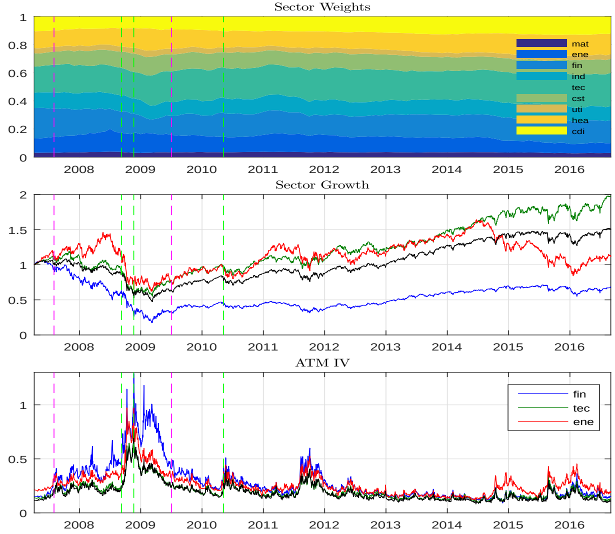


Figure: Sector weights over time. Pink vertical lines indicate Financial crisis. Green vertical lines: 08-Sep-08, 20-Nov-08, and 06-May-10.

Implementation Details

- Daily frequency, τ is at least 30 days, or closest available
- Estimate RNDs for S and each X_j from traded options on SPY and $d = 9$ Sector ETFs
- Estimate RNDs nonparametrically with **Positive Convolution Approximation (PCA)**, Bondarenko (2003)
- Discretize each distribution into $n = 1000$ equiprobable returns and arrange them in $n \times (d+1)$ matrix:

$$\mathbf{M} = [X_1, \dots, X_d, -S] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1d} & -s_1 \\ x_{21} & x_{22} & \dots & x_{2d} & -s_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nd} & -s_n \end{bmatrix}.$$

- Apply BRA on matrix $\mathbf{M}$ to infer dependence structure

Implementation Details

- Compute **various dependence statistics**:
  - **Pairwise correlations** and their value-weighted average
  - **Correlations conditional** on various events $\rho(R_i, R_j \mid Scenario)$, which can depend on the aggregate market or other factors:
    - localized or "corridor" correlation: $Scenario = \{q_1 \leqslant R_S \leqslant q_2\}$ for some quantiles q_1, q_2
    - **Down and Up correlations**: Let R_S^M be the median of R_S

$$\rho_{i,S}^{d,\mathbb{Q}} = \text{corr}^{\mathbb{Q}}(R_i, R_S \mid R_S \leqslant R_S^M)$$

$$\rho_{i,S}^{u,\mathbb{Q}} = \text{corr}^{\mathbb{Q}}(R_i, R_S \mid R_S > R_S^M),$$

  - Also **Spearman's rho** – not affected by changes in marginal distributions (not sensitive to changes in volatility)

$$\text{Spearman's rho}(R_i, R_j) = \rho(F_i(R_i), F_j(R_j))$$

  - Other **tail indices**

Selective Date: 08-Sep-2008

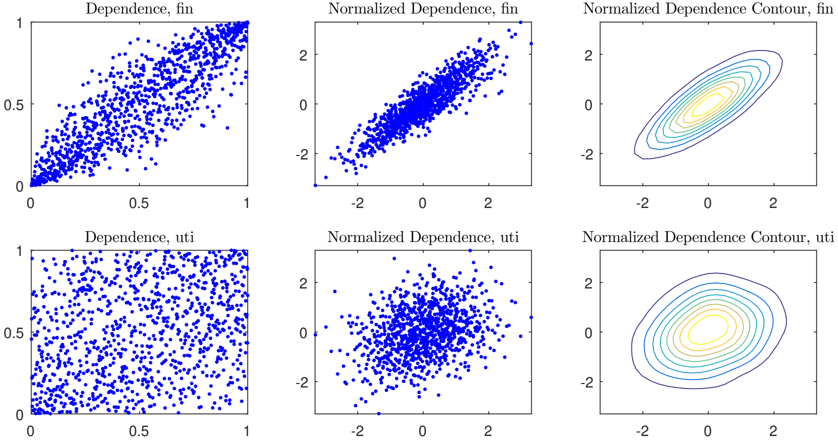


Figure: Implied Dependence.

Selective Date: 08-Sep-2008

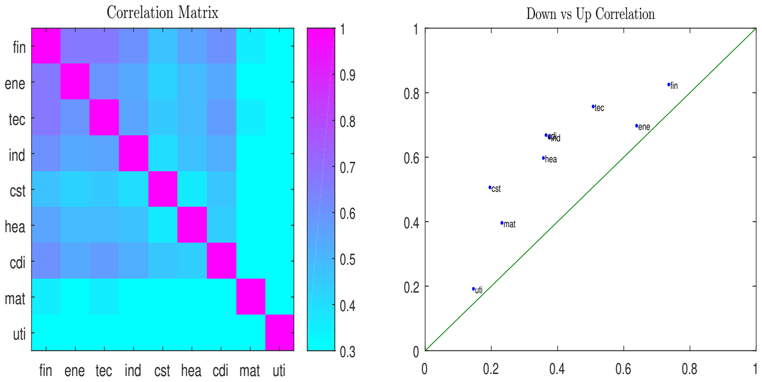


Figure: Implied Correlations.

Selective Date: 20-Nov-2008

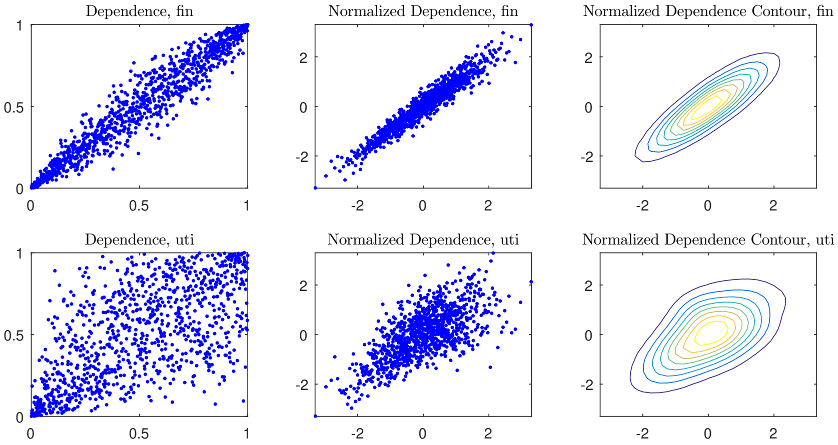


Figure: Implied Dependence.

Selective Date: 20-Nov-2008

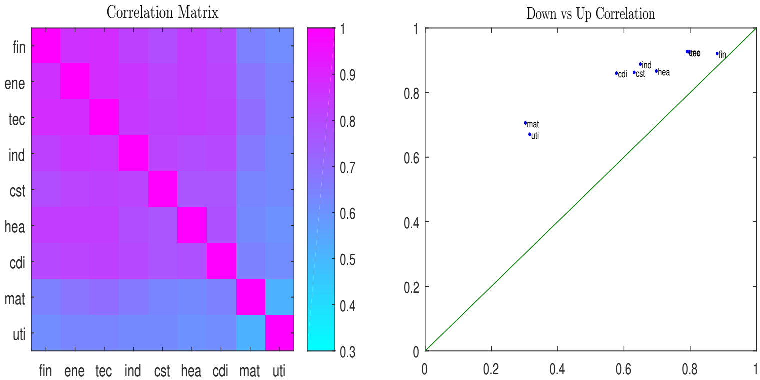


Figure: Implied Correlations.

Up and down average pairwise correlations

From **option prices**, we estimate:

$$\rho_{i,j}^{g,\mathbb{Q}} = \text{corr}^{\mathbb{Q}}(R_i, R_j)$$

$$\rho_{i,j}^{d,\mathbb{Q}} = \text{corr}^{\mathbb{Q}}(R_i, R_j \mid R_S \leqslant R_S^M)$$

and

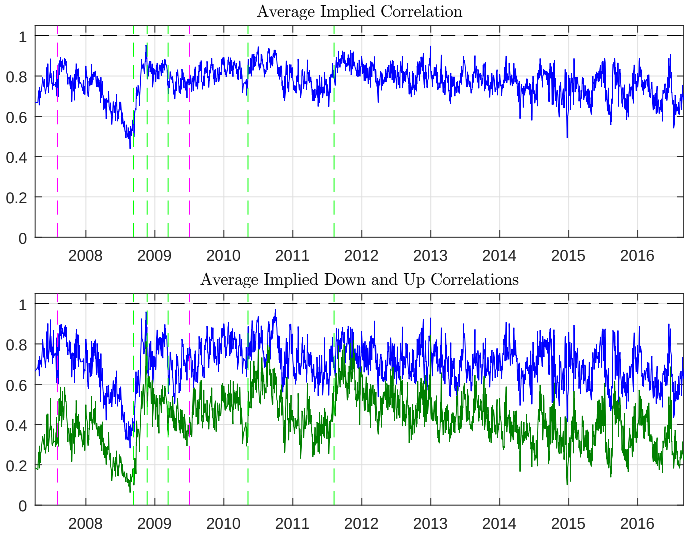
$$\rho_{i,j}^{u,\mathbb{Q}} = \text{corr}^{\mathbb{Q}}(R_i, R_j \mid R_S > R_S^M),$$

We then average

$$\rho^{x,\mathbb{Q}} = \frac{\sum_{i<j} \pi_i \pi_j \rho_{i,j}^{x,\mathbb{Q}}}{\sum_{i<j} \pi_i \pi_j},$$

with $\pi_i = \omega_i \sigma_i$

Implied Correlation



Up and down correlation risk premia

From **option prices**, we estimate:

$$\rho_{i,j}^{d,\mathbb{Q}} = \text{corr}^{\mathbb{Q}}(R_i, R_j \mid R_S \leqslant R_S^M)$$

and

$$\rho_{i,j}^{u,\mathbb{Q}} = \text{corr}^{\mathbb{Q}}(R_i, R_j \mid R_S > R_S^M),$$

From corresponding **stock prices** daily returns

$$\rho_{i,j}^{d,\mathbb{P}} = \text{corr}^{\mathbb{P}}(R_i, R_j \mid R_S \leqslant R_S^M)$$

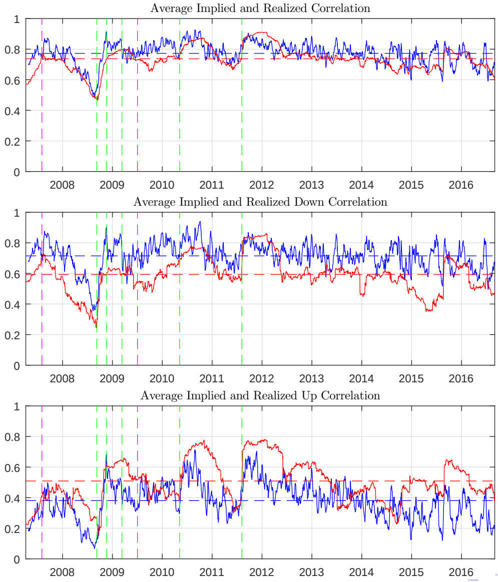
and

$$\rho_{i,j}^{u,\mathbb{P}} = \text{corr}^{\mathbb{P}}(R_i, R_j \mid R_S > R_S^M),$$

Correlation risk premium (global, up and down):

$$\rho_{i,j}^{g,\mathbb{P}} - \rho_{i,j}^{g,\mathbb{Q}}, \quad \rho_{i,j}^{u,\mathbb{P}} - \rho_{i,j}^{u,\mathbb{Q}}, \quad \rho_{i,j}^{d,\mathbb{P}} - \rho_{i,j}^{d,\mathbb{Q}}$$

Implied and Realized Correlation



Results

What we observe

$$\rho_{i,j}^{u,\mathbb{Q}} < \rho_{i,j}^{u,\mathbb{P}} < \rho_{i,j}^{d,\mathbb{P}} < \rho_{i,j}^{d,\mathbb{Q}}$$

Asymmetry under $\mathbb{P}$ was observed in the literature: Longin and Solnik (JOF 2001), Ang and Bekaert (RFS 2002), Hong, Tu and Zhou (RFS 2007), Jondeau (CSDA, 2016)... higher correlations in “bear markets”

Under $\mathbb{Q}$, this asymmetry is **amplified** and we give evidence that this asymmetry in the correlations comes from an **asymmetry in the dependence** and **not** from properties of the **marginal** distributions.

Margins or Dependence?

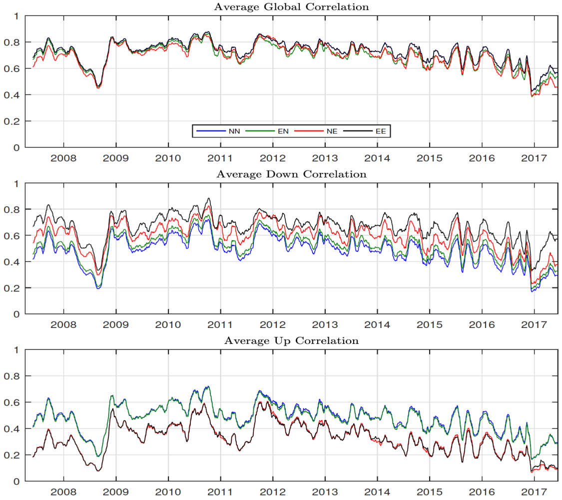


Figure 4.10: **Implied Correlations**. Average implied global, down, and up correlations are computed for the four cases (NN, EN, NE, EE), where the first letter denotes the type of margins (Normal or Empirical) and the second letter denotes the type of the copula (Normal or Empirical). Statistics are plotted as 1-month moving averages.

Additional Elements To Be Found in the Paper

- Implied dependence is **non-Gaussian**, **time-varying**, and **asymmetric**
- Global Correlation Risk Premium **disappears** when computed with Spearman's Rho, whereas the Down (resp. Up) Correlation Risk Premium stays significantly negative (resp. positive)
- **Alternative semi-parametric approach** to our model-free approach to model the joint distribution of assets in the risk-neutral world:
  - Fit **margins** with **model-free** approach
  - Fit **dependence** using a two-parameter **Skewed Normal Copula**

  Model sufficiently flexible to re-obtain the results on the global, down and up correlation risk premia

Conclusions on the Analysis of the Correlation Risk Premium

- A **novel algorithm** to infer the dependence among variables given their marginal distributions and distribution of the sum
- Consistent with **maximum entropy**. This is a desirable property: a dependence with lower entropy would mean that we use information that we do not possess
- **Application** to S&P 500 Sector options:
  - Implied dependence is **non-Gaussian**, time-varying, and asymmetric
  - **Down correlation is larger than Up correlation**
  - Correlation risk premium: **Down** (strongly negative), **Up** (positive), **Global** (negative)
  - Parsimonious multivariate model with a two-parameter copula
  - Evidence of extreme events / left tail dependence
  - **Correlation indices** (down, up), improving on CBOE index

Other Potential Applications

A number of potential **applications**:

- Identify **properties of a "good" multivariate model** to reproduce option prices available in the market (such as stochastic correlation, asymmetry between average up and down correlation, etc).
- A new approach to price any **path-independent multivariate derivatives** (basket options and correlation swaps). Joint work with Oleg Bondarenko and Steven Vanduffel.
- Detection of **arbitrage** opportunities – Dispersion arbitrage
- Disentangle modelling of **volatility** (margins) and of the **dependence** (copula)
- New forward-looking indicators of **contagion/tail risk**
- Covariance matrix estimation / Optimal **portfolio** construction

Outline

Part 1: The Rearrangement Algorithm

- Minimizing variance of a sum with full dependence uncertainty
- Variance bounds

Part 2: Application to Model-Risk Assessment, e.g., Uncertainty on Value-at-Risk

- With 2 risks and full dependence uncertainty
- With d risks and full dependence uncertainty

Part 3: Adding information on dependence

- Moment constraints
- Information on a subset...

Part 4: Using the RA to infer dependence

- Add information about the sum of the risks
- Application to explain the correlation risk premium
- Application to multivariate option pricing

Part 5: Improved Rearrangement Algorithm

Part V
Improved block rearrangement algorithm
with Jinghui Chen, Ludger Rüschendorf and Steven Vanduffel

Block rearrangement algorithm (BRA)

$d = 4$ variables: X_1, X_2, X_3, X_4, $n = 5$ values with probability $\frac{1}{5}$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \\ 5 & 5 & 5 & 5 \end{bmatrix}$$

The yellow block size $r_t = 1$: $\begin{cases} X_1 \downarrow X_2 + X_3 + X_4 \\ X_2 \downarrow X_1 + X_3 + X_4 \\ X_3 \downarrow X_1 + X_2 + X_4 \\ X_4 \downarrow X_1 + X_2 + X_3 \end{cases}$

Block rearrangement algorithm (BRA)

$d = 4$ variables: X_1, X_2, X_3, X_4, $n = 5$ values with probability $\frac{1}{5}$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \\ 5 & 5 & 5 & 5 \end{bmatrix}$$

The yellow block size $r_t = 1$:
$$\begin{cases} X_1 \downarrow X_2 + X_3 + X_4 \\ X_2 \downarrow X_1 + X_3 + X_4 \\ X_3 \downarrow X_1 + X_2 + X_4 \\ X_4 \downarrow X_1 + X_2 + X_3 \end{cases}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \\ 5 & 5 & 5 & 5 \end{bmatrix}$$

The yellow block size $r_t = 2$:
$$\begin{cases} X_1 + X_2 \downarrow X_3 + X_4 \\ X_1 + X_3 \downarrow X_2 + X_4 \\ X_1 + X_4 \downarrow X_2 + X_3 \end{cases}$$

Block rearrangement algorithm (BRA)

$d = 4$ variables: X_1, X_2, X_3, X_4, $n = 5$ values with probability $\frac{1}{5}$

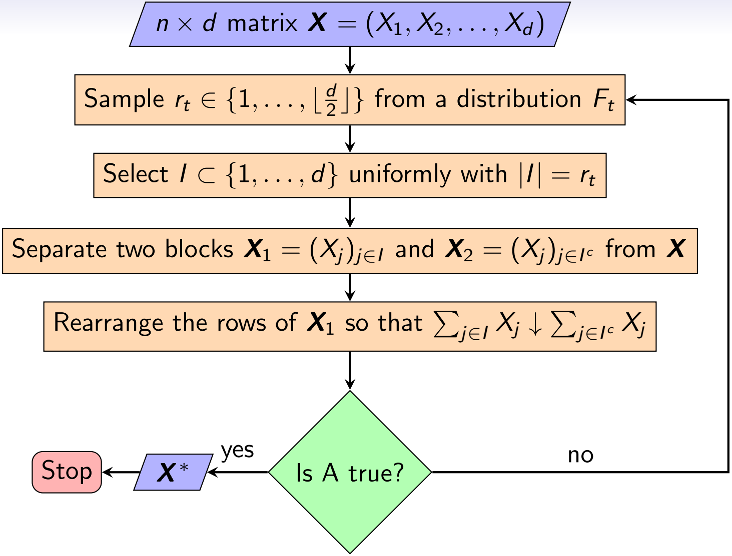
$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \\ 5 & 5 & 5 & 5 \end{bmatrix}$$

The yellow block size $r_t = 1$: $\begin{cases} X_1 \downarrow X_2 + X_3 + X_4 \\ X_2 \downarrow X_1 + X_3 + X_4 \\ X_3 \downarrow X_1 + X_2 + X_4 \\ X_4 \downarrow X_1 + X_2 + X_3 \end{cases}$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \\ 5 & 5 & 5 & 5 \end{bmatrix}$$

The yellow block size $r_t = 2$: $\begin{cases} X_1 + X_2 \downarrow X_3 + X_4 \\ X_1 + X_3 \downarrow X_2 + X_4 \\ X_1 + X_4 \downarrow X_2 + X_3 \end{cases}$

Applications: Model risk on VaR, TVaR, variance and so on



A: $\sum_{j \in I} X_j \downarrow \sum_{j \in I^c} X_j$ for all possible I.

We expect that the performance of BRA may be affected by two factors:

1. the cardinality of subset I in each step, i.e., the number r_t of columns of $\boldsymbol{X}_1$;
2. the maximum number of iterations T.

Definition (BRA Unif)

A BRA is called BRA Unif if each F_t is a discrete uniform distribution with support $\{1, 2, \ldots, \lfloor \frac{d}{2} \rfloor\}$.

when $d = 4$:
BRA Unif: $\mathbb{P}(r_t = 1) = \mathbb{P}(r_t = 2) = \frac{1}{2}$ at each BRA step
standard BRA: $\mathbb{P}(r_t = 1) = \frac{4}{7}$, $\mathbb{P}(r_t = 2) = \frac{3}{7}$

To measure the performance of BRA, we use

$$\delta_t = \log \mathrm{Var}(X_1 + X_2 + \cdots + X_d)$$

to denote the log variance after t steps of BRA.

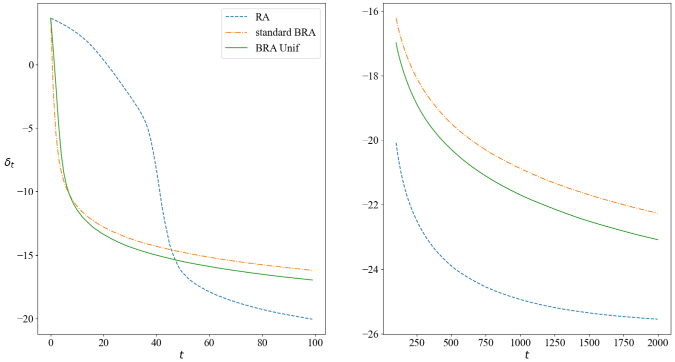


Figure: **Uniform Risks:** δ_t with $k = 100$, $T = 2000$, $n = 1000$ and $d = 500$. The left figure displays the δ_t during the first 100 steps, while the right displays the δ_t after 100 steps.

Our observations suggest the need for a BRA design that behaves similarly to the standard BRA at the beginning, and more like the RA towards the end, achieving better performance overall.

Our observations suggest the need for a BRA design that behaves similarly to the standard BRA at the beginning, and more like the RA towards the end, achieving better performance overall. when

$d = 100$:
RA: $\mathbb{P}(r_t = 1) = 1$
standard BRA: $\mathbb{P}(r_t = 1) = \frac{100}{2^{99}-1} \approx 0$ and $\mathbb{P}(r_t = 50) = 7.96\%$
BRA Unif: $\mathbb{P}(r_t = 1) = \frac{1}{50}$ and $\mathbb{P}(r_t = 50) = \frac{1}{50}$

BRA Beta

Definition (BRA Beta)

A BRA is called BRA Beta if F_t is the distribution where a random variable, $r_t \sim F_t$, takes integer parts of numbers sampled from Beta(α_t, β_t). The parameters α_t and β_t are

$$\begin{aligned} \alpha_t &= A - \left(\frac{t-1}{T-1}\right)^{\frac{1}{B}}(A-1), \\ \beta_t &= 1 + \left(\frac{t-1}{T-1}\right)^{\frac{1}{B}}(A-1), \end{aligned} \tag{1}$$

where A and B are two constants.

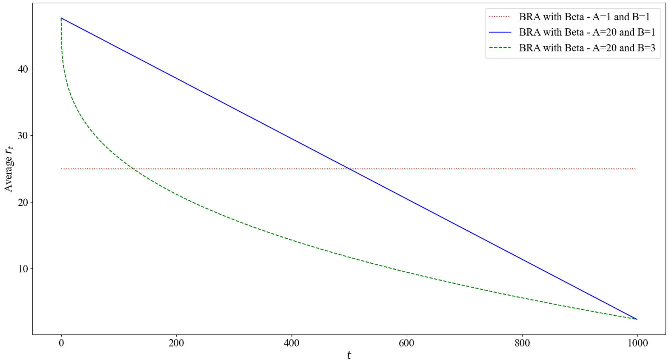


Figure: Average r_t of the Beta distributions as a function of t. The graph shows the average r_t from the corresponding Beta distribution for $d = 100$, $T = 1000$ and some examples of A and B.

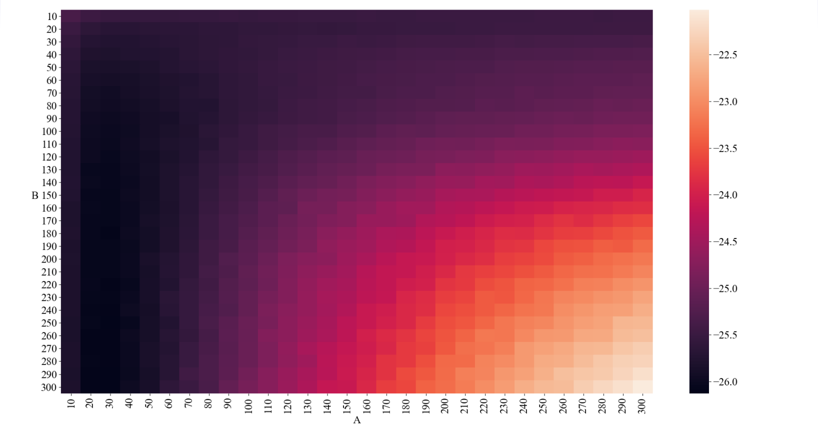


Figure: **Uniform risks:** The heatmaps of δ_T with $k = 100$, $T = 2000$, $n = 1000$ and $d = 100$ when implementing the BRA Beta for different A and B.

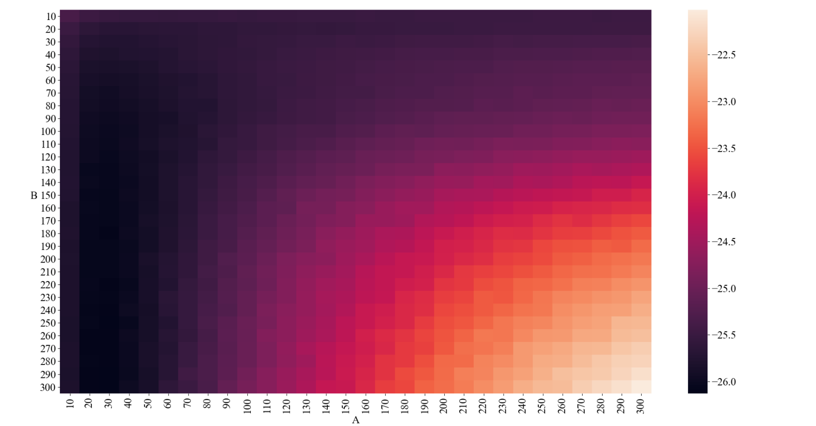


Figure: **Uniform risks:** The heatmaps of δ_T with $k = 100$, $T = 2000$, $n = 1000$ and $d = 100$ when implementing the BRA Beta for different A and B.

The best choices for A and B are $A = 0.3d$ and $B = 50$.

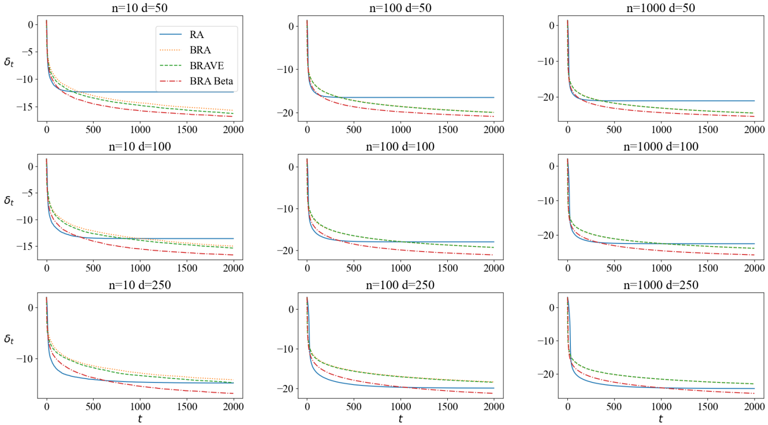


Figure: **Uniform risks:** The effect of four types of BRA on the trajectory of δ_t with $k = 100$ and $T = 2000$ as a function of t.


THANK YOU

Robustness to Initial Conditions (back)

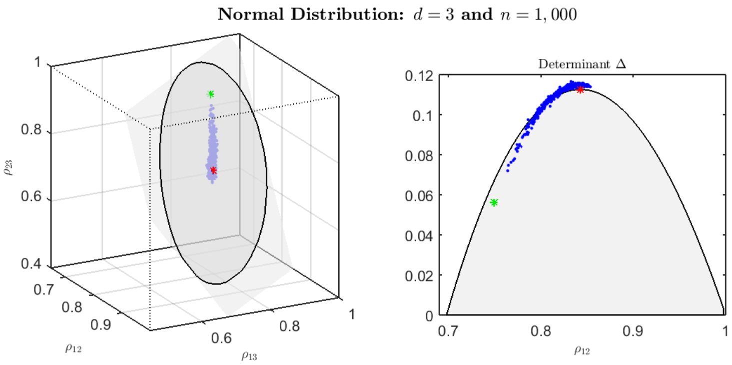


Figure: $K = 500$ blue dots correspond to different runs of BRA. Each run starts at a particular solution (green star), but with 2 random rows swapped. Shaded gray area is constrained set $\mathcal{C}(\rho^{imp})$, red star is maximal correlation matrix R_M.

Robustness to Initial Conditions (back)

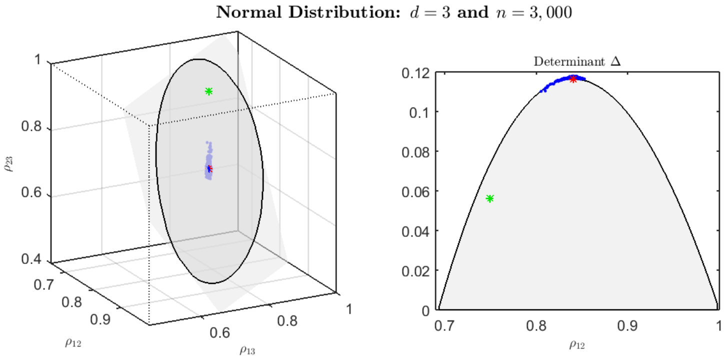


Figure: $K = 500$ blue dots correspond to different runs of BRA. Each run starts at a particular solution (green star), but with 6 random rows swapped. Shaded gray area is constrained set $\mathcal{C}(\rho^{imp})$, red star is maximal correlation matrix R_M.

Robustness to Initial Conditions (back)

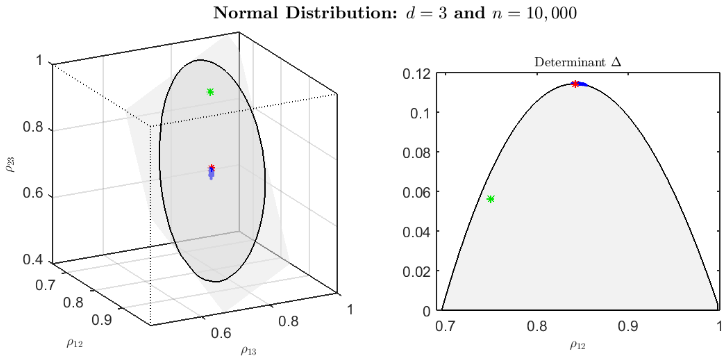


Figure: $K = 500$ blue dots correspond to different runs of BRA. Each run starts at a particular solution (green star), but with 20 random rows swapped. Shaded gray area is constrained set $\mathcal{C}(\rho^{imp})$, red star is maximal correlation matrix R_M.

Robustness to Distributional Assumptions (back)

- A d-dimensional random vector $\mathbf{X}$ is a normal mean-variance mixture, if $\mathbf{X} \sim \boldsymbol{\mu} + Y\boldsymbol{\gamma} + \sqrt{Y}\mathbf{Z}$ where $\mathbf{Z} \sim N_d(0, \mathbf{W})$, $Y \geqslant 0$ is a scalar random variable independent of $\mathbf{Z}$, and $\boldsymbol{\gamma} \in \mathbb{R}^d$ and $\boldsymbol{\mu} \in \mathbb{R}^d$ are constants.
- We consider a special case where Y is Inverse Gamma, $Y \sim IG(\nu/2, \nu/2)$. This corresponds to a Skewed-t distribution $\mathbf{X} \sim Skew_d(\nu, \boldsymbol{\mu}, \mathbf{W}, \gamma)$
- The sum S as well as the components X_i $(i = 1, 2, \ldots, d)$ follow one-dimensional Skewed-t distribution. In particular,

$$S \sim Skew_1\left(\nu, \sum_i \mu_i, \mathbf{1W1}^t, \sum_i \gamma_i\right).$$

Multivariate Skewed-t Distribution (back)

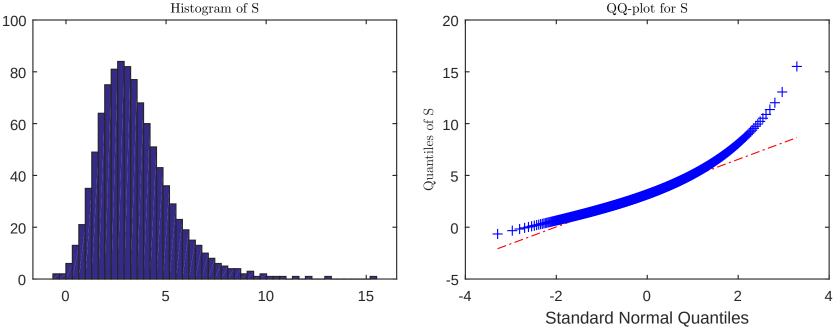


Figure: Histogram and QQ-plot for sum S generated with multivariate Skewed-t distribution when $d = 3$ and $n = 1000$.

Stability of BRA: Multivariate Skewed-t Distribution (back)

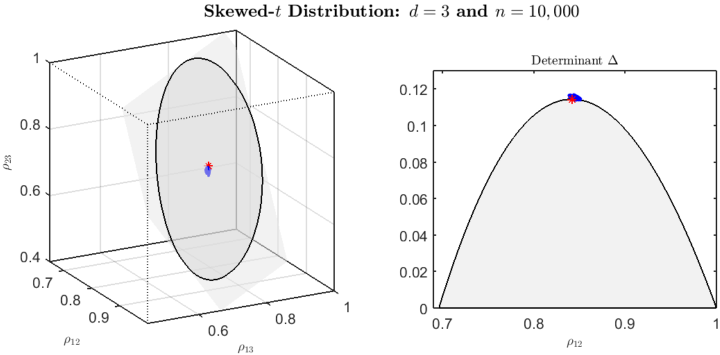
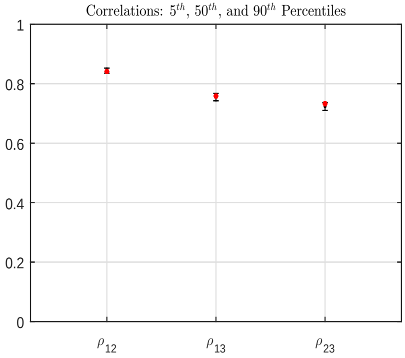


Figure: $K = 500$ blue dots correspond to different runs of the BRA. The shaded gray area is the constrained set $\mathcal{C}(\rho^{imp})$; the red star is the maximal matrix R_M. The left panel shows realizations of the correlations ρ_{12}, ρ_{13}, and ρ_{23}. The right panel shows the relation Δ versus ρ_{12}.

Pairwise correlations (back)

We recover the dependence among the variables including pairwise correlations

$d = 3$



$d = 10$