Quantile Estimation in Structural Reliability with Incomplete Dependence Structure

Nazih Benoumechiara [1 2]

Gérard Biau [1] Bertrand Michel [3] Philippe Saint-Pierre [4]
Roman Sueur [2] Nicolas Bousquet [1,5] Bertrand Iooss [2,4]

[1]Laboratoire de Probabilités, Statistique et Modélisation,
Sorbonne Université / UPMC, Paris.

[2]Performances, Risques Industriels et Surveillance pour la Maintenance et l'Exploitation (PRISME),
EDF R&D, Chatou.

[3]Ecole Centrale de Nantes (ECN),
Nantes.

[4]Institut de Mathématique de Toulouse (IMT),
Université Paul Sabatier, Toulouse.

[5]Quantmetry R&D,
Paris.

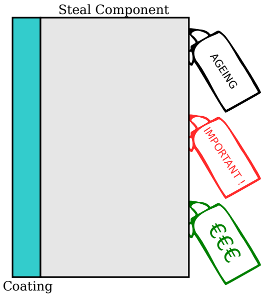
Wednesday, March 21th 2018

Contents

Contents

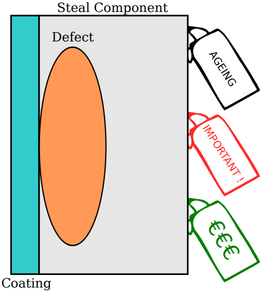
Industrial Context

- Important to safety component exposed to an ageing phenomenon



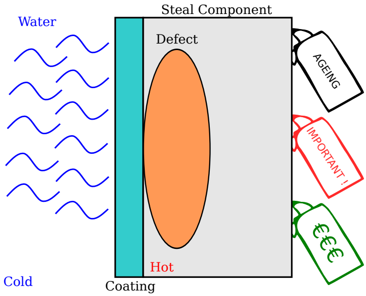
Industrial Context

- Important to safety component exposed to an ageing phenomenon
- Potential manufacturing defects



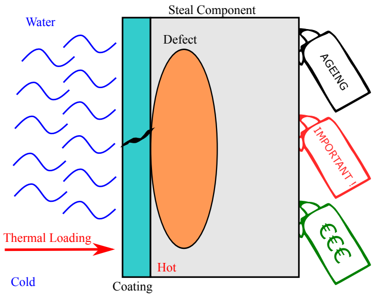
Industrial Context

- Important to safety component exposed to an ageing phenomenon
- Potential manufacturing defects
- Extreme thermo-mechanical constraints



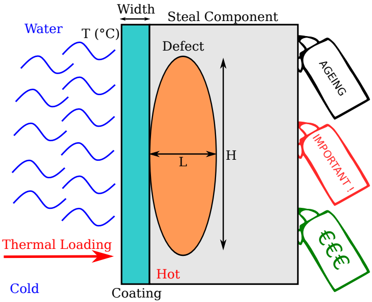
Industrial Context

- Important to safety component exposed to an ageing phenomenon
- Potential manufacturing defects
- Extreme thermo-mechanical constraints
- Jeopardize the structure



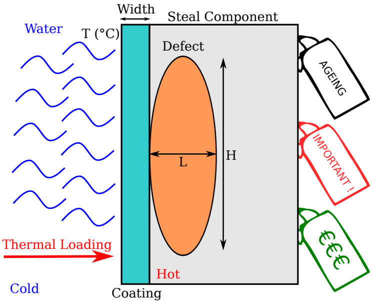
Industrial Context

- Important to safety component exposed to an ageing phenomenon
- Potential manufacturing defects
- Extreme thermo-mechanical constraints
- Jeopardize the structure
- Several parameters influence the structural safety

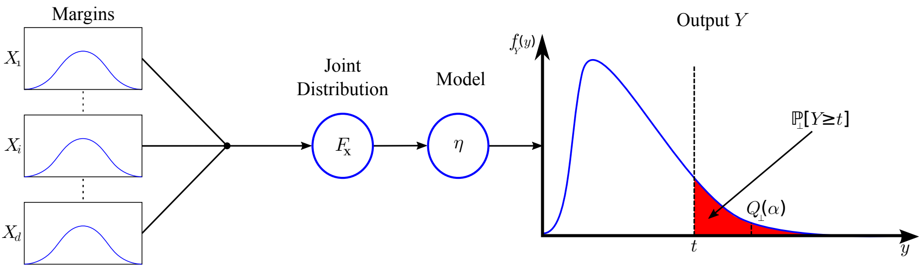


Industrial Context

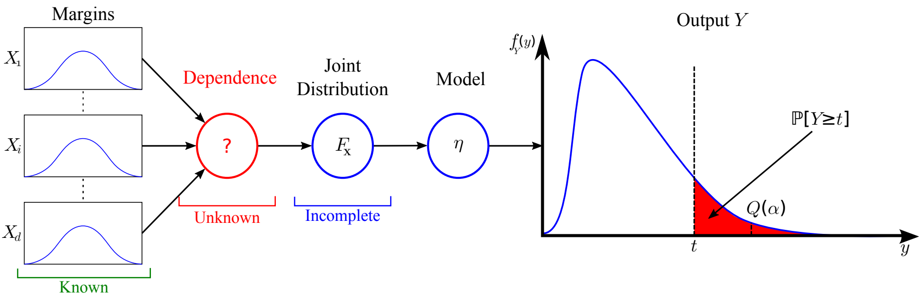
- Important to safety component exposed to an ageing phenomenon
- Potential manufacturing defects
- Extreme thermo-mechanical constraints
- Jeopardize the structure
- Several parameters influence the structural safety
- The parameters are uncertain



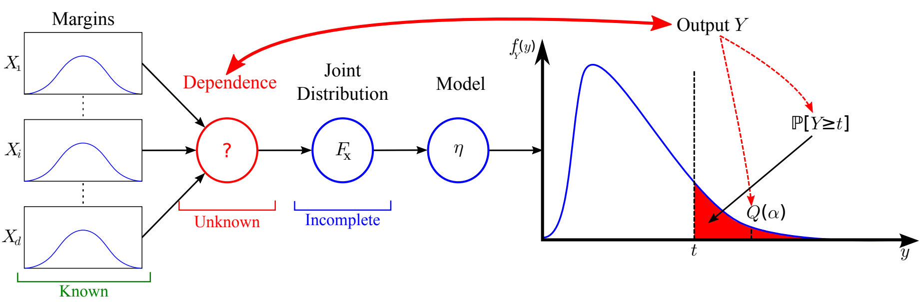
Structural Reliability



Structural Reliability



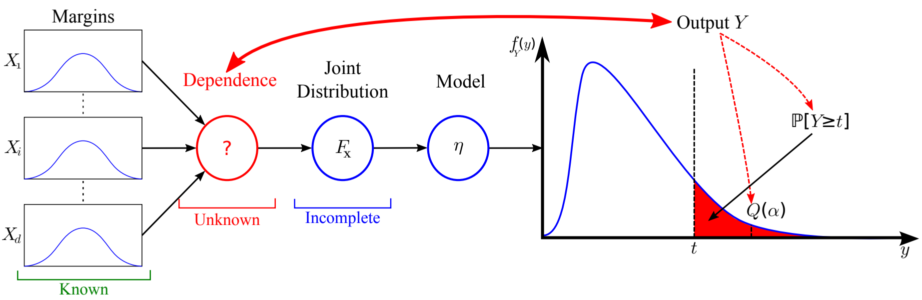
Structural Reliability



Question:

- How important the dependence structure of $\mathbf{X}$ is on the quantity of interest of Y?

Structural Reliability



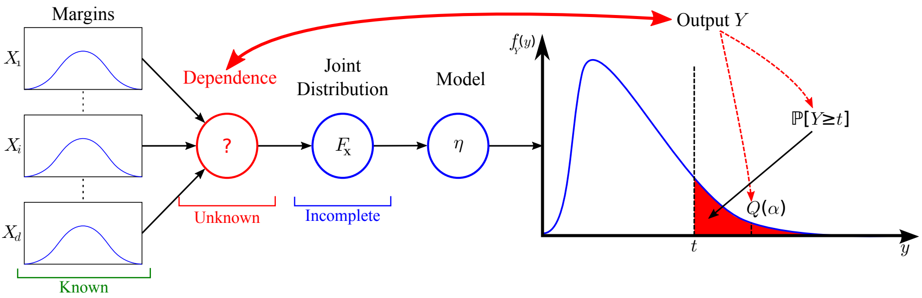
Question:

- How important the dependence structure of $\mathbf{X}$ is on the quantity of interest of Y?

Objectives:

- Inform about the importance of the dependence structure.
- Bound the quantity of interest: e.g. $\mathbb{P}_{\perp}[Y \geq t] \leq \mathbb{P}^*[Y \leq t]$.

Structural Reliability



Question:

- How important the dependence structure of $\mathbf{X}$ is on the quantity of interest of Y?

Objectives:

- Inform about the importance of the dependence structure.
- Bound the quantity of interest: e.g. $\mathbb{P}_{\perp}[Y \geq t] \leq \mathbb{P}^{*}[Y \leq t]$.

How to describe the dependence structure?

Copulas

The dependence structure is described by a **parametric** copula $C_{\boldsymbol{\theta}}$ with $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$ such as[5,6]

$$F_{\mathbf{X}}(\mathbf{x}) = C_{\boldsymbol{\theta}}(F_1(x_1), \ldots, F_d(x_d)).$$

Family	$C_\theta(u, v)$	Θ	Kendall's τ
Independent	uv	/	/
Gaussian	$\Phi_\theta(\Phi^{-1}(u), \Phi^{-1}(v))$	$[-1, 1]$	$\frac{2 \arcsin \theta}{\pi}$
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$[0, \infty]$	$\frac{\theta}{2+\theta}$
Gumbel	$\exp\left\{-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}\right\}$	$[1, \infty]$	$1 - \frac{1}{\theta}$

Table: Example of copula families.

[5] Roger B Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.
[6] Abe Sklar. *Fonctions de répartition à n dimensions et leurs marges*. Vol. 8. ISUP, 1959, pp. 229–231.

Copulas

The dependence structure is described by a **parametric** copula $C_{\boldsymbol{\theta}}$ with $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$ such as[5,6]

$$F_{\mathbf{X}}(\mathbf{x}) = C_{\boldsymbol{\theta}}(F_1(x_1), \ldots, F_d(x_d)).$$

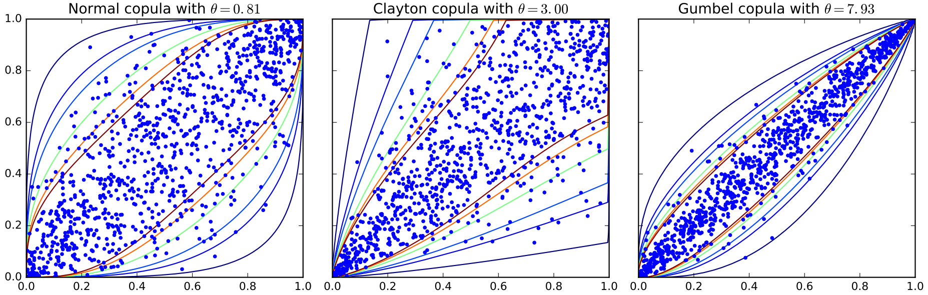


Figure: Example of copula densities with $\tau = 0.6$.

[5] Roger B Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.
[6] Abe Sklar. *Fonctions de répartition à n dimensions et leurs marges*. Vol. 8. ISUP, 1959, pp. 229–231.

Copulas

The dependence structure is described by a **parametric** copula $C_{\boldsymbol{\theta}}$ with $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$ such as[5,6]

$$F_{\mathbf{X}}(\mathbf{x}) = C_{\boldsymbol{\theta}}(F_1(x_1), \ldots, F_d(x_d)).$$

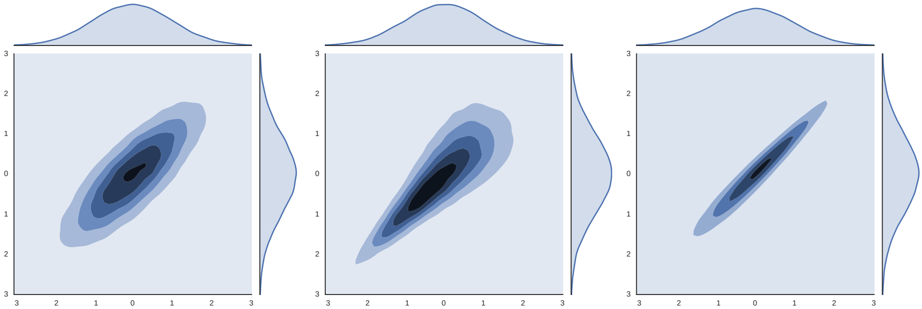
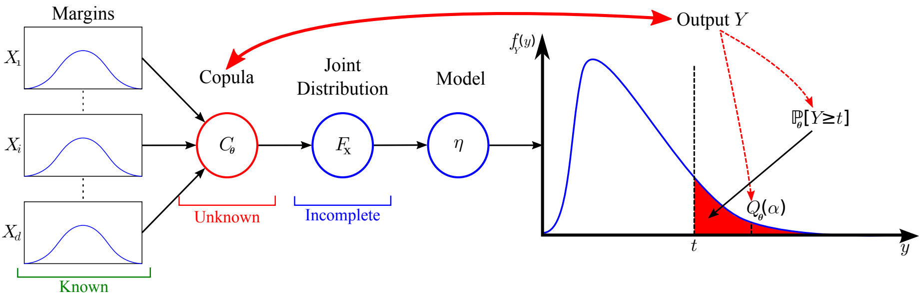
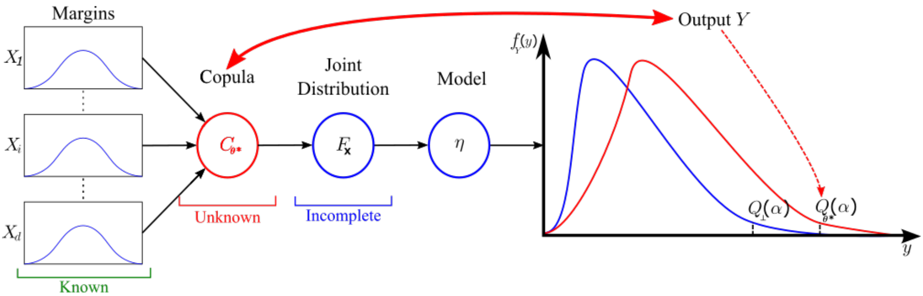


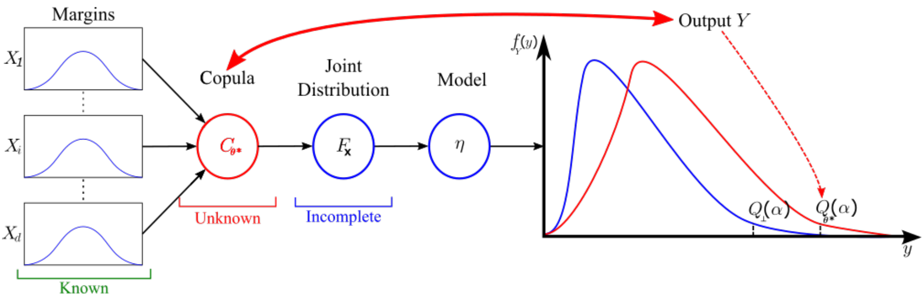
Figure: Example of joints p.d.f with Gaussian margins and $\tau = 0.6$.

[5] Roger B Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.
[6] Abe Sklar. *Fonctions de répartition à n dimensions et leurs marges*. Vol. 8. ISUP, 1959, pp. 229–231.


Margins
X_1
X_i
X_d
Known
Copula
C_θ
Unknown
Joint Distribution
$F_{\mathbf{X}}$
Incomplete
Model
η
Output Y
$f_Y(y)$
$\mathbb{P}_\theta[Y \geq t]$
$Q_\theta(\alpha)$
t
y



We are interested on the **output quantile**.



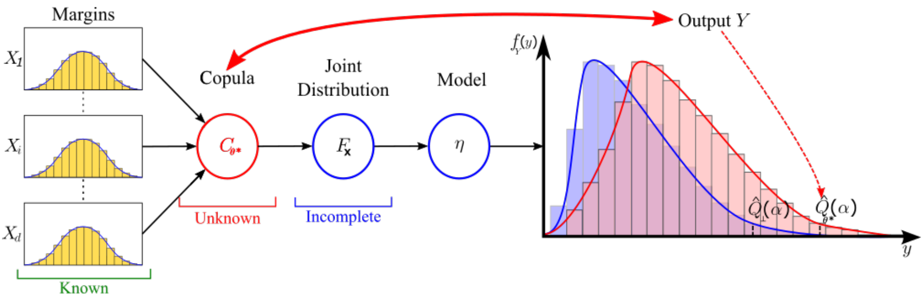
We are interested on the **output quantile**.
For a given probability $\alpha \in (0, 1)$ and a parametric copula $C_{\boldsymbol{\theta}}$ such that $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$, we state the maximization problem[1]

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\operatorname{argmax}}\, Q_{\boldsymbol{\theta}}(\alpha),$$

which gives the upper bound

$$Q_{\boldsymbol{\theta}^*}(\alpha) \geq Q_{\perp}(\alpha).$$

1 Or minimization



We are interested on the **output quantile**.
For a given probability $\alpha \in (0, 1)$ and a parametric copula $C_{\boldsymbol{\theta}}$ such that $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$, we state the maximization problem[1]

$$\boldsymbol{\theta}^* = \operatorname*{argmax}_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} Q_{\boldsymbol{\theta}}(\alpha),$$

which gives the upper bound

$$Q_{\boldsymbol{\theta}^*}(\alpha) \geq Q_{\perp}(\alpha).$$

Because Y is not explicitly known, $Q_{\boldsymbol{\theta}}(\alpha)$ is estimated:

$$\hat{\boldsymbol{\theta}}_n = \operatorname*{argmax}_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \widehat{Q}_{n,\boldsymbol{\theta}}(\alpha).$$

1 Or minimization

Contents

Related Studies

Several studies showed the influence of dependencies[3,7].

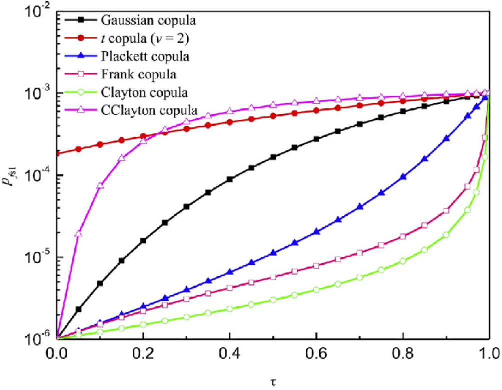


Figure: Variation of the output probability for different copula families.

[3] Mircea Grigoriu and Carl Turkstra. "Safety of structural systems with correlated resistances". In: *Applied Mathematical Modelling* 3.2 (1979), pp. 130–136.
[7] Xiao-Song Tang et al. "Impact of copulas for modeling bivariate distributions on system reliability". In: *Structural safety* 44 (2013), pp. 80–90.

The worst case is not always at the edge

> “ ***Fallacy 3.*** *The worst case VaR (quantile) for a linear portfolio $X + Y$ occurs when $\rho(X, Y)$ is maximal, i.e. X and Y are comonotonic.*[2] ”

For example, we consider:

- $X_1 \sim \mathcal{N}(0,1)$, $X_2 \sim \mathcal{N}(-2,1)$ and different copula families with $\tau \in [-1,1]$
- $\eta(x_1, x_2) = x_1^2 x_2^2 - x_1 x_2$

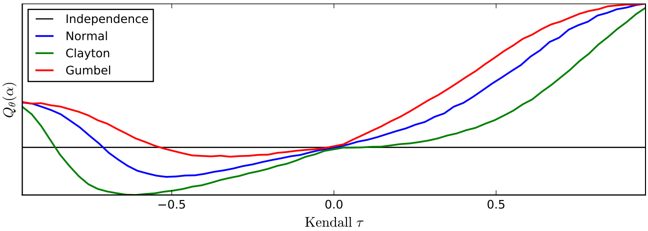


Figure: Variation of the output quantile for different copula families and $\alpha = 5\%$.

[2] Paul Embrechts, Alexander McNeil, and Daniel Straumann. “Correlation and dependence in risk management: properties and pitfalls”. In: *Risk management: value at risk and beyond* (2002), pp. 176–223.

In practice, $\boldsymbol{\theta}$ is discretized using a thin grid $\boldsymbol{\Theta}_K$ of size K:

$$\hat{\boldsymbol{\theta}}_{n,K} = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}_K}{\operatorname{argmax}} \, \widehat{Q}_{n,\boldsymbol{\theta}}(\alpha).$$

In practice, $\boldsymbol{\theta}$ is discretized using a thin grid $\boldsymbol{\Theta}_K$ of size K:

$$\hat{\boldsymbol{\theta}}_{n,K} = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}_K}{\text{argmax}}\, \widehat{Q}_{n,\boldsymbol{\theta}}(\alpha).$$

Theorem 1 (Consistency of $\hat{\boldsymbol{\theta}}_{n,K}$)

As K tends to infinity and under regularity assumptions of η and $F_{\mathbf{X}}$, for a given $\alpha \in (0, 1)$ and for all $\varepsilon > 0$ we have

$$P\left(\left|\widehat{Q}_{\hat{\boldsymbol{\theta}}_{n,K}}(\alpha) - Q_{\boldsymbol{\theta}^*}(\alpha)\right| > \varepsilon\right) \xrightarrow{n \to \infty} 0.$$

Moreover, if $Q_{\boldsymbol{\theta}}$ is uniquely minimized at $\boldsymbol{\theta}^*$, then for all $h > 0$ we have

$$\mathbb{P}[|\hat{\boldsymbol{\theta}}_{n,K} - \boldsymbol{\theta}^*| > h] \xrightarrow{n \to \infty} 0.$$

How to model a multivariate copula?

How to model a multivariate copula?

- Elliptical Copulas: the worst case correlation matrix
  - Pros: intuitive, simple to implement
  - Cons: assumption of linear correlations, no tails dependencies, ...
- Archimedian Copulas
  - Pros: simple, tail dependencies
  - Cons: not flexibile with $\boldsymbol{\theta}$
- Vine Copulas
  - Pros: flexibility
  - Cons: complexity

How to model a multivariate copula?

- Elliptical Copulas: the worst case correlation matrix
  - Pros: intuitive, simple to implement
  - Cons: assumption of linear correlations, no tails dependencies, ...
- Archimedian Copulas
  - Pros: simple, tail dependencies
  - Cons: not flexibile with $\boldsymbol{\theta}$
- **Vine Copulas**
  - Pros: flexibility
  - Cons: complexity

Regular Vines

The joint density $f(x_1, \ldots, x_d)$ can be represented by a product of pair-copula densities and marginal densities[4].

[4] Harry Joe. "Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters". In: *Lecture Notes-Monograph Series* (1996), pp. 120–141.

Regular Vines

The joint density $f(x_1, \dots, x_d)$ can be represented by a product of pair-copula densities and marginal densities[4].

For example in $d = 4$. One possible decomposition of $f(x_1, x_2, x_3, x_4)$ is:

$f(x_1, x_2, x_3, x_4) = f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4)_{\text{(margins)}}$

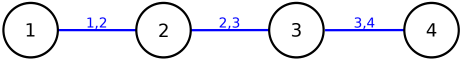
[4] Harry Joe. "Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters". In: *Lecture Notes-Monograph Series* (1996), pp. 120–141.

Regular Vines

The joint density $f(x_1, \ldots, x_d)$ can be represented by a product of pair-copula densities and marginal densities[4].

For example in $d = 4$. One possible decomposition of $f(x_1, x_2, x_3, x_4)$ is:

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) = & f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4)_{\text{(margins)}} \\
& \times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \cdot c_{34}(F_3(x_3), F_4(x_4))_{\text{(unconditional pairs)}}
\end{aligned}
$$



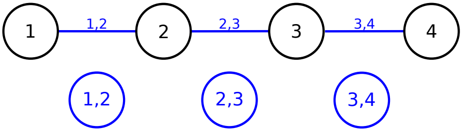
[4] Harry Joe. "Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters". In: *Lecture Notes-Monograph Series* (1996), pp. 120–141.

Regular Vines

The joint density $f(x_1, \ldots, x_d)$ can be represented by a product of pair-copula densities and marginal densities[4].

For example in $d = 4$. One possible decomposition of $f(x_1, x_2, x_3, x_4)$ is:

$$f(x_1, x_2, x_3, x_4) = f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4)_{\text{(margins)}}$$
$$\times\, c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \cdot c_{34}(F_3(x_3), F_4(x_4))_{\text{(unconditional pairs)}}$$



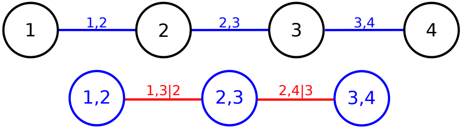
[4] Harry Joe. "Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters". In: *Lecture Notes-Monograph Series* (1996), pp. 120–141.

Regular Vines

The joint density $f(x_1, \dots, x_d)$ can be represented by a product of pair-copula densities and marginal densities[4].

For example in $d = 4$. One possible decomposition of $f(x_1, x_2, x_3, x_4)$ is:

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) = {} & f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4)_{\text{(margins)}} \\
& \times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \cdot c_{34}(F_3(x_3), F_4(x_4))_{\text{(unconditional pairs)}} \\
& \times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \cdot c_{24|3}(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3))_{\text{(conditional pair)}}
\end{aligned}
$$



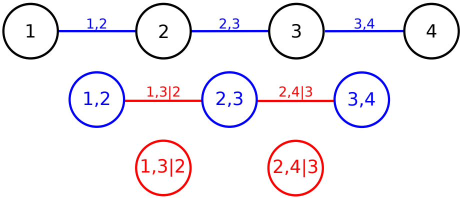
[4] Harry Joe. "Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters". In: *Lecture Notes-Monograph Series* (1996), pp. 120–141.

Regular Vines

The joint density $f(x_1, \dots, x_d)$ can be represented by a product of pair-copula densities and marginal densities[4].

For example in $d = 4$. One possible decomposition of $f(x_1, x_2, x_3, x_4)$ is:

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) = {} & f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4)_{\text{(margins)}} \\
& \times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \cdot c_{34}(F_3(x_3), F_4(x_4))_{\text{(unconditional pairs)}} \\
& \times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \cdot c_{24|3}(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3))_{\text{(conditional pair)}}
\end{aligned}
$$



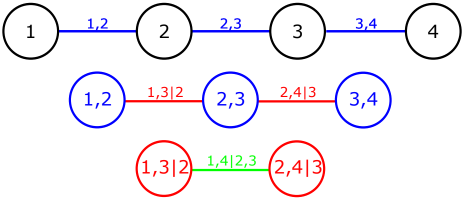
[4] Harry Joe. "Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters". In: *Lecture Notes-Monograph Series* (1996), pp. 120–141.

Regular Vines

The joint density $f(x_1, \ldots, x_d)$ can be represented by a product of pair-copula densities and marginal densities[4].

For example in $d = 4$. One possible decomposition of $f(x_1, x_2, x_3, x_4)$ is:

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) = & f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4)_{\text{(margins)}} \\
& \times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \cdot c_{34}(F_3(x_3), F_4(x_4))_{\text{(unconditional pairs)}} \\
& \times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \cdot c_{24|3}(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3))_{\text{(conditional pair)}} \\
& \times c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3))_{\text{(conditional pair)}}
\end{aligned}
$$



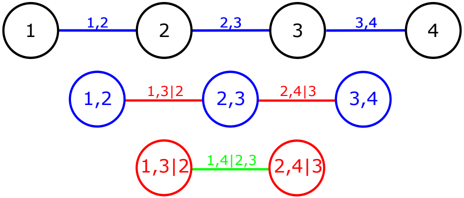
[4] Harry Joe. "Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters". In: *Lecture Notes-Monograph Series* (1996), pp. 120–141.

Regular Vines

Problem: There is very large number of possible Vine decompositions : $\binom{d}{2} \times (n-2)! \times 2^{\binom{d-2}{2}}$.

For example, when $d = 6$, there are 23.040 possible R-vines.

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) = {} & f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4) \text{(margins)} \\
& \times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \cdot c_{34}(F_3(x_3), F_4(x_4)) \text{(unconditional pairs)} \\
& \times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \cdot c_{24|3}(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3)) \text{(conditional pair)} \\
& \times c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)) \text{(conditional pair)}
\end{aligned}
$$

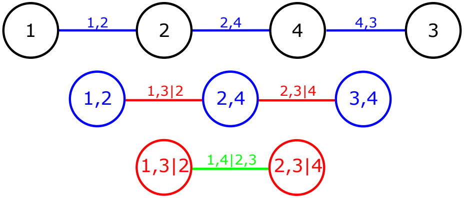


Regular Vines

Problem: There is very large number of possible Vine decompositions : $\binom{d}{2} \times (n-2)! \times 2^{\binom{d-2}{2}}$.

For example, when $d = 6$, there are 23.040 possible R-vines.

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) = & f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4) \text{(margins)} \\
& \times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{24}(F_2(x_2), F_4(x_4)) \cdot c_{34}(F_3(x_3), F_4(x_4)) \text{(unconditional pairs)} \\
& \times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \cdot c_{23|4}(F_{2|4}(x_2|x_4), F_{3|4}(x_3|x_4)) \text{(conditional pair)} \\
& \times c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)) \text{(conditional pair)}
\end{aligned}
$$



Example: Grid-search

- The model:

$$Y = -\sum_{j=1}^{d} \beta_j X_j,$$

  where $\beta_j = 10^{\frac{j}{d-1}}$.
- The marginal distributions: Generalized Pareto with $\sigma = 10$ and $\xi = 0.75$.
- The copula families: $d(d-1)/2$ Gaussian copulas.

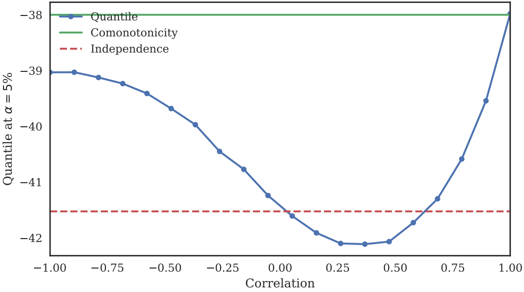


Figure: Quantile variation with the correlation in dimension $d = 2$

Example: Grid-search

- The model:

$$Y = -\sum_{j=1}^{d} \beta_j X_j,$$

where $\beta_j = 10^{\frac{j}{d-1}}$.

- The marginal distributions: Generalized Pareto with $\sigma = 10$ and $\xi = 0.75$.
- The copula families: $d(d-1)/2$ Gaussian copulas.
- The vine structure: a C-vine.

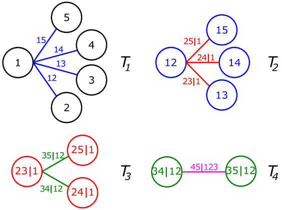


Figure: C-vine illustration in dimension $d = 5$

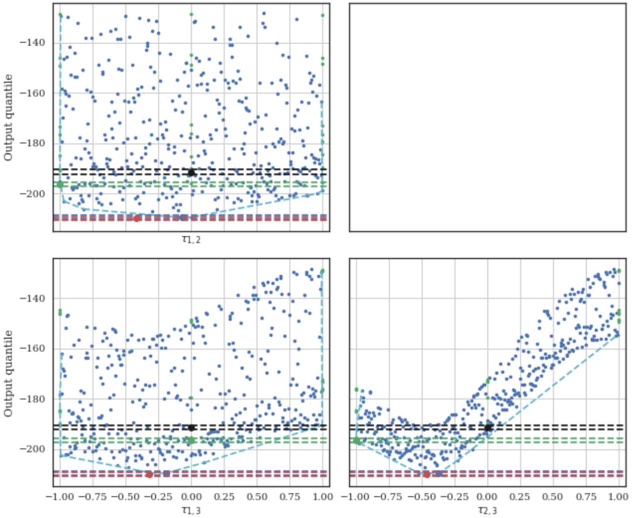


Figure: Quantile estimations with the dependence of each pairs in dimension $d = 3$.

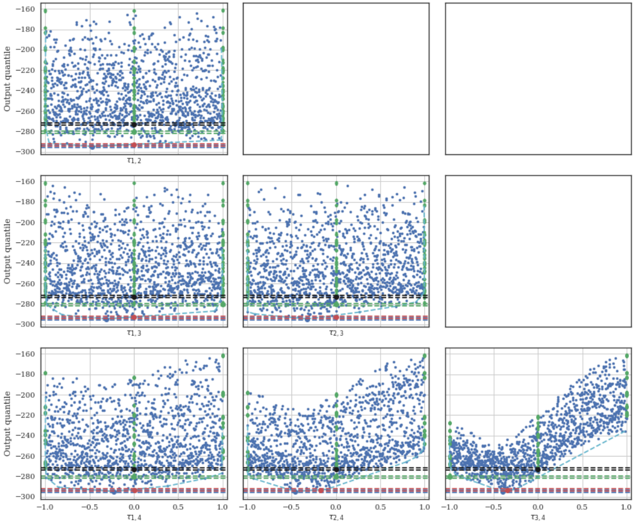


Figure: Quantile estimations with the dependence of each pairs in dimension $d = 4$.

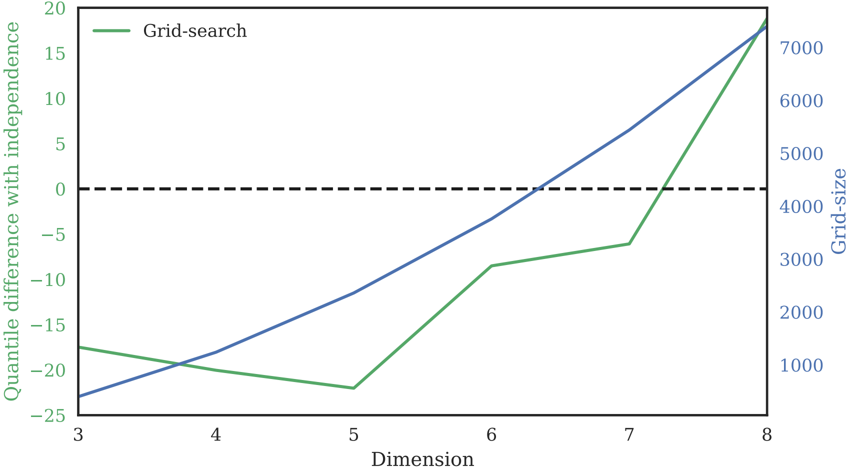


Figure: $Q_{\perp\!\!\!\perp}(\alpha) - Q_{\hat{\theta}_n}(\alpha)$

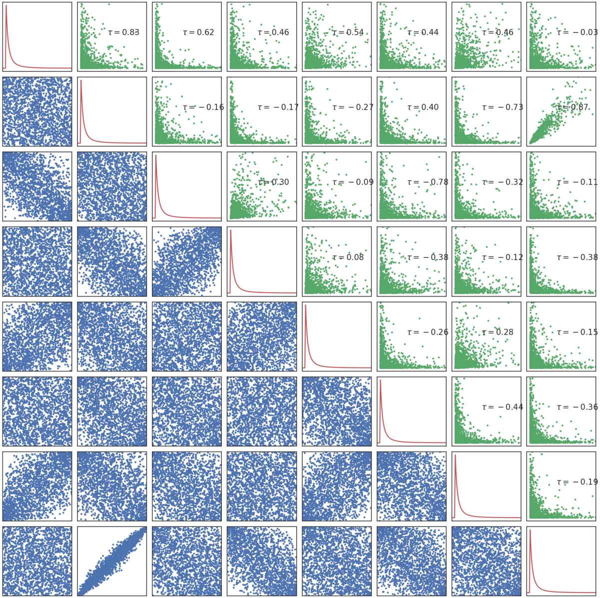


Figure: Sampling of $\mathbf{X}$ at the obtained minimum for $d = 8$

Grid-search with R-vines

- How can we chose the vine structure?
- Are all the pairs influential on the minimization?
- Can we simplify the dependence structure?

Algorithm 3: Generating a vine structure from a given list of indexed pairs Ω_k

Data: Ω_k, d

Result: A vine structure $\mathcal{V}$.

$\Omega_k^{init} = \Omega_k$;

$k = 1$;

do

  /* initialize $\mathcal{V}$ with a first empty tree */

  $N_1 = (1, \dots, d)$;

  $E_1 = ()$;

  $\mathcal{V} = ((N_1, E_1))$;

  /* filling $\mathcal{V}$ with the list of selected pairs Ω_k */

  $\mathcal{V} = \mathtt{Fill}(\mathcal{V},\ \Omega_k,\ d)$; // See Algorithm 4

  /* determining a permutation of Ω_{-k} that fills $\mathcal{V}$ */

  for $\Omega^{\pi}_{-k} \in \pi(\Omega_{-k})$ **do**

    /* filling $\mathcal{V}$ with the candidate pairs Ω^{π}_{-k} */

    $\mathcal{V}_{\pi} = \mathtt{Fill}(\mathcal{V},\ \Omega^{\pi}_{-k},\ d)$; // See Algorithm 4

    if $\mathcal{V}_{\pi}$ is a R-vine **then**

      /* a permutation worked $\rightarrow$ we quit the loop */

      break

  $\mathcal{V} = \mathcal{V}_{\pi}$;

  if $\mathcal{V}$ is not a R-vine **then**

    /* filling did not work $\rightarrow$ permute initial list Ω_k^{init} */

    Get Ω_k by inverting pairs of $(\Omega_k^{init}$;

    $k = k + 1$;

while $\mathcal{V}$ is not a R-vine;

Illustration in $d = 4$ with $d(d-1)/2 = 6$ pairs

- Set of all pairs: $\Omega = \{(i, j) : 1 \leq i, j \leq d\}$
- Selected pairs: Ω_k
- Candidate pairs: $\Omega_{-k} = \Omega \backslash \Omega_k$

Illustration in $d = 4$ with $d(d-1)/2 = 6$ pairs

- Set of all pairs: $\Omega = \{(i,j) : 1 \leq i, j \leq d\}$
- Selected pairs: Ω_k
- Candidate pairs: $\Omega_{-k} = \Omega \backslash \Omega_k$
- $k = 0$: $\Omega_k = \emptyset$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(1,4),(2,4)\}$

k = 0

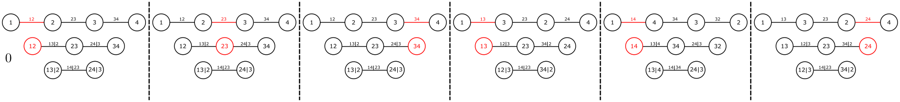


Illustration in $d = 4$ with $d(d-1)/2 = 6$ pairs

- Set of all pairs: $\Omega = \{(i,j) : 1 \leq i,j \leq d\}$
- Selected pairs: Ω_k
- Candidate pairs: $\Omega_{-k} = \Omega \backslash \Omega_k$
- $k = 0$: $\Omega_k = \emptyset$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(1,4),(2,4)\}$

k = 0

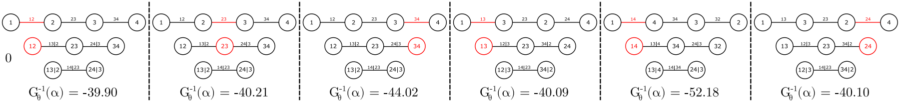


Illustration in $d = 4$ with $d(d-1)/2 = 6$ pairs

- Set of all pairs: $\Omega = \{(i,j) : 1 \leq i,j \leq d\}$
- Selected pairs: Ω_k
- Candidate pairs: $\Omega_{-k} = \Omega \backslash \Omega_k$

- $k = 0$: $\Omega_k = \emptyset$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(1,4),(2,4)\}$
- $k = 1$: $\Omega_k = \{(1,4)\}$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(2,4)\}$

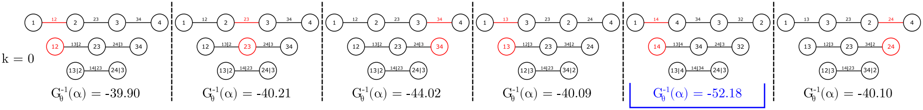


Illustration in $d = 4$ with $d(d-1)/2 = 6$ pairs

- Set of all pairs: $\Omega = \{(i,j) : 1 \leq i,j \leq d\}$
- Selected pairs: Ω_k
- Candidate pairs: $\Omega_{-k} = \Omega \backslash \Omega_k$

- $k = 0$: $\Omega_k = \emptyset$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(1,4),(2,4)\}$
- $k = 1$: $\Omega_k = \{(1,4)\}$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(2,4)\}$

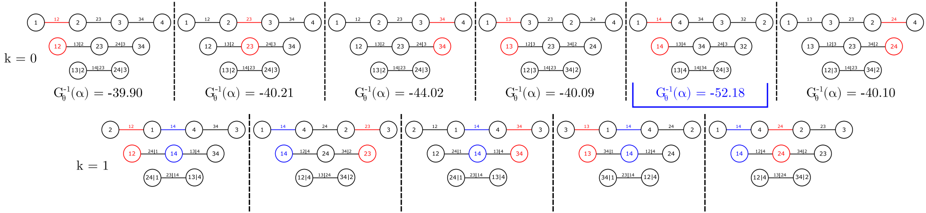


Illustration in $d = 4$ with $d(d-1)/2 = 6$ pairs

- Set of all pairs: $\Omega = \{(i,j) : 1 \leq i,j \leq d\}$
- Selected pairs: Ω_k
- Candidate pairs: $\Omega_{-k} = \Omega \backslash \Omega_k$

- $k = 0$: $\Omega_k = \emptyset$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(1,4),(2,4)\}$
- $k = 1$: $\Omega_k = \{(1,4)\}$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(2,4)\}$
- $k = 2$: $\Omega_k = \{(1,4),(3,4)\}$, $\Omega_{-k} = \{(1,2),(2,3),(1,3),(2,4)\}$

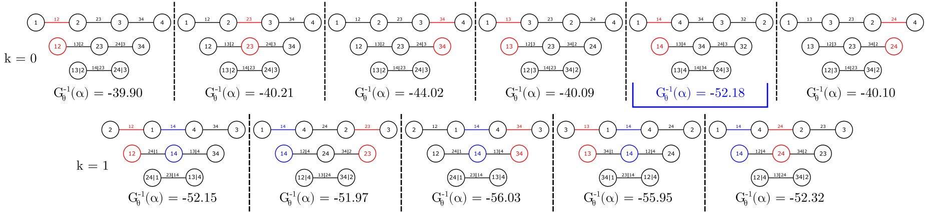


Illustration in $d = 4$ with $d(d-1)/2 = 6$ pairs

- Set of all pairs: $\Omega = \{(i,j) : 1 \leq i,j \leq d\}$
- Selected pairs: Ω_k
- Candidate pairs: $\Omega_{-k} = \Omega \backslash \Omega_k$

- $k = 0$: $\Omega_k = \emptyset$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(1,4),(2,4)\}$
- $k = 1$: $\Omega_k = \{(1,4)\}$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(2,4)\}$
- $k = 2$: $\Omega_k = \{(1,4),(3,4)\}$, $\Omega_{-k} = \{(1,2),(2,3),(1,3),(2,4)\}$
- Results: $\Omega_k = \{(1,4),(3,4),(2,4)\}$

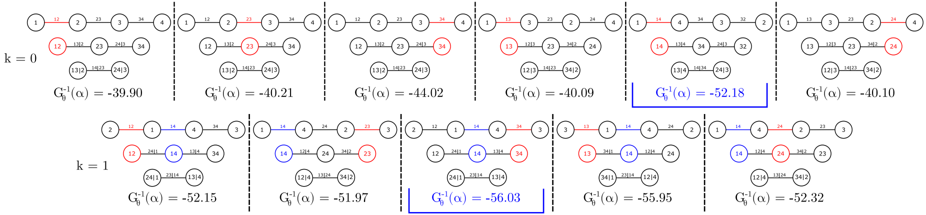
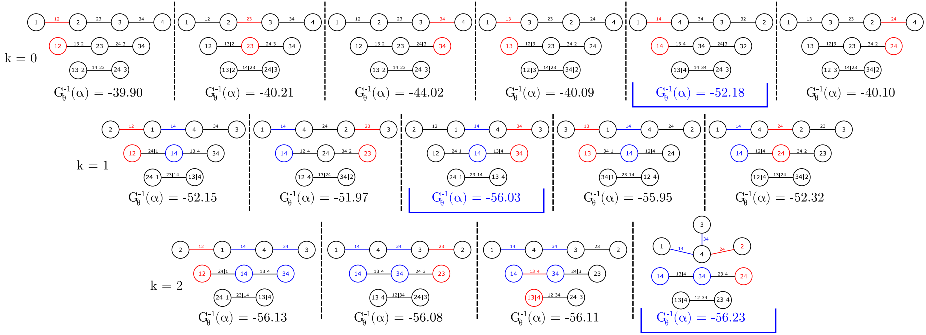


Illustration in $d = 4$ with $d(d-1)/2 = 6$ pairs

- Set of all pairs: $\Omega = \{(i,j) : 1 \leq i,j \leq d\}$
- Selected pairs: Ω_k
- Candidate pairs: $\Omega_{-k} = \Omega \backslash \Omega_k$

- $k = 0$: $\Omega_k = \emptyset$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(1,4),(2,4)\}$
- $k = 1$: $\Omega_k = \{(1,4)\}$, $\Omega_{-k} = \{(1,2),(2,3),(3,4),(1,3),(2,4)\}$
- $k = 2$: $\Omega_k = \{(1,4),(3,4)\}$, $\Omega_{-k} = \{(1,2),(2,3),(1,3),(2,4)\}$
- Results: $\Omega_k = \{(1,4),(3,4),(2,4)\}$



Back to the example in $d = 8$

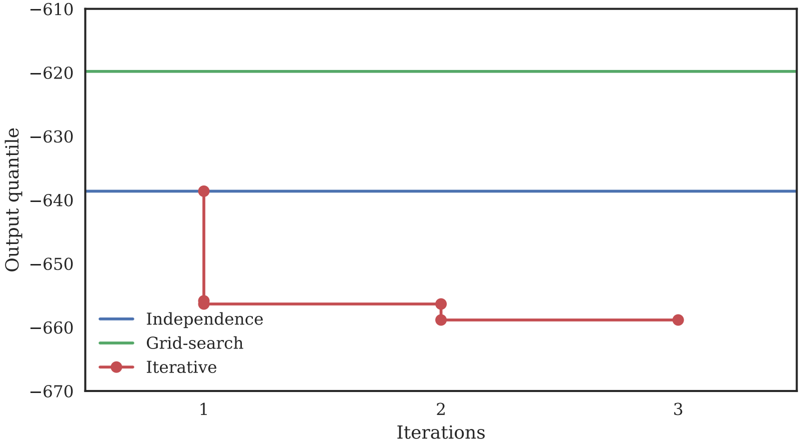
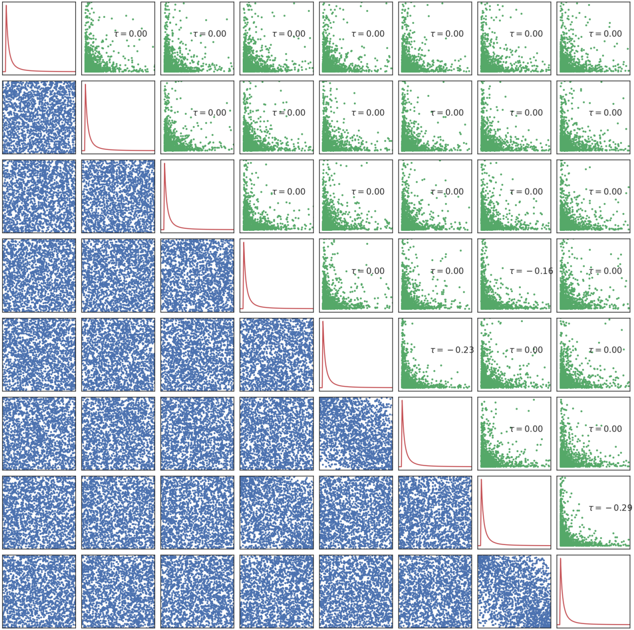
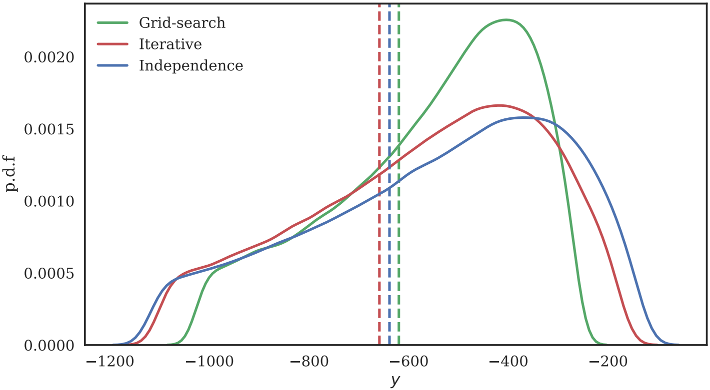


Figure: Resulting output quantiles of each method


$\tau = 0.00$
$\tau = -0.16$
$\tau = -0.23$
$\tau = -0.29$

How it changes the output distributions?



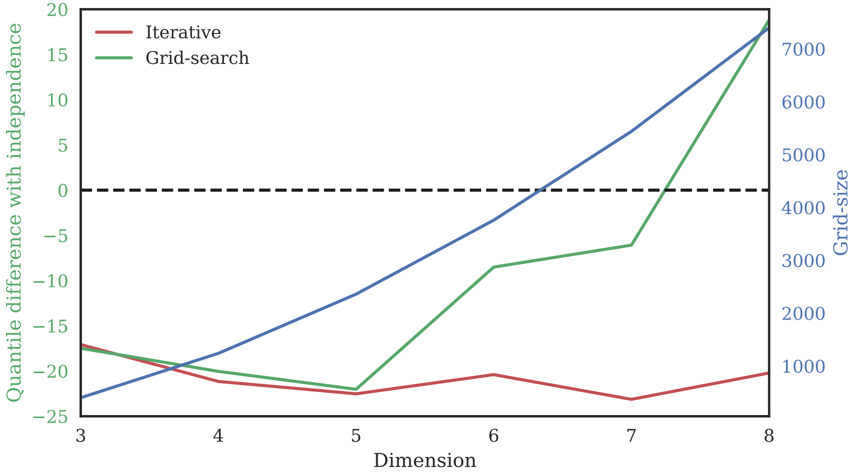
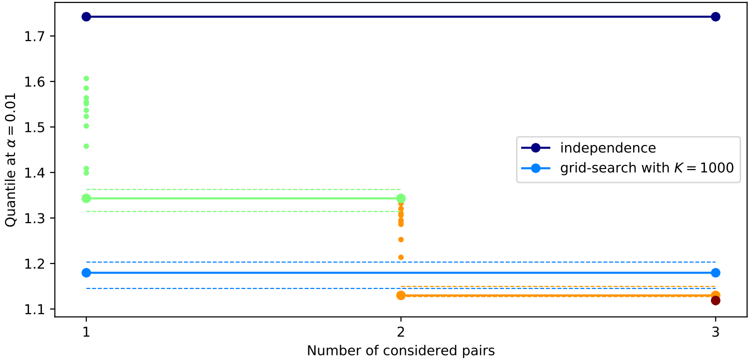


Figure: $Q_{\perp\!\!\!\perp}(\alpha) - Q_{\hat{\theta}_n}(\alpha)$

Results

For $d = 6$, $p = 13$ (2 known dependences) and bounds on $\boldsymbol{\Theta}$. Moreover, $Q_{\boldsymbol{\theta}}(\alpha)$ is monotonic with $\boldsymbol{\theta}$.



Costs for $n = 15000$.

- Grid Search: $n \times K = 15 \times 10^6$
- Sequential: $\sum_{l=1}^{l_{max}} (p - l + 1) \times k(l) = 6.5 \times 10^6$

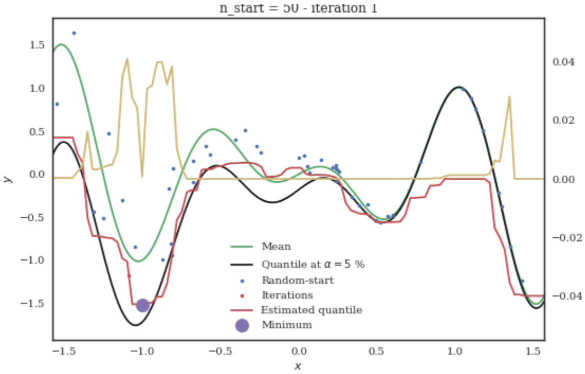
Contents

Conclusion

- Dependencies can have a significant impact on the model output Y and the targeted quantity of interest.
- We used a grid-search strategy to determine a parametric copula that minimizes the output quantile of a model
- We used regular vines to model multivariate copulas
- We proposed a sequential strategy to minimize the quantile by selecting the most influential pairs of variables
- Library `dep-impact` available on
  - PyPi: https://pypi.python.org/pypi/dep-impact
  - Github: https://github.com/nazben/dep-impact

To Infinity and Beyond

- Stay with parametric copulas: Bayesian optimization



- Using non parametric dependence structure?

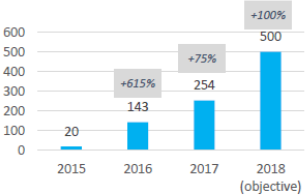
DATA SCIENCE GAME

1 Qualification *Online*

April-June 2018

2 Finals *In Paris*

September 2018



Registered Teams

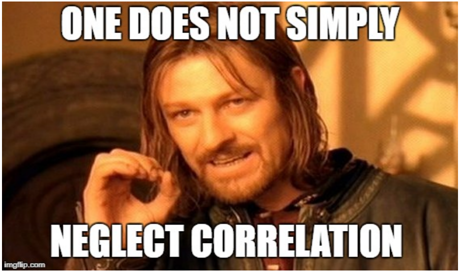
20

Finalist Teams

DATA SCIENCE GAME

2016 Finalists

2017 Finalists



Thank You!