



Detecting and modeling worst-case dependence structures between random inputs of computational reliability models

Nazih Benoumechiara, Bertrand Michel, Philippe Saint-Pierre, Nicolas Bousquet

► To cite this version:

Nazih Benoumechiara, Bertrand Michel, Philippe Saint-Pierre, Nicolas Bousquet. Detecting and modeling worst-case dependence structures between random inputs of computational reliability models. 2018. hal-01779484

HAL Id: hal-01779484

<https://hal.science/hal-01779484>

Preprint submitted on 26 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Detecting and modeling worst-case dependence structures between random inputs of computational reliability models

Nazih Benoumechiara^{1, 4}, Bertrand Michel², Philippe Saint-Pierre³, and Nicolas Bousquet^{1, 5}

¹Sorbonne Université, Laboratoire de Probabilités, Statistique et Modélisation, 75005 Paris, France

²LMJL-UMR6629, École Centrale de Nantes, Nantes, France

³Institut de Mathématiques de Toulouse, UMR 5219 Université de Toulouse UPS IMT, F-31062 Toulouse Cedex 9, France

⁴EDF R&D, Chatou, France

⁵Quantmetry, Paris, France

Abstract

Uncertain information on input parameters of reliability models is usually modeled by considering these parameters as random, and described by marginal distributions and a dependence structure of these variables. In numerous real-world applications, while information is mainly provided by marginal distributions, typically from samples, little is really known on the dependence structure itself. Faced with this problem of incomplete or missing information, risk studies are often conducted by considering independence of input variables, at the risk of including irrelevant situations. This approach is especially used when reliability functions are considered as black-box computational models. Such analyses remain weakened in absence of in-depth model exploration, at the possible price of a strong risk misestimation. Considering the frequent case where the reliability output is a quantile, this article provides a methodology to improve risk assessment, by exploring a set of pessimistic dependencies using a copula-based strategy. In dimension greater than two, a greedy algorithm is provided to build input regular vine copulas reaching a minimum quantile to which a reliability admissible limit value can be compared, by selecting pairwise components of sensitive influence on the result. The strategy is tested over toy models and a real industrial case-study. The results highlight that current approaches can provide non-conservative results, and that a nontrivial dependence structure can be exhibited to define a worst-case scenario.

1 Introduction

Many industrial companies, like energy producers or vehicle and aircraft manufacturers, have to ensure a high level of safety for their facilities or products. In each case, the structural reliability of certain so-called critical components plays an essential role in overall safety. For reasons related to the fact that these critical components are highly reliable, and that real robustness tests can be very expensive or even hardly feasible, structural reliability studies generally use simulation tools [16,36]. The physical phenomenon of interest being reproduced by a numerical model η (roughly speaking, a *computer code*), such

studies are based on the calculation of a reliability indicator based on the comparison of $y = \eta(\mathbf{x})$ and a safety margin, where \mathbf{x} corresponds to a set of input parameters influencing the risk. In the framework of this article, such models are considered as black box and can be explored only by simulation means.

While the problems of checking the validity of η and selecting inputs $\mathbf{x} \in \chi \subseteq \mathbb{R}^d$ are addressed by an increasing methodological corpus [4,13], a perennial issue is the modeling of \mathbf{x} . Differing from the specification of η itself, this input vector is known with uncertainty, either because the number of experiments to estimate is limited, or because some inputs reflect intrinsically variable phenomena [52]. In most cases, these epistemic and aleatory uncertainties are jointly modeled by probability distributions [28]. Consecutively, the reliability indicator is often defined as the probability that y be lower than a threshold (*failure probability*), or a limit quantile for y . This article focuses on this last indicator, which provides an upper or lower bound of the mean effect of the output variable uncertainty.

Therefore the modeling of \mathbf{x} stands on the assessment of a joint probability distribution with support χ , divided between marginal and dependencies features. Though information on each dimension of \mathbf{x} can often be accessible experimentally or using physical or expert knowledge [8], the dependence structure between the component of \mathbf{x} remains generally unknown. Typically, statistical data are only available per dimension, but not available for two or more dimensions simultaneously. For this reason, most of robustness studies are conducted by sampling within independent marginal distributions. Doing so, reliability engineers try to capture input situations that minimize the reliability indicator. Such situations are defined as so-called *worst cases*. However, the assumption of independence between inputs has been severely criticized since the works by [25] and [55], who showed that output failure probabilities of industrial systems can significantly vary and be underestimated if the input dependencies are neglected. More generally, [53,54] showed that tail dependencies between inputs can have major expected effects on the uncertainty analysis results.

Returning to a probabilist framework, and beyond structural reliability, the problem of defining a worst-case scenario by selecting a joint input distribution, from incomplete information, is a topical issue encountered in many fields. In decision-making problems, [49] proposed a general definition of the worst case distribution as the minimizer of an expected cost among a set of possible distributions. More recently, [3] extended this approach to account for incomplete dependence information. These theoretical works, that propose selection rules over the infinite set of all possible joint distributions, remain hard to apply in practice. Recent applied works made use of copulas [43] to model dependencies between stochastic inputs [53,54], following other researchers confronted to similar problems in various fields: finance [12], structural safety [24], environmental sciences [50] or medicine [5]. These studies mainly consider bivariate copulas, which makes theses analysis effective only when two random variables are correlated. Cases where a greater number of variables is involved were explored by [31], who used *vine copulas* to approach complex multidimensional correlation problems in structural reliability. A vine copula is a graphical representation of the *pair-copula construction* (PCC), proposed by [33], which defines a multidimensional dependence structure using conditional bivariate copulas. Various class of vines exist (see [15] for a review), and among them the regular vines (R-vines) introduced by [6,7] are known for their appealing computational properties, while inference on

PCC is usually demanding [18,27].

R-vine parametric copulas seem promising to improve the search for a worst-case dependence between stochastic inputs, while keeping the benefits of a small number of parameters, as favoring inference and conducting simple sensitivity analyses *a posteriori*. To our knowledge, however, no practical methodology has been yet proposed to this end for which the notion of worst case is defined by the minimization of an output quantile. This is the subject of this article. More precisely, the aim of this research is to determine a parametric copula over \mathbf{x} , close to the worst case dependence structure, which is associated to a minimum value of the quantile of the distribution of y . Given a vine structure defined by a parameter vector, the optimization problem involves to conduct empirical quantile estimations for each value of this vector in a finite set of interest (chosen as a grid). The proposed methodology stands on an encompassing greedy algorithm exploring copula structures, which integrates several sub-algorithms of increasing complexity and is based on some simplifying assumptions. These algorithms are made available in the Python library `dep-impact` [9].

The article is therefore organized as follows. Section 2 introduces the framework and studies the consistency of a statistical estimation of the minimum quantile, given an input copula family and a growing sequence of grids. A preliminary study of the influence of the dependence structure, specific to quantile minimization, is conducted in Section 3 as a first application of this statistical optimization. The wider problem of selecting copulas in high-dimensional settings using a sequence of quantile minimization is considered in Section 4. While the choice of R-vines is defended, a sparsity hypothesis is made to diminish the computational burden, according to which only a limited number of pairwise dependencies is influent on the result. A greedy algorithm is proposed to carry out the complete procedure of optimization and modeling. This heuristic is tested in Section 5 over toy examples, using simulation, and a real industrial case-study. The results highlight that worst-case scenarios produced by this algorithm are often bivariate copulas reaching the Fréchet-Hoeffding bounds [22,29] (describing perfect dependence between variables), as it could be expected in monotonic frameworks, but that other nontrivial copulas can be exhibited in alternative situations. Results and avenues for future research are extensively discussed in the last section of this article. We also refer to Appendix A and B for supplementary material on consistency proofs, on R-vine copulas and on R-vine iterative construction.

2 Minimization of the quantile of the output distribution

This section introduces a general framework for the calculation of the minimum quantile of the output distribution of a computational model, when the input distribution can be taken from a large family of distributions, each one corresponding to a particular choice of dependencies between the input variables.

2.1 A general framework for the computation of the minimum quantile

To be general, let us consider a computer code which takes a vector $\mathbf{x} \in \chi \subseteq \mathbb{R}^d$ as an input and produces a real quantity y in output. This code is represented by a deterministic function $\eta : \mathbb{R}^d \rightarrow \mathbb{R}$ such that $\eta(\mathbf{x}) = y$. The sets \mathbb{R} and \mathbb{R}^d are endowed with their Borel sigma algebras and we assume that η is measurable. The general expression of the function η is unknown but for some vector $\mathbf{x} \in \mathbb{R}^d$ it is assumed that the quantity $\eta(\mathbf{x})$ can always be computed. In particular, the derivatives of η , when they exist, are never assumed to be known. Let P_1, \dots, P_d be a fixed family of d distributions, all supported on \mathbb{R} . We introduce the set $\mathcal{D}(P_1, \dots, P_d)$ of all multivariate distributions P on \mathbb{R}^d such that the marginal distributions of P are all equal to the $(P_j)_{j=1\dots d}$. Henceforth, we use the shorter notation \mathcal{D} for $\mathcal{D}(P_1, \dots, P_d)$.

For some $P \in \mathcal{D}$, let G be the cumulative distribution function of the model output. In other terms dG is the push-forward measure of P by η . For $\alpha \in (0, 1)$, let G^{-1} be the α -quantile of the output distribution:

$$G^{-1}(\alpha) := \inf\{y \in \mathbb{R} : G(y) \geq \alpha\}. \quad (1)$$

For the rest of this document, we denote as *output quantile* the α -quantile of the output distribution.

In many real situations, the function η corresponds to a known physical phenomenon. The input variables \mathbf{x} of the model are subject to uncertainties and are quantified by the distribution P . The propagation of these uncertainties leads to the calculation of the output quantile, which defines an overall risk. Due to the difficulties to gather information, it is common to have this distribution incompletely defined and only known through its marginal distributions. Therefore, the set \mathcal{D} corresponds to all the possible distributions that are only known through their marginal distributions $(P_j)_{j=1\dots d}$. In a reliability study, it is essential to avoid underestimating the risk. In such a situation, we might consider a more pessimistic computation of the quantile. We define as the *worst quantile*, the minimum value of the quantile by considering all the possible input distributions $P \in \mathcal{D}$. This conservative approach consists in minimizing $G^{-1}(\alpha)$ over the family \mathcal{D} such as

$$G^{-1\star}(\alpha) := \min_{P \in \mathcal{D}} G^{-1}(\alpha). \quad (2)$$

Since the function η has no closed form in general, it is not possible to give a simple expression of $G^{-1}(\alpha)$ in function of the distribution P , and consequently the minimum $G^{-1\star}(\alpha)$ does not have a simple expression too. In this paper we propose to study a simpler problem than (2), by minimizing $G^{-1}(\alpha)$ over a subset of \mathcal{D} . This subset is a family of distributions $(P_\theta)_{\theta \in \Theta}$ associated to a parametric family of copula $(C_\theta)_{\theta \in \Theta}$, where Θ is a compact set of \mathbb{R}^p and p is the number of copula parameters.

2.2 Copula-based approach

We introduce the real-values random vector $\mathbf{X} = (X_1, \dots, X_d) \in \mathbb{R}^d$ associated to the distribution P_θ . Each component X_j , for $j = 1, \dots, d$, is a real-value random variable with distribution P_j . A copula describes the dependence structure between a group of random

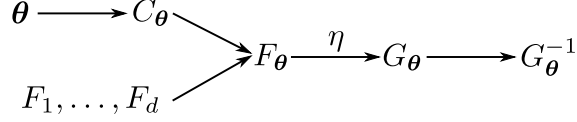


Figure 1: Illustration of the link between the dependence parameter θ and the quantile function G_θ^{-1} . The joint CDF F_θ is obtained using (3) from a copula C_θ and marginal CDF's $(F_j)_{j=1}^d$. The push-forward of F_θ through the model η leads to the CDF G_θ and quantile function G_θ^{-1} of the output distribution.

variables. Formally, a copula is a multidimensional continuous cumulative distribution function (CDF) linking the margins of \mathbf{X} to its joint distribution. Sklar's Theorem [51] states that every joint distribution F_θ associated to the measure P_θ can be written as

$$F_\theta(\mathbf{x}) = C_\theta(F_1(x_1), \dots, F_d(x_d)), \quad (3)$$

with some appropriate d -dimensional copula C_θ with parameter $\theta \in \Theta$ and the marginal CDF's $F_j(x_j) = \mathbb{P}[X_j \leq x_j]$. If all marginal distributions are continuous functions, then there exists a unique copula satisfying

$$C_\theta(u_1, \dots, u_d) = F_\theta(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$$

where $u_j = F_j(x_j)$. For F_θ absolutely continuous with strictly increasing marginal distributions, one can derive (3) to obtain the joint density of \mathbf{X} :

$$f_\theta(\mathbf{x}) = c_\theta(F_1(x_1), \dots, F_d(x_d)) \prod_{j=1}^d f_j(x_j), \quad (4)$$

where c_θ denotes the copula density function of C_θ and $f_j(x_j)$ are the marginal densities of \mathbf{X} . Numerous parametric copula families are available and are based on different dependence structures. Most of these families have bidimensional dependencies, but some can be extended to higher dimensions. However, these extensions have a lack of flexibility and cannot describe all types of dependencies [43]. To overcome these difficulties, tools like vine copulas [32] (described in Section 4) combine bivariate copulas, from different families, to create a multidimensional copula.

Let G_θ and G_θ^{-1} be respectively the CDF and quantile function of the push-forward distribution of P_θ by η (see Figure 1). For a given parametric family of copula $(C_\theta)_{\theta \in \Theta}$ and a given $\alpha \in (0, 1)$, the minimum output quantile for a given copula is defined by

$$G_C^{-1*}(\alpha) := \inf_{\theta \in \Theta} G_\theta^{-1}(\alpha) \quad (5)$$

and if it exists, we consider a minimum

$$\theta_C^* \in \operatorname{argmin}_{\theta \in \Theta} G_\theta^{-1}(\alpha). \quad (6)$$

We call this quantity the minimum quantile parameter or worst dependence structure.

Note that there is no reason for $G_\theta^{-1}(\alpha)$ to be a convex function of θ . The use of gradient descent algorithms is thus not straightforward in this context. Moreover, the gradient of $\theta \rightarrow G_\theta^{-1}$ is unknown and only zero-order optimization methods can be applied to solve (6). For this reason, in the following of this section, we analyze the basic approach which

consists in estimating θ_C^* by approximating Θ with a finite regular grid Θ_N of cardinality N . Therefore, for a given parametric copula $(C_\theta)_{\theta \in \Theta}$ and a given $\alpha \in (0, 1)$, we restrict the problem (6) to

$$\theta_N^* \in \underset{\theta \in \Theta_N}{\operatorname{argmin}} G_\theta^{-1}(\alpha). \quad (7)$$

2.3 Estimation with a grid search strategy

In the restricted problem (7), the greater N , the closer θ_N^* to the minimum θ^* of Θ ; obviously the convergence rate should depend on the regularity of the function η and on the regularity of the quantile function $\theta \mapsto G_\theta^{-1}(\alpha)$. Because η has no closed form, the quantile function $G_\theta^{-1}(\alpha)$ has no explicit expression. The minimizer θ_N^* can be estimated by coupling the simulation of independent and identically distributed (i.i.d) data (Y_1, \dots, Y_n) , defined as realizations of the model output random variable $Y := \eta(\mathbf{X})$ with distribution dG_θ , with a minimization of the empirical quantile over Θ_N .

For θ taking a value over the grid Θ_N , the empirical CDF of Y is defined for any $y \in \mathbb{R}$ by

$$\hat{G}_\theta(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{Y_i \leq y}. \quad (8)$$

The corresponding empirical quantile function $\hat{G}_\theta^{-1}(\alpha)$ is defined as in (1) by replacing G with its empirical estimate. For a given α , the worst quantile on the fixed grid Θ_N is given by

$$\min_{\theta \in \Theta_N} G_\theta^{-1}(\alpha).$$

and can be estimated by replacing the quantile function with its empirical function:

$$\min_{\theta \in \Theta_N} \hat{G}_\theta^{-1}(\alpha). \quad (9)$$

Finally the estimation of the minimum quantile parameter over the grid Θ_N is denoted by

$$\hat{\theta}_N = \underset{\theta \in \Theta_N}{\operatorname{argmin}} \hat{G}_\theta^{-1}(\alpha). \quad (10)$$

The construction of the grid Θ_N can be difficult because Θ can be unbounded (e.g. $\Theta = [1, \infty]$ for a Gumbel copula). To tackle this issue, we chose to construct Θ_N among a normalized space using a concordance measure, which is bounded in $[-1, 1]$ and does not rely on the marginal distributions. We chose the commonly used Kendall rank correlation coefficient (or Kendall's tau) [34] as a concordance measure to create this transitory space. This non-linear coefficient $\tau \in [-1, 1]$ is related to the copula function as follows:

$$\tau = 4 \int_{-1}^1 \int_{-1}^1 C_\theta(u_1, u_2) dC(u_1, u_2) - 1.$$

For many copula families, this relation is much more explicit (see for instance [23]). Therefore, the finite grid is created among $[-1, 1]^p$ and each element of this grid is converted to the copula parameter θ . Moreover, the use of concordance measures gives a normalized expression of the strength of dependencies for all pairs of variables, independently of the used copula families.

The consistency of estimators (9) and (10) is studied in next section, under general regularity and geometric assumptions on η and the functional $\theta \mapsto P_\theta$.

2.4 Consistency of worst quantile-related estimators

In this section, we give consistency results of the estimators $\min_{\theta \in \Theta_N} \widehat{G}_\theta^{-1}(\alpha)$ and $\widehat{\theta}_N$, for a growing sequence of grids on the domain Θ . For easier reading, we skip some definitions needed for our assumptions. Section A in Appendix provides a more complete presentation, including the formal definition of the modulus of increase of the quantile function.

Let α be a fixed value in $(0, 1)$. To approximate Θ , we consider a sequence of finite discrete grids $(\Theta_N)_{N \geq 1}$ on Θ where N is the cardinal of Θ_N and such that

$$\sup_{\theta \in \Theta, \theta' \in \Theta_N} \|\theta - \theta'\|_2 \rightarrow 0 \quad \text{as } N \text{ tends to infinity.} \quad (11)$$

We first introduce technical hypotheses required for the consistency result which are commented further in the text.

Assumption A. For all $\theta \in \Theta$, the distribution P_θ admits a density f_θ for the Lebesgue measure and the copula C_θ admits a density c_θ for the Lebesgue measure on $[0, 1]^d$ such that

$$\begin{aligned} \Theta \times [0, 1]^d &\longrightarrow \mathbb{R} \\ \theta \times (x_1, \dots, x_d) &\longrightarrow c_\theta(x_1, \dots, x_d) \end{aligned}$$

is a continuous function.

Assumption B. For all $\theta \in \Theta$, G_θ is a continuous function.

Assumption C. For all $\theta \in \Theta$, G_θ is strictly increasing and the modulus of increase of G_θ at $G_\theta^{-1}(\alpha)$ is lower bounded by a positive function $\underline{\epsilon}_\Theta$.

Assumption D. There exists an unique $\theta^* \in \Theta$ minimizing $\theta \mapsto G_\theta^{-1}(\alpha)$.

Let $(N_n)_{n \geq 1}$ be a sequence of integers such that $N_n \lesssim n^\beta$ for some $\beta > 0$. For every $n \geq 1$ we consider the grid Θ_{N_n} and for every $\theta \in \Theta_{N_n}$ we compute the empirical quantile $\widehat{G}_\theta^{-1}(\alpha)$ from a sample of n i.i.d variables Y_1, \dots, Y_n with $Y_i = \eta(\mathbf{X}_i)$, where the \mathbf{X}_i 's are i.i.d. random vectors with distribution P_θ . We then introduce the extremum estimator

$$\widehat{\theta} := \widehat{\theta}_{N_n}. \quad (12)$$

Theorem 1. Under Assumptions A, B and C, for all $\varepsilon > 0$ we have

$$P \left(\left| \widehat{G}_{\widehat{\theta}}^{-1}(\alpha) - G_C^{-1}(\alpha) \right| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0. \quad (13)$$

Moreover, if Assumption D is also satisfied, then for all $h > 0$ we have

$$\mathbb{P}[|\widehat{\theta} - \theta^*| > h] \xrightarrow{n \rightarrow \infty} 0$$

(proof given in Appendix A).

It would be possible to provide rates of convergence for this extremum quantile and for θ^* at the price of more technical proofs, by considering also the dimension metric of the

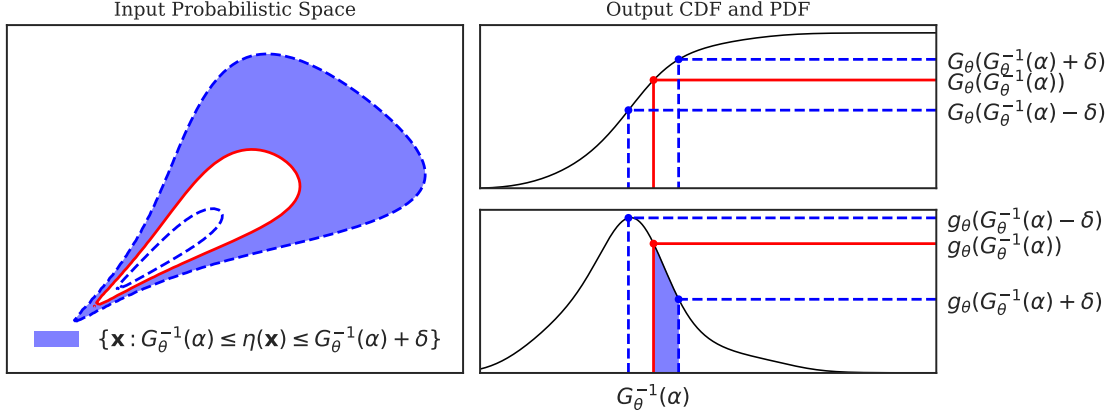


Figure 2: Pre-image (left) and image (right) of a modulus of increase of G_θ at the point $G_\theta^{-1}(\alpha)$ for a deviation $\pm\delta$.

domain Θ and the modulus of increase of the function $\theta \mapsto G_\theta(\alpha)$ (see for instance the proofs of Theorems 1 and 2 in [11] for an illustration of such computations). It would be also possible to derive similar results for alternative extremum quantities. One first example, useful in many applications, would be to estimate some risk probability by determining an extremum $\inf_{\theta \in \Theta} G_\theta(y)$ of the CDF for a fixed y .

This consistency result could also be extended for regular functional of G_θ or G_θ^{-1} , such that

$$\inf_{\theta \in \Theta} \int_{y \geq y_0} G_\theta(y) dy \quad \text{or} \quad \inf_{\theta \in \Theta} \int_{\alpha \geq \alpha_0} G_\theta^{-1}(\alpha) d\alpha,$$

for some fixed values y_0 and α_0 . Extending our results for such quantities is possible essentially because the Dvoretzky-Kiefer-Wolfowitz (DKW) inequality [19], used in the proof, gives a uniform control on the estimation of the CDF and the quantile function.

We now discuss the three first assumptions and provide some geometric and probabilistic interpretations of them. Assumption **A** requires some regularity of the input distribution with respect to θ . This is indeed necessary to locate the minimum of the quantile. Assumption **B** and **C** ensure that the output quantile function G_θ^{-1} has a regular behavior in a neighborhood of the computed quantile $G_\theta^{-1}(\alpha)$. Assumption **B** ensures that the output distribution dG_θ has no Dirac masses whereas Assumption **C** ensures that there is no area of null mass inside the domain of dG_θ .

Figure 2 illustrates Assumption **B** with a possible configuration of the input distribution. For $\theta \in \Theta$ and $\delta > 0$, we consider a small neighborhood $[G_\theta^{-1}(\alpha) - \delta, G_\theta^{-1}(\alpha) + \delta]$ of $G_\theta^{-1}(\alpha)$, and the pre-image of this neighborhood. The two right figures are the CDF G_θ (top) and PDF g_θ (bottom) of the output variable Y for a given θ . The figure at the left hand represents the contours of the pre-image in the input space. The red plain line is the level set $\eta^{-1}(G_\theta^{-1}(\alpha))$ and the dot blue line is the perturbed level set $\eta^{-1}(G_\theta^{-1}(\alpha) \pm \delta)$. The blue area in the right figure corresponds to $[G_\theta^{-1}(\alpha) - \delta, G_\theta^{-1}(\alpha) + \delta]$ and the pre-image of this neighborhood is the blue area in the left figure. Assumption **B** requires that the mass of the blue domain is lower bounded by a positive function $\underline{\varepsilon}_\Theta(\delta)$ that does not depend on θ .

It is possible to give sufficient conditions on the input distribution F_{θ} and on the geometry of the code η to obtain Assumptions **B** and **C**. Using the definition of the modulus of continuity from Equation (A.1) in Appendix, it comes

$$\begin{aligned}\epsilon_{G_{\theta}}(\delta, G_{\theta}^{-1}(\alpha)) &= \max \left[\int_{\{G_{\theta}^{-1}(\alpha) \leq g \leq G_{\theta}^{-1}(\alpha) + \delta\}} f_{\theta}(\mathbf{x}) d\lambda(\mathbf{x}); \int_{\{G_{\theta}^{-1}(\alpha) - \delta \leq g \leq G_{\theta}^{-1}(\alpha)\}} f_{\theta}(\mathbf{x}) d\lambda(\mathbf{x}) \right] \\ &\geq \int_{\{G_{\theta}^{-1}(\alpha) \leq g \leq G_{\theta}^{-1}(\alpha) + \delta\}} f_{\theta}(\mathbf{x}) d\lambda(\mathbf{x})\end{aligned}$$

Assume that the code η is a Lipschitz and differentiable function with no null derivatives almost everywhere in the neighborhood of $G_{\theta}^{-1}(\alpha)$. Then, using the *coarea formula* (see for instance [21], Section 3.4.4, Proposition 3), we find that

$$\epsilon_{G_{\theta}}(\delta, G_{\theta}^{-1}(\alpha)) \geq \int_{G_{\theta}^{-1}(\alpha)}^{G_{\theta}^{-1}(\alpha) + \delta} \left[\int_{\eta^{-1}\{u\}} \frac{f_{\theta}}{\|\nabla \eta\|} d\mathcal{H}^{d-1} \right] du,$$

where \mathcal{H}^{d-1} is the $d - 1$ dimensional Hausdorff measure (see for instance Chapter 2 in [21]). If the copula and the code are such that there exists a constant I such that for any $\theta \in \Theta$ and any u in the support of dG_{θ}

$$\int_{\eta^{-1}\{u\}} f_{\theta} d\mathcal{H}^{d-1} \leq I,$$

then we find that

$$\epsilon_{G_{\theta}}(\delta, G_{\theta}^{-1}(\alpha)) \geq \delta \frac{I}{\|\nabla \eta\|_{\infty}}.$$

Note that $\|\nabla \eta\|_{\infty} < \infty$ since η is assumed to be Lipschitz. We have proved that Assumption **C** is satisfied in this context. Finally, by rewriting again the co-area formula for $G_{\theta}(y)$, we find that Assumption **B** is satisfied as soon as the set of stationary points ($\|\nabla \eta(x)\| = 0$) of all level set $\eta^{-1}\{u\}$ has null mass for the Hausdorff measure.

In conclusion, we see that for smooth copulas, Assumptions **C** and **B** mainly depend on the regularity of the code, by requiring on one side that η does not oscillate too much and on the other side that the set of stationary points does not have a positive mass on the level sets of η .

3 A preliminary study of the copula influence on quantile minimization

This section is dedicated to a preliminary exploration of the influence of copula structure on the behavior of the worst quantile, illustrated with toy examples. Especially, while it could be expected that $G_{\theta}^{-1}(\alpha)$ is a monotonic function with θ , and that the minimum can be reached for a trivial copula (i.e., reaching the Fréchet-Hoeffding bounds). Our experiments show that this behavior is not systematic.

3.1 About the copula choice

One of the most common approaches to model the dependence between random variables is to assume linear correlations feeding a Gaussian copula. In this case, the problem is

reduced by determining the correlation matrix of \mathbf{X} that minimizes $G_{\theta}^{-1}(\alpha)$. However, the positive semi-definite constraint on the correlation matrix makes the exploration difficult and the minimization harder when the problem dimension increases. Moreover, such a Gaussian assumption is very restrictive and is inappropriate for simulating heavy tail dependencies [37]. Still in this elliptical configuration, the t -copulas [17] can be used to counterpart these problems. Nevertheless, tail dependencies are symmetric and with equal strengths for each pair of variables. Another alternative is to consider multivariate Archimedean copulas [40] which are great tools to describe asymmetric tail dependencies. However, only one parameter governs the strength of the dependence among all the pairs, which is very restrictive and not flexible in high dimension. For a same correlation measure between two random variables, multiple copulas can be fitted and lead to a different distribution of Y .

It is clear that the copula choice of \mathbf{X} has a strong impact on the distribution of Y (see for instance [53]). Therefore, various copula types should be tested to determine the most conservative configuration. In the following, we may consider a flexible approach setting by modeling the input multivariate distribution using regular vine copulas (R-vines). The necessary basics of R-vines are introduced in Section 4.1 and detailed in Appendix 1.

3.2 About the monotony of the quantile

For many simple case studies case studies, the worst quantile is reached for perfect dependencies (Fréchet-Hoeffding bounds). More generally, when the function has a monotonic behavior with respect to many variables, it is likely that the minimum output quantile is reached at the boundary of Θ . This phenomenon is observed for various physical systems.

To illustrate this phenomenon, we consider a simplified academic model that simulates the overflow of a river over a dike that protects industrial facilities. The river overflow S is described by

$$S = H_d + C_b - Z_v - H \quad \text{with} \quad H = \left(\frac{Q}{BK_s \sqrt{\frac{Z_m - Z_v}{L}}} \right)^{0.6}, \quad (14)$$

such as, when $S < 0$, a flooding occurs. The involved parameters of (14) are physical characteristics of the river and the dike (e.g., flow rate, height of the dike) which are described by random variables with known marginal distributions. See [30] for more information. For a given risk α , we aim at quantifying the associated overflow's height describe by the α -quantile of S . We extend this model by supposing that the friction (Strickler-Manning) coefficient K_s and the maximal annual flow rate Q are dependent with an unknown dependence structure. To show the influence of a possible correlation between K_s and Q on the quantile of S , we describe their dependence structure with multiple copula families.

The Figure 3 shows the variation of the estimated quantile of S (with a large sample size) in function of the Kendall coefficient τ between K_s and Q for different copula families. We observe different slopes of variation for the different copula families, with lower quantile values for the copulas with heavy tail dependencies (i.e., Clayton, Joe). At independence ($\tau = 0$) and for the counter-monotonic configuration ($\tau = -1$), the quantile values of these

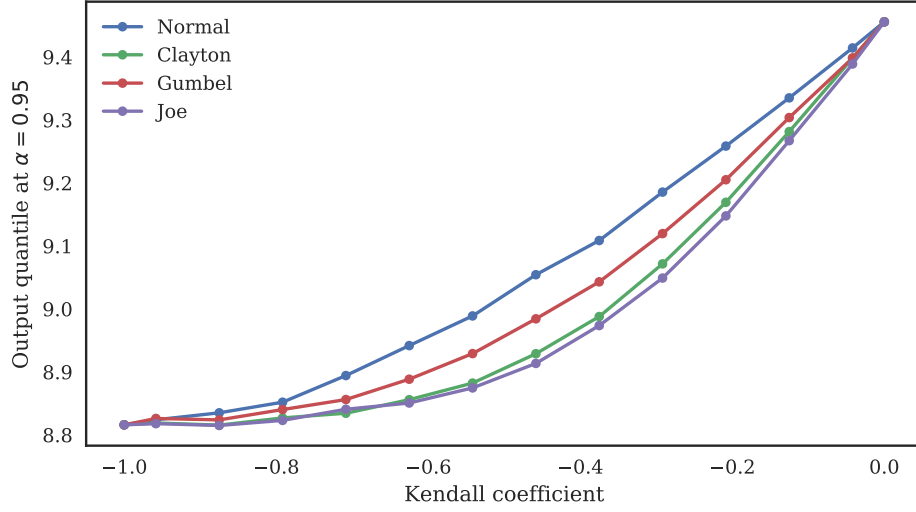


Figure 3: Variation of the quantile of the overflow distribution with the Kendall coefficient τ for $\alpha = 95\%$ and different copula families (Gaussian, Clayton, Gumbel and Joe).

families are obviously equivalent. This variation is slight and the quantile is still above zero, but this shows how the dependencies can influence the results of a reliability problem. This illustration shows that the minimum is reached at the boundary of the exploration space, where the two variables are perfectly correlated.

We can take advantage of this observation to speed up the algorithms presented in the next sections by exploring only the boundaries of Θ . However, assuming that the minimum is reached on the boundary of Θ is a strong assumption that can be unsatisfied in some applications. See Fallacy 3 of [20] for a highlight of this pitfall.

To illustrate this statement, we now give a counter example in the bidimensional setting. We assume uniform marginal distributions for the input such that $X_1 \sim \mathcal{U}(-3, 1)$ and $X_2 \sim \mathcal{U}(-1, 3)$, and we consider the model function

$$\eta(x_1, x_2) = 0.58x_1^2x_2^2 - x_1x_2 - x_1 - x_2. \quad (15)$$

The same experience as for Figure 3 is established and the results are shown in Figure 4. The slopes of the quantile estimations with the Kendall coefficient, for each copula families, are quite different than the results of Figure 3. We observe that the quantile is not monotonic with the Kendall coefficient and its minimum is not reached at the boundary, but for $\tau \approx 0.5$. Moreover, the Gaussian copula is the family that minimizes the most the quantile. It shows that copula with tail dependencies are not always the most penalizing.

A second example, inspired from Example 6 of [20], also shows that the worst case dependence structure in an additive problem is not necessary for perfectly correlated variables. We consider a simple portfolio optimization problem with two random variables X_1 and X_2 with generalized Pareto distributions such as $F_1(x) = F_2(x) = \frac{x}{1+x}$. We aim at maximizing the profit of the portfolio, which is equivalent as minimizing the following additive model function

$$\eta(X_1, X_2) = -(X_1 + X_2). \quad (16)$$

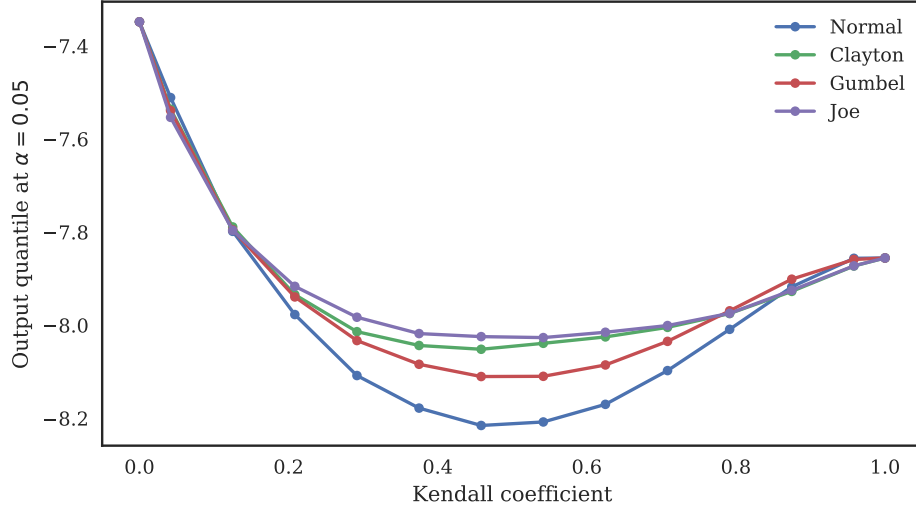


Figure 4: Variation of the output quantile with the Kendall coefficient τ for $\alpha = 5\%$ and different copula families (Gaussian, Clayton, Gumbel and Joe).

We consider the median ($\alpha = 0.5$) of the output as an efficiency measure. The Figure 5 shows the output median in function of the Kendall coefficient τ between X_1 and X_2 . Just like the previous example, we observe a non-monotonic slope of the median in function of τ . The variation can be significant and the minimum is obtained at $\tau \approx 0.53$ for the heavy tail copula families (i.e., Clayton and Joe). The phenomenon can be explained by the marginal distributions of the random variables, which are close Pareto distributions. A large correlation seems to diminish the influence of the tails, which gives a higher quantile value. This explains why the minimum is obtained for a dependence structure other than independence or the perfect dependence.

Therefore, these examples show that the worst quantile can be reached for other configurations than the perfect dependencies.

4 Quantile minimization and choice of penalized correlation structure

This section first provides a rationale for choosing the so-called R-vine structure as a preferential copula structure for modeling the variety of correlations between inputs. Then, the search for a minimum quantile is presented in two times. Subsection 4.2 proposes an exhaustive grid-search algorithm for estimating this quantile when the R-vine copula structure is fixed with a given pair-copula families and indexed by the parameter vector θ . Subsection 4.3 extends this rigid framework by permitting the search of particular sub-copula pairwise structures, such that the minimization be more significant. In each situation, examples are provided to demonstrate the feasibility of the approach.

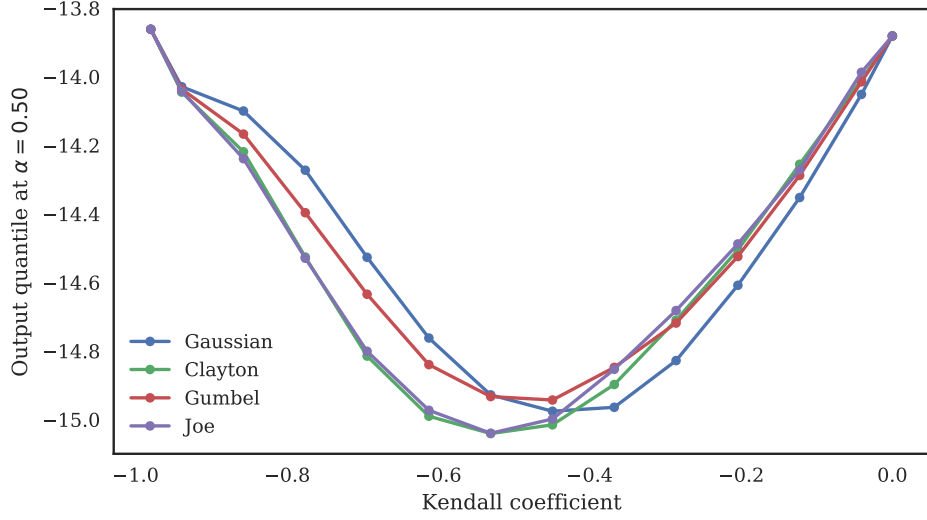


Figure 5: Variation of the portfolio median with the Kendall coefficient τ for different copula families.

4.1 A rationale for R-vine copula structures

Representing multi-dimensional dependence structures in high dimensional settings is a challenging problem. For the following definition, we simplify the expressions by omitting the use of θ : $f = f_\theta$, $F = F_\theta$ and $c = c_\theta$. By recursive conditioning, the joint density can be written as a product of conditioning distributions such as

$$f(x_1, \dots, x_d) = f_1(x_1) \cdot f_{2|1}(x_2|x_1) \cdot f_{3|1,2}(x_3|x_1, x_2) \cdots f_{d|1,2,\dots,d-1}(x_d|x_1, x_2, \dots, x_{d-1}). \quad (17)$$

For clarity reason, we now simplify the expression with $f_{3|1,2} = f_{3|1,2}(x_3|x_1, x_2)$ and so on for other orders. From (4), the conditioning densities of (17) can be rewritten as products of conditioning copula and marginal densities. For example, in a case of three variables and using (17), one possible decomposition of the the joint density can be written as

$$f(x_1, x_2, x_3) = f_1 \cdot f_{2|1} \cdot f_{3|1,2}. \quad (18)$$

Using (4), the reformulation of $f_{3|1,2}$ leads to

$$f_{3|1,2} = \frac{f_{1,3|2}}{f_{1|2}} = \frac{c_{1,3|2} \cdot f_{1|2} \cdot f_{3|2}}{f_{1|2}} = c_{1,3|2} \cdot f_{3|2} \quad (19)$$

where $c_{1,3|2} = c_{1,3|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2))$. By developing $f_{3|1,2}$ in the same way, we find that

$$f_{3|1,2} = c_{1,3|2} \cdot c_{2,3} \cdot f_3. \quad (20)$$

Thus, by replacing the expression of $f_{3|1,2}$ in (18) and doing the same procedure for $f_{2|1}$, the joint density can be written as

$$f(x_1, x_2, x_3) = f_1 \cdot f_2 \cdot f_3 \cdot c_{1,2} \cdot c_{2,3} \cdot c_{1,3|2}. \quad (21)$$

This final representation of the joint density based on pair-copulas has been developed in [33] and is called the pair-copula construction (PCC). The resulting copula represented

by the product of conditional copulas in (21) offers a very flexible way to construct high-dimensional copulas. However, it is not unique; indeed, (17) has numerous decomposition forms and it increases with the dimension.

To describe all such possible constructions in an efficient way, [6,7] introduced the vine models. This graphical tool, based on a sequence of trees, gives a specific way to decompose the multivariate probability distribution. Basically, a vine model is defined by

- a structure of trees which can be represented by a matrix [41],
- a copula family for each pair of the structure,
- a parameter for each pair-copula.

A R-vine is the general construction of a vine model, but particular cases exist such as the D-vines and C-vines, described in Appendix B. Vine models were deeply studied in terms of density estimation and model selection using maximum likelihood [2], sequential estimation [18,35], truncation [1] and Bayesian techniques [26]. Their popularity and well-known flexibility led us to use R-vines in this article, despite the fact that in our context we are looking for a conservative form and not to select the most appropriate form with given data, in absence of correlated observations.

4.2 Estimating a minimum quantile from a given R-vine

4.2.1 Grid-search algorithm

Let $\Omega = \{(i, j) : 1 \leq i, j \leq d\}$ be the set of all the possible pairs of \mathbf{X} , in a d -dimensional problem. The number of pairs p is associated to the size of Ω such as $p = |\Omega| = \binom{d}{2} = d(d-1)/2$. We define \mathcal{V} as the vine structure and we consider fixed copula families for each pair. In this article, we only consider single parameter pair-copulas, such that the parameter $\boldsymbol{\theta}$ is a p -dimensional parameter vector with a definition space $\Theta := \prod_{(i,j) \in \Omega} \Theta_{i,j}$ where $\Theta_{i,j}$ is the parameter space for the pair-copula of the pair (i, j) . However, the methodology can easily be extended to multi-parameter pair-copulas. Note that a pair-copula can be conditioned to other variables, depending on its position in the vine structure \mathcal{V} . Thus, the input distribution $dF_{\boldsymbol{\theta}}(\mathcal{V})$ is defined by the vine structure \mathcal{V} , the copula families and the parameter $\boldsymbol{\theta}$. Also note that the copula parameter $\boldsymbol{\theta}$ is associated to the R-vine structure \mathcal{V} (i.e., $\boldsymbol{\theta} = \boldsymbol{\theta}_{\mathcal{V}}$), see Section 4.2.2. For the sake of clarity, we simplify the notation to $\boldsymbol{\theta}$ only.

The most direct approach to estimate the minimum quantile is the Exhaustive Grid-Search

algorithm, described by the following pseudo-code.

Algorithm 1: Exhaustive Grid-Search algorithm to minimize the output quantile.

Data: A vine structure \mathcal{V} , a fixed grid Θ_N , a sample size n

```

1 for  $\theta \in \Theta_N$  do
2   1. Simulate a sample  $\{\mathbf{X}_i\}_{i=1}^n$  according to  $dF_\theta(\mathcal{V})$ ;
3   2. Evaluate  $\{Y_i = \eta(\mathbf{X}_i)\}_{i=1}^n$ ;
4   3. Compute  $\hat{G}_\theta^{-1}(\alpha)$ : empirical quantile of  $\{Y_i\}_{i=1}^n$ ;

```

Result: $\min_{\theta \in \Theta_N} \hat{G}_\theta^{-1}(\alpha)$

For a given vine structure \mathcal{V} , copula families, a grid Θ_N and a sample size n , three steps are needed for each $\theta \in \Theta_N$. The first step simulates an input sample $\{\mathbf{X}_i\}_{i=1}^n$ according to the distribution $dF_\theta(\mathcal{V})$ for a given sample size n . The second evaluates the sample through the model η . The third estimates the output quantile from the resulting sample $\{Y_i = \eta(\mathbf{X}_i)\}_{i=1}^n$. The minimum quantile is took among the results of each loop.

4.2.2 Influence of the vine structure

Using R-vines, the dependence parameter θ is associated to the vine structure \mathcal{V} . Due to the hierarchy of the vine structure, some pair-copulas are conditioned to other variables and thus for their parameters. As an illustration, let us consider two vine structures with the two following copula densities, with the same simplified expressions as for (21):

$$c_{\mathcal{V}_1}(x_1, x_2, x_3, x_4) = c_{\theta_{1,3}} \cdot c_{\theta_{1,2}} \cdot c_{\theta_{2,4}} \cdot c_{\theta_{2,3|1}} \cdot c_{\theta_{1,4|2}} \cdot c_{\theta_{3,4|1,2}} \quad (22)$$

$$c_{\mathcal{V}_2}(x_1, x_2, x_3, x_4) = c_{\theta_{1,3}} \cdot c_{\theta_{3,4}} \cdot c_{\theta_{2,4}} \cdot c_{\theta_{1,4|3}} \cdot c_{\theta_{2,3|4}} \cdot c_{\theta_{1,2|3,4}}. \quad (23)$$

The difference between these densities is the conditioning of some pairs, the dependence parameters of theses vines are $\theta_{\mathcal{V}_1} = [\theta_{1,2}, \theta_{1,3}, \theta_{1,4|2}, \theta_{2,3|1}, \theta_{2,4}, \theta_{3,4|1,2}]$ and $\theta_{\mathcal{V}_2} = [\theta_{1,2|3,4}, \theta_{1,3}, \theta_{1,4|3}, \theta_{2,3|4}, \theta_{2,4}, \theta_{3,4}]$. Applying the same grid for these two vines may give different results due to the conditioning order from the vine structure. For example, if the pair X_3 - X_4 is very influential on minimizing the output quantile, it would be more difficult to find a minimum with \mathcal{V}_1 than \mathcal{V}_2 due to the conditioning of the pair with X_1 and X_2 in \mathcal{V}_1 . However, if the grid is thin enough, the minimum from these two vines should be equivalent.

To counter this difficulty, one possible option consists in randomly permuting the indexes of the variables and repeating the algorithm several times to visit different vines structures.

4.2.3 Computational cost

For one given R-vine structure and one fixed copula family at each pair, the overall cost of the method is equal to nN . However, as explained in § 2.4, the finite grid Θ_N , should be thin enough to reasonably explore Θ . Therefore, N should increase with the number of dimensions d and more specifically with the number of pairs $p = \binom{d}{2}$. A natural form for N would be to write it as $N = \gamma^p$, where $\gamma \in \mathbb{R}^+$. Thus, the overall cost of the exhaustive

grid-search would be equal to $n\gamma^{\binom{d}{2}}$. The cost is in $\mathcal{O}(\gamma^{d^2})$ which makes the method hardly scalable when the dimension d increases.

4.3 Iterative search for a penalizing R-vine structure: a greedy heuristic based on pairwise copula

4.3.1 Going further in quantile minimization

With Algorithm 1, the previous subsection proposes an exhaustive grid-search strategy to determine a R-vine copula $C_{\hat{\theta}}$ such that the associated output quantile $G_{\hat{\theta}}^{-1}(\alpha)$ be the smallest (and also the most conservative in a structural reliability context). This approach remains however limited in practice since $C_{\hat{\theta}}$ for fixed pair-copula families (e.g., Archimedean or max-stable copulas) and \mathcal{V} which is a member of the set \mathcal{F}_d of all the possible d -dimensional R-vine structure. Intuitively, a more reliable approach to quantile minimization should be based on mixing this estimation method with a selection among all members of the finite set \mathcal{F}_d , as well for the copula families. It is indeed likely that searching within an associative class of copulas like Archimedean ones, allowing modeling dependence in arbitrarily high dimensions, be a too rigid choice for estimating the minimum $G_{\hat{\theta}}^{-1}(\alpha)$.

A minimum quantile can probably be found using a R-vine structure defined by conditional pairwise sub-copulas (according to (21)) that are not part of the same rigid structure. However, a brute force exploration of \mathcal{F}_d would be conducted at an exponential cost increasing with d [42]. If we also consider the large computational cost of an exhaustive grid-search for a large number of dependent variables (as explained in § 4.2), this approach is not feasible in practice for high dimensions.

For this reason, it is proposed to extend Algorithm 1 by a greedy *heuristic* that dynamically selects the most influential correlations between variables while limiting the search to pairwise correlations. Doing so, minimizing the output quantile can be conducted in a reasonable computational time. Therefore the selected d -dimensional vine structure would be filled with independent pair-copulas except for the pairs that are influential on the minimization.

This working limitation, interpreted as a sparsity constraint, is based on the following assumption: it is hypothesized that only few pairs of variables have real influences on the minimization. It is close in spirit to the main assumption of global sensitivity analysis applied to computer models, according to which only a limited number of random variables has a major impact on the output [30,48].

4.3.2 General principle

The method basically relies on an iterative algorithm exploring pairwise correlations between the uniform random variables $U_j = F_j^{-1}(X_j)$ and progressively building a non-trivial R-vine structure, adding one pair of variable to the structure at each iteration. Starting

at step $k = 0$ from the simple independent copula

$$C_{\theta^{(0)}}(u_1, \dots, u_d) = \prod_{j=1}^d u_j,$$

the algorithm finally stops at a given step $k = K$ while proposing a new copula $C_{\theta^{(K)}}$ associated to a R-vine structure \mathcal{V}_K mostly composed of independent pair-copulas.

At each iteration k , we denote by Ω_k the *selected pairs* which are considered non-trivial (non-independent) due to their influence on the quantile minimization. Let $\Omega_{-k} = \Omega \setminus \Omega_k$ be the *candidate pairs*, which were not the remaining pairs, which influence on the minimization is still to be tested and are still considered independent. We also consider \mathbf{B} as a set of candidate copula families. The pseudo-code of Algorithm 2 shows in detail how this iterative exploration and building is conducted. More algorithms in Appendix B.2 described how to construct a vine structure with a given list of indexed pairs of variable.

4.3.3 Example

Consider the four-dimensional ($d = 4$) situation such as $\mathbf{X} = (X_1, \dots, X_4)$ where, for to the sake of simplicity, all marginal distributions of \mathbf{X} are assumed to be uniform on $[0, 1]$. We consider a simple additive model described by

$$\eta(\mathbf{X}) = 30X_1 + 10X_3 + 100X_4. \quad (24)$$

For an additive model and uniform margins, the output quantile is monotonic with the dependence parameters (see Section 3.2) which locates the minimum quantile at the edge of Θ . Thus, Step 1.b. of Algorithm 2 is simplified by considering only Fréchet-Hoeffding copulas in the exploration.

In this illustration we consider $\alpha = 0.1$ and we select $n = 300,000$ large enough in order to have a great quantile estimation and the algorithm stops at $K = 3$. Figure 6 shows, for each iteration k , the $p - k$ vine structures that have been created by the algorithm. The red nodes and edges are the candidate pairs $(i, j) \in \Omega_{-k}$ and the blue nodes and edges are the selected pairs Ω_k . At iteration $k = 0$, the selected pair is $(1, 4)$ with an estimated minimum quantile of -52.18 . At iteration $k = 1$, the second selected pair is $(3, 4)$ with an estimated minimum quantile of -56.03 . At iteration $k = 2$, the third selected pair is $(2, 4)$ with an estimated minimum quantile of -56.23 .

We observe that X_4 appears in all the selected pairs. This is not surprising since X_4 is the most influential variable with the largest coefficient in (24). The algorithm considers D-vines by default, but this is important for the first iterations since most of the pairs are independent. When it is possible, the algorithm creates a vine such as the selected pairs and the candidate pair are in the first trees. For example, the fourth vine at iteration $k = 2$ with the candidate pair $(2, 4)$ shows a R-vine structure that respects the ranking of the listed pairs. However, the third vine at iteration $k = 2$ for the candidate pair $(1, 3)$ along with the selected pairs $\{(1, 4), (3, 4)\}$ could respect the ranking and set all the pairs in the first tree altogether. Thus, using Algorithm 3 in Appendix, a valid vine structure is determined by placing the candidate pair $(1, 3)$ in the next tree.

Algorithm 2: Minimization of the output quantile and estimation of $\theta^{(K)}$ over an increasing family of R-vine structures.

```

1 Initialization:
2 Iteration:  $k = 0$ ;
3 Selected pairs:  $\Omega_0 = \emptyset$ ;
4 Selected families:  $\mathbf{B}_0 = \emptyset$ ;
5 while  $k \leq K$  do
6   Copula parameter space of the selected pairs:  $\Theta_k = \prod_{(i,j) \in \Omega_k} \Theta_{i,j}$ ;
7   1. Explore the set of candidate pairs  $\Omega_{-k}$ ;
8   for  $(i, j) \in \Omega_{-k}$  do
9     a. Create a vine structure  $\mathcal{V}_{(i,j)}$  using the procedure of Section B.2 applied to the list
        $\Omega_k \cup (i, j)$ ;
10    b. Explore the set of candidate families  $\mathbf{B}$ ;
11    for  $\mathcal{B} \in \mathbf{B}$  do
12      Apply Algorithm 1 with the pair-copula families  $\mathcal{B} \cup \mathbf{B}_k$ ;
        (i) Define a  $(k+1)$ -dimensional grid  $\Delta_{i,j}$  of  $\Theta_k \times \Theta_{i,j}$  with cardinality  $N_k$ ;
        (ii) Select the minimum over the grid  $\Delta_{i,j}$ :
            
$$\hat{\theta}_{\mathcal{B}} = \operatorname{argmin}_{\theta_{\mathcal{B}} \in \Delta_{i,j}} \left\{ \hat{G}_{\theta_{\mathcal{B}}}^{-1}(\alpha) \right\}.$$

13    c. Select the minimum among  $\mathbf{B}$ 
            
$$\mathcal{B}_{i,j} = \operatorname{argmin}_{\mathcal{B} \in \mathbf{B}} \left\{ \hat{G}_{\hat{\theta}_{\mathcal{B}}}^{-1}(\alpha) \right\}$$

            
$$\hat{\theta}_{i,j} = \hat{\theta}_{\mathcal{B}_{i,j}}$$

14  2. Select the minimum among  $\Omega_{-k}$ 
            
$$(i, j)^{(k)} = \operatorname{argmin}_{(i,j) \in \Omega_{-k}} \left\{ \hat{G}_{\hat{\theta}_{i,j}}^{-1}(\alpha) \right\},$$

            
$$\mathcal{V}^{(k)} = \mathcal{V}_{(i,j)^{(k)}},$$

            
$$\hat{\theta}^{(k)} = \hat{\theta}_{(i,j)^{(k)}}$$

            
$$\mathcal{B}^{(k)} = \mathcal{B}_{(i,j)^{(k)}}$$

15  3. Check the stopping condition;
16  if  $\hat{G}_{\hat{\theta}^{(k)}}^{-1}(\alpha) \geq \hat{G}_{\hat{\theta}^{(k-1)}}^{-1}(\alpha)$  then
17     $K = k - 1$ ;
18  else
19    Extend the list of selected pairs:  $\Omega_k = \Omega_k \cup (i, j)^{(k)}$  and families:  $\mathbf{B}_k = \mathbf{B}_k \cup \mathcal{B}^{(k)}$ ;
20    if  $k < K$  and computational budget not reached then
21      New iteration:  $k = k + 1$ ;

```

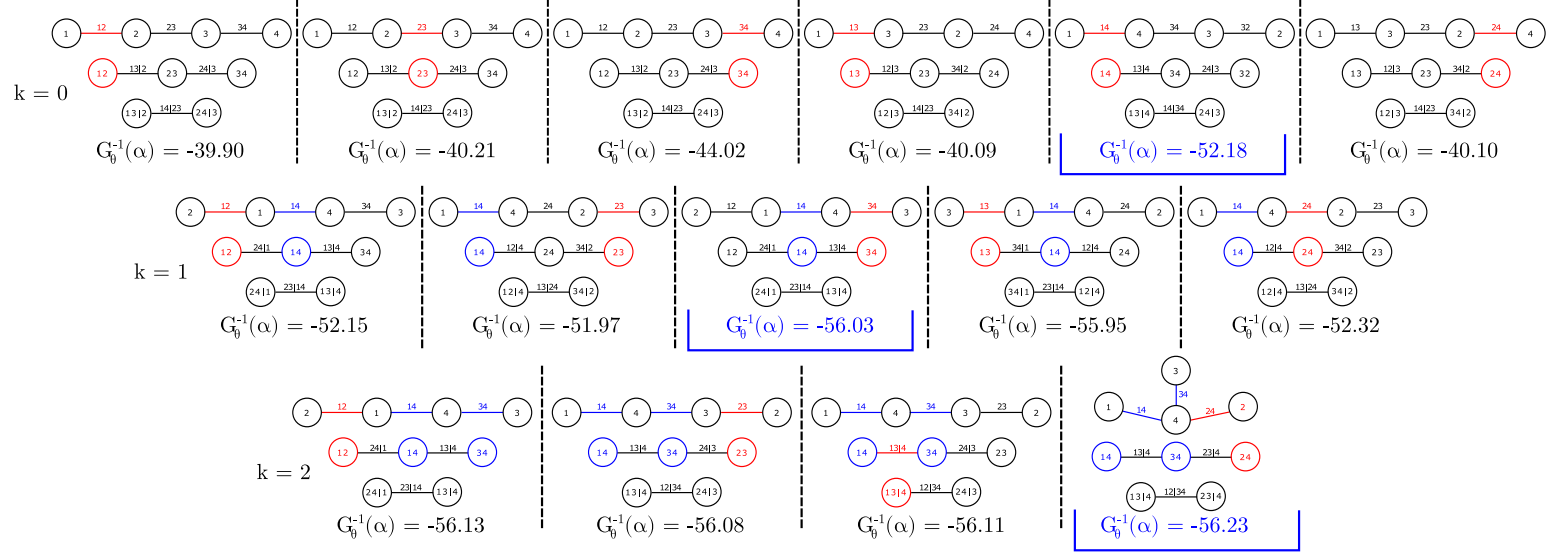


Figure 6: Illustration of the vine structures created during the 3 iterations of the algorithm for the example of Section 4.3.3. The candidate and selected pairs are respectively represented in red and blue. The quantile associated to the selected pair of each iteration is written in blue.

4.3.4 Computational cost

The number of model evaluations is influenced by several characteristics from the probabilistic model and from the algorithm. Let $|\mathbf{B}|$ be the number of family candidates. The total number of runs is

$$N = |\mathbf{B}| \frac{n}{2} \sum_{k=0}^K N_k \times (d(d-1) - 2k). \quad (25)$$

The sum corresponds to the necessary iterations to determine the influential pairs. The maximum possible cost is if all the pairs are equivalently influential (i.e., $K = p = d(d-1)/2$), which would be extremely high. The term nN_k is the cost from the grid-search quantile minimization at step 2. of the algorithm. The greater N_k is and the better the exploration of $\Theta_k \cup \Theta_{i,j}$. Because the dimension of Θ_k increases at each iteration k , it is normal that N_k should also increase with k (e.g. $N_k = \gamma^{\beta k}$, where γ and β are constants). Also, the greater n is and the better the quantile estimations. The second term is the cost from the input dimension d which influences the number of candidate pairs Ω_{-k} at each iteration k .

Extensions can be implemented to reduce the computational cost such as removing from Ω , the pairs that are not sufficiently improving the minimization.

5 Applications

The previously proposed methodology is applied to a toy example and a real industrial case-study. It is worth to mention that these experiments (and future ones) can be conducted again using the Python library `dep-impact` [9], in which are encoded all the procedures of estimation and optimization presented here.

5.1 Numerical example

We pursue and extend the portfolio example considered in Section 3.2 and illustrated on Figure 5. The numerical model η is now defined by the weighted sum

$$Y = \eta(\mathbf{X}) = -\beta \mathbf{X}^T = -\sum_{j=1}^d \beta_j X_j, \quad (26)$$

where the $\beta = (\beta_1, \dots, \beta_d)$ is a vector of constant weights. The margins of the random vector \mathbf{X} follow the same generalized Pareto distribution with scale σ and shape parameter ξ . Note that the bivariate example in Section 3.2 considered $\beta = 1$ and the distribution parameters as $\sigma = 1$ and $\xi = 1$. In the following examples, we aim at minimizing the median ($\alpha = 0.5$) of the output distribution. We chose to fix the marginal distribution's parameters at $\sigma = 10$ and $\xi = 0.75$, and we set the constant vector β to a base-10 increasing sequence such that $\beta = (10^{1/d}, 10^{2/d}, \dots, 10)$. This choice of weights aims to give more influence to the latest components of \mathbf{X} on Y . Thus, some pairs of variables should be more important in the minimization of the output quantile, as required by the sparsity constraint. We also took n large enough to estimate the output quantile with high precision (i.e. $n = 300,000$).

For all these experiments the results from the different methods can be compared.

- Method 1: the grid-search approach with an optimized LHS sampling [39] inside Θ and a random vine structure,
- Method 2: the iterative algorithm with an increasing grid-size of $N_k = 25^*(k+1)^2$.

The Method 1 is established with the same computational budget as Method 2.

5.1.1 Dimension 3

In a three dimensional problem, only three pairs of variables ($p = 3$) are involved in the dependence structure. The sampling size of Θ in Method 1 is set to 400, which is great enough to explore a three dimensional space. The results are displayed on Figure 7: the estimated quantiles from Method 1 (blue dots) with a convex hull (blue dot line) and the quantile at independence (dark point) are provided. It also highlights the minimum estimated quantiles from Methods 1 and 2 which are respectively represented in blue and red points. We also show in green point, the minimum quantile by considering only the Fréchet-Hoeffding bounds. For each minimum, the 95 % bootstrap confidence intervals is displayed in dot lines.

This low dimensional problem confirms the non-monotonic form of the quantile with the dependence parameter, in particular for the variation of the quantile in function of $\tau_{2,3}$. As expected, the pair X_2 - X_3 is more influential on the output quantile due to the large weights on X_2 and X_3 . The minimum values obtained by each method are still lower than the results given by an independent configuration. The minimum using Fréchet-Hoeffding bounds is also provided to show that the minimum is not at the boundary of Θ . Method 1 and 2 have very similar minimum results.

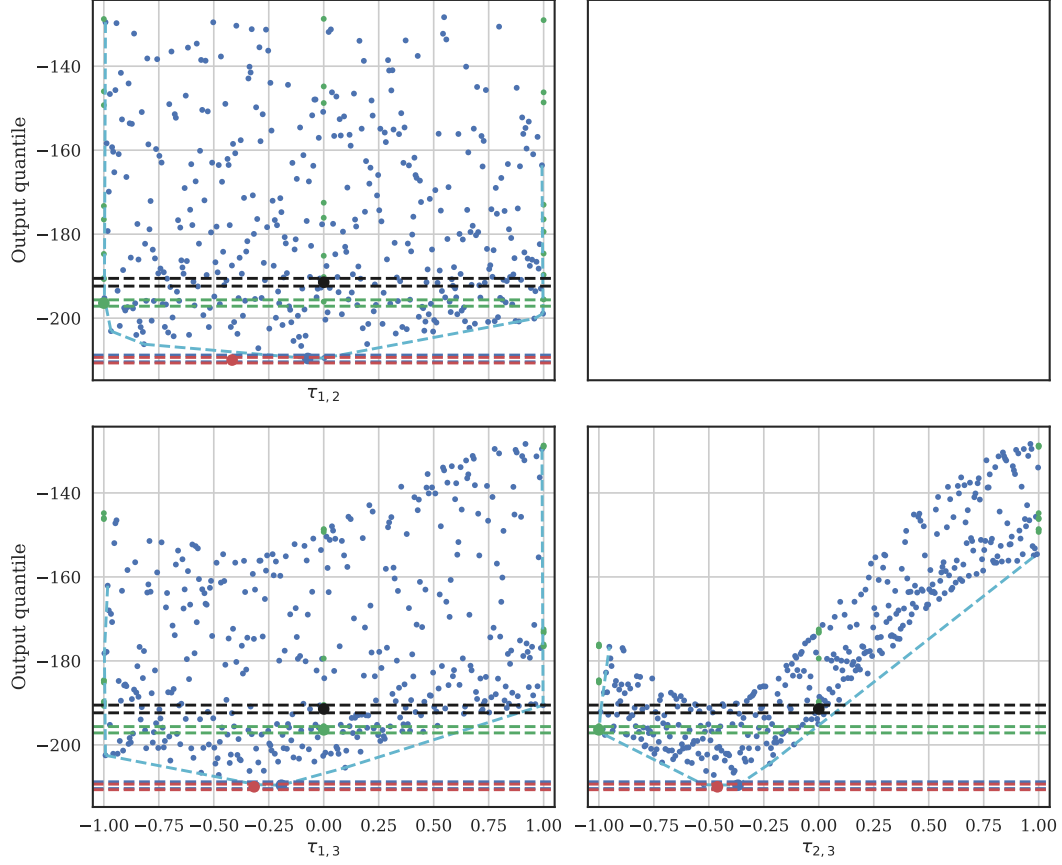


Figure 7: Matrix plot of the output median in function of the Kendall coefficient of each pair. The blue dots represents the estimated quantiles of Method 1. The black point is the quantile at independence and the minimum of Method 1 and 2 are the red and blue points, which are equivalent here. The green point is the the minimum with only Fréchet-Hoeffding bounds. The 95 % bootstrap confidence intervals are displayed in dot lines.

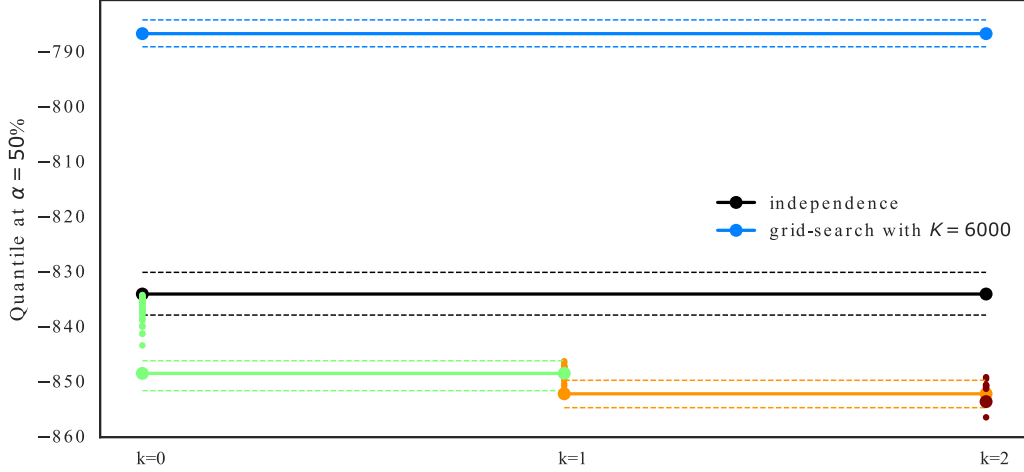


Figure 8: Minimum quantile results with the iteration k of the iterative procedure. The quantile at independence is shown in dark line. The minimum quantiles from Method 1 is shown in blue lines. The other lines and dots colors are the results from Method 2. For each iteration, the small dots are the estimated quantiles of all candidates and the point is the minimum. The 95 % bootstrap confidence intervals are also displayed for the independence and each minimums.

5.1.2 Dimension 10

To illustrate the advantages of the iterative procedure, we now consider $d = 10$. In this example, we chose to only consider a Gaussian family for the set of pair-copula family candidates. The sampling size for the exploration of Θ in Method 1 is set to 6,000. Experimental results are summarized over Figure 8, by displaying the minimum quantiles in function of the iteration k of Method 2. The quantile at independence is shown in dark line, the minimum estimated quantile from Method 1 is shown in blue line and the other lines are the minimum quantiles at each iteration of the algorithm, all with their 95% bootstrap confidence interval. We display at each iteration the minimum quantiles of each candidate pair in small dots.

The minimum result from Method 1 is even higher than the quantile at independence. This is due to the very large number of pairs ($p = 45$) that makes the exploration of Θ extremely difficult. On the other hand, Method 2 (iterative algorithm) is definitely better and significantly decreases the quantile value even at the first iteration (for only one dependent pair). The results are slightly improved with the iterations. We observe at the last iteration that the results from the candidate pairs are slightly higher than the minimum from the previous iteration. It is due to the choice of N_k which does not increase enough with the iterations to correctly explore Θ_k , which also increases with the iterations.

5.1.3 Using multiple pair-copula family candidates

To show the importance of testing multiple copula families, we consider $d = 6$ and three tests of Method 2 (iterative procedure). The Figure 9 shows the minimum from the iterative results using three sets of family candidates: a set of Gaussian and Clayton in red ($\mathbf{B}^1 = \{G, C\}$), Gaussian only in green ($\mathbf{B}^2 = \{G\}$), and Clayton only in yellow

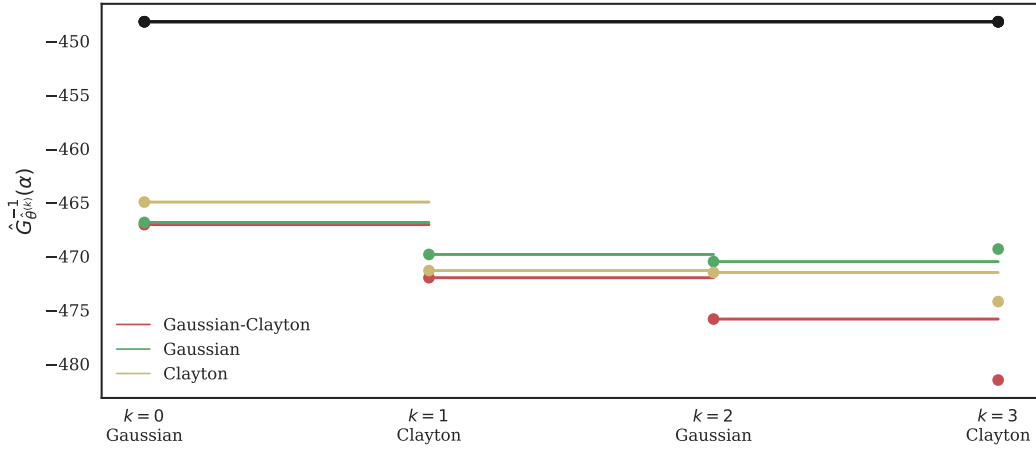


Figure 9: Quantile minimization for different set of family candidates \mathbf{B} . The dark line shows the quantile at independence. The minimum at each iteration for the family candidates sets \mathbf{B}^1 , \mathbf{B}^2 and \mathbf{B}^3 respectively in red, green and yellow.

($\mathbf{B}^3 = \{C\}$). We also display below the iteration number, the selected family for \mathbf{B}^1 .

At iteration $k = 0$, the algorithm with the set \mathbf{B}^1 has selected the Gaussian copula as the selected pair and the result is as expected equivalent as for the set \mathbf{B}^2 . At next iteration, a Clayton copula has been selected for algorithm with the set \mathbf{B}^1 , which slightly improves the minimization compared to the others. The improvement start at iteration $k = 2$ where the Algorithm with the set \mathbf{B}^1 minimizes more the output quantile than the other sets with only one copula family. At the last iteration, the algorithm with set \mathbf{B}^1 selected a mix between Gaussian an Clayton families. This diversity seems to lead to better results than using only one family for every pairs. Testing multiple families is an interesting feature of the algorithm and is something that cannot be feasible for the grid-search approach. However, the cost for \mathbf{B}^1 is twice larger than for the other methods.

5.2 Industrial Application

5.2.1 Context

We consider an industrial component belonging to a production unit. This component must maintain its integrity even in case of an accidental situation. For this reason, it is subject to a justification procedure by regulation authorities, in order to demonstrate its ability to withstand severe operating conditions. This undesirable event consists in the concomitance of three different factors:

- the potential existence of small and undetectable manufacturing defects ;
- the exposition of the structure to an ageing phenomenon harming the material which progressively diminishes its mechanical resistance throughout its lifespan ;
- the occurrence of an accidental event generating severe constraints on the structure.

If combined, these three factors might lead to the initiation of a crack within the structure. Since no failure was observed until now, a structural reliability study should be conducted to check the safety of the structure. To do so a thermal-mechanical code $\eta : \mathbb{R}^d \rightarrow \mathbb{R}^+$ was used, which calculates the ratio between the resistance and the stress acting on the component during a simulated accident. The numerical model depends on parameters affected by uncertainties quantified throughout numerous mechanical tests. Nevertheless, these experiments are mostly established individually and only few experiments involves simultaneously two parameters.

5.2.2 Probabilistic model

For this problem, we introduce $d = 6$ random variables with predefined marginal distributions $(P_j)_{j=1\dots d}$. The dependence structure is however unknown. From the 15 pairs of variables, only the dependencies of two pairs are known: one is independent and the other follows a Gumbel copula with parameter 2.27. Therefore, we consider $p = 13$ pairs of variables with unknown dependencies.

Given expert feedbacks, we restricted the exploration space Θ by defining bounds for each pair of variables $(i, j) \in \Omega$ such that

$$T_{c_{i,j}}(\tau_{i,j}^-) \leq \theta_{i,j} \leq T_{c_{i,j}}(\tau_{i,j}^+),$$

where $\tau_{i,j}^-$ and $\tau_{i,j}^+$ are respectively the upper and lower kendall's correlation coefficient bounds for the dependence of the pair (i, j) and $T_{c_{i,j}}$ is the transformation from Kendall's tau value to the copula parameter for the associated copula $c_{i,j}$. This choice enables to explore only realistic dependence structures. For these experiments we only considered Gaussian copulas.

5.2.3 Results

We consider the quantile at $\alpha = 0.01$ as a quantity of interest. A first experiment is established with the incomplete probability structure: only the two pairs with a known dependence structure and all others at independence. Two other experiments are established: an exhaustive grid-search approach with a given vine structure and an iterative procedure with a maximum budget equivalent to the grid-search. A grid-size of 1000 is chosen with $n = 20,000$.

The results are displayed in Figure 10. and has the same description as Figure 8. The quantile for the incomplete probability structure is approximately at 1.8. The grid-search and the iterative approaches found dependence structures leading to output quantile values close to 1.2 and 1.1 respectively. The minimum quantile from the iterative procedure is slightly lower than the grid-search approach. The problem dimension is not big enough to create make a significant difference between the methods. However, the resulting dependence structure from the iterative method is greatly simplified with only four pairs of variables, in addition to the already known pair.

This result highlights the risk of having an incomplete dependence structure in a reliability

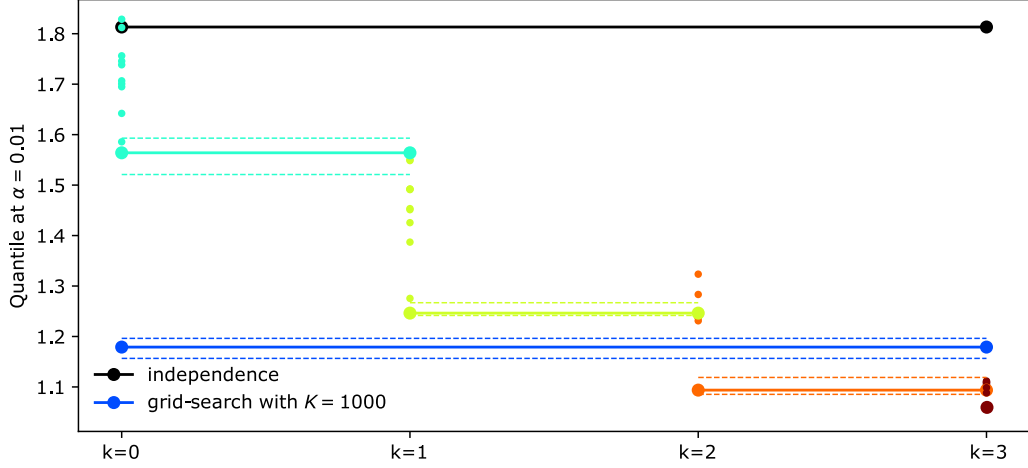


Figure 10: Minimization of the output quantile using a grid-search and the iterative procedure for $\alpha = 1\%$. The description is the same as in Figure 8.

problem. In this application, the critical limit (*safety margin*) of the considered industrial component is 1. With the incomplete distribution of \mathbf{X} , the output quantile is very high compared to the critical limit and states a high reliability of the component. Unfortunately, if we consider worst-case dependence structures, the output quantile is significantly minimized and becomes closer to the critical limit. Thus, if the true dependence structure is close to the obtained worst case dependence structure, the risk of over estimating the output quantile can be important.

6 Conclusion and discussion

Incomplete information on inputs is an issue frequently encountered in structural reliability. Because safety analyses are mostly based on propagating random uncertainties through black-box computer models, the selection of a conservative dependence structure between input components appears as a requirement to define probabilistic worst cases. This article takes a first step towards such a methodology, by proposing a greedy, heuristic algorithm that explores a set of possible dependencies, taking advantage of the pair-copula construction (PCC) of multivariate probability distributions. Results of experiments conducted on toy and a real models illustrate the good behavior of the procedure: in situations where the monotonicity of the considered risk indicator (the output quantile) with respect to the inputs is postulated, a minimum value for the risk indicator is obtained using Fréchet-Hoeffding bounds. Nonetheless, it is possible to exhibit situations where the algorithm detect other and more conservative dependence structures. This first step required a number of hypotheses and approximations that pave the way for future research. Besides, some perspectives arise from additional technical results.

It would be interesting to improve the statistical estimation of the minimum quantile, given a dependence structure, by checking the hypotheses underlying the convergence results of Theorem 1. Checking and relaxing these hypotheses should be conducted in relation with expert knowledge on the computer model η and, possibly, a numerical exploration of its regularity properties. The grid search estimation strategy promoted in Section

2.3 arises from the lack of information about the convexity and the gradient of $\theta \rightarrow G_{\theta}^{-1}(\alpha)$. However, the method remains basic and stochastic recursive algorithms, such as the Robbins-Munro algorithm [47], can be proposed and tested as possibly more powerful (faster) alternatives.

A significant issue, is the computational cost of the exploration of possible dependence structures. Reducing this cost while increasing the completeness of this exploration should be a main concern of future works. Guiding the exploration in the space of conditional bivariate copulas using enriching criteria and possible expert knowledge can facilitate the minimization. The Algorithm 2 can also be improved using nonparametric bootstrap. This would quantify the estimation quality of the selected minimum quantile of each iteration. Note however that a seducing feature of an iterative procedure is the a priori possibility of its adaptation to situations where the computational model η is time-consuming. In such cases, it is likely that Bayesian global optimization methods based on replacing the computer model by a surrogate model (e.g., a kriging-based meta-model) [45] should be explored, keeping in mind that nontrivial conservative correlations – losses of quantile monotonicity – can be due to edge effects (e.g., discontinuities) characterizing the computational model itself.

We noticed in our experiments on real case-studies that expert knowledge remains difficult to incorporate otherwise that using association and concordance measures, mainly since we are lacking of representation tools (e.g., visual) of the properties of multivariate laws that provide intelligible diagnostics. A first step towards the efficient incorporation of expert knowledge could be to automatize the visualization of the obtained vine structures, to simplify judgements about their realism.

Finally, another approach to consider could be to address the optimization problem (2) within the more general framework of optimal transport theory, and to take advantage of the many ongoing works in this area. Indeed, the problem (2) can be seen as a multi-marginal optimal transport problem (see [46] for an overview). When $d = 2$, it corresponds respectively to the classical optimal transport problems of Monge and Kantorovich [57]. However, the multimarginal theory is not as well understood as for the bimarginal case, and developing efficient algorithms for solving this problem remains also a challenging issue [46].

Acknowledgements

The authors express their thanks to Gerard Biau (Sorbonne Université), Bertrand Iooss and Roman Sueur (EDF R&D) for their useful comments and advice during the writing of this article.

References

- [1] Aas, K., Czado, C., and Brechmann, E. C. (2012). Truncated regular vines in high dimensions with application to financial data. *Canadian Journal of Statistics*, 40(1):68–85.

- [2] Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. Insurance, Mathematics and Economics, 44:182–198.
- [3] Agrawal, S., Ding, Y., Saberi, A., and Ye, Y. (2012). Price of correlations in stochastic optimization. Operations Research, 60(1):150–162.
- [4] Bayarri, M., Berger, J., Paulo, R., Sacks, J., Cafeo, J., Cavendish, J., Lin, C., and Tu, J. (2007). A framework for validation of computer models. Technometrics, 49:138–154.
- [5] Beaudoin, D. and Lakhel-Chaieb, L. (2008). Archimedean copula model selection under dependent truncation. Statistics in medicine, 27(22):4440–4454.
- [6] Bedford, T. and Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial intelligence, 32(1):245–268.
- [7] Bedford, T. and Cooke, R. M. (2002). Vines: A new graphical model for dependent random variables. Annals of Statistics, 30(4):1031–1068.
- [8] Bedford, T., Quigley, J., and Walls, L. (2006). Expert elicitation for reliable system design. Statistical Science, 21(4):428–450.
- [9] Benoumechiara, N. (2015–). dep-impact: Uncertainty quantification under incomplete probability information with Python.
- [10] Bobkov, S. and Ledoux, M. (2014). One-dimensional empirical measures, order statistics and kantorovich transport distances. preprint.
- [11] Chazal, F., Massart, P., and Michel, B. (2015). Rates of convergence for robust geometric inference. to appear in Electronic Journal of Statistics, arXiv preprint arXiv:1505.07602.
- [12] Cherubini, U., Luciano, E., and Vecchiato, W. (2004). Copula methods in finance. John Wiley & Sons.
- [13] Council, N. R. (2012). Assessing the Reliability of Complex Models: Mathematical and statistical foundations of Verification, Validation and Uncertainty Quantification. Washington, DC: The National Academies Press.
- [14] Csorgo, M. (1983). Quantile processes with statistical applications, volume 42. SIAM.
- [15] Czado, C. (2010). Pair-copula constructions of multivariate copulas. In Copula theory and its applications, pages 93–109. Springer.
- [16] de Rocquigny, E., Devictor, N., and Tarantola, S. (2008). Uncertainty in industrial practice: a guide to quantitative uncertainty management. John Wiley & Sons.
- [17] Demarta, S. and McNeil, A. J. (2005). The t copula and related copulas. International statistical review, 73(1):111–129.
- [18] Dissmann, J., Brechmann, E. C., Czado, C., and Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. Computational Statistics & Data Analysis, 59:52–69.
- [19] Dvoretzky, A., Kiefer, J., and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. The Annals of Mathematical Statistics, pages 642–669.
- [20] Embrechts, P., McNeil, A., and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. Risk management: value at risk and beyond, pages 176–223.

- [21] Evans, L. C. and Gariepy, R. F. (2015). Measure theory and fine properties of functions. CRC press.
- [22] Frechet, M. (1951). Sur les tableaux de correlation dont les marges sont donnees. AIIm. Univ.
- [23] Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. North American actuarial journal, 2(1):1–25.
- [24] Goda, K. (2010). Statistical modeling of joint probability distribution using copula: Application to peak and permanent displacement seismic demands. Structural Safety, 32(2):112–123.
- [25] Grigoriu, M. and Turkstra, C. (1979). Safety of structural systems with correlated resistances. Applied Mathematical Modelling, 3(2):130–136.
- [26] Gruber, L. and Czado, C. (2015). Sequential bayesian model selection of regular vine copulas. Bayesian Analysis, 10(4):937–963.
- [27] Haff, I. H. (May 9-11, 2016, Oslo, Norway). How to select a good vine. International FocuStat Workshop on Focused Information Criteria and Related Themes.
- [28] Helton, J. (2011). Quantification of margins and uncertainties: conceptual and computational basis. Reliability Engineering and System Safety, 96:976–1013.
- [29] Hoeffding, W. (1940). Scale-invariant correlation theory. Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin, 5(3):181–233.
- [30] Iooss, B. and Lemaître, P. (2015). A review on global sensitivity analysis methods. In Uncertainty Management in Simulation-Optimization of Complex Systems, pages 101–122. Springer.
- [31] Jiang, C., Zhang, W., Han, X., Ni, B., and Song, L. (2015). A vine-copula-based reliability analysis method for structures with multidimensional correlation. Journal of Mechanical Design, 137(6):061405.
- [32] Joe, H. (1994). Multivariate extreme-value distributions with applications to environmental data. Canadian Journal of Statistics, 22(1):47–64.
- [33] Joe, H. (1996). Families of m -variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. Lecture Notes-Monograph Series, pages 120–141.
- [34] Kendall, M. G. (1938). A new measure of rank correlation. Biometrika, 30(1/2):81–93.
- [35] Kurowicka, D. (2011). Optimal truncation of vines. In Kurowicka, D. and Joe, H., editors, Dependence Modeling: Vine Copula Handbook. World Scientific Publishing Co.
- [36] Lemaire, M., Chateauneuf, A., and Mitteau, J.-C. (2010). Structural reliability. Wiley.
- [37] Malevergne, Y., Sornette, D., et al. (2003). Testing the gaussian copula hypothesis for financial assets dependences. Quantitative Finance, 3(4):231–250.
- [38] Massart, P. (1990). The tight constant in the dvoretzky-kiefer-wolfowitz inequality. The Annals of Probability, pages 1269–1283.
- [39] McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, 21(2):239–245.
- [40] McNeil, A. J. and Nešlehová, J. (2009). Multivariate archimedean copulas, d-monotone functions and l-norm symmetric distributions. The Annals of Statistics, pages 3059–3097.
- [41] Morales Nápoles, O. (2010). Bayesian belief nets and vines in aviation safety and other applications. PhD thesis, TU Delft, Delft University of Technology.

- [42] Morales-Nápoles, O. (2011). Counting vines. Dependence modeling: Vine copula handbook, pages 189–218.
- [43] Nelsen, R. B. (2007). An introduction to copulas. Springer Science & Business Media.
- [44] Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. Handbook of econometrics, 4:2111–2245.
- [45] Osborne, M., Garnett, R., and Roberts, S. (2009). Gaussian processes for global optimization. Proceedings of the 3rd International Conference on Learning and Intelligent Optimization.
- [46] Pass, Brendan (2015). Multi-marginal optimal transport: Theory and applications. ESAIM: M2AN, 49(6):1771–1790.
- [47] Robbins, H. and Monro, S. (1951). A stochastic approximation method. The annals of mathematical statistics, pages 400–407.
- [48] Saltelli, A., Chan, K., Scott, E. M., et al. (2000). Sensitivity analysis, volume 1. Wiley New York.
- [49] Scarf, H., Arrow, K., and Karlin, S. (1958). A min-max solution of an inventory problem. Studies in the mathematical theory of inventory and production, 10(2):201.
- [50] Schoelzel, C. and Friederichs, P. (2008). Multivariate non-normally distributed random variables in climate research—introduction to the copula approach. Nonlin. Processes Geophys., 15(5):761–772.
- [51] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges, volume 8. ISUP.
- [52] T., N. and T., A. (2003). Models and model uncertainty in the context of risk analysis. Reliability Engineering and System Safety, 79:309–317.
- [53] Tang, X.-S., Li, D.-Q., Rong, G., Phoon, K.-K., and Zhou, C.-B. (2013). Impact of copula selection on geotechnical reliability under incomplete probability information. Computers and Geotechnics, 49:264–278.
- [54] Tang, X.-S., Li, D.-Q., Zhou, C.-B., and Phoon, K.-K. (2015). Copula-based approaches for evaluating slope reliability under incomplete probability information. Structural Safety, 52:90–99.
- [55] Thoft-Christensen, P. and Sørensen, J. D. (1982). Reliability of structural systems with correlated elements. Applied Mathematical Modelling, 6(3):171–178.
- [56] Van der Vaart, A. W. (2000). Asymptotic statistics, volume 3. Cambridge University Press.
- [57] Villani, C. (2008). Optimal transport: old and new, volume 338. Springer Science & Business Media.

A Proof of the consistency result

The consistency of the estimator $\hat{\theta}$ requires some regularity of the function $\theta \mapsto G_\theta$. This regularity can be also expressed in term of modulus of increase of the function $\theta \mapsto G_\theta^{-1}(\alpha)$, on which some useful definitions and connections with the modulus of continuity are reminded.

A.1 Modulus of increase of a cumulative distribution function

Let us recall that a modulus of continuity is any real-extended valued function $\omega : [0, \infty) \mapsto [0, \infty)$ such that $\lim_{\delta \rightarrow 0} \omega(\delta) = \omega(0) = 0$. The function $f : \mathbb{R} \mapsto \mathbb{R}$ admits ω as modulus of continuity if for any $(x, x') \in \mathbb{R}^2$,

$$|f(x) - f(x')| \leq \omega(|x - x'|).$$

Similarly, for some $x \in \mathbb{R}$, the function f admits ω as a *local* modulus of continuity if for any $x' \in \mathbb{R}^2$,

$$|f(x) - f(x')| \leq \omega(|x - x'|).$$

To control the deviation of the empirical quantile in the proof of Proposition 1 further, we consider the modulus of continuity of the quantile functions $G^{-1} : [0, 1] \rightarrow \mathbb{R}$ where G is a distribution function on \mathbb{R} . The quantile function being an increasing function, the *exact local modulus of continuity* of the quantile function G^{-1} at $\alpha \in (0, 1)$ can be defined as

$$\omega_{G^{-1}}(\epsilon, \alpha) := \max \left(G^{-1}(\alpha + \epsilon) - G^{-1}(\alpha), G^{-1}(\alpha) - G^{-1}(\alpha - \epsilon) \right).$$

In the proof of Proposition 1, we note that the continuity of a quantile function G^{-1} can be connected to the increase of the distribution function G (see also for instance Section A in [10]). Using the fact that the distribution function is increasing, we introduce the *local modulus of increase* of the distribution function ϵ_G at $y = G^{-1}(\alpha) \in \mathbb{R}$ as:

$$\epsilon_G(\delta, y) := \min (G(y + \delta) - G(y), G(y) - G(y - \delta)).$$

A.2 Proofs

The estimator $\hat{\theta}$ defined in (12) is an extremum-estimator (see for instance Section 2.1 of [44]). The main ingredient to prove the consistency of this estimator is the uniform convergence in probability of the families of the empirical quantiles $(\hat{G}_{\theta}^{-1}(\alpha))_{\theta \in \Theta_{K_n}}$ over the family of grids Θ_{K_n} .

Proposition 1. *Let Θ_{K_n} be defined as in Theorem 1. Let assume that **B** and **C** are both satisfied. Then, for all $\epsilon > 0$,*

$$\mathbb{P} \left[\sup_{\theta \in \Theta_{K_n}} |\hat{G}_{\theta}^{-1}(\alpha) - G_{\theta}^{-1}(\alpha)| > \epsilon \right] \xrightarrow{n \rightarrow \infty} 0.$$

Proof of Proposition 1. We first make the connection between the local continuity of the quantile function G_{θ}^{-1} and the local increase of the distribution function G_{θ} . According to Assumption **C**, we have that for any $\epsilon \in (0, \max((1 - \alpha), \alpha))$, for any $\delta > 0$ and for any $\theta \in \Theta$,

$$(*) : \begin{cases} G_{\theta}^{-1}(\alpha + \epsilon) - G_{\theta}^{-1}(\alpha) < \delta \\ G_{\theta}^{-1}(\alpha) - G_{\theta}^{-1}(\alpha - \epsilon) < \delta \end{cases} \implies \begin{cases} G_{\theta}(G_{\theta}^{-1}(\alpha + \epsilon)) < G_{\theta}(G_{\theta}^{-1}(\alpha) + \delta) \\ G_{\theta}(G_{\theta}^{-1}(\alpha) - \delta) < G_{\theta}(G_{\theta}^{-1}(\alpha - \epsilon)) \end{cases}.$$

Next, using basic properties of quantile functions (see for instance point **ii** of Lemma 21.1 in [56]) together with Assumption **B**, we find that

$$G_{\theta}(G_{\theta}^{-1}(\alpha + \epsilon)) = \alpha + \epsilon = G_{\theta}(G_{\theta}^{-1}(\alpha)) + \epsilon$$

and

$$G_{\boldsymbol{\theta}} \left(G_{\boldsymbol{\theta}}^{-1}(\alpha - \epsilon) \right) = \alpha - \epsilon = G_{\boldsymbol{\theta}} \left(G_{\boldsymbol{\theta}}^{-1}(\alpha) \right) - \epsilon.$$

Thus,

$$(*) \implies \begin{cases} G_{\boldsymbol{\theta}} \left(G_{\boldsymbol{\theta}}^{-1}(\alpha) + \delta \right) - G_{\boldsymbol{\theta}} \left(G_{\boldsymbol{\theta}}^{-1}(\alpha) \right) > \epsilon \\ G_{\boldsymbol{\theta}} \left(G_{\boldsymbol{\theta}}^{-1}(\alpha) \right) - G_{\boldsymbol{\theta}} \left(G_{\boldsymbol{\theta}}^{-1}(\alpha) - \delta \right) > \epsilon \end{cases}.$$

We have shown that any $\epsilon \in (0, \max((1 - \alpha), \alpha))$, for any $\delta > 0$ and for any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$,

$$\omega_{G_{\boldsymbol{\theta}}^{-1}}(\epsilon, \alpha) > \delta \implies \epsilon_{G_{\boldsymbol{\theta}}}(\delta, G_{\boldsymbol{\theta}}^{-1}(\alpha)) < \epsilon. \quad (27)$$

We now prove the proposition. For any $n \geq 1$ and any $\epsilon > 0$, we have

$$\begin{aligned} \mathbb{P} \left(\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} |\hat{G}_{\boldsymbol{\theta}}^{-1}(\alpha) - G_{\boldsymbol{\theta}}^{-1}(\alpha)| > \epsilon \right) &= \mathbb{P} \left(\bigcup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} \{ |\hat{G}_{\boldsymbol{\theta}}^{-1}(\alpha) - G_{\boldsymbol{\theta}}^{-1}(\alpha)| > \epsilon \} \right) \\ &\leq \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} P_{\boldsymbol{\theta}} \left(|\hat{G}_{\boldsymbol{\theta}}^{-1}(\alpha) - G_{\boldsymbol{\theta}}^{-1}(\alpha)| > \epsilon \right). \end{aligned} \quad (28)$$

Let ξ_1, \dots, ξ_n be n i.i.d. uniform random variables. The uniform empirical distribution function is defined by

$$\mathbb{U}(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\xi_i \leq t} \quad \text{for } 0 \leq t \leq 1.$$

The inverse uniform empirical distribution function is the function

$$\mathbb{U}_n^{-1}(u) = \inf \{ t \mid \mathbb{G}_n(t) > u \} \quad \text{for } 0 \leq u \leq 1.$$

The empirical distribution function $\hat{G}_{\boldsymbol{\theta}}$ can be rewritten as (see for instance [56]):

$$\hat{G}_{\boldsymbol{\theta}}(y) \stackrel{\mathcal{L}}{=} \mathbb{U}_n(G_{\boldsymbol{\theta}}(y))$$

and as well for the quantile function,

$$\hat{G}_{\boldsymbol{\theta}}^{-1}(\alpha) \stackrel{\mathcal{L}}{=} G_{\boldsymbol{\theta}}^{-1}(\mathbb{U}_n^{-1}(\alpha)).$$

From Inequality (28), we obtain

$$\sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} P_{\boldsymbol{\theta}} \left(|\hat{G}_{\boldsymbol{\theta}}^{-1}(\alpha) - G_{\boldsymbol{\theta}}^{-1}(\alpha)| > \epsilon \right) = \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} P_{\boldsymbol{\theta}} \left(|G_{\boldsymbol{\theta}}^{-1}(\mathbb{U}_n^{-1}(\alpha)) - G_{\boldsymbol{\theta}}^{-1}(\alpha)| > \epsilon \right) \quad (29)$$

By definition of the local modulus of continuity $\omega_{G_{\boldsymbol{\theta}}^{-1}}$ of the quantile function $G_{\boldsymbol{\theta}}^{-1}$ at α , we have

$$|G_{\boldsymbol{\theta}}^{-1}(\mathbb{U}_n^{-1}(\alpha)) - G_{\boldsymbol{\theta}}^{-1}(\alpha)| \leq \omega_{G_{\boldsymbol{\theta}}^{-1}}(|\mathbb{U}_n^{-1}(\alpha) - \alpha|, \alpha). \quad (30)$$

Therefore, by replacing (30) in (29) and using (27), we obtain

$$\begin{aligned} \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} P_{\boldsymbol{\theta}} \left(|G_{\boldsymbol{\theta}}^{-1}(\mathbb{U}_n^{-1}(\alpha)) - G_{\boldsymbol{\theta}}^{-1}(\alpha)| > \epsilon \right) &\leq \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} P_{\boldsymbol{\theta}} \left(\omega_{G_{\boldsymbol{\theta}}^{-1}}(|\mathbb{U}_n^{-1}(\alpha) - \alpha|, \alpha) > \epsilon \right) \\ &\leq \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} P_{\boldsymbol{\theta}} \left(\epsilon_{G_{\boldsymbol{\theta}}}(\epsilon, G_{\boldsymbol{\theta}}^{-1}(\alpha)) < |\mathbb{U}_n^{-1}(\alpha) - \alpha| \right). \end{aligned}$$

Assumption **C** then yields

$$\mathbb{P} \left(\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} |\hat{G}_{\boldsymbol{\theta}}^{-1}(\alpha) - G_{\boldsymbol{\theta}}^{-1}(\alpha)| > \epsilon \right) \leq K_n \mathbb{P} \left(|\mathbb{U}_n^{-1}(\alpha) - \alpha| > \underline{\epsilon}_{\boldsymbol{\Theta}}(\epsilon) \right). \quad (31)$$

The DKW inequality [19] gives an upper bound of the probability of an uniform empirical process $\{|\mathbb{U}_n(\alpha) - \alpha|\}$. As well for an uniform empirical quantile process $\{|\mathbb{U}_n^{-1}(\alpha) - \alpha|\}$ (see for example Section 1.4.1 of [14]), such as $\forall \lambda > 0$:

$$\mathbb{P} \left(\sup_{\alpha \in [0,1]} |\mathbb{U}_n^{-1}(\alpha) - \alpha| \geq \lambda \right) \leq C \exp(-2n\lambda^2).$$

Moreover, [38] proved that one can take $C = 2$. Therefore, Equation (31) can be bounded using the DKW and

$$\mathbb{P} \left(\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} |\hat{G}_{\boldsymbol{\theta}}^{-1}(\alpha) - G_{\boldsymbol{\theta}}^{-1}(\alpha)| > \epsilon \right) \leq 2K_n \exp \left[-2n\underline{\epsilon}_{\boldsymbol{\Theta}}^2(\epsilon) \right] \xrightarrow{n \rightarrow \infty} 0$$

since $K_N \lesssim n^\beta$. □

A second requirement to get the consistency of the extremum estimator is the regularity of $\boldsymbol{\theta} \mapsto G_{\boldsymbol{\theta}}^{-1}(\alpha)$. This is shown in the next proposition.

Proposition 2. *Under Assumptions **A**, **B** and **C**, the function*

$$\begin{aligned} \boldsymbol{\Theta} &\longrightarrow \overline{Im(\eta)} \\ \boldsymbol{\theta} &\longrightarrow G_{\boldsymbol{\theta}}^{-1}(\alpha) \end{aligned}$$

is continuous in $\boldsymbol{\theta}$ over $\boldsymbol{\Theta}$.

Proof of Proposition 2. According to Assumption **A**, for any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, the distribution $P_{\boldsymbol{\theta}}$ admits a density function $f_{\boldsymbol{\theta}}$ with respect to the Lebesgue measure on \mathbb{R}^d such that

$$f_{\boldsymbol{\theta}}(x_1, \dots, x_d) = c_{\boldsymbol{\theta}} (F_1(x_1, \dots, x_d)) f_1(x_1) \dots f_d(x_d),$$

where f_j is the marginal density function of X_j , for $j = 1, \dots, d$ and the Lebesgue measure on \mathbb{R} . Moreover, for any $\mathbf{x} \in \mathbb{R}^d$, the function $\boldsymbol{\theta} \rightarrow f_{\boldsymbol{\theta}}(\mathbf{x})$ is continuous in $\boldsymbol{\theta}$ over $\boldsymbol{\Theta}$.

The domain $\boldsymbol{\Theta} \times [0, 1]^p$ is a compact set and according to Assumption **A**, there exists a constant \bar{c} such that $\forall (\boldsymbol{\theta}, \mathbf{u}) \in \boldsymbol{\Theta} \times [0, 1]^d$, $c_{\boldsymbol{\theta}}(\mathbf{u}) \leq \bar{c}$. Consequently, we have

$$|f_{\boldsymbol{\theta}}(x_1, \dots, x_d)| \leq \bar{c} \prod_{i=1}^d f_i(x_i). \quad (32)$$

For $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and for any $h > 0$, we denote $y_h = G_{\boldsymbol{\theta}+h}^{-1}(\alpha)$. According to Assumption **B** we have $\alpha = G_{\boldsymbol{\theta}+h}(y_h)$ and thus,

$$\begin{aligned} G_{\boldsymbol{\theta}}^{-1}(\alpha) - G_{\boldsymbol{\theta}+h}^{-1}(\alpha) &= G_{\boldsymbol{\theta}}^{-1}(\alpha) - y_h \\ &= G_{\boldsymbol{\theta}}^{-1}(G_{\boldsymbol{\theta}+h}(y_h)) - G_{\boldsymbol{\theta}}^{-1}(G_{\boldsymbol{\theta}}(y_h)) \end{aligned} \quad (33)$$

Now, using Assumption **C**, we have that $G_{\boldsymbol{\theta}}$ is strictly increasing in the neighborhood of $G_{\boldsymbol{\theta}}^{-1}(\alpha)$ and thus $G_{\boldsymbol{\theta}}^{-1}$ is continuous in the neighborhood of α . Note that

$$\begin{aligned} |G_{\boldsymbol{\theta}+h}(y_h) - G_{\boldsymbol{\theta}}(y_h)| &= \left| \int_{\mathbb{R}^d} \mathbb{1}_{\eta(\mathbf{x}) \leq y_h} dF_{\boldsymbol{\theta}+h}(\mathbf{x}) - \int_{\mathbb{R}^d} \mathbb{1}_{\eta(\mathbf{x}) \leq y_h} dF_{\boldsymbol{\theta}}(\mathbf{x}) \right| \\ &= \left| \int_{\mathbb{R}^d} [f_{\boldsymbol{\theta}+h}(\mathbf{x}) - f_{\boldsymbol{\theta}}(\mathbf{x})] \mathbb{1}_{\eta(\mathbf{x}) \leq y_h} d\lambda(\mathbf{x}) \right| \\ &\leq \int_{\mathbb{R}^d} |f_{\boldsymbol{\theta}+h}(\mathbf{x}) - f_{\boldsymbol{\theta}}(\mathbf{x})| d\lambda(\mathbf{x}) \end{aligned}$$

We then apply a standard dominated convergence theorem using (32) to get that

$$G_{\boldsymbol{\theta}+h}(y_h) - G_{\boldsymbol{\theta}}(y_h) \xrightarrow{h \rightarrow 0} 0.$$

This, with (33) and with the continuity of $\boldsymbol{\theta} \mapsto G_{\boldsymbol{\theta}}$, shows that

$$G_{\boldsymbol{\theta}}^{-1}(\alpha) - G_{\boldsymbol{\theta}+h}^{-1}(\alpha) \xrightarrow{h \rightarrow 0} 0.$$

□

We are now in position to prove Theorem 1.

Proof of Theorem 1. Under Assumptions **B** and **C**, Proposition 1 directly gives that for any $\varepsilon > 0$,

$$P \left(\left| \inf_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} \widehat{G}_{\boldsymbol{\theta}}^{-1}(\alpha) - \inf_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} G_{\boldsymbol{\theta}}^{-1}(\alpha) \right| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0$$

which means that

$$P \left(\left| \widehat{G}_{\hat{\boldsymbol{\theta}}}^{-1}(\alpha) - \inf_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} G_{\boldsymbol{\theta}}^{-1}(\alpha) \right| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0. \quad (34)$$

If Assumption **A** is also satisfied, Proposition 2 together with (11) give that $\inf_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} G_{\boldsymbol{\theta}}^{-1}(\alpha)$ tends to $\inf_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} G_{\boldsymbol{\theta}}^{-1}(\alpha)$ as n tends to infinity. Thus

$$\inf_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{K_n}} G_{\boldsymbol{\theta}}^{-1}(\alpha) \xrightarrow{n \rightarrow \infty} G_C^{-1*}(\alpha) = G_{\boldsymbol{\theta}_C^*}^{-1}(\alpha) \quad (35)$$

We then derive (13) from (34) and (35).

We now assume that Assumption **D** is also satisfied. Let $\boldsymbol{\theta}^*$ be the unique minimizer of $\boldsymbol{\theta} \mapsto G_{\boldsymbol{\theta}}^{-1}(\alpha)$. Let $h > 0$ such that $B(\boldsymbol{\theta}^*, h)^c := \{\boldsymbol{\theta} \in \boldsymbol{\Theta} : \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2 \geq \varepsilon\}$ is not empty. According to Proposition 2 and using the fact that $\boldsymbol{\Theta}$ is compact, we have

$$\sup_{\boldsymbol{\theta} \in B(\boldsymbol{\theta}^*, h)^c} |G_{\boldsymbol{\theta}}^{-1}(\alpha) - G_{\boldsymbol{\theta}^*}^{-1}(\alpha)| > 0. \quad (36)$$

Consequently, for any $\forall h > 0$ small enough, there exists $\epsilon > 0$ such that

$$|G_{\boldsymbol{\theta}}^{-1}(\alpha) - G_{\boldsymbol{\theta}^*}^{-1}(\alpha)| \leq \epsilon \implies |\boldsymbol{\theta} - \boldsymbol{\theta}^*| < h \quad (37)$$

Let $h > 0$ and take ϵ such that (37) is satisfied for h . According to Proposition 1, $\widehat{G}_{\hat{\boldsymbol{\theta}}}^{-1}(\alpha) - G_{\hat{\boldsymbol{\theta}}}^{-1}(\alpha)$ tends to zero in probability as n tends to infinity. This, with (13), shows that

$$P \left(\left| G_{\hat{\boldsymbol{\theta}}}^{-1}(\alpha) - G_{\boldsymbol{\theta}^*}^{-1}(\alpha) \right| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0.$$

We conclude using (37). □

B Vine copulas

B.1 Definition

A vine model describes a d -dimensional pair-copula construction (PCC) and is a sequence of linked trees where the nodes and edges correspond to the $d(d-1)/2$ pair-copulas. According to Definition 1 from [6], a vine structure is composed of $d-1$ trees T_1, \dots, T_{d-1} with several conditions.

Definition 1 (R-vine). *The sequence $\mathcal{V} = (T_1, \dots, T_{d-1})$ is an R-vine on n elements if*

1. T_1 is a tree with nodes $N_1 = \{1, \dots, d\}$ and a set of edges denoted E_1 .
2. For $i = 2, \dots, d-1$, T_i is a tree with nodes $N_i = E_{i-1}$ and edges set E_i .
3. For $i = 2, \dots, d-1$ and $\{a, b\} \in E_i$ with $a = \{a_1, a_2\}$ and $b = \{b_1, b_2\}$ it must hold that $\#(a \cap b) = 1$ (proximity condition).

Each tree T_i is composed of $d-i+1$ nodes which are linked by $d-i$ edges for $i = 1, \dots, d-1$. A node in a tree T_i must be an edge in the tree T_{i-1} , for $i = 2, \dots, d-1$. Two nodes in a tree T_i can be joined if their respective edges in tree T_{i-1} share a common node, for $i = 2, \dots, d-1$. The proximity condition, suggests that two nodes connected by an edge should share one variable from the conditioned set. The *conditioning set* and *conditioned set* are defined in Definition 2 along with the *complete union*. The complete union of an edge e is a set of all unique variables contained in e .

Definition 2 (Complete union, conditioning and conditioned sets of an edge). *Let A_e be the complete union of an edge $e = \{a, b\} \in E_k$ in a tree T_k of a regular vine \mathcal{V} ,*

$$A_e = \{v \in N_1 | \exists e_i \in E_i, i = 1, \dots, k-1, \text{ such that } v \in e_i \in \dots \in e_{k-1} \in e\}.$$

The conditioning set associated with edge $e = \{a, b\}$ is $D(e) := A_a \cap A_b$ and the conditioned sets associated with edge e are $i(e) := A_a \setminus D(e)$ and $j(e) := A_b \setminus D(e)$. Here, $A \setminus B := A \cap B^c$ and B^c is the complement of B .

The conditioned and conditioning sets of an edge $e = \{a, b\}$ are respectively the symmetric difference and the intersection of the complete unions of a and b . The conditioned and conditioning sets of all edges of \mathcal{V} are collected in a set called *constraint set*. Each element of this set is composed of a pair of indices corresponding to the conditioned set and a set containing indices corresponding to the conditioning set, as shown in Definition 3.

Definition 3 (Constraint set). *The constrain set for \mathcal{V} is a set:*

$$C\mathcal{V} = \{(\{i(e), j(e)\}, D_e) | e \in E_i, e = \{a, b\}, i = 1, \dots, d-1\}$$

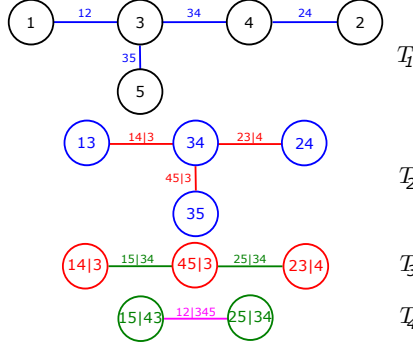


Figure 11: R-vine structure for $d = 5$.

The pair-copula in the first tree characterize pairwise unconditional dependencies, while the pair-copula in higher order trees model the conditional dependency between two variables given a set of variables. The number of conditioning variables grows with the tree order. Note that a PCC where all trees have a path-like structure define the D-vine subclass while the star-like structures correspond to C-vine subclass. All other vine structures are called regular vines (R-vines) [6].

We illustrate the concept of a vine model with a $d = 5$ dimensional example. For clarity reasons, we use the same simplifications as in Section 4.2 which consider for instance $f_1 = f_1(x_1)$, $f_2 = f_2(x_2)$ and so on for higher order and conditioning. One possible PCC can be written for this 5-dimensional configuration:

$$\begin{aligned}
 f(x_1, x_2, x_3, x_4, x_5) &= f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot f_5 \text{ (margins)} \\
 &\quad (\text{unconditional pairs}) \times c_{12} \cdot c_{35} \cdot c_{34} \cdot c_{24} \\
 &\quad (\text{1st conditional pair}) \times c_{14|3} \cdot c_{23|4} \cdot c_{45|3} \\
 &\quad (\text{2nd conditional pair}) \times c_{15|34} \cdot c_{25|34} \\
 &\quad (\text{3rd conditional pair}) \times c_{12|345}.
 \end{aligned} \tag{38}$$

The vine structure associated to (38) is illustrated in Figure 11. This graphical model considerably simplify the understanding and we observe that this model is a R-vine because there is no specific constraints on the trees.

A re-labeling of the variables can lead to a large number of different PCC. [42] calculated the number of possible vine structures with the dimension d and shows that it becomes extremely large for high dimension problems. We illustrate below, using the same $d = 5$ di-

mensional example, two other PCC densities:

$$\begin{aligned}
 f_D &= f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot f_5 \\
 &\quad \times c_{12} \cdot c_{23} \cdot c_{34} \cdot c_{45} \\
 &\quad \times c_{13|2} \cdot c_{24|3} \cdot c_{35|4} \\
 &\quad \times c_{14|23} \cdot c_{25|34} \\
 &\quad \times c_{15|234}
 \end{aligned} \tag{39}$$

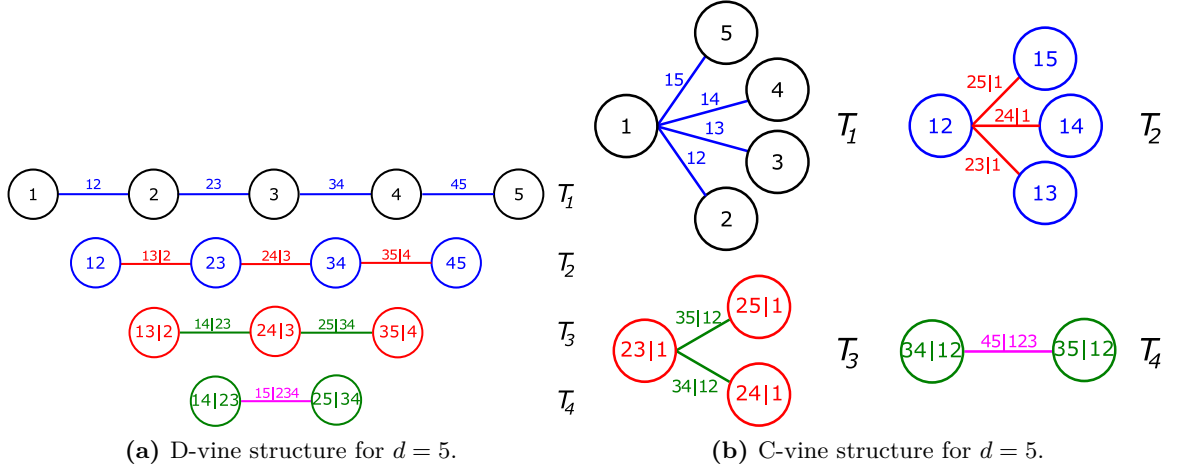


Figure 12: D-vine and C-vine structure for $d = 5$.

$$\begin{aligned}
f_C &= f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot f_5 \\
&\times c_{12} \cdot c_{13} \cdot c_{14} \cdot c_{15} \\
&\times c_{23|1} \cdot c_{24|1} \cdot c_{25|1} \\
&\times c_{34|12} \cdot c_{35|12} \\
&\times c_{45|123}
\end{aligned}
\tag{40}$$

where (39) and (40) respectively correspond

to D-vine and C-vine structures and are represented in Figures 12a and 12b. As we can see in these examples, the D-vine have a constraint on each tree that gives a path-like arrangement of the nodes. The C-vine on the other hand only has one node connected to all others for each tree.

An efficient way to store the information of a vine structure is proposed in [41] and is called a R-vine array. The approach uses the specification of a lower triangular matrix where the entries belong to $1, \dots, d$. Such matrix representation allows to directly derive the tree structure (or equivalently the associated PCC distribution). For more details, see [41].

B.2 Generating R-vine from an indexed list of pairs

The iterative procedure proposed in Section 4.3, described by Algorithm 2, minimizes the output quantile by iteratively determining the pairs of variables that influences the most the quantile minimization. At each iteration of the algorithm (step 1.a), a new vine structure is created by considering the list of influential pairs. The specificity of this vine creation is to consider the ranking of the list by placing the most influential pairs in the first trees of the R-vine. Thus, we describe in this section how to generate vine structure with the constraint of a given list of indexed pairs to fill in the structure.

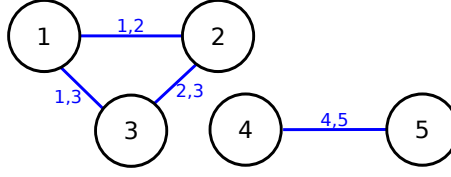


Figure 13: Example: first tree of a non valid vine structure for $d = 5$ that does lead to a single connected tree.

B.2.1 The algorithm

We consider the same notation as in Algorithm 2. Creating a vine structure from a given indexed list of pairs Ω_k is not straightforward. The difficulties come from respecting the ranking of Ω_k and the respect of the R-vine conditions. Indeed, the pairs cannot be append in the structure easily. The vine structure must respect these conditions, which can be sometime very restrictive. The procedure we proposed is detailed by the pseudo-code of Algorithm 3 and can be greatly simplified in these few key steps:

1. fill \mathcal{V} with the list Ω_k ,
2. fill \mathcal{V} with a permutation of Ω_{-k} ,
3. if \mathcal{V} is not a R-vine, then permute Ω_k and restart at step 1.

In step 1 and 2, the *filling* procedure, detailed in Algorithm 4, successively *adds* the pairs of a list in the trees of a vine structure. Adding a pair (i, j) in a tree T_l associates (i, j) with the conditioned set and determine a possible conditioning set D from the previous tree such as a possible edge is $i, j|D$.

In step 2, because the ordering of Ω_{-k} is not important in the filling of \mathcal{V} , the permutation of Ω_{-k} aims at finding a ranking such as \mathcal{V} leads to a R-vine.

In step 3, when the previous step did not succeeded and the resulting \mathcal{V} is not a R-vine structure, then the ranking of Ω_k is not possible and must be changed. The permutation of some elements of Ω_k must be done such as the ranking of the most influential pairs remains as close as possible to the initial one.

B.2.2 Example

For illustration, let's create a $d = 5$ dimensional vine structure with the given list of pairs $\Omega_k = ((1, 2), (1, 3), (2, 3), (4, 5), (2, 4), (1, 5))$ using Algorithm 3. Using the original list Ω_k , the *Fill* function may fail at line 7 of Algorithm 3, and more precisely, at line 15 of Algorithm 4. Indeed, the first tree of \mathcal{V} does not validate the R-vine conditions. The tree is illustrated in Figure 13 and as we can see, the nodes are not all connected into one single tree. Therefore, we permuted the list Ω_k by exchanging the pairs $(2, 4)$ and $(4, 5)$, as shown in Figure 14. This permutation now leads to a vine structure that respects the new ranked list $\Omega_k = ((1, 2), (1, 3), (2, 3), (2, 4), (4, 5), (1, 5))$.

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} .

```

1  $\Omega_k^{init} = \Omega_k$ ;
2  $k = 1$ ;
3 do
    /* initialize  $\mathcal{V}$  with a first empty tree */
4      $N_1 = (1, \dots, d)$ ;
5      $E_1 = ()$ ;
6      $\mathcal{V} = ((N_1, E_1))$ ;
    /* filling  $\mathcal{V}$  with the list of selected pairs  $\Omega_k$  */
7      $\mathcal{V} = \text{Fill}(\mathcal{V}, \Omega_k, d)$ ; // See Algorithm 4
    /* determining a permutation of  $\Omega_{-k}$  that fills  $\mathcal{V}$  */
8     for  $\Omega_{-k}^\pi \in \pi(\Omega_{-k})$  do
        /* filling  $\mathcal{V}$  with the candidate pairs  $\Omega_{-k}^\pi$  */
9          $\mathcal{V}_\pi = \text{Fill}(\mathcal{V}, \Omega_{-k}^\pi, d)$ ; // See Algorithm 4
10        if  $\mathcal{V}_\pi$  is a R-vine then
11            /* a permutation worked  $\rightarrow$  we quit the loop */
12            break
12         $\mathcal{V} = \mathcal{V}_\pi$ ;
13    if  $\mathcal{V}$  is not a R-vine then
14        /* filling did not work  $\rightarrow$  permute initial list  $\Omega_k^{init}$  */
15        Get  $\Omega_k$  by inverting pairs of  $(\Omega_k^{init},$ 
16         $k = k + 1$ ;
16 while  $\mathcal{V}$  is not a R-vine;

```

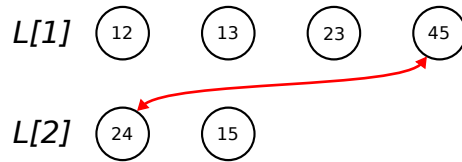


Figure 14: Example: exchange of elements of Ω_k in order to lead to a valid vine structure.

Algorithm 4: Filling a vine structure with a given list

```

1 function Fill( $\mathcal{V}$ ,  $\Omega_k$ ,  $d$ ):
    /*  $\mathcal{V}$ : an incomplete vine structure , */
    /*  $\Omega_k$ : a list of indexed pairs */
    /*  $d$ : the input dimension. */
2      $l = |\mathcal{V}|$ ; // number of existing trees
3      $(T_1, \dots, T_l) = \mathcal{V}$ ;
4      $k = |T_l|$ ; // number of existing nodes in last tree
    /* loop over the list of pairs */
5     for  $(i, j) \in \Omega_k$  do
6          $D = \emptyset$ ;
7         if  $l \geq 2$  then
            /* conditioning set is only computed from  $T_2$  */
8              $D = \text{FindConditioningSet}((i, j), N_{l-1})$ ; // See Algorithm 5
9             if  $D = \emptyset$  then
                /* no conditioning set found  $\rightarrow$  not possible */
10                return False
11             $E_l = E_l \cup i, j | D$ ; // add new edge in  $E_l$ 
12             $T_l = (N_l, E_l)$ ; // update current tree
13             $\mathcal{V} = (T_1, \dots, T_l)$ ;
14            if  $k \geq d - l$  then
                /* if tree  $T_l$  is complete */
15                if  $\mathcal{V}$  does not fulfill the R-vine conditions then
                    /* the vine structure  $\mathcal{V}$  is not valid */
16                    return False
17                 $k = 1$ ;
18                 $l = l + 1$ ;
19                 $N_l = E_{l-1}$ ; // nodes of next tree are the edges of previous tree
20            else
21                 $k = k + 1$ ;
22    return  $\mathcal{V}$ 

```

Algorithm 5: Gets the conditioning set of a given conditioned set

```

1 function FindConditioningSet( $(i, j)$ ,  $N_-$ ):
    /*  $(i, j)$ : the conditioned set, */
    /*  $N_-$ : list of nodes from the previous tree. */
2      $D = \emptyset$ ;
3     for  $a, b \in N_-$ , with  $a \neq b$  do
4         if  $i \in a$  and  $j \in b$  then
5             if  $j \notin A_a$  and  $i \notin A_b$  then
                /* See Definition 2 */
6                  $D = A_a \cap A_b$ ;
7                 break;
8     return  $D$ 

```
