



A novel multivariate risk measure: the Kendall VaR

Matthieu Garcin, Dominique Guegan, Bertrand Hassani

► To cite this version:

Matthieu Garcin, Dominique Guegan, Bertrand Hassani. A novel multivariate risk measure: the Kendall VaR. 2018. halshs-01467857v2

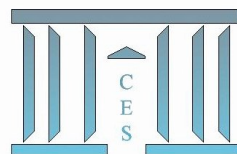
HAL Id: halshs-01467857

<https://shs.hal.science/halshs-01467857v2>

Submitted on 6 Jul 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.



A novel multivariate risk measure: the Kendall VaR

Matthieu GARCIN, Dominique GUEGAN, Bertrand HASSANI

2017.08R

Version révisée



A novel multivariate risk measure: the Kendall VaR

Matthieu Garcin*, Dominique Guégan†, Bertrand Hassani‡

April 13, 2018

Abstract

The definition of multivariate Value at Risk is a challenging problem, whose most common solutions are given by the lower- and upper-orthant VaRs, which are based on copulas: the lower-orthant VaR is indeed the quantile of the multivariate distribution function, whereas the upper-orthant VaR is the quantile of the multivariate survival function. In this paper we introduce a new multivariate Value at Risk, referred to as the Kendall Value at Risk, which links the copula approach to an alternative definition of multivariate quantiles, known as the quantile surface, which is not used in finance, to our knowledge. We more precisely transform the notion of orthant VaR thanks to the Kendall function so as to get a multivariate VaR, that is to say a set of loss vectors, with some advantageous properties compared to the standard orthant VaR: *i/* it is based on a total order, *ii/* the probability level of the VaR is consistent with the probability measure of the set of the more severe loss vectors, *iii/* the d -dimensional VaRs based on the distribution function or on the survival function have vectors in common, which conciliate both upper- and lower-orthant approaches. We quantify the differences between this new Kendall VaR and orthant VaRs. In particular, we show that Kendall VaRs are less (respectively more) conservative than lower-orthant (resp. upper-orthant) VaRs. The definition and the properties of the Kendall VaR are illustrated using Gumbel and Clayton copulas with lognormal marginal distributions and several levels of risk.¹

Keywords – Value at Risk, multivariate quantile, risk measure, Kendall function, copula, total order

1 Introduction

For the random variable X , whose realisation is the loss of a unique asset or of a portfolio (so that the realisation of $-X$ is a gain), the Value at Risk at the level α , defined for a given horizon, is the univariate quantile of probability α of X . This metric allows estimating and comparing the risk of portfolios, experienced as the severity of a loss. However, the aggregation of the risks may rely on an arbitrary criterion and be overly simplistic. As a consequence, a multivariate analysis

*Natixis Asset Management and LabEx ReFi, 21 quai d'Austerlitz, 75013 Paris, France, matthieu.garcin@m4x.org, corresponding author

†Université Paris 1 Panthéon-Sorbonne, LabEx ReFi and IPAG, 106 bd de l'Hôpital, 75013 Paris, France, dguegan@univ-paris1.fr

‡Grupo Santander, Université Paris 1 Panthéon-Sorbonne and LabEx ReFi, CES

¹This work was achieved through the Laboratory of Excellence on Financial Regulation (Labex ReFi) supported by PRES heSam under the reference ANR10LABX0095. It benefited from a French government support managed by the National Research Agency (ANR) within the project Investissements d'Avenir Paris Nouveaux Mondes (investments for the future Paris New Worlds) under the reference ANR11IDEX00060

of the risks may be preferred to the risk of the aggregated value of a portfolio. For example, the decomposition of the risk into different asset classes (the loss for stocks, bonds, currencies, etc., expressed separately instead of a sole aggregated loss) can legitimate this approach. The computation of capital requirements also leads to the necessity of using a multivariate Value at Risk [25].

In a broader perspective, the risk must be expressed as a vector when the risk components cannot be aggregated. It may be useful in other fields than in finance. For example, in meteorology, the definition of what severe weather events are depends on the combination of unaggregatable parameters such as speed of wind, quantity of precipitation, temperature, cloud cover [33]. It may also be relevant in finance, because assets of different markets can be associated to different liquidity and transaction costs and therefore it might not be possible to aggregate them.

The multivariate analysis of the risk relies on a multivariate probability distribution of losses. A useful tool for representing multivariate distributions is the copula, which is a d -dimensional distribution function $[0, 1]^d \rightarrow [0, 1]$, where $d \in \mathbb{N}$ is the number of underlying random variables. Copulas are therefore used to describe the dependence between random variables. They have been applied in many fields, mostly in finance, but also in hydrology [20], in astronomy [41, 40] or in telecommunication networks [37, 19].

Embrechts and Puccetti [16] have been the firsts to deal with the notion of multivariate Value at Risk. Let G be the multivariate probability distribution of a random loss vector $X = (X_1, \dots, X_d)'$, that is to say the copula-based probability distribution:

$$G : (y_1, \dots, y_d) \in \mathbb{R}^d \mapsto \mathbb{P}[X_1 \leq y_1, \dots, X_d \leq y_d].$$

If G_1, \dots, G_d are the d univariate marginal distribution functions of X , then Sklar's theorem affirms the existence of a copula C such that $G(y_1, \dots, y_d) = C(G_1(y_1), \dots, G_d(y_d))$ [43]. Embrechts and Puccetti [16] defined the multivariate Value at Risk of X as the set of vectors belonging to the boundary of the α -level set of G . They distinguished the lower-orthant Value at Risk, $\underline{\text{VaR}}_\alpha(G)$, and the upper-orthant Value at Risk, $\overline{\text{VaR}}_\alpha(G)$. The lower-orthant Value at Risk is defined by $\underline{\text{VaR}}_\alpha(G) = \partial\{y \in \mathbb{R}^d | G(y) \geq \alpha\}$, where ∂ denotes the boundary of the mentioned set, whereas the upper-orthant Value at Risk is $\overline{\text{VaR}}_\alpha(G) = \partial\{y \in \mathbb{R}^d | \overline{G}(y) \leq 1 - \alpha\}$, where \overline{G} is the survival function associated to G . In general, the lower-orthant Value at Risk is more conservative than the upper-orthant Value at Risk.² If $(y_1, \dots, y_d)' \in \underline{\text{VaR}}_\alpha(G)$, then

$$\mathbb{P}[X_1 \leq y_1, \dots, X_d \leq y_d] = \alpha.$$

By the way, the lower-orthant approach is based on a partial order in \mathbb{R}^d , since only the lower-left quadrant of y is dominated by y according to this order. Consequently, two loss vectors can be incomparable for this setting. In addition, if y and z belong to $\underline{\text{VaR}}_\alpha(G)$ and y' is in the lower-left quadrant of y , then y and z are seen as equally severe and y is seen as more severe than y' . Indeed, the set of loss vectors less severe than y has a probability measure α and this set is in fact the lower-left quadrant of y . However, according to this definition of the Value at Risk, it is not enough for asserting that the loss vector z is more severe than y' , if y' is not in the lower-left quadrant of z . Paradoxically, a lower level is associated to y' than to z , but both vectors are not directly comparable because no one is in the lower-left quadrant of the other. For these reasons, another definition of the multivariate Value at Risk inducing a total order of \mathbb{R}^d would be preferable. This total order would mean that the probability measure of all the vectors of probability level lower than α is exactly α . This is the aim of this paper.

² This means that if a loss vector y belongs both to $\overline{\text{VaR}}_\alpha(G)$ and $\underline{\text{VaR}}_{\alpha'}(G)$, then the probability level α' associated by the lower-orthant Value at risk to y is lower than the probability level α associated by the upper-orthant Value at risk to y .

For this purpose, we introduce a new definition of the multivariate Value at Risk, as an alternative to the orthant VaRs, based on another interpretation of what a multivariate Value at Risk should be. We call it the Kendall Value at Risk since it uses the Kendall distribution function: $K : t \in [0, 1] \mapsto \mathbb{P}[G(X) \leq t]$, where G is the multivariate probability distribution of the random vector X . Then, the Kendall Value at Risk³ is $\text{VaR}_\alpha^K(G) = \partial\{y \in \mathbb{R}^d | K(G(y)) \geq \alpha\}$. We will see to which extent it differs from $\underline{\text{VaR}}_\alpha(G)$ and $\overline{\text{VaR}}_\alpha(G)$. In particular, the Kendall VaR is less conservative than the lower-orthant VaR and more conservative than the upper-orthant VaR: indeed, if a loss vector y belongs to $\underline{\text{VaR}}_\alpha(G)$, $\overline{\text{VaR}}_{\alpha'}(G)$ and $\text{VaR}_{\alpha''}^K(G)$, then the probability levels associated to this loss are ordered in the following way: $\alpha \leq \alpha'' \leq \alpha'$. More precisely, our approach unifies the concepts of lower-orthant and upper-orthant VaRs, since the Kendall VaRs based on the distribution function or on the survival function have some vectors in common. Both orthant VaRs are based on a partial order and nothing indicates which orthant VaR should be preferred. The Kendall VaR provides a rigorous way of defining a consensus between both orthant bounds.

The approach of Embrechts and Puccetti [16] has been used as such in the literature [11], or has presented some modifications. For example, Cousin and Di Bernardino [12] have replaced both the sets defined by the lower-orthant and the upper-orthant Value at Risk by their expected value, so that the Value at Risk is only defined by a simple vector instead of an infinite set of vectors. Similarly to the unidimensional Value at Risk, this approach has the advantage to define a total order, but only for $\text{VaR}^{-1}([0, 1])$, which is a subset of \mathbb{R}^d . The methodological choice of Cousin and Di Bernardino makes it impossible to compare some pairs of vectors: it is hence relevant in the case of the orthant Value at Risk but not in our framework, where total order plays a significant role.

Our total-order multivariate Value at Risk is also linked to alternative definitions of multivariate quantile, which are not used in finance, to our knowledge, such as the quantile surface. Ordering real univariate data is quite easy. However, when dealing with multivariate data, no consensus arises about what should order statistics and quantiles be. In particular, the question of quantiles of multivariate distributions has led to numerous interpretations, often inspired by analogies with different ways of defining the quantiles of a univariate distribution. These interpretations are also thought to fit with a particular framework. In our case, the random vector from which we want to define a quantile is a vector of absolute or relative losses. Such quantiles must enable to define a risk measure, as does the VaR in the univariate case. But the definition of a multivariate quantile will induce some particularities, which are worth noting:

- (P1) Contrarily to what happens in the univariate case, a multivariate quantile does not correspond to only one value of the random vector but to a set of values. In addition, when dealing with multivariate quantiles, magnitude is not the only feature to be considered. The direction of vectors is indeed important and several directions are possible. For example, for a bivariate quantile, we may choose the four following directions: $(-1, -1)'$, $(-1, 1)'$, $(1, 1)'$ and $(1, -1)'$ [4].
- (P2) The notion of aggregation has to be questioned, since aggregation smooths micro-structure of losses. Should the vectors of relative losses $(-1\%, 1\%, -1\%, 2\%, -1\%)'$ and $(0\%, 0\%, 0\%, 0\%, 0\%)'$ be aggregated in 0%, and therefore be considered as equal-severity vectors? Or is it possible to have a vectorial interpretation of losses, that is not limited to the sum of the coordinates?
- (P3) As there is no consensus about what a multidimensional quantile is, several approaches can be considered. However, a good method must enable an easy comparison of d -dimensional

³ We define as well another Kendall Value at Risk similarly to the upper-orthant Value at Risk, as $\overline{\text{VaR}}_\alpha^K(G) = \partial\{y \in \mathbb{R}^d | \mathcal{K}(\overline{G}(y)) \leq 1 - \alpha\}$, where $\mathcal{K} : t \in [0, 1] \mapsto \mathbb{P}[\overline{G}(X) \leq t]$. We will show that for not too restrictive assumptions, $\overline{\text{VaR}}_\alpha^K(G) = \text{VaR}_\alpha^K(G)$.

vectors in order to determine which of two loss vectors is more severe. If not, the *quantile* can lead to probabilistic misinterpretations, mainly if $d \geq 2$. In other words, this quantile must induce a total order: for any pair of loss vectors (y, z) , y and z should either be equally severe or y should be more severe (or respectively less severe) than z and, in this case, the risk level associated to y (resp. z) should be the probability measure of all less severe loss vectors than y (resp. z), including z (resp. y). Mathematically speaking, if a multivariate VaR of level α is noted V_α , then α should be solution of the equation

$$\alpha = \mathbb{P}\{z \in V_{\alpha'} | \alpha' < \alpha\}.$$

In Section 2, we present various interpretations of multivariate quantiles. In Section 3, we focus on the new Value at Risk that we introduce in this paper: the Kendall Value at Risk. We compare it theoretically with lower- and upper-orthant VaRs. In Section 4, we present the results of an application to simulated data. Section 5 concludes.

2 Multivariate quantiles

Among the possible definitions of a multivariate quantile, we can cite the following ones. Further details can be found in the survey of Serfling [42].

1. *Spatial quantile.* This approach was introduced by Abdous and Theodorescu [1]. They extended to the multivariate case the fact that the p -quantile of a univariate random variable X is any solution of the minimization problem [17]:

$$\inf_{y \in \mathbb{R}} \mathbb{E} \left\{ \frac{|X - y| + (2p - 1)(X - y)}{2} - \frac{|X| + (2p - 1)X}{2} \right\}. \quad (1)$$

Then, they define a p -quantile of a real multivariate random variable $X = (X_1, \dots, X_d)'$ as any solution of the minimization problem:

$$\inf_{y \in \mathbb{R}^d} \mathbb{E} \{ \|X - y\|_{n,p} - \|X\|_{n,p} \}, \quad (2)$$

where $\|\cdot\|_{n,p}$ has some norm-like properties, while not being a norm, and is defined by:

$$\|X\|_{n,p} = \left(\sum_{i=1}^d \left| \frac{|X_i| + (2p - 1)X_i}{2} \right|^n \right)^{1/n}.$$

In the context of Value at Risk, such a quantile definition has the advantage of being different from a simple aggregation. However, the interpretation of what the quantile means in terms of ordering losses does not seem straightforward. In addition, the calculation takes much time if the dimension is high because of the intricate estimation of multidimensional densities. This method has been applied for quantile prediction and Value at Risk for financial data [24]. However, this method leads to a probabilistic misinterpretation (P3).

2. *Geometric quantile.* The multidimensional extension of equation (1) in equation (2) can lead to another interpretation than the spatial quantile. In the geometric quantile [8], the probability p is not a simple real number but a vector, \tilde{p} , so that the quantity $\tilde{u} = 2\tilde{p} - 1$ belongs to the unit ball $B^d = \{u \in \mathbb{R}^d, |u| < 1\}$, which is a direct extension of the unidimensional case $u = 2p - 1 \in (-1, 1)$. As a consequence, the u -quantile, for $u \in B^d$, is any solution of the minimization problem:

$$\inf_{y \in \mathbb{R}^d} \mathbb{E} \{ \|X - y\|_u - \|X\|_u \},$$

where $\|\cdot\|_u$ is defined by:

$$\|X\|_u = \|X\| + \langle u, X \rangle,$$

with the help of the Euclidean inner product $\langle \cdot, \cdot \rangle$ and Euclidean norm $\|\cdot\|$. Then, two readings of the minimization problem are possible. On the one hand, we get rid of the difficult interpretation of a multidimensional probability, \tilde{p} , by only considering its magnitude. In this case, the p -quantile, where $p \in (0, 1)$, is described by the set of geometric u -quantiles, where $u \in B^d$ and $|u| = p$. But this center-outward approach is not well adapted to our financial risk-measure framework. On the other hand, we can favour a particular direction and choose a single vector of norm 1, u^* . It represents either the ideal asset return or the most feared loss. Then, a p -quantile can be defined as a geometric pu^* -quantile. The latter interpretation allows easier ordering of the vectors than the spatial quantile does. However, the same curse of dimensionality and probabilistic misinterpretation (P3) as in the spatial-quantile approach arise. We find several applications of this method in the literature, for example in medicine [9] or in ecology [7].

3. *Quantiles based on inversions of mappings.* In the unidimensional framework, a quantile is defined as the generalized inverse of the cumulated distribution function. If one defines a mapping G from \mathbb{R}^d to \mathbb{R}^d or from \mathbb{R}^d to \mathbb{R} , then inversions can also define a quantile [32]. This method is linked to multivariate ordering based on a scalarization, which is the ordering of vectors by comparing scalars, such as isolated coordinates or a function of a linear combination of coordinates [38], or any mapping [44]. For some choice of G , probabilistic misinterpretations (P3) may arise. This method has been applied to Value at Risk by Embrecht and Puccetti [16] as cited above. For example, if the copula and therefore the joint distribution of the coordinates is known, the p -quantile $y = (y_1, \dots, y_d) \in \mathbb{R}^d$, if it exists, is simply defined by $y = G^{-1}(p)$, in which:

$$G(y) = \mathbb{P}(X_1 \leq y_1, \dots, X_d \leq y_d) = p,$$

where $X = (X_1, \dots, X_d)'$ is a real random vector. It has to be noted that, in this copula-inversion-based approach, the probability associated to a quantile consists in the weight of the lower-left quadrant of the quantile. Then, for two distinct vectors y and z having the same image p by G , the lower-left quadrants are different and the probabilistic interpretation of what a Value at Risk is becomes complicated since (P3) does not necessarily hold. Indeed, the union of both lower-left quadrants is a set of less severe loss vectors than y and z , but its probability measure is strictly higher than p .

4. *Center-outward quantile surface.* If one is given a statistical depth function⁴, the center of the distribution is defined as the maximal-depth vector⁵. Then, a natural ordering arises based on the depth. Thus, quantiles are concentric regions around the center [34, 45]. More precisely, given a probability $p \in (0, 1)$, the p -quantile of a distribution G is the set of vectors of depth α_p , which is defined such that the probability associated to G of the region of vectors of depth D higher than α_p is p : $p = \mathbb{P}_G(D(X, G) \geq \alpha_p)$, where X is a random vector of distribution G . Thanks to this definition, no probabilistic misinterpretation (P3) is possible. It is an advantage on the methods previously presented. However, this center-outward approach is not adapted to our financial risk-measure purpose. Indeed, with such a method, a vector

⁴ For example, if we consider the so-called likelihood depth, the depth function is simply the probability density [18]. Otherwise, the Mahalanobis depth [35], $D^M(y, G)$, of a vector y with respect to a multivariate probability distribution G is defined by

$$D^M(y, G) = \left[1 + (y - \mu_G)' \Sigma_G^{-1} (y - \mu_G) \right]^{-1},$$

where μ_G and Σ_G are respectively the mean vector and the covariance matrix of a random vector of distribution G .

⁵ The center is also sometimes called median of the distribution, but it is quite different from a statistical median in the sense of a 0.5-quantile.

containing only negative coordinates may be in the same quantile as a vector containing only positive coordinates, whereas in finance the direction of the vectors matters (P1).

5. *Modified quantile surface.* We propose to modify the center-outward quantile surface so as to focus rather on gains and losses than on the center of the distribution. Instead of determining first a spatial median, we associate a metric for each vector. Vectors with the same metric are gathered in an equivalence class: they are supposed to be equivalently risky. We can then compare each pair of classes and define an order resulting in the definition of a multivariate probability distribution and thus of quantiles. We propose two methods to achieve this goal, depending on the definition of the equivalence classes.

- ▷ We choose a vector u^* corresponding to the ideal gain, such that, in the empirical distribution, no vector of returns will provide higher utility to the portfolio manager than u^* . We then define a function of dissimilarity with the ideal gain. Contrarily to a depth function, a dissimilarity function gives the lowest value to the vector of interest, u^* . For example, we can use, $d^M(y, G, u^*)$, the squared Mahalanobis distance to u^* for a vector y with respect to the multivariate probability distribution G , defined by

$$d^M(y, G, u^*) = (y - u^*)' \Sigma_G^{-1} (y - u^*),$$

where Σ_G is the covariance matrix of a multivariate random variable of probability distribution G .

- ▷ The Kendall probability distribution

$$K : t \in [0, 1] \mapsto \mathbb{P}[G(X) \leq t],$$

where G is the multivariate probability distribution of the random vector X , defines natural equivalence classes [36]. If vectors y and y' are such that $G(y) = G(y')$, then the loss vector y is as severe as y' and these loss vectors are more severe than every loss vector y'' such that $G(y'') < G(y)$. Contrarily to the orthant Value at Risk, we affirm that y is a vector belonging to the set of the Value at Risk of probability $K(G(y))$, rather than of probability $G(y)$, which is lower than $K(G(y))$ by construction. We base our new definition on the Kendall stochastic ordering [36] instead of the traditional product ordering. The first one is a total order, whereas the second one is only partial. An explanatory illustration is provided in Figure 1.

Then, whatever the methodological choice for the equivalence classes, the definition of the quantiles follows the center-outward quantile-surface method, which provides a consistent probabilistic interpretation (P3).

Among all the presented notions of multivariate quantile, the modified quantile surface is the most relevant one in the perspective of Value at Risk. However, this method is based on the choice of an arbitrary scalarization. Among all the possible choices, we focus on the modified quantile surface based on Kendall function. This scalarization is indeed interesting since:

- ▷ it is based on the distribution function of the variable,
- ▷ it provides us with quantiles made of sets of vectors identical to the sets of vectors in the orthant approach, but simply associated to another probability level,
- ▷ it conciliates both lower-orthant and upper-orthant approaches as the probability level associated to a given vector for the Kendall quantile is between probability levels associated to lower- and upper-orthant quantiles,

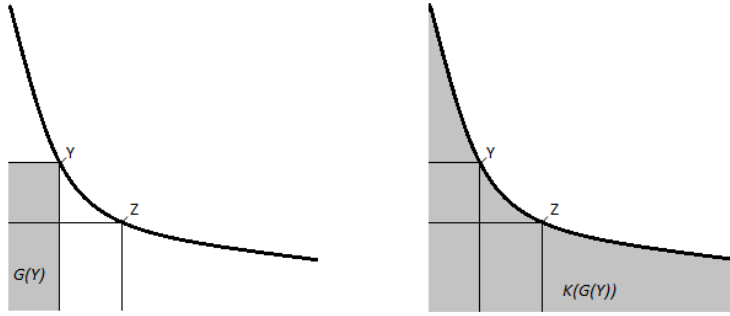


Figure 1: On the left, the thick line is a set of two-dimensional vectors having the same lower-left cumulated probability. In particular, $G(y) = G(z)$, which is the probability measure of the lower-left quadrant of y or z . However, some vectors of the lower-left quadrant of z are not in the lower-left quadrant of y and therefore cannot be compared to y in terms of loss severity. On the right, the multivariate probability distribution only leads to the definition of equivalence classes. Therefore, every loss vector in the grey zone is a less severe loss than every vector on the thick line. The vectors dominated by z are the same than those dominated by y . The probability associated with y and z is therefore the probability measure of all the grey zone, that is $K(G(y))$, which is equal to $K(G(z))$ and which is greater than $G(y)$.

- ▷ the level associated to a set of vectors is exactly the probability measure of all the loss vectors considered as less severe, that is to say vectors associated to a lower probability level.

We call it *Kendall Value at Risk*. We make the choice to focus on the Kendall Value at Risk in the rest of the paper.

3 Kendall Value at Risk and Upper-Orthant Kendall VaR

Two approaches relying on Kendall distributions are presented in the next subsection. Drawing a parallel with lower and upper-orthant VaRs presented above, it is possible to formalise a lower and an upper-orthant Kendall VaR respectively labelled Kendall VaR and upper-orthant Kendall VaR in what follows.

3.1 Definitions

1. *Kendall VaR*: As introduced above, the Kendall distribution function is $K : t \in [0, 1] \mapsto \mathbb{P}[G(X) \leq t]$, where G is the multivariate probability distribution of the random vector X , associated to a given copula. The Kendall function has been used for example to estimate Archimedean copulas [21], or for creating hierarchical Kendall copulas, which deal with high-dimension problems [5]. Thanks to this function, we define the Kendall Value at Risk.

Definition 1. For a random vector X of dimension d , the Kendall Value at Risk of probability $\alpha \in [0, 1]$, denoted VaR_α^K , is the boundary set of the set of vectors $y \in \mathbb{R}^d$ such that $K(G(y)) \geq \alpha$, where G is the multivariate distribution of X and K the corresponding Kendall function:

$$\text{VaR}_\alpha^K(G) = \partial\{y \in \mathbb{R}^d | K(G(y)) \geq \alpha\}.$$

2. *Upper-orthant Kendall VaR*: Similarly to the distinction between lower-orthant and upper-orthant Value at Risk, we can make a distinction between two kinds of Kendall Value at Risk,

based either on the multivariate distribution function G or the corresponding survival function \bar{G} . We thus introduce in Definition 2 another Kendall function, $\mathcal{K} : t \in [0, 1] \mapsto \mathbb{P}[\bar{G}(X) \leq t]$. We stress the fact that \mathcal{K} is neither the survival Kendall function associated to G , nor K itself, but it is the standard Kendall function associated to \bar{G} .

Definition 2. For a random vector X of dimension d , the upper-orthant Kendall Value at Risk of probability $\alpha \in [0, 1]$, denoted $\overline{\text{VaR}}_\alpha^\mathcal{K}$, is the boundary set of the set of vectors $y \in \mathbb{R}^d$ such that $\mathcal{K}(\bar{G}(y)) \leq 1 - \alpha$, where \bar{G} is the survival function associated to the multivariate distribution G of X and \mathcal{K} is the Kendall function corresponding to \bar{G} , that is $\mathcal{K} : t \in [0, 1] \mapsto \mathbb{P}[\bar{G}(X) \leq t]$:

$$\overline{\text{VaR}}_\alpha^\mathcal{K}(G) = \partial\{y \in \mathbb{R}^d | \mathcal{K}(\bar{G}(y)) \leq 1 - \alpha\}.$$

In fact, these different definitions of a multivariate Value at Risk are linked, as exposed in the following proposition. Indeed, contrarily to lower-orthant and upper-orthant VaRs, both Kendall VaRs have some vectors in common.

Proposition 1. Let $\alpha \in [0, 1]$, G be a non-atomic multivariate distribution function of dimension $d \in \mathbb{N}$ having a density function whose support is \mathbb{R}^d , with K the Kendall function, \bar{G} the survival distribution, both associated to G , and \mathcal{K} the Kendall function of \bar{G} as introduced in Definition 2. Then:

$$\overline{\text{VaR}}_\alpha^\mathcal{K}(G) \cap \text{VaR}_\alpha^K(G) \neq \emptyset.$$

Proof. Let $L = \{y \in \mathbb{R}^d | K(G(y)) \geq \alpha\}$ and $U = \{y \in \mathbb{R}^d | \mathcal{K}(\bar{G}(y)) \leq 1 - \alpha\}$. Let g be the probability measure associated to G . If $\overline{\text{VaR}}_\alpha^\mathcal{K}(G) \cap \text{VaR}_\alpha^K(G) = \emptyset$, there are two possibilities:

1. $L \cap U$ is an infinite and closed set. Since $\overline{\text{VaR}}_\alpha^\mathcal{K}(G) \cap \text{VaR}_\alpha^K(G) = \emptyset$, every vector is at least in one of the two sets L and U . The probability measure of $L \cap U$ is then strictly positive, but it is contradictory to the fact that the measure of \mathbb{R}^d , equal to 1, is then $g(L) + g(U) - g(L \cap U) = \alpha + (1 - \alpha) - g(L \cap U) = 1 - g(L \cap U)$.
2. The set S of vectors, such that $\forall y \in S, \mathcal{K}(\bar{G}(y)) > 1 - \alpha$ and $K(G(y)) < \alpha$, is an infinite and closed set. Then, similarly to the previous case, $1 = g(L) + g(U) + g(S) = 1 + g(S)$, which is contradictory with the fact that its probability measure is expected to be strictly greater than 0.

□

Remark 1. It is noteworthy to mention that though the (lower-orthant) Kendall VaR is directly applicable as copulas can be adjusted to data or distributions, the upper-orthant Kendall VaR is both more complicated to interpret and not directly applicable. This is the reason why the term Kendall VaR has been kept instead of lower-orthant Kendall VaR. Besides, as a consequence, in Section 4 we will only focus on the Kendall VaR.

3.2 Properties

It is well known that the univariate Value at Risk is – in general – not a coherent risk measure since it is not sub-additive [2, 30]. Nevertheless, as this property depends on the underlying distribution, sub-additivity can arise [14, 26, 27]. Thus, the absence of sub-additivity can be a limitation for the use of the VaR in certain cases. With the multivariate approach, the question is opened. A

definition of what a consistent risk measure is in the multivariate case can be provided based on the choice of a partial order [31]. As the question of multivariate sub-additivity is open and beyond this article's scope, we think that it is here more fruitful to focus on other more specific properties of the Kendall VaR. In particular, it is interesting to understand the difference between the Kendall Value at Risk and the lower-orthant Value at Risk.

As we mentioned, the probability associated to a vector by the Kendall function is higher (respectively lower) than in the approach of the lower-orthant (resp. upper-orthant) Value at Risk. It can be shown by the Fréchet-Hoeffding bounds that $t \leq K(t) \leq 1$ [21] and that $t \leq \mathcal{K}(t) \leq 1$ as well. For a vector y and a multivariate probability distribution G , the lower-orthant approach links y to the level of probability $G(y)$, whereas the Kendall approach associates it to a probability $K(G(y))$, which is therefore in the interval $[G(y), 1]$. In other words, the Kendall VaR is less conservative than the lower-orthant VaR. We can compare both VaRs in the following manner:

Proposition 2. *Let K be strictly monotonic on a neighbourhood of a given probability $\alpha \in [0, 1]$. Then, the Kendall Value at Risk and the lower-orthant Value at Risk are linked by:*

$$\underline{\text{VaR}}_\alpha(G) = \text{VaR}_{K(\alpha)}^K(G).$$

Proof. Let $A = \{y \in \mathbb{R}^d | G(y) \geq \alpha\}$ and $B = \{y \in \mathbb{R}^d | K(G(y)) \geq K(\alpha)\}$.

- ▷ K is an increasing function, since it is a probability distribution function. Therefore, it comes immediately that $A \subset B$.
- ▷ The reciprocal inclusion does not hold in general. However, with the assumption of strict monotonicity of K in a neighbourhood V_α of α , the restriction of K to V_α is invertible. Let $y \in B$ and $\alpha' \in V_\alpha$ such that $\alpha' < \alpha$ (it does not exist if $\alpha = 0$ but this case is besides trivial), so that $K(\alpha') < K(\alpha)$. Let assume $y \notin A$. Then $G(y) < \alpha$. Two cases arise. First, if $G(y) \leq \alpha'$, then $K(G(y)) \leq K(\alpha') < K(\alpha)$, what is contradictory with the assumption $y \in B$. Second, if $G(y) \in (\alpha', \alpha)$, then $G(y)$ is in V_α , so that $K(G(y)) < K(\alpha)$: the contradiction also holds. Therefore the assumption $y \notin A$ was absurd and we can conclude that $B \subset A$.
- ▷ Finally $A = B$.

As a consequence, when considering the definition of both VaRs, we get:

$$\begin{aligned} \underline{\text{VaR}}_\alpha(G) &= \partial\{y \in \mathbb{R}^d | G(y) \geq \alpha\} \\ &= \partial\{y \in \mathbb{R}^d | K(G(y)) \geq K(\alpha)\} \\ &= \text{VaR}_{K(\alpha)}^K(G). \end{aligned}$$

□

Similarly, we can show that the upper-orthant Kendall VaR (which has a non-empty intersection with the standard Kendall VaR as stated in Proposition 1.) is more conservative than the upper-orthant VaR. It is the meaning of the next proposition, since \mathcal{K} being a growing function we have $1 - \mathcal{K}(1 - \alpha) \leq \alpha$.

Proposition 3. *Let K be strictly monotonic on a neighbourhood of a given probability $\alpha \in [0, 1]$. Then, the survival Kendall Value at Risk and the upper-orthant Value at Risk are linked by:*

$$\overline{\text{VaR}}_\alpha(G) = \overline{\text{VaR}}_{1-\mathcal{K}(1-\alpha)}^\mathcal{K}(G).$$

The proof is similar to the one of Proposition 2. We can also compare the level associated to the standard Kendall VaR to the level associated to the upper-orthant VaR and also between the upper-orthant Kendall VaR and the lower-orthant VaR.

Proposition 4. *Let $\alpha, \alpha', \alpha'' \in [0, 1]$ and G a probability distribution with no atoms.*

1. *If $\text{VaR}_\alpha(G) \cap \overline{\text{VaR}}_{\alpha''}^K(G) \neq \emptyset$, then $\alpha'' \geq \alpha$.*
2. *If $\overline{\text{VaR}}_{\alpha'}(G) \cap \text{VaR}_{\alpha''}^K(G) \neq \emptyset$, then $\alpha'' \leq \alpha'$.*

Proof. We prove the second assertion, the proof for the first one being similar. Whatever y and z in $\text{VaR}_{\alpha''}^K(G)$, z cannot be in the interior of the upper orthant of y . Indeed, in such a case, $G(z) > G(y)$ or, if $G(z) = G(y)$, y would not be on the border of $A = \{x \in \mathbb{R}^d | K(G(x)) \geq \alpha''\}$ since all the lower orthant of z , in the interior of which is y , belongs to A .

Let $y \in \overline{\text{VaR}}_{\alpha'}(G) \cap \text{VaR}_{\alpha''}^K(G)$. The probability measure of the upper orthant U of y is $1 - \alpha'$. Since no vector of A is in the interior of U , the probability measure of A , which is equal to α'' , is lower than the measure of the complement set of U , since G has no atoms. Therefore, $\alpha'' \leq 1 - (1 - \alpha') = \alpha'$. \square

The message conveyed by Propositions 2, 3 and 4 is that the Kendal VaRs are a compromise between both orthant VaRs: they are more conservative than one and less conservative than the other.

The interesting metric for comparing the Kendall VaR and the lower-orthant VaR is given by the positive function $r : \alpha \in [0, 1] \mapsto K(\alpha) - \alpha$. It is the difference of risk associated to a same vector by the Kendall VaR and by the lower-orthant VaR, for a given level of risk. In other words, for a probability α , $\text{VaR}_\alpha(G)$ is a set of vectors corresponding to this risk α ; for the same set of vectors, the Kendall Value at Risk associates another level of risk, which is $K(\alpha)$ according to Proposition 2; $r(\alpha)$ denotes this difference of the estimated risk. For example, the Gumbel copula in example 1 leads to $r(\alpha) = -\frac{\alpha \log(\alpha)}{\theta}$. More generally, r can be linked to the Kendall rank correlation coefficient, known as Kendall's tau coefficient.

Proposition 5. *The average difference between the probabilities associated to the Kendall function and to the sole copula, for d -dimensional vectors and a continuous copula, is:*

$$\int_0^1 r(\alpha) d\alpha = (1 - \tau) \left(\frac{1}{2} - \frac{1}{2^d} \right),$$

where τ is the Kendall rank correlation coefficient.

Proof. Kendall's tau and the Kendall function are linked by the following relation, for a continuous copula [22]:

$$\tau = \frac{2^d - 1 - 2^d \int_0^1 K(\alpha) d\alpha}{2^{d-1} - 1}.$$

Therefore:

$$\begin{aligned} \int_0^1 r(\alpha) d\alpha &= \int_0^1 (K(\alpha) - \alpha) d\alpha \\ &= \frac{2^d - 1 - (2^{d-1} - 1)\tau}{2^d} - \frac{1}{2} \\ &= (1 - \tau) \left(\frac{1}{2} - \frac{1}{2^d} \right). \end{aligned}$$

\square

In the bivariate case, this average difference is $(1 - \tau)/4$, which belongs to $[0, 1/2]$ due to the fact that $\tau \in [-1, 1]$. When d tends toward infinity, the average difference increases concomitantly with the dimension d , up to $(1 - \tau)/2 \in [0, 1]$. The case of the independent copula, for which $\tau = 0$, leads to an average r of $(1/2) - (1/2)^d$, whose value, $1/4$ for $d = 2$, progressively increases with the dimension up to $1/2$. It confirms the analysis made in example 2. If we consider comonotonic coordinates, then $\tau = 1$ and the average r is equal to zero, whatever the dimension d . Graphically, it corresponds to a case where all the vectors dominated by a reference vector belong to the lower-left quadrant of this reference vector: the order implied by the orthant VaRs, which is partial in general, is total in this particular case and there is no difference between the orthant and the Kendall VaRs. On the opposite, if the coordinates are countermonotonic, then $\tau = -1$ and the average r reaches its maximum: $1 - (1/2)^{d-1}$, which goes from $1/2$, for $d = 2$, to 1, when d goes toward infinity.

We can quantify as well the difference between the probability associated to a vector by the upper-orthant method and by the upper-orthant Kendall method: $\bar{r} : \alpha \in [0, 1] \mapsto \alpha - (1 - \mathcal{K}(1 - \alpha))$, which is a positive function. Proposition 6 states that the average twist of the risk level between the upper-orthant VaR and the upper-orthant Kendall VaR is, in absolute value, exactly the same as the average twist between the lower-orthant VaR and the standard Kendall VaR. In the framework of Proposition 1 where the upper-orthant and the standard Kendall VaRs have a non-empty intersection for a given risk level, the vectors belonging to both Kendall VaRs can thus be seen as a balanced compromise between lower- and upper-orthant VaRs, as they twist as much, in absolute value, the risk associated to both in average over all the possible risk levels. Nevertheless, for a particular level of risk, the Kendall VaR can be closer to the one or to the other.

Proposition 6. *The average difference between the probabilities associated to the sole survival copula and to the Kendall function of the survival copula, for d -dimensional vectors and a continuous copula, is:*

$$\int_0^1 \bar{r}(\alpha) d\alpha = (1 - \tau) \left(\frac{1}{2} - \frac{1}{2^d} \right),$$

where τ is the Kendall rank correlation coefficient.

Proof. By a change of variable, we have:

$$\begin{aligned} \int_0^1 \bar{r}(\alpha) d\alpha &= \int_0^1 (\mathcal{K}(\alpha) - \alpha) d\alpha \\ &= \int_0^1 \mathcal{K}(\alpha) d\alpha - \frac{1}{2}. \end{aligned}$$

Moreover, what we note \mathcal{K} is, according to Definition 2, the Kendall function corresponding to the survival distribution function. It can thus be written in terms of the Kendall's tau of the survival copula, $\bar{\tau}$:

$$\int_0^1 \mathcal{K}(\alpha) d\alpha = \frac{2^d - 1 - (2^{d-1} - 1)\bar{\tau}}{2^d}.$$

Besides, we know that the Kendall's tau of the survival copula is equal to the Kendall's tau of the copula itself [23], so that $\bar{\tau} = \tau$. This leads immediately to the result stated in the proposition. \square

In addition to the average error, we can calculate an upper bound, when α tends toward 0 or 1, of the limit error $r(\alpha)$ and more widely of $R(\alpha)$, which is the probability distortion between the lower-orthant and the upper-orthant VaRs. If $x \in \underline{\text{VaR}}_\alpha(G) \cap \bar{\text{VaR}}_{\alpha'}(G)$, the probability distortion is $\alpha' - \alpha$. It depends on α but also on the choice of x in $\underline{\text{VaR}}_\alpha(G)$. To fix ideas, we will focus on a particular x corresponding to equal marginal probabilities: $x = (G_1^{-1}(u), \dots, G_d^{-1}(u))$, for $u \in [0, 1]$ well chosen so as to have $x \in \underline{\text{VaR}}_\alpha(G)$. So we have $u = \mathcal{C}^{-1}(\alpha)$, where $\mathcal{C} : v \in [0, 1] \mapsto C(v, \dots, v)$,

with C the copula associated to the joint distribution G . This choice is possible only if C is invertible. Then, we define R by:

$$R : \alpha \in [0, 1] \mapsto 1 - \alpha - \overline{G}(G_1^{-1}(C^{-1}(\alpha)), \dots, G_d^{-1}(C^{-1}(\alpha))). \quad (3)$$

In Proposition 7, we propose an upper bound for $R(\alpha)/\alpha$.

This distortion $R(\alpha)$ is directly linked to the notion of tail dependence. For a bivariate variable, the lower tail dependence λ_L is the following limit, if it exists: $\lim_{\alpha \rightarrow 0} \mathbb{P}(X_1 \leq G_1^{-1}(\alpha) | X_2 \leq G_2^{-1}(\alpha))$. Thanks to Bayes' rule, this expression is symmetric in each component of the vector. Moreover, it only depends on the copula and not on the marginals. In higher dimension, one can define several lower tail dependence parameters corresponding to various choices of subsets $I_k \subset \{1, \dots, d\}$ of size k : $\lambda_{L, I_k} = \lim_{\alpha \rightarrow 0} \mathbb{P}(X_i \leq G_i^{-1}(\alpha), \forall i \in I_k | X_j \leq G_j^{-1}(\alpha), \forall j \in \bar{I}_k)$ [13, 15]. Contrarily to the case $d = 2$, this expression depends in general on the composition of I_k and not only on its cardinal. We will limit to a particular case of exchangeable copulas, for which $\lambda_{L, I_k} = \lambda_{L, I'_k}$ if $|I_k| = |I'_k|$. This assumption is in particular verified for Archimedean copulas [15] and we then write $\lambda_{L, k}$ instead of λ_{L, I_k} . Symmetrically, one can define upper tail dependence parameters. For instance, for bivariate variables, it is $\lambda_U = \lim_{\alpha \rightarrow 1} \mathbb{P}(X_1 > G_1^{-1}(\alpha) | X_2 > G_2^{-1}(\alpha))$, if the limit exists.

Proposition 7. *Let R be defined as in equation (3) for an exchangeable copula C such that $C : v \in [0, 1] \mapsto C(v, \dots, v)$ is invertible. If all the lower and upper tail dependence parameters exist and are noted $\lambda_{L, k}$ and $\lambda_{U, k}$, for $k \in \{1, \dots, d-1\}$, then the asymptotic difference $R(\alpha)$ between the probabilities associated to the lower- and upper-orthant VaRs is such that:*

$$\lim_{\alpha \rightarrow 0} \frac{R(\alpha)}{\alpha} \leq \frac{1}{\lambda_{L, d-1}} \sum_{k=1}^{d-1} \binom{d}{k} (1 - \lambda_{L, k}) \quad (4)$$

with equality only if the lower tail dependence parameters are all equal to 1, and

$$\lim_{\alpha \rightarrow 1} \frac{R(\alpha)}{1 - \alpha - R(\alpha)} \leq \frac{1}{\lambda_{U, d-1}} \sum_{k=1}^{d-1} \binom{d}{k} (1 - \lambda_{U, k}), \quad (5)$$

with equality only if the upper tail dependence parameters are all equal to 1.

Proof. First, we look for $u \in [0, 1]$ such that $\alpha \underset{\alpha \rightarrow 0}{\sim} G(G_1^{-1}(u), \dots, G_d^{-1}(u))$. By using the corresponding copula, this is equivalent to $\alpha \underset{\alpha \rightarrow 0}{\sim} C(u, \dots, u)$. By Bayes' rule, we thus should have

$$\alpha \underset{\alpha \rightarrow 0}{\sim} u \lambda_{L, d-1}.$$

Therefore, we define u as $\alpha / \lambda_{L, d-1}$.

Then, we define a vector $(x_1, \dots, x_d) = (G_1^{-1}(u), \dots, G_d^{-1}(u))$. Since $G(G_1^{-1}(u), \dots, G_d^{-1}(u)) = \alpha$, this vector belongs to $\text{VaR}_\alpha(G)$. It also belongs to $\text{VaR}_{\alpha + R(\alpha)}(G)$, by definition of $R(\alpha)$. By the way, the probability measure of the complement of the upper right quadrant of this vector (x_1, \dots, x_d) , that is to say $\alpha + R(\alpha)$, is such that:

$$\begin{aligned} \alpha + R(\alpha) &= 1 - \mathbb{P}(X_1 > x_1, \dots, X_d > x_d) \\ &= 1 - \mathbb{P}(G_1(X_1) > u, \dots, G_d(X_d) > u) \\ &= 1 - \mathbb{P}(U_1 > u, \dots, U_d > u), \end{aligned}$$

with U_1, \dots, U_d uniform variables linked by the same copula C as X_1, \dots, X_d . Then

$$\begin{aligned} \alpha + R(\alpha) &= \mathbb{P}(U_1 \leq u, \dots, U_d \leq u) \\ &\quad + \sum_{i=1}^d \mathbb{P}(U_1 \leq u, \dots, U_i > u, \dots, U_d \leq u) \\ &\quad + \sum_{i=1}^d \sum_{j=1, j \neq i}^d \mathbb{P}(U_1 \leq u, \dots, U_i > u, \dots, U_j > u, \dots, U_d \leq u) \\ &\quad + \dots \end{aligned}$$

whose asymptote as $\alpha \rightarrow 0$ is $\sum_{k=0}^{d-1} \binom{d}{k} P_{d-k}$, where $P_k = \mathbb{P}(U_1 \leq u, \dots, U_k \leq u, U_{k+1} > u, \dots, U_d > u)$, thanks to the assumption that the lower tail dependence parameter is constant for a given size of I_k , whatever the composition of the subset I_k .

Last, we observe that $P_d = C(u, \dots, u) \underset{\alpha \rightarrow 0}{\sim} \alpha$ and that, for $k \geq 1$,

$$\begin{aligned} P_k &= \mathbb{P}(U_{k+1} > u, \dots, U_d > u | U_1 \leq u, \dots, U_k \leq u) \mathbb{P}(U_1 \leq u, \dots, U_k \leq u) \\ &= (1 - \mathbb{P}(U_{k+1} \leq u, \dots, U_d \leq u | U_1 \leq u, \dots, U_k \leq u)) \mathbb{P}(U_1 \leq u, \dots, U_k \leq u), \end{aligned}$$

thanks to Bayes' rule. The asymptote of P_k , as $\alpha \rightarrow 0$, is thus $(1 - \lambda_{L,d-k}) \mathbb{P}(U_1 \leq u, \dots, U_k \leq u) \leq (1 - \lambda_{L,d-k})u$ according to the upper Fréchet-Hoeffding bound, with equality only if $\lambda_{L,d-k} = 1$ or, if focusing on the Fréchet-Hoeffding inequality, if the variables are comonotonic, what also implies that lower-tail dependence parameters are equal to 1. This leads to equation (4).

Concerning equation (5), we observe that the upper tail dependence parameters of a random vector are equal to the lower tail dependence of the opposite of the vector. We can thus directly apply equation (4) to $(-X_1, \dots, -X_d)$, for a probability level $\bar{\alpha}$, when it tends toward 0, a difference function \bar{R} and lower tail dependence parameters $\bar{\lambda}_{L,k} = \lambda_{U,k}$. For a vector $(-x_1, \dots, -x_d)$ belonging to the lower-orthant VaR of level $\bar{\alpha}$ of the distribution of $(-X_1, \dots, -X_d)$, if the probability measure of its upper orthant is α , then $\bar{\alpha} = 1 - (\alpha + R(\alpha))$, noting that this upper orthant is the lower orthant of (x_1, \dots, x_d) . Moreover, $\bar{R}(\bar{\alpha})$ is 1 minus the probability measure of both the lower and upper orthants, therefore $\bar{R}(\bar{\alpha}) = 1 - \alpha - \bar{\alpha} = R(\alpha)$. \square

Proposition 7 gives upper bound for the difference of probability associated to the lower-orthant and the upper-orthant approaches. Naturally, the level of risk associated to the corresponding Kendall VaR is between lower-orthant and upper-orthant measures. In particular, $r(\alpha) \leq R(\alpha)$, what provides us with an upper bound for $r(\alpha)$. When $d = 2$, inequalities in Proposition 7 are simplified and upper bounds in equations (4) and (5) are respectively $2(\lambda_L^{-1} - 1)$ and $2(\lambda_U^{-1} - 1)$. As special cases, if the lower tail dependence is strong, λ_L is close to 1 and the upper bound in equation (4) is close to 0: lower-orthant, upper-orthant and Kendall VaRs are very close in the lower tail. On the contrary, when the lower tail dependence is weak, λ_L is close to 0 and the upper bound in equation (4) tends to infinity: lower-orthant, upper-orthant and Kendall VaRs are very disparate in the lower tail.

In Proposition 8, we will use the result of Proposition 7 in the particular framework of Archimedean copulas with regularly varying generators.

3.3 Kendall Value at Risk for an Archimedean copula

In this section, we assume that the multivariate distribution of the random vector X of dimension d is provided by an Archimedean copula C of generator ϕ :

$$C : (u_1, \dots, u_d) \in [0, 1]^d \mapsto \phi^{-1} \left(\sum_{j=1}^d \phi(u_j) \right).$$

It is a wide class of copulas which includes the following copulas: independent, Gumbel, Clayton, Frank, Joe and Ali-Mikhail-Haq, among others. Moreover, this framework leads to simple expressions for the Kendall function, so that it is an interesting illustration to our theory.⁶ We make

⁶ It is known that Archimedean copulas can be difficult to use in high dimension for estimation purpose, nevertheless the vine approach permits to bypass this problem. Vine copulas are indeed based on nested bivariate copulas instead of a sole high-dimension copula [10, 25, 28, 29]. Statistical selection techniques may help to truncate the vine so as to reduce the dimension of the problem in a relevant way [6]. For non-Archimedean copulas, semi-parametric methods may be used to estimate the Kendall function [39].

some assumptions on ϕ :

- ▷ $\phi : (0, 1] \rightarrow [0, \infty)$,
- ▷ $\phi(1) = 0$,
- ▷ $(-1)^i (\phi^{-1})^{(i)}(x) > 0$ for all $1 \leq i \leq d$ and all $x \geq 0$,⁷
- ▷ $\lim_{t \rightarrow 0+} \phi(t)^i (\phi^{-1})^{(i)}(\phi(t)) = 0$ for all $1 \leq i \leq d - 1$.

Then, the Kendall distribution function can be derived as:

$$K : t \in (0, 1] \mapsto t + \sum_{i=1}^{d-1} \frac{(-\phi(t))^i}{i!} (\phi^{-1})^{(i)}(\phi(t)), \quad (6)$$

where $f^{(i)}$ denotes the i -th derivative of f [3, 22]. We now apply this formula in two examples.

Example 1. *The Gumbel copula is an Archimedean copula of parameter $\theta \geq 1$, generated by the function*

$$\phi : t \mapsto (-\log(t))^\theta.$$

When $\theta = 1$, the Gumbel copula is equal to the independent copula. The inverse generator is $\phi^{-1}(x) = \exp(-x^{1/\theta})$. According to equation (6), if we consider the bivariate case, the Kendall function is:

$$K : t \mapsto t - \frac{t \log(t)}{\theta}.$$

We show in Figure 2 how the Kendall function behaves when θ changes: the greater θ , the closer the Kendall function and the identity. In particular, when θ tends toward infinity, K converges toward the identity, so that the Kendall VaR and the lower-orthant VaR are equal in this limit case.

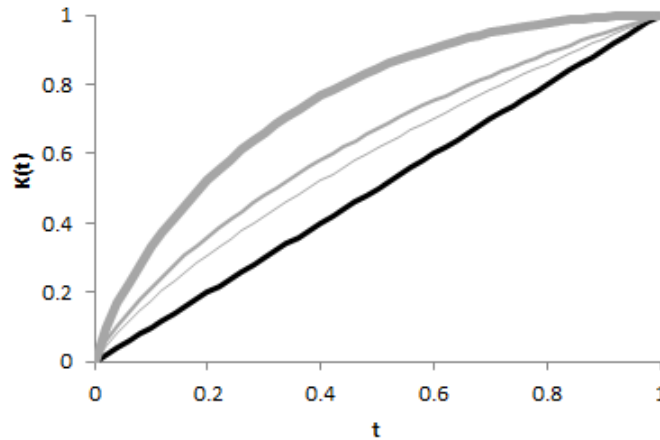


Figure 2: Kendall function (in grey) of the Gumbel copula for $\theta = 1$ (thick line), $\theta = 2$ (medium line) and $\theta = 3$ (thin line). The greater the difference between the Kendall function and the identity (in black, corresponding to $\theta \rightarrow \infty$), the greater the difference between the Kendall VaR and the lower-orthant VaR.

⁷ In particular, ϕ is strictly decreasing and convex.

Example 2. The independent copula leads to easy formulas in higher dimension. It is a particular case of the Gumbel copula with $\theta = 1$. According to equation (6), for a dimension $d \geq 2$, we get the following formula for K :

$$K : t \mapsto t \left(1 + \sum_{i=1}^{d-1} \frac{(-\log(t))^i}{i!} \right).$$

When d goes to infinity, $\sum_{i=1}^{d-1} \frac{(-\log(t))^i}{i!}$ tends toward $-1 + e^{-\log(t)} = -1 + 1/t$, for every $t \neq 0$. Therefore, the limit behaviour of K , for $d \rightarrow \infty$ is a discontinuous function, equal to 0 for $t = 0$ and equal to 1 everywhere else. It leads to the maximal difference possible between the Kendall VaR and the lower-orthant VaR. We see in Figure 3 the Kendall function for various values of d .

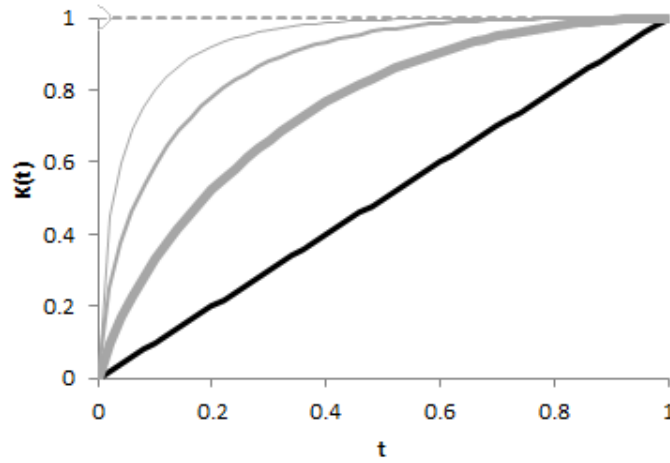


Figure 3: Kendall function (in grey) of the independent copula for $d = 2$ (thick line), $d = 3$ (medium line), $d = 4$ (thin line) and the limit case $d \rightarrow \infty$ (dotted line). The greater the difference between the Kendall function and the identity (in black), the greater the difference between the Kendall VaR and the lower-orthant VaR.

We saw in Proposition 7, that the probability distortion between orthant and Kendall approaches was linked to the tail dependence. This is true whatever the copula. In the case of Archimedean copulas, we have an additional result, allowing to link the probability distortion to the regular variations of the inverse generator. We first recall the definition of a regularly varying function:

Definition 3. A function f is regularly varying at 0, with index ρ , if

$$\forall s > 0, \lim_{x \rightarrow 0^+} \frac{f(sx)}{f(x)} = s^\rho.$$

We then note $f \in \mathcal{RV}_\rho(0)$.

Proposition 8. If, for a given bivariate Archimedean copula, the inverse generator $\phi^{-1} \in \mathcal{RV}_{-\rho}(0)$, with $\rho > 0$, then the asymptotic difference $R(\alpha)$ between the probabilities associated to the lower- and upper-orthant VaRs, as defined in equation (3), is such that:

$$\lim_{\alpha \rightarrow 0} \frac{R(\alpha)}{\alpha} \leq \sum_{k=1}^{d-1} \binom{d}{k} \left(d^{1/\rho} - (d-k)^{1/\rho} \right). \quad (7)$$

Proposition 8 is a direct consequence of Proposition 7 with Theorem 2.1 of lower tail coefficients in [15]. In equation (7), for bivariate variables, the upper bound is $2(2^{1/\rho} - 1)$. The faster varying ϕ^{-1} , the greater $|\rho|$ and the closer to zero the upper bound of equation (7). On the contrary, slowly varying inverse generators are associated to a big probability distortion between lower-orthant, upper-orthant and Kendall approaches.

4 Application and Results

In this section we apply the methodology presented above and evaluate the Kendall VaR using various copulas. This section is divided in three subsections. The first illustrates the characteristics of the copulas we are working with. In a second subsection, we illustrate the probability transformation implied by the Kendall distribution. Finally, in a third subsection we present and compare orthant and Kendall VaRs.

Indeed, in order for us to properly analyse the results of the Kendall VaR, it is necessary to analyse them with respect to the form of the copula underlying the Kendall distribution used to transform the percentile. Note that in the following applications we focus on Archimedean copulas, as their properties are of particular interest; for instance, these are non-linear, may capture different dependence patterns and may behave as extreme-value copulas.

4.1 Archimedean Copula

In this first section we briefly illustrate the characteristics of Archimedean copulas, in particular their non-linearity. This non-linearity is visible for the four bivariate Archimedean copulas presented in Figure 4, for instance the Gumbel which is upper-tail dependent, i.e. extreme positive events have a tendency to occur simultaneously while others are independent, the Clayton copula which is lower-tail dependent, i.e. extreme negative events have a tendency to occur simultaneously while others are independent, the Frank copula, which is more body-centered, i.e. events present in the body are more dependent than those present in the tails, and the Joe copula, which is also upper-tail dependent but with a dependency magnitude far steeper than in the Gumbel copula case.

Figures 5 and 6 illustrate both the Clayton and Gumbel copulas density and distribution functions exhibiting asymmetric behavior. The asymmetry is of particular importance to capture asymmetric shocks in financial markets as we will demonstrate in a companion paper.

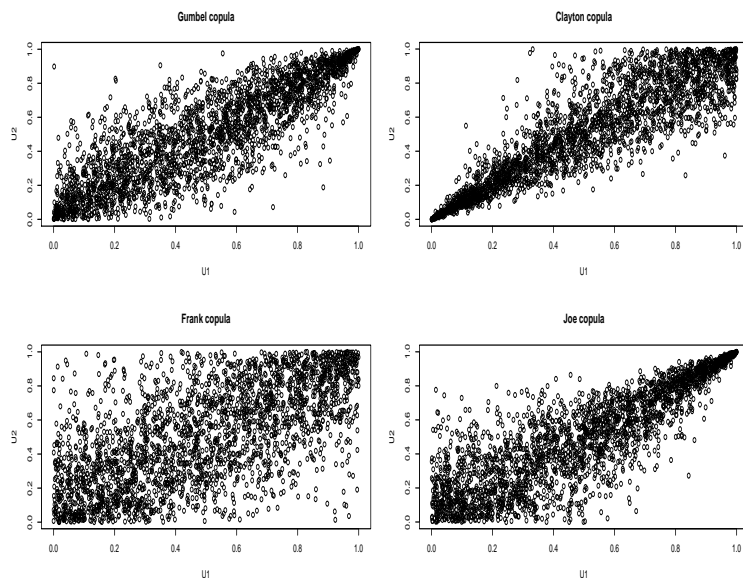


Figure 4: Archimedean copula illustrations: This figure represents four Archimedean copulas, the Gumbel which is upper-tail dependent, the Clayton which is lower-tail dependent, the Frank copula which is more body centered and the Joe which is also upper-tail dependent.

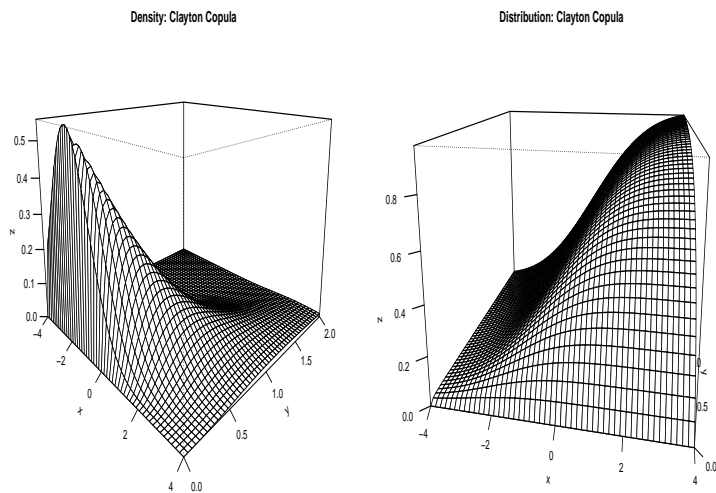


Figure 5: This figure presents the density and distribution functions of a Clayton copula whose margins are a Gaussian distribution $N(0,2)$ and an exponential distribution with parameter 2. We observe that the functions are asymmetric.

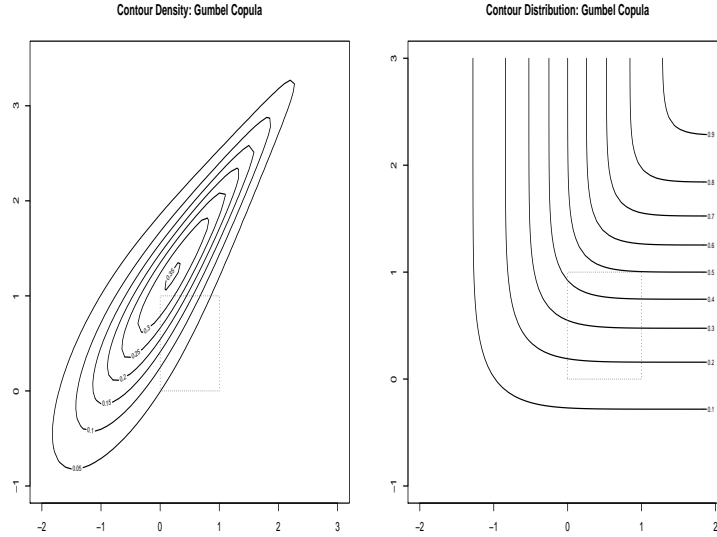


Figure 6: This figure presents the density and distribution functions of a Gumbel copula whose margins are Gaussian, $N(0,1)$ and $N(1,1)$ respectively. Here, a contour function has been used to transform the figure in three dimensions in two. Note that the distribution function on the right is a first representation of the lower-orthant quantiles presented below.

4.2 Kendall Distribution

In Section 3.1, we introduce the Kendall distribution which relies on an underlying copula. Figure 7 shows how the Kendall distribution evolves with the type of copula we are using. Each Kendall distribution link corresponds to a set of percentiles and the bissector represents the percentiles of the orthant approach. It is interesting to note that the shape of the Kendall distribution obtained with a particular copula is consistent with the scatter plots of the copulas presented in Figure 4. In other words, if the copula captures an upper-tail dependance behaviour, the Kendall distribution inflexion point is located in the left tail of the distribution (see Kendall distribution relying on a Gumbel or a Joe copula). However, if the copula captures a lower-tail dependance behaviour, the Kendall distribution inflexion point is located in the right tail of the distribution (see Kendall distribution relying on a Clayton copula). The Kendall transformation breaks the linearity provided inside the orthant approach.

Our interpretation of the behaviour of the Kendall distribution is that it captures the properties of the copula in its construction and subsequently in the calculation of the corresponding quantile while the standard orthant methodology loses a part of the information calculating the quantiles, as the orthant methodology partially suppress the effect of the non-linear dependency.

In Figure 8, we present the dynamic twist of Kendall distributions built using the four copulas presented in Figure 4. Here, the parameters of the four copulas evolved from 1 to 25, and the Kendall distributions obtained are respectively plotted on four figures. As discussed earlier, we see that the Kendall distribution corresponds to a probability twist.

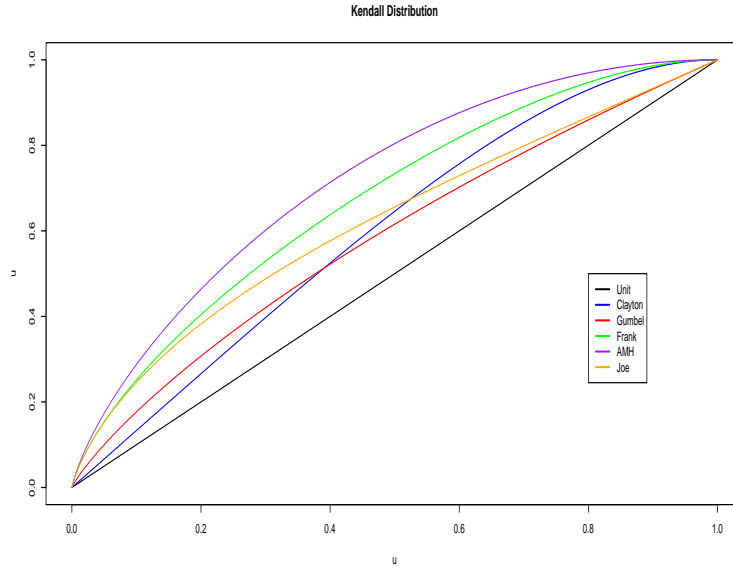


Figure 7: This figure presents the various Kendall distributions forms obtained using multiple Archimedean copulas, for instance the Clayton, the Gumbel, the Frank, the Ali-Mikhail-Haq and the Joe copulas. The transformations are compared to the bisector of the unit square representing the percentile of the orthant approach.

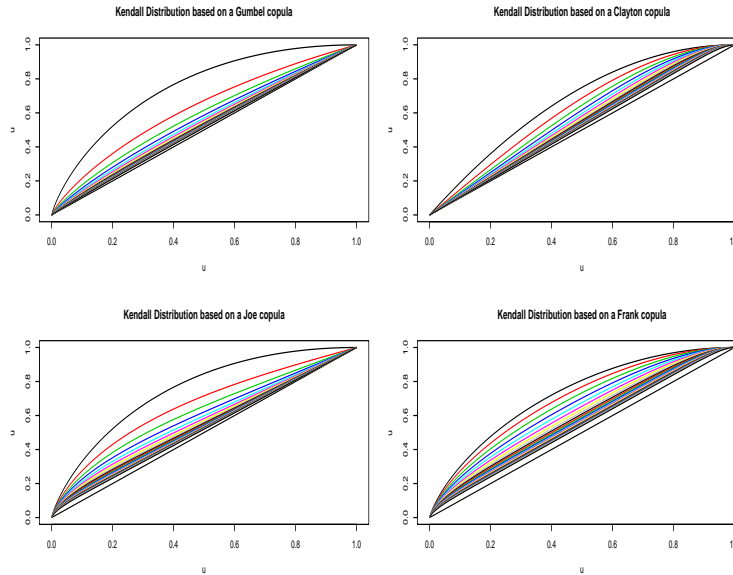


Figure 8: This figure presents the Kendall distributions obtained considering four Archimedean copulas, for instance Clayton, Gumbel, Frank and Joe. The parameters of each copula are slowly varying from 1 to 25. The transformation are compared to the bisector of the unit square representative of the percentile of orthant approach.

4.3 Multivariate VaR representation and calculation

In this section, we compare and discuss the sets of VaRs obtained from both the lower-orthant strategy and the Kendall approach. Recall that multivariate VaRs will be represented by vectors. To initiate our experimentation, we build two copula functions, for instance the Clayton and the Gumbel copulas, with parameters equal to 3. Both have been constructed using two lognormal marginal distributions, with the following sets of parameters ($\mu = 5, \sigma = 2$) and ($\mu = 8, \sigma = 1.2$).

Figures 9⁸ and 10 represent the lower-orthant VaRs obtained using the Clayton copula presented in the previous paragraph. Lower-orthant VaRs are obtained calculating all the combinations of all pairs of margins providing an identical bivariate probability.

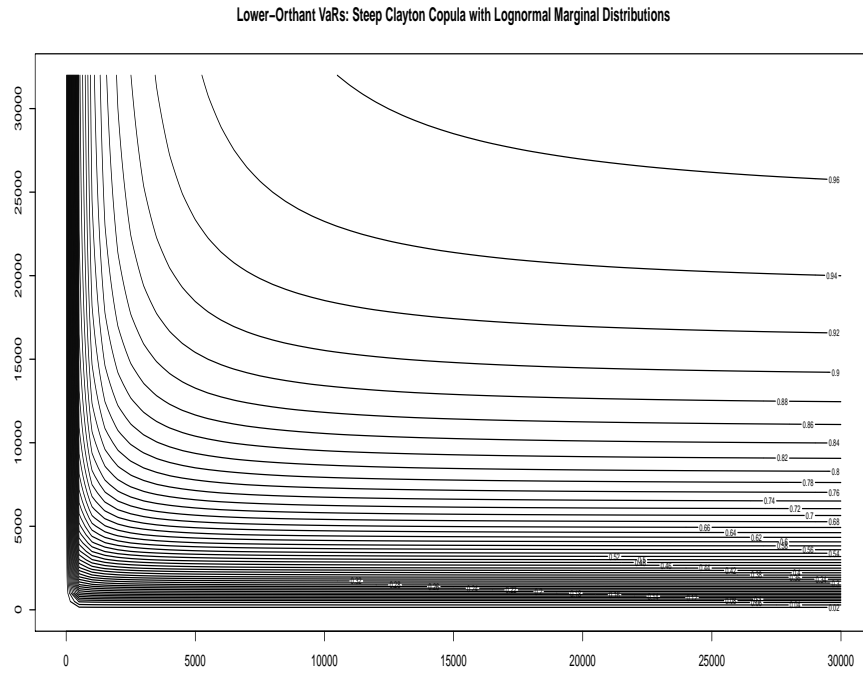


Figure 9: This figure presents the quantile function of a steep Clayton copula with two different lognormal marginal distributions. Here each quantile is represented by a vector. Each axis represents the projection of the multivariate quantile on each marginal distribution. Each line of the graph represent a percentile, going from 0 to 96%.

As analyzed in the previous subsection, the Kendall distribution transforms the natural probabilities taking into account the shape of the copula. This transformation allows us to calculate in a similar fashion the lower-orthant VaR, the Kendall VaR transforming the lower-orthant percentile into the Kendall one.

Figures 11, 12 and 13 compare lower-orthant, upper-orthant and Kendall VaRs using the Clayton and the Gumbel copulas built as explained in the first paragraph of this subsection. In these figures,

⁸We have observed that the result is similar with a lot of copulas.

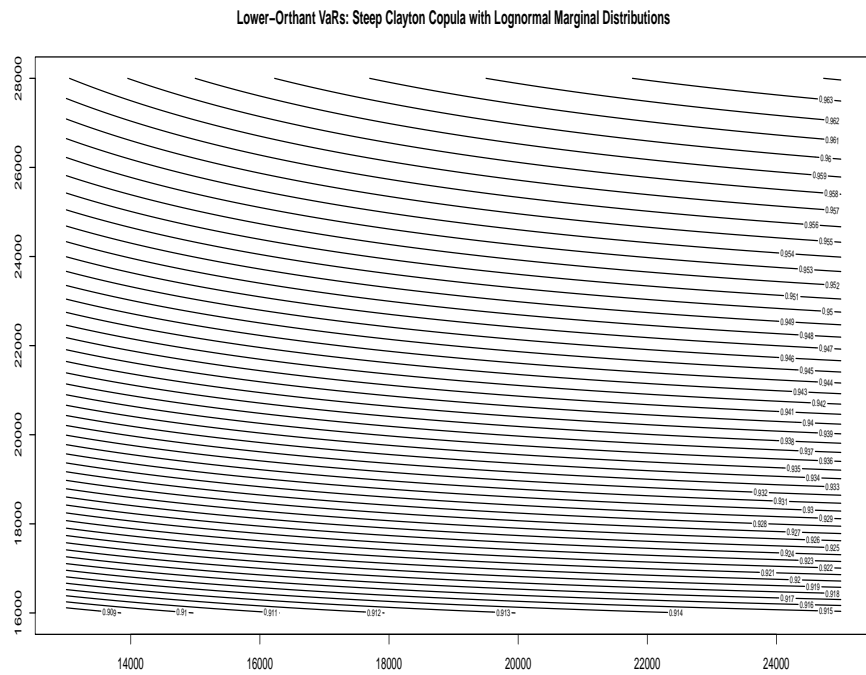


Figure 10: This figure presents the quantile function of a steep Clayton copula with two different lognormal marginal distributions. Here each quantile is represented by a vector. In this figure, we focus on the highest quantiles. Each axis represents the projection of the multivariate quantile on each marginal distribution. Each line of the graph represent a percentile, going from 90.9% to 96.4%.

the dotted lines represent the Kendall VaR, the continuous lines located above⁹ the dotted lines represent the lower-orthant VaR, and the continuous lines located below, the upper-orthant VaR. On each figure, the VaRs are given at the same percentile, but we see that the Kendall VaR is not equivalent to the lower-orthant as the Kendall distributions twist the probabilities. As a result, in our case the lower-orthant equivalent of the Kendall VaRs given a percentile α is much lower than the lower-orthant ones obtained at the same percentile α . On the contrary, upper-orthant VaRs are much less conservative. Indeed, Figures 11, 12 and 13 compare lower-orthant VaRs obtained for α respectively equal to 86%, 98% and 71% (approximately) with their Kendall equivalent, i.e. for $K(\alpha)$ also respectively equal to 86%, 98% and 71%. To obtain the previous $K(\alpha)$, α has to respectively be equal to 80%, 89% and 56%. Besides, Figures 11 and 12 provide a comparison of the Kendall VaR respectively obtained for $K(\alpha)$ equal to 62% and 65%, with upper-orthant VaRs obtained for $1 - \alpha$ respectively equal to 38% and 35%. Figure 12 provides upper-orthant, Kendall and lower-orthant VaR on the same graph, but here the upper-orthant is obtained for $1 - \alpha$ equal to 29%.

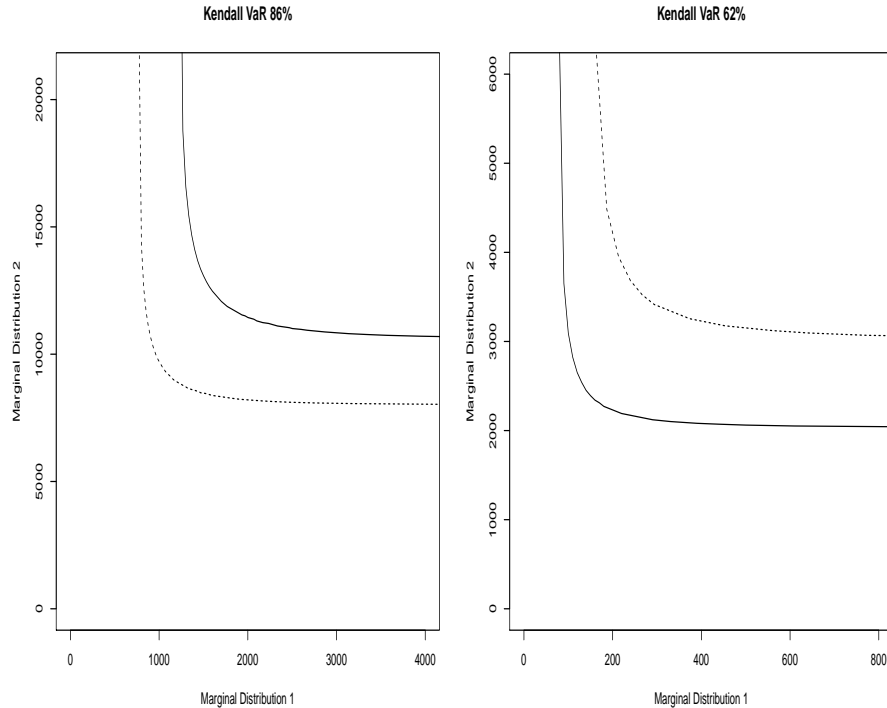


Figure 11: This figure represents the Kendall VaR obtained using a Gumbel copula with two different lognormal marginal distributions. The figure on the left compares the 86% Kendall VaR to its lower-orthant equivalent. We observe that the Kendall VaR is inferior to the lower-orthant one. The figure on the right compares the 62% Kendall VaR to the upper-orthant VaR computed from the survival copula of the Gumbel. Here, we see that the Kendall VaR is superior to the upper-orthant one.

With respect to the marginal distribution selected and the copulas chosen, the difference between the lower-orthant quantile and the Kendall one is not linear, i.e. the quantile may be closer to each

⁹ The comparison between these sets of vectors must be understood in the sense of the lexicographical order for each pair of vectors.

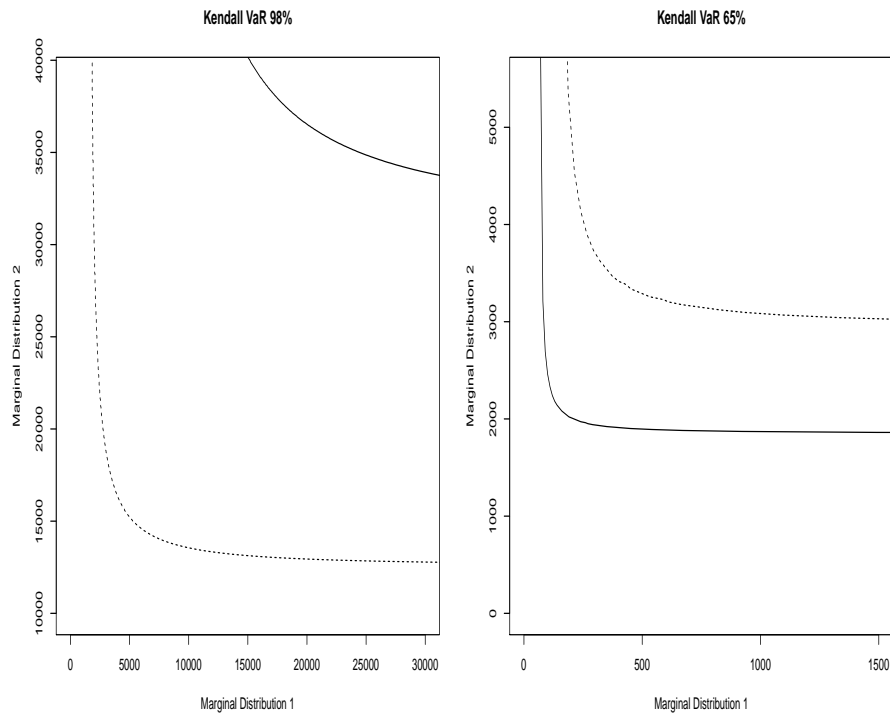


Figure 12: This figure represents the Kendall VaR obtained using a Clayton copula with two different lognormal marginal distributions. The figure on the left compares the 98% Kendall VaR to its lower-orthant equivalent. We observe that the Kendall VaR is inferior to the lower-orthant one. The figure on the right compares the 65% Kendall VaR to the upper-orthant VaR computed from the survival copula of the Clayton. Here, we see that the Kendall VaR is superior to the upper-orthant one.

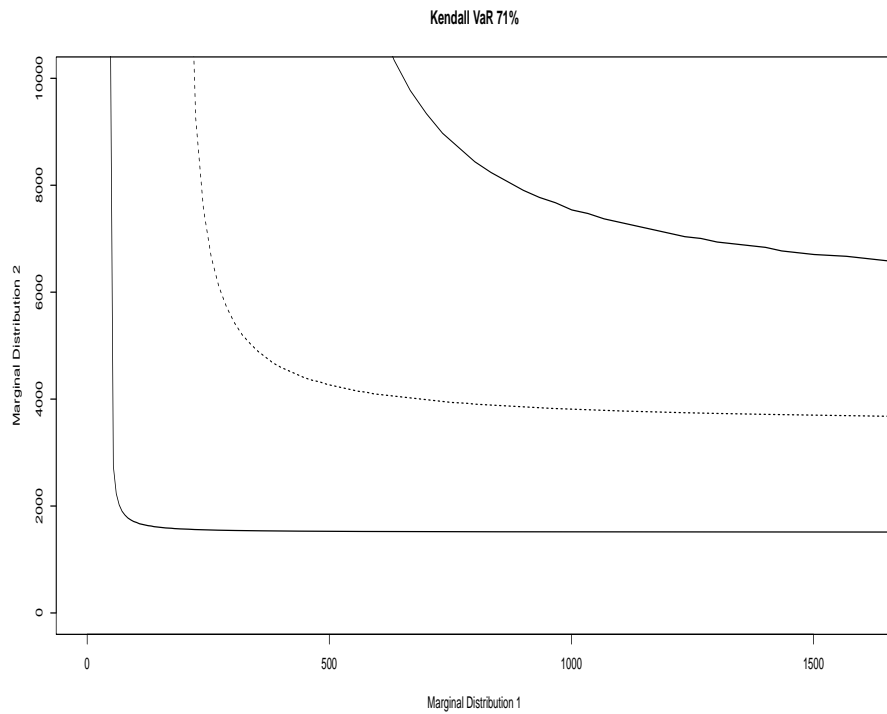


Figure 13: This figure represents the Kendall VaR obtained using a Clayton copula with two different lognormal marginal distributions. The figure compares the 71% Kendall VaR to its lower-orthant equivalent and to the upper-orthant obtained using the survival copula of the Clayton used here. We observe that the Kendall VaR is inferior to the lower-orthant one and superior to the upper-orthant.

other depending on the dependence structure selected (both in terms of copula and parameters). This statement is consistent with what we observed in Figure 7.

The next statement is valid for both the Kendall and the orthant VaRs. On the last five figures presented, for instance Figures 9, 10, 11, 12 and 13, it is very important to note that taking each point of any multivariate quantile represented by the lines on the figures above and summing each couple of projections, this sum is not constant as it evolves along the line. This phenomenon is a consequence of the non-linearity of the dependence structure captured by the copula used. Indeed, if we consider the Gumbel copula which is upper tail dependent, (i.e. extreme positive events have a tendency to occur simultaneously), the bivariate VaR at a given percentile (or multivariate VaR in a dimension superior to 2) is represented by a matrix where each line represents different combinations having the same probability¹⁰ but not necessarily the same projected values. In other words, for a given percentile, a bivariate VaR can be represented by a very large projection on a first marginal distribution axis and a very low one on the second marginal distribution axis, or two projections of approximately the same magnitude, but the sum of these two different combinations will be completely different. This particular phenomenon will have to be bore in mind in case either of the methodologies proposed here is implemented. Indeed, if we take the example of the use of multivariate risk measures for financial institutions' regulatory capital, then these methodologies will lead to both capital calculation and allocation issues as all the combination obtained at a given percentile are equally valid.

5 Conclusion

In this paper, introducing the Kendall VaR, the issues related to the formalisation and the calculation of a multivariate VaR have been addressed. Indeed, the Kendall VaR allows the calculation of a multivariate VaR using the probability transformation implied by the Kendall distribution. For instance, the Kendall distribution captures the intrinsic characteristics of the dependence architecture represented by the selected copula (non-linearity, upper- or lower-tail dependence etc.) and transfers it in one dimension. Therefore the Kendall distribution allows operating a percentile transformation. The Kendall VaR is more conservative than the upper-orthant VaR and less conservative than the lower-orthant VaR.

In a first part of the paper, after introducing this new approach to build a multivariate risk measure, that we call the Kendall VaR, we provide a simple relationship between the Kendall VaR and the orthant VaRs, which allows us defining the Kendall VaR as a compromise between the bounds represented by both orthant VaRs. Finally we quantify, the difference between the Kendall VaR and the orthant VaRs and we also link this asymptotic difference to tail dependence parameters.

In a second part, we illustrate, with different examples the fact that for any multivariate distribution based on a copula, the Kendall VaR can be constructed and plotted gathering all the sets of values of identical probability.

Besides, the Kendall VaR created is compared to the orthant VaRs. The probability of the latter has not suffered any modifications and therefore partially loses the information captured by the copula. Indeed, the Kendall VaR is more appropriate than the orthant VaRs when copulas capturing non-linear dependencies are used, as the Kendall VaR does not lose any information with respect to the property of interest. We notice that the Kendall VaR and the orthant VaRs are

¹⁰ A point of each line.

equal in case of comonotonicity.

Building and comparing the two approaches, we observed that the non-linearity of the copulas implies that the sums of each set representing a given percentile are not constant. This phenomenon will have an important impact if any of these methodologies is used within financial institutions (for instance banks or insurance companies), as if these approaches are used to evaluate the diversified capital pertaining to the various risks faced by them, the accurate value of the capital as well as the allocation of this one will be problematic. Indeed, multiple sets of values will be representative of the same level of risks going from one end to the other. In terms of applications, this result provides a variety of possible interpretations, it will be the purpose of a companion paper.

References

- [1] ABDOUS, B. AND THEODORESCU, R. (1992), *Note on the spatial quantile of a random vector*, Statistics and probability letters, 13, 4: 333-336
- [2] ARTZNER, PH., DELBAEN, F., EBER, J.-M. AND HEATH, D. (1998), *Coherent measures of risk*, Finance and Stochastics, 9, 3: 203-228
- [3] BARBE, P., GENEST, C., GHOUDI, K. AND RÉMILLARD, B. (1996), *On Kendall's process*, Journal of multivariate analysis, 58, 2: 197-229
- [4] BELZUNCE, F., CASTAÑO, A., OLIVERA-CERVANTES, A. AND SUÁREZ-LLORENS, A. (2007), *Quantile curves and dependence structure for bivariate distributions*, Computational statistics and data analysis, 51, 1: 5112-5129
- [5] BRECHMANN, E. (2014), *Hierarchical Kendall copulas: Properties and inference*, Canadian journal of statistics, 42, 1: 78-108
- [6] BRECHMANN, E., CZADO, C. AND AAS, K. (2012), *Truncated regular vines in high dimensions with application to financial data*, Canadian journal of statistics, 40, 1: 68-85
- [7] CHAOUCH, M., GANNOUN, A. AND SARACCO, J. (2009), *Estimation de quantiles géométriques conditionnels et non conditionnels* Journal de la société française de statistique, 150, 2: 1-27
- [8] CHAUDHURI, P. (1996), *On a geometric notion of quantiles for multivariate data*, Journal of the American statistical association, 91, 434: 862-872
- [9] CHENG, Y. AND DE GOOIJER, J. (2004), *On the u -th Geometric Conditional Quantile*, Tinbergen institute discussion paper
- [10] COOKE, R. M. (1997), *Markov and entropy properties of tree-and vine-dependent variables*, Proceedings of the ASA section of Bayesian statistical science, 27
- [11] COSSETTE, H., MAILHOT, M., MARCEAU, É. AND MESFIOUI, M. (2013), *Bivariate lower and upper orthant value-at-risk*, European actuarial journal, 3, 2: 321-357
- [12] COUSIN, A. AND DI BERNARDINO, E. (2013), *On multivariate extensions of value-at-risk*, Journal of multivariate analysis, 119, 1: 32-46

- [13] DE LUCA, G. AND RIVIECCIO, G. (2012), *Multivariate tail dependence coefficients for Archimedean copulae*, in *Advanced statistical methods for the analysis of large data-sets* (pp. 287-296), Springer, Berlin, Heidelberg
- [14] DELBAEN, F. (2002), *Coherent risk measures on general probability spaces*, in *Advances in finance and stochastics. Essays in honour of Dieter Sondermann*, Springer, Berlin
- [15] DI BERNARDINO, E. AND RULLIÈRE, D. (2016), *On tail dependence coefficients of transformed multivariate Archimedean copulas*, Fuzzy sets and systems, 284: 89-112
- [16] EMBRECHTS, P. AND PUC CETTI, G. (2006), *Bounds for functions of multivariate risks*, Journal of multivariate analysis, 97, 2: 526-547
- [17] FERGUSON, T. (1967), *Mathematical statistics: A decision theoretic approach*, Academic press, New York
- [18] FRAIMAN, R. AND MELOCHE, J. (1999), *Multivariate L-estimation*, Test, Sociedad de estadística e investigación operativa, 8, 2: 255-317
- [19] GARCIN, M. AND GUÉGAN, D. (2012), *Extreme values of random or chaotic discretization steps and connected networks*, Applied mathematical sciences, 6, 119: 5901-5926
- [20] GENEST, C. AND FAVRE, A. (2007), *Everything you always wanted to know about copula modeling but were afraid to ask*, Journal of hydrologic engineering, 12, 4: 347-368
- [21] GENEST, C. AND RIVEST, L. (1993), *Statistical inference procedures for bivariate Archimedean copulas*, Journal of the American statistical association, 88, 423: 1034-1043
- [22] GENEST, C. AND RIVEST, L. (2001), *On the multivariate probability integral transformation*, Statistics and probability letters, 53, 4: 391-399
- [23] GEORGES, P., LAMY, A.G., NICOLAS, E., QUIBEL, G. AND RONCALLI, T. (2001), *Multivariate survival modelling: a unified approach with copulas*, Available at SSRN 1032559
- [24] DE GOOIJER, J., GANNOUN, A. AND ZEROM, D. (2006), *A multivariate quantile predictor*, Communications in statistics - theory and methods, 35, 1: 133-147
- [25] GUÉGAN, D. AND HASSANI, B. K. (2013), *Multivariate VaRs for operational risk capital computation: a vine structure approach*, International journal of risk assessment and management, 17, 2: 148-170
- [26] GUÉGAN, D. AND HASSANI, B. K. (2015), *Risk or Regulatory capital? Bringing distributions back in the foreground*, Publications du CES Paris1 Panthéon - Sorbonne, WP 2015-46
- [27] GUÉGAN, D. AND HASSANI, B. K. (2016), *More Accurate Measurement for Enhanced Controls: VaR vs ES?*, Publications du CES Paris1 Panthéon - Sorbonne, WP 2016-15
- [28] GUÉGAN, D. AND MAUGIS, P.-A. (2010), *An econometric study of vine copulas*, Available at SSRN 1590296
- [29] GUÉGAN, D. AND MAUGIS, P.-A. (2010), *New prospects on vines*, Insurance markets and companies: analyses and actuarial computations, 1, 1: 4-11
- [30] GUÉGAN, D. AND TARRANT, W. (2012), *On the necessity of five risk measures*, Annals of finance, 8, 4: 533-552
- [31] JOUINI, E., MEDDEB, M. AND TOUZI, N. (2004), *Vector-valued coherent risk measures*, Finance and stochastics, 8, 4: 531-552

- [32] KOLTCHINSKII, V. (1997), *M-estimation, convexity and quantiles*, The annals of statistics, 25, 2: 435-477
- [33] LALAURETTE, F. (2003), *Early detection of abnormal weather conditions using a probabilistic extreme forecast index*, Quarterly journal of the royal meteorological society, 129, 594: 3037-3057
- [34] LIU, R., PARELIUS, J. AND SINGH, K. (1999), *Multivariate analysis by data depth: descriptive statistics, graphics and inference (with discussions and rejoinder)*, The annals of statistics, 27, 3: 783-858
- [35] MAHALANOBIS, P. (1936), *On the generalized distance in statistics*, Proceedings of the national institute of sciences (Calcutta), 2, 1: 49-55
- [36] NELSEN, R. B., QUESADA-MOLINA, J. J., RODRÍGUEZ-LALLENA, J. A. AND ÚBEDA-FLORES, M. (2003), *Kendall distribution functions*, Statistics and probability letters, 65, 3: 263-268
- [37] NEUHÄUSER, D., HIRSCH, C., GLOAGUEN, C. AND SCHMIDT, V. (2015), *Joint distributions for total lengths of shortest-path trees in telecommunication networks*, Annals of telecommunications, 70, 5-6: 221-232
- [38] NOYAN, N. AND RUDOLF, G. (2013), *Optimization with multivariate conditional value-at-risk constraints*, Operations research, 61, 4: 990-1013
- [39] SALVADORI, G., DURANTE, F. AND PERRONE, E. (2013), *Semi-parametric approximation of Kendall's distribution function and multivariate return periods*, Journal de la société française de statistique, 154,1: 151-173
- [40] SATO, M., ICHIKI, K. AND TAKEUCHI, T. (2011), *Copula cosmology: Constructing a likelihood function*, Physical review D, 83, 2: 023501
- [41] SCHERRER, R., BERLIND, A., MAO, Q. AND MCBRIDE, C. (2009), *From finance to cosmology: The copula of large-scale structure*, The astrophysical journal letters, 708, 1: L9
- [42] SERFLING, R. (2002), *Quantile functions for multivariate analysis: approaches and applications*, Statistica Neerlandica, 56, 2: 214-232
- [43] SKLAR, A. (1959), *Fonctions de répartition à n dimensions et leurs marges*, Publ. inst. statist. univ. Paris, 8: 229-231
- [44] THORARINSDOTTIR, T.L., SCHEUERER, M. AND HEINZ, C. (2016), *Assessing the calibration of high-dimensional ensemble forecasts using rank histograms*, Journal of computational and graphical statistics, 25, 1: 105-122
- [45] ZUO, Y. AND SERFLING, R. (2000), *General notions of statistical depth function*, The annals of statistics, 28, 2: 461-482