A CONSISTENT ESTIMATOR TO THE ORTHANT-BASED TAIL VALUE-AT-RISK[☆,☆☆]

Nicholas Beck[1] and Mélina Mailhot[2,*]

Abstract. In this paper, we address the estimation of multivariate value-at-risk (VaR) and tail value-at-risk (TVaR). We recall definitions for the bivariate lower and upper orthant VaR and bivariate lower and upper orthant TVaR, presented in Cossette *et al.* [*Eur. Actuar. J.* **3** (2013) 321–357 or *Methodol. Comput. Appl. Probab.* (2014) 1–22]. Here, we present estimators for both these measures extended to an arbitrary dimension $d \geq 2$ and establish the consistency of our estimator for the lower and upper orthant TVaR in any dimension. We demonstrate these results by providing numerical examples that compare our estimator to theoretical results for both simulated and real data.

1. Introduction

Evaluating risk is a crucial task in many industries, particularly in finance and insurance. Regulators establish solvency capital requirements and risk allocation in order to protect stakeholders and customers. With the increasing strictness of solvency requirements in recent years, see for instance OSFI [18] or Solvency II [11] regulations, more research is becoming available on the development of the appropriate measures. In the univariate setting, for a random variable (rv) X and cumulative distribution function (cdf) F_X, the value-at-risk (VaR) at level $\alpha \in (0, 1)$ defined by

$$\text{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\},$$

is often used for capital allocation requirements. However, it has become increasingly criticized for its lack of coherence, as defined by Artzner *et al.* [1]. Specifically, VaR is not a subadditive risk measures. Moreover, it provides little information about the tail of a distribution. As an alternative to VaR, other measures exist that do recognize the risk in the tail and are subadditive. One such measure is the tail value-at-risk (TVaR), defined

[1] Department of Mathematics and Statistics, McGill University, Montreal, QC, Canada.

[2] Department of Mathematics and Statistics, Concordia University, Montreal, QC, Canada.

[*] Corresponding author: melina.mailhot@concordia.ca

by

$$\mathrm{TVaR}_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 \mathrm{VaR}_u(X)\mathrm{d}u.$$

TVaR allows one to calculate the average loss at level α given that the loss is greater than the α-level VaR.

Recently, there has been an ample amount of research conducted on the development of extensions of VaR and TVaR, as well as other risk measures, in the multivariate setting. Embrechts and Puccetti [10] were first to introduce the notion of multivariate VaR through use of level sets. Jouini *et al.* [13] presents the multivariate worst conditional expectation (WCE), from an axiomatic representation of multivariate risk measures. In [14], using the framework established in [13], WCE is studied and the Generalized WCE is defined. Vector-valued risk measures for conditional tail expectation (CTE), conditional VaR (CVaR) and VaR are presented in [3], [16] and [6], respectively, the last of which differing from the level set definition of [10].

Much work has also been devoted to the efficient computation of these values in the empirical setting. In particular, the VaR and CVaR have received a lot of attention. Rockafellar and Uryasev [19] develop the idea of optimizing CVaR (and by consequence, VaR) for continuous distributions. Specifically, they show that calculation of CVaR can be written in terms of a convex optimization problem making it easy to calculate and estimate. Additionally, the methods developed here allow for one to skip calculating VaR entirely, which can drastically decrease computation time. Rockafellar and Uryasev [20] further develop these concepts in the case of more complicated distributions which might be more representative of real world situations. Bardou *et al.* [2] take these ideas a step further by first developing a stochastic gradient descent algorithm in tandem with the CVaR representation of Rockafellar and Uryasev [19, 20]. They continue to streamline the algorithm by incorporating importance sampling. Specifically, they allow for a moving sequence of significance levels which can further improve efficiency and decrease estimation error. These methodologies allow them to establish several central limit theorem (CLT) type results for their estimation procedures.

In this paper, our focus is on the consistent estimation of the multivariate lower and upper orthant VaR and TVaR. The lower (upper) orthant VaR was introduced by Embrechts and Puccetti [10], in the bivariate context, as a level set generated by the cdf (survival function (sf)) of a random vector **X**. Both the orthant VaR and TVaR were later developed by Cossette *et al.* [4, 5], respectively, with multivariate extensions of each discussed in [15]. In the univariate context, estimators of VaR and TVaR can be written

$$\mathrm{VaR}_\alpha^n(X) = \inf\{x \in \mathbb{R} : F_n(x) \geq \alpha\}, \tag{1.1}$$

and

$$\mathrm{TVaR}_\alpha^n(X) = \frac{1}{1-\alpha} \int_\alpha^1 \mathrm{VaR}_u^n(X)\mathrm{d}u, \tag{1.2}$$

where F_n is the empirical cdf. Using the reparametrization of the orthant VaR introduced by Cossette *et al.* [4], we may write the estimators for the multivariate orthant VaR and TVaR in a similar fashion to (1.1) and (1.2), respectively. We will call these the empirical orthant VaR and TVaR. With consistency results established by Di Bernardino *et al.* [9], we can prove that the empirical orthant VaR is consistent for the true multivariate orthant VaR. Finally, the consistency of the empirical orthant TVaR is shown using the dominated convergence theorem. Note that we focus on consistency in the lower orthant, however, the results are easily extendable to the upper orthant. To estimate our measures, we will use Monte Carlo simulations. However, many of the techniques outlined above could certainly be incorporated and provide an avenue for further research.

The paper is organized as follows. In Section 2, we present the orthant-based VaR and recall some results that will be needed to establish the consistency of its estimator. We end the section by introducing the orthant-based TVaR, as defined by Cossette *et al.* [5]. In Section 3, we turn our attention to the estimation of the orthant-based TVaR. Here, our estimator is introduced and its almost sure consistency is established. Section 4

demonstrates a few examples of this estimator with simulated data as well as comparisons to the multivariate CTE of Di Bernardino *et al.* [9] using the lossALAE data set of Frees and Valdez [12]. Concluding remarks and further research interests are given in Section 5.

2. NOTATION AND PRELIMINARIES

We introduce some of the notation and definitions that will be of use throughout the paper. We consider the random vector $\mathbf{X} = (X_1, \ldots, X_d) \in \mathbb{R}_+^d$ of dependent risks with joint cdf F and joint sf $\bar{F}$. Additionally, we define the random vector $\mathbf{X}_{\backslash i} = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_d)$ with cdf $F_{\backslash i}$ and sf $\bar{F}_{\backslash i}$, $i = 1, \ldots, d$. Similarly, for $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}_+^d$, a realization of $\mathbf{X}$, consider the vector $\boldsymbol{x}_{\backslash i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d)$ and the functions

$$F_{\boldsymbol{x}_{\backslash i}} : x_i \mapsto F(\boldsymbol{x}) \text{ and } \bar{F}_{\boldsymbol{x}_{\backslash i}} : x_i \mapsto \bar{F}(\boldsymbol{x}).$$

Now, we define the lower and upper orthant VaR. As our applications are in actuarial science, we are often dealing with losses that are positively valued. Therefore, we have restricted ourselves to the positive orthant of $\mathbb{R}^d$.

Definition 2.1. For a random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with joint cdf F, we define the lower orthant VaR by

$$\underline{\mathrm{VaR}}_\alpha(\mathbf{X}) = \partial\{\boldsymbol{x} \in \mathbb{R}_+^d : F(\boldsymbol{x}) \geq \alpha\}$$

$$= \bigcup_{i=1}^d \left\{ \left(x_1, \ldots, x_{i-1}, \underline{\mathrm{VaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}), x_{i+1}, \ldots, x_d\right) : x_j \geq \mathrm{VaR}_\alpha(X_j), \ \forall j \neq i \right\}, \tag{2.1}$$

where ∂A represents the boundary of a set A. Moreover, the lower orthant VaR for a rv X_i when $\mathbf{X}_{\backslash i} = \boldsymbol{x}_{\backslash i}$ is given by

$$\underline{\mathrm{VaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}) = \inf\{x_i \in \mathbb{R}_+ : F_{\boldsymbol{x}_{\backslash i}}(x_i) \geq \alpha\}.$$

Similarly, we define the upper orthant VaR as follows.

Definition 2.2. For a random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with joint sf $\bar{F}$ we define the upper orthant VaR by

$$\overline{\mathrm{VaR}}_\alpha(\mathbf{X}) = \partial\{\boldsymbol{x} \in \mathbb{R}_+^d : \bar{F}(\boldsymbol{x}) \leq 1 - \alpha\}$$

$$= \bigcup_{i=1}^d \left\{ \left(x_1, \ldots, x_{i-1}, \overline{\mathrm{VaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}), x_{i+1}, \ldots, x_d\right) : x_j < \mathrm{VaR}_\alpha(X_j), \ \forall j \neq i \right\}, \tag{2.2}$$

where

$$\overline{\mathrm{VaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}) = \inf\{x_i \in \mathbb{R}_+ : \bar{F}_{\boldsymbol{x}_{\backslash i}}(x_i) \leq 1 - \alpha\}$$

is the upper orthant VaR for a rv X_i when $\mathbf{X}_{\backslash i} = \boldsymbol{x}_{\backslash i}$.

Remark 2.3. It should be noted that both the existence and finiteness of $\underline{\mathrm{VaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X})$ and $\overline{\mathrm{VaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X})$ is linked directly to the existence and finiteness of $\mathrm{VaR}_\alpha(X_i)$ for $i = 1, \ldots, d$. Recall that $\mathrm{VaR}_\alpha(X)$ exists and is finite for all $\alpha \in (0, 1)$ but may take values at infinity for $\alpha = 1$. Thus, the existence of $\underline{\mathrm{VaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X})$ and $\overline{\mathrm{VaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X})$ is guaranteed given the existence of its univariate counterparts. Similarly, $\underline{\mathrm{VaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X})$ and

$\overline{\text{VaR}}_{\alpha,\boldsymbol{x}\setminus i}(\mathbf{X})$ are finite, except asymptotically, where Cossette *et al.* [4] showed that

$$\lim_{x_j \to \text{VaR}_\alpha(X_j)} \underline{\text{VaR}}_{\alpha,\boldsymbol{x}\setminus i}(\mathbf{X}) = u_{X_i}, \ \ \forall j \neq i,$$

where $u_{X_i} = \sup\{x : x \in \text{supp}(X_i)\}$ is the essential supremum of X_i. A similar result exists for $\overline{\text{VaR}}_{\alpha,\boldsymbol{x}\setminus i}(\mathbf{X})$ and the essential infimum of X_i.

Remark 2.4. It can be seen that both $F_{\boldsymbol{x}\setminus i}(x_i) = F(\boldsymbol{x})$ and $\bar{F}_{\boldsymbol{x}\setminus i}(x_i) = \bar{F}(\boldsymbol{x})$. However, this notation is used throughout to emphasize the process when calculating orthant VaR and TVaR. For these calculations, one fixes all random variables but one and calculates the required value for the final unfixed random variable.

Remark 2.5. For simplicity, we will restrict ourselves to formulations (2.1) and (2.2). Additionally, for brevity, we rewrite each point in the unions of (2.1) and (2.2) by $\left(\boldsymbol{x}_{\setminus i}, \underline{\text{VaR}}_{\alpha,\boldsymbol{x}\setminus i}(\mathbf{X})\right)$ and $\left(\boldsymbol{x}_{\setminus i}, \overline{\text{VaR}}_{\alpha,\boldsymbol{x}\setminus i}(\mathbf{X})\right)$, respectively, keeping note that we are still preserving the proper order of the coordinates.

Next, we define the Hausdorff distance. This measure, used in Di Bernardino *et al.* [9] to establish the consistency of their empirical orthant VaR, is crucial to establishing the consistency of our empirical orthant TVaR.

Definition 2.6. Consider the ball of radius ρ centered at a point $\boldsymbol{x}$ denoted by $B(\boldsymbol{x}, \rho) = \{\boldsymbol{y} \in \mathbb{R}^d : \delta(\boldsymbol{x}, \boldsymbol{y}) \leq \rho\}$, where δ is the Euclidean distance. For a set S, the ball of radius ρ about S is denoted $B(S, \rho) = \cup_{\boldsymbol{x} \in S} B(\boldsymbol{x}, \rho)$. Then, for A_1, A_2 compact sets in the measure space $(\mathbb{R}^2_+, \delta)$, the Hausdorff distance between these sets is represented by

$$d_H(A_1, A_2) = \inf\left\{\rho > 0 : A_1 \subset B(A_2, \rho), A_2 \subset B(A_1, \rho)\right\}$$

$$= \max\left\{\sup_{\boldsymbol{x} \in A_1} \delta(\boldsymbol{x}, A_2), \sup_{\boldsymbol{x} \in A_2} \delta(\boldsymbol{x}, A_1)\right\},$$

where $\delta(\boldsymbol{x}, A_1) = \min_{\boldsymbol{y} \in A_1} \|\boldsymbol{x} - \boldsymbol{y}\|$.

Since VaR can approach infinity, the Hausdorff distance may be infinite. To curb this issue, we introduce a truncated version of $\underline{\text{VaR}}_\alpha(\mathbf{X})$, that is

$$\underline{\text{VaR}}_\alpha(\mathbf{X})^{T_n} = \left\{\left(\boldsymbol{x}_{\setminus i}, \underline{\text{VaR}}_{\alpha,\boldsymbol{x}\setminus i}(\mathbf{X})\right) : \text{VaR}_\alpha(X_j) \leq x_j \leq T_n, \ \ \forall j \neq i\right\}, \tag{2.3}$$

$i = 1, \ldots, d$ where $\{T_n\}_{n \in \mathbb{N}}$ is an increasing sequence. In essence, the sequence T_n is a method for controlling the orthant VaR at lower sample sizes, ensuring finite Hausdorff distance and allowing us to use the results established by Di Bernadino and Laloë [8]. Moreover, since $T_n \xrightarrow[n \to \infty]{} \infty$ we have

$$\underline{\text{VaR}}_\alpha(\mathbf{X})^{T_n} \xrightarrow[n \to \infty]{} \underline{\text{VaR}}_\alpha(\mathbf{X}).$$

Next, we introduce a few expressions that will be of use throughout the paper. For $\alpha \in (0, 1)$ and $r, \zeta > 0$, consider the ball

$$E = B\left(\left\{\boldsymbol{x} \in \mathbb{R}^d_+ : |F(\boldsymbol{x}) - \alpha| \leq r\right\}, \zeta\right).$$

Additionally, consider the infimum of the Euclidean norm of the gradient vector evaluated at $\boldsymbol{x}$ for a twice differentiable function F, denoted $m^\nabla = \inf_{\boldsymbol{x} \in E} \|(\nabla F)_{\boldsymbol{x}}\|$. Similarly, define by $M_H = \sup_{\boldsymbol{x} \in E} \|(HF_{\boldsymbol{x}})\|$, the supremum of the matrix norm, induced by the Euclidean norm, of the Hessian matrix evaluated at $\boldsymbol{x}$. Finally, we

set $\{F = \alpha\} = \left\{ \boldsymbol{x} \in \mathbb{R}_+^d : F(\boldsymbol{x}) = \alpha \right\}$ and $\{F = \alpha\}^T = \left\{ \boldsymbol{x} \in [0, T]^d : F(\boldsymbol{x}) = \alpha \right\}$, which is used in the following assumption.

Assumption H. There exists $\gamma > 0$ and $A > 0$, such that if $|\alpha_1 - \alpha_2| \leq \gamma$, then for all $T > 0$ such that $\{F = \alpha_1\}^T \neq \emptyset$ and $\{F = \alpha_2\}^T \neq \emptyset$,

$$d_H(\{F = \alpha_1\}^T, \{F = \alpha_2\}^T) \leq A|\alpha_1 - \alpha_2|.$$

For details on the conditions, see Cuevas *et al.* [7]. This assumption removes the possibility of horizontal asymptotes in F. With some extra conditions on F, the constant A can be specified. Note that this is a multivariate extension of Proposition 2.1 from [9], which can be found in [8].

Proposition 2.7. *Let $\alpha_1 \in (0, 1)$ and F twice differentiable on $\mathbb{R}_+^d$. Assume there exists $r > 0$ and $\zeta > 0$ such that $m^\nabla > 0$ and $M_H < \infty$. Then F satisfies Assumption H, with $A = \frac{2}{m^\nabla}$.*

To end the section, we introduce the truncated L^∞ norm which will be of use in the remaining sections. For two functions f and g,

$$\|f - g\|_\infty^T = \sup_{\boldsymbol{x} \in [0,T]^d} |f(\boldsymbol{x}) - g(\boldsymbol{x})|,$$

for some $T \in \mathbb{R}^+$.

2.1. Lower orthant tail value-at-risk

We present multivariate lower and upper orthant TVaR, introduced by Cossette *et al.* [5]. These measures provide a natural extension to the univariate TVaR. They are easy to calculate and are represented by curves, which provides flexibility in applications. We define the multivariate lower and upper orthant TVaR below.

Definition 2.8. For a random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with joint cdf F, the lower orthant TVaR of $\mathbf{X}$ at level $\alpha \in (0, 1)$ is defined by

$$\underline{\text{TVaR}}_\alpha(\mathbf{X}) = \bigcup_{i=1}^d \left\{ (x_1, \ldots, x_{i-1}, \underline{\text{TVaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}), x_{i+1}, \ldots, x_d) : x_j > \text{VaR}_\alpha(X_j), \ \forall \, j \neq i \right\},$$

where

$$\underline{\text{TVaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}) = \mathbb{E}\left[X_i | X_i > \underline{\text{VaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}), \mathbf{X}_{\backslash i} \leq \boldsymbol{x}_{\backslash i} \right], \ x_j \geq \text{VaR}_\alpha(X_j),$$

for all $j = 1, 2, \ldots, d, \ i \neq j$.

The following proposition provides an interesting parallel for $\underline{\text{TVaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X})$ to the univariate TVaR.

Proposition 2.9. *For all $x_i \geq \text{VaR}_\alpha(X_i)$, $i = 1, 2, \ldots, d$,*

$$\underline{\text{TVaR}}_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}) = \frac{1}{F_{\backslash i}(\boldsymbol{x}_{\backslash i}) - \alpha} \int_\alpha^{F_{\backslash i}(\boldsymbol{x}_{\backslash i})} \underline{\text{VaR}}_{u, \boldsymbol{x}_{\backslash i}}(\mathbf{X}) \mathrm{d}u, \quad i = 1, 2, \ldots, d. \tag{2.4}$$

Proof. Refer to [5]. □

In Figure 1, we provide an illustration of the lower orthant VaR and TVaR in relation to their univariate counterparts.
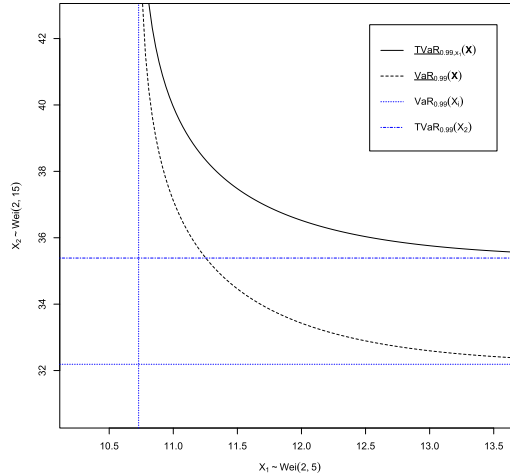
FIGURE 1. $x_1 \mapsto \underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})$ and $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ at $\alpha = 0.99$ for $X_1 \sim \mathrm{Weibull}(2,5)$ and $X_2 \sim \mathrm{Weibull}(2,15)$ linked by a Frank copula with Kendall's $\tau = 0.5$.

2.2. Upper orthant tail value-at-risk

Considering the lower orthant TVaR, which uses the joint cdf and looks at the upper tail of a distribution, is not always appropriate. For instance, when interested in returns on stocks one may be concerned with small or, in the case of $\mathbb{R}^d$, negative values. To this end, it would be beneficial to study the lower tail of a distribution. In this case one considers the upper orthant TVaR. For a multivariate sf $\bar{F}$, we can define the upper orthant TVaR as follows.

Definition 2.10. For a random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with multivariate sf $\bar{F}$, the upper orthant TVaR of $\mathbf{X}$ at level $\alpha \in (0,1)$ is defined by

$$\overline{\mathrm{TVaR}}_\alpha(\mathbf{X}) = \bigcup_{i=1}^{d} \left\{ \left(x_1, \ldots, x_{i-1}, \overline{\mathrm{TVaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}), x_{i+1}, \ldots, x_d\right) : x_j \le \mathrm{VaR}_\alpha(X_j),\ \forall\, j \ne i \right\},$$

where

$$\overline{\mathrm{TVaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}) = \mathbb{E}\left[X_i | X_i > \overline{\mathrm{TVaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}), \mathbf{X}_{\setminus i} \ge \boldsymbol{x}_{\setminus i}\right],\ x_j \le \mathrm{VaR}_\alpha(X_j),$$

for all $j = 1, \ldots, d,\ i \ne j$.

Similarly to $\underline{\mathrm{TVaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X})$, we can reformulate the upper orthant TVaR in terms of $\overline{\mathrm{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X})$.

Proposition 2.11. *For all* $x_i \le \mathrm{VaR}_\alpha(X_i)$, $i = 1, \ldots, d$,

$$\overline{\mathrm{TVaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}) = \frac{1}{1-\alpha} \int_\alpha^1 \overline{\mathrm{VaR}}_{u,\boldsymbol{x}_{\setminus i}}(\mathbf{X}) \mathrm{d}u,\ \ i = 1, \ldots, d. \tag{2.5}$$

Proof. Refer to [5].   □

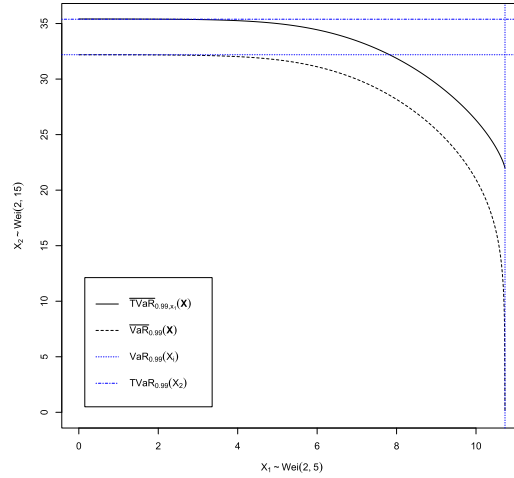In Figure 2, we provide an example of the upper orthant VaR and TVaR in relation to their univariate counterparts.

FIGURE 2. $x_1 \mapsto \overline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})$ and $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$ at $\alpha = 0.99$ for $X_1 \sim$ Weibull$(2, 5)$ and $X_2 \sim$ Weibull$(2, 15)$ linked by a Frank copula with Kendall's $\tau = 0.5$

3. ESTIMATION OF THE MULTIVARIATE ORTHANT TVAR

With the formulation of Propositions 2.9 and 2.11, we introduce our estimators, beginning with the lower orthant TVaR.

3.1. Lower orthant TVaR

Definition 3.1. Consider a series of observations $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_d)$ with $\mathbf{X}_i = (x_{1i}, \ldots, x_{ni})$ and $\mathbf{X}_{\setminus i} = (\mathbf{X}_1, \ldots, \mathbf{X}_{i-1}, \mathbf{X}_{i+1}, \ldots, \mathbf{X}_d)$, $i = 1, \ldots, d$. Denote F_n and $F_{n,\setminus i}$, the empirical cdf's (ecdf) for $\mathbf{X}$ and $\mathbf{X}_{\setminus i}$, $i = 1, \ldots, d$, respectively. We define the estimator for the lower orthant TVaR for fixed $\boldsymbol{x}_{\setminus i}$, $i = 1, \ldots, d$, by

$$\underline{\mathrm{TVaR}}^n_{\alpha, \boldsymbol{x}_{\setminus i}}(\mathbf{X}) = \frac{1}{F_{n,\setminus i}(\boldsymbol{x}_{\setminus i}) - \alpha} \int_\alpha^{F_{n,\setminus i}(\boldsymbol{x}_{\setminus i})} \underline{\mathrm{VaR}}^n_{u, \boldsymbol{x}_{\setminus i}}(\mathbf{X}) \mathrm{d}u \tag{3.1}$$

$$= \sum_{j=1}^m \frac{\underline{\mathrm{VaR}}^n_{u_j, \boldsymbol{x}_{\setminus i}}(\mathbf{X}) \cdot s}{F_{n,\setminus i}(\boldsymbol{x}_{\setminus i}) - \alpha},$$

for $m \in \mathbb{N}$ large enough. Setting $s = \frac{F_{n,\setminus i}(\boldsymbol{x}_{\setminus i}) - \alpha}{m}$ and $u_j = \alpha + j \cdot s$, (3.1) simplifies to

$$\underline{\mathrm{TVaR}}^n_{\alpha, \boldsymbol{x}_{\setminus i}}(\mathbf{X}) = \sum_{j=1}^m \frac{\underline{\mathrm{VaR}}^n_{u_j, \boldsymbol{x}_{\setminus i}}(\mathbf{X})}{m},$$

where $\underline{\mathrm{VaR}}^n_{u, \boldsymbol{x}_{\setminus i}}(\mathbf{X}) = \inf \left\{ x_i \in \mathbb{R}_+ : F_{n, \boldsymbol{x}_{\setminus i}}(x_i) \geq u \right\}$ is the empirical lower orthant VaR for a given $\boldsymbol{x}_{\setminus i}$ and $F_{n, \boldsymbol{x}_{\setminus i}} : x_i \mapsto F_n(\boldsymbol{x})$.

3.2. Upper orthant TVaR

When considering the multivariate sf, we have the estimator for $\overline{\mathrm{TVaR}}_{\alpha, x_1}(\mathbf{X})$.

Definition 3.2. Consider a series of observations $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_d)$ with $\mathbf{X}_i = (x_{1i}, \ldots, x_{ni})$ and $\mathbf{X}_{\setminus i} = (\mathbf{X}_1, \ldots, \mathbf{X}_{i-1}, \mathbf{X}_{i+1}, \ldots, X_d)$, $i = 1, \ldots, d$. Denote F_n and $F_{n,\setminus i}$, the empirical sf's (esf) for $\mathbf{X}$ and $\mathbf{X}_{\setminus i}$,

$i = 1, \ldots, d$, respectively. We define the estimator for the lower orthant TVaR for fixed $\boldsymbol{x}_{\backslash i}$, $i = 1, \ldots, d$, by

$$\overline{\mathrm{TVaR}}^n_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}) = \frac{1}{1-\alpha} \int_\alpha^1 \overline{\mathrm{VaR}}^n_{u, \boldsymbol{x}_{\backslash i}}(\mathbf{X}) \mathrm{d}u \tag{3.2}$$
$$= \sum_{j=1}^m \frac{\overline{\mathrm{VaR}}^n_{u_j, x_{\backslash i}}(\mathbf{X}) \cdot s}{1-\alpha},$$

for $m \in \mathbb{N}$ large enough. Setting $s = \frac{1-\alpha}{m}$ and $u_j = \alpha + j \cdot s$ (3.2) simplifies to

$$\overline{\mathrm{TVaR}}^n_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}) = \sum_{j=1}^m \frac{\overline{\mathrm{VaR}}^n_{u_j, \boldsymbol{x}_{\backslash i}}(\mathbf{X})}{m},$$

where $\overline{\mathrm{VaR}}^n_{u, \boldsymbol{x}_{\backslash i}}(\mathbf{X}) = \inf\left\{ x_i \in \mathbb{R}_+ : \bar{F}_{n, x_{\backslash i}}(x_i) \leq 1 - u \right\}$ is the empirical upper orthant VaR for a given $\boldsymbol{x}_{\backslash i}$ and $\bar{F}_{n, \boldsymbol{x}_{\backslash i}} : x_i \mapsto \bar{F}_n(\boldsymbol{x})$.

3.3. Consistency

When working with estimators, the property of consistency is crucial. With enough data, a consistent estimator allows one to accurately estimate the value of interest without specification, and the possibility of misspecification, of a statistical model. We focus on the estimator of the lower orthant TVaR. To show that (3.1) is in fact a consistent estimator for (2.9), we use the dominated convergence theorem. Our first step will be to establish the pointwise convergence of the function $\underline{\mathrm{VaR}}^n_{\alpha, x_1}(\mathbf{X})$ to $\underline{\mathrm{VaR}}_{\alpha, x_1}(\mathbf{X})$. To do so, we begin by establishing the consistency in Hausdorff distance of the empirical multivariate lower orthant VaR, represented by

$$\underline{\mathrm{VaR}}^n_\alpha(\mathbf{X}) = \partial\{\boldsymbol{x} \in \mathbb{R}_+^d : F_n(\boldsymbol{x}) \geq \alpha\},$$

as seen in Cuevas *et al.* [7]. Due to the restriction of compactness from the Hausdorff distance, we establish a truncated version of $\underline{\mathrm{VaR}}^n_\alpha(\mathbf{X})$ which is analogous to (2.3), introduced in Di Bernardino *et al.* [9], defined by

$$\underline{\mathrm{VaR}}^n_\alpha(\mathbf{X})^{T_n} = \partial\left\{ \boldsymbol{x} \in [0, T_n]^d : F_n(\boldsymbol{x}) \geq \alpha \right\}$$
$$= \bigcup_{i=1}^d \left\{ \left(\boldsymbol{x}_{\backslash i}, \underline{\mathrm{VaR}}^n_{\alpha, \boldsymbol{x}_{\backslash i}}(\mathbf{X}) \right) : \mathrm{VaR}^n_\alpha(X_j) \leq x_j \leq T_n \ \ \forall j \neq i \right\}.$$

The following theorem establishes the consistency in Hausdorff distance. This is an extension of the proof for the bivariate case provided in Di Bernardino *et al.* [9] to $\mathbb{R}^d$, which was originally described in Cuevas *et al.* [7].

Theorem 3.3. *Let $\alpha_1 \in (0, 1)$. Let F be twice differentiable on $\mathbb{R}^d$. Assume there exists $r > 0$, $\zeta > 0$ such that $m^\nabla > 0$ and $M_H < \infty$. Let $T_1 > 0$ such that for all α_2 with $|\alpha_2 - \alpha_1| \leq r$, $\underline{\mathrm{VaR}}_{\alpha_2}(\mathbf{X})^{T_1} \neq \emptyset$. Let $(T_n)_{n \in \mathbb{N}}$ be an increasing sequence of positive values. Assume that, for each n, F_n is continuous with probability one and that*

$$\|F - F_n\|_\infty \to 0 \ \ \text{with probability one (wp1)}.$$

Then,

$$d_H(\underline{\mathrm{VaR}}_{\alpha_1}(\mathbf{X})^{T_n}, \underline{\mathrm{VaR}}^n_{\alpha_1}(\mathbf{X})^{T_n}) = \mathcal{O}(\|F - F_n\|_\infty), \ \ wp1.$$

We note that the quality of the estimator $\underline{\text{VaR}}_\alpha^n(\mathbf{X})$ is linked to the choice of F_n. This will be discussed further in Section 4. Using Theorem 3.3, we establish the pointwise convergence of $\underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}^n(\mathbf{X})$ to $\underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X})$.

Corollary 3.4. *Under the assumptions of Theorem 3.3, and the representation of $\underline{\text{VaR}}_\alpha(\mathbf{X})$ given by (2.1), we have that*

$$\underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}^n(\mathbf{X}) \xrightarrow[n\to\infty]{wp1} \underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}),$$

pointwise for all $\boldsymbol{x}_{\setminus i}$.

Before proving Corollary 3.4, we provide the following definition and lemma:

Definition 3.5. For M, a compact metric space, we denote the convergence of sets $A, A_n \in M$ as $A_n \rightsquigarrow A$ in the power set of M, denoted 2^M, if for every $\epsilon > 0$, we have that $A_n \subseteq B(A, \epsilon)$ and $A \subseteq B(A_n, \epsilon)$, for all large enough n.

Lemma 3.6. *If sets A and A_n satisfy $A_n \rightsquigarrow A$, then if $x \in A$, there exists $x_n \in A_n$ with $x_n \to x$.*

Proof. Take $x \in A$ and let $x_n \in A_n$ denote the point minimizing the distance from x to A_n. Since $A \subseteq B(A_n, \epsilon)$, for all large enough n, we conclude that $\delta(x, x_n) = \delta(x, A_n)$ is eventually smaller than ϵ, for all $\epsilon > 0$, so $\delta(x, x_n) \to 0$. $\qquad\square$

Now we may prove Corollary 3.4.

Proof. It should be clear from Theorem 3.3, that for

$$A_i = \left\{ \left(\boldsymbol{x}_{\setminus i}, \underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}) \right) \right\} \quad \text{and} \quad A_{i,n} = \left\{ \left(\boldsymbol{x}_{\setminus i}, \underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}^n(\mathbf{X}) \right) \right\},$$

$x_j \geq \text{VaR}_\alpha(X_j)$, $j \neq i$ and $i = 1, \ldots, d$, one has $A_{i,n}^T \rightsquigarrow A_i^T$ for all $T \in \mathbb{R}$ large enough. Now, take a point $\left(\boldsymbol{x}_{\setminus i}, \underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}) \right) \in A_i$. Next, choose $T^* \in \mathbb{R}$ such that $\left(\boldsymbol{x}_{\setminus i}, \underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}) \right) \in A_i^{T^*}$. By Lemma 3.6, there exists a corresponding point in $A_{i,n}^{T^*}$ which converges to $\left(\boldsymbol{x}_{\setminus i}, \underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}) \right)$. From the assumptions of Theorem 3.3, we have that each $\boldsymbol{x}_{\setminus i}$ corresponds to a unique value of $\underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X})$ and $\underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}^n(\mathbf{X})$. Thus, we know this point will be $\left(\boldsymbol{x}_{\setminus i}, \underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}^n(\mathbf{X}) \right) \in A_{i,n}^{T^*}$. Finally, we have that $\left(\boldsymbol{x}_{\setminus i}, \underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}) \right) \xrightarrow[n\to\infty]{wp1} \left(\boldsymbol{x}_{\setminus i}, \underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}^n(\mathbf{X}) \right)$ and therefore

$$\underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}^n(\mathbf{X}) \xrightarrow[n\to\infty]{wp1} \underline{\text{VaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}),$$

pointwise, as required. $\qquad\square$

Now, we prove the consistency of $\underline{\text{TVaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}^n(\mathbf{X})$. This result is important because it allows risk managers in industry to assess the risks associated to their assets and liabilities without the risk of mispecification of the underlying model and with a negligible loss in accuracy.

Theorem 3.7. *Under the assumptions of Theorem 3.3 and a consistent estimator $F_{n,\setminus i}$ for $F_{\setminus i}$, we have that $\underline{\text{TVaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}^n(\mathbf{X})$ is strongly consistent for $\underline{\text{TVaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X})$, that is*

$$\underline{\text{TVaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}^n(\mathbf{X}) \xrightarrow[n\to\infty]{wp1} \underline{\text{TVaR}}_{\alpha,\boldsymbol{x}_{\setminus i}}(\mathbf{X}).$$

Proof. We will use the dominated convergence theorem. In particular we aim to show

$$\lim_{n\to\infty} \underline{\text{TVaR}}^n_{\alpha,\boldsymbol{x}_{\backslash i}}(\mathbf{X}) = \lim_{n\to\infty} \int h_{n,\alpha,\boldsymbol{x}_{\backslash i}}(u)\mathrm{d}u = \underline{\text{TVaR}}_{\alpha,\boldsymbol{x}_{\backslash i}}(\mathbf{X})$$

where

$$h_{n,\alpha,\boldsymbol{x}_{\backslash i}}(u) = \frac{\underline{\text{VaR}}^n_{u,\boldsymbol{x}_{\backslash i}}(\mathbf{X})\mathbb{1}_{[\alpha,F_{n,\backslash i}(\boldsymbol{x}_{\backslash i})]}(u)}{F_{n,\backslash i}(\boldsymbol{x}_{\backslash i}) - \alpha}.$$

By Corollary 3.4 we have that

$$\underline{\text{VaR}}^n_{u,\boldsymbol{x}_{\backslash i}}(\mathbf{X}) \xrightarrow[n\to\infty]{wp1} \underline{\text{VaR}}_{u,\boldsymbol{x}_{\backslash i}}(\mathbf{X}),$$

pointwise, $\forall\, u \in (0,1)$. Next, by hypothesis, one has that

$$F_{n,\backslash i} \xrightarrow[n\to\infty]{wp1} F_{\backslash i},$$

and so

$$\mathbb{1}_{[F_{n,\backslash i}(\boldsymbol{x}_{\backslash i}),\alpha]}(u) = \begin{cases} 1, & u \in [\alpha, F_{n,\backslash i}(\boldsymbol{x}_{\backslash i})] \\ 0, & \text{otherwise} \end{cases}$$

$$\xrightarrow[n\to\infty]{wp1} \begin{cases} 1, & u \in [\alpha, F_{\backslash i}(\boldsymbol{x}_{\backslash i})] \\ 0, & \text{otherwise} \end{cases}$$

$$= \mathbb{1}_{[F_{\backslash i}(\boldsymbol{x}_{\backslash i}),\alpha]}(u).$$

Therefore,

$$h_{n,\alpha,\boldsymbol{x}_{\backslash i}}(u) \xrightarrow[n\to\infty]{wp1} \frac{\underline{\text{VaR}}_{u,\boldsymbol{x}_{\backslash i}}(\mathbf{X})\mathbb{1}_{[\alpha,F_{\backslash i}(\boldsymbol{x}_{\backslash i})]}(u)}{F_{\backslash i}(\boldsymbol{x}_{\backslash i}) - \alpha},$$

pointwise. Next, consider the function

$$g_{\alpha,\boldsymbol{x}_{\backslash i}}(u) = \frac{\underline{\text{VaR}}_{u,\boldsymbol{x}_{\backslash i}}(\mathbf{X})\mathbb{1}_{[\alpha,F_{\backslash i}(\boldsymbol{x}_{\backslash i})+\epsilon]}(u)}{F_{\backslash i}(\boldsymbol{x}_{\backslash i}) - \alpha} + b\mathbb{1}_{[\alpha,F_{\backslash i}(\boldsymbol{x}_{\backslash i})+\epsilon]}(u),$$

for some $b,\epsilon > 0$ and $u \in (0,1)$. It can be seen that $h_{n,\alpha,\boldsymbol{x}_{\backslash i}}(u) \leq g_{\alpha,\boldsymbol{x}_{\backslash i}}(u)$, for all $u \in (0,1)$ and for n large enough, and that $g_{\alpha,\boldsymbol{x}_{\backslash i}}$ is integrable. Finally, by the dominated convergence theorem

$$\begin{aligned} \lim_{n\to\infty} \underline{\text{TVaR}}^n_{\alpha,\boldsymbol{x}_{\backslash i}}(\mathbf{X}) &= \lim_{n\to\infty} \int \frac{\underline{\text{VaR}}^n_{u,\boldsymbol{x}_{\backslash i}}(\mathbf{X})\mathbb{1}_{[\alpha,F_{n,\backslash i}(\boldsymbol{x}_{\backslash i})]}(u)}{F_{n,\backslash i}(\boldsymbol{x}_{\backslash i}) - \alpha}\mathrm{d}u \\ &= \int \frac{\underline{\text{VaR}}_{u,\boldsymbol{x}_{\backslash i}}(\mathbf{X})\mathbb{1}_{[\alpha,F_{\backslash i}(\boldsymbol{x}_{\backslash i})]}(u)}{F_{\backslash i}(\boldsymbol{x}_{\backslash i}) - \alpha}\mathrm{d}u \\ &= \frac{1}{F_{\backslash i}(\boldsymbol{x}_{\backslash i}) - \alpha} \int_\alpha^{F_{\backslash i}(\boldsymbol{x}_{\backslash i})} \underline{\text{VaR}}_{u,\boldsymbol{x}_{\backslash i}}(\mathbf{X})\mathrm{d}u \\ &= \underline{\text{TVaR}}_{\alpha,\boldsymbol{x}_{\backslash i}}(\mathbf{X}), \end{aligned}$$
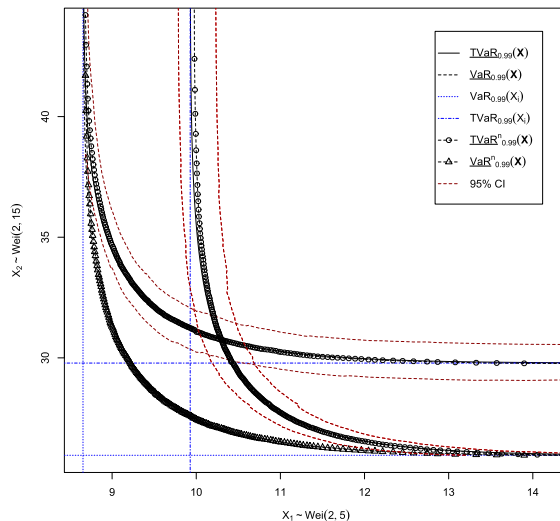
FIGURE 3. Simulation study for $\underline{\mathrm{TVaR}}_{0.95}(\mathbf{X})$
.

establishing the consistency of the empirical TVaR. □

4. EXAMPLES

In this section, examples demonstrating the estimators' performances are presented. The first example presents confidence intervals for the lower orthant estimator, as well as a comparison between the empirical copula and a smoothed version. The second example presents a simulation study of the estimator for the upper orthant TVaR. Finally, we provide comparisons to the vector-valued CTE developed in Di Bernardino *et al.* [9] for both simulated data and the lossALAE data set.

It should be noted that the calculations presented below were completed using R statistical software. While this choice makes sense from an ease of implementation standpoint, as many of the functions required are already built in, the software is notorious for being drastically slower than other languages.

4.1. Simulation-lower orthant

We begin with a study of $\underline{\mathrm{TVaR}}_\alpha^n(\mathbf{X})$. A simulation study was run to estimate $\underline{\mathrm{TVaR}}_{0.95}(\mathbf{X})$. In particular, 50 simulations were run each with a sample size of $n = 4000$ and $m = 250$ (this choice of m will be kept constant throughout the remainder of the paper). Marginally, the random variables were Weibull with $X_1 \sim \mathrm{Weibull}(2, 5)$ and $X_2 \sim \mathrm{Weibull}(2, 15)$. The dependence was described by a Frank copula with a Kendall's tau of $\tau = 0.5$. The results of the simulations are presented in Figure 3. As can be seen, the differences between $\underline{\mathrm{VaR}}_{0.95}^n(\mathbf{X})$, $\underline{\mathrm{TVaR}}_{0.95}^n(\mathbf{X})$ and their theoretical counterparts are mostly negligible.

As previously mentioned, the choice of empirical distribution function is crucial when discussing the consistency of the proposed estimator $\underline{\mathrm{TVaR}}_\alpha^n(\mathbf{X})$. To this end, we ran the simulations presented in Figure 4 to compare two choices of ecdf. The first choice was the standard empirical copula, C_n, which does not satisfy the conditions of Theorem 3.3. The second choice was the multilinear extension, C_n^*, as described in Nešlehová [17], which does satisfy the conditions of Theorem 3.3. Fifty simulations were run at the $\alpha = 0.99$ level with sample sizes $n = 1000$ and $n = 4000$, and the distributional settings were identical to those used for Figure 3, described above. Additionally, note that only $\underline{\mathrm{TVaR}}_{0.99, x_1}(\mathbf{X})$ was calculated for this comparison.
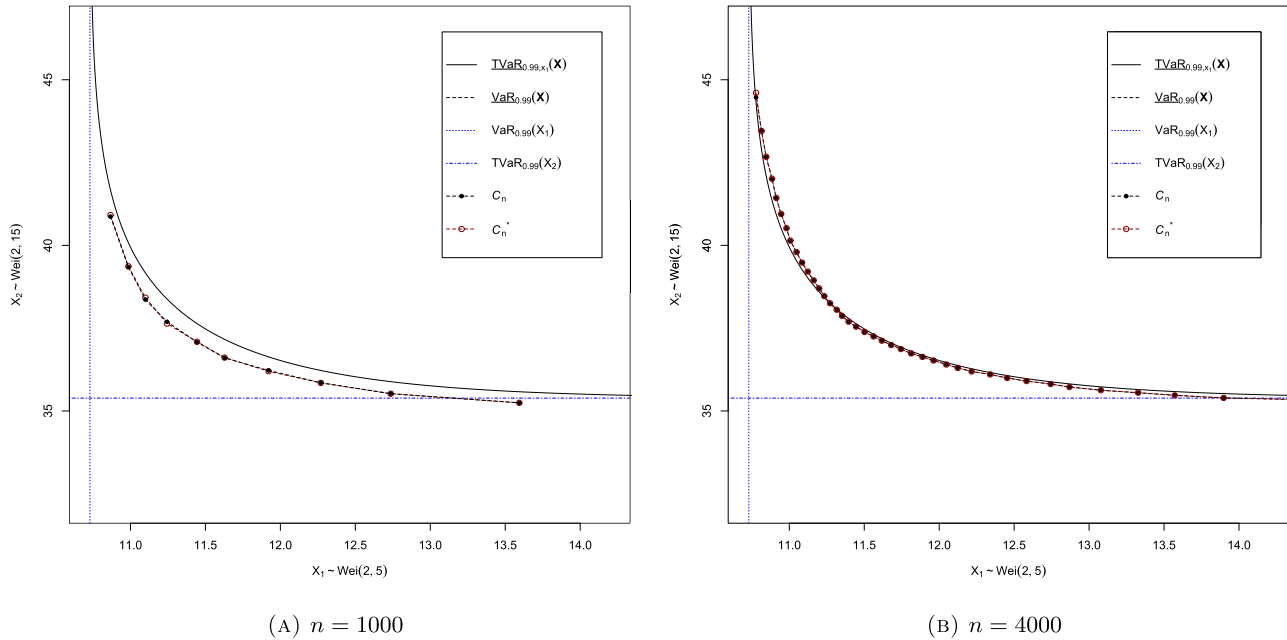
(A) $n = 1000$ (B) $n = 4000$

FIGURE 4. Comparison of $\underline{\mathrm{TVaR}}^n_{0.99,x_1}(\mathbf{X})$ using C_n *vs.* C_n^*.

TABLE 1. Comparison of SSE for multilinear and empirical copula.

$n = 1000$		$n = 4000$	
C_n	C_n^*	C_n	C_n^*
3.483	3.372	2.088	2.181

Both sets of simulations show that the fits of the empirical curves are very good, demonstrating the consistency of the proposed estimators. Moreover, it can be seen the smoothed multilinear empirical copula provides no noticeable improvement over the standard empirical copula. The sum of squared errors, $\mathrm{SSE} = \sum_{i=1}^n \left(\underline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X}) - \underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X}) \right)^2$, for each simulation in Figure 4 and presented in Table 1, demonstrate that the differences are again minimal. While the multilinear extension does satisfy the assumptions set forth in Theorem 3.3, the computational complexity is much higher. Therefore, the remaining analyses were completed using C_n.

4.2. Simulation-upper orthant

Our next simulation study was performed on the upper orthant, with results presented in Figure 5. Here, sample sizes of $n = 100$ and 250 were used. The reason behind the large discrepancy of sample sizes when compared to the study of the lower orthant boils down to the amount of data that is thrown away. When evaluating $\underline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X})$, only $100(1 - \alpha)\%$ of the data is used. In the case of Figure 4, this means 99% of the data is discarded. However, when computing $\overline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X})$, $100\alpha\%$ of the data is considered in the calculations (99% of the data in this simulation).

In Figure 5, one may see that there is a noticeable difference between $\overline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})$ and $\overline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X})$, for both sample sizes. The inversion of the empirical survival copula is more demanding than inverting C_n, requiring additional estimation, which could explain the less accurate results. Use of the multilinear extension
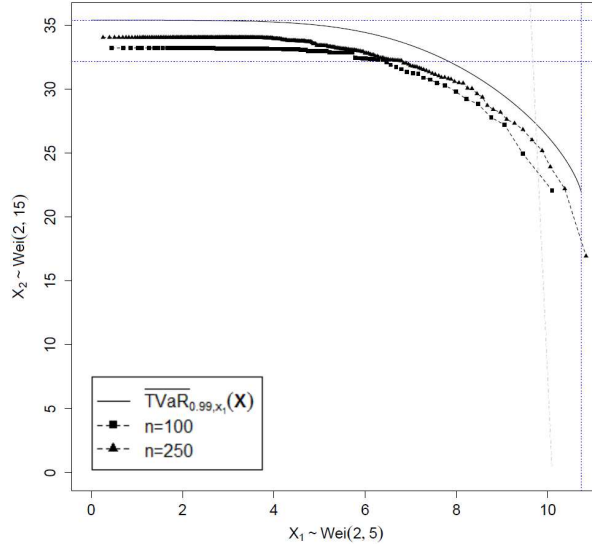
FIGURE 5. Simulation study for Weibull margins linked by Frank Copula with Kendall's tau, $\tau = 0.5$. Fifty simulations of $n = 100$ and 250 generated data pairs were run.

in the upper orthant could potentially improve fit. However, these calculations would take several weeks with a high performance processor and as no improvement was seen in the lower orthant, we do not consider this case.

4.3. Capital allocation comparison

As both $\underline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ are infinite collections of points, from a decision making standpoint, a strategy for establishing the appropriate allocation is necessary. Here, we refer to Mailhot and Mesfioui [15] for their discussion on capital allocation based on multivariate lower and upper orthant TVaR. In particular, we focus on VaR and TVaR-based orthogonal projections in the bivariate case. First, the VaR-based projection is the allocation pair $\left(x_1^V, \underline{\mathrm{TVaR}}_{\alpha, x_1^V}(\mathbf{X})\right)$ where

$$x_1^V = \min_{x_1 \geq \mathrm{VaR}_\alpha(X_1)} \left\{ \left(\underline{\mathrm{VaR}}_{\alpha, x_1}(\mathbf{X}) - \mathrm{VaR}_\alpha(X_2)\right)^2 + (x_1 - \mathrm{VaR}_\alpha(X_1))^2 \right\}. \tag{4.1}$$

Next, the TVaR-based projection gives the pair $\left(x_1^T, \underline{\mathrm{TVaR}}_{\alpha, x_1^T}(\mathbf{X})\right)$ such that

$$x_1^T = \min_{x_1 \geq \mathrm{VaR}_\alpha(X_1)} \left\{ \left(\underline{\mathrm{TVaR}}_{\alpha, x_1}(\mathbf{X}) - \mathrm{TVaR}_\alpha(X_2)\right)^2 + (x_1 - \mathrm{VaR}_\alpha(X_1))^2 \right\}. \tag{4.2}$$

Note that analogous allocations can be done in terms of x_2. Here, we compare our empirical estimator for $\underline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ to the consistent estimator of the multivariate lower orthant CTE, $\mathrm{CTE}_\alpha^n(\mathbf{X})$, as outlined in Di Bernardino *et al.* [9]. As both measures are based on the same multivariate VaR and consistency argument, they are a natural choice for comparison. However, one key difference in these measures is that, as mentioned, $\underline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ is an infinite collection of points whereas $\mathrm{CTE}_\alpha(\mathbf{X})$ is the same dimension as the random variable of interest. To this end, we made use of the allocation methods outlined in (4.1) and (4.2) to compare. We used the same sample as was used in Section 4.1 and the results are presented in Table 2. As can be seen, the allocation pairs coming from $\underline{\mathrm{TVaR}}_\alpha^n(\mathbf{X})$ are very close to the estimates of $\mathrm{CTE}_\alpha(\mathbf{X})$. While the allocation methods used produce very similar results, $\underline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ provides an infinite set of possibilities from which to draw

TABLE 2. Comparison of empirical measures on simulated data.

α	$\left(x_1^V, \underline{\text{TVaR}}_{\alpha,x_1^V}^n(\mathbf{X})\right)$	$\left(x_1^T, \underline{\text{TVaR}}_{\alpha,x_1^T}^n(\mathbf{X})\right)$	$\left(\underline{\text{TVaR}}_{\alpha,x_2^V}^n(\mathbf{X}), x_2^V\right)$	$\left(\underline{\text{TVaR}}_{\alpha,x_2^T}^n(\mathbf{X}), x_2^T\right)$	$\text{CTE}_\alpha^n(\mathbf{X})$
0.950	$(10.178, 30.827)$	$(10.175, 27.187)$	$(11.352, 27.071)$	$(11.343, 27.130)$	$(10.511, 31.275)$
0.990	$(12.028, 36.382)$	$(12.124, 36.302)$	$(13.047, 32.982)$	$(13.038, 33.093)$	$(12.299, 36.461)$
0.995	$(12.720, 38.500)$	$(12.869, 38.342)$	$(13.610, 35.134)$	$(13.678, 35.155)$	$(13.207, 38.294)$

TABLE 3. Comparison of empirical measures on LossALAE data set.

α	$\left(x_1^V, \underline{\text{TVaR}}_{\alpha,x_1^V}^n(\mathbf{X})\right)$	$\left(x_1^T, \underline{\text{TVaR}}_{\alpha,x_1^T}^n(\mathbf{X})\right)$	$\left(\underline{\text{TVaR}}_{\alpha,x_2^V}^n(\mathbf{X}), x_2^V\right)$	$\left(\underline{\text{TVaR}}_{\alpha,x_2^T}^n(\mathbf{X}), x_2^T\right)$	$\text{CTE}_\alpha^n(\mathbf{X})$
0.950	$(210\,000, 153\,281)$	$(210\,000, 144\,899)$	$(373\,158, 81\,128)$	$(384\,772.7, 72\,060)$	$(533\,281.7, 132\,637.7)$
0.990	$(500\,000, 274\,223)$	$(500\,000, 274\,223)$	$(1\,104\,683, 134\,743)$	$(1\,104\,683, 160\,265)$	$(1\,043\,399, 254\,461)$
0.995	$(750\,000, 448\,858)$	$(750\,000, 448\,858)$	$(1\,138\,139, 306\,072)$	$(1\,138\,139, 306\,072)$	NA

the allocation pair. This will allow insurers and regulators alike to adjust their allocation needs accordingly, based on internal objectives or by factoring other outside information, providing a distinct advantage to using $\underline{\text{TVaR}}_\alpha^n(\mathbf{X})$ over $\text{CTE}_\alpha^n(\mathbf{X})$. Next, the measures were compared using the lossALAE data set in Table 3. Here, a much larger discrepancy between the TVaR and CTE is observed at each significance level. Additionally, the choice of allocation method and whether one fixes x_1 or x_2 is very significant. The empirical CTE is often noticeably larger than the TVaR across the board. However, it is noted that at the 99.5% level, $\text{CTE}_{0.995}^n(\mathbf{X})$ does not exist, as the 99.5% empirical level set, $L_{0.995}^n$, is empty. Such problems do not occur with $\underline{\text{TVaR}}_{0.995}^n(\mathbf{X})$.

5. Conclusion

In this paper, the estimation of dependent tail risks in dimension $d \geq 2$ is addressed. In particular, a strongly consistent estimator to the orthant-based TVaR for d-dimensional dependent random vectors is developed. Through simulation studies, it is shown that the methods presented here allow for accurate evaluation of tail risks without having to assume a statistical model, removing the risk of misspecification. Additionally, this method does not require the aggregation of risks. While evaluating the risks of random sums can still preserve the underlying dependence of the individual random variables, it can result in the underlying risks compensating for each other, which may not be apt in today's more conservative economic landscape. The ability to estimate individual values for these risks, while maintaining the dependence between them can allow for more flexibility. In particular, this added flexibility can prove beneficial for operational risk management, where individual financial institutions can have vastly different makeups and goals, such as reducing operational loss or exposure to future risks. While at higher dimensions the curse of dimensionality is certainly present, methods of projection-based capital allocation, particularly those involving VaR-based projections, can be used to curb these computational issues and provide comparable results to measures seen in the literature. Incorporating the techniques of Bardou *et al.* [2] might also prove beneficial as they could speed up computation time and lay the ground work for a CLT type result.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, coherent measures of risk. *Math. Finance* **9** (1999) 203–228.

[2] O. Bardou, N. Frikha and G. Pagès, Computing VaR and CVaR using stochastic approximation and adaptive unconstrained importance sampling. *Monte Carlo Method. Appl.* **15** (2009) 173–210.

[3] B. Brahimi, B. Fatah and Y. Djabrane, Copula conditional tail expectation for multivariate financial risks. Arab J Math. Seci. **24** (2018) 82–100.

[4] H. Cossette, M. Mailhot, É. Marceau and M. Mesfioui, Bivariate lower and upper orthant value-at-risk. *Eur. Actuar. J.* **3** (2013) 321–357.

[5] H. Cossette, M. Mailhot, É. Marceau and M. Mesfioui, Vector-valued tail value-at-risk and capital allocation. *Methodol. Comput. Appl. Probab.* **18** (2016) 653–674.

[6] A. Cousin and E. Di Bernardino, On multivariate extensions of value-at-risk. *J. Multivar. Anal.* **119** (2013) 32–46.

[7] A. Cuevas, W. González-Manteiga and A. Rodríguez-Casal, Plug-in estimation of general level sets. *Aust. N. Z. J. Stat.* **48** (2006) 7–19.

[8] E. Di Bernadino and T. Laloë, Estimating level sets of a distribution function using a plug-in method: a multidimensional extension. Preprint arXiv:1202.2035 (2012).

[9] E. Di Bernardino, T. Laloë, V. Maume-Deschamps and C. Prieur, Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory. *ESAIM: Probab. Stat.* **17** (2013) 236–256.

[10] P. Embrechts and G. Puccetti, Bounds for functions of multivariate risks. *J. Multivar. Anal.* **97** (2006) 526–547.

[11] European Commission, Solvency II (2014), http://ec.europa.eu/finance/insurance/solvency/index_en.htm.

[12] E.W. Frees and E.A. Valdez, Understanding relationships using copulas. *North Am. Actuar. J.* **2** (1998) 1–25.

[13] E. Jouini, M. Meddeb and N. Touzi, Vector-valued coherent risk measures. *Finance Stoch.* **8** (2004) 531–552.

[14] E. Lépinette and I.B. Tahar, Vector-valued coherent risk measure processes. *Int. J. Theor. Appl. Finance* **17** (2014) 1450011.

[15] M. Mailhot and M. Mesfioui, Multivariate TVaR-based risk decomposition for vector-valued portfolios. *Risks* **4** (2016) 33.

[16] G. Mainik and E. Schaanning, On dependence consistency of covar and some other systemic risk measures. Stat. Risk Model. **31** (2014) 49–77.

[17] J. Nešlehová, On rank correlation measures for non-continuous random variables. *J. Multivar. Anal.* **98** (2007) 544–567.

[18] OSFI, Minimum capital test for federally regulated property and casualty insurance companies. (2015), http://www.osfi-bsif.gc.ca/Eng/fi-if/rg-ro/gdn-ort/gl-ld/Pages/mct2015.aspx.

[19] R.T. Rockafellar and S. Uryasev, Optimization of conditional value-at-risk (2000).

[20] R.T. Rockafellar and S. Uryasev, Conditional value-at-risk for general loss distributions. *J. Bank. Finance* **26** (2002) 1443–1471.