



**HAL**  
open science

# The Hoeffding-Sobol decomposition in extreme value theory. Exploring the asymptotic dependence structure.

Cécile Mercadier, Olivier Roustant

## ► To cite this version:

Cécile Mercadier, Olivier Roustant. The Hoeffding-Sobol decomposition in extreme value theory. Exploring the asymptotic dependence structure.. 2017. hal-01649596v1

**HAL Id: hal-01649596**

**<https://hal.science/hal-01649596v1>**

Preprint submitted on 27 Nov 2017 (v1), last revised 7 Mar 2019 (v3)

**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.

# The Hoeffding-Sobol decomposition in extreme value theory. Exploring the asymptotic dependence structure.

Cécile MERCADIER  
mercadier@math.univ-lyon1.fr

Olivier ROUSTANT  
roustant@emse.fr

Univ Lyon,  
Université Claude Bernard Lyon 1,  
CNRS UMR 5208,  
Institut Camille Jordan,  
43 blvd. du 11 novembre 1918,  
F-69622 Villeurbanne cedex  
France

Mines Saint-Étienne,  
UMR CNRS 6158,  
LIMOS,  
F-42023 Saint-Étienne,  
France

November 27, 2017

## Abstract

All characterizations of non degenerate multivariate tail dependence structure are both functional and non finite dimensional, as the stable tail dependence function  $\ell$ . Taking advantage of the Hoeffding-Sobol decomposition of  $\ell$ , we derive new measures to summarize the strength of dependence in a multivariate extreme value analysis. The tail superset importance coefficients provide a pairwise ordering of the asymptotic dependence structure. We then define the tail dependograph in which vertices represent components of the vector of interest and where edge weights are proportional to the tail superset importance coefficients. For inference, as soon as an estimator of  $\ell$  is given, it yields an estimation of the tail dependograph. In particular, the empirical tail superset importance coefficient is rank-based statistics and its asymptotic behavior is stated. These new concepts are illustrated with several examples including theoretical models through simulation and real data, which shows that our methodology works well in practice.

**Keywords:** Global sensitivity analysis, Hoeffding-Sobol decomposition, Multivariate extreme value analysis, Pairwise indices, Tail dependency graph.

## 1 Introduction

In extreme value theory, a central question that emerges when studying several variables simultaneously is that of the tail dependence modeling. For instance, abnormal rise of water generated by a storm are a consequence of the behavior of other variables as low pressure, strong wind, large storm, angle to coast and are stressed also by some covariate values as concavity of coastline and gently sloped shelves. As pointed out in [Cai et al. \(2013\)](#), asymptotic independence with positive association might be relevant for the link between surge and wind force, at least for this particular location on the French coastal zone. More recently, [Kereszturi et al. \(2016\)](#) performed a wide study taking into account North Sea sea-state and wave height at several locations with the covariates wave direction and distance between locations.

They generally observe asymptotic independence but conclude asymptotic dependence also in few cases. Temperatures at several locations would be an environmental example of strong asymptotic dependence. See the recent study of [Mhalla et al. \(2017\)](#) including time and altitude covariates for US dataset. In a financial context, one can think about contagion between stock indices across markets as investigated in [Peng and Ng \(2012\)](#). All the previous examples reflect the possible variation for the strength of tail dependence between variables, taking into account the nature of the real dataset or depending the value of some covariates.

Under multivariate max-stable modeling, there exists various ways to describe the extremal dependence structure: extreme value copula, dependence functions, intensity measure, spectral measure. . . that, of course, have strong connections. We refer to [Fougères \(2004\)](#) and [Segers \(2012\)](#) for a comprehensive description of these notions. Here the focus is done on the stable tail dependence function. See [Ressel \(2013\)](#) for its complete characterization. It has some nice properties: homogeneity of order 1, convexity, identity on the axis, and tractable upper and lower bounds for independence and complete positive dependence respectively. Our choice is also motivated by the fact that tail independence induces additivity properties to the stable tail dependence function. However, this is a functional characterization of the tail dependence structure with no finite-dimensional parametrization. And, as always, it is interesting to summarize all the relevant information throughout well chosen indices. Since the original idea of Tiago de Oliveira, and the name proposed by [Smith \(1990\)](#), many authors have studied or just used the extremal coefficients. For the dimension  $d$  of a vector under study, one has  $2^d - d - 1$  coefficients that, for a given subset of  $\{1, \dots, d\}$ , might be interpreted as the number of asymptotically independent components among this subset.

The main goal of this paper is to introduce new indices, called *tail superset importance coefficients*, for such a multivariate extreme value analysis. Let  $\ell$  stand for the stable tail dependence function. Roughly speaking, our indices are derived from the decomposition of the variance of  $\ell(\mathbf{U})$  at some generic random vector  $\mathbf{U}$ , whereas the already known coefficients focus on the evaluation of  $\ell$  at points on diagonals. Moreover, they are now only pairwise so reduce to a collection of  $d(d-1)/2$  indices. Their definition starts from the Hoeffding-Sobol decomposition (of the function  $\ell$ ) that is a very well-known tool in computer experiments. See [Hoeffding \(1948\)](#) for the pionner construction and [Sobol' \(1993\)](#) for its practical efficient use. It detects indeed some additive structure in a response but also ranks the corresponding entries thank's to a variance decomposition. In our setting, we only have natural entries represented by the  $d$  variables jointly studied but there is no response. The idea is to define a pseudo-response as  $\ell(\mathbf{U})$  with the wish to rank dependence in the asymptotic joint structure. The new measures allow to decompose the tail dependence structure but above all rank the pairs of components in terms of their global contribution to tail dependence.

Furthermore, superset importance coefficients of [Hooker \(2004\)](#) and [Liu and Owen \(2006\)](#) are not the only tool from sensitivity analysis we try to transpose to an extreme value setting. We also want to highlight the features of the FANOVA graph introduced by [Muehlenstaedt et al. \(2012\)](#). Applied to the stable tail dependence function, we call it the *tail dependograph* since it graphically represents, on an undirected but weighted graph, the strength of asymptotic dependence within a vector. In this graph, the edge thickness between two vertices is proportional to the force of tail dependence in the corresponding bivariate model but also cumulates the information of the asymptotic dependence structure for any multivariate model that contains this pair.

If instead of applying the Hoeffding-Sobol decomposition on the true (and most of the time unknown)

stable tail dependence function, its estimation is used, one gets a decomposition of the estimated variance of  $\ell(\mathbf{U})$ . As a consequence, one obtains an estimation of the tail superset importance coefficients and, obviously, an estimate of the tail dependograph. The empirical estimator of the stable tail dependence function, first proposed by [Huang \(1992\)](#), leads to explicit rank-based estimators of the tail superset importance coefficients. Under a second order condition, similar to the setting assumed in [Fougères et al. \(2015\)](#), and taking advantage of the asymptotic distributional expansion of the empirical estimator of  $\ell$ , we state the asymptotic behaviour of the empirical tail superset importance coefficients. Normalized versions of the new indices are also introduced and consistent rank-based estimators are provided.

The remainder of the article is organized as follows: In [Section 2](#), we remind some tools from global sensitivity analysis. Then, in [Section 3](#), we derive new coefficients and dependency graphs for extreme value analysis. In [Section 4](#), the inference is done through a non parametric approach. The asymptotic behaviour of these empirical indices is also explored at the end of the section. We provide in [Section 5](#) a wide numerical illustration of the new concepts. Some concluding remarks end the paper before collecting proofs in a dedicated section.

## 2 Background on global sensitivity analysis

Global Sensitivity Analysis is a branch of control theory and computer sciences. Such analysis measures how sensitive the outcome quantity  $f(\mathbf{u})$  is to variation of individual input variables  $u_i$ . A variance decomposition determines which of the multiple input parameters drive most of the variation in the outcome. The current section deals with reminders from functional decomposition to computational issues to consequences in terms of variance.

### 2.1 The Hoeffding-Sobol Decomposition

Let  $d$  be an integer greater or equal to 2. Let  $f$  be a square integrable real valued function defined on  $\mathbb{R}^d$ . The Hoeffding-Sobol decomposition aims at representing  $f$  as

$$f(u_1, \dots, u_d) = f_\emptyset + \sum_{i=1}^d f_i(u_i) + \sum_{i < j} f_{ij}(u_i, u_j) + \dots + f_{1\dots d}(u_1, \dots, u_d) \quad (1)$$

where, for a subset  $I$  of  $\{1, \dots, d\}$ , the subfunction  $f_I$  is  $|I|$ -dimensional. The construction of the decomposition can be explained both stochastically and recursively. Let  $\mathbf{U} = (U_1, \dots, U_d)$  be a random vector with independent margins. Its distribution is denoted by  $\mu = \mu_1 \otimes \dots \otimes \mu_d$ . We thus assumed above that  $f \in \mathbb{L}^2(\mathbb{R}^d, \mu)$ . The random vector  $\mathbf{U}$  allows to make a presentation using mathematical expectation in place of classical integrals. The recursive formula is the following

$$f_I(\mathbf{u}_I) = \mathbb{E}[f(\mathbf{U}) | \mathbf{U}_I = \mathbf{u}_I] - \sum_{I' \subsetneq I} f_{I'}(\mathbf{u}_{I'})$$

where  $\mathbf{u}_I$  concatenates the components of  $\mathbf{u}$  which indices are included in  $I$ . For a non-empty subset of indices  $I \subset \{1, \dots, d\}$ , one has in addition  $\mathbb{E}[f_I(\mathbf{U}_I) | \mathbf{U}_{I'}] = 0$  when  $I' = \emptyset$  or  $I' \subsetneq I$ . This property implies that the orthogonal projection (in  $\mathbb{L}^2$  sense) of the terms onto strict subspaces are zero, which guarantees the unicity of the decomposition.

The terms in the decomposition can be interpreted as follows. The constant  $f_\emptyset = \mathbb{E}[f(\mathbf{U})]$  is the global (un-

conditional) mean,  $f_i(u_i) = \mathbb{E}[f(\mathbf{U})|U_i = u_i] - f_\emptyset$  represents the main effect of component  $\{i\}$ ,  $f_{ij}(u_i, u_j) = \mathbb{E}[f(\mathbf{U})|U_i = u_i, U_j = u_j] - f_i(u_i) - f_j(u_j) - f_\emptyset$  captures the second-order interaction from the pair of components  $\{i, j\}$ , and so on. In the previous formulae, the mathematical expectation refers to the integral with respect to  $d\mu$  in  $f_\emptyset$ , to  $\otimes_{t \neq i} d\mu_t$  in the definition of  $f_i$ , to  $\otimes_{t \neq i, j} d\mu_t$  in that of  $f_{ij}$ .

The name of this decomposition arises first from its creation by [Hoeffding \(1948\)](#) and second from its use for sensitivity analysis by [Sobol' \(1993\)](#). The reader could also be interested by [Efron and Stein \(1981\)](#) and [van der Vaart \(1998; Section 11.4\)](#). More recently, an extension to dependent entries has been obtained by [Chastaing et al. \(2012\)](#) under specific conditions. Before going further, it is important to avoid any possible confusion. Here, we only refer to the pionner setting where the distribution of  $\mathbf{U}$  is a product measure. On the contrary, all the dependence we talk about in this paper will be stored in the function  $\ell$  on which we will focus from [Section 3.1](#).

## 2.2 A variance decomposition

The Hoeffding-Sobol decomposition has also the advantage of a variance decomposition

$$\text{var}(f(\mathbf{U})) = \sum_{I \subseteq \{1, \dots, d\}} \text{var}(f_I(\mathbf{U}_I)),$$

mainly justified by the condition  $\mathbb{E}[f_I(\mathbf{U}_I)|\mathbf{U}_{I'}] = 0$  for any subset  $I'$  strictly included in  $I$ . As already pointed out, it implies that all the terms of the decomposition are centered and orthogonal. One generally refers to the variance decomposition with the following notation

$$D(f) = \sum_{I \subseteq \{1, \dots, d\}} D_I(f) \tag{2}$$

where  $D(f)$  denotes the global variance  $\text{var}(f(\mathbf{U}))$  and  $D_I(f)$  the variance corresponding to the subset  $I$ , that is  $\text{var}(f_I(\mathbf{U}_I))$  the variance of the term  $f_I$ . For instance,  $D_{\{i\}}(f)$  is the variance of  $f(\mathbf{U})$  due to the  $i$ th component only. Similarly,  $D_{\{i, j\}}(f)$  is that due to the combined effect of components  $\{i, j\}$ .

One combination of such variances is of prime interest. It is defined by

$$D_{\{i, j\}}^{\text{super}}(f) = \sum_{I \supseteq \{i, j\}} D_I(f)$$

and called the superset importance of the pair  $\{i, j\}$ . See [Hooker \(2004\)](#) and [Liu and Owen \(2006\)](#). This coefficient is positive and smaller or equal to  $D(f)$ . Such index allows discovering additive structures in multivariate functions. Indeed, under continuity assumptions on  $\mu$  and  $f$ ,  $D_{\{i, j\}}^{\text{super}}(f) = 0$  implies that  $f$  does not depend simultaneously on  $i$  and  $j$  in its ANOVA decomposition, and thus  $f$  is additive with respect to  $u_i$  and  $u_j$ .

Taking into account the decomposition of the variance  $D(f)$ , scaled versions have been introduced by [Sobol' \(1993\)](#). The superset sensitivity indices, that belong to the interval  $[0, 1]$ , are given by

$$S_{\{i, j\}}^{\text{super}}(f) = \frac{D_{\{i, j\}}^{\text{super}}(f)}{D(f)}.$$

### 2.3 On computational aspects

When the function of interest  $f$  is known, the Hoeffding-Sobol decomposition (1) is not necessarily derived obviously. Let us introduce some notation. Let  $(\mathbf{e}_1, \dots, \mathbf{e}_d)$  stand for the canonical basis of  $\mathbb{R}^d$ . Then  $\mathbf{u} + (v - u_i)\mathbf{e}_i$  and  $\mathbf{u} + (v - u_i)\mathbf{e}_i + (w - u_j)\mathbf{e}_j$  are the vectors both obtained from  $\mathbf{u}$ , with  $u_i$  replaced by  $v$ , and additionally in the second case with  $u_j$  replaced by  $w$ . We denote the corresponding evaluations of  $f$  at these points by  $T_i[f](\mathbf{u}, v) := f(\mathbf{u} + (v - u_i)\mathbf{e}_i)$ , and  $T_{i,j}[f](\mathbf{u}, v, w) := f(\mathbf{u} + (v - u_i)\mathbf{e}_i + (w - u_j)\mathbf{e}_j)$ . Then, from Liu and Owen (2006), we have

$$D_{\{i,j\}}^{\text{super}}(f) = \frac{1}{4} \int_{[0,1]^{d+2}} \{f(\mathbf{u}) - T_i[f](\mathbf{u}, v) - T_j[f](\mathbf{u}, w) + T_{i,j}[f](\mathbf{u}, v, w)\}^2 d\mu(\mathbf{u})d\mu_i(v)d\mu_j(w). \quad (3)$$

As noticed in the latter reference, Formula (3) can be viewed as an extension to the continuous case of the interaction sum of squares in a discrete ANOVA (Cornfield and Tukey, 1956; p. 916)

$$\text{SS}_{\text{inter}} = n \sum_i \sum_j (\bar{y}_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y} \dots)^2.$$

Equality (3) is powerful. On the one hand, it provides an information that is not just limited to  $\{i, j\}$  but that is linked to all subsets  $I$  from  $\{1, \dots, d\}$  that contain this pair. On the other hand, it does not require the knowledge of all subfunctions  $f_I$  induced in the initial definition of  $D_{\{i,j\}}^{\text{super}}(f)$ . In other words, this integral representation allows to compute

$$\sum_{I \supseteq \{i,j\}} \text{var}(f_I(\mathbf{U}_I))$$

without identifying the terms of the decomposition  $\{f_I, I \supseteq \{i, j\}\}$ .

Again, even if the superset importance coefficient associated to any pair has an integral formula, it is not easy to compute explicitly. Hopefully, Formula (3) admits Monte Carlo approximations as estimates. It can be written as

$$\widehat{D}_{\{i,j\}}^{\text{super}}(f) = \frac{1}{4N} \sum_{s=1}^N \left( f(\mathbf{u}^{[s]}) - T_i[f](\mathbf{u}^{[s]}, v^{[s]}) - T_j[f](\mathbf{u}^{[s]}, w^{[s]}) + T_{i,j}[f](\mathbf{u}^{[s]}, v^{[s]}, w^{[s]}) \right)^2 \quad (4)$$

where all  $\mathbf{u}^{[s]}, v^{[s]}, w^{[s]}$  are i.i.d. random vectors or variables drawn from  $\otimes_{t=1, \dots, d} d\mu_t$ ,  $d\mu_i$  and  $d\mu_j$  respectively. Its statistical properties have been investigated in Fruth et al. (2014): it is unbiased and asymptotically normal. More interestingly, it is asymptotically efficient in a class of models with exchangeable variables, indicating that it has the smallest within-class variance. Furthermore,  $\widehat{D}_{\{i,j\}}^{\text{super}}(f)$  vanishes when the theoretical  $D_{\{i,j\}}^{\text{super}}(f)$  is zero: the Monte Carlo error of estimation under this singular case is always zero, which is remarkable.

Finally, the scaled version  $S_{\{i,j\}}^{\text{super}}(f)$  can also be approximated with Monte Carlo replications, by dividing  $\widehat{D}_{\{i,j\}}^{\text{super}}(f)$  by an usual estimator of the overall variance  $D(f)$ .

## 3 New measures for extreme value analysis

The main idea of this paper is to apply what precedes to a functional characterization of the tail dependence structure. We start this section by some backgrounds on the stable tail dependence function, which is

additive under tail independence. Then we derive both new coefficients and a new graph for multivariate extreme value analysis. A small numerical illustration is reported at the end of this section.

### 3.1 The stable tail dependence function

For any arbitrary dimension  $d \geq 2$ , let  $\mathbf{X} = (X^{(1)}, \dots, X^{(d)})$  be a multivariate vector with continuous marginal cumulative distribution functions denoted by  $F_1, \dots, F_d$ . When it exists, the stable tail dependence function (s.t.d.f.) is defined at the point  $\mathbf{u} = (u_1, \dots, u_d)$  by

$$\ell(\mathbf{u}) = \lim_{z \rightarrow \infty} z \mathbb{P} \left( 1 - F_1(X^{(1)}) \leq u_1/z \text{ or } \dots \text{ or } 1 - F_d(X^{(d)}) \leq u_d/z \right)$$

or equivalently by

$$\ell(\mathbf{u}) = \lim_{z \rightarrow \infty} z \left\{ 1 - F \left( F_1^{-1}(1 - u_1/z), \dots, F_d^{-1}(1 - u_d/z) \right) \right\}. \quad (5)$$

This function furnishes a full characterization of the extremal dependence between  $X^{(1)}, \dots, X^{(d)}$ . It is not the only measure for modeling the dependence between extreme components. As such, it is regularly used in the literature in a modeling bid, for instance with parametric construction, or for inference concerns. For the sake of completeness, let us recall the characterization of such asymptotic dependence function. Again  $(\mathbf{e}_1, \dots, \mathbf{e}_d)$  stands for the canonical basis of  $\mathbb{R}^d$ . For a function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ , let  $\Delta_{\mathbf{a}}^{\mathbf{b}} \varphi$  be the cross finite difference of order  $d$  of  $\varphi$  at points  $\mathbf{a}$  and  $\mathbf{b}$  of  $\mathbb{R}^d$ , defined as follows:

$$\begin{aligned} \Delta_{\mathbf{a}}^{\mathbf{b}} \varphi &:= \sum_{I \subseteq \{1, \dots, d\}} (-1)^{d-|I|} \varphi \left( \sum_{i \in I} b_i \mathbf{e}_i + \sum_{i \in \{1, \dots, d\} \setminus I} a_i \mathbf{e}_i \right) \\ &= \varphi(\mathbf{b}) - \varphi(a_1, b_2, \dots, b_d) - \dots - \varphi(b_1, b_2, \dots, b_{d-1}, a_d) \\ &\quad + \varphi(a_1, a_2, b_3, \dots, b_d) + \dots + \varphi(b_1, b_2, \dots, b_{d-2}, a_{d-1}, a_d) - \dots - \dots + (-1)^d \varphi(\mathbf{a}). \end{aligned}$$

**Theorem 3.1** (Ressel (2013)). *The function  $\ell : [0, \infty]^d \rightarrow [0, \infty]$  is a stable tail dependence function if and only if  $\ell$  is homogeneous of order 1,  $\ell(\mathbf{e}_i) = 1$  for  $i \in \{1, \dots, d\}$  and  $\ell$  is fully  $d$ -max decreasing, that is to say:  $\Delta_{\mathbf{a}+h}^{\mathbf{a}} \ell \leq 0$  for all  $\mathbf{a}, h \in [0, \infty]^d$  and, when say  $p$  variables are fixed,  $\ell$  remains a fully  $(d-p)$ -max decreasing function in the remaining variables.*

More tractable properties are well-known. The s.t.d.f. is continuous, convex and, as already said, homogeneous of order 1. Furthermore, from (5) for instance, one gets

$$u_1 \vee \dots \vee u_d \leq \ell(\mathbf{u}) \leq u_1 + \dots + u_d \quad (6)$$

corresponding to *asymptotic full dependence* (lower bound, denoted *AD*) and *asymptotic independence* (upper bound, denoted *AI*).

## 3.2 The tail superset importance coefficients

In the present setting, the random vector  $\mathbf{U}$  from Section 2 is an abstract construction. We will further introduce the following assumption

$$(H_\mu) \quad \text{The measure } \mu \text{ is continuous with support } [0, 1]^d,$$

where  $\mu = \mu_1 \otimes \dots \otimes \mu_d$  refers to the distribution of  $\mathbf{U}$ . First, the requirement on the support allows to substitute the function  $f$  by  $\ell$  in Equation (1) or (2). Indeed, the s.t.d.f. restricted to  $[0, 1]^d$  becomes square integrable, without loss of information by homogeneity of  $\ell$ . Second, the continuity assumption is useful to deduce that a term of variance zero is null, as soon as the function itself is continuous, which is the case of  $\ell$ .

**Definition 1** (Tail superset importance coefficient). *The tail superset importance coefficient is a superset importance indice  $D_{\{i,j\}}^{super}(\ell)$  where  $\ell$  is a stable tail dependence function and where  $\mu$  satisfies  $(H_\mu)$ . Its scaled version is  $S_{\{i,j\}}^{super}(\ell)$ .*

Even though it is associated to a random vector  $\mathbf{X}$ , the function  $\ell$  is a deterministic function. Because  $\ell$  is generally unknown, we provide here a new modelling tool. For instance, under AI, all tail superset importance coefficients vanish. However, these indices are also helpful in less exaggerated situation. By looking at their sorted values, one can trace the pair with the strongest asymptotic dependence until the one with the weakest (in the superset sense, not in the bivariate one). These coefficients can be directly compared, and applied for dimension reduction purposes.

## 3.3 The tail dependograph

We introduce now a new weighted graph which weights are directly given by the tail superset importance coefficients. An *undirected and weighted graph* is a list  $G = (V, \mathbf{g})$  of the set of vertices  $V = \{1, \dots, d\}$  and the collection  $\mathbf{g} = \{g_{1;2}, g_{1;3}, \dots, g_{1;d}, \dots, g_{d-1;d}\}$  of the  $d(d-1)/2$  edge weights. Graphically, we represent the vertex number inside a bubble and plot a segment between  $\{i\}$  and  $\{j\}$  whose width varies proportionally to  $g_{i;j}$ . For a *dependency graph* of a random vector  $\mathbf{X} = (X^{(1)}, \dots, X^{(d)})$ ,  $V$  represents the components of  $\mathbf{X}$  and  $\mathbf{g}$  collects some positive dependence measures between pairs.

When the weight of an edge is close to zero, it indicates some independence. Moreover, ordered values of edge weights reflect ordered strength of dependence between the associated pairs. In this paper, we adopt the graph representation given in the R packages `sensitivity` (Pujol et al., 2017) and `fanovaGraph` (Fruth et al., 2015): the weights are not printed along edges but a visual indication is given through the edges thickness.

We now introduce a new dependency graph for exploring tail dependence.

**Definition 2** (Tail dependency graph). *An undirected and weighted graph  $G$  will be known as a tail dependency graph if the following holds: the vectors  $\mathbf{X}_A = \{X^{(a)}, a \in A\}$  and  $\mathbf{X}_B = \{X^{(a)}, a \in B\}$  are asymptotically independent when there is no edge from any vertex in  $A$  to any vertex in  $B$ , for any two subsets of vertices  $A$  and  $B$  from  $V$ .*

The new indices, namely the tail superset importance coefficients of Section 3.2, are suitable measures to play the role of edge weights of tail dependency graphs.



**Definition 3** (Tail dependograph). *The tail dependograph is a tail dependency graph  $G$  with edge weights given by  $D_{\{i,j\}}^{super}(\ell)$ , namely the tail superset importance coefficient from Definition 1. Similarly, the weights for the scaled tail dependograph are given by the tail sensitivity indices  $S_{\{i,j\}}^{super}(\ell)$ .*

Notice that the tail dependograph corresponds to a FANOVA graph, introduced in [Muehlenstaedt et al. \(2012\)](#), when applied to the s.t.d.f.  $\ell$ .

Section 3.5 shows two examples of tail dependographs when  $d = 3$  and  $d = 4$ . More illustrations for  $d = 4, 10$  and  $21$  are given in Sections 5.1 and 5.3.

### 3.4 Two competitors for tail graphical summaries

To provide the reader with a better understanding of the interest of the tail dependograph, it is useful to compare it with two other graphs relying on well-known summary measures for tail dependence, namely the extremal coefficients. For a s.t.d.f.  $\ell$ , the extremal coefficient associated to subset  $I$  from  $\{1, \dots, d\}$  is given as  $\theta_I(\ell) := \ell(\sum_{i \in I} \mathbf{e}_i)$ . From (6) one should keep in mind that

$$1 \leq \theta_I(\ell) \leq |I|, \quad (7)$$

where the lower (resp. upper) bound corresponds to AD (resp. AI).

The  $\lambda$ -graph is a tail dependency graph with the edge weights given by

$$\lambda_{\{i,j\}} := 2 - \theta_{\{i,j\}}(\ell) = 2 - \ell(\mathbf{e}_i + \mathbf{e}_j).$$

From (7) all the weights lie in the interval  $[0, 1]$ , ranging from AI to AD. In particular they are positive and satisfy the tail dependency graph property. By extension to supersets of a pair, we define *the  $\lambda$ -supergraph* as the tail dependency graph  $G$  with weights

$$\lambda_{\{i,j\}}^{super} := \sum_{I \supseteq \{i,j\}} \{|I| - \theta_I(\ell)\} = (d+2)2^{d-3} - \sum_{I \supseteq \{i,j\}} \ell\left(\sum_{i \in I} \mathbf{e}_i\right).$$

Similarly to the  $\lambda$ -graph, (7) ensures that the weights are valid. They range from 0 (AI) to  $d2^{d-3}$  (AD), and could be easily normalized to range within  $[0, 1]$ . The value  $(d+2)2^{d-3}$  as well as the upper bound  $d2^{d-3}$  have been derived from standard combinatorial results, with details given in the proofs' section.

### 3.5 Illustrating the tail dependograph on the asymmetric logistic model

Among the parametric subfamilies of s.t.d.f., one of the well-known form is the asymmetric logistic model. Introduced in full generality by [Tawn \(1990\)](#), one has

$$\ell(\mathbf{u}) = \sum_{b \in B} \left( \sum_{i \in b} (\beta_{i,b} u_i)^{1/\alpha_b} \right)^{\alpha_b}$$

where  $B$  is the set of all non-empty subsets of  $\{1, \dots, d\}$ , the dependence parameters  $\alpha_b$  lie in  $(0, 1]$  and the asymmetric collection of weights  $\beta_{i,b}$  are coefficients from  $[0, 1]$  satisfying  $\sum_{b \in B: b \ni i} \beta_{i,b} = 1$ . For standard Fréchet univariate margins, the simulation of vectors with such multivariate dependence is included in the

R package `evd` of [Stephenson \(2002\)](#) and detailed in [Stephenson \(2003\)](#).

Let us consider two examples extracted from the help file of the function `rmvevd` of the package `evd`. For the dimension  $d = 3$ , let s.t.d.f. (8) be the asymmetric logistic model with the following family of coefficients

$$\begin{aligned} (\beta_{1,\{1\}}, \beta_{2,\{2\}}, \beta_{3,\{3\}}) &= (.4, .1, .6) \quad (\beta_{1,\{12\}}, \beta_{2,\{12\}}) = (.3, .2) \text{ with } \alpha_{\{12\}} = .6, \\ (\beta_{1,\{13\}}, \beta_{3,\{13\}}) &= (.1, .1) \text{ with } \alpha_{\{13\}} = .5, \quad (\beta_{2,\{23\}}, \beta_{3,\{23\}}) = (.4, .1) \text{ with } \alpha_{\{23\}} = .8, \\ &\text{and } (\beta_{1,\{123\}}, \beta_{2,\{123\}}, \beta_{3,\{123\}}) = (.2, .3, .2) \text{ with } \alpha_{\{123\}} = .3. \end{aligned} \quad (8)$$

Below we provide a graph representation of the s.t.d.f., through the tail superset indices in Figure 1(a), the  $\lambda$ -graph in (b) and  $\lambda$ -supergraph in (c).

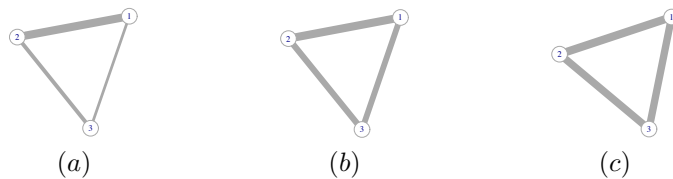


Figure 1: The tail dependograph (a), the  $\lambda$ -graph (b), and the  $\lambda$ -supergraph (c) for s.t.d.f. (8).

Corresponding numerical values are available in Table 1.

Pair $\{i, j\}$	$\{1, 2\}$	$\{1, 3\}$	$\{2, 3\}$
$D_{\{i,j\}}^{\text{super}}(\ell)$	0.0026	0.0008	0.0012
$\lambda_{\{i,j\}}(\ell)$	0.2946	0.2123	0.2229
$\lambda_{\{i,j\}}^{\text{super}}(\ell)$	0.8736	0.7913	0.8019

Table 1: Values of theoretical coefficients for  $\ell$  being the s.t.d.f. (8).

The tail dependograph points out a strict ordering between the three pairs thanks to tail superset indices, that is not visible by looking at the two other graphs. Focusing on the  $\lambda$ -graph, one can isolate the pair  $\{1, 2\}$  from the two others. Looking at the  $\lambda$ -supergraph, we would conclude that the three variables are exchangeable with respect to this criterion. Of course, the  $\lambda$ -graph remains interesting since it relies on coefficients  $\theta_I$  that are interpretable through the equality

$$\mathbb{P}\left(X^{(i)} \leq F_i^{-1}(p), \text{ for all } i \in I\right) = p^{\theta_I(\ell)}, \quad (9)$$

written above when  $F$  the distribution of  $\mathbf{X}$  is a max-stable distribution. With the tail superset coefficients, we lose this kind of interpretation. Instead a global variance  $D(\ell)$  is a reference to which each  $D_{\{i,j\}}^{\text{super}}(\ell)$  can be compared so that they also can be compared one to another directly.

We illustrate now the tail dependograph in dimension 4 under s.t.d.f. (10) with the following family of coefficients

$$\begin{aligned} (\beta_{1,\{123\}}, \beta_{2,\{123\}}, \beta_{3,\{123\}}) &= (.2, .1, .2), \quad (\beta_{1,\{124\}}, \beta_{2,\{124\}}, \beta_{4,\{124\}}) = (.1, .1, .2), \\ (\beta_{1,\{134\}}, \beta_{2,\{134\}}, \beta_{3,\{134\}}) &= (.3, .4, .1), \quad (\beta_{2,\{234\}}, \beta_{3,\{234\}}, \beta_{4,\{234\}}) = (.2, .2, .2), \\ &\text{and } (\beta_{1,\{1234\}}, \beta_{2,\{1234\}}, \beta_{3,\{1234\}}, \beta_{4,\{1234\}}) = (.4, .6, .2, .5) \\ &\text{with } \alpha_{\{123\}} = \alpha_{\{124\}} = \alpha_{\{134\}} = \alpha_{\{234\}} = \alpha_{\{1234\}} = .7. \end{aligned} \quad (10)$$

We plot the three graphs in Figure 2. Again, the tail superset indices reveals details in the tail

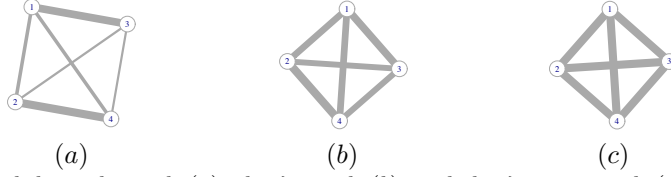


Figure 2: The tail dependograph (a), the  $\lambda$ -graph (b), and the  $\lambda$ -supergraph (c) for s.t.d.f. (10).

dependence structure that are not perceptible with the two competitors.

## 4 Nonparametric statistical inference

Let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  be  $n$  independent copies of the  $d$ -variate vector  $\mathbf{X}$ . For  $t = 1, \dots, d$  and  $s = 1, \dots, n$ , we denote by  $X_s^{(t)}$  the  $t$ -th component of the vector  $\mathbf{X}_s$ . Let  $k$  be a positive integer smaller than  $n$ . In extreme value theory,  $k$  represents the number of largest values considered in inference procedures. When  $k$  is too large, the estimation relies on some observations that are not extreme, introducing some bias. On the contrary, when  $k$  is too small, there is no bias but the variance of the estimator is large. Nice asymptotic properties are based on such balance. In this section we derive non parametric estimators of the main measures of the paper: the indices  $D_{\{i,j\}}^{\text{super}}(\ell)$  in the first part, and the whole Hoeffding-Sobol decomposition in the second one. The particular case of uniform margins for  $\mathbf{U}$  is then handled. Finally, asymptotic results are addressed at the end of the section.

### 4.1 Empirical versions of the tail superset importance coefficients

Taking into account Section 2.3, once an estimator  $\hat{\ell}$  for  $\ell$  is given, then the computation of a tail superset importance coefficient is possible by substituing  $f$  with  $\hat{\ell}$  in Formula (4). Consequently, the estimating issue could already appear settled. More interestingly, exact expressions can be derived for new indices  $D_{\{i,j\}}(\hat{\ell})$  when the empirical counterpart of the true function  $\ell$  is used.

Recall first the natural estimation of the stable tail dependence function, given by Huang (1992),

$$\hat{\ell}_{k,n}(\mathbf{u}) = \frac{1}{k} \sum_{s=1}^n \mathbf{1} \left\{ X_s^{(1)} \geq X_{n-[ku_1]+1,n}^{(1)} \text{ or } \dots \text{ or } X_s^{(d)} \geq X_{n-[ku_d]+1,n}^{(d)} \right\},$$

where  $z$  has been replaced by  $n/k$  in (5),  $F$  by its empirical version, and  $F_t^{-1}(1 - z^{-1}u_t)$  for  $t = 1, \dots, d$  by  $X_{n-[nz^{-1}u_t]+1,n}^{(t)}$ . This estimator can be written as

$$\hat{\ell}_{k,n}(\mathbf{u}) = \frac{1}{k} \sum_{s=1}^n \mathbf{1} \left\{ u_1 \geq \tilde{R}_s^{(1)} \text{ or } \dots \text{ or } u_d \geq \tilde{R}_s^{(d)} \right\}, \quad (11)$$

in terms of  $\tilde{R}_s^{(t)} := (n - R_s^{(t)} + 1)/k$  where  $R_s^{(t)}$  denotes the rank of  $X_s^{(t)}$  among  $X_1^{(t)}, \dots, X_n^{(t)}$ .

We now give the expression of the empirical version of the tail superset importance coefficients in terms of the rank observations.

**Proposition 4.1.** *The empirical tail superset importance coefficient can be expressed as*

$$D_{\{i,j\}}^{\text{super}}(\hat{\ell}_{k,n}) = \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \left\{ \prod_{t \in \{i,j\}} \left( \mu_t \left( \tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)} \right) - \mu_t \left( \tilde{R}_s^{(t)} \right) \mu_t \left( \tilde{R}_{s'}^{(t)} \right) \right) \prod_{t \notin \{i,j\}} \mu_t \left( \tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)} \right) \right\}.$$

The empirical scaled tail superset importance coefficient is estimated by  $\hat{S}_{k,n;\{i,j\}}^{\text{super}} := \frac{D_{\{i,j\}}^{\text{super}}(\hat{\ell}_{k,n})}{D(\hat{\ell}_{k,n})}$  where

$$D(\hat{\ell}_{k,n}) = \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \left\{ \prod_{t=1}^d \mu_t \left( \tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)} \right) - \prod_{t=1}^d \mu_t \left( \tilde{R}_s^{(t)} \right) \mu_t \left( \tilde{R}_{s'}^{(t)} \right) \right\}.$$

Proposition 4.1 provides accurate versions of these coefficients. Indeed, these ranked based expressions will prevent from Monte Carlo error estimation that one would obtain applying Section 2.3 to the empirical estimator of  $\ell$  given by (11).

## 4.2 Complete empirical Hoeffding-Sobol decomposition

The main difference between previous sections and the current one remains on the fact that any subset  $I$  from  $\{1, \dots, d\}$  is considered here. One could have start this paper by defining the new summaries associated to any subsets, and not only the pairs. But the reason comes from our strong wish to introduce a new graph for extreme analysis, that is only feasible with weights associated to pairs. The coefficient  $D_I^{\text{super}}(\ell)$  generalizes the pair version and stands for  $\sum_{J \supseteq I} D_J(\ell)$ . The global sensitivity indices are also extended to any subset  $I$  by

$$S_I(\ell) = \frac{D_I(\ell)}{D(\ell)}$$

and

$$S_I^{\text{super}}(\ell) = \frac{D_I^{\text{super}}(\ell)}{D(\ell)}.$$

These two ratios satisfy

$$0 \leq S_I(\ell) \leq S_I^{\text{super}}(\ell) \leq 1$$

with extreme behavior  $S_I^{\text{super}}(\ell) = 0$  or  $S_I(\ell) = 1$  that characterize respectively the situation where  $\ell$  does not depend on  $I$  or where  $\ell$  only depends on  $I$ .

The Hoeffding-Sobol decomposition of the function  $\hat{\ell}_{k,n}$  is given by

$$\hat{\ell}_{k,n}(u_1, \dots, u_d) = \hat{\ell}_{k,n;\emptyset} + \sum_{i=1}^d \hat{\ell}_{k,n;\{i\}}(u_i) + \sum_{i < j} \hat{\ell}_{k,n;\{i,j\}}(u_i, u_j) + \dots + \hat{\ell}_{k,n;\{1,\dots,d\}}(u_1, \dots, u_d)$$

where the subfunctions are centered and orthogonal as explained in Section 2.1. The framework under study is fairly unusual in global sensitivity analysis since most of the empirical computation can be done, all the terms have an explicit expression as stated below.

**Theorem 4.2.** *Let  $I \subseteq \{1, \dots, d\}$  be non empty. The associated subfunction in the decomposition is*

$$\hat{\ell}_{k,n;I}(\mathbf{u}_I) = -\frac{1}{k} \sum_{s=1}^n \left\{ \prod_{t \in I} \left( \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\} - \mu_t \left( \tilde{R}_s^{(t)} \right) \right) \prod_{t \notin I} \mu_t \left( \tilde{R}_s^{(t)} \right) \right\},$$

and its variance, which corresponds to an estimate of  $D_I(\ell)$ , has the following form

$$D_I(\hat{\ell}_{k,n}) = \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \left( \prod_{t \in I} \left\{ \mu_t \left( \tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)} \right) - \mu_t \left( \tilde{R}_s^{(t)} \right) \mu_t \left( \tilde{R}_{s'}^{(t)} \right) \right\} \prod_{t \notin I} \mu_t \left( \tilde{R}_s^{(t)} \right) \mu_t \left( \tilde{R}_{s'}^{(t)} \right) \right).$$

The superset part of  $\hat{\ell}_{n,k}(\mathbf{u})$  associated to the subset  $I$ , that is the part in Hoeffding-Sobol decomposition that contains the variables  $\{u_t, t \in I\}$  is given by

$$\hat{\ell}_{k,n;I}^{super}(\mathbf{u}) = -\frac{1}{k} \sum_{s=1}^n \prod_{t \in I} \left( \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\} - \mu_t \left( \tilde{R}_s^{(t)} \right) \right) \prod_{t \notin I} \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\},$$

with variance

$$D_I^{super}(\hat{\ell}_{k,n}) = \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \left\{ \prod_{t \in I} \left( \mu_t \left( \tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)} \right) - \mu_t \left( \tilde{R}_s^{(t)} \right) \mu_t \left( \tilde{R}_{s'}^{(t)} \right) \right) \prod_{t \notin I} \mu_t \left( \tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)} \right) \right\}.$$

Explicit expressions are then immediately deduced for the estimators of  $S_I(\ell)$  and  $S_I^{super}(\ell)$  defined as

$$\hat{S}_{k,n;I} := \frac{D_I(\hat{\ell}_{k,n})}{D(\hat{\ell}_{k,n})} \quad \text{and} \quad \hat{S}_{k,n;I}^{super} := \frac{D_I^{super}(\hat{\ell}_{k,n})}{D(\hat{\ell}_{k,n})},$$

by plugging the expression of  $D(\hat{\ell}_{k,n})$  given in Proposition 4.1. Additionnaly, the constant term in the Hoeffding-Sobol decomposition can be written as

$$\hat{\ell}_{k,n;\emptyset} = \frac{n}{k} - \frac{1}{k} \sum_{s=1}^n \prod_{t=1}^d \mu_t \left( \tilde{R}_s^{(t)} \right).$$

By construction, the inequality  $D_I(\ell) \leq D_I^{super}(\ell)$  is verified by the empirical estimates which expressions are stated above. This can be checked directly by observing that

$$\mu_t \left( \tilde{R}_s^{(t)} \right) \mu_t \left( \tilde{R}_{s'}^{(t)} \right) = \mu_t \left( \tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)} \right) \mu_t \left( \tilde{R}_s^{(t)} \vee \tilde{R}_{s'}^{(t)} \right) \leq \mu_t \left( \tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)} \right).$$

### 4.3 Under standard uniform distributions

The issue of the optimal choice for the marginal distribution of  $\mathbf{U}$  will not be raised in this paper. However, it is natural to consider the uniform distribution. By enforcing  $\mathbf{u}$  to be in the unit cube, introducing

$$\bar{R}_s^{(t)} := \tilde{R}_s^{(t)} \wedge 1$$

allows to write

$$\hat{\ell}_{k,n}(\mathbf{u}) = \frac{1}{k} \sum_{s=1}^n \mathbf{1}\left\{ u_1 \geq \bar{R}_s^{(1)} \text{ or } \dots \text{ or } u_d \geq \bar{R}_s^{(d)} \right\}.$$

When each probability measure  $\mu_t$  is the standard uniform distribution, all preceding expressions can be simplified thanks to the equalities

$$\mu_t \left( \tilde{R}_s^{(t)} \right) = \bar{R}_s^{(t)}$$

and

$$\mu_t \left( \tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)} \right) = \bar{R}_s^{(t)} \wedge \bar{R}_{s'}^{(t)} .$$

In particular, the empirical tail superset importance coefficient is

$$D_{\{i,j\}}^{\text{super}}(\hat{\ell}_{k,n}) = \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \left\{ \prod_{t \in \{i,j\}} \left( \bar{R}_s^{(t)} \wedge \bar{R}_{s'}^{(t)} - \bar{R}_s^{(t)} \bar{R}_{s'}^{(t)} \right) \prod_{t \notin \{i,j\}} \bar{R}_s^{(t)} \wedge \bar{R}_{s'}^{(t)} \right\}$$

and its scaled version becomes

$$\hat{S}_{k,n;\{i,j\}}^{\text{super}} = \frac{\sum_{s=1}^n \sum_{s'=1}^n \left\{ \prod_{t \in \{i,j\}} \left( \bar{R}_s^{(t)} \wedge \bar{R}_{s'}^{(t)} - \bar{R}_s^{(t)} \bar{R}_{s'}^{(t)} \right) \prod_{t \notin \{i,j\}} \bar{R}_s^{(t)} \wedge \bar{R}_{s'}^{(t)} \right\}}{\sum_{s=1}^n \sum_{s'=1}^n \left\{ \prod_{t=1}^d \bar{R}_s^{(t)} \wedge \bar{R}_{s'}^{(t)} - \prod_{t=1}^d \bar{R}_s^{(t)} \bar{R}_{s'}^{(t)} \right\}} .$$

## 4.4 Consistency

The large sample behavior of the new empirical indices are now stated. In order to prove consistency of the empirical tail superset importance coefficients, we only need to focus on settings where the estimator  $\hat{\ell}_{k,n}$  is uniformly consistent for the estimation of  $\ell$ .

### The multivariate domain of attraction assumption

The multivariate distribution function  $F$  with continuous marginal distribution functions  $F_t$  is such that the limit

$$\lim_{z \rightarrow \infty} z \left\{ 1 - F \left( F_1^{-1}(1 - u_1/z), \dots, F_d^{-1}(1 - u_d/z) \right) \right\} = \ell(\mathbf{u}) \quad (12)$$

exists for all  $\mathbf{u}$  in the positive orthant.

**Theorem 4.3** (Theorem 7.2.1 of [de Haan and Ferreira \(2006\)](#)). *Let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  be  $n$  independent copies of the random vector  $\mathbf{X}$ . Assume (12) holds and let  $k$  be an intermediate sequence:  $k = k(n)$  such that  $k \rightarrow \infty$  as  $n \rightarrow \infty$  but with  $k = o(n)$ . Then, as  $n$  tends to infinity,*

$$\sup_{\mathbf{u} \in [0,1]^d} \left| \hat{\ell}_{k,n}(\mathbf{u}) - \ell(\mathbf{u}) \right| \xrightarrow{\mathbb{P}} 0 .$$

This uniform convergence is valid for any subfunction of the Hoeffding-Sobol decomposition, as stated in the next result.

**Corollary 4.4.** *Under conditions of Theorem 4.3, all the empirical subfunctions defined in Theorem 4.2 are uniformly consistent*

$$\sup_{\mathbf{u}_I \in [0,1]^{|I|}} \left| \hat{\ell}_{k,n;I}(\mathbf{u}_I) - \ell_I(\mathbf{u}_I) \right| \xrightarrow{\mathbb{P}} 0 .$$

As a consequence, the empirical variances  $D_I(\hat{\ell}_{k,n})$  or  $D_I^{\text{super}}(\hat{\ell}_{k,n})$  and the empirical ratios  $\hat{S}_{k,n;I}$  or  $\hat{S}_{k,n;I}^{\text{super}}$  from Theorem 4.2 are consistent for their equivalent theoretical forms.

In particular, the (scaled) tail superset coefficients, whose empirical versions are detailed in Proposition 4.1, satisfy

$$D_{\{i,j\}}^{\text{super}}(\hat{\ell}_{k,n}) \xrightarrow{\mathbb{P}} D_{\{i,j\}}^{\text{super}}(\ell) \quad \text{and} \quad \hat{S}_{k,n;\{i,j\}}^{\text{super}} \xrightarrow{\mathbb{P}} S_{\{i,j\}}^{\text{super}}(\ell) ,$$

so that the empirical (scaled) dependograph is consistent.

## 4.5 Asymptotic normality

We need to assume some additional conditions in order to establish some asymptotics for the distribution of the empirical tail superset importance coefficients.

### The second order condition

There exists a function  $\alpha(\cdot)$  that tends to zero with constant sign and  $m$  that is not proportional to  $\ell$  such that

$$\lim_{z \rightarrow \infty} \frac{z \{1 - F(F_1^{-1}(1 - u_1/z), \dots, F_d^{-1}(1 - u_d/z))\} - \ell(\mathbf{u})}{\alpha(z)} = m(\mathbf{u}) \quad (13)$$

uniformly on  $[0, 1]^d$ . Note that this condition quantifies the speed of convergence in the multivariate domain of attraction assumption.

The next result can be seen as the concatenation of several statements already established in the literature. See for instance (de Haan and Ferreira, 2006; Theorem 7.2.2) for a very similar result. It provides a general setting to describe the asymptotic distribution of the function  $\hat{\ell}_{k,n}(\cdot)$ .

**Lemma 4.5.** *Let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  be  $n$  independent copies of  $\mathbf{X}$ . Assume (13) holds. Suppose further that the first order partial derivative of  $\ell$  (denoted by  $\partial_t \ell$  for  $t = 1, \dots, d$ ) exists and that  $\partial_t \ell$  is continuous on the set of points  $\{\mathbf{u} = (u_1, \dots, u_d) \in \mathbb{R}_+^d : u_t > 0\}$ .*

*Let  $k$  be an intermediate sequence such that  $\sqrt{k}\alpha(n/k) \rightarrow \alpha$ . Then as  $n$  tends to infinity,*

$$\sqrt{k} \left\{ \hat{\ell}_{k,n}(\mathbf{u}) - \ell(\mathbf{u}) \right\} \xrightarrow{d} Z_\ell(\mathbf{u}) + \alpha m(\mathbf{u}), \quad (14)$$

*in the space of functions in  $[0, 1]^d$  that are right-continuous and have left-hand limits. The process  $Z_\ell$  above can be stochastically represented as follows*

$$Z_\ell(\mathbf{u}) := W_\ell(\mathbf{u}) - \sum_{j=1}^d W_\ell(u_j \mathbf{e}_j) \partial_j \ell(\mathbf{u}). \quad (15)$$

*The process  $W_\ell$  is a continuous centered Gaussian process with  $\mathbb{E}[W_\ell(\mathbf{u})W_\ell(\tilde{\mathbf{u}})] = \nu\{H(\mathbf{u}) \cap H(\tilde{\mathbf{u}})\}$  given in terms of the intensity measure  $\nu$  defined by  $\nu\{\tilde{\mathbf{u}} \in \mathbb{R}_+^d : \text{there exists } t \text{ such that } \tilde{u}_t > u_t\} := \ell(\mathbf{u})$ , and  $H(\mathbf{u}) = \{\tilde{\mathbf{u}} \in \mathbb{R}_+^d : \text{there exists } t \text{ such that } 0 \leq \tilde{u}_t \leq u_t\}$ .*

The asymptotic distributional expansion (15) is inherited by subfunctions of the ANOVA decomposition, as stated in the next result.

**Corollary 4.6.** *Suppose that the assumptions of Lemma 4.5 hold. Assume moreover that the function  $m$  in (13) and that the sample paths of the stochastic process (15) are both square integrable with respect to  $\otimes_{t=1}^d d\mu_t$ . Then, as  $n$  tends to infinity,*

$$\sqrt{k} \left\{ \hat{\ell}_{k,n,I}(\mathbf{u}_I) - \ell_I(\mathbf{u}_I) \right\} \xrightarrow{d} Z_{\ell,I}(\mathbf{u}_I) + \alpha m_I(\mathbf{u}_I)$$

*where the stochastic process  $Z_{\ell,I} : \omega \in \Omega \mapsto \{Z_\ell(\omega)\}_I$  corresponds to the path of the subfunction associated to subset  $I$  in the Hoeffding-Sobol decomposition of  $Z_\ell(\omega)$ .*

*Moreover,*

$$\sqrt{k} \left( D_I(\hat{\ell}_{k,n}) - D_I(\ell) \right) \xrightarrow{d} 2\mathbb{E}[\ell_I(\mathbf{U}_I)\{Z_{\ell,I}(\mathbf{U}_I) + \alpha m_I(\mathbf{U}_I)\}]$$

*where the expectation refers to the integral with respect of  $\otimes_{t \in I} d\mu_t$ . Finally, if  $\ell_I \equiv 0$  that is equivalent*

to  $D_I(\ell) = 0$ , one gets

$$kD_I(\hat{\ell}_{k,n}) \xrightarrow{d} D_I(Z_\ell + \alpha m).$$

Let us point out that by adding terms from Corollary 4.6, one also gets, if  $D_I^{\text{super}}(\ell) > 0$ ,

$$\sqrt{k} \left( D_I^{\text{super}}(\hat{\ell}_{k,n}) - D_I^{\text{super}}(\ell) \right) \xrightarrow{d} \sum_{J \supseteq I} 2\mathbb{E}[\ell_J(\mathbf{U}_J) \{Z_{\ell,J}(\mathbf{U}_J) + \alpha m_J(\mathbf{U}_J)\}].$$

The next proposition states again the asymptotic normality for the empirical tail superset coefficients. Another limiting expression is thus obtained, depending only on the function  $\ell$ , and not on the subfunctions  $\{\ell_J, J \supseteq I\}$  as above.

**Proposition 4.7.** *Under assumptions of Lemma 4.5, as  $n$  tends to infinity,*

$$\sqrt{k} \left( D_I^{\text{super}}(\hat{\ell}_{k,n}) - D_I^{\text{super}}(\ell) \right) \xrightarrow{d} \frac{1}{2^{|I|-1}} \int \Delta_I[\ell, Z_\ell + \alpha m](\mathbf{u}_I, \mathbf{v}) d\mathbf{u}_I d\mathbf{v}$$

where

$$\Delta_I[\ell, Z](\mathbf{u}_I, \mathbf{v}) = \left( \sum_{J \subseteq I} (-1)^{|I \setminus J|} \ell(\mathbf{u}_J, \mathbf{v}_{-J}) \right) \left( \sum_{K \subseteq I} (-1)^{|I \setminus K|} Z(\mathbf{u}_K, \mathbf{v}_{-K}) \right).$$

If  $\sum_{J \subseteq I} (-1)^{|I \setminus J|} \ell(\mathbf{u}_J, \mathbf{v}_{-J}) \equiv 0$  which is equivalent to  $D_I^{\text{super}}(\ell) = 0$  then

$$kD_I^{\text{super}}(\hat{\ell}_{k,n}) \rightarrow D_I^{\text{super}}(Z_\ell + \alpha m).$$

In particular, when  $I$  stands for the pair  $\{i, j\}$ , one derives the asymptotic normality for the indices involved in the empirical tail dependograph. More precisely, the convergence of  $D_{\{i,j\}}^{\text{super}}(\hat{\ell}_{k,n})$  to  $D_{\{i,j\}}^{\text{super}}(\ell)$  leads to two different regimes. In general, one achieves the classical root square rate of convergence. But when the true tail superset importance coefficient vanishes, then the rate is faster.

One can note that each factor in the definition above of  $\Delta_I$  is closely related to that of the cross finite difference introduced before Theorem 3.1. Indeed, as soon as  $\varphi(\cdot) = \ell(\cdot, \mathbf{v}_{\{1, \dots, d\} \setminus I})$

$$\sum_{J \subseteq I} (-1)^{|I \setminus J|} \ell(\mathbf{u}_J, \mathbf{v}_{-J}) = \Delta \varphi_{\mathbf{v}_{I \setminus J}}^{\mathbf{u}_J}.$$

## 5 Some illustrations

We exhibit first a setting where classical tail measures fail in finding extreme dependence. On simulation, we study the performance of the non parametric estimators. In particular, the order of tail superset importance between pairs is rather good for reasonable sample sizes. Finally, the last section illustrates the new indices and graphs on real data.

### 5.1 Revealing asymptotic dependence under asymmetric models

We exhibit now an example where  $\delta$  and  $\lambda$  do not give the same information. Let us consider  $d = 4$  and the parametric asymmetric model with s.t.d.f. given as

$$\ell(u_1, u_2, u_3, u_4) = (1 - \psi)(u_1 + u_2) + \psi \left( u_1^{1/\alpha} + u_2^{1/\alpha} \right)^\alpha + (1 - \tilde{\psi})(u_3 + u_4) + \tilde{\psi} \left( u_3^{1/\tilde{\alpha}} + u_4^{1/\tilde{\alpha}} \right)^{\tilde{\alpha}} \quad (16)$$



Each bivariate component is a mixture of AI with the logistic model of Gumbel introduced by [Tawn \(1990\)](#). Fix  $(\psi, \alpha) = (0.2, 0.2)$  and  $(\tilde{\psi}, \tilde{\alpha}) = (0.8, 0.83)$ . The pairplots  $(X_1, X_2)$  and  $(X_3, X_4)$  are given respectively on the left and right of [Figure 3](#) under Gumbel margins.

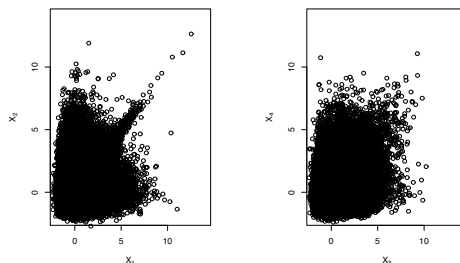


Figure 3: Two pair plots for  $\mathbf{X}$  a 4-dimensional extreme value random vector with Gumbel margins and s.t.d.f. (16). Left:  $(X_1, X_2)$ . Right:  $(X_3, X_4)$ .

We have  $\lambda_{12} \simeq 0.17 < \lambda_{34} \simeq 0.18$  whereas  $D_{12}^{\text{super}} \simeq 3.09 \times 10^{-4} > D_{34}^{\text{super}} \simeq 1.81 \times 10^{-4}$ . The new index better detects the strongest tail dependent pair even in a small proportion. In terms of graphs, this example leads to [Figure 4](#), where the difference between the two graphs' messages is eloquent.

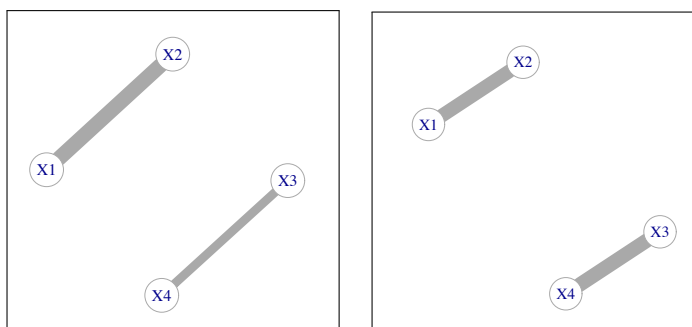


Figure 4: The tail dependograph (left) and the  $\lambda$ -graph (right) for s.t.d.f. (16).

## 5.2 Finite sample performances in estimating the tail superset indices

We consider again the s.t.d.f. (8) from [Section 3.5](#). Recall that numerical values of several indices were collected in [Table 1](#). Our goal is now to estimate the pairwise coefficients  $\lambda_{\{i,j\}}$  and  $D_{\{i,j\}}^{\text{super}}$  on trivariate samples. We assume the distribution to be trivariate max-stable with Fréchet margins.

Computed on several repetitions, we first record the proportion of trials that ordered correctly a triplet of theoretical coefficients. It means that a trial is a success when  $D_{\{1,2\}}^{\text{super}}(\hat{\ell}_{k,n}) > D_{\{2,3\}}^{\text{super}}(\hat{\ell}_{k,n}) > D_{\{1,3\}}^{\text{super}}(\hat{\ell}_{k,n})$  or respectively when  $\lambda_{\{1,2\}}(\hat{\ell}_{k,n}) > \lambda_{\{2,3\}}(\hat{\ell}_{k,n}) > \lambda_{\{1,3\}}(\hat{\ell}_{k,n})$  depending on the criterion we are examining. Making the sample size  $n$  varying, we get the curves presented in [Figure 5](#). The values for the experiment are:  $n \in \{100, 200, \dots, 1000\}$ ,  $N = 1000$  repetitions and several choices for  $k$  have been taken. We only plot two of them:  $k = 50$  (plain lines) and  $k = n/2$  (dotted lines), that roughly speaking correspond to the worst and best cases.

To furnish most of the details, we also provide the boxplots associated to the estimation of each triplets in both situation:  $k = 50$  and  $k = n/2$  when  $n = 1000$  in [Figure 6](#). It is worth noting that even if the coefficients are not so well estimated when  $k$  is large (see [Figure 6 \(b\) and \(d\)](#)), which is a classical

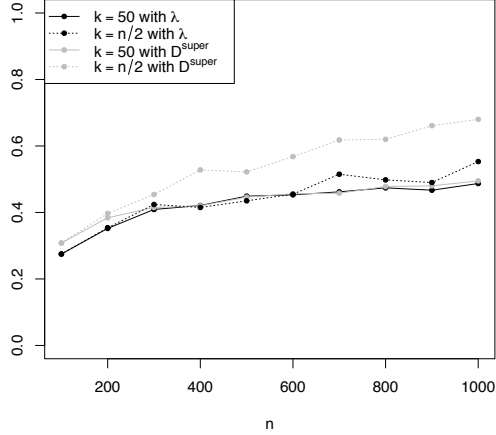


Figure 5: Proportion of well-ordered triplets of theoretical coefficients. For the first (resp. second) line of Table 1, results are in grey (resp. black).

observation in extreme, the proportion of well-ordered triplets remains good, and even increases compared to the value obtained for a small  $k$  (see Figure 6 (a) and (c)).

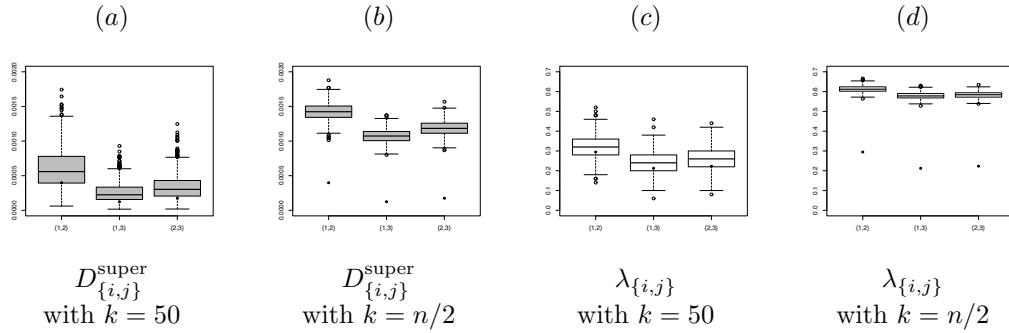


Figure 6: Boxplots when  $n = 1000$  for the estimation of  $D_{\{i,j\}}^{\text{super}}$  (left) and  $\lambda_{\{i,j\}}$  (right). For both type of indices, the first triplet is obtained with  $k = 50$  and the second with  $k = n/2$ . Black points represent true values, already given in Table 1.

### 5.3 Selection of the strongest tail dependences in large dimension

We want to motivate here the use of the tail superset importance coefficients for selecting the pairs with strongest asymptotic dependence. One should not understand this in the sense of strongest extreme bivariate structure of dependence. Indeed, recall that the tail superset importance coefficient of a pair measures the contribution in the global tail dependence not only through the pair itself, but also through all its supersets.

#### 5.3.1 Log-returns contagion

Let consider the close prices of ten stock indices from package `qrmdata` (Hofert and Hornik, 2016): S&P 500 (SP500), Dow Jones (DJ), NASDAQ 100 (NASD), Swiss Market Index (SMI), Euro Stoxx 50 (EURS), Cotation Assistée en Continu (CAC), Deutscher Aktienindex (DAX), Hang Seng Index (HSI), Shanghai

Stock Exchange Composite Index (SSEC), and the NIKKEI (NIKK). Focusing on log-returns for the period 1990 to 2015, we compute renormalized yearly maxima. We thus get a 10-dimensional vector with unit Fréchet marginal distributions and joint distribution characterized by an unknown s.t.d.f.  $\ell$ . The complete tail dependograph is provided in Figure 7(a). Because of the size of the problem  $d = 10$ , we get 45 edges. Of course, the tail indices are not all of the same importance and one could be interested by partial dependographs with the intention of dimension reduction or guidelines for construction of parametric models. Consequently, we also furnish two additional plots. Figure 7(b) shows the values of the tail superset importance coefficients  $D_{\{i,j\}}^{\text{super}}$  in a decreasing order while Figure 7(c) represents the size of a partial tail dependograph (number of variables involved by at least one edge) with respect to the number of the largest tail indices that are plotted.

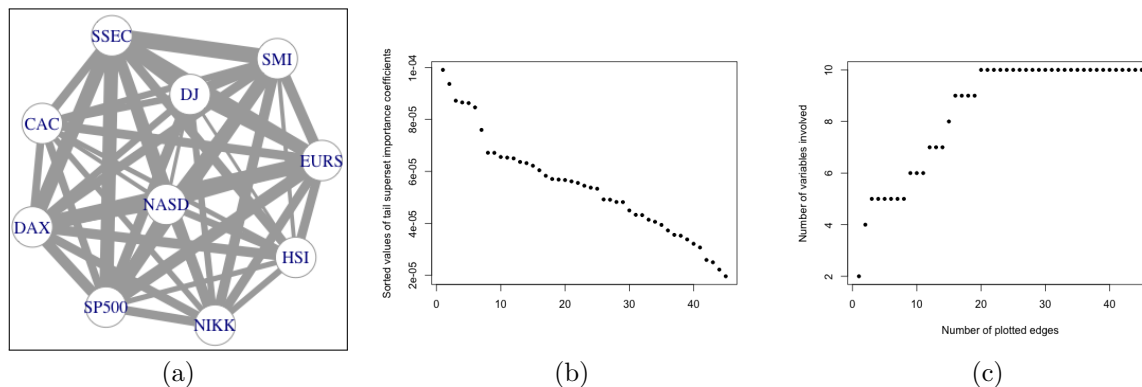


Figure 7: For ten log-returns over the period 1990-2015: (a) Complete empirical tail dependograph (b) Sorted estimated values of  $D_{\{i,j\}}^{\text{super}}$  (c) Number of variables involved in a partial tail dependograph with respect to the number of plotted edges.

The largest empirical tail superset coefficient is obtained for the pair SSEC – SMI. The second largest links EURS and DAX. The eight largest empirical tail superset coefficients concern the group  $\{SMI, SSEC, SP500, EURS, DAX\}$ . Then DJ joins through its tail superset dependence with SSEC, as illustrated in Figure 8(a). One has to wait for the 12th largest (resp. 15th, 16th) value of these empirical measures to see a connection of NIKK (resp. HSI, NASD) with the group through SSEC (resp. DAX, SSEC). Finally the 20th largest value of the empirical tail superset coefficients links CAC with the rest of the group via EURS, as plotted in Figure 8(b).

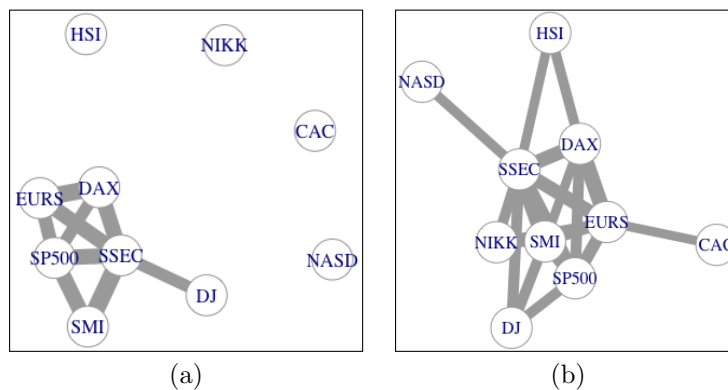


Figure 8: Partial empirical tail dependographs for ten log-returns over the period 1990-2015: (a) The nine largest empirical tail superset coefficients (b) the twenty largest.

We do not want to put one methodology up and not another to select variables. Our wish is to provide new indices, to compute or estimate them, and to share the interest for a number of related issues based on the assessment of these values. Selection with respect to a criterion is a very general question that is beyond the scope of the underlying study. For that reason, we do not conclude in measuring the effective dimension of the tail dependence structure.

### 5.3.2 Temperatures in France

The yearly maxima temperatures of 21 continental cities of France have been recorded during the period 1946-2000. We plot in Figure 9 two partial dependographs and the complete one. Of course, looking at all the empirical tail superset informations simultaneously on Figure 9(c) is uninformative because of the dimension of the problem here. However, the strongest connections can be observed in the North part, and after that, a strong dependence structure appears in the South, as one can see in Figure 9(a). In the partial tail dependograph of Figure 9(b), links between North and South are already present.

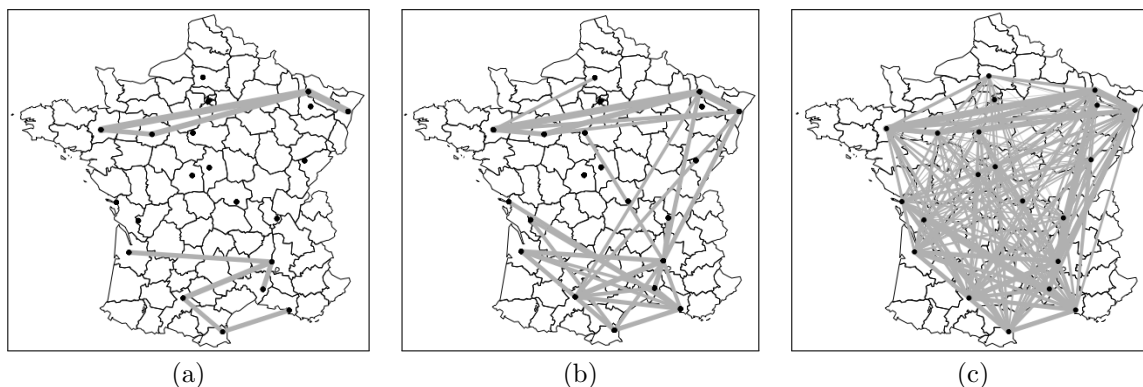


Figure 9: Partial and complete empirical tail dependographs for temperatures of twenty one French continental cities over the period 1946-2000: (a) The nine largest empirical tail superset coefficients (b) the thirty largest (c) all of them.

## Concluding remarks and perspectives

We have introduced new measures to explore the asymptotic dependence structure in extreme value theory. These indices, called tail superset importance coefficients, are based on the Hoeffding-Sobol (ANOVA) decomposition of the stable tail dependence function (s.t.d.f.). They quantify the (proportion of) variance of the s.t.d.f. explained by all supersets of pairs of variables. In other words, for a given pair, the tail superset importance index informs not only on the asymptotic bivariate structure but also on any tail dependence structure that contains this pair.

In a  $d$ -variate setting, it is remarkable that such information is available at a quadratic cost with respect to the number of variables ( $d(d-1)/2$  pairs), compared to the exponential number of subsets of  $\{1, \dots, d\}$ . Furthermore, these coefficients can be visualized on a graph, called tail dependograph. On the other hand, one loses the powerful probabilistic interpretation (9) of the family of  $\theta_I$ 's coefficients but one gains a comparison via a global variance decomposition.

On the theoretical level, we derived several results for statistical inference. First, analytical expressions

are available when the s.t.d.f. is replaced by its empirical version. Second, we proved consistency of the tail superset importance coefficient estimates under usual assumptions ensuring uniform consistency for the empirical s.t.d.f. Asymptotic normality is also proved, with the unusual feature that the speed of convergence depends on the nullity of the (true) value:  $O(k^{-1/2})$  for a non-zero index, and  $O(k^{-1})$  for a null value.

We further investigated the behavior of the new measures, compared to other competitors based on standard extremal coefficients. We observed that some features are better detected by the new coefficients, such as asymmetry, in dependent models for extremes. On the other hand, first trials indicate that tail superset importance indices do not improve on testing asymptotic independence. Thus, they may be more useful to describe the dependence structure of extremes, via the tail dependograph.

Several questions could perhaps be answered with these new measures. First, the effective dimension of the asymptotic distribution under the assumption of max-domain of attraction. Second, a multivariate summary of the asymptotic dependence, similar to the multivariate tau for instance. Third, a real guide to construct nested model. Other topics related to the estimation of the tail superset importance coefficients could be handled. We restricted our study to the empirical estimator of the stable tail dependence function. Other choices are possible, as the asymptotically unbiased version of [Fougères et al. \(2015\)](#) or the M-estimator of [Einmahl et al. \(2016\)](#) for instance. Finally, the inference of the tail superset coefficients inside parametric models could be investigated in future research.

## Acknowledgements

The authors would like to thank Christian Genest (McGill University, Montréal, Canada) for fruitful discussions. The first author would also like to thank Roland Denis and Benoît Fabrigès (Institut Camille Jordan, Université de Lyon, France) whose discussions and workshop induced great progress in the codes associated to this paper.

## Proofs

*Details for Section 3.4.* We prove here the combinatorial formulas used in the definition of the  $\lambda$ -supergraph. First, starting from the identity

$$(1+x)^d = \sum_{I \subseteq \{1, \dots, d\}} x^{|I|},$$

and computing the derivative at 1, one gets

$$\sum_{I \subseteq \{1, \dots, d\}} |I| = d2^{d-1}.$$

Thus, by parameterizing the supersets  $I$  of a pair  $\{i, j\}$  by  $J = I \setminus \{i, j\}$ , we have

$$\sum_{I \supseteq \{i, j\}} |I| = \sum_{J \subseteq \{1, \dots, d-2\}} \{|J| + 2\} = (d-2)2^{d-3} + 2 \times 2^{d-2} = (d+2)2^{d-3}.$$

Consequently, the upper bound of the weight edges is given by  $\sum_{I \supseteq \{i, j\}} \{|I| - 1\} = d2^{d-3}$ .  $\square$

For the proofs, we need to provide first a technical preliminary lemma in sensitivity analysis.

**Lemma 5.1.** Let  $f$  be a tensor-product function  $f(\mathbf{u}) = \prod_{t=1}^d f_t(u_t)$ . Assume moreover the framework of Hoeffding-Sobol decomposition:  $u_1, \dots, u_d$  are viewed as independent random variables, and  $f(\mathbf{u})$  is square integrable. For any subset  $I \subseteq \{1, \dots, d\}$ , denote by  $f_I(\mathbf{u}_I)$  the term indexed by  $I$  in the Hoeffding-Sobol decomposition of  $f(\mathbf{u})$ . Finally, set  $m_t = \mathbb{E}[f_t(u_t)]$ . Then

$$f_I(\mathbf{u}_I) = \prod_{t \in I} \{f_t(u_t) - m_t\} \prod_{t \notin I} m_t$$

and its variance is

$$D_I(f) = \prod_{t \in I} \text{Var}(f_t(u_t)) \prod_{t \notin I} m_t^2.$$

Furthermore, for a given set  $I \subseteq \{1, \dots, d\}$ , the superset part of  $f(\mathbf{u})$ , i.e. the part in Hoeffding-Sobol decomposition that contains the variables  $u_t, t \in I$  is given by

$$f_I^{\text{super}}(\mathbf{u}) = \prod_{t \in I} \{f_t(u_t) - m_t\} \prod_{t \notin I} f_t(u_t) \quad (17)$$

and its variance is

$$D_I^{\text{super}}(f) = \prod_{t \in I} \text{Var}(f_t(u_t)) \prod_{t \notin I} \mathbb{E}(f_t(u_t)^2).$$

*Proof of Lemma 5.1.* For the first part of the lemma, following Sobol' (2003), observe that the Hoeffding-Sobol decomposition of  $f(\mathbf{u})$  is obtained by expanding the product

$$f(\mathbf{u}) = \prod_{t=1}^d \{(f_t(u_t) - m_t) + m_t\} = \sum_{I \subseteq \{1, \dots, d\}} f_I(\mathbf{u}_I),$$

with  $f_I(\mathbf{u}_I) = \prod_{t \in I} \{f_t(u_t) - m_t\} \prod_{t \notin I} m_t$ . Such  $f_I$ 's satisfy the condition  $\mathbb{E}[f_I(\mathbf{u}_I) | \mathbf{u}_J] = 0$  for each strict subset  $J \subsetneq I$  so that the result follows from unicity of the decomposition.

Let us now fix a set  $I \subseteq \{1, \dots, d\}$ . By expanding the product

$$f(\mathbf{u}) = \prod_{t \in I} \{(f_t(u_t) - m_t) + m_t\} \prod_{t \notin I} f_t(u_t),$$

the term  $\prod_{t \in I} (f_t(u_t) - m_t) \prod_{t \notin I} f_t(u_t)$  corresponds to all supersets  $J$  of  $I$  in the Hoeffding-Sobol decomposition of  $f(\mathbf{u})$ . Indeed the  $f_J$ 's are obtained through the same expansion mechanism by picking the first term  $(f_t(u_t) - m_t)$  in the parenthesis  $\{(f_t(u_t) - m_t) + m_t\}$  when  $t$  belongs to  $J \supseteq I$ . This can also be checked by using the analytical expression of  $f_J$  proved above:

$$\begin{aligned} f_I^{\text{super}}(\mathbf{u}) &= \sum_{J \supseteq I} f_J(\mathbf{u}_J) = \sum_{J \supseteq I} \prod_{t \in J} \{f_t(u_t) - m_t\} \prod_{t \notin J} m_t \\ &= \prod_{t \in I} (f_t(u_t) - m_t) \left\{ \sum_{J \supseteq I} \prod_{t \in J \setminus I} (f_t(u_t) - m_t) \prod_{t \notin J} m_t \right\} \\ &= \prod_{t \in I} (f_t(u_t) - m_t) \left\{ \sum_{K \subseteq \{1, \dots, d\} \setminus I} \prod_{t \in K} (f_t(u_t) - m_t) \prod_{t \notin K} m_t \right\} = \prod_{t \in I} (f_t(u_t) - m_t) \prod_{t \notin I} f_t(u_t). \end{aligned}$$

Finally the variance expressions are direct, remembering that each  $f_I$  is centered and all the  $u_t$  are assumed to be independent.  $\square$

*Proof of Proposition 4.1.* We have:

$$\hat{\ell}_{k,n}(\mathbf{u}) = \frac{1}{k} \sum_{s=1}^n \left( 1 - \prod_{t=1}^d \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\} \right) = \frac{n}{k} - \frac{1}{k} \sum_{s=1}^n \prod_{t=1}^d \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\}. \quad (18)$$

Let us now fix a pair  $\{i, j\}$ . The constant term  $n/k$  does not play a role in the superset importance. The second term is a sum of tensor-product functions. Thus, from Lemma 5.1, its superset part is obtained by centering the terms in the products which involve the considered pair. Namely, from (17), we get

$$\hat{\ell}_{k,n;\{i,j\}}^{\text{super}}(\mathbf{u}) = -\frac{1}{k} \sum_{s=1}^n \prod_{t=1}^d \left( \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\} - \mathbf{1}\{t \in \{i, j\}\} \mu_t \left( \tilde{R}_s^{(t)} \right) \right).$$

Observing that the term is centered, the variance reduces to the expectation of its square, that is

$$\begin{aligned} \left( \hat{\ell}_{k,n;\{i,j\}}^{\text{super}}(\mathbf{u}) \right)^2 &= \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \prod_{t=1}^d \left( \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\} - \mathbf{1}\{t \in \{i, j\}\} \mu_t \left( \tilde{R}_s^{(t)} \right) \right) \\ &\quad \times \left( \mathbf{1}\{u_t < \tilde{R}_{s'}^{(t)}\} - \mathbf{1}\{t \in \{i, j\}\} \mu_t \left( \tilde{R}_{s'}^{(t)} \right) \right). \end{aligned}$$

A direct computation of the expectation then gives the announced expression

$$D_{\{i,j\}}^{\text{super}}(\hat{\ell}_{k,n}) = \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \prod_{t=1}^d \left( \mu_t \left( \tilde{R}_s^{(t)} \wedge \tilde{R}_{s'}^{(t)} \right) - \mathbf{1}\{t \in \{i, j\}\} \mu_t \left( \tilde{R}_s^{(t)} \right) \mu_t \left( \tilde{R}_{s'}^{(t)} \right) \right).$$

To get the global variance, we write from (18)

$$\begin{aligned} D(\hat{\ell}_{k,n}) &= \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \text{Cov} \left( \prod_{t=1}^d \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\}, \prod_{t=1}^d \mathbf{1}\{u_t < \tilde{R}_{s'}^{(t)}\} \right) \\ &= \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \left\{ \mathbb{E} \left( \prod_{t=1}^d \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\} \mathbf{1}\{u_t < \tilde{R}_{s'}^{(t)}\} \right) \right. \\ &\quad \left. - \mathbb{E} \left( \prod_{t=1}^d \mathbf{1}\{u_t < \tilde{R}_s^{(t)}\} \right) \mathbb{E} \left( \prod_{t=1}^d \mathbf{1}\{u_t < \tilde{R}_{s'}^{(t)}\} \right) \right\}, \end{aligned}$$

which yields the result by using the independence of the  $u_t$ 's.  $\square$

*Proof of Theorem 4.2.* Lemma 5.1 can be applied to  $\hat{\ell}_{k,n}(\cdot) - n/k$  since it is a sum of tensor-product functions. The results lead directly to  $\hat{\ell}_{k,n;I}$  for any non empty subset  $I$  and the term  $n/k$  has to be added in order to recover the constant term in the decomposition.  $\square$

*Proof of Corollary 4.4.* Recall that the Hoeffding-Sobol decomposition relies on the recursive construction, written below in its empirical version,

$$\hat{\ell}_{k,n;I}(\mathbf{u}_I) = \mathbb{E}[\hat{\ell}_{k,n}(\mathbf{U}) | \mathbf{U}_I = \mathbf{u}_I] - \sum_{I' \subsetneq I} \hat{\ell}_{k,n;I'}(\mathbf{u}_{I'}).$$

The proof is by induction on  $I$ . Notice first that

$$|\hat{\ell}_{k,n;\emptyset} - \ell_\emptyset| = |\mathbb{E}[\hat{\ell}_{k,n}(\mathbf{U}) - \ell(\mathbf{U})]| \leq \sup_{\mathbf{u} \in [0,1]^d} |\hat{\ell}_{k,n}(\mathbf{u}) - \ell(\mathbf{u})| \xrightarrow{\mathbb{P}} 0.$$

Now, let  $I$  be a non empty subset of  $\{1, \dots, d\}$  and assume that for any subset  $I' \subsetneq I$

$$\sup_{\mathbf{u}_{I'} \in [0,1]^{|I'|}} |\hat{\ell}_{k,n;I'}(\mathbf{u}_{I'}) - \ell_{I'}(\mathbf{u}_{I'})| \xrightarrow{\mathbb{P}} 0.$$

Applying twice the recursive formula, one gets

$$\sup_{\mathbf{u}_I \in [0,1]^{|I|}} |\hat{\ell}_{k,n;I}(\mathbf{u}_I) - \ell_I(\mathbf{u}_I)| \leq \sup_{\mathbf{u}_I \in [0,1]^{|I|}} \left| \mathbb{E}[\hat{\ell}_{k,n}(\mathbf{U}) - \ell(\mathbf{U}) | \mathbf{U}_I = \mathbf{u}_I] \right| + \sum_{I' \subsetneq I} \sup_{\mathbf{u}_{I'} \in [0,1]^{|I'|}} |\hat{\ell}_{k,n;I'}(\mathbf{u}_{I'}) - \ell_{I'}(\mathbf{u}_{I'})|$$

that both converge to zero in probability. The consistency of the global variance or that of the variances associated to subsets are a direct consequence of what precedes. Nevertheless, note that the existence and the consistency of the empirical global sensitivity indices  $\hat{S}_{k,n;I}$  and  $\hat{S}_{k,n;I}^{\text{super}}$  rely on the fact that the empirical global variance  $D(\hat{\ell}_{k,n})$  and the limiting global variance  $D(\ell)$  never vanish in the extreme value theory setting.  $\square$

*Proof of Lemma 4.5.* The difference between this lemma and Theorem 7.2.2 of [de Haan and Ferreira \(2006\)](#) is very thin. It is concentrated on the specification of the intermediate sequence  $k$ . However, all the arguments apply here.  $\square$

*Proof of Corollary 4.6.* As soon as the Hoeffding-Sobol decomposition of the process  $Z_\ell + \lambda m$  has some sense, there is no difficulty to make  $\hat{\ell}_{k,n;I}$  inherit the asymptotic normality property from  $\hat{\ell}_{k,n}$ . It suffices to apply the asymptotic distributional expansion

$$\hat{\ell}_{k,n} \simeq^d \ell + \frac{Z_\ell}{\sqrt{k}} + \alpha(n/k)m + o(1/\sqrt{k}) + o(\alpha(n/k))$$

and make use of the recursive definition  $\hat{\ell}_{k,n;I}(\mathbf{u}_I) = \mathbb{E}[\hat{\ell}_{k,n}(\mathbf{U}) | \mathbf{U}_I = \mathbf{u}_I] - \sum_{I' \subsetneq I} \hat{\ell}_{k,n;I'}(\mathbf{u}_{I'})$ .  $\square$

*Proof of Proposition 4.7.* This result is obtained by the fact that any superset importance coefficient has an integral representation. Indeed, from (9) in [Liu and Owen \(2006\)](#), one knows that

$$D_I^{\text{super}}(\ell) = \frac{1}{2^{|I|}} \int \left( \sum_{J \subseteq I} (-1)^{|I \setminus J|} \ell(\mathbf{u}_J, \mathbf{v}_{-J}) \right)^2 d\mu_J(\mathbf{u}_J) d\mu_{-J}(\mathbf{v}_{-J}).$$

Applying this formula to the empirical version  $\hat{\ell}_{k,n}$  and using the expansion  $\hat{\ell}_{k,n} \simeq \ell + Z_\ell/\sqrt{k} + \alpha(n/k)m$ , provided by Lemma 4.5, allows to write

$$\begin{aligned} D_I^{\text{super}}(\hat{\ell}_{k,n}) &= D_I^{\text{super}}(\ell) + \frac{1}{k} D_I^{\text{super}}(Z_\ell) + \alpha^2(n/k) D_I^{\text{super}}(m) \\ &\quad + \frac{1}{2^{|I|} - 1} \int \Delta_I[\ell, Z_\ell/\sqrt{k}](\mathbf{u}_I, \mathbf{v}) + \Delta_I[\ell, \alpha(n/k)m](\mathbf{u}_I, \mathbf{v}) + \Delta_I[Z_\ell/\sqrt{k}, \alpha(n/k)m](\mathbf{u}_I, \mathbf{v}) \\ &\quad + o(1/\sqrt{k}) + o(\alpha^2(n/k)) + o(\alpha(n/k)/\sqrt{k}), \end{aligned}$$

where the integral is with respect to the measure  $d\mu_I(\mathbf{u}_I) d\mu(\mathbf{v})$ . The result follows from the asymptotic behavior of  $\alpha$  that is:  $\alpha(n/k) \sim \alpha/\sqrt{k}$ .  $\square$



## References

- J. J. Cai, A.-L. Fougères, and C. Mercadier. Environmental data: multivariate extreme value theory in practice. *Journal of Soc. Fr. Stat.*, 154(2):178–199, 2013.
- G. Chastaing, F. Gamboa, and C. Prieur. Generalized Hoeffding-Sobol decomposition for dependent variables - application to sensitivity analysis. *Electron. J. Statist.*, 6:2420–2448, 2012.
- J. Cornfield and J. W. Tukey. Average values of mean squares in factorials. *Ann. Math. Statist.*, 27(4):907–949, 12 1956.
- L. de Haan and A. Ferreira. *Extreme value theory. An introduction*. Springer Series in Operations Research and Financial Engineering. New York, NY: Springer., 2006.
- B. Efron and C. Stein. The jackknife estimate of variance. *The Annals of Statistics*, 9(3):586–596, 1981.
- J. H. J. Einmahl, A. Kiriliouk, A. Krajina, and J. Segers. An M-estimator of spatial tail dependence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78(1):275–298, 2016.
- A.-L. Fougères. *Multivariate Extremes, In: Extreme Values in Finance, Telecommunications, and the Environment*, ed. Finkenstädt, B. and Rootzén, H. Monographs on Statistics and Applied Probability 99, Chapter 7, pp. 373-388. Chapman and Hall/CRC, 2004.
- A.-L. Fougères, L. de Haan, and C. Mercadier. Bias correction in multivariate extremes. *The Annals of Statistics*, 43(2):903–934, 2015.
- J. Fruth, O. Roustant, and S. Kuhnt. Total interaction index: A variance-based sensitivity index for interaction screening. *Journal of Statistical Planning and Inference*, 147:212–223, 2014.
- J. Fruth, T. Muehlenstaedt, O. Roustant, and M. Jastrow. *fanovaGraph: Building Kriging Models from FANOVA Graphs*, 2015. R package version 1.4.8.
- W. Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics*, 19:293–325, 1948. ISSN 0003-4851.
- M. Hofert and K. Hornik. *qrmdata: Data Sets for Quantitative Risk Management Practice*, 2016. R package version 2016-01-03-1.
- G. Hooker. Discovering additive structure in black box functions. In *Proceedings of KDD 2004*, pages 575–580. ACM DL, 2004.
- X. Huang. Statistics of bivariate extremes. PhD Thesis, Erasmus University Rotterdam, Tinbergen Institute Research series No. 22, 1992.
- M. Kereszturi, J. Tawn, and Ph. Jonathan. Assessing extremal dependence of north sea storm severity. *Ocean Engineering*, 118:242 – 259, 2016.
- R. Liu and A. B. Owen. Estimating mean dimensionality of analysis of variance decompositions. *J. Amer. Statist. Assoc.*, 101(474):712–721, 2006. ISSN 0162-1459.
- L. Mhalla, V. Chavez-Demoulin, and Ph. Naveau. Non-linear models for extremal dependence. *Journal of Multivariate Analysis*, 159:49 – 66, 2017.

- T. Muehlenstaedt, O. Roustant, L. Carraro, and S. Kuhnt. Data-driven kriging models based on FANOVA-decomposition. *Statistics and Computing*, 22(3):723–738, 2012.
- Yue Peng and Wing Lon Ng. Analysing financial contagion and asymmetric market dependence with volatility indices via copulas. *Annals of Finance*, 8(1):49–74, Feb 2012.
- G. Pujol, B. Iooss, A. Janon, with contributions from, K. Boumhaout, S. Da Veiga, T. Delage, J. Fruth, L. Gilquin, J. Guillaume, L. Le Gratiet, P. Lemaitre, B. L. Nelson, F. Monari, R. Oomen, B. Ramos, O. Roustant, E. Song, J. Staum, T. Touati, and F. Weber. *sensitivity: Global Sensitivity Analysis of Model Outputs*, 2017. R package version 1.14.0.
- P. Ressel. Homogeneous distributions, and a spectral representation of classical mean values and stable tail dependence functions. *Journal of Multivariate Analysis*, 117:246 – 256, 2013.
- J. Segers. Max-stable models for multivariate extremes. *REVSTAT – Statistical Journal*, 10(1):61–82, 2012.
- R. L. Smith. Max-stable processes and spatial extremes. Dept. of Math., Univ. of Surrey, Guildford GU2 5XH, England, 1990.
- I. M. Sobol'. Sensitivity estimates for nonlinear mathematical models. *Math. Modeling Comput. Experiment*, 1(4):407–414 (1995), 1993. ISSN 1061-7590.
- I. M. Sobol'. Theorems and examples on high dimensional model representation. *Reliability Engineering and System Safety*, 79:187–193, 2003.
- A. G. Stephenson. evd: Extreme value distributions. *R News*, 2(2), June 2002.
- A. G. Stephenson. Simulating multivariate extreme value distributions of logistic type. *Extremes*, 6(1): 49–59, Mar 2003.
- J. A. Tawn. Modelling multivariate extreme value distributions. *Biometrika*, 77(2):245–253, 1990.
- A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.