DE GRUYTER Depend. Model. 2021; 9:179–198

Research Article Open Access

Cécile Mercadier* and Paul Ressel

Hoeffding–Sobol decomposition of homogeneous co-survival functions: from Choquet representation to extreme value theory application

https://doi.org/10.1515/demo-2021-0108
Received April 18, 2021; accepted August 26, 2021

Abstract: The paper investigates the Hoeffding–Sobol decomposition of homogeneous co-survival functions. For this class, the Choquet representation is transferred to the terms of the functional decomposition, and in addition to their individual variances, or to the superset combinations of those. The domain of integration in the resulting formulae is reduced in comparison with the already known expressions. When the function under study is the stable tail dependence function of a random vector, ranking these superset indices corresponds to clustering the components of the random vector with respect to their asymptotic dependence. Their Choquet representation is the main ingredient in deriving a sharp upper bound for the quantities involved in the tail dependograph, a graph in extreme value theory that summarizes asymptotic dependence.

Keywords: Hoeffding–Sobol decomposition, co-survival function, spectral representation, stable tail dependence function, multivariate extreme value modeling.

MSC: 26A48, 26B99, 44A30, 62G32, 62H05

1 Introduction

Let $f : [0, 1]^d \to \mathbb{R}$ be a function in $L^2([0, 1]^d, \lambda)$ where $\lambda = \prod_{i=1}^d \lambda_i$ is a product of probability measures on [0, 1]. One way to understand the structural form of the d-variables function f is to decompose it into functions of increasing complexity. This is precisely what allows the functional analysis of variance (FANOVA). It relies on the Hoeffding–Sobol decomposition

$$f(\mathbf{x}) = \sum_{u\subseteq\{1,\dots,d\}} f_u(\mathbf{x}) \tag{1}$$

where

$$f_u(\mathbf{x}) = \sum_{v\subseteq u} (-1)^{|u\setminus v|} \int f(\mathbf{x}) d\lambda_{-v}(\mathbf{x}) \tag{2}$$

for $d\lambda_u(\mathbf{x}) = \prod_{i\in u} d\lambda_i(x_i)$ and $-v = \{1, \dots, d\}\setminus v$. See [5, 18, 19]. The term f_u only depends on the components of $\mathbf{x}$ associated with u. The constant term $f_\varnothing$ is equal to $\int f d\lambda$ and the global variance is given by $\sigma^2 = \int (f - f_\varnothing)^2 d\lambda$. Set $\sigma_u^2 = \int f_u^2 d\lambda_u$ and $\sigma_\varnothing^2 = 0$. Then, from orthogonality arguments (see, for instance, [2]), the term f_u is centered (except for the empty set) and the FANOVA expression relies on the equality

$$\sigma^2 = \sum_{u\subseteq\{1,\dots,d\}} \sigma_u^2 \, .$$

***Corresponding Author: Cécile Mercadier:** Université de Lyon, Université Claude Bernard Lyon 1, Institut Camille Jordan, UMR CNRS 5208, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne cedex, France, E-mail: mercadier@math.univ-lyon1.fr
Paul Ressel: Kath. Universität Eichstätt-Ingolstadt, Ostenstraße 26-28, 85072 Eichstätt, Germany, E-mail: paul.ressel@ku.de

Interest in the individual variances σ_u^2, and more particularly their ratio to the total variance σ_u^2/σ^2, traces back to [18] and [12]. The current research problems in Global Sensitivity Analysis (GSA) are varied in nature. Our concern in this paper is not improvements for estimation, cost-saving, construction of surrogate models, or other practical but no less crucial aspects or perspectives ; see rather [11, 13] and references contained therein for an overall and recent assessment. The main goal here is to reveal simplified theoretical expressions for the quantity σ_u^2 within a specific class of functions. Knowing such quantities σ_u^2 allows to order the importance of the input variables $x_1, \dots, x_d$ with respect to the global variance of f, the function under study. Reducing the number of variables of interest in f is one of the main consequences of this hierarchical ranking.

In this paper we will concentrate on homogeneous co-survival functions. Classical examples to keep in mind are the power norms, defined for $t \geq 1$ by $\psi_t(\mathbf{x}) := \left(\sum_{i=1}^d x_i^t\right)^{1/t}$ for $\mathbf{x} \in [0, \infty[^d$. More generally, if there exists μ a non-negative Radon measure on $[0, \infty]^d \setminus \{\infty\}$ such that $\psi(\mathbf{x}) = \mu(\{\mathbf{y} \in [0, \infty]^d | \mathbf{y} \not\geq \mathbf{x}\})$ for all $\mathbf{x} \in \mathbb{R}_+^d$ then $\psi : \mathbb{R}_+^d \to \mathbb{R}_+$ is said to be a *co-survival function*. The class of co-survival functions additionally assumed to be homogeneous, is in a one-one correspondence (modulo value at $(1, \dots, 1)$) with probability measures ν on $C = \{\mathbf{w} = (w_1, \dots, w_d) \in [0, 1]^d | \max(w_1, \dots, w_d) = 1\}$. Indeed, the spectral representation

$$\psi(x_1, \dots, x_d) = \psi(1, \dots, 1) \int_C \max(x_1 w_1, \dots, x_d w_d) d\nu(\mathbf{w}) \tag{3}$$

is stated in [16, Theorem 2]. Some details are given in Section 2 to make the paper almost self-contained.

In extreme value theory, stable tail dependence functions (stdf), usually denoted by ℓ, play a central role to describe the asymptotic dependence between components of a random vector $\mathbf{X} = (X_1, \dots, X_d)$. Assuming the existence of a multivariate domain of attraction for the componentwise maxima of $\mathbf{X}$ is a classical starting point. This is equivalently written as

$$\lim_{t \to \infty} t \left(1 - F(F_1^{-1}(1 - x_1/t), \dots, F_d^{-1}(1 - x_d/t))\right) = \ell(x_1, \dots, x_d)$$

in terms of $F, F_1, \dots, F_d$ the cumulative distribution functions of $\mathbf{X}, X_1, \dots, X_d$. More details on multivariate extreme value theory can be found, e.g., in [1, 3, 4, 7, 10]. As pointed out in [16], the stdfs are particular cases of homogeneous co-survival functions. The corresponding probability measures ν in (3) must satisfy d constraints induced by the fact that a stdf equals 1 at unit vectors. A graph based on the Hoeffding–Sobol decomposition of a stdf, called the tail dependograph, has been introduced in [10]. It reveals the asymptotic dependence structure of the random vector $\mathbf{X}$ through the structural analysis of the function ℓ. Tail superset indices, which are the superset combination of individual variances, are of prime interest in the tail dependograph. Their pairwise values define the thickness of the edges.

The aim of this paper is twofold. On the one hand, we shall establish a simplified expression for the individual variances σ_u^2, as for their superset combinations, when the function under study is a homogeneous co-survival function. Their resulting Choquet representation thus provide new test cases for GSA. On the other hand, we will apply these results to stdfs so that upper bounds for the tail superset indices will be obtained. Proving this majorization initially motivated the current study.

The paper is organized as follows. We first investigate the class of homogeneous co-survival functions: in Section 2, the expression of the FANOVA effect ψ_u and the corresponding variance σ_u^2 are written as integrals of rank-one tensors (which are products of univariate functions in each of the input parameters, as defined by [8]). The numerical performance of our results is analyzed at the end of this part. As an application, the study focuses on stdfs in Section 3. The new expressions allow to derive some sharp upper bounds for the tail superset indices. All proofs are postponed to Section 4. Finally, the last lines summarize conclusions and references.

Notation. Let $\vee$ and $\wedge$ stand respectively for the maximum and the minimum. Set $x_+ = x \vee 0$. The indicator 1_A equals 1 on A and 0 on A^c. Set $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^d$. The vector $\mathbf{z}^u$ is the concatenation of z_i for $i \in u$ so that $(\mathbf{z}^u, \mathbf{x}^{-u}) = \sum_{i\in u} z_i\mathbf{e}_i + \sum_{i\notin u} x_i\mathbf{e}_i$ in the canonical basis $(\mathbf{e}_1, \dots, \mathbf{e}_d)$. Binary operations are understood componentwise, e.g. $\mathbf{x}\cdot\mathbf{w} = (x_1w_1, \dots, x_dw_d)$, $\mathbf{x}\vee\mathbf{w} = (x_1\vee w_1, \dots, x_d\vee w_d)$, $s/\mathbf{w} = (s/w_1, \dots, s/w_d)$ and $1_{s\geq\mathbf{x}_u} = \prod_{i\in u} 1_{s\geq x_i}$ for $s\in[0,1]$. Throughout the paper $[\mathbf{x},\infty]^c := \{\mathbf{y}\in[0,\infty]^d|\mathbf{y}\not\geq\mathbf{x}\}$. Let $\lambda = \prod_{i=1}^d \lambda_i$ be an arbitrary product of probability measures on $[0,1]$. For positive v and w, when s and t lie in $[0,1]$, let $K_i(w,v;s,t) = \lambda_i([0,(s/w)\wedge(t/v)\wedge 1]) = \int_0^1 1_{(s/w)\wedge(t/v)\geq x}\, d\lambda_i(x)$ and let $K_i(w;s)$ stand for $K_i(w,w;s,s)$. The notation ψ is used for a homogeneous co-survival function whereas ℓ represents a stdf.

2 FANOVA of homogeneous co-survival function

In this section, the functional decomposition is explored under a new setting by considering homogeneous co-survival functions. Before stating our main result, we give a description of the class under study. It is worth noticing that focusing on the unit hypercube $[0,1]^d$ is not restrictive by homogeneity assumption.

2.1 Choquet representation of homogeneous co-survival functions

Similar to distribution functions, also co-survival functions are essentially characterized by a special multivariate monotonicity property. First, we introduce a notation. Let $A_1, \dots, A_d$ be non-empty sets, $A = A_1\times\cdots\times A_d$, and let $f: A\to\mathbb{R}$ be any function. Then for $\mathbf{x},\mathbf{z}\in A$ we put

$$D_{\mathbf{z}}^{\mathbf{x}} f := \sum_{u\subseteq\{1,\dots,d\}} (-1)^{|u|} f(\mathbf{z}^u, \mathbf{x}^{-u})\ .$$

Moreover, for a non-empty subset $u\subsetneq\{1,\dots,d\}$ and for $\mathbf{x}^{-u}\in\prod_{j\in -u} A_j$, let us define on $\prod_{j\in u} A_j$

$$f(\cdot, \mathbf{x}^{-u})(\mathbf{z}^u) := f(\mathbf{z}^u, \mathbf{x}^{-u})\ .$$

If $A_j\subseteq\overline{\mathbb{R}}$ for all j, the function f is called $\mathbf{1}_d$–alternating[1] if $D_{\mathbf{z}}^{\mathbf{x}} f\leq 0$ for $\mathbf{x}\leq\mathbf{z}$ (both in A), and if this inequality also holds whenever some of the variables are fixed, for the function of the remaining variables, i.e. if for each non-empty subset $v\subsetneq\{1,\dots,d\}$, for each $\mathbf{y}\in\prod_{j\in -v} A_j$ and any $\mathbf{x}^v\leq\mathbf{z}^v$ both in $\prod_{j\in v} A_j$, we have

$$D_{\mathbf{z}^v}^{\mathbf{x}^v} f(\cdot, \mathbf{y})\leq 0\ .$$

See [17] for a detailed presentation of this concept.

Let $f: \mathbb{R}_+^d\to\mathbb{R}_+$ be the co-survival function of μ, a non-negative Radon measure on $[0,\infty]^d\setminus\{\infty\}$, i.e. for any $\mathbf{x}\in\mathbb{R}_+^d$

$$f(\mathbf{x}) = \mu([\mathbf{x},\infty]^c)\ .$$

If the reader is not familiar with Radon measures, one should only keep in mind that this assumption ensures that f is well defined and $f(\mathrm{x})$ finite for any $\mathrm{x}\in\mathbb{R}_+^d$.
By Theorem 3 in [16] one knows that it is equivalent to assuming f $\mathbf{1}_d$-alternating, left continuous, and $f(\mathbf{0}) = 0$. Moreover, for any $\mathbf{0}\leq\mathbf{x}<\mathbf{z}$ in $\mathbb{R}_+^d$

$$D_{\mathbf{z}}^{\mathbf{x}} f = -\mu([\mathbf{x},\mathbf{z}[)$$

by an application of the inclusion/exclusion principle. Now, if f is additionally assumed to be homogeneous, that is $f(t\mathbf{x}) = tf(\mathbf{x})$ for any positive t and vector $\mathbf{x}$ then the measure μ is homogeneous: $\mu(tA) = t\mu(A)$ for any positive t and measurable subset A (and reciprocally). Note that any homogeneous $\mathbf{1}_d$-alternating function $f: \mathbb{R}_+^d\to\mathbb{R}$ is automatically continuous, non-negative, with $f(\mathbf{0}) = 0$.

1 This notion was first introduced in [15] under the name *fully d-max-decreasing*.

An important example of a homogeneous measure is given by the image $\lambda_{\mathbf{w}}$ of the Lebesgue measure λ on $\mathbb{R}_+$ under the mapping $s \mapsto s/\mathbf{w}$, where $\mathbf{w} \in C$. The co-survival function of $\lambda_{\mathbf{w}}$ is then

$$\begin{aligned}\lambda_{\mathbf{w}}([\mathbf{x}, \infty]^c) &= \lambda(\{s \in \mathbb{R}_+ | s/\mathbf{w} \ngeq \mathbf{x}\}) \\ &= \lambda(\{s \in \mathbb{R}_+ | \exists i \le d, s/w_i < x_i\}) \\ &= \lambda(\{s \in \mathbb{R}_+ | s < \max_{i=1,\dots,d} (x_i w_i)\}) \\ &= \max(\mathbf{x} \cdot \mathbf{w}) \,.\end{aligned}$$

These functions will play a decisive role in the following, since they are the "building stones" of all homogeneous co-survival functions. More precisely, consider the set of all normalized functions discussed above

$$K := \{\psi : \mathbb{R}_+^d \to \mathbb{R} | \psi \text{ is } \mathbf{1}_d\text{-alternating, homogeneous and } \psi(\mathbf{1}) = 1\} \,.$$

Then K is obviously convex and compact (with respect to pointwise convergence). It turns out that K is even a simplex, with

$$\{\mathbf{x} \mapsto \max(\mathbf{x} \cdot \mathbf{w}) | \mathbf{w} \in C\} = \mathrm{ex}(K)$$

as its set of extreme points, and this set is closed (so compact as well) ; see [16, Theorem 4 (ii)]. In other words, K is a so-called Bauer simplex, i.e. for each $\psi \in K$ the representing probability measure on $\mathrm{ex}(K)$ guaranteed by Krein–Milman's theorem, is unique. The resulting integral representation is also called Choquet representation. So, for each $\mathbf{1}_d$-alternating and homogeneous ψ on $\mathbb{R}_+^d$, $\psi \not\equiv 0$, there is a unique probability measure ν on C such that

$$\psi(\mathbf{x}) = \psi(\mathbf{1}) \int_C \max(\mathbf{x} \cdot \mathbf{w}) d\nu(\mathbf{w}), \quad \mathbf{x} \in \mathbb{R}_+^d \,.$$

It is easily seen that ψ is the co-survival function of the measure $\mu := \psi(\mathbf{1}) \int_C \lambda_{\mathbf{w}} d\nu(\mathbf{w})$.

2.2 Expression of Sobol effects and associated variances

The main result of this paper is stated below. It says that Sobol effects ψ_u (as their variances) have rather simpler expressions in comparison with (2) when ψ is a homogeneous co-survival function. Indeed, they are expressed as integrals on $C \times [0, 1]$ of rank-one functions. Recall that $\mathbf{1} = (1, \dots, 1)$ in $\mathbb{R}^d$.

Theorem 1. *Let ψ be a homogeneous co-survival function (3) associated with a spectral probability measure ν on C. Then, the term ψ_u in the Hoeffding–Sobol decomposition with respect to λ satisfies on* $[0, 1]^d$

$$\psi_u(\mathbf{x}) = -\psi(\mathbf{1}) \int_C \left\{ \int_0^1 \prod_{i \in u} \left(1_{s \ge x_i w_i} - K_i(w_i; s)\right) \prod_{i \notin u} K_i(w_i; s)\, ds \right\} d\nu(\mathbf{w})$$

for any non-empty subset u of $\{1, \dots, d\}$ *and*

$$\psi_\varnothing = \psi(\mathbf{1}) \left(1 - \int_C \left\{ \int_0^1 \prod_{i=1}^d K_i(w_i; s) ds \right\} d\nu(\mathbf{w}) \right) \,.$$

Its corresponding variance $\int \psi_u^2$ has the following expression

$$\sigma_u^2 = \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{i \notin u} K_i(w_i; s) K_i(v_i; t) \prod_{i \in u} \left(K_i(w_i, v_i; s, t) - K_i(w_i; s) K_i(v_i; t)\right) \,.$$

Furthermore,

$$\sigma^2 = -(\psi(\mathbf{1}) - \psi_\varnothing)^2 + \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{i=1}^d K_i(w_i, v_i; s, t) \,.$$

The link of Sobol effects ψ_u and their variances σ_u^2 with the spectral measure ν has been made explicit. The main ingredient for proving the previous theorem is to remark that the spectral representation of ψ can be written as an integral of rank-one tensors. Then, all Sobol effects ψ_u and corresponding variances σ_u^2 inherit the same form by application of the Fubini–Tonelli theorem. As can be seen through Formula (2), the variance σ_u^2 is usually computed as an alternating combination of cumulated variances. It thus suffers from accumulation of estimation error, overall as d becomes larger. Theorem 1 offers a setting where the numerical complexity of σ_u^2 is the same as that of σ^2 or other well-known quantities discussed in Subsection 2.3.

Example 1. *If the measure λ corresponds to the product of Lebesgue measures $d\lambda(\mathbf{x}) = dx_1 \cdots dx_d$ then $K_i(w, v; s, t) = (s/w) \wedge (t/v) \wedge 1$. Under this measure λ, consider $\psi(\mathbf{x}) = \max(\mathbf{x})$, a particular extreme point. By Theorem 1, with the probability measure $\nu = \delta_\mathbf{1}$ on C, one obtains $\psi_\varnothing = d/(d+1)$, $\sigma^2 = d/((d+1)^2(d+2))$,*

$$\psi_u(\mathbf{x}) = -\int_0^1 \prod_{i\in u} (1_{s\geq x_i} - s)\, s^{d-|u|}\, ds$$

and

$$\sigma_u^2 = 2\int_0^1 t^{d-|u|}(1-t)^{|u|}\left\{\int_0^t s^d ds\right\} dt = \frac{2(2d-|u|+1)!\,|u|!}{(d+1)(2d+2)!} .$$

Example 2. *Consider an extreme point of the convex and compact set K (mentioned in Subsection 2.1), precisely $\psi(\mathbf{x}) = \max(\mathbf{x}\cdot\mathbf{w})$ with $\mathbf{w} \in C$. It is worth noticing that Theorem 1 furnishes the expressions of the variances σ^2 and σ_u^2 as integrals on $[0,1]^2$ of a product of d univariate functions. In comparison with their original definitions, already mentioned in the introduction, this provides an important gain: The number of integrals is reduced (it is no longer an alternating sum) and the domain of integration is smaller. Under a precise value of $\mathbf{w}$, the calculations would give exact expressions after very tedious efforts. One could numerically approximate them by Monte-Carlo procedures on $[0,1]^2$ instead.*
With λ as the two dimensional Lebesgue measure, we focus here on these extreme points in the bivariate setting. For $\mathbf{w} = (w, 1)$, we obtain

$$\psi_\varnothing = 1/2 + w^2/6 \quad \text{and} \quad \sigma^2 = 1/12 - w^2/6 - w^4/36 + w^3/6 ,$$

with the following decomposition of σ^2

$$\sigma^2_{\{1\}} = w^4/45, \quad \sigma^2_{\{2\}} = -w^2/6 + 2w^3/15 - w^4/36 + 1/12, \quad \sigma^2_{\{1,2\}} = w^3/30 - w^4/45 .$$

2.3 Consequences for cumulated variances

It turns out that several combinations of variances are of prime interest in order to characterize the importance of a subset u of variables. Justifications can be found in [9, 18] in the case of

$$\mathcal{I}_u^2 = \sum_{v\subseteq u}\sigma_v^2 = \int_{[0,1]^{2d-|u|}} \psi(\mathbf{x}^u, \mathbf{x}^{-u})\psi(\mathbf{x}^u, \mathbf{z}^{-u})d\lambda(\mathbf{x})d\lambda_{-u}(\mathbf{z}) - \psi_\varnothing^2 \tag{4}$$

and

$$\overline{\tau}_u^2 = \sum_{v\cap u\neq\varnothing}\sigma_v^2 = \sigma^2 + \psi_\varnothing^2 - \int_{[0,1]^{d+|u|}} \psi(\mathbf{x}^u, \mathbf{x}^{-u})\psi(\mathbf{z}^u, \mathbf{x}^{-u})d\lambda(\mathbf{x})d\lambda_u(\mathbf{z}) . \tag{5}$$

We see immediately that $0 \leq \mathcal{I}_u^2 \leq \overline{\tau}_u^2 \leq \sigma^2$ and $\mathcal{I}_u^2 + \overline{\tau}_{-u}^2 = \sigma^2$. Finally, [6] examined the meaning of the sum over the supersets of u

$$\Upsilon_u^2 = \sum_{v\supseteq u}\sigma_v^2 . \tag{6}$$

Ranking based on the superset quantities Υ_u^2 takes into account the importance of $\mathbf{x}^u$ but additionnaly that of any vector containing these $|u|$ variables. Formulae depending on the spectral measure are now derived for these three types of cumulated variances. The next corollary asserts that they are also written as integrals of rank-one tensors.

Corollary 1. *Let ψ be a homogeneous co-survival function (3) associated with a spectral probability measure ν on C. Then,*

$$\mathcal{I}_u^2 = -(\psi(\mathbf{1}) - \psi_\varnothing)^2 + \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{i \in u} K_i(w_i, v_i; s, t) \prod_{i \notin u} K_i(w_i; s) K_i(v_i; t)$$

$$\begin{aligned}\overline{\tau}_u^2 = \psi(\mathbf{1})(2\psi_\varnothing - \psi(\mathbf{1})) + \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \\ \left(\prod_{i \in u} K_i(w_i, v_i; s, t) - \prod_{i \in u} K_i(w_i; s) K_i(v_i; t) \right) \prod_{i \notin u} K_i(w_i, v_i; s, t)\end{aligned}$$

$$\Upsilon_u^2 = \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{i \notin u} K_i(w_i, v_i; s, t) \prod_{i \in u} \left(K_i(w_i, v_i; s, t) - K_i(w_i; s) K_i(v_i; t) \right)$$

The Choquet representation of Υ_u^2 will play a crucial role in the proof of the upper bound stated in the extreme value theory setting at the end of the paper.

Example 3. *Consider again Example 1. One obtains easily*

$$\mathcal{I}_u^2 = 2 \int_0^1 t^{d-|u|} \left\{ \int_0^t s^d ds \right\} dt - (1 - \psi_\varnothing)^2 = \frac{|u|}{(d+1)^2 (2d - |u| + 2)} ,$$

$$\overline{\tau}_u^2 = \sigma^2 + \psi_\varnothing^2 - 2 \int_0^1 t^{|u|} \left\{ \int_0^t s^d ds \right\} dt = \frac{d}{d+2} - \frac{2}{(d+1)(d+|u|+2)} ,$$

and

$$\Upsilon_u^2 = 2 \int_0^1 (1-t)^{|u|} \left\{ \int_0^t s^d ds \right\} dt = \frac{2\, d!\, |u|!}{(d+|u|+2)!} .$$

In the opinion of the authors the current example (as its first part Example 1) looks promising for being a convenient test function. It provides a simple but non trivial function which has known individual variances as well as cumulated and global ones, for any dimension d.

In [9, Theorem 1] the following identity is shown

$$\Upsilon_u^2 = 2^{-|u|} \int_{[0,1]^{d+|u|}} (D_{\mathbf{z}^u}^{\mathbf{x}^u} \psi(\cdot, \mathbf{z}^{-u}))^2 d\mathbf{x}^u d\mathbf{z} \tag{7}$$

where $\psi(\cdot, \mathbf{z}^{-u}) : [0,1]^u \to \mathbb{R}$ is defined by $\psi(\cdot, \mathbf{z}^{-u})(\mathbf{x}^u) := \psi(\mathbf{x}^u, \mathbf{z}^{-u})$. The gain of the expression of Υ_u^2 claimed in Corollary 1 can be questioned with regard to the dimension of the domain of integration. Similar comments hold for $\mathcal{I}_u^2$ and τ_u^2 with reference formulae (4) and (5). But, it does not exist a direct formula of σ_u^2, i.e. based on ψ, except from inversion of (6) for instance. It yields

$$\sigma_u^2 = \sum_{v \supseteq u} (-1)^{|v \setminus u|} \Upsilon_v^2 = \sum_{v \supseteq u} (-1)^{|v \setminus u|} 2^{-|v|} \int_{[0,1]^{d+|v|}} (D_{\mathbf{z}^v}^{\mathbf{x}^v} \psi(\cdot, \mathbf{z}^{-v}))^2 d\mathbf{x}^v d\mathbf{z} . \tag{8}$$

Comparing the already known formula (8) with our result associated with σ_u^2 in Theorem 1 makes the interest of our expressions more obvious. In Theorem 1 indeed, it is no longer expressed as an alternating sum of integrals. We have reduced the dimension of integration. Nevertheless, to be also numerically convincing, a wide comparison between the estimation of σ_u^2 derived from (8) and from Theorem 1 will now be offered. The same comparison is done for the estimation of Υ_u^2, Formula (6) competing with the one from Corollary 1.

2.4 Numerical illustrations

For the sake of simplicity, we assume here that the distribution of the entries are known and fixed as uniforms. Our goal is to compare the effectiveness of the new formulae, obtained for homogeneous co-survival functions, with the already known and general ones. Both are integrations approximated by Monte-Carlo procedures, but neither the domain of integration nor the complexity of the integrand are the same. Our choices must assess impartiality. One possibility is to compare the estimation obtained after a given common executing time. However, this depends strongly on the way the integrands are coded. We thus decide to fix the Monte-Carlo size N on the unit interval.

We will first restrict ourselves to the case of the max function

$$\psi(\mathbf{x}) = \max(\mathbf{x})$$

for which the exact values are known (see Examples 1 and 3). This corresponds to a one-point measure. The comparison will therefore be broken down according to the value of d and the size of u with respect to d. Completely arbitrarily we set the following values $d = 5$ or $d = 10$; then $u = \{1, 2\}$ or $u = \{1, \dots, d\}$. Both σ_u^2 and Υ_u^2 will be estimated. The measures are based on the absolute mean error obtained over n replicates and defined as

$$AME := \frac{1}{n}\sum_{i=1}^{n} |\hat{\theta}_{i,N} - \theta_0|$$

where θ_0 is the true value and $\hat{\theta}_{i,N}$ is the i-th estimate. The number of replicates here is $n = 50$.

Table 1: AME for the estimation of σ_u^2 when $\psi(\mathbf{x}) = \max(\mathbf{x})$. Missing value – refers to exceeding the time limit.

	$d = 5$		$d = 10$	
	$u = \{1, 2\}$	$u = \{1, \dots, d\}$	$u = \{1, 2\}$	$u = \{1, \dots, d\}$
Formula (8) $N = 100$	48.25×10^{-5}	2.6×10^{-5}	17.52×10^{-5}	24.86×10^{-9}
Theorem 1 $N = 100$	8.34×10^{-5}	1.22×10^{-5}	1.19×10^{-5}	6.22×10^{-9}
Formula (8) $N = 10,000$	–	–	–	–
Theorem 1 $N = 10,000$	0.86×10^{-5}	0.11×10^{-5}	0.1×10^{-5}	0.75×10^{-9}

The level of accuracy is the same on each column of Table 1 in order to facilitate the comparison. Two values for N have been handled: $N = 100$ and $N = 10,000$. However, for the largest value, the time limit has been reached using the already known formula.

Table 2: AME for the estimation of Υ_u^2 when $\psi(\mathbf{x}) = \max(\mathbf{x})$. Missing value – refers to exceeding the time limit.

	$d = 5$		$d = 10$	
	$u = \{1, 2\}$	$u = \{1, \dots, d\}$	$u = \{1, 2\}$	$u = \{1, \dots, d\}$
Formula (7) $N = 1000$	12.15×10^{-5}	79.41×10^{-5}	31.18×10^{-6}	159×10^{-10}
Corollary 1 $N = 1000$	7.15×10^{-5}	0.40×10^{-5}	14.71×10^{-6}	24.39×10^{-10}
Formula (7) $N = 10,000$	5.43×10^{-5}	0.26×10^{-5}	12.07×10^{-6}	–
Corollary 1 $N = 10,000$	1.99×10^{-5}	0.11×10^{-5}	3.68×10^{-6}	7.53×10^{-10}

In Table 2 only one estimation has not been obtained because of time exceedance. Again, the level of accuracy is the same on each column to facilitate the comparison.

Let us now consider another homogeneous co-survival function associated with a discrete probability measure $\nu = \sum_{k=1}^{m} p_k \delta_{\mathbf{w}_k}$ where each $\mathbf{w}_k$ lies in C and $p_1 + \ldots + p_m = 1$ so that

$$\psi(\mathbf{x}) = \sum_{k=1}^{m} p_k \max(\mathbf{x} \cdot \mathbf{w}_k) \ . \tag{9}$$

Fix arbitrarily $m = 15$ and $d = 5$. The weights, chosen at random, are

$$(p_1, \ldots, p_m) = (0.04, 0.08, 0.12, 0.05, 0.02, 0.10, 0.11, 0.01, 0.12, 0.13, 0.06, 0.03, 0.10, 0.01, 0.02)$$

and the associated locations $(\mathbf{w}_1, \ldots, \mathbf{w}_m)$ are

$$\begin{pmatrix} 0.11 & 1.00 & 0.52 & 0.21 & 0.38 & 1.00 & 1.00 & 0.36 & 1.00 & 0.18 & 0.18 & 0.20 & 0.17 & 0.02 & 0.31 \\ 0.62 & 0.81 & 0.59 & 0.52 & 1.00 & 0.56 & 0.59 & 0.08 & 0.15 & 0.10 & 0.35 & 1.00 & 0.56 & 0.43 & 1.00 \\ 0.67 & 0.84 & 1.00 & 0.24 & 0.43 & 0.69 & 0.12 & 0.20 & 0.09 & 0.71 & 0.62 & 0.31 & 1.00 & 0.37 & 0.04 \\ 1.00 & 0.65 & 0.64 & 0.41 & 0.76 & 0.74 & 0.57 & 1.00 & 0.49 & 1.00 & 1.00 & 0.54 & 0.42 & 1.00 & 0.44 \\ 0.32 & 0.37 & 0.03 & 1.00 & 0.02 & 0.11 & 0.50 & 0.70 & 0.18 & 0.16 & 0.75 & 0.03 & 0.11 & 0.18 & 0.29 \end{pmatrix} .$$

Since the true values are not easily computable, we only provide a graphical comparison of the resulting boxplots obtained from $n = 50$ repetitions.

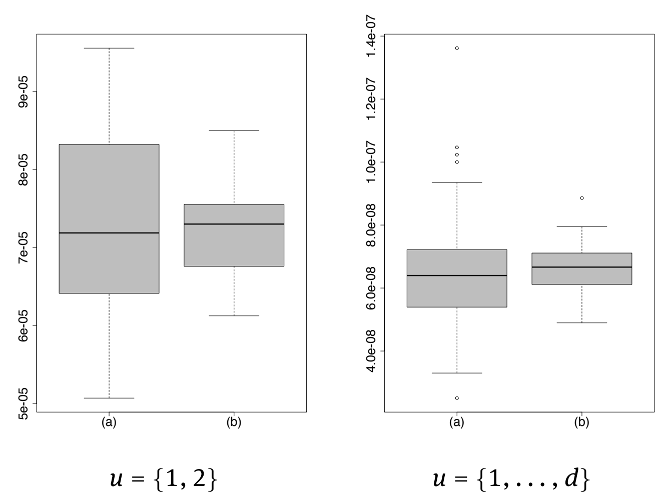

Figure 1: Estimation of σ_u^2 when ψ is given by (9) for $u = \{1, 2\}$ on the left panel and $u = \{1, \ldots, d\}$ on the right. The boxplot (a) is associated to the well-known Formula (8) whereas (b) refers to the new one stated in Theorem 1.

As expected, this numerical study shows that the estimation from the new formulae is more accurate. This is nothing more than the illustration of the domain of integration being reduced. The reader should be aware that recent studies in GSA provided new methods compared to the classical Monte-Carlo procedure. Going further with a comparison based on pick-freeze method or any other refinement would clearly exceed our ambitions in this paper.

3 Statistical applications in extreme value theory

In the following, we focus on the Hoeffding–Sobol representation of a stable tail dependence function (stdf). A homogeneous co-survival function ℓ is a stdf iff $\ell(\mathbf{e}_1) = \ldots = \ell(\mathbf{e}_d) = 1$ i.e., it is associated with a probability

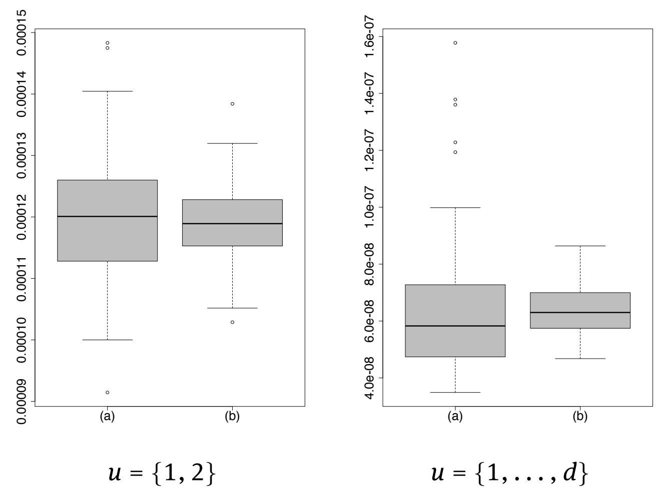

Figure 2: Estimation of Υ_u^2 when ψ is given by (9) for $u = \{1, 2\}$ on the left panel and $u = \{1, \ldots, d\}$ on the right. The boxplot (a) is associated to the well-known Formula (7) whereas (b) refers to the new one stated in Corollary 1.

measure ν on C satisfying

$$\int_C w_i d\nu(\mathbf{w}) = 1/\ell(\mathbf{1}) \qquad \forall\, i = 1, \ldots, d\ .$$

Denote by μ the measure such that $\ell(\mathbf{x}) = \mu([\mathbf{x}, \infty]^c)$. This measure μ is closely related to the so-called exponent measure $\mu_\star$ introduced in [14, Section 5.4.1] for instance. In fact, for any $\mathbf{x} \in \mathbb{R}_+^d$

$$\mu([\mathbf{x}, \infty]^c) = \mu_\star([0, 1/\mathbf{x}]^c)\ .$$

This means that $\mu_\star$ is the image of μ under $\mathbf{x} \mapsto 1/\mathbf{x}$, so that μ is directly homogeneous (as is ℓ) when $\mu_\star$ is inversely homogeneous: $\mu_\star(tA) = t^{-1}\mu_\star(A)$ for any positive t and any measurable set A of $[0, \infty]^d \setminus \{\mathbf{0}\}$. Whereas the characterization of stdfs was shown relatively late [16, Theorem 6], their integral representation was known long before: it goes back essentially to [4]. Most of the use of their integral representation has been done under the L_1 or L_2-norm on $\mathbb{R}_+^d$. But as emphasized by de Haan and Resnick, it is an arbitrary choice. As seen in Section 2, the extreme points of K (functions $\mathbf{x} \mapsto \max(\mathbf{x} \cdot \mathbf{w})$ for $\mathbf{w} \in C$) combined with the max-norm were natural choices here.

The main objective of this section is to analyze the theoretical aspect of the functional decompositon for stdfs with respect to the Lebesgue measure $d\lambda(\mathbf{x}) = dx_1 \ldots dx_d$. As mentioned in the introduction, this idea has been introduced in [10] but the focus was on the meaning of Υ_u^2 in this context, named as tail superset indices, and on their estimation. To illustrate their importance in multivariate extreme value modeling, let us focus for instance on the comparison $\Upsilon_{\{i,j\}}^2(\ell) < \Upsilon_{\{h,k\}}^2(\ell)$. This means that *the asymptotic dependence between components X_i and X_j themselves added to the asymptotic dependence between the pair (X_i, X_j) and the $d-2$ remaining variables* is weaker than its equivalent in h, k. Reducing the dimension of the asymptotic dependence structure consists in selecting subsets u according to their tail superset indices Υ_u^2.

Below, we first obtain a simplified expression for these indices by application of Corollary 1 to ℓ. Then, we deduce an upper bound for the tail superset indices. The section is ended by a short discussion.

3.1 Tail superset indices

The tail dependograph introduced in [10] starts from a non-oriented graph whose vertices represent components of the random vector $\mathbf{X}$ in the domain of attraction of ℓ. The edge between i and j is drawn proportionally to the pairwise superset indices $\Upsilon_{\{i,j\}}^2$ of ℓ. This index measures the strength of asymptotic dependence between the components X_i and X_j, not only in their associated bivariate model (X_i, X_j), but in the complete model **X**. A thick line reveals a strong asymptotic dependence between corresponding components, whereas at the opposite, such index vanishes when the asymptotic dependence is null. The present paper thus offers a theoretical expression of the tail dependograph indices as

$$\Upsilon^2_{\{i,j\}}(\ell) = \ell(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{k\neq i,j} \frac{s}{w_k} \wedge \frac{t}{v_k} \wedge 1$$
$$\left(\frac{s}{w_i} \wedge \frac{t}{v_i} \wedge 1 - \left\{\frac{s}{w_i} \wedge 1\right\} \cdot \left\{\frac{t}{v_i} \wedge 1\right\}\right) \left(\frac{s}{w_j} \wedge \frac{t}{v_j} \wedge 1 - \left\{\frac{s}{w_j} \wedge 1\right\} \cdot \left\{\frac{t}{v_j} \wedge 1\right\}\right) .$$

Pairwise indices are perhaps the most important since their value on a graph is easily represented by the thickness of a segment. However, more general indices can be defined and an application of the previous section also provides the representation of $\Upsilon^2_u(\ell)$ as follows

$$\ell(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{i\notin u} \frac{s}{w_i} \wedge \frac{t}{v_i} \wedge 1 \prod_{i\in u} \left(\frac{s}{w_i} \wedge \frac{t}{v_i} \wedge 1 - \left\{\frac{s}{w_i} \wedge 1\right\} \cdot \left\{\frac{t}{v_i} \wedge 1\right\}\right) .$$

Examples. The asymptotic independence occurs when $\ell^+(\mathbf{x}) := \sum_{i=1}^d x_i$ so that $\ell^+(\mathbf{1}) = d$ and $\nu = (\sum_{i=1}^d \delta_{\mathbf{e}_i})/d$. All the terms in the integrand of $\Upsilon^2_{\{i,j\}}(\ell^+)$ cancel since at least the term depending on i or on j (or both) will be reduced to $(1-1)$. As a consequence $\Upsilon^2_u(\ell^+) = 0$ as soon as $|u| \geq 2$.

The asymptotic complete dependence corresponds to $\ell^\vee(\mathbf{x}) := \max(\mathbf{x})$ so that $\ell^\vee(\mathbf{1}) = 1$ and $\nu = \delta_{\mathbf{1}}$. All indices of interest $\Upsilon^2_{\{i,j\}}(\ell^\vee)$ are equal to

$$\Upsilon^2_{\{i,j\}}(\ell^\vee) = \int_0^1 ds \int_0^1 dt (s\wedge t)^{d-2} (s \wedge t - s \cdot t)^2 = \frac{2}{d+1} \int_0^1 (1-t)^2 t^{d+1} dt = \frac{4}{(d+1)(d+2)(d+3)(d+4)} .$$

For in between strengths of asymptotic dependence, one can use logistic extreme value models. Symmetric versions

$$\ell(\mathbf{x}) = \left(x_1^{1/r} + \ldots + x_d^{1/r}\right)^r$$

for $r \in (0,1)$ are obtained with $\ell(\mathbf{1}) = d^r$ and

$$\nu(d\mathbf{w}) = \sum_{i=1}^d q_i(\mathbf{w}_{-i}) 1_{w_i=1, \mathbf{w}_{-i}\in[0,1]^{d-1}} d\mathbf{w}$$

where for instance

$$q_d(\mathbf{w}_{-d}) = c_r \left(\sum_{i=1}^{d-1} w_i^{-1/r} + 1\right)^{r-d} \left(\prod_{i=1}^{d-1} w_i\right)^{-(1/r+1)}$$

as pointed out in [16]. Unfortunately, the expressions obtained in the present paper do not allow a real simplification under such models.

3.2 Upper bounds for tail superset indices

Some simple computations allow to obtain the following lower and upper bounds.

Lemma 1. *Let ℓ be a d-variate stable tail dependence function. Then,*

$$\frac{d}{d+1} \leq \ell_\varnothing \leq \frac{d}{2}$$

and

$$\frac{d}{d+2} \leq \sigma^2(\ell) + \ell^2_\varnothing \leq \frac{d}{12} + \frac{d^2}{4} ,$$

implying

$$\sigma^2(\ell) \le \frac{d\,(3d^3 + 7d^2 - 7d + 1)}{12(d+1)^2}\,.$$

Moreover, for any subset u of $\{1,\dots,d\}$, *set* $\ell^{\vee,u}(\mathbf{x}) = \max_{i\in u} x_i$. *Then,*

$$\mathcal{I}_u^2(\ell^{\vee,u}) + (\ell_\varnothing^{\vee,u})^2 \le \mathcal{I}_u^2(\ell) + \ell_\varnothing^2 \le \frac{d^2}{4} + \frac{|u|}{12}\,.$$

The lower bound is given by

$$\mathcal{I}_u^2(\ell^{\vee,u}) + (\ell_\varnothing^{\vee,u})^2 = \frac{\sigma^2(\ell^{\vee,u}) + (\ell_\varnothing^{\vee,u})^2}{(d-|u|+1)^2} + \frac{2\ell_\varnothing^{\vee,u}(d-|u|)}{d-|u|+1} + \frac{(d-|u|)^2}{(d-|u|+1)^2}$$

where

$$\ell_\varnothing^{\vee,u} = \frac{|u|}{|u|+1} \quad \textit{and} \quad \sigma^2(\ell^{\vee,u}) = \frac{|u|}{(|u|+1)^2(|u|+2)}\,.$$

Hence, one can derive lower and upper bounds for $\mathcal{I}_u^2$ of a stable tail dependence function in any dimension, and for any size of the subset u, whenever $\ell_\varnothing$ is controlled. However, it doesn't provide a very simple way to deduce bounds for σ_u^2 or, more interestingly in the tail dependograph context, for Υ_u^2. The following result answers this question.

Theorem 2. *Let* ℓ *be a d-variate stable tail dependence function. Then, for any non-empty* $u \subseteq \{1,\dots,d\}$

$$\Upsilon_u^2(\ell) \le \Upsilon_u^2(\ell^{\vee,u}) = \frac{2(|u|!)^2}{(2|u|+2)!}$$

where $\ell^{\vee,u}(\mathbf{x}^u) = \max_{i\in u} x_i$.
If ℓ *is a d-variate stdf with equality* $\Upsilon_u^2(\ell) = \Upsilon_u^2(\ell^{\vee,u})$ *for a given* $\varnothing \ne u \subseteq \{1,\dots,d\}$, *then its projection on the variables* $\mathbf{x}^u$ *is equal to* $\ell(\mathbf{x}^u, \mathbf{0}^{-u}) = \ell^{\vee,u}(\mathbf{x}^u) = \max_{i\in u} x_i$.
In particular, if ℓ *is a d-variate stdf with equality* $\Upsilon_{\{1,\dots,d\}}^2(\ell) = \frac{2(d!)^2}{(2d+2)!}$ *then* $\ell(\mathbf{x}) = \max(\mathbf{x})$.

The proof is postponed to Section 4. However, note that it relies on the following preliminary result.

Proposition 1. *Let* $f : \mathbb{R}_+^d \to \mathbb{R}$ *be* $\mathbf{1}_d$*-alternating and let u be a non-empty subset of* $\{1,\dots,d\}$. *Then,*

$$\Upsilon_u^2(f) \le \Upsilon_u^2(f^{[u]})$$

with $f^{[u]}(\mathbf{z}^u) := f(\mathbf{z}^u, \mathbf{0}^{-u})$.

The authors conjectured the sharp upper bound in Theorem 2 a long time ago but the rigorous proof was only made possible after transferring the Choquet representation of the function ℓ to its indices as investigated in Section 2. The optimization problem dealt with in Theorem 2 might be looked at in the broader perspective of maximizing a convex functional over a compact convex set (which need not be a simplex). Bauer's maximum principle ensures that the maximal value is attained in an extreme point, in our case in an extreme stable tail dependence function. It does however give no hint to localize such a point nor to its uniqueness. Our statement in Theorem 2 answers completely the question: it asserts the existence, the uniqueness and the location (and so finds the maximal value) of the maximization problem.
The following statement is included in the proof of Theorem 2.

Corollary 2. *Let* ψ *be a homogeneous co-survival function (3) associated with a spectral probability measure* ν *on C. Then,*

$$\Upsilon_{\{1,\dots,d\}}^2(\psi) \le \psi(\mathbf{1})^2 \left(\int_C \prod_{i=1}^d w_i^{1/d}\, d\nu(\mathbf{w})\right)^2 \frac{2(d!)^2}{(2d+2)!}\,.$$

3.3 Practical meaning and use

Taking into account the bounds provided by Theorem 2, the tail superset indices are normalized after multiplication by 90, so that the corresponding affinity matrix is $[90\,\Upsilon^2_{\{i,j\}}(\ell)]_{1\le i,j\le d}$, on which classical clustering algorithms and analyses can be performed.
However, even if this pairwise normalization is correct, the use of this bound is more powerful when comparing subsets with different sizes. Indeed, thanks to the renormalization $\Upsilon^2_u(\ell)/\Upsilon^2_u(\ell^{\vee,u})$ due to Theorem 2 all the renormalized superset tail indices can now be compared even if the subsets u have unequal sizes. The effective dimension of the asymptotic dependence structure could be defined as

$$\Delta(\ell) := \operatorname{argmax}_{|u|, u\subseteq\{1,\dots,d\}} \left\{ \frac{(2|u|+2)!\,\Upsilon^2_u(\ell)}{2(|u|!)^2} \right\} .$$

In the asymptotic independent case, $\Delta(\ell^+) = 1$ since $\Upsilon^2_{\{i,j\}}$ always vanishes: the entire additive component of ℓ^+ explains the whole variance. In the asymptotic complete dependent case, $\Delta(\ell^\vee) = d$ by application of Theorem 2. Now, for models in between, rules can be easily defined: select a subset u that achieves the maximization, or remove (in the asymptotic dependence modeling) subsets associated with small values of the previous brace.

Let us provide an example. Consider ℓ a trivariate stdf with value at (x, y, z) given by

$$\ell(x,y,z) = ((0.36x)^{\frac{1}{0.8}} + (0.35y)^{\frac{1}{0.8}})^{0.8} + ((0.37x)^{\frac{1}{0.44}} + (0.38z)^{\frac{1}{0.44}})^{0.44} + ((0.32y)^{\frac{1}{0.67}} + (0.30z)^{\frac{1}{0.67}})^{0.67} + ((0.27x)^{\frac{1}{0.04}} + (0.33y)^{\frac{1}{0.04}} + (0.32z)^{\frac{1}{0.04}})^{0.04} . \tag{10}$$

It is an asymmetric extreme value logistic model. Its associated tail dependograph is drawn below

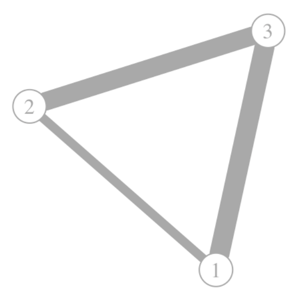

Figure 3: The Tail Dependograph of the stdf (10).

A quick estimation gives $\Upsilon^2_{12} = 1.700684 \times 10^{-4}$, $\Upsilon^2_{13} = 1.625145 \times 10^{-4}$, $\Upsilon^2_{23} = \mathbf{2.909913} \times 10^{-4}$ and $\Upsilon^2_{123} = 1.391676{\times}10^{-4}$. These values are not completely comparable since the sizes of the subsets are unequal. Their single use does not indicate the effective dimension of the function ℓ. Applying our bound, one obtains $\Upsilon^2_{12}(\ell)\cdot\Upsilon^{-2}_{12}(\ell^{\vee,12}) \simeq 0.0153$, $\Upsilon^2_{13}(\ell)\cdot\Upsilon^{-2}_{13}(\ell^{\vee,13}) \simeq 0.0146$, $\Upsilon^2_{23}(\ell)\cdot\Upsilon^{-2}_{23}(\ell^{\vee,23}) \simeq 0.0262$ and $\Upsilon^2_{123}(\ell)\cdot\Upsilon^{-2}_{123}(\ell^{\vee,123}) \simeq \mathbf{0.0779}$. These calculations reveal that the strength of asymptotic dependence, when modelled by ℓ, between the three components is closer to the possible maximal value, than its pairwise equivalent. Thus, the effective dimension of ℓ is 3 and it is not 2, according to this criteria. In other words, one should not simplify the representation of ℓ by combining only bivariate terms. The knowledge of our bound is crucial to construct this reasoning.

4 Proofs

Proof of Theorem 1. The proof relies on the combination of

$$\psi(\mathbf{x}) = \psi(\mathbf{1}) \int_C \max(\mathbf{x} \cdot \mathbf{w}) d\nu(\mathbf{w}) = \psi(\mathbf{1}) - \psi(\mathbf{1}) \int_C d\nu(\mathbf{w}) \int_0^1 ds 1_{s \geq \mathbf{x} \cdot \mathbf{w}}$$

for $\mathbf{x} \in [0,1]^d$ with the equality

$$\sum_{v \subseteq u} \psi_v(\mathbf{x}) = \int_{[0,1]^{-u}} \psi(\mathbf{x}) d\lambda_{-u}(\mathbf{x}) .$$

Then, applying the Fubini–Tonelli theorem yields

$$\begin{aligned}\sum_{v \subseteq u} \psi_v(\mathbf{x}) &= 2^{|u|}\psi(\mathbf{1}) - \psi(\mathbf{1}) \int_{[0,1]^{-u}} d\lambda_{-u}(\mathbf{x}) \int_C d\nu(\mathbf{w}) \int_0^1 ds 1_{s \geq \mathbf{x} \cdot \mathbf{w}} \\ &= 2^{|u|}\psi(\mathbf{1}) - \psi(\mathbf{1}) \int_C \left\{ \int_0^1 \prod_{i \in u} 1_{s \geq x_i w_i} \prod_{i \notin u} K_i(w_i; s)\, ds \right\} d\nu(\mathbf{w}) .\end{aligned}$$

Consequently, if $u = \varnothing$

$$\psi_\varnothing = \psi(\mathbf{1}) - \psi(\mathbf{1}) \int_C \left\{ \int_0^1 \prod_{i=1}^d K_i(w_i; s)\, ds \right\} d\nu(\mathbf{w})$$

whereas if $u \neq \varnothing$

$$\begin{aligned}\psi_u(\mathbf{x}) &= \sum_{v \subseteq u} (-1)^{|u \setminus v|} \left\{ \sum_{\tilde{v} \subseteq v} \psi_{\tilde{v}}(\mathbf{x}) \right\} \\ &= -\psi(\mathbf{1}) \int_C \left\{ \int_0^1 \sum_{v \subseteq u} (-1)^{|u \setminus v|} \prod_{i \in v} 1_{s \geq x_i w_i} \prod_{i \notin v} K_i(w_i; s)\, ds \right\} d\nu(\mathbf{w}) \\ &= -\psi(\mathbf{1}) \int_C \int_0^1 \left(\sum_{v \subseteq u} (-1)^{|u \setminus v|} \prod_{i \in v} 1_{s \geq x_i w_i} \prod_{i \in u \setminus v} K_i(w_i; s) \right) \prod_{i \notin u} K_i(w_i; s)\, ds d\nu(\mathbf{w}) \\ &= -\psi(\mathbf{1}) \int_C \left\{ \int_0^1 \prod_{i \in u} \left(1_{s \geq x_i w_i} - K_i(w_i; s)\right) \prod_{i \notin u} K_i(w_i; s)\, ds \right\} d\nu(\mathbf{w}) .\end{aligned}$$

For non-empty u, the term ψ_u is centered so that its variance σ_u^2 is also the second order moment.

$$\begin{aligned}\int_{[0,1]^{|u|}} \psi_u^2(\mathbf{x}_u) d\lambda_u(\mathbf{x}) &= \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{i \notin u} K_i(w_i; s) K_i(v_i; t) \\ &\quad \int_{[0,1]^{|u|}} \prod_{i \in u} \left(1_{s \geq x_i w_i} - K_i(w_i; s)\right) \left(1_{t \geq x_i v_i} - K_i(v_i; t)\right) d\lambda_u(\mathbf{x}) \\ &= \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{i \notin u} K_i(w_i; s) K_i(v_i; t) \\ &\quad \prod_{i \in u} \int_0^1 \left(1_{s \geq x_i w_i} - K_i(w_i; s)\right) \left(1_{t \geq x_i v_i} - K_i(v_i; t)\right) d\lambda_i(x_i) \\ &= \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{i \notin u} K_i(w_i; s) K_i(v_i; t) \\ &\quad \prod_{i \in u} \left(K_i(w_i, v_i; s, t) - K_i(w_i; s) K_i(v_i; t)\right) .\end{aligned}$$

The last assertion comes from the computation of $\int \psi^2 d\lambda(\mathbf{x}) = \sigma^2 + \psi_\varnothing^2$. More precisely,

$$\sigma^2 + \psi_\varnothing^2 = \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \int_{[0,1]^d} (1 - 1_{s \geq \mathbf{x} \cdot \mathbf{w}})(1 - 1_{t \geq \mathbf{x} \cdot \mathbf{v}}) d\lambda(\mathbf{x})$$
$$= \psi(\mathbf{1})(2\psi_\varnothing - \psi(\mathbf{1})) + \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{i=1}^d K_i(w_i, v_i; s, t)$$

using the fact that $\psi(\mathbf{1})^2 \int_C \int_{[0,1]^d} \int_{[0,1]} 1_{s \geq \mathbf{x} \cdot \mathbf{w}} \, d\nu(\mathbf{w}) d\lambda(\mathbf{x}) ds = \psi(\mathbf{1})^2 - \psi(\mathbf{1})\psi_\varnothing$. The result follows. □

Proof of Corollary 1. Following [18] one knows that

$$\psi_\varnothing^2 + \mathfrak{I}_u^2 = \int_{[0,1]^{2d-|u|}} \psi(\mathbf{x}^u, \mathbf{x}^{-u}) \psi(\mathbf{x}^u, \mathbf{z}^{-u}) d\lambda(\mathbf{x}) d\lambda_{-u}(\mathbf{z})$$
$$= \psi(\mathbf{1})^2 \int_{[0,1]^{2d-|u|}} d\lambda(\mathbf{x}) d\lambda_{-u}(\mathbf{z}) \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt$$
$$(1 - 1_{s \geq \mathbf{x}^u \cdot \mathbf{w}^u} 1_{s \geq \mathbf{x}^{-u} \cdot \mathbf{w}^{-u}})(1 - 1_{t \geq \mathbf{x}^u \cdot \mathbf{v}^u} 1_{t \geq \mathbf{z}^{-u} \cdot \mathbf{v}^{-u}})$$
$$= \psi(\mathbf{1})(2\psi_\varnothing - \psi(\mathbf{1})) + \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt$$
$$\int_{[0,1]^{2d-|u|}} d\lambda(\mathbf{x}) d\lambda_{-u}(\mathbf{z}) 1_{s \geq \mathbf{x}^u \cdot \mathbf{w}^u} 1_{s \geq \mathbf{x}^{-u} \cdot \mathbf{w}^{-u}} 1_{t \geq \mathbf{x}^u \cdot \mathbf{v}^u} 1_{t \geq \mathbf{z}^{-u} \cdot \mathbf{v}^{-u}}$$
$$= \psi(\mathbf{1})(2\psi_\varnothing - \psi(\mathbf{1})) + \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt$$
$$\prod_{i \in u} K_i(w_i, v_i; s, t) \prod_{i \notin u} K_i(w_i; s) K_i(v_i; t) \ .$$

Again, starting from

$$\sigma^2 + \psi_\varnothing^2 - \overline{\tau}_u^2 = \int_{[0,1]^{d+|u|}} \psi(\mathbf{x}^u, \mathbf{x}^{-u}) \psi(\mathbf{z}^u, \mathbf{x}^{-u}) d\lambda(\mathbf{x}) d\lambda_u(\mathbf{z})$$
$$= \psi(\mathbf{1})^2 \int_{[0,1]^{2d-|u|}} d\lambda(\mathbf{x}) d\lambda_u(\mathbf{z}) \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt$$
$$(1 - 1_{s \geq \mathbf{x}^u \cdot \mathbf{w}^u} 1_{s \geq \mathbf{x}^{-u} \cdot \mathbf{w}^{-u}})(1 - 1_{t \geq \mathbf{z}^u \cdot \mathbf{v}^u} 1_{t \geq \mathbf{x}^{-u} \cdot \mathbf{v}^{-u}})$$
$$= \psi(\mathbf{1})(2\psi_\varnothing - \psi(\mathbf{1})) + \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt$$
$$\prod_{i \in u} K_i(w_i; s) K_i(v_i; t) \prod_{i \notin u} K_i(w_i, v_i; s, t) \ .$$

Now, applying (7),

$$\Upsilon_u^2 = \frac{1}{2^{|u|}} \int_{[0,1]^{|u|+d}} \left(\sum_{a \subseteq u} (-1)^{|u \setminus a|} \psi(\mathbf{x}^a, \mathbf{z}^{-a}) \right)^2 d\lambda_u(\mathbf{x}) d\lambda(\mathbf{z}) \tag{11}$$
$$= \frac{1}{2^{|u|}} \sum_{a, a' \subseteq u} (-1)^{|u \setminus a| + |u \setminus a'|} \int_{[0,1]^{|u|+d}} \psi(\mathbf{x}^a, \mathbf{z}^{-a}) \psi(\mathbf{x}^{a'}, \mathbf{z}^{-a'}) d\lambda_u(\mathbf{x}) d\lambda(\mathbf{z})$$
$$= \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \int_{[0,1]^{|u|+d}} d\lambda_u(\mathbf{x}) d\lambda(\mathbf{z})$$
$$\frac{1}{2^{|u|}} \sum_{a, a' \subseteq u} (-1)^{|u \setminus a| + |u \setminus a'|} (1 - 1_{s \geq \mathbf{x}^a \cdot \mathbf{w}^a} 1_{s \geq \mathbf{z}^{-a} \cdot \mathbf{w}^{-a}})(1 - 1_{t \geq \mathbf{x}^{a'} \cdot \mathbf{v}^{a'}} 1_{t \geq \mathbf{z}^{-a'} \cdot \mathbf{v}^{-a'}})$$
$$= \psi(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \int_{[0,1]^{|u|+d}} d\lambda_u(\mathbf{x}) d\lambda(\mathbf{z})$$
$$\frac{1}{2^{|u|}} \sum_{a, a' \subseteq u} (-1)^{|u \setminus a| + |u \setminus a'|} 1_{s \geq \mathbf{x}^a \cdot \mathbf{w}^a} 1_{s \geq \mathbf{z}^{-a} \cdot \mathbf{w}^{-a}} 1_{t \geq \mathbf{x}^{a'} \cdot \mathbf{v}^{a'}} 1_{t \geq \mathbf{z}^{-a'} \cdot \mathbf{v}^{-a'}}$$

since $\sum_{a\subseteq u}(-1)^{|u\setminus a|} = 0$ as soon as u is non-empty. As usual, we let $a\Delta b = (a\cup b)\setminus(a\cap b)$ be the symmetric difference of the subsets a and b. As a consequence,

$$\begin{aligned}\Upsilon_u^2 &= \psi(\mathbf{1})^2\int_C d\nu(\mathbf{w})\int_C d\nu(\mathbf{v})\int_0^1 ds\int_0^1 dt\\ &\quad \frac{1}{2^{|u|}}\sum_{a,a'\subseteq u}(-1)^{|u\setminus a|+|u\setminus a'|}\prod_{i\in a\Delta a'}K_i(w_i;s)K_i(v_i;t)\prod_{i\notin a\Delta a'}K_i(w_i,v_i;s,t)\\ &= \psi(\mathbf{1})^2\int_C d\nu(\mathbf{w})\int_C d\nu(\mathbf{v})\int_0^1 ds\int_0^1 dt\sum_{a\subseteq u}(-1)^{|a|}\prod_{i\in a}K_i(w_i;s)K_i(v_i;t)\prod_{i\notin a}K_i(w_i,v_i;s,t)\end{aligned}$$

where the last equality comes from the general fact

$$\sum_{a\subseteq u,b\subseteq u}(-1)^{|u\setminus a|+|u\setminus b|}f(a\Delta b) = 2^{|u|}\sum_{a\subseteq u}(-1)^{|a|}f(a)\ .$$

Finally, one obtains

$$\Upsilon_u^2 = \psi(\mathbf{1})^2\int_C d\nu(\mathbf{w})\int_C d\nu(\mathbf{v})\int_0^1 ds\int_0^1 dt\prod_{i\notin u}K_i(w_i,v_i;s,t)\prod_{i\in u}(K_i(w_i,v_i;s,t)-K_i(w_i;s)K_i(v_i;t))\ .$$

□

Proof of Proposition 1. Prior to proving Proposition 1, we review some preliminary results.

Lemma 2. *Let $A_1,\dots,A_d$ be non-empty sets, $A := A_1\times\dots\times A_d$, and $f : A\to\mathbb{R}$. For $\mathbf{x}$ and $\mathbf{z}$ both in A and for any subset $v\subseteq\{1,\dots,d\}$ then*

$$D_{(\mathbf{z}^v,\mathbf{x}^{-v})}^{(\mathbf{x}^v,\mathbf{z}^{-v})}f = (-1)^{d-|v|}D_{\mathbf{z}}^{\mathbf{x}}f\ .$$

In particular,

$$D_{\mathbf{x}}^{\mathbf{z}}f = (-1)^d D_{\mathbf{z}}^{\mathbf{x}}f\ .$$

If $x_i = z_i$ for some $i\in\{1,\dots,d\}$ then $D_{\mathbf{z}}^{\mathbf{x}}f = 0$.

Proof of Lemma 2. For $\mathbf{x}$ and $\mathbf{z}$ both in A write $\mathbf{x}' := (x_1,\dots,x_{d-1})$ and $\mathbf{z}' := (z_1,\dots,z_{d-1})$. Define $\rho : A_d\to\mathbb{R}$ by

$$\rho(t) := D_{\mathbf{z}'}^{\mathbf{x}'}f(\cdot,t)\ .$$

Then

$$\begin{aligned}D_{\mathbf{z}}^{\mathbf{x}}f &= \sum_{u\subseteq\{1,\dots,d\}}(-1)^{|u|}f(\mathbf{z}^u,\mathbf{x}^{-u})\\ &= \sum_{u\subseteq\{1,\dots,d-1\}}(-1)^{|u|}f(\mathbf{z}^u,\mathbf{x}^{-u}) + \sum_{v\subseteq\{1,\dots,d-1\},u=v\cup\{d\}}(-1)^{|u|}f(\mathbf{z}^u,\mathbf{x}^{-u})\\ &= \sum_{u\subseteq\{1,\dots,d-1\}}(-1)^{|u|}f(\mathbf{z}^u,\mathbf{x}^{-(u\cup\{d\})},x_d) - \sum_{u\subseteq\{1,\dots,d-1\}}(-1)^{|u|}f(\mathbf{z}^u,\mathbf{x}^{-(u\cup\{d\})},z_d)\\ &= \rho(x_d)-\rho(z_d)\\ &= D_{z_d}^{x_d}\rho\ .\end{aligned}$$

It implies $D_{\mathbf{z}}^{\mathbf{x}}f = 0$ whenever $x_d = z_d$. Moreover, we also see that

$$D_{(\mathbf{z}',x_d)}^{(\mathbf{x}',z_d)}f = -D_{\mathbf{z}}^{\mathbf{x}}f\ .$$

The analogue results hold for any coordinate $i\le d$. The conclusion follows by obvious iteration. □

Lemma 3. *Let $A_1,\dots,A_d\subseteq\overline{\mathbb{R}}$ be non-empty, $A := A_1\times\dots\times A_d$, and suppose $f : A\to\mathbb{R}$ to be $\mathbf{1}_d$-alternating. Then, for any $\mathbf{x}'$ and $\mathbf{z}'$ both in $\prod_{j=1}^{d-1}A_j$ the function*

$$r(t) := \left|D_{\mathbf{z}'}^{\mathbf{x}'}f(\cdot,t)\right|$$

is decreasing on A_d.

Proof of Lemma 3. By Lemma 2, we may assume that $\mathbf{x}' \le \mathbf{z}'$. For $s \le t$ both in A_d we have

$$0 \ge D_{(\mathbf{z}',t)}^{(\mathbf{x}',s)} f = D_{\mathbf{z}'}^{\mathbf{x}'} f(\cdot, s) - D_{\mathbf{z}'}^{\mathbf{x}'} f(\cdot, t)$$

where both terms on the right hand side are non-positive. Hence

$$0 \ge D_{\mathbf{z}'}^{\mathbf{x}'} f(\cdot, t) \ge D_{\mathbf{z}'}^{\mathbf{x}'} f(\cdot, s)$$

and so $r(t) \le r(s)$. □

We now go back to the proof of Proposition 1. By iterating Lemma 3, we deduce that

$$\left| \sum_{v \subseteq u} (-1)^{|u \setminus v|} f(\mathbf{x}^v, \mathbf{z}^{-v}) \right| = \left| D_{\mathbf{z}^u}^{\mathbf{x}^u} f(\cdot, \mathbf{z}^{-u}) \right| \le \left| D_{\mathbf{z}^u}^{\mathbf{x}^u} f(\cdot, \mathbf{0}^{-u}) \right| = \left| \sum_{v \subseteq u} (-1)^{|u \setminus v|} f(\mathbf{x}^v, \mathbf{z}^{u \setminus v}, \mathbf{0}^{-u}) \right| .$$

It yields

$$\Upsilon_u^2(f) \le \frac{1}{2^{|u|}} \int \left(\sum_{v \subseteq u} (-1)^{|u \setminus v|} f(\mathbf{x}^v, \mathbf{z}^{u \setminus v}, \mathbf{0}^{-u}) \right)^2 d\mathbf{x}^u d\mathbf{z}^u = \Upsilon_u^2(f^{[u]}) ,$$

where $f^{[u]}(\mathbf{z}^u) = f(\mathbf{z}^u, \mathbf{0}^{-u})$. □

Proof of Lemma 1. Recall that $\ell^+(\mathbf{x})$ and $\ell^\vee(\mathbf{x})$ now stand for $\sum_{i=1}^d x_i$ and $\max(\mathbf{x})$ respectively. Set also $\ell^{\vee,u}(\mathbf{x}) = \max_{i \in u} x_i$. Stable tail dependence functions ℓ have the well-known property

$$\ell^\vee \le \ell \le \ell^+ . \tag{12}$$

To prove (12) recall the equality $\ell(\mathbf{x}) = \mu([\mathbf{x}, \boldsymbol{\infty}]^c)$. Then, the inclusion

$$[x_i \mathbf{e}_i, \boldsymbol{\infty}]^c \subseteq [\mathbf{x}, \boldsymbol{\infty}]^c = \bigcup_{i \le d} \{\mathbf{y} | y_i < x_i\}$$

leads to the result since ℓ is the identity on each axis. Indeed, ℓ is homogeneous and equals one at the canonical basis vectors. The inequality (12) is easily transferred to first and second order moments

$$\ell_\varnothing^\vee \le \ell_\varnothing \le \ell_\varnothing^+$$

and

$$\sigma^2(\ell^\vee) + (\ell_\varnothing^\vee)^2 \le \sigma^2(\ell) + \ell_\varnothing^2 \le \sigma^2(\ell^+) + (\ell_\varnothing^+)^2 .$$

Precisely, it yields

$$\frac{d}{d+1} \le \ell_\varnothing \le \frac{d}{2}$$

and

$$\frac{d}{d+2} \le \sigma^2(\ell) + \ell_\varnothing^2 \le \frac{d}{12} + \frac{d^2}{4} ,$$

implying

$$\sigma^2(\ell) \le \frac{d\,(3d^3 + 7d^2 - 7d + 1)}{12(d+1)^2} .$$

From (12), one can also prove that

$$\frac{(\max_{i \in u} x_i)^{d-|u|+1} + d - |u|}{d - |u| + 1} \le \sum_{v \subseteq u} \ell_v(\mathbf{x}) = \int \ell(\mathbf{x}) d\mathbf{x}_{-u} \le d/2 + \sum_{i \in u} \{x_i - 1/2\} .$$

Indeed, the left hand term comes from

$$\mathbb{E}[\max(m, \mathbf{Y}_{-u})] = \int_0^1 \max(m, y)(d - |u|) y^{d-|u|-1} dy$$

where $m = \max(\mathbf{x}_u)$ and $\mathbf{Y}_{-u}$ is sampled from identical standard uniform. Taking moment of order 2, one obtains

$$\mathcal{I}_u^2(\ell^{\vee,u}) + (\ell_\varnothing^{\vee,u})^2 \leq \mathcal{I}_u^2(\ell) + \ell_\varnothing^2 \leq \mathcal{I}_u^2(\ell^+) + (\ell_\varnothing^+)^2 \ .$$

Precise computations give

$$\mathcal{I}_u^2(\ell^+) + (\ell_\varnothing^+)^2 = \frac{d^2}{4} + \frac{|u|}{12}$$

and

$$\mathcal{I}_u^2(\ell^{\vee,u}) + (\ell_\varnothing^{\vee,u})^2 = \frac{\sigma^2(\ell^{\vee,u}) + (\ell_\varnothing^{\vee,u})^2}{(d-|u|+1)^2} + \frac{2\ell_\varnothing^{\vee,u}(d-|u|)}{d-|u|+1} + \frac{(d-|u|)^2}{(d-|u|+1)^2}$$

where

$$\ell_\varnothing^{\vee,u} = \frac{|u|}{|u|+1}$$

and

$$\sigma^2(\ell^{\vee,u}) = \frac{|u|}{(|u|+1)^2(|u|+2)}$$

are deduced from Example 1 (just replace d by $|u|$). □

Proof of Theorem 2. Proposition 1 allows us to focus on the majorization associated to the largest subset u in $\{1,\dots,d\}$. Then, for ℓ a d-variate stdf, we are just interested in an upper bound for $\Upsilon^2_{\{1,\dots,d\}}$. Let us introduce the following notation $h_{\mathbf{w}}(\mathbf{x}) = \max(\mathbf{x}\cdot\mathbf{w})$ so that the spectral representation can be written as

$$\ell(\mathbf{x}) = \ell(\mathbf{1}) \int_C h_{\mathbf{w}}(\mathbf{x}) d\nu(\mathbf{w}) \ .$$

From linearity, one obtains

$$D_{\mathbf{z}}^{\mathbf{x}}\ell = \ell(\mathbf{1}) \int_C D_{\mathbf{z}}^{\mathbf{x}} h_{\mathbf{w}} d\nu(\mathbf{w}) \ .$$

Now, recall $\mu_{\mathbf{w}}$ is the image of the Lebesgue measure λ on the half-line $\mathbb{R}_+$ under the mapping $s \mapsto s/\mathbf{w}$, and the fact that $h_{\mathbf{w}}(\mathbf{x}) = \mu_{\mathbf{w}}([\mathbf{x},\infty]^c)$ as discussed in Subsection 2.1. Consequently,

$$|D_{\mathbf{z}}^{\mathbf{x}} h_{\mathbf{w}}| = \mu_{\mathbf{w}}([\mathbf{x}\wedge\mathbf{z}, \mathbf{x}\vee\mathbf{z}]) = \int_0^1 1_{\mathbf{x}\wedge\mathbf{z}\leq s/\mathbf{w}\leq\mathbf{x}\vee\mathbf{z}} ds \ .$$

Then, from [9, Theorem 1] and what precedes, one can write

$$\begin{aligned}
\Upsilon^2_{\{1,\dots,d\}}(\ell) &= 2^{-d} \int_{[0,1]^d}\int_{[0,1]^d} (D_{\mathbf{z}}^{\mathbf{x}}\ell)^2 d\mathbf{x}d\mathbf{z} \\
&\leq 2^{-d}\ell(\mathbf{1})^2 \int_{[0,1]^d}\int_{[0,1]^d} \left(\int_C |D_{\mathbf{z}}^{\mathbf{x}} h_{\mathbf{w}}| d\nu(\mathbf{w})\right)\left(\int_C |D_{\mathbf{z}}^{\mathbf{x}} h_{\mathbf{v}}| d\nu(\mathbf{v})\right) d\mathbf{x}d\mathbf{z} \ .
\end{aligned}$$

The Fubini–Tonelli theorem leads to

$$\begin{aligned}
&\int_{[0,1]^d}\int_{[0,1]^d} |D_{\mathbf{z}}^{\mathbf{x}} h_{\mathbf{w}}||D_{\mathbf{z}}^{\mathbf{x}} h_{\mathbf{v}}| d\mathbf{x}d\mathbf{z} \\
&= \int_0^1\int_0^1 \left(\int_{[0,1]^d}\int_{[0,1]^d} 1_{\mathbf{x}\wedge\mathbf{z}\leq s/\mathbf{w}\leq\mathbf{x}\vee\mathbf{z}} 1_{\mathbf{x}\wedge\mathbf{z}\leq t/\mathbf{v}\leq\mathbf{x}\vee\mathbf{z}} d\mathbf{x}d\mathbf{z}\right) dsdt \\
&= 2^d \int_0^1\int_0^1 \prod_{i=1}^d \left(\frac{s}{w_i}\wedge\frac{t}{v_i}\right)\left(1-\frac{s}{w_i}\vee\frac{t}{v_i}\right)_+ dsdt \\
&\leq 2^d \prod_{i=1}^d \left(\int_0^1\int_0^1 \left(\frac{s}{w_i}\wedge\frac{t}{v_i}\right)^d \left(1-\frac{s}{w_i}\vee\frac{t}{v_i}\right)_+^d dsdt\right)^{1/d} \\
&= 2^d \prod_{i=1}^d \left(\int_0^1\int_0^1 (s\wedge t)^d (1-s\vee t)^d\, w_i v_i dsdt\right)^{1/d} \\
&= 2^d \left(\prod_{i=1}^d w_i^{1/d} v_i^{1/d}\right)\left(\int_0^1\int_0^1 (s\wedge t)^d (1-s\vee t)^d\, dsdt\right)
\end{aligned}$$

where the inequality comes from the generalization of Hölder's inequality: for any positive numbers p_i such that $\frac{1}{p_1} + \ldots + \frac{1}{p_n} = 1$

$$\int |f_1 \ldots f_n|\, d\mu \le \left(\int |f_1|^{p_1}\, d\mu\right)^{1/p_1} \cdots \left(\int |f_n|^{p_n}\, d\mu\right)^{1/p_n} . \tag{13}$$

Combining the last results gives us

$$\begin{aligned} \Upsilon^2_{\{1,\ldots,d\}}(\ell) &\le \ell(\mathbf{1})^2 \left(\int_C \prod_{i=1}^d w_i^{1/d} dv(\mathbf{w})\right)^2 \left(\int_0^1\int_0^1 (s\wedge t)^d (1-s\vee t)^d\, dsdt\right) \\ &\le \ell(\mathbf{1})^2 \left(\prod_{i=1}^d \int_C w_i dv(\mathbf{w})\right)^{2/d} \left(\int_0^1\int_0^1 (s\wedge t)^d (1-s\vee t)^d\, dsdt\right) \end{aligned}$$

invoking again (13). Since a stable tail dependence function satisfies $\ell(\mathbf{e}_i) = 1$, we have $\int_C w_i dv(\mathbf{w}) = 1/\ell(\mathbf{1})$. Finally

$$\Upsilon^2_{\{1,\ldots,d\}}(\ell) \le \int_0^1\int_0^1 (s\wedge t)^d (1-s\vee t)^d\, dsdt = \Upsilon^2_{\{1,\ldots,d\}}(h_{\mathbf{1}}) = \frac{2(d!)^2}{(2d+2)!}$$

by application of Corollary 1 with $\psi(\mathbf{1}) = 1$, $v = \delta_{\mathbf{1}}$, and Example 3.
Assume now the second assertion of Theorem 2 and recall that if ℓ is a d-variate stdf then $\ell^{[u]}$ is a $|u|$-variate stdf. Combining the assumption with what precedes and Proposition 1, we obtain

$$\Upsilon^2_u(\ell^{\vee,u}) = \Upsilon^2_u(\ell) \le \Upsilon^2_u(\ell^{[u]}) \le \Upsilon^2_u(\ell^{\vee,u})$$

so that $\Upsilon^2_u(\ell^{[u]}) = \Upsilon_u(\ell^{\vee,u})$. The problem is now the same as the last statement of Theorem 2 for $d = |u|$. The result will thus follow if one can prove it directly.
Let ℓ be any stdf where the maximal value is attained. Then the inequalities after the inequality (13) are in fact equalities, in particular

$$\int_C \prod_{i=1}^d w_i^{1/d} dv(\mathbf{w}) = \left(\prod_{i=1}^d \int_C w_i dv(\mathbf{w})\right)^{1/d}$$

implying by Lemma 4 (below) that the functions $\mathbf{w} \mapsto w_i^{1/d}$, $1 \le i \le d$, are proportional, as are then also $\mathbf{w} \mapsto w_i$. Since $\int_C w_i dv(\mathbf{w}) = 1/\ell(\mathbf{1})$ for each i, we see that v-almost surely the components w_i are equal, i.e. v is concentrated on the diagonal $\{\mathbf{w} | w_1 = w_2 = \ldots = w_d\}$, and the only $\mathbf{w} \in C$ with this property is $\mathbf{w} = \mathbf{1}$. Consequently, $v = \delta_{\mathbf{1}}$ and $\ell(\mathbf{1}) = 1$. In other words, $\ell(\mathbf{x}) = \max(x_1, \ldots, x_d)$. □

Lemma 4. *Consider in the generalized Hölder inequality (13) the special case* $p_1 = \ldots = p_n = 1/n$, $f_i \ge 0$ *and* $0 < \int f_i^n d\mu < \infty$ *for all* i. *Then*

$$\int f_1 \ldots f_n\, d\mu \le \left(\prod_{i=1}^n \int f_i^n d\mu\right)^{1/n} \tag{14}$$

with equality if and only if the functions f_i *are proportional, i.e.* $f_i = \alpha_i f_1$ *for* $i = 2, \ldots, n$ *where* $\alpha_i > 0$.

Proof of Lemma 4. If $f_i = \alpha_i f_1$ for all i ($\alpha_1 = 1$) then both sides in (14) have the same value $\int f_1^n d\mu \cdot \prod_{i=1}^n \alpha_i$. Supposing now equality in (14), we proceed by induction. For $n = 2$ the inequality (14) is the Cauchy-Schwarz inequality and it is well-known that f_1 and f_2 are proportional in case of equality. We assume now the validity of our assertion for some $n \ge 2$ and consider $n+1$ functions $f_1, \ldots, f_{n+1}$. Hölder's inequality for two functions $g, h \ge 0$ reads

$$\int gh d\mu \le \left(\int g^p d\mu\right)^{1/p} \left(\int h^q d\mu\right)^{1/q}$$

with $p, q \geq 1$ such that $1/p + 1/q = 1$, and with equality iff g^p and h^q are proportional. It will be applied to $g := f_1 \cdots f_n$, $h := f_{n+1}$, $p = (n+1)/n$ and $q = n+1$

$$\int (f_1 \cdots f_n) f_{n+1} d\mu \leq \left(\int (f_1 \cdots f_n)^{\frac{n+1}{n}} d\mu\right)^{\frac{n}{n+1}} \left(\int f_{n+1}^{n+1} d\mu\right)^{\frac{1}{n+1}}$$
$$\leq \left(\prod_{i=1}^{n} \left(\int f_i^{n+1} d\mu\right)^{\frac{1}{n}}\right)^{\frac{n}{n+1}} \left(\int f_{n+1}^{n+1} d\mu\right)^{\frac{1}{n+1}} = \left(\prod_{i=1}^{n+1} \int f_i^{n+1} d\mu\right)^{\frac{n}{n+1}}$$

where the second majorization is obtained by applying the induction hypotheses to $f_1^{(n+1)/n}, \ldots, f_n^{(n+1)/n}$. These two inequalities are by assumption equalities, and lead to

$$(f_1 \cdots f_n)^{\frac{n+1}{n}} = \beta_1 f_{n+1}^{n+1}$$

as well as $f_2^{n+1} = \beta_2 f_1^{n+1}, \ldots, f_n^{n+1} = \beta_n f_1^{n+1}$ for positive $\beta_1, \beta_2, \ldots, \beta_n$, i.e.

$$\beta_1^{1/(n+1)} f_{n+1} = (\beta_2 \cdots \beta_n)^{1/(n+1)} \cdot f_1 \ ,$$

showing $f_1, \ldots, f_{n+1}$ to be proportional. □

Concluding remarks

The Choquet representation of homogeneous co-survival functions, shown to be unique in [16], is the source of all results in this paper. Then, the Fubini–Tonelli theorem appears as the main ingredient in transposing the spectral expressions to several forms of combined variances.

Illustrated through Monte-Carlo comparisons, the coverage accuracy is significantly improved. However, this just illustrates the contraction of the domain of integration.

As a natural example, the function that summarizes the tail dependence structure in extreme value theory, namely the stable tail dependence function, received successfully the application of the new results. Furthermore, the generalization of Hölder's inequality associated with more tricky justifications from multivariate monotonicity yielded a sharp upper bound to the quantity of interest.

Finally, it may be worth pursuing the consequences for measuring the effective asymptotic dependence dimension of a random vector.

Acknowledgements: We thank three referees for their comments and suggestions, which helped to improve our manuscript, and the editorial board for managing the overall review. The first author thanks her colleague Guillaume Aubrun for generously participating in some discussions on convexity.

Conflict of interest statement: Authors state no conflict of interest.

References

[1] Drees, H. and X. Huang (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *J. Multivariate Anal. 64*(1), 25–46.
[2] Efron, B. and C. Stein (1981). The jackknife estimate of variance. *Ann. Statist. 9*(3), 586–596.
[3] Fougères, A.-L. , L. de Haan, and C. Mercadier (2015). Bias correction in multivariate extremes. *Ann. Statist. 43*(2), 903–934.
[4] de Haan, L. and S. I. Resnick (1977). Limit theory for multivariate sample extremes. *Z. Wahr. verw. Geb. 40*, 317–337.
[5] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist. 19*(3), 293–325.

[6] Hooker, G. (2004). Discovering additive structure in black box functions. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge Discovery and Data Mining*, pp. 575–580.

[7] Huang, X. (1992). *Statistics of Bivariate Extremes*. PhD Thesis, Erasmus University Rotterdam.

[8] Konakli, K. and B. Sudret (2016). Global sensitivity analysis using low-rank tensor approximations. *Reliab. Eng. Syst. Saf. 156*, 64–83.

[9] Liu, R. and A. B. Owen (2006). Estimating mean dimensionality of analysis of variance decompositions. *J. Amer. Statist. Assoc. 101*(474), 712–721.

[10] Mercadier, C. and O. Roustant (2019). The tail dependograph. *Extremes 22*, 343–372.

[11] Ökten, G. and Y. Liu (2021). Randomized quasi-monte carlo methods in global sensitivity analysis. *Reliab. Eng. Syst. Saf. 210*, 107520, 9 pages.

[12] Owen, A. B. (2003). The dimension distribution and quadrature test functions. *Stat. Sin. 13*, 1–17.

[13] Razavi, S., A. Jakeman, A. Saltelli, C. Prieur, B. Iooss, E. Borgonovo, E. Plischke, S. Lo Piano, T. Iwanaga, W. Becker, S. Tarantola, J. H. Guillaume, J. Jakeman, H. Gupta, N. Melillo, G. Rabitti, V. Chabridon, Q. Duan, X. Sun, S. Smith, R. Sheikholeslami, N. Hosseini, M. Asadzadeh, A. Puy, S. Kucherenko, and H. R. Maier, (2021). The future of sensitivity analysis: An essential discipline for systems modeling and policy support, *Environ. Model. Softw. 137*, 104954, 22 pages.

[14] Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.

[15] Ressel, P. (2012). Functions operating on multivariate distribution and survival functions—With applications to classical mean-values and to copulas. *J. Multivariate Anal. 105*(1), 55–67.

[16] Ressel, P. (2013). Homogeneous distributions–And a spectral representation of classical mean values and stable tail dependence functions. *J. Multivariate Anal. 117*, 246–256.

[17] Ressel, P. (2019). Copulas, stable tail dependence functions, and multivariate monotonicity. *Depend. Model. 7*, 247–258.

[18] Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Math. Modeling Comput. Experiment 1*(4), 407–414.

[19] van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.