



HAL
open science

Distortions of multivariate risk measures: a level-sets based approach

Elena Di Bernardino, Didier Rullière

► **To cite this version:**

Elena Di Bernardino, Didier Rullière. Distortions of multivariate risk measures: a level-sets based approach. 2012. hal-00756387

HAL Id: hal-00756387

<https://hal.science/hal-00756387>

Preprint submitted on 22 Nov 2012

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.

Distortions of multivariate risk measures: a level-sets based approach

Elena Di Bernardino¹, Didier Rullière²

Abstract

In this paper, we propose a parametric model for multivariate distributions. The model is based on distortion functions, i.e. some transformations of a multivariate distribution which permit to generate new families of multivariate distribution functions. We derive some properties of considered distortions. A suitable proximity indicator between level curves is introduced in order to evaluate the quality of candidate distortion parameters. Using this proximity indicator and properties of distorted level curves, we give a specific estimation procedure. The estimation algorithm is mainly relying on straightforward univariate optimizations, and we finally get parametric representations of both multivariate distribution functions and associated level curves. Our results are motivated by applications in multivariate risk theory. The methodology is illustrated on real examples.

Keywords: Multivariate probability distortions, Level sets estimation, Iterated compositions, Hyperbolic conversion functions, Multivariate risk measures.

Introduction

Multivariate distributions can be modeled using parametric marginal distribution functions and parametric copulas, or directly by some specific parametric expressions of the multivariate cumulative distribution function. However some problems of such parametric forms can arise. We underline, for instance, the difficulty (see for example Bienvenüe and Rullière, 2012, in the univariate case)

- to fit non-regular or multimodal distributions using classical unimodal distributions;
- to change the number of parameters and to improve a fit to observed data;
- to estimate the parameters when their number is large;
- to get analytical expressions for level curves of the distribution functions.

In order to overcome these shortcomings, in this paper we will define multivariate distributions by using probability distortions. We show that using distortions has several advantages compared to using classical parametric multivariate distributions. In particular we focus on:

- the possibility to get analytical expressions both for the multivariate distribution function and for the associated level-curves;
- the huge variety of reachable distributions (multimodal, etc.), with the possibility to improve the fit by adding parameters (via distortions composition), and the possibility to converge to any target in dimension 1;

¹CNAM, Paris, Département IMATH, 292 rue Saint-Martin, Paris Cedex 03, France, elena.di_bernardino@cnam.fr

²Université de Lyon, Université Lyon 1, ISFA, Laboratoire SAF, 50 avenue Tony Garnier, 69366 Lyon, France, rulliere@univ-lyon1.fr

- some estimation facilities, with the possibility to get very good initial values for parameters even with a large number of parameters.

Using parametric distortions will also have some advantages compared to nonparametric fitting procedures like multivariate splines (e.g. see Awanou *et al.*, 2005) or kernel estimation (e.g. see Charpentier *et al.* 2006). Among advantages of the considered distortions, one can cite the possibility to ensure regular conditions (like monotonicity and positiveness of density function), and the possibility to get analytical expressions for both cumulative distribution function and for level curves.

Applications in multivariate risk theory:

Our results are motivated by applications in multivariate risk theory. In the following we detail the connection between the multivariate level-curves and the recent developments of risk theory.

During the last decades, researchers joined efforts to properly quantify and manage risks. Traditionally, risk measures are thought of as mappings from a set of real-valued random variables to the real numbers. However, it is often insufficient to consider a single real measure to quantify risks created by business activities, especially if the latter are affected by other external risk factors. Consequently we need a generalization of univariate classical risk measures in dimension higher than one. In the recent literature, an intuitive and immediate generalization of the *Value-at-Risk* measure in the case of a d -dimensional loss distribution function F is represented by its α -quantile curves, i.e. the boundary of its α -level set. More precisely, let F be a d -dimensional distribution function, then we define:

$$\text{VaR}_\alpha(F) = \{x \in \mathbb{R}_+^d : F(x) = \alpha\}, \text{ for } \alpha \in (0, 1).$$

This generalization of the *Value-at-Risk* may be useful to quantify dependent risks: dependent environmental risks factors in same geographic area, several dependent business lines of a financial institution (e.g. see Chebana and Ouarda, 2011; Cousin and Di Bernardino, 2012). The interest reader is also referred to Tibiletti (1993), Belzunce *et al.* (2007), Embrechts and Puccetti (2006), Nappo and Spizzichino (2009), Di Bernardino *et al.* (2011).

Starting from these considerations, we are interested in parametric representations of multivariate distribution function F and associated multivariate *Value-at-Risk* $\text{VaR}_\alpha(F)$ (i.e. level curves of F). They may be of great importance in many practical studies of multivariate risks because they analytically describe a multivariate risk-area (see Embrechts and Puccetti, 2006).

Some effort has been made in recent literature to estimate the level-curves of a multivariate distribution function. A non-parametric plug-in estimation of $\text{VaR}_\alpha(F)$ is given by Di Bernardino *et al.* (2011). de Haan and Huang (1995) model a risk-problem of flood, in the bivariate setting, using an extremal estimator of $\text{VaR}_\alpha(F)$, for F a bivariate distribution function. In this paper we provide an estimation algorithm to estimate multivariate distribution function F and associated $\text{VaR}_\alpha(F)$. Differently from the literature cited above, our procedure gives explicit parametric-analytical representations for estimated F and $\text{VaR}_\alpha(F)$. This feature can be relevant from a practical point of view.

Organization of the paper:

The paper is organized as follows. In Section 1, we introduce some notations, tools and technical assumptions. In Section 2, we present some transformations of a multivariate distribution function F which permit to generate new families of multivariate distribution functions. In particular, we introduce *external distortions* (see Section 2.1), and *internal* ones (see Section 2.2). In Section 3 we propose an estimation procedure in the case of piecewise linear distortions. In Section 4, we deal with estimation in the case of differentiable distortions. Numerical applications are presented in Section 5 using hyperbolic distortion functions. Finally, two real case studies are illustrated in Sections 5.2 and 5.3.

1. Basic notions and preliminaries

Assume that we have d underlying risks described by a d -dimensional nonnegative real-valued random vector $X = (X_1, \dots, X_d)$. Denote its multivariate distribution function by $F : \mathbb{R}_+^d \rightarrow [0, 1]$ with univariate margins $F_i(x_i) = P(X_i \leq x_i)$, for $i = 1, \dots, d$. Sklar's Theorem (1959) is a well-known result which states that for any random vector X , its multivariate distribution function has the representation

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)),$$

where C is called the *copula* function. Effectively, it is a distribution function on the d -cube $[0, 1]^d$ with uniform margins and it links the univariate margins to their full multivariate distribution. In the case where we have a continuous random vector, we know that $U_i = F_i(X_i)$ is an uniform random variable so that we can write

$$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)),$$

to be the unique copula associated with X , with quantile functions F_i^{-1} defined by:

$$F_i^{-1}(p) = \inf\{x \in \mathbb{R}_+ : F_i(x) \geq p\}, \quad \text{for } p \in (0, 1).$$

Denote \mathcal{F} the class of d -dimensional distribution functions F such that $F : \mathbb{R}_+^d \rightarrow [0, 1]$ be a continuous partially increasing function³ of a non-negative absolutely-continuous random vector⁴ $X = (X_1, \dots, X_d)$ (with respect to Lebesgue measure λ on \mathbb{R}^d). In the following, $F \in \mathcal{F}$ will be said to satisfy *regularity conditions*.

Let $L(\alpha) = \{x \in \mathbb{R}_+^d : F(x) \geq \alpha\}$ be the *upper α -level set* of F , for $\alpha \in (0, 1)$ and $d \geq 2$. Furthermore, for any set $A \subset \mathbb{R}_+^d$ we denote by ∂A its boundary.

Note that, under the *regularity conditions*, $\partial L(\alpha) = \{x \in \mathbb{R}_+^d : F(x) = \alpha\}$ has Lebesgue-measure zero in \mathbb{R}_+^d (e.g., see Property 3 in Tibiletti, 1990). We call $\partial L(\alpha)$ the *α -level curve* of distribution F . For instance, if $d = 2$, each $\partial L(\alpha)$ for $\alpha \in (0, 1)$, is identified by a decreasing curve in the plane \mathbb{R}_+^2 (e.g., see Theorem 2 in Rossi, 1973; Section 2 in Tibiletti, 1991).

We now recall the notion of *absolutely monotonic function* that will be useful later. The interested reader is referred also to Valdez and Xiao (2011). Assume that n is a non-negative integer.

Definition 1.1 (Absolute monotony) *A real function $g(t)$ is said to be absolutely monotonic, of order n , on an interval I if the following conditions are satisfied:*

- g is continuous on I ; and
- g has non-negative derivatives of orders up to, and including, n , i.e., $g^{(k)}(t) \geq 0$, for all t on the interior of I and for $k = 0, 1, \dots, n$.

Remark 1 If g and h are both absolutely monotonic of order n on an interval I and h is defined on I such that $h(t)$ is on the interior of I for all t on the interior of I , then the composite function $g \circ h(t)$ is also absolutely monotonic of order n on I .

³A function $F(x_1, \dots, x_d)$ is partially increasing on $\mathbb{R}_+^d \setminus \{0\}$ if the functions of one variable $g(\cdot) = F(x_1, \dots, x_{j-1}, \cdot, x_{j+1}, \dots, x_d)$ are increasing. About properties of partially increasing multivariate distribution functions we refer the interested reader to Rossi (1973) and Tibiletti (1991).

⁴We restrict ourselves to \mathbb{R}_+^d because, in our applications, components of d -dimensional vectors correspond to random losses and are then valued in \mathbb{R}_+ . However extensions of our results in the case of multivariate distribution function on the entire space \mathbb{R}^d are possible.

2. Considered distortions

The aim of this section is to introduce some simple transformations of a multivariate distribution function F which permit to generate new families of multivariate distribution functions. Firstly we will be interested in distortion applied to the function F (called external distortions and leading to a distorted function \tilde{F}^{ext}), secondly we will include also the distortions of marginal components of F . In this last case we will finally obtain a global distorted function \tilde{F} . The interested reader is referred, for instance, to Christian Genest in the conference “Distributions with Given Marginals and Statistical Modeling” (Barcelona, July 17-20, 2000) and Durrleman *et al.* (2000).

Definition 2.1 (Regular external distortions) *Let $T : [0, 1] \rightarrow [0, 1]$ be a continuous and increasing function on the interval $[0, 1]$, with $T(0) = 0$, $T(1) = 1$, such that $\forall F \in \mathcal{F}$, $T \circ F \in \mathcal{F}$. Denote by \mathcal{T} the set of such functions T .*

Proposition 2.1 (A sufficient regular condition) *If T is an absolutely monotonic function of order d on the interval $[0, 1]$, such that $T(0) = 0$ and $T(1) = 1$, then $T \in \mathcal{T}$.*

Proof: Let $F \in \mathcal{F}$. If T is absolutely monotonic function on the interval $[0, 1]$ then in particular T is continuous and increasing. So from Definition 3.6 (using the absolutely monotonic property) and 3.4 (using the continuous and increasing property) in Valdez and Xiao (2011) we obtain that $T \circ F \in \mathcal{F}$. \square

We remark that assumptions of Proposition 2.1 are sufficient, but not necessary.

2.1. External distortion

In this section we introduce the external distorted multivariate distribution function \tilde{F}^{ext} .

Definition 2.2 (Externally distorted distribution) *Let $F \in \mathcal{F}$ a d -dimensional distribution function that satisfies the regularity conditions. If $T \in \mathcal{T}$ we define the external distorted multivariate distribution function by $\tilde{F}^{ext}(x) = T \circ F(x)$ with $x \in \mathbb{R}_+^d$.*

Furthermore we introduce the following distortion for the upper α -level set of F .

Proposition 2.2 (External distortion for upper α -level set) *Let $\alpha \in (0, 1)$ and $L(\alpha) = \{x \in \mathbb{R}_+^d : F(x) \geq \alpha\}$. If $T \in \mathcal{T}$, it holds that:*

$$\tilde{L}^{ext}(\alpha) = \{x \in \mathbb{R}_+^d : \tilde{F}^{ext}(x) \geq \alpha\} = L(T^{-1}(\alpha)).$$

Proof: Since $\tilde{F}^{ext}(x) = T \circ F(x)$ and T is invertible in $[0, 1]$ then

$$\tilde{L}^{ext}(\alpha) = \{x \in \mathbb{R}_+^d : T \circ F(x) \geq \alpha\} = \{x \in \mathbb{R}_+^d : F(x) \geq T^{-1}(\alpha)\}.$$

Hence the result. \square

Remark 2 If T is absolutely monotonic so that it fulfils sufficient assumptions of Proposition 2.1, T^{-1} is a concave function. Then in particular $\alpha := T^{-1}(\beta) > \beta$, for $\beta \in (0, 1)$. This means the transformation T transforms the upper β -level set of F into the upper α -level set, with higher risk level α .

Furthermore, we remark that using an external distortion we are able to modify the dependence structure of the random vector X . Indeed, for simplicity in the case $d = 2$, we easily obtain the following property:

Remark 3 (Dependence impact of external distortion) Let X and Y be two independent real random variables (i.e., $F(x, y) = F_X(x)F_Y(y)$) and T be a distortion such that $T \in \mathcal{T}$. Let (\tilde{X}, \tilde{Y}) be a vector with distribution function $\tilde{F}^{ext}(x, y) = T \circ F(x, y)$. Since in general $T \circ (F_X(x)F_Y(y)) \neq (T \circ F_X(x)) \cdot (T \circ F_Y(y))$, \tilde{X} and \tilde{Y} are not necessarily independent.

2.2. Global distortion

We now include to the externally distorted distribution \tilde{F}^{ext} also the distortions of marginal components of F . In this last case we will finally obtain a global distorted distribution function \tilde{F} .

Definition 2.3 (Distorted distribution) Let F be a d -dimensional distribution function. Let $T \in \mathcal{T}$ and $T_i : [0, 1] \rightarrow [0, 1]$ be a continuous non-decreasing function, such that $T_i(0) = 0$, $T_i(1) = 1$, for $i = 1, \dots, d$. We introduce the global distorted distribution function \tilde{F} of F as :

$$\tilde{F}(x_1, \dots, x_d) = T \circ C(T_1^{-1}F_1(x_1), \dots, T_d^{-1}F_d(x_d)).$$

The interested reader is also referred to Valdez and Xiao (2011) (Definitions 3.2 and 3.6) and Charpentier (2008). They deal with the particular case $T = T_1 = \dots = T_d$.

Proposition 2.3 (Distorted margins) Denote by $\tilde{F}_1, \dots, \tilde{F}_d$ the marginal distributions of the joint distribution \tilde{F} . Then,

$$\tilde{F}_i = T \circ T_i^{-1} \circ F_i, \text{ for } i = 1, \dots, d.$$

In analogy with Proposition 2.2, we now analyze the impact of the global distortion on the α -level curves of F .

Proposition 2.4 (Distorted α -level curves) Let F be a d -dimensional distribution function and \tilde{F} be as in Definition 2.3. Let $\partial\tilde{L}(\alpha) = \{x \in \mathbb{R}_+^d : \tilde{F}(x) = \alpha\}$. A parametric expression for this distorted α -level curve is:

$$\partial\tilde{L}(\alpha) = \{(F_1^{-1} \circ T_1(u_1), \dots, F_d^{-1} \circ T_d(u_d)), (u_1, \dots, u_d) \in (0, 1)^d, C(u_1, \dots, u_d) = T^{-1}(\alpha)\}. \quad (1)$$

Or equivalently, as a function of marginal distorted distributions \tilde{F}_i , $i = 1, \dots, d$:

$$\partial\tilde{L}(\alpha) = \psi_T(\partial L(T^{-1}(\alpha))), \quad \forall \alpha \in (0, 1), \quad (2)$$

with $\psi_T(x) = (\tilde{F}_1^{-1} \circ T \circ F_1(x_1), \dots, \tilde{F}_d^{-1} \circ T \circ F_d(x_d))$, where the function $\psi_T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $x = (x_1, \dots, x_d) \in \mathbb{R}^d$.

Proof: It sufficient to remark that

$$\tilde{L}(\alpha) = \{(x_1, \dots, x_d) : C(F_1(x'_1), \dots, F_d(x'_d)) \geq T^{-1}(\alpha), F_1(x'_1) = T_1^{-1} \circ F_1(x_1), \dots, F_d(x'_d) = T_d^{-1} \circ F_d(x_d)\}.$$

Hence

$$\partial\tilde{L}(\alpha) = \psi_{T_1, \dots, T_d}(\partial L(T^{-1}(\alpha))), \quad \forall \alpha \in (0, 1), \quad (3)$$

where $\psi_{T_1, \dots, T_d} : \mathbb{R}^d \rightarrow \mathbb{R}^d$,

$$x \mapsto \psi_{T_1, \dots, T_d}(x) = (F_1^{-1} \circ T_1 \circ F_1(x_1), \dots, F_d^{-1} \circ T_d \circ F_d(x_d)),$$

with $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. For Equation (1) one can write:

$$\begin{aligned}\partial \tilde{L}(\alpha) &= \psi_{T_1, \dots, T_d}(\partial L(T^{-1}(\alpha))) \\ &= \{(F_1^{-1} \circ T_1 \circ F_1(x_1), \dots, F_d^{-1} \circ T_d \circ F_d(x_d)), (x_1, \dots, x_d) \in \partial L(T^{-1}(\alpha))\} \\ &= \{(F_1^{-1} \circ T_1 \circ F_1(x_1), \dots, F_d^{-1} \circ T_d \circ F_d(x_d)), C(F_1(x_1), \dots, F_d(x_d)) = T^{-1}(\alpha)\} \\ &= \{(F_1^{-1} \circ T_1(u_1), \dots, F_d^{-1} \circ T_d(u_d)), (u_1, \dots, u_d) \in (0, 1)^d, C(u_1, \dots, u_d) = T^{-1}(\alpha)\}.\end{aligned}$$

For Equation (2), the result comes down from Propositions 2.3 and Equation (3). \square

If initial copula C can be inverted, that is $\{(u_1, \dots, u_d) \in (0, 1)^d, C(u_1, \dots, u_d) = T^{-1}(\alpha)\}$ is given, and if F_1, \dots, F_d admit analytic expressions for inverse distributions then Equation (1) can be particularly useful to get an analytic expression for α -level curves.

Proposition 2.5 (Distorted copula) Denote $\tilde{C}(u_1, \dots, u_d)$ the distorted copula such that $\tilde{F}(x_1, \dots, x_d) = \tilde{C}(\tilde{F}_1(x_1), \dots, \tilde{F}_d(x_d))$. Then copula \tilde{C} only depends on external distortion T , i.e.

$$\tilde{C}(u_1, \dots, u_d) = T(C(T^{-1}(u_1), \dots, T^{-1}(u_d))).$$

Proof: The result comes down from Proposition 2.3.

Proposition 2.6 (Regular condition in bivariate case) Denote by C the initial copula. In the bivariate case, assume C is twice differentiable w.r.t. x and y . Write $C_x(x, y) = \frac{\partial}{\partial x} C(x, y)$, $C_y(x, y) = \frac{\partial}{\partial y} C(x, y)$ and $C_{xy}(x, y) = \frac{\partial^2}{\partial x \partial y} C(x, y)$. Assume T is a continuous and increasing function, and that T is differentiable with respective first and second order derivatives T' and T'' . Then $T \in \mathcal{T}$ if and only if:

$$T'(C(x, y))C_{xy}(x, y) + T''(C(x, y))C_x(x, y)C_y(x, y) \geq 0, \quad \text{for all } x, y \in [0, 1].$$

In particular, this condition is satisfied if T is absolutely monotonic of order 2. In the case $C(x, y) = xy$ and this condition becomes:

$$T'(z) + zT''(z) \geq 0, \quad \text{for all } z \in [0, 1]. \quad (4)$$

Proof: Write $\tilde{C}(u, v) = T(C(T^{-1}(u), T^{-1}(v)))$. Asking that $\frac{\partial}{\partial u} \frac{\partial}{\partial v} \tilde{C}(u, v) \geq 0$, the condition holds. \square

This last condition in Equation (4) obviously holds when T is absolutely monotonic of order 2, i.e. when $T'(x)$ and $T''(x)$ are positive functions of x . However, this condition also holds in some cases where $T(x)$ is a concave function of x . This theoretical aspect will be illustrated in Section 5.

3. Estimation

Here, we aim at estimating a target-level curve of a global distorted d -dimensional distribution function. Firstly we consider the estimation of the internal distortions. If target marginal distributions \tilde{F}_i , for $i = 1, \dots, d$, are given then the internal distortions can be easily deduced from the external one T .

3.1. Methodology

Estimation of internal distortions

As a consequence of Proposition 2.3, when target distorted marginal distributions \tilde{F}_i are known, and when the global distortion T is given, we have:

$$T_i = F_i \circ \tilde{F}_i^{-1} \circ T, \quad \text{for } i = 1, \dots, d. \quad (5)$$

It follows that the main problem in our procedure is to estimate the external distortion T .

Estimation of external distortion

We now aim at estimating the external distortion T . When margins are perfectly fitted to given targets \tilde{F}_i , $i = 1, \dots, d$, using Equation (2), the level-curves only depend on the external distortion T . For this reason we will denote in the following $\partial\tilde{L}_T(\beta)$ the distorted β level-curve:

$$\partial\tilde{L}_T(\beta) = \psi_T(\partial L(T^{-1}(\beta))), \quad \forall \beta \in (0, 1), \quad (6)$$

where the function $\psi_T : \mathbb{R}^d \rightarrow \mathbb{R}^d$,

$$x \mapsto \psi_T(x) = (\tilde{F}_1^{-1} \circ T \circ F_1(x_1), \dots, \tilde{F}_d^{-1} \circ T \circ F_d(x_d)),$$

with $x = (x_1, \dots, x_d) \in \mathbb{R}^d$.

Then, we can summarize some key-points for the estimation of the external distortion T :

- The distorted β level-curve $\partial\tilde{L}_T(\beta)$ only depends on the external distortion T .
- Finding a good distortion T is strongly related to a good choice of level $\alpha := T^{-1}(\beta)$.
- Then we need an indicator of the quality of a candidate level-curve.

Following these considerations, firstly we define a suitable proximity indicator (PI) (see Section 3.2), that quantifies if a candidate level curve is good or not. An algorithm with a practical fit of these distortions in the case of piecewise linear external distortion is given in Section 3.3.

3.2. A suitable proximity indicator (PI)

Let \mathbf{X}^* a d -dimensional random vector with distribution F^* that satisfies *regularity conditions*. Let $\partial L^*(\alpha)$ be the α -level curve of F^* . Frequently α -level curve $\partial L^*(\alpha)$ are difficult to obtain analytically. For this reason we will try to obtain quantities relying on F^* rather than ∂L^* .

In order to evaluate the proximity of a “candidate set” γ to the target-curve $\partial L^*(\alpha)$ we introduce the following proximity indicator:

$$PI_{F^*,\alpha}(\gamma) = \mathbb{E}_{\mathbf{X}^*}[(F^*(\mathbf{X}^*) - \alpha)^2 \mid \mathbf{X}^* \in \gamma]. \quad (7)$$

Trivially, if $\gamma \subseteq \partial L^*(\alpha)$, then $PI_{F^*,\alpha}(\gamma) = 0$. As an interesting property we point out that the integral expression of formula (7) only depends on the analytical expression of F^* and not on $\partial L^*(\alpha)$. For an illustration of the proximity indicator in (7) the interested reader is remanded to Example 1 below.

Example 1 (An explicit proximity indicator in a bivariate setting) Let uniform marginal distributions in $[0, 1]$ and the bivariate Farlie-Gumbel-Morgenstern copula structure, i.e., $F^*(x, y) = xy + \theta xy(1-x)(1-y)$, where $\theta \in [-1, 1]$ (see Nelsen, 1999, Example 3.12). Let the “candidate set” be $\gamma_\alpha = \{(x, y) \in [0, 1]^2 : xy = \alpha\}$, for $\alpha \in (0, 1)$. In this case we can obtain a closed formula for $PI_{F^*,\alpha}(\gamma)$, that depends only to $\theta \in [-1, 1]$ and $\alpha \in (0, 1)$.

Recall that Farlie-Gumbel-Morgenstern copula reduces to the independence case when $\theta = 0$, and it can only model relatively weak dependence in the two extreme cases ($\theta = \pm 1$) (e.g. Example 3.12, Nelsen, 1999). Then in particular $PI_{F^*,\alpha}(\gamma) = 0$, for $\theta = 0$. In this sense, Figure 1 shows that $PI_{F^*,\alpha}(\gamma)$ takes small values for θ close to zero ($\theta = 0.18$). Conversely for θ close to ± 1 ($\theta = -0.99$ and $\theta = 0.85$) we have a greater range of values of $PI_{F^*,\alpha}(\gamma)$. Furthermore we remark that $PI_{F^*,\alpha}(\gamma) = 0$ for $\alpha = 0$ and $\alpha = 1$, $PI_{F^*,\alpha}(\gamma) > 0$ for all $\alpha \in (0, 1)$ and it is a continuous function with respect to α (see Figure 1).

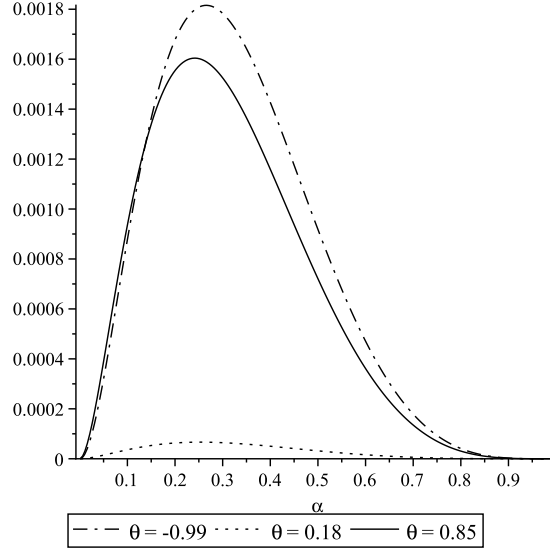


Figure 1: $PI_{F^*, \alpha}(\gamma)$, for $\theta = -0.99$ (dashed line), $\theta = 0.18$ (dotted line), $\theta = 0.85$ (full line), with $F^*(x, y) = xy + \theta xy(1-x)(1-y)$ and $\gamma = \{(x, y) \in [0, 1]^2 : xy = \alpha\}$.

In the case where we do not dispose of an explicit formula for the proximity indicator of a candidate set, we have to estimate it. If F^* is supposed to be unknown one can propose the estimated version of the proximity indicator in (7) as:

$$\widehat{PI}_{F^*, \alpha}(\gamma) = PI_{\widehat{F}^*, \alpha}(\gamma),$$

where \widehat{F}^* is a consistent smooth estimator of F^* admitting a density function. For instance given a d -dimensional sample $\{x_1^*, \dots, x_n^*\}$, with corresponding empirical distribution function F_n^* , the smooth estimator \widehat{F}^* could be $f_K * F_n^*$, where f_K is a suitable kernel density function (e.g. see Chacón and Rodríguez-Casal, 2010). Furthermore if γ is a discrete set of points, then one can write :

$$\widehat{PI}_{F^*, \alpha}(\gamma) = \frac{\sum_{i=1}^n \sum_{x \in \gamma} (\widehat{F}^*(x) - \alpha)^2 f_K\left(\frac{x-x_i^*}{h}\right)}{\sum_{i=1}^n \sum_{x \in \gamma} f_K\left(\frac{x-x_i^*}{h}\right)},$$

In order to build a discrete candidate set from parametric level curve, we introduce following notation that will be useful for numerical algorithms.

Definition 3.1 (Discretization of a level curve) We denote by $\partial L_g(\alpha)$ any finite subset of $\partial L(\alpha)$:

$$\partial L_g(\alpha) = \{(F_1^{-1}(\alpha_1), \dots, F_d^{-1}(\alpha_d)), (\alpha_1, \dots, \alpha_d) \in \mathcal{C}_{g, \alpha}\}.$$

where $\mathcal{C}_{g, \alpha}$ is a finite subset of $\{(\alpha_1, \dots, \alpha_d) \in (0, 1)^d : C(\alpha_1, \dots, \alpha_d) = \alpha\}$.

In the independence case, one can propose for example, with $g \in \mathbb{N}^*$:

$$\partial L_g(\alpha) = \{(F_1^{-1}(\alpha^{p_1}), \dots, F_d^{-1}(\alpha^{p_d})), (p_1, \dots, p_d) \in P_g\},$$

with $P_g = \left\{ \left(\frac{i_1}{g+1}, \dots, \frac{i_d}{g+1} \right), (i_1, \dots, i_d) \in \{1, \dots, g\}^d, i_1 + \dots + i_d = g + 1 \right\}.$

The set P_g is a discrete subset of the unit simplex in $(0, 1)^d$. Since components of its vectors are summing to one, this ensures $\partial L_g(\alpha) \subset \partial L(\alpha)$.

3.3. Estimation algorithm

Assume that T has a parametric form and belongs to the class of regular distortions \mathcal{T} . The best fitted distorted curve for level β can be defined as

$$\tilde{L}^{opt}(\beta) = \tilde{L}_{T^{opt}}(\beta), \text{ with } T^{opt} = \operatorname{arginf}_{T \in \mathcal{T}} PI_{F^*, \beta} [\tilde{L}_T(\beta)],$$

and using (6), we get

$$\partial \tilde{L}^{opt}(\beta) = \psi_T(\partial L(T^{-1}(\beta))), \text{ with } T = \operatorname{arginf}_{T \in \mathcal{T}} PI_{F^*, \beta} [\psi_T(\partial L(T^{-1}(\beta)))].$$

If T has a lot of parameters, the optimization procedure have to face a dimensionality problem (cf. Richard E. Bellman, *the curse of dimensionality*). Here, the problem comes from the fact that ψ_T depends on the searched distortion T .

Assume that at a step ν of the algorithm, we have a suitable approximation ψ_{T^ν} of ψ_T , where T^ν is a known distortion. One can define at this step ν the corresponding fitted level-curve:

$$\tilde{L}_\nu^{opt}(\beta) = \psi_{T^\nu}(\partial L(T^{-1}(\beta))), \text{ with } T = \operatorname{arginf}_{T \in \mathcal{T}} PI_{F^*, \beta} [\psi_{T^\nu}(\partial L(T^{-1}(\beta)))] \quad (8)$$

$$= \psi_{T^\nu}(\partial L(\alpha^\nu)), \text{ with } \alpha^\nu = \operatorname{arginf}_{a \in (0,1)} PI_{F^*, \beta} [\psi_{T^\nu}(\partial L(a))]. \quad (9)$$

The optimization in Equation (9) is far more easier to solve because it is a one-dimensional optimization, whereas optimization in Equation (8) is an optimization relying on all parameters of T .

This trick simply rely on the fact when ψ_{T^ν} is known, finding the best distortion T is equivalent to find the best level α^ν , simply by setting $\alpha^\nu = T^{-1}(\beta)$. As a consequence, the distortion T is then assumed to pass through the point (α^ν, β) .

As a summary, given chosen distorted levels β_j , we aim at finding optimal corresponding levels α_j , and thus the optimal distortion T passing through of all points (α_j, β_j) , $j = 1, \dots, m$. Such a distortion passing through all points of a set Ω will be written \bar{T}_Ω , as detailed in following definition.

Definition 3.2 (Piecewise linear distortion) *Let $\Omega = \{\omega_1, \dots, \omega_m\}$, $\omega_j \in (0, 1)^2$, $j = 1, \dots, m$, $m \in \mathbb{N}^*$. We denote by $\bar{T}_\Omega(x)$ a given nonparametric function linking point $(0, 0)$, points of Ω , and point $(1, 1)$. Typically, $\bar{T}_\Omega(x)$ can be piecewise linear, or piecewise linear after a change of scale.*

Starting from these considerations and Equation (9), we can now introduce the *procedure of estimation of $\bar{T}_\Omega(x)$* , detailed in Algorithm 1.

A necessary condition for this algorithm to be applied is that $\bar{T}_\Omega(x)$ is an increasing function for x , that is for all $i \in \{1, \dots, m\}$, $\alpha_i^\nu \geq \alpha_{i-1}^\nu$, where $\alpha_0^\nu = 0$ and $\nu = \nu_{\max}$. This algorithm is motivated by the following consideration. In the case where the levels of Ω_ν converge to the respective levels of a set Ω , then the distortion T_ν converges toward a distortion T .

To summarize, Algorithm 1 gives a fitted nonparametric external distortion \bar{T} , and corresponding fitted internal distortions T_1, \dots, T_d .

Algorithm 1 Fit of piecewise linear distortions

Let $0 < \beta_1 < \dots < \beta_m < 1$ be a set of levels.

Set $\Omega_0 = \{(\beta_j, \beta_j), j = 1, \dots, m\}$.

for ν varying from 1 to ν_{\max}

 Compute external fitted distortion of step ν , $T^\nu(x) = \bar{T}_{\Omega_{\nu-1}}(x)$.

for j varying from 1 to m

 Compute $\alpha_j^\nu = \operatorname{arginf}_{a \in (0,1)} \widehat{PI}_{F^*, \beta_j}[\gamma_\nu(a)]$, with $\gamma_\nu(a) = \psi_{T^\nu}(\partial L_g(a))$.

end for

 Set $\Omega_\nu = \{(\alpha_j^\nu, \beta_j)\}_{j \in \{1, \dots, m\}}$.

end for

Set $\Omega = \Omega_{\nu_{\max}}$.

Get external distortion $T = \bar{T}_\Omega$.

Get internal distortions $T_i = F_i \circ \tilde{F}_i^{-1} \circ T$, $i \in \{1, \dots, d\}$.

Remark 4 We summarize here some advantages of Algorithm 1:

1. Using this procedure our d -dimensional problem is decomposed in a sequence of *univariate optimization problems*. We just have to optimize separately m univariate parameters α_i , corresponding to the levels of the level-curves.
2. Once distortions \bar{T}, T_1, \dots, T_d estimated, one can provide analytical expressions for both \tilde{F} and \tilde{L} . This is an interesting aspect of the procedure above because, frequently, the analytical expression of \tilde{L} is quite difficult to obtain.
3. Algorithm 1 gives a non parametric estimation of \bar{T} and corresponding internal distortions T_1, \dots, T_d . Smoothed parametric version of these distortions will be easy to find (see Section 4.3).

In order to get smooth fitted distortions satisfying regular conditions, we propose hereafter some numerical applications using hyperbolic distortions.

4. Smooth estimation

4.1. A particular class of distortion functions

We take back from Bienvenüe and Rullière (2012) the following key notion of conversion function and distortion function :

Definition 4.1 (Conversion and distortion functions) *Let f any bijective increasing function from \mathbb{R} to \mathbb{R} . It is said to be a conversion function. Furthermore we define the function $T_f : [0, 1] \rightarrow [0, 1]$ such that*

$$T_f(u) = \begin{cases} 0 & \text{if } u = 0, \\ \operatorname{logit}^{-1}(f(\operatorname{logit}(u))) & \text{if } 0 < u < 1, \\ 1 & \text{if } u = 1. \end{cases}$$

Remark that the distortions function are chosen in a way to be easily invertible. In particular in a way such that $T_f \circ T_g = T_{f \circ g}$, $T_f^{-1} = T_{f^{-1}}$. These readily invertible distortions help simulation of the distorted distributions (see Bienvenüe and Rullière, 2012).

4.2. Hyperbolic conversion functions

In this section we consider this particular class of conversion function, which correspond to functions that are defined in Bienvenüe and Rullière (2012).

Definition 4.2 (A class of hyperbole) *The considered hyperbole H is*

$$H_{m,h,\rho_1,\rho_2,\eta}(x) = m - h + (e^{\rho_1} + e^{\rho_2})\frac{x - m - h}{2} - (e^{\rho_1} - e^{\rho_2})\sqrt{\left(\frac{x - m - h}{2}\right)^2 + e^{\eta - \frac{\rho_1 + \rho_2}{2}}}. \quad (10)$$

with $m, h, \rho_1, \rho_2 \in \mathbb{R}$, and one smoothing parameter $\eta \in \mathbb{R}$.

After some calculations, one can check that

$$H_{m,h,\rho_1,\rho_2,\eta}^{-1}(x) = H_{m,-h,-\rho_1,-\rho_2,\eta}(x).$$

In the following we illustrate with some examples the hyperbolic conversion functions proposed in Definition 4.2.

Example 2 (Global distortion: a bivariate case) We provide an illustration of a bivariate global distorted distribution $\tilde{F}(x, y)$ (see Definition 2.3). As internal distortion we consider the hyperbolic conversion functions proposed in Definition 4.2 with $m = 0.536, h = 0, \rho_1 = -\rho_2 = 0.321, \eta = 5$.

Let now $C(u, v) = uv$ and $F_1(x) = F_2(x) = 1 - e^{-x}$. Then $F(x, y) = (1 - e^{-x})(1 - e^{-y})$. Using Definition 2.3, we consider the bivariate distorted distribution function :

$$\tilde{F}(x, y) = T \circ C(T^{-1}F_1(x), T^{-1}F_2(y)),$$

where $T = T_f$, with $f = H_{m,h,\rho_1,\rho_2,\eta}$. We denote \tilde{f} the associated density function to \tilde{F} . For graphical illustration see Figure 2.

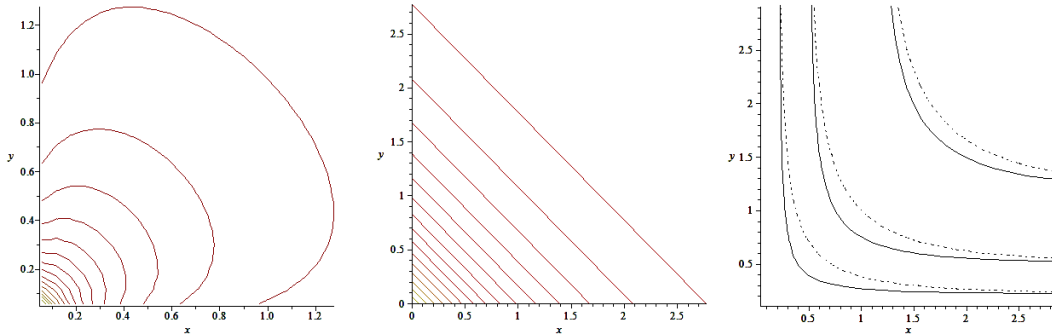


Figure 2: (Left) level curves of distorted density $\tilde{f}(x, y)$; (Center) level curves of non-distorted independent density $f(x, y) = e^{-x} e^{-y}$; (Right) level curves for $\alpha = 0.2, 0.4, 0.7$ of the distorted distribution $\tilde{F}(x, y)$ (full curves) and the non-distorted independent distribution (i.e., $F(x, y) = (1 - e^{-x})(1 - e^{-y})$) (dashed curves). Parameters: $m = 0.536, h = 0, \rho_1 = -\rho_2 = 0.321, \eta = 5$.

In the following we illustrate a global distortion in the case of a 3-dimensional distribution function.

Example 3 (Global distortion: a 3-dimensional case) We provide an illustration of a 3-dimensional global distorted distribution $\tilde{F}(x, y, z)$ (see Definition 2.3). We follow the same approach to the bivariate example above. As internal distortion we consider the hyperbolic conversion functions proposed in Definition 4.2 with $m = 0.5, h = 0, \rho_1 = -\rho_2 = 0.91, \eta = 3$. We remark that with this choice of parameters we deal with an absolutely monotonic external distortion, of order 3. Using Definition 2.3, we consider the 3-dimensional distorted distribution function :

$$\tilde{F}(x, y, z) = T \circ C(T^{-1}F_1(x), T^{-1}F_2(y), T^{-1}F_3(z)),$$

where $T = T_f$, with $f = H_{m,h,\rho_1,\rho_2,\eta}$. For graphical illustration see Figure 3.

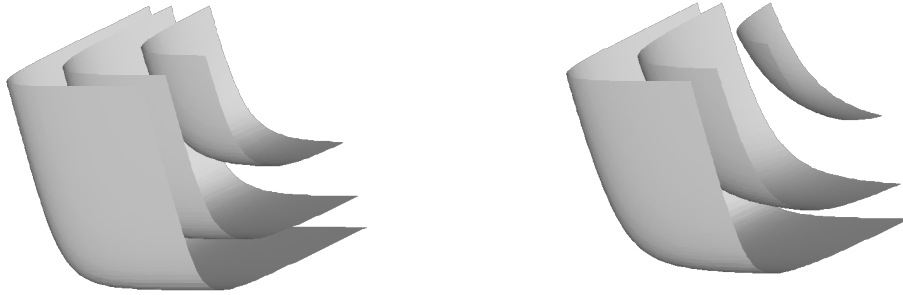


Figure 3: (Left) level curves for $\alpha = 0.1, 0.4, 0.7$ of the 3-dimensional distorted distribution $\tilde{F}(x, y, z)$, for parameters $m = 0.5, h = 0, \rho_1 = -\rho_2 = 0.91, \eta = 3$. (Right) level curves for $\alpha = 0.1, 0.4, 0.7$ of the 3-dimensional non-distorted independent distribution with exponential marginals with parameter 1.

4.3. Smooth estimation algorithm

We consider the generic hyperbolic conversion function defined in Equation (10). First remark that when the smoothing parameter η tends to $-\infty$, the hyperbole H tends to the angle function:

$$A_{m,h,\rho_1,\rho_2}(x) = m - h + (x - m - h) (e^{\rho_1} 1_{\{x < m+h\}} + e^{\rho_2} 1_{\{x > m+h\}}). \quad (11)$$

As remarked in Bienvenüe and Rullière (2012), it thus appears that hyperbolic distortions have the advantage of being smooth versions of angle functions. They show in their paper that initial parameters for the estimation are easy to obtain with angle compositions.

Definition 4.3 (Composite distortions) Let $k \in \mathbb{N}$. Consider $\eta \in \mathbb{R}$ and a given parameter vector $\theta = (m, h, \rho_1, \rho_2, a_1, r_1, \dots, a_k, r_k)$ if $k \geq 1$, or $\theta = (m, h, \rho_1, \rho_2)$ if $k = 0$. We define the angle composite distortion \mathcal{A}_θ as:

$$\mathcal{A}_\theta = T_{f_\theta}, \text{ with } f_\theta = \begin{cases} A_{a_k,0,0,r_k} \circ \dots \circ A_{a_1,0,0,r_1} \circ A_{m,h,\rho_1,\rho_2} & \text{if } k \geq 1, \\ A_{m,h,\rho_1,\rho_2} & \text{if } k = 0, \end{cases}$$

and the hyperbolic composite distortion $\mathcal{H}_{\theta,\eta}$ as:

$$\mathcal{H}_{\theta,\eta} = T_{f_{\theta,\eta}}, \text{ with } f_{\theta,\eta} = \begin{cases} H_{a_k,0,0,r_k,\eta} \circ \dots \circ H_{a_1,0,0,r_1,\eta} \circ H_{m,h,\rho_1,\rho_2,\eta} & \text{if } k \geq 1, \\ H_{m,h,\rho_1,\rho_2,\eta} & \text{if } k = 0, \end{cases}$$

where A_{m,h,ρ_1,ρ_2} is given in Equation (11), and where $H_{m,h,\rho_1,\rho_2,\eta}$ is defined in Definition 4.2.

Remark 5 (Inverse composite distortions) Let $k \in \mathbb{N}$. Consider $\eta \in \mathbb{R}$ and a given parameter vector $\theta = (m, h, \rho_1, \rho_2, a_1, r_1, \dots, a_k, r_k)$ if $k \geq 1$, or $\theta = (m, h, \rho_1, \rho_2)$ if $k = 0$. Since $T_f^{-1} = T_{f^{-1}}$, the angle composite distortion \mathcal{A}_θ^{-1} is such that:

$$\mathcal{A}_\theta^{-1} = T_{f_\theta}, \text{ with } f_\theta = \begin{cases} A_{m,-h,-\rho_1,-\rho_2} \circ A_{a_1,0,0,-r_1} \circ \dots \circ A_{a_k,0,0,-r_k} & \text{if } k \geq 1, \\ A_{m,-h,-\rho_1,-\rho_2} & \text{if } k = 0. \end{cases}$$

The hyperbolic inverse composite distortion $\mathcal{H}_{\theta,\eta}$ is such that:

$$\mathcal{H}_{\theta,\eta}^{-1} = T_{f_{\theta,\eta}}, \text{ with } f_{\theta,\eta} = \begin{cases} H_{m,-h,-\rho_1,-\rho_2,\eta} \circ H_{a_1,0,0,-r_1,\eta} \circ \dots \circ H_{a_k,0,0,-r_k,\eta} & \text{if } k \geq 1, \\ H_{m,-h,-\rho_1,-\rho_2,\eta} & \text{if } k = 0, \end{cases}$$

Definition 4.4 (Suited parameters from Ω) Let $k \in \mathbb{N}$. Consider one given set $\Omega = \{\omega_1, \dots, \omega_{3+k}\}$, $\omega_j \in (0, 1)^2$. Denote by u_j and v_j the two respective components of each ω_j in the logit scale, such that $\omega_j = (\text{logit}^{-1}u_j, \text{logit}^{-1}v_j)$, $j \in \{1, \dots, 3+k\}$. Assume that u_j and v_j are increasing sequences of j . We define:

$$\Theta(\Omega) = \begin{cases} (m, h, \rho_1, \rho_2, a_1, r_1, \dots, a_k, r_k) & \text{if } k \geq 1, \\ (m, h, \rho_1, \rho_2) & \text{if } k = 0. \end{cases}$$

where $m = \frac{u_2+v_2}{2}$, $h = \frac{u_2-v_2}{2}$, $\rho_1 = \ln\left(\frac{v_2-v_1}{u_2-u_1}\right)$, $\rho_2 = \ln\left(\frac{v_3-v_2}{u_3-u_2}\right)$, $r_k = \ln\left(\frac{v_{3+k}-v_{2+k}}{u_{3+k}-u_{2+k}} \frac{u_{2+k}-u_{1+k}}{v_{2+k}-v_{1+k}}\right)$, $a_k = v_{2+k}$, $k \geq 1$.

Proposition 4.1 (Suited composite distortions) Let $k \in \mathbb{N}$. Consider one given set $\Omega = \{\omega_1, \dots, \omega_{3+k}\}$, $\omega_j \in (0, 1)^2$ and a smoothing parameter $\eta \in \mathbb{R}$. Set $\theta = \Theta(\Omega)$, then

- the distortion $\mathcal{A}_\theta(x)$ is piecewise linear in the logit scale and will be called *logit-piecewise linear*. It links point $(0, 0)$, points of Ω , and point $(1, 1)$, so that it fulfils conditions of Definition 3.2.
- the distortion $\mathcal{H}_{\theta,\eta}$ converges pointwise to \mathcal{A}_θ as η tends to $-\infty$. It results that the continuous and differentiable distortion $\mathcal{H}_{\theta,\eta}$ can fit as precisely as desired the set of points Ω when η tends to $-\infty$.

Proof: The first result is proved in Bienvenüe and Rullière (2012). It simply comes from the fact that $\mathcal{A}_{\Theta(\Omega)}(u_j) = v_j$ for all $j \in \{1, \dots, 3+k\}$, where $\omega_j = (\text{logit}^{-1}u_j, \text{logit}^{-1}v_j)$. The convergence of the hyperbole composite distortion toward the angle composite distortion is straightforward and also evoked in Bienvenüe and Rullière (2012). \square

Using Algorithm 1, one can find a piecewise or logit-piecewise linear distortion for the external distortion and corresponding set Ω . Furthermore, using Proposition 2.3, one can build the corresponding internal distortions and easily propose suited piecewise or logit-piecewise linear internal distortions and corresponding sets Ω_i , for $i = 1, \dots, d$.

The corresponding parameters of these sets Ω and $\Omega_1, \dots, \Omega_d$ are given by Definition 4.4. These parameters $\theta = \Theta(\Omega)$ and $\theta_1 = \Theta(\Omega_1), \dots, \theta_d = \Theta(\Omega_d)$ constitute initial values for distortions of smoothed hyperbolic external and internal distortions. Choosing a convenient smoothing parameter η , one can define a complete vector parameter:

$$\vec{\Theta} = (\theta_1, \dots, \theta_d, \theta, \eta).$$

As seen previously, from now, the estimation of $\vec{\Theta}$ relies exclusively on univariate optimizations.

From estimated vector $\vec{\Theta}$, the corresponding estimated distribution is:

$$\tilde{F}_{\vec{\Theta}}(x_1, \dots, x_d) = \mathcal{H}_{\theta, \eta} \circ C(\mathcal{H}_{\theta_1, \eta}^{-1} \circ F_1(x_1), \dots, \mathcal{H}_{\theta_d, \eta}^{-1} \circ F_d(x_d)). \quad (12)$$

and the parametric α level-curves are given by Proposition 2.4

$$\partial \tilde{L}_{\vec{\Theta}}(\alpha) = \{(F_1^{-1} \circ \mathcal{H}_{\theta_1, \eta}(u_1), \dots, F_d^{-1} \circ \mathcal{H}_{\theta_d, \eta}(u_d)), (u_1, \dots, u_d) \in (0, 1)^d, C(u_1, \dots, u_d) = \mathcal{H}_{\theta, \eta}^{-1}(\alpha)\}. \quad (13)$$

Since initial parameters of $\vec{\Theta}$ are usually close to optimal values, all these parameters can be improved using standard local optimization algorithm, like gradient descent methods:

$$\vec{\Theta}^* = \operatorname{argmax}_{\vec{\Theta}} \ln \mathcal{L}(\vec{\Theta}), \quad (14)$$

with \mathcal{L} the likelihood on the considered data.

Remark 6 Since these initial values are usually close to optimal values, the numerical optimization easily converges. It is important to notice that this optimization could be done directly in theory. However in practice, starting from piecewise linear distortions is necessary to ensure that the optimization converges.

In the following we detail the algorithm-procedure to get smoothed version of distortions obtained in Algorithm 1.

Algorithm 2 Smooth estimation algorithm

Piecewise linear external distortion

Let $0 < \beta_1 < \dots < \beta_m < 1$ be a set of levels,

Obtain distortion levels Ω from Algorithm 1,

Obtain suited parameters $\theta = \Theta(\Omega)$ from Definition 4.4 and parametric distortion $T = \mathcal{A}_\theta$.

Piecewise linear internal distortions

for i varying from 1 to d

Let $0 < \alpha_1^{(i)} < \dots < \alpha_{m_i}^{(i)} < 1$ be a set of levels associated to the i th-marginal F_i ,

Obtain distortion levels $\Omega_i = \left\{ (\alpha_j^{(i)}, T_i(\alpha_j^{(i)})) \right\}_{j=1, \dots, m_i}$, from Equation (5),

Obtain suited parameters $\theta_i = \Theta(\Omega_i)$ from Definition 4.4 and distortions $T_i = \mathcal{A}_{\theta_i}$.

end for

Smoothed distortions

Choose a smoothing parameter η ,

Improve all parameters including η by Equation (14),

Check validity condition $T \in \mathcal{T}$,

Get parametric expression for \tilde{F} and $\partial \tilde{L}$ by Equations (12) and (13).

4.4. Refinements

In some particular numerical cases, the choice of the input values of the algorithms may be difficult to do. This section presents some propositions to ease this choice.

Choice of initial levels: Concerning the choice of initial levels β_j , a problem is that if T is very concave or convex, then levels $\alpha_j = T^{-1}(\beta_j)$ may be concentrated on a small interval of $(0, 1)$, e.g. $(0.90, 1)$, even for values of β_j spread quite uniformly on $(0, 1)$.

Proposed solution: One can modify levels β_j in order to ensure that both α_j and β_j are reasonably spread over the whole interval $(0, 1)$. Let Δ be the diagonal of equation $y = x$. Assuming that the projection of (α_j, β_j) on Δ is (δ_j, δ_j) , one gets $\beta_j = -\alpha_j + 2\delta_j$. Given a distortion T and a set of levels $0 < \delta_1 < \dots < \delta_m < 1$, e.g. $\delta_j = j/(m+1)$, one can define

$$\begin{cases} \alpha_j(T) &= a \text{ such that } T(a) = -a + 2\delta_j, \\ \beta_j(T) &= -\alpha_j(T) + 2\delta_j. \end{cases}$$

Replacing initial levels β_j by $\beta_j(T^\nu)$ at the beginning of each step ν in Algorithm 1 allows to get a better distribution of elements of the final set Ω into $(0, 1)^2$. Similar considerations can be applied to each internal distortion T_i : initial levels $\alpha_j^{(i)}$ can be replaced by $\alpha_j(T_i)$ in Algorithm 2, for $j = 1, \dots, m_i$.

Choice of initial smoothing parameter: Consider a given angle function A_{m,h,ρ_1,ρ_2} with apex in $(m+h, m-h)$. Replacing this angle by a smooth hyperbole $H_{m,h,\rho_1,\rho_2,\eta}$ causes some problems. First, the hyperbole is not passing through the point $(m+h, m-h)$ as the angle function. Second, when η is large, the hyperbole gets far from the angle function. In summary, the angle is set to pass through a well chosen point $(m+h, m-h)$, but the hyperbole gets far from this point when smoothing parameter η increases. In numerical illustrations, the optimization of parameter η thus leads to small values of this parameter, high absolute derivatives of distribution functions at some points, and an insufficient smoothing of the final distortions.

Proposed solution: Define a new hyperbole H^* as

$$H_{m,h,\rho_1,\rho_2,\eta}^* = H_{m,h+\delta(\rho_1,\rho_2,\eta),\rho_1,\rho_2,\eta}$$

with $\delta(\rho_1, \rho_2, \eta) = \frac{e^{\rho_2} - e^{\rho_1}}{\sqrt{(1+e^{\rho_1})(1+e^{\rho_2})}} e^{-\frac{\rho_1+\rho_2}{4}} e^{\eta/2}$.

One can check after some calculations that:

- $H_{m,h,\rho_1,\rho_2,\eta}^*(m+h) = m-h$, so that the new hyperbole H^* is passing through the point $(m+h, m-h)$.
- $\delta(-\rho_1, -\rho_2, \eta) = -\delta(\rho_1, \rho_2, \eta)$, so that $H_{m,h,\rho_1,\rho_2,\eta}^{*-1} = H_{m,-h,-\rho_1,-\rho_2,\eta}^*$.

Given a smoothing parameter η , the proposed improvement is to choose $H_{m,h,\rho_1,\rho_2,\eta}^*$ as a smooth version of the angle A_{m,h,ρ_1,ρ_2} , instead of $H_{m,h,\rho_1,\rho_2,\eta}$. Corresponding distortion composition will be denoted \mathcal{H}^* instead of \mathcal{H} . Using \mathcal{H}^* instead of \mathcal{H} , an univariate maximization of the log-likelihood on the data leads to a good choice of the parameter η .

Choice of interpolation function \bar{T}_Ω : Piecewise linear distortion \bar{T}_Ω is used in Algorithm 1. It gives interpolations of T , when T is passing through all points of Ω . The most simple solution is to choose a simple linear interpolation of points of Ω . The problem is that when T is very concave or convex, a simple linear interpolation may be quite far from the real value of T .

Proposed solution: The proposed solution consists in replacing \bar{T}_Ω by $\mathcal{A}_{\Theta(\Omega)}$, or eventually by a smoothed version $\mathcal{H}_{\Theta(\Omega),\eta}^*$ for a given parameter η .

5. Numerical applications

5.1. A simple bivariate model

In the following numerical illustrations we consider a very basic independent exponential model where:

$$F(x_1, x_2) = F_1(x_1)F_2(x_2),$$

with $F_1(x) = F_2(x) = 1 - e^{-x}$, $x \in \mathbb{R}^+$. We aim at distorting this basic bivariate distribution in order to fit real data. Once distortions T, T_1, \dots, T_d estimated, we obtain analytical expressions for \tilde{F} and $\partial\tilde{L}(\alpha)$:

$$\tilde{F}(x_1, x_2) = T(T_1^{-1}(1 - e^{-x_1}) \cdot T_2^{-1}(1 - e^{-x_2})), \quad \text{for } x_1, x_2 \geq 0,$$

and for any $\alpha \in (0, 1)$, from Proposition 2.4,

$$\partial\tilde{L}(\alpha) = \left\{ \left(-\log(1 - T_1(u)), -\log\left(1 - T_2\left(\frac{T^{-1}(\alpha)}{u}\right)\right) \right), u \in (T^{-1}(\alpha), 1) \right\}. \quad (15)$$

Distorted marginals in this simple bivariate model are given by:

$$\tilde{F}_1(x) = T \circ T_1^{-1}(1 - e^{-x}) \quad \text{and} \quad \tilde{F}_2(x) = T \circ T_2^{-1}(1 - e^{-x}),$$

and corresponding distorted copula is

$$\tilde{C}(u, v) = T(T^{-1}(u) \cdot T^{-1}(v)).$$

In this independent case the regular condition of Proposition 2.6 becomes:

$$T'(x) + xT''(x) \geq 0, \quad \text{for any } x \in [0, 1]. \quad (16)$$

This simple initial independent exponential multivariate distribution is very far from the data that will be considered. It has been chosen to demonstrate the ability of the estimation procedure to fit real data even when starting with a very poor initial distribution.

5.2. Real case study: Loss-ALAE data

In this section we present a real case for which we illustrate the behavior of our algorithm, presented in Sections 3 and 4. We consider the uncensored **Loss-ALAE data** in the logarithmic scale (for details see Frees and Valdez, 1998). The data size is $n = 1500$. Each claim consists of an indemnity payment (the loss, X) and an allocated loss adjustment expense (ALAE, Y). Examples of ALAE are the fees paid to outside attorneys, experts, and investigators used to defend claims. We now implement Algorithm 1 and 2 on Loss-ALAE data.

External distortion T :

From Algorithm 1 we obtain the piecewise linear external distortion $\bar{T}_\Omega(x)$, for $x \in [0, 1]$ (see Figure 4, black line). Furthermore, from set Ω , by applying Proposition 4.4, the suited angle composition-distortion is:

$$\mathcal{A}_\theta, \quad \text{with } \theta = \Theta(\Omega),$$

with the following vector of parameters $\theta = (m, h, \rho_1, \rho_2)$:

Parameters	m	h	ρ_1	ρ_2
θ	-0.541	-0.551	-0.467	-0.146

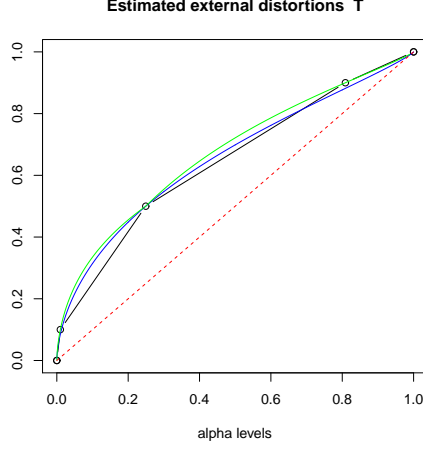


Figure 4: $T(x) = \bar{T}_\Omega(x)$ (black line) and $\mathcal{A}_\theta(x)$ (green line), $\mathcal{H}_{\theta,\eta}^*(x)$ (blue line), with $\eta = 0.18$. Red line is bisectrix of the first quadrant. We get $\Omega = \{(0.00987, 0.1), (0.24935, 0.5), (0.80883, 0.9)\}$

For a graphical representation of \mathcal{A}_θ see Figure 4, green line.

We check that despite the concavity of T , it satisfies the regular condition of Equation (16) on points such that T is differentiable.

Internal distortions T_i :

From Proposition 4.4 we obtain that the estimated marginal distortions T_1 and T_2 can be represented using angle compositions $T_1 = \mathcal{A}_{\theta_1}$ and $T_2 = \mathcal{A}_{\theta_2}$ with associated parameters:

Parameters	m	h	ρ_1	ρ_2
θ_1	3.901	-5.297	-0.524	-0.278
θ_2	3.457	-4.961	-0.581	-0.511

Remark that $T_1^{-1} = \mathcal{A}_{\theta_1}^{-1}$ and $T_2^{-1} = \mathcal{A}_{\theta_2}^{-1}$ can be directly obtained by Remark 5.

Including smoothing parameters:

We now include smoothing hyperbolic distortions indeed of angle compositions. For these data we use the hyperbolic function:

$$H_{m,h,\rho_1,\rho_2,\eta}^* = H_{m,h+\delta(\rho_1,\rho_2,\eta),\rho_1,\rho_2,\eta} \text{ with } \delta(\rho_1,\rho_2,\eta) = \frac{e^{\rho_2} - e^{\rho_1}}{\sqrt{(1+e^{\rho_1})(1+e^{\rho_2})}} e^{-\frac{\rho_1+\rho_2}{4}} e^{\eta/2}, \quad (17)$$

and $H_{m,h,\rho_1,\rho_2,\eta}$ as in (10). In this case we choose a smoothing parameter for the external distortion (i.e. η) and two smoothing parameters the internal ones (i.e. η_1 and η_2). Using the global optimization presented in Equation (14) we get: $\eta = 0.18, \eta_1 = \eta_2 = -2$. This value was chosen as the initial value for a global maximization with respect to the whole vector $\vec{\Theta}$ (see Equation (14)). In this case we obtain:

Parameters	m	h	ρ_1	ρ_2	η
θ	-0.549	-0.55	-0.471	-0.145	0.18
θ_1	3.912	-5.298	-0.525	-0.268	-2
θ_2	3.458	-4.974	-0.594	-0.544	-2

Furthermore, from Proposition 2.3, we get $\tilde{F}_i = T \circ T_i^{-1} \circ F_i$, for $i = 1, 2$, where T, T_i^{-1} are obtained above and F_i are the known initial marginals (in this case $F_1(x) = F_2(x) = 1 - e^{-x}$). The results are drawn in Figure 5 below.

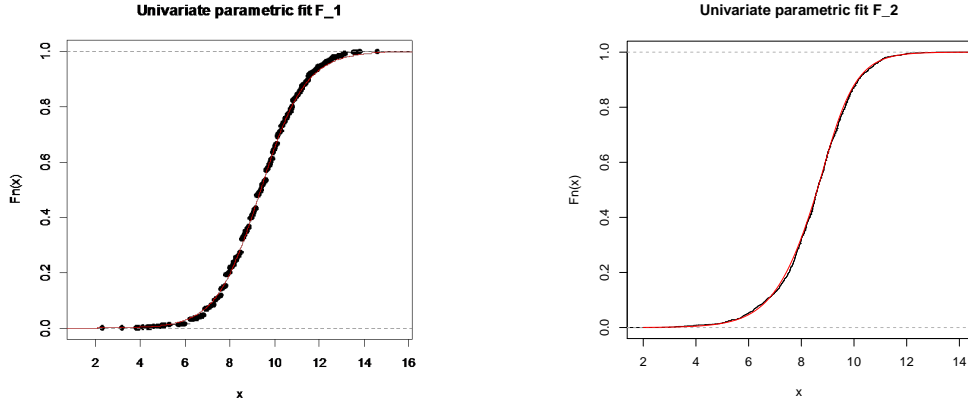


Figure 5: (Left) \tilde{F}_1 (red) and the empirical distribution function of Loss data (black). (Right) \tilde{F}_2 (red) and the empirical distribution function of ALAE data (black).

We check that the smoothed distortion T satisfies the regular condition of Equation (16) on all points of $(0, 1)$. For a graphical representation of the smoothed distortion T , i.e. $\mathcal{H}_{\theta, \eta}^*$, with $\eta = 0.18$, see Figure 4, blue line. The log-likelihood of this model on the Loss-ALAE data is -21804.01 . In the literature the fit of Loss-Alae data has received some attention. For instance Frees and Valdez (1998) fit on these data a Gumbel-Hougaard copula with parameter $\theta = 1.453$; Loss \sim Pareto with parameter $(14.036, 1.122)$ and ALAE \sim Pareto with parameter $(14.219, 2.118)$. The obtained log-likelihood on the Loss-Alae data of the Frees and Valdez's model is : -49075.82 . Furthermore in Klugman and Parsa (1999), they propose a fitted model on Loss-Alae data such that the log-likelihood is -31767.9 . Firstly, we remark that these results belong to the same order of magnitude. However, our model seems best fit the considered data.

In Figure 6, we compare our method with the empirical level curves defined by $\{(x, y) \in \mathbb{R}_+^2 : F_n(x, y) = \alpha\}$, where F_n is the bivariate empirical estimator of F , for $\alpha = 0.5, 0.8$. We have drawn the level curves $\partial \tilde{L}_{\mathcal{A}_\theta}(\alpha)$ obtained using a composition of angles (as described by Algorithm 1) and $\partial \tilde{L}_{\mathcal{H}_{\theta, \eta}^*}(\alpha)$, using smoothed hyperbolic external and internal distortions in (17), (as described by Algorithm 2), see Equation (6). Note that the smoothing parameter η is not really crucial with respect to the quality of our estimation, even if it improves the estimation of the level-curves. However it is important in order to obtain differentiable representations of level curves $\partial \tilde{L}(\alpha)$. In Figure 7(left) we draw the final distorted distribution $\tilde{F}(x_1, x_2)$ and its level-curves; in Figure 7(right) we present the level-curves of distorted density function $\tilde{f}(x_1, x_2)$. In particular Figure 7(right) illustrates the good quality of the fit of proposed model on the bivariate Loss-ALAE data.

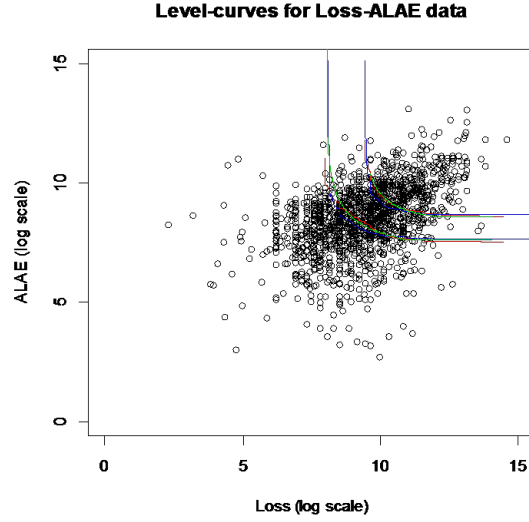


Figure 6: ALAE versus Loss data (in logarithmic scale). $\partial\tilde{L}_{\mathcal{A}_\theta}(\alpha)$ (red curves); $\partial\tilde{L}_{\mathcal{H}_{\theta,\eta}^*}(\alpha)$ (green curves); Empirical level curves $\{(x, y) \in \mathbb{R}_+^2 : F_n(x, y) = \alpha\}$ (blue curves); for $\alpha = 0.5, 0.8$.

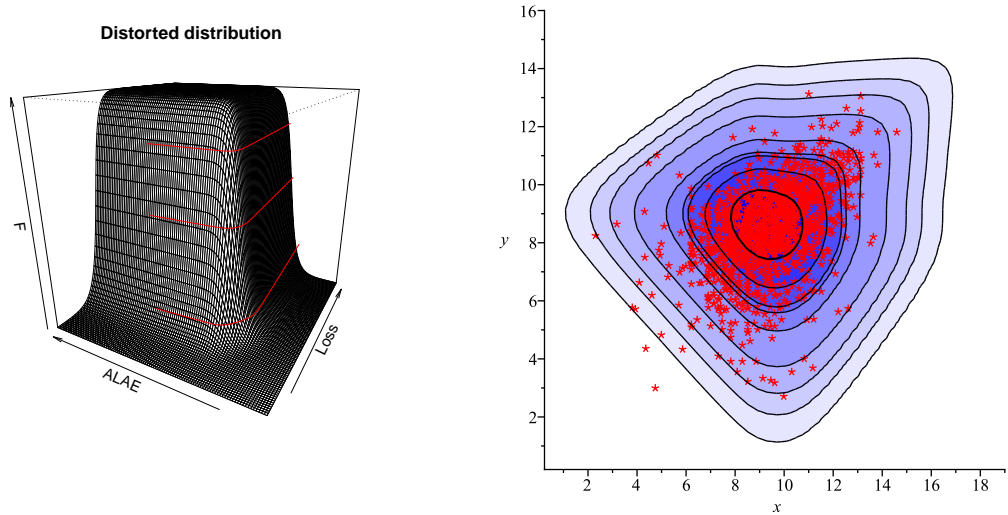


Figure 7: Distorted distribution $\tilde{F}(x_1, x_2)$ whit associated distorted level curves (red curves) (left); Level curves of distorted density $\tilde{f}(x_1, x_2)$ and Loss-ALAE data (red points) (right).

5.3. Real case study: Old Faithful Geyser Data

In this section we study a second real case that presents an interesting bimodal behavior. This data concern 272 eruptions of the Old Faithful geyser in Yellowstone National Park. Each observation consists of two measurements: the duration (in min) of the eruption (X), and the waiting time (in min) before the next eruption (Y). This waiting time has been divided by 20 in order to get the same order of magnitude on both axis. The interested reader is referred for instance to Obereder *et al.* (2007), Biernacki *et al.*

(2007). We following the same estimation procedure presented in Section 5.2 above, for Loss-Alae data. From Algorithm 1 we obtain the piecewise linear external distortion $\bar{T}_\Omega(x)$, for $x \in [0, 1]$ (see Figure 8, black line) and the suited angle composition-distortion is:

$$A_{a_4,0,0,r_4} \circ A_{a_3,0,0,r_3} \circ A_{a_2,0,0,r_2} \circ A_{a_1,0,0,r_1} \circ A_{m,h,\rho_1,\rho_2},$$

with $\theta = \Theta(\Omega)$ the following vector of parameters:

Parameters	m	h	ρ_1	ρ_2	a_1	a_2	a_3	a_4	r_1	r_2	r_3	r_4
θ	-6.081	5.328	0.863	2.061	-3.915	-2.251	-0.754	0	-1.614	-1.216	-0.391	0.661

For a graphical representation of \mathcal{A}_θ see Figure 8, green line. The estimated marginal distortions T_1 and T_2 can be represented using angle compositions $T_1 = \mathcal{A}_{\theta_1}$ and $T_2 = \mathcal{A}_{\theta_2}$ with associated parameters:

Parameters	m	h	ρ_1	ρ_2	a_1	a_2	a_3	a_4	r_1	r_2	r_3	r_4
θ_1	2.526	0.651	-2.213	-1.169	2.314	2.317	3.808	4.556	0.115	1.039	-1.474	-0.785
θ_2	2.509	-0.067	-1.683	-2.011	2.631	2.665	2.966	3.623	0.215	0.353	0.491	-1.014

As in the Loss-ALAE model (see Section 5.2) we now include the smoothing hyperbolic distortions indeed of angle compositions. For these data we use the hyperbolic function in Equation (17). Using the global optimization presented in Equation (14) we get: $\eta = 1, \eta_1 = \eta_2 = -7.5$. The log-likelihood of our final smoothed model on the Old Faithful geyser data is -1072 . In the literature the fit of Old Faithful geyser data has received some attention (see for instance Obereder *et al.*, 2007). Biernacki *et al.* (2007) fit on these data a Gaussian mixture model. In this case, they obtain a log-likelihood equal to -1124 . As in Section 5.2, we remark that these results belong to the same order of magnitude. However, our model seems best fit the considered data. For a graphical representation of the smoothed distortion T , i.e. $\mathcal{H}_{\theta,\eta}^*$, with $\eta = 1$, see Figure 8, blue line.

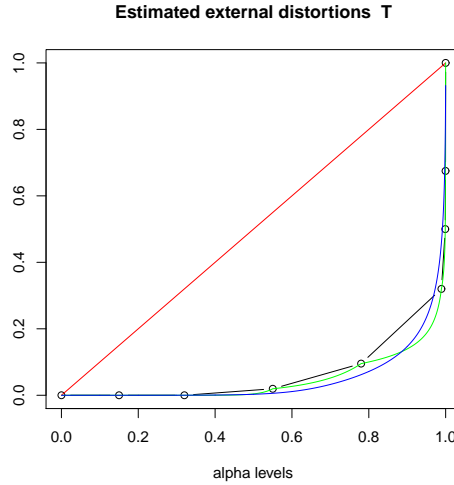


Figure 8: $T(x) = \bar{T}_\Omega(x)$ (black line), $\mathcal{A}_\theta(x)$ (green line), $\mathcal{H}_{\theta,\eta}^*(x)$ (blue line). Red line is bisectrix of the first quadrant.

Furthermore, from Proposition 2.3, we get $\tilde{F}_i = T \circ T_i^{-1} \circ F_i$, for $i = 1, 2$, where T, T_i^{-1} are obtained above and F_i are the known initial marginals (in this case $F_1(x) = F_2(x) = 1 - e^{-x}$). The results are drawn in Figure 9 below.

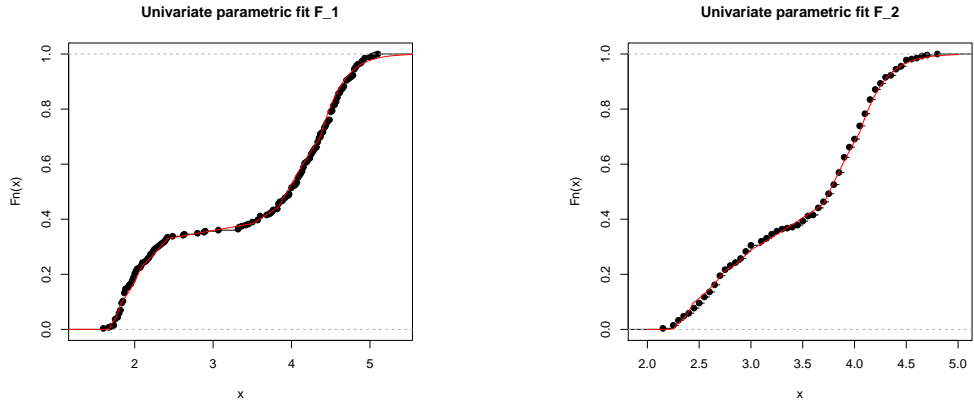


Figure 9: (Left) \tilde{F}_1 (red) and the empirical distribution function of the duration (in min) of the eruption data (black). (Right) \tilde{F}_2 (red) and the empirical distribution function of the waiting time (in min) data (black).

Finally, in Figure 10 (left) we draw the final distorted distribution $\tilde{F}(x_1, x_2)$ and its level-curves; in Figure 10 (right) we present the level-curves of distorted density function $\tilde{f}(x_1, x_2)$. In particular Figure 10 (right) illustrates the good quality of the fit of proposed model on the bivariate Old Faithful geyser data. We remark how our distortions can be able to capture the bimodal behavior of these data (see Figure 10, right). Despite its observable complexity, the proposed parametric fit also allows to get a relatively simple parametric expression for level curves, using Equation (15) (see Figure 10, left).

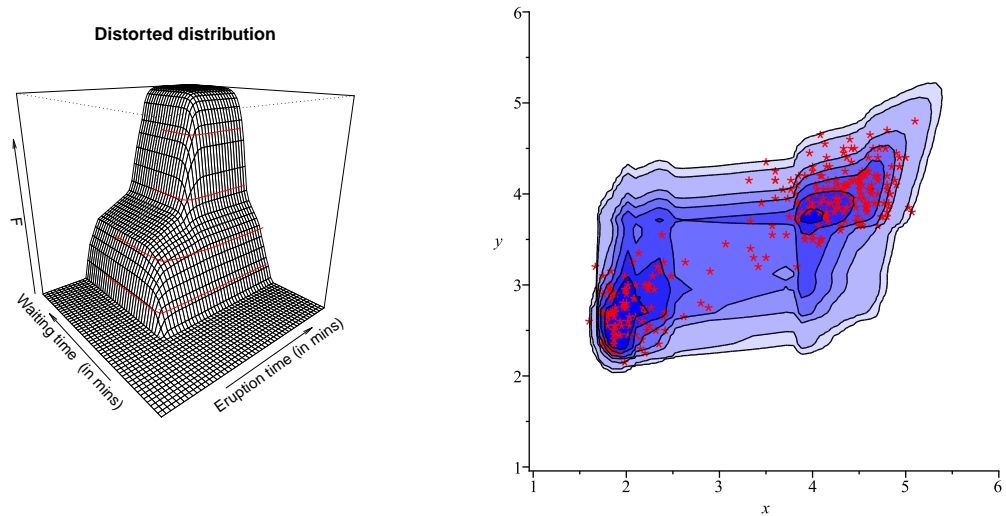


Figure 10: Distorted distribution $\tilde{F}(x_1, x_2)$ with associated distorted level curves (red curves) (left); Level curves of distorted density $\tilde{f}(x_1, x_2)$ and Old Faithful geyser data (red points) (right).

Conclusion

Probability distortions allow to build new classes of multivariate distribution functions. Impacts of such distortions on *multivariate level curves* have been studied. Among interesting properties, we have seen that external distortions are only acting on the *univariate* level of a level curve. Finding the best distorted levels on a particular data directly gives external distortion point values, and only requires *univariate* optimizations. Following this idea, we have proposed two algorithms: Algorithm 1 gives estimated piecewise linear distortions to fit a given data, and Algorithm 2 gives smoothed version of these distortions by using hyperbolic compositions in the logit scale.

The theoretical properties of considered distortions are thus very helpful for fitting distorted multivariate distributions. Applications on real data have emphasis the advantages of this methodology. First, despite the variable and possibly high number of parameters, the estimation is straightforward and only relying on univariate optimizations. Second, the methodology benefit from the theoretical (and practical) possibility to fit marginal distributions as precisely as desired. Third, it leads to parametric expressions of both cumulative distribution functions and level curves, which can be useful for risk measures. Fourth, numerical results on real data easily lead to a better likelihood than other parametric representations of the same data in some recent studies.

Some interesting perspectives would be a better characterization of necessary and sufficient conditions for regular external distortions, some work on the optimal initial levels to be chosen for the proposed algorithms, and on how to choose or reduce the number of parameters. The use of the parametric level curves for risk measures also opens a large research field.

Acknowledgements: This work has been partially supported by the French National Research Agency (ANR) under the reference ANR-08BLAN-0314-01. Part of this work also benefit from the support of the MIRACCLE-GICC project.

References

References

- [1] G. Awanou, M.J. Lai, and P. Wenston. The multivariate spline method for scattered data fitting and numerical solutions of partial differential equations. *Wavelets and Splines: Athens*, pages 24–74, 2005.
- [2] F. Belzunce, A. Castaño, A. Olvera-Cervantes, and A. Suárez-Llorens. Quantile curves and dependence structure for bivariate distributions. *Computational Statistics & Data Analysis*, 51(10):5112–5129, 2007.
- [3] A. Bienvenüe and D. Rullière. Iterative adjustment of survival functions by composed probability distortions. *The Geneva Risk and Insurance Review*, 37(2):156–179, 2011.
- [4] Alexis Bienvenüe and Didier Rullière. On hyperbolic iterated distortions for the adjustment of survival functions. In Cira Perna and Marilena Sibillo, editors, *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pages 35–42. Springer Milan, 2012.
- [5] Christophe Biernacki, Gilles Celeux, and Gérard Govaert. Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Comput. Statist. Data Anal.*, 41(3-4):561–575, 2003. Recent developments in mixture models (Hamburg, 2001).
- [6] Jose E. Chacon and Alberto Rodriguez-Casal. A note on the universal consistency of the kernel distribution function estimator. *Statistics & Probability Letters*, 80(17-18):1414–1419, 2010.
- [7] Fermanian J.D. Charpentier, Arthur and O. Scaillet. The estimation of copulas: theory and practice. In *Copulas: from theory to application in finance*. Risk Books; 1 edition, 2006.

- [8] F. Chebana and T.B.M.J. Ouarda. Multivariate quantiles in hydrological frequency analysis. *Environmetrics*, 22(1):63–78, 2011.
- [9] A. Cousin and E. Di Bernardino. A multivariate extension of value-at-risk and conditional-tail-expectation. *Submitted*, 2012.
- [10] Laurens de Haan and Xin Huang. Large quantile estimation in a multivariate setting. *J. Multivariate Anal.*, 53(2):247–263, 1995.
- [11] Elena Di Bernardino, Thomas Laloë, Vronique Maume-Deschamps, and Clmentine Prieur. Plug-in estimation of level sets in a non-compact setting with applications in multivariable risk theory. *ESAIM: Probability and Statistics*, <http://dx.doi.org/10.1051/ps/2011161>, 2011.
- [12] V. Durrleman, A. Nikeghbali, and T. Roncalli. A simple transformation of copulas. Technical report, Groupe de Research Operationnelle Credit Lyonnais, 2000.
- [13] Paul Embrechts and Giovanni Puccetti. Bounds for functions of multivariate risks. *Journal of Multivariate Analysis*, 97(2):526–547, 2006.
- [14] Edward W. Frees and Emiliano A. Valdez. Understanding relationships using copulas. *North American Actuarial Journal*, 2(1):1–25, 1998.
- [15] Stuart A. Klugman and Rahul Parsa. Fitting bivariate loss distributions with copulas. *Insurance Math. Econom.*, 24(1-2):139–148, 1999. 1st IME Conference (Amsterdam, 1997).
- [16] Giovanna Nappo and Fabio Spizzichino. Kendall distributions and level sets in bivariate exchangeable survival models. *Information Sciences*, 179:2878–2890, August 2009.
- [17] Roger B. Nelsen. *An introduction to copulas*, volume 139 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1999.
- [18] Andreas Obereder, Otmar Scherzer, and Arne Kovac. Bivariate density estimation using BV regularisation. *Comput. Statist. Data Anal.*, 51(12):5622–5634, 2007.
- [19] C. Rossi. Sulle curve di livello di una superficie di ripartizione in due variabili. *Giornale dell'Istituto Italiano degli Attuari*, 36:87–108, 1973.
- [20] L. Tibiletti. Sulla quasi concavita delle funzioni di ripartizione n-dimensionali - on quasi-concavity of n-dimensional distribution functions. In *Atti del XV convegno A.M.A.S.E.S.*, pages 503–515, 1991.
- [21] Luisa Tibiletti. Gli insiemi di livello delle funzioni di ripartizione n-dimensionali: un'applicazione statistica (Level sets of n-dimensional distribution functions: a statistical application). *Quaderni dell'Istituto di Matematica Finanziaria dell'universit di Torino: terza serie*, 56:1–11, 1990.
- [22] Luisa Tibiletti. On a new notion of multidimensional quantile. *Metron. International Journal of Statistics*, 51(3-4):77–83, 1993.
- [23] Emiliano A. Valdez and Yugu Xiao. On the distortion of a copula and its margins. *Scandinavian Actuarial Journal*, 2011. accepted for publication and forthcoming.