On the Distortion of a Copula and its Margins

Valdez, Emiliano A.

University of Connecticut

23 December 2009

On the distortion of a copula and its margins

Emiliano A. Valdez [*†] Yugu Xiao [‡]

December 23, 2009

Abstract

This article examines the notion of distortion of copulas, a natural extension of distortion within the univariate framework. We study three approaches to this extension: (1) distortion of the margins alone while keeping the original copula structure, (2) distortion of the margins while simultaneously altering the copula structure, and (3) synchronized distortion of the copula and its margins. When applying distortion within the multivariate framework, it is important to preserve the properties of a copula function. For the first two approaches, this is a rather straightforward result, however for the third approach, the proof has been exquisitely constructed in Morillas (2005). These three approaches of multivariate distortion unify the different types of multivariate distortion that have scarcely scattered in the literature. Our contribution in this paper is to further consider this unifying framework: we give numerous examples to illustrate and we examine their properties particularly with some aspects of ordering multivariate risks. The extension of multivariate distortion can be practically implemented in risk management where there is a need to perform aggregation and attribution of portfolios of correlated risks. Furthermore, ancillary to the results discussed in this article, we are able to generalize the formula developed by Genest and Rivest (2001) for computing the distribution of the probability integral transformation of a random vector and extend it to the case within the distortion framework.

Keywords: Multivariate distortion, ordering of risks, probability integral transformation.

[*]Corresponding author: valdez@math.uconn.edu (phone: +1 860-486-3596). Acknowledgement: The authors would like to thank Christian Genest, Roger Nelsen, and Shaun Wang for sharing and communicating their ideas about distortion of copulas during the conceptualization stages of developing this paper, and Michel Denuit for assisting us clarify some concepts on supermodular functions.

[†]Department of Mathematics, University of Connecticut, Storrs, Connecticut, USA.

[‡]School of Statistics, Remnin University of China, Beijing, P.R. China.

1 Preliminaries

Assume that we have an underlying risk described by an n-dimensional real-valued random vector $\mathbf{X} = (X_1, \ldots, X_n)$ on a well-defined probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is the sample space, $\mathcal{F} = \sigma(X_1, \ldots, X_n)$ is the smallest σ-algebra generated from $(X_1, \ldots, X_n)$, and $\mathbb{P}$ is the probability measure. Denote its multivariate distribution function by $F_{\mathbf{X}}$ belonging to the Fréchet space $\mathcal{R}_n(F_1, \ldots, F_n)$ of random variables with univariate margins $F_i(x_i) = \mathbb{P}(X_i \leq x_i)$ for $i = 1, \ldots, n$. The theorem by Sklar (1959) is a well-known result which states that for any random vector $\mathbf{X}$, its multivariate distribution function has the representation

$$F_{\mathbf{X}}\left(x_1, \ldots, x_n\right) = \mathbb{P}\left(X_1 \leq x_1, \ldots, X_n \leq x_n\right) = C\left(F_1(x_1), \ldots, F_n(X_n)\right), \tag{1}$$

where C is called the *copula* function. Effectively, it is a distribution function on the n-cube $[0,1]^n$ 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 the transformation $U_i = F_i(X_i)$ leads to a uniform random variable so that we can write

$$C(u_1, \ldots, u_n) = F_{\mathbf{X}}(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)) \tag{2}$$

to be the unique copula structure associated with $\mathbf{X}$ with quantile functions F_i^{-1} defined by

$$F_i^{-1}(p) = \inf\left\{x \in \mathbb{R} \mid F_X(x) \geq p\right\}, \qquad p \in [0,1]. \tag{3}$$

It is well-known that members of the Fréchet space $\mathcal{R}_n(F_1, \ldots, F_n)$ are bounded according to

$$\max\left[\sum_{i=1}^{n} F_i(x_i) - n + 1, 0\right] \leq F_{\mathbf{X}}\left(x_1, \ldots, x_n\right) \leq \min\left[F_1(x_1), \ldots, F_n(x_n)\right]$$

where the bounds are respectively called the Fréchet lower and upper bounds. The copula function in (2) is therefore bounded by

$$\max\left(\sum_{i=1}^{n} u_i - n + 1, 0\right) \leq C(u_1, \ldots, u_n) \leq \min(u_1, \ldots, u_n). \tag{4}$$

The Fréchet lower bound in (4) does not satisfy properties of a copula in the case where $n > 2$, but the Fréchet upper bound does for all n referred to as the comonotonic copula and will be denoted by C_U. The independence copula will be denoted by $C_I(u_1, \ldots, u_n) = \prod_{i=1}^{n} u_i$.

In summary, we define a copula $C : [0,1]^n \to [0,1]$ to be a multivariate distribution function whose univariate margins are uniform on $[0,1]$. Its important properties can be summarized below:

- $C(u_1, \ldots, u_n)$ must be increasing in each component u_k.
- $C(u_1, \ldots, u_{k-1}, 0, u_{k+1}, \ldots, u_n) = 0$, for all $k = 1, \ldots, n$.
- $C(1, \ldots, 1, u_k, 1, \ldots, 1) = u_k$, for all $k = 1, \ldots, n$.
- the *rectangle inequality* which leads us to

  $$\sum_{i_1=1}^{2} \cdots \sum_{i_n=1}^{2} (-1)^{i_1 + \cdots + i_n} C(u_{1i_1}, \ldots, u_{ni_n}) \geq 0$$

  for all $u_i \in [0,1]$, $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ with $a_i \leq b_i$, and $u_{i1} = a_i$ and $u_{i2} = b_i$.

There are several classes of copulas that have been examined and studied in the literature. For our purposes, consider the two families of copulas most commonly studied: elliptical copulas and Archimedean copulas. Elliptical copulas are those types of copulas derived from the family of multivariate elliptical distributions. See, for example, Landsman and Valdez (2003). Because the joint distribution functions of elliptical random vectors can only be implicitly expressed, members of elliptical copulas have copula forms that can only be implicitly expressed. Their primary advantages include their flexibility to model tail (or extreme) probabilities and the straightforward procedures to simulate from them even for high dimensions. Two prime examples are the normal and t copulas.

A normal copula has the form

$$C^n_{\mathbf{R}}(u_1, \ldots, u_n) = \mathbf{\Phi}_{\mathbf{R}}(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)), \tag{5}$$

where Φ is the distribution function of a standard univariate normal, $\mathbf{\Phi}_{\mathbf{R}}$ is the joint distribution function of an n-dimensional normal random vector $\mathbf{X} \sim \mathrm{N}_n(\mathbf{0}, \mathbf{R})$ with $\mathbf{R}$, the correlation matrix. The case where $\mathbf{R} = \mathbf{I}_n$, the identity matrix, results in independence, and $\mathbf{R} = \mathbf{J}_n$, the exchange matrix, gives comonotonicity. The exchange matrix consists of 1's on the counterdiagonal and 0's everywhere else.

A t-copula has the form

$$C^t_{\nu,\mathbf{R}}(u_1, \ldots, u_n) = \mathbf{t}_{\nu,\mathbf{R}}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_n)), \tag{6}$$

where t_ν is the distribution function of a standard univariate student t with ν degrees of freedom, $\mathbf{t}_{\nu,\mathbf{R}}$ is the joint distribution of an n-dimensional t random vector $\mathbf{X} \sim \mathrm{t}_n(\nu, \mathbf{0}, \mathbf{R})$ with $\mathbf{R}$, the correlation matrix. The case where $\mathbf{R} = \mathbf{J}_n$ gives comonotonicity, but $\mathbf{R} = \mathbf{I}_n$ does not result in independence.

An *Archimedean* copula has the form

$$C_\psi(u_1, \ldots, u_n) = \psi^{-1}(\psi(u_1) + \cdots + \psi(u_n)), \tag{7}$$

for some function ψ (called the Archimedean generator) satisfying:

- $\psi(1) = 0$;
- ψ is decreasing; and
- ψ is convex.

To ensure that we get a legitimate copula for higher dimensions say n, ψ^{-1} must be *completely monotonic* of order n; these are functions with higher derivatives, provided they exist, that alternate in signs up to and including n. An important source of Archimedean generators is the inverse of a Laplace transform of distribution functions. In Feller (1971), p. 439, a function φ on $[0, \infty]$ is the Laplace transform of a distribution function F if and only if φ is completely monotonic with $\varphi(0) = 1$.

The independence copula can be be viewed as a special Archimedean copula with generator $\psi(t) = -\log(t)$. Other well-known generators within the Archimedean class are:

- Clayton copula: $\psi(t) = t^{-\alpha} - 1$, for $\alpha > 0$;
- Ali-Mikhail-Haq copula: $\psi(t) = \log\left(\dfrac{1 - \alpha(1-t)}{t}\right)$, for $-1 \leq \alpha \leq 1$;

- Gumbel-Hougaard copula: $\psi(t) = (-\log(t))^\alpha$, for $\alpha \geq 1$; and
- Frank copula: $\psi(t) = -\log\left(\frac{e^{-\alpha t}-1}{e^{-\alpha}-1}\right)$ for $\alpha > 0$.

The resulting copulas using formula (7) based on these generators are straightforward to derive. Many other types, as well as additional characterizations, of this special class of copulas can be found in Nelsen (2006).

To learn more about copulas, please refer to Joe (1997) and Nelsen (2006). For applications in actuarial science, insurance and finance, please see Cherubini et al. (2004), Frees and Valdez (1998), Klugman and Parsa (1999) and McNeil et al. (2005).

This introductory section on "Preliminaries" serves as a basic introduction to copulas, their properties and some illustrative examples. While they are more extensively covered in many of the sources we cite above, we provide here the foundation so that this article can be self-contained for our purposes. The main purpose of this paper is examine the extension of univariate distortion to the multivariate framework. Because most of the applications we are interested in relate to risks, we confine ourselves to non-negative random vectors, i.e. each component of the vector is a non-negative random variable. In Section 2, we provide a basic foundation of distortion within the univariate framework to provide a prelude to subsequent sections. In Section 3, we consider the extension to three different kinds of multivariate distortion. In Section 4, we examine some properties of multivariate distortion in the particular case when we order risks in higher dimensions. In Section 5, we extend the results of Genest and Rivest (2001) on the distribution of the probability integral transform in two directions: one is to extend their probability formula to more than two dimensions and another is to generalize it to integral transforms with distortion of the third kind. Finally, we conclude in Section 6.

2 Univariate distortion

For this and subsequent sections, we assume X is a non-negative random variable, on a well-defined probability space, with distribution function $F_X \in \mathcal{R}^+(F_X)$, the Frechet space of non-negative random variables with distribution function F_X.

Convex functions will play a very important role in the distortion of univariate distribution functions. Let I be an interval on the real line $\mathbb{R}$ and g, a mapping from I to $\mathbb{R}$, i.e. $g : I \to \mathbb{R}$. We say that g is *convex* if for all $t_1, t_2 \in I$ and for any $\alpha \in [0, 1]$, we have

$$g(\alpha t_1 + (1-\alpha)t_2) \leq \alpha g(t_1) + (1-\alpha)g(t_2). \tag{8}$$

The simplest form of visualizing a convex function is that its graph will always fall below its chords. That is, if you take any two points on the graph of g and connect them to form a line, the line will always lie above the graph of g. If the inequality in (8) is reversed, then we say the function is *concave*. Clearly, if g is convex, then $-g$ is concave. In addition, if g is convex, then g^{-1} is concave.

Suppose that g is differentiable everywhere on I. A necessary and sufficient condition for g to be convex is that its derivative g' is non-decreasing. Thus, in the case where its second derivative, g'' exists everywhere on I, a necessary and sufficient condition for g to be convex is that $g''(t) \geq 0$ for all $t \in I$. For more about convex functions, see Roberts and Varberg (1973).

An important property of convex functions is the Jensen's inequality. Let X have mean $\mathbb{E}(X)$ that exists. According to *Jensen's inequality*, we have

$$\mathbb{E}[g(X)] \geq g[\mathbb{E}(X)] \tag{9}$$

for any convex function g.

Definition 2.1 ***(Distortion function)*** *Assume that $g:[0,1]\to[0,1]$. We say g is a distortion function if it satisfies the following properties:*

- *$g(0)=0$ and $g(1)=1$; and*
- *g is continuous and non-decreasing.*

In the case where g is convex, then we have what we call a *convex distortion function*. On the other hand, if g is concave, we have what we call *concave distortion function*.

Definition 2.2 ***(Probability distortion)*** *Let g be a distortion function and X a random variable with distribution function F_X. Then the transformation of the distribution function with*

$$F_{X^*}(x)=g[F_X(x)]=g\circ F_X(x) \tag{10}$$

is the distribution function of X^* that leads to a probability distortion of X to X^*.

In insurance pricing and in financial risk management, transformation of the distribution function typically represents a change in the probability measure. To illustrate in actuarial science, Wang (1996) defines a premium principle based on the concept of distortion function motivated by Yaari's dual theory of choice under risk; see Yaari (1987). The distortion premium principle associated with the distortion function g is then defined to be

$$\pi_g(X)=\mathbb{E}(X^*), \tag{11}$$

the expectation under the distorted probability measure. Note that for insurance premium purposes, this distorted expectation must be at least equal to the expectation under the original probability measure. Such is the case only when g is convex. To see this, if g is indeed convex, then direct application of Jensen's inequality in (9) leads us to

$$g[F_X(x)]\le F_X(x)$$

from which it follows that

$$\int_0^\infty [1-g[F_X(x)]]dx \ge \int_0^\infty [1-F_X(x)]dx.$$

Thus, clearly,

$$\pi_g(X)-\mathbb{E}(X)\ge 0, \tag{12}$$

and this difference is often referred to as the risk premium. In the actuarial literature, it is a more common practice to distort the survival function, $S_X(x)=1-F_X(x)$, instead of the distribution function. If we distort F_X with a distortion function g, this implies that

$$S_{X^*}(x)=1-g[1-S_X(x)].$$

By defining the function

$$\tilde{g}(t)=1-g(1-t) \tag{13}$$

Table 1: *Some Examples of Distortion Functions*

Distortion	Functional form $g(t)$	Inverse form $g^{-1}(s)$	Convex constraints	Concave constraints
Proportional hazard	$t^{1/\gamma}$	s^{γ}	$\gamma \geq 1$	$0 < \gamma \leq 1$
Exponential	$\dfrac{1-e^{-\gamma t}}{1-e^{-\gamma}}$	$\log[1-s(1-e^{-\gamma}]$	$\gamma < 0$	$\gamma > 0$
Logarithmic	$\dfrac{1}{\gamma}\log[1-t(1-e^{\gamma})]$	$\dfrac{e^{\gamma t-1}}{e^{\gamma-1}}$	$\gamma < 0$	$\gamma > 0$
Wang transform	$\Phi[\Phi^{-1}(t)+\gamma]$	$\Phi[\Phi^{-1}(s)-\gamma]$	$\gamma \leq 0$	$\gamma \geq 0$
Dual-power	$1-(1-t)^{\gamma}$	$1-(1-s)^{1/\gamma}$	$\gamma \leq 1$	$\gamma \geq 1$

Note: The convex/concave constraints are for the function $g(t)$.

for which it may be called the *conjugate* of the distortion function g, this demonstrates the equivalence of the distortion between the distribution and survival functions. Clearly, if g is convex, then its conjugate $\tilde{g}$ is concave, and vice-versa. However, in higher dimensions as we shall see in subsequent sections, it is more imperative to distort distribution functions.

Examples of distortion functions are summarized above in Table (1). Several other distortion functions can be found in Morillas (2005).

Mathematical theories of risk assume that probability distributions for risks under consideration are known without ambiguity. In practice, for example, we estimate these probability distributions usually based on limited data. As a result, parameter uncertainty is always present. To illustrate, consider an insurance risk random variable X such that conditional on the risk parameter γ, its distribution is Exponential with parameter γ: $F_X(x|\gamma) = 1 - \exp(-\gamma x)$. If γ has a Gamma distribution with a *scale* and *shape* parameters λ and α, respectively, the unconditional distribution of X is a Pareto distribution expressed as

$$F_X(x) = 1 - (1 + \lambda x)^{-\alpha}.$$

See Frees and Valdez (1998). One can therefore effectively think of introducing uncertainty in the parameter of the distribution as a distortion. As a matter of fact, one can easily derive the corresponding distortion function in this case:

$$g(t) = 1 - (1 + \log(1-t)^{-\lambda/\gamma})^{-\alpha}.$$

Note that this distortion function is neither strictly convex nor concave.

3 Distortion of copulas

First, let us consider absolutely monotonic functions. Assume n is a non-negative integer.

Definition 3.1 ***(Absolutely monotonic function)*** *A 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 $k = 0, 1, \ldots, n$.*

We can simply say that a function is absolutely monotonic if it is absolutely monotonic of order n for all non-negative integer n. Absolutely monotonic functions are typically studied in connection with Laplace transforms, see e.g. Widder (1946). Linear combinations and products of absolutely monotonic functions are also absolutely monotonic functions. Clearly following immediately from the definition, if g is absolutely monotonic of order n on I, then its k-th derivative is absolutely monotonic of order $n-k$. Finally, 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 $g \circ h(t) = g(h(t))$ is also absolutely monotonic of order n on I.

A sufficient condition for a function to be absolutely monotonic is given by: If g is absolutely monotonic of order n on I, then it must be non-negative, non-decreasing, convex and continuous everywhere on I. For proof of this and results stated in the previous paragraph, please consult Widder (1946).

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be an n-dimensional real-valued random vector with corresponding copula denoted by $C_{\mathbf{X}}(u_1, \ldots, u_n)$ with $u_i = F_i(x_i)$. We now consider three different approaches to extending distortion within the multivariate framework.

Definition 3.2 ***(Distortion of the first kind)*** *Let $g_1, \ldots, g_n$ be n distortion functions. Then the transformation of the copula associated with $\mathbf{X}$ defined by*

$$C_{\mathbf{X}}(u_1^*, \ldots, u_n^*) = C_{\mathbf{X}}(g_1(u_1), \ldots, g_n(u_n))$$

induces a multivariate probability distortion of $\mathbf{X}$ to $\mathbf{X}^*$.

We shall call this a *copula distortion of the first kind*. This type of a distortion leads to a simple distortion of the margins while preserving the copula structure. An example of this type is the multivariate extension of the Wang transform constructed by Kijima (2006). Because the copula is preserved, this type of a distortion does not lead us to a construction of a new copula but it does lead us to the construction of a new multivariate distribution function. Consider the following example.

Example 3.3 ***(Multivariate Burr I)*** *Consider the Weibull margins*

$$F_i(x_i) = 1 - \exp(-x_i^k), \qquad x_i \geq 0, k > 0, \tag{14}$$

for $i = 1, \ldots, n$, linked with a legitimate copula, for example, a Clayton copula defined by

$$C_{\mathbf{X}}(u_1, \ldots, u_n) = \left[\sum_{i=1}^{n} u_i^{-\alpha} - n + 1\right]^{-1/\alpha}. \tag{15}$$

One can therefore view this as a multivariate Weibull distribution. Using the distortion function

$$g(t) = 1 - (1 - \log(1-t))^{-\gamma}, \qquad \gamma > 0, \tag{16}$$

this leads us to the Burr margins

$$F_i^*(x_i) = 1 - [(1 + x_i^k)]^{-\gamma}, \qquad x_i \geq 0, k > 0, \gamma > 0. \tag{17}$$

The result of the distortion is a simple transformation to a multivariate Burr distribution.

Definition 3.4 ***(Distortion of the second kind)*** *Let $g_1, \ldots, g_n$ be n distortion functions. Then the transformation of the copula associated with $\mathbf{X}$ defined by*

$$\widehat{C}(u_1^*, \ldots, u_n^*) = \widehat{C}(g_1(u_1), \ldots, g_n(u_n)),$$

where $\widehat{C}$ is a copula function, induces a multivariate probability distortion of $\mathbf{X}$ to $\widehat{\mathbf{X}}$.

We shall call this a *copula distortion of the second kind*. This definition leads to a simultaneous distortion of the margins and the copula structure. Just as in the distortion of the first kind, this does not lead us to a construction of a new copula, but rather a specification of a new copula. In the case where we preserve the copula structure, i.e. $\widehat{C} = C_{\mathbf{X}}$, then we recover the distortion of the first kind. Furthermore, similar to the first kind, it also leads us to the construction of a new multivariate distribution function.

Example 3.5 ***(Multivariate Burr II)*** *Following up on example (3.3), we can similarly distort the margins from a Weibull to a Burr distribution, but transform the copula structure based on the Gumbel-Hougaard which has the form*

$$\widehat{C}(u_1, \ldots, u_n) = \exp\left\{ -\left[\sum_{i=1}^{n} (-\log u_i)^\alpha \right]^{1/\alpha} \right\}. \tag{18}$$

The result of this distortion can be viewed as yet another multivariate Burr distribution.

Definition 3.6 ***(Distortion of the third kind)*** *Let g be a distortion function with an inverse function g^{-1} that is absolutely monotonic of order n on the interval $[0, 1]$. Then the transformation of the copula associated with $\mathbf{X}$ defined by*

$$C_g(u_1, \ldots, u_n) = g^{-1}(C_{\mathbf{X}}(g(u_1), \ldots, g(u_n)))$$

induces a multivariate probability distortion of $\mathbf{X}$ to $\widetilde{\mathbf{X}}$.

We shall call this a *copula distortion of the third kind*. The function C_g that is induced by this distortion indeed satisfies the properties of a copula function and is then the copula that is associated with the distorted random variable $\widetilde{\mathbf{X}}$ and therefore it can be written as

$$C_g(u_1, \ldots, u_n) = C_{\widetilde{\mathbf{X}}}(u_1, \ldots, u_n).$$

For proof, see Morillas (2005). This definition leads to a synchronized distortion of the margins and the copula structure. Unlike the first two kinds, this leads us to a new method of constructing new copulas from a given one. Furthermore, because $C_{\mathbf{X}}$ is a copula, we have that

$$\begin{aligned} C_g(1, \ldots, 1, u_k, 1, \ldots, 1) &= g^{-1}(C_{\mathbf{X}}(g(1), \ldots, g(1), g(u_k), g(1), \ldots, g(1)) \\ &= g^{-1}(C_{\mathbf{X}}(1, \ldots, 1, g(u_k), 1, \ldots, 1) \\ &= g^{-1}(g(u_k)) = u_k, \end{aligned}$$

for all $k = 1, \ldots, n$. This implies that a synchronized distortion preserves the margins; it simply distorts the dependence structure.

Note that in Definition 3.6, the function g^{-1} must be absolutely monotonic which implies that it must be convex. As a consequence, g must be concave. Except for the case of the "proportional hazard", all the distortion functions in Table 1 have inverses that are absolutely monotonic. We provide the proof for the case of the "Wang Transform"; the rest are straightforward to prove. For the "proportional hazard", it is also easy to show that when the parameter γ is a non-negative integer, it satisfies absolute monotonicity.

Example 3.7 ***(Distortion of the comonotonic copula)*** *It is interesting to note that if we distort the comonotonic copula defined by*

$$C_U(u_1, \ldots, u_n) = \min(u_1, \ldots, u_n), \tag{19}$$

we recover the same comonotonic copula. This is straightforward to show by noting that because g is increasing, if $u_i = \min(u_1, \ldots, u_n)$ for some $i = 1, \ldots, n$, then $g(u_i) = \min(g(u_1), \ldots, g(u_n))$. Therefore, it follows that

$$C_g(u_1, \ldots, u_n) = g^{-1}(C_U(g(u_1), \ldots, g(u_n))) = \min(u_1, \ldots, u_n).$$

This result is also an immediate consequence of the fact that distortion of the third kind preserves the margins.

Example 3.8 ***(Generalized multivariate Wang distortion)*** *We note in the appendix that the inverse of the Wang transform is an absolutely monotonic function. Thus, we have the following generalization of the multivariate Wang transform:*

$$C_w(u_1, \ldots, u_n) = \Phi\left\{\Phi^{-1}\left[\mathbf{\Phi_R}(\Phi[\Phi^{-1}(u_1) + \gamma], \ldots, \Phi[\Phi^{-1}(u_n) + \gamma])\right] - \gamma\right\}, \tag{20}$$

provided $\gamma \geq 0$. An even further generalization is to additionally apply distortion of the first kind to the copula in (20) with $u_i = \Phi[\Phi^{-1}(u_i^*) + \gamma_i]$ for $i = 1, \ldots, n$ where $\gamma_i \geq 0$. This leads us to the following multivariate Wang transform:

$$C_w(u_1^*, \ldots, u_n^*) = \Phi\left\{\Phi^{-1}\left[\mathbf{\Phi_R}(\Phi[\Phi^{-1}(u_1^*) + \gamma_1^*], \ldots, \Phi[\Phi^{-1}(u_n^*) + \gamma_n^*])\right] - \gamma\right\}, \tag{21}$$

where, in our context, $\gamma_i^* = \gamma_i + \gamma$.

Example 3.9 ***(Distortion of composite functions)*** *Suppose g_1 and g_2 are two distortion functions with respective inverses g_1^{-1} and g_2^{-1} that are both absolutely monotonic. Define the composite function $g = g_1 \circ g_2$ so that g is itself a distortion function with $g^{-1} = g_2^1 \circ g_1^{-1}$. From the property of absolutely monotone, since both g_1^{-1} and g_2^{-1} are absolutely monotonic, so with g^{-1}. Therefore, distortion of composite functions with absolutely monotonic inverses leads to a distortion of copula of the third kind.*

We notice that the absolutely monotone requirement for g^{-1} is only a sufficient condition for a distortion of a copula of the third kind as defined in Definition 3.6. However, it is not a necessary condition as the following two examples illustrate.

Example 3.10 ***(Distortion of Archimedean copulas)*** *Suppose g is a distortion function with inverse g^{-1} that is absolutely monotonic. Consider applying this distortion on an Archimedean copula of the form as defined in (7). We find that*

$$\begin{aligned} C_g(u_1,\ldots,u_n) &= g^{-1}(C_\psi(g(u_1),\ldots,g(u_n))) \\ &= g^{-1}[\psi^{-1}(\psi(g(u_1))+\cdots+\psi(g(u_n)))] \\ &= \psi_g^{-1}(\psi_g(u_1)+\cdots+\psi_g(u_n)) \\ &= C_{\psi_g}(u_1,\ldots,u_n) \end{aligned}$$

where $\psi_g = \psi \circ g$ is the composite function of ψ and g. For this to be a legitimate Archimedean copula, ψ_g^{-1} must be completely monotonic. This is in conflict with the requirement in example (3.9) where for it to be a legitimate copula, it must be absolutely monotonic. Clearly, it cannot be both completely and absolutely monotonic at the same time. See also the comment of Morillas (2005) on page 183.

There are many examples we can use to demonstrate distortion within the class of Archimedean copulas based on example (3.10). Consider the proportional hazard distortion of the Independence copula. Here, we have $\psi(t) = -\log(t)$ and $g(t) = t^{1/\gamma}$ so that $\psi_g(t) = -\log(t^{1/\gamma})$ so that this distortion preserves Independence. A proportional hazard distortion of the Clayton copula leads us to an Archimedean generator $\psi_g = t^{-\alpha/\gamma} - 1$, another Clayton copula. Finally, it is easy to demonstrate that in the logarithmic distortion of the Independence copula, we derive the Frank copula.

Example 3.11 ***(Distortion of the Independence to Archimedean copulas)*** *Consider the distortion function defined by*

$$g(t) = \exp(-\psi(t)), \tag{22}$$

where ψ is a function with inverse ψ^{-1} that is completely monotonic. Clearly, the inverse of g is

$$g^{-1}(t) = \psi^{-1}(-\log(t)). \tag{23}$$

If we apply the distortion of the third kind, it is straightforward to show it yields to an Archimedean copula of the form as exactly defined in (7). This distortion can indeed be viewed as a special case of Example 3.10 where the distortion function g is the composite function of two distortion functions $g_1(t) = \exp(-t)$ and $g_2(t) = \psi(t)$.

4 Multivariate ordering of risks with distortion

In this and subsequent sections, we shall consider only non-negative random vectors. Let $\mathbf{X} = (X_1,\ldots,X_n)$ and $\mathbf{Y} = (Y_1,\ldots,Y_n)$ be two n-dimensional random vectors and $\mathbf{x} = (x_1,\ldots,x_n)$ and $\mathbf{y} = (y_1,\ldots,y_n)$ be two n-dimensional real-valued vectors.

Definition 4.1 ***(Supermodular function)*** *A function $h : \mathbb{R}^n \to \mathbb{R}$ is called supermodular, if for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*

$$h(\mathbf{x} \vee \mathbf{y}) + h(\mathbf{x} \wedge \mathbf{y}) \geq h(\mathbf{x}) + h(\mathbf{y}) \tag{24}$$

where the operators $\vee$ and $\wedge$ denotes componentwise maximum and minimum of $\mathbf{x}$ and $\mathbf{y}$, respectively defined as:

$$\mathbf{x} \vee \mathbf{y} = (\max(x_1,y_1),\ldots,\max(x_n,y_n)) \text{ and } \mathbf{x} \wedge \mathbf{y} = (\min(x_1,y_1),\ldots,\min(x_n,y_n)).$$

As pointed out in Denuit et al. (2005), the condition for supermodularity is equivalent to satisfying

$$h(x_1, \ldots, x_i + \delta, \ldots, x_j + \varepsilon, \ldots, x_n) - h(x_1, \ldots, x_i + \delta, \ldots, x_j, \ldots, x_n)$$
$$- h(x_1, \ldots, x_i, \ldots, x_j + \varepsilon, \ldots, x_n) + h(x_1, \ldots, x_n) \geq 0,$$

for any $\delta > 0$, $\varepsilon > 0$, $0 \leq i < j \leq n$, from which an insurance interpretation of supermodularity can best be drawn from. If $h(x_1, \ldots, x_n)$ represents the loss to the insurer when $x_1, \ldots, x_n$ denote the individual claims coming from n policies, then the supermodularity of the function has the implication of a worse loss to the insurer, given that an increase of a single claim increases the values of some of the other claims. More about supermodular functions and their properties can be found in Bäuerle (1997) and Müller and Scarsini (2000).

Definition 4.2 *(Supermodular order)* $\mathbf{X}$ *is said to be smaller than* $\mathbf{Y}$ *in the supermodular order, denoted by* $\mathbf{X} \leq_{sm} \mathbf{Y}$*, if*

$$\mathbb{E}[h(\mathbf{X})] \leq \mathbb{E}[h(\mathbf{Y})]$$

for all measurable supermodular functions h *for which the expectations exist.*

Lemma 4.3 *Suppose that* $\mathbf{X}$ *and* $\mathbf{Y}$ *have joint distribution functions* $F_{\mathbf{X}}$ *and* $F_{\mathbf{Y}}$*, respectively, and let* $\mathbf{X}$ *have the univariate margins denoted by* $F_1, \ldots, F_n$*. If* $\mathbf{X} \leq_{sm} \mathbf{Y}$*, then there exist two copulas* $C_{\mathbf{X}}$ *and* $C_{\mathbf{Y}}$ *such that*

$$F_{\mathbf{X}}(x_1, \ldots, x_n) = C_{\mathbf{X}}(F_1(x_1), \ldots, F_n(x_n)) \quad \text{and} \quad F_{\mathbf{Y}}(y_1, \ldots, y_n) = C_{\mathbf{Y}}(F_1(y_1), \ldots, F_n(y_n)).$$

Proof. According to Theorem 3.4 in Müller and Scarsini (2000), since $\mathbf{X} \leq_{sm} \mathbf{Y}$, then $\mathbf{X}$ and $\mathbf{Y}$ have the same margins. The existence of the copulas immediately follows from Sklar (1959). ■

Proposition 4.4 ***(Distortion of the first kind preserves supermodular order)*** *Let* $\mathbf{X}^*$ *and* $\mathbf{Y}^*$ *be the two random vectors induced by the distortion according to Definition 3.2 with respective joint distribution functions* $F_{\mathbf{X}^*}(x_1, \ldots, x_n) = C_{\mathbf{X}}(g_1(F_1(x_1)), \ldots, g_n(F_n(x_n)))$ *and* $F_{\mathbf{Y}^*}(y_1, \ldots, y_n) = C_{\mathbf{Y}}(g_1(F_1(y_1)), \ldots, g_n(F_n(y_n)))$*. The following holds true:*

$$\mathbf{X} \leq_{sm} \mathbf{Y} \ \textit{ implies } \ \mathbf{X}^* \leq_{sm} \mathbf{Y}^*. \tag{25}$$

Proof. Let $U_i = g_i \circ F_i(X_i^*)$ for $i = 1, \ldots, n$. Clearly, each U_i is uniform on $[0, 1]$ and the joint distribution of $(U_1, \ldots, U_n)$ is $C_{\mathbf{X}}(u_1, \ldots, u_n)$. For any measurable supermodular function h,

$$\begin{aligned}
\mathbb{E}[h(\mathbf{X}^*)] &= \mathbb{E}[h((g_1 \circ F_1)^{-1}(U_1), \ldots, (g_n \circ F_n)^{-1}(U_n))] \\
&= \mathbb{E}[h(F_1^{-1} \circ g_1^{-1} \circ F_1 \circ F_1^{-1}(U_1), \ldots, F_n^{-1} \circ g_n^{-1} \circ F_n \circ F_n^{-1}(U_n))] \\
&= \mathbb{E}[h^*(F_1^{-1}(U_1), \ldots, F_n^{-1}(U_n))] \\
&= \mathbb{E}[h^*(\mathbf{X})],
\end{aligned}$$

where $h^*(x_1, \ldots, x_n) = h(F_1^{-1} \circ g_1^{-1} \circ F_1(x_1), \cdots, F_n^{-1} \circ g_n^{-1} \circ F_n(x_n))]$. According to Lemma 2.1 in Bäuerle (1997), since for each $i = 1, \ldots, n$, $F_i^{-1} \circ g_i^{-1} \circ F_i(x_i)$ is increasing, the composition $h[F_1^{-1} \circ g_1^{-1} \circ F_1(\cdot), \ldots, F_n^{-1} \circ g_n^{-1} \circ F_n(\cdot)]$ is a supermodular function. It therefore follows $\mathbf{X}^* \leq_{sm} \mathbf{Y}^*$ which implies that $\mathbb{E}[h^*(\mathbf{X})] \leq \mathbb{E}[h^*(\mathbf{Y})]$. Subsequently, we have $\mathbb{E}[h(\mathbf{X}^*)] \leq \mathbb{E}[h(\mathbf{Y}^*)]$ so that the desired result follows. ■

We remark that the multivariate Wang transform constructed in Kijima (2006) is an example of a distortion of the first kind and from Proposition 4.4, we find that it preserves supermodular order.

Lemma 4.5 *Let $C_g(u_1, \ldots, u_n) = g^{-1}(C_{\mathbf{X}}(g(u_1), \ldots, g(u_n)))$ where g^{-1} is an absolutely monotonic distortion function. If the copula $C_{\mathbf{X}}(u_1, \ldots, u_n)$ has second order derivatives for all $(u_1, \ldots, u_n) \in [0,1]^n$, then the relative density between $C_g(F_1(x_1), \ldots, F_n(x_n))$ and $C_{\mathbf{X}}(F_1^*(x_1), \ldots, F_n^*(x_n))$ has the expression*

$$\frac{dC_g(F_1(x_1), \ldots, F_n(x_n))}{dC_{\mathbf{X}}(F_1^*(x_1), \ldots, F_n^*(x_n))} = \frac{1}{g'(g^{-1}(C(F_1^*(x_1), \ldots, F_n^*(x_n))))} \tag{26}$$

where $F_i^*(x_i) = g(F_i(x_i))$ for $i = 1, \ldots, n$, and it is a supermodular function.

Proof. Straightforward differentiation leads us to the first part of the lemma:

$$\begin{aligned} \frac{dC_g(F_1(x_1), \ldots, F_n(x_n))}{dC_{\mathbf{X}}(F_1^*(x_1), \ldots, F_n^*(x_n))} &= \frac{dg^{-1}(C_{\mathbf{X}}(F_1^*(x_1), \ldots, F_n^*(x_n))}{dC_{\mathbf{X}}(F_1^*(x_1), \ldots, F_n^*(x_n))} \\ &= \frac{1}{g'(g^{-1}(C_{\mathbf{X}}(F_1^*(x_1), \ldots, F_n^*(x_n))))}. \end{aligned}$$

Since, by assumption, the copula $C_{\mathbf{X}}$ is twice differentiable, and we know that $\frac{\partial^2 C_{\mathbf{X}}(u_1, \cdots, u_n)}{\partial u_i \partial u_j} \geq 0$ for any $1 \leq i < j \leq n$, then it follows from Theorem 2.2 in Müller and Scarsini (2000) that $C_{\mathbf{X}}$ is supermodular. In addition, the composition $C_{\mathbf{X}}(F_1^*(\cdot), \ldots, F_n^*(\cdot)))$ is also supermodular. Because g^{-1} is an absolutely monotonic distortion function, this implies that its derivative $dg^{-1}(t)/dt$ is a convex function. Finally, from Lemma 2.1 in Bäuerle (1997), we have that

$$\frac{1}{g'(g^{-1}(C_{\mathbf{X}}(F_1^*(x_1), \ldots, F_n^*(x_n))))}$$

is supermodular which gives us the desired result. ■

The following proposition gives us a condition for supermodularity so that supermodular order is preserved for distortion of the third kind. Let us consider the following lemma needed to prove the subsequent condition for supermodularity.

Lemma 4.6 ***(Product of supermodular functions)*** *Suppose v and w are functions $\mathbb{R}^n \to \mathbb{R}$, both of which are non-negative, non-decreasing and supermodular. Their product vw is supermodular.*

Proof. Define $v_{11} = v(x_1, \ldots, x_i + \delta, \ldots, x_j + \varepsilon, \ldots, x_n)$, $v_{10} = v(x_1, \ldots, x_i + \delta, \ldots, x_j, \ldots, x_n)$, $v_{01} = v(x_1, \ldots, x_i, \ldots, x_j + \varepsilon, \ldots, x_n)$ and $v_{00} = v(x_1, \ldots, x_n)$, for any $\delta > 0$, $\varepsilon > 0$, and $0 \leq i < j \leq n$. Similar definition holds for w. Because

$$\begin{aligned} & v_{11}w_{11} - v_{01}w_{01} - v_{10}w_{10} + v_{00}w_{00} \\ & = [v_{01}(w_{11} - w_{01}) - v_{00}(w_{10} - w_{00})] + [w_{11}(v_{11} - v_{01}) - w_{10}(v_{10} - v_{00})] \\ & = [v_{01}(w_{11} - w_{01} - w_{10} + w_{00}) + (v_{01} - v_{00})(w_{10} - w_{00})] \\ & \quad + [(w_{11} - w_{10})(v_{11} - v_{01}) + w_{10}(v_{11} - v_{01} - v_{10} + v_{00})] \geq 0, \end{aligned}$$

we conclude that vw is supermodular. ■

Proposition 4.7 ***(Condition for supermodularity)*** *Suppose that $\mathbf{X}$ and $\mathbf{Y}$ have respective associated copulas $C_{\mathbf{X}}$ and $C_{\mathbf{Y}}$ with absolutely continuous margins. Denote by $\widetilde{\mathbf{X}}$ and $\widetilde{\mathbf{Y}}$ the random vectors induced by an absolutely monotonic distortion function g^{-1} according to Definition 3.6. Let $h : \mathbb{R}^n \to \mathbb{R}$ be a non-negative, non-decreasing and supermodular function with $\mathbb{E}[h(\mathbf{X})]$ that exists. If $\mathbf{X} \leq_{sm} \mathbf{Y}$, then $\mathbb{E}[h(\widetilde{\mathbf{X}})] \leq \mathbb{E}[h(\widetilde{\mathbf{Y}})]$.*

Proof. According to Proposition 4.6 in Morillas (2005), there exists a sequence of n-dimensional copulas $\{C_k\}$ which converges uniformly to $C_{\mathbf{X}}$ and that each C_k has nth-derivatives that exist. Define the sequence

$$h_k(x_1, \ldots, x_n) = \frac{1}{g'(g^{-1}(C_k(F_1^*(x_1), \ldots, F_n^*(x_n))))}.$$

From Lemma 4.5, each h_k in the sequence is supermodular. Because h_k is non-negative and non-decreasing, from Lemma 4.6, the product $h(x_1, \ldots, x_n)h_k(x_1, \cdots, x_n)$ is supermodular. By Proposition 4.4, if $\mathbf{X} \leq_{sm} \mathbf{Y}$, then $\mathbf{X}^* \leq_{sm} \mathbf{Y}^*$. Thus, it follows that

$$\mathbb{E}[h(\mathbf{X}^*)h_k(\mathbf{X}^*)] \leq \mathbb{E}[h(\mathbf{Y}^*)h_k(\mathbf{Y}^*)].$$

Note that g^{-1} is an absolutely monotonic distortion function on $[0, 1]$. This implies that $g^{-1}(t) \leq t$ and $dg^{-1}(t)/dt \leq g^{-1}(t)/t \leq 1$, and that we have $F_i^{-1} \circ g^{-1} \circ F_i(y) \leq y$ and $h_k(x_1, \ldots, x_n) \leq C_k(F_1^*(x_1), \ldots, F_n^*(x_n)) \leq 1$. Thus, we have

$$\mathbb{E}[h(\mathbf{X}^*)h_k(\mathbf{X}^*)] \leq \mathbb{E}[h(\mathbf{X}^*)] \leq \mathbb{E}[h(\mathbf{X})],$$

and by the Dominated Convergence Theorem and from formula (26),

$$\mathbb{E}[h(\widetilde{\mathbf{X}})] = \lim_{k\to\infty} \mathbb{E}[h(\mathbf{X}^*)h_k(\mathbf{X}^*)] \leq \lim_{k\to\infty} \mathbb{E}[h(\mathbf{Y}^*)h_k(\mathbf{Y}^*)] = \mathbb{E}[h(\widetilde{\mathbf{Y}})],$$

which gives us the desired result. ■

Corollary 4.8 ***(Distortion of the third kind preserves supermodular order)*** *Suppose that $\mathbf{X}$ and $\mathbf{Y}$ have respective associated copulas $C_{\mathbf{X}}$ and $C_{\mathbf{Y}}$ with absolutely continuous margins. The following holds true:*

$$\mathbf{X} \leq_{sm} \mathbf{Y} \ \textit{implies} \ \widetilde{\mathbf{X}} \leq_{sm} \widetilde{\mathbf{Y}}. \tag{27}$$

Proof. First, for any bounded, continuous and non-decreasing supermodular function $h : \mathbb{R}^n \to \mathbb{R}$, assuming $h(x_1, \ldots, x_n) \geq M$, define $v(x_1, \ldots, x_n) = h(x_1, \ldots, x_n) - M$. Then $v(x_1, \ldots, x_n)$ is non-negative, non-decreasing and supermodular. From Proposition 4.7, we have $\mathbb{E}[v(\widetilde{\mathbf{X}})] \leq \mathbb{E}[v(\widetilde{\mathbf{Y}})]$ so that $\mathbb{E}[h(\widetilde{\mathbf{X}})] \leq \mathbb{E}[h(\widetilde{\mathbf{Y}})]$. According to Theorem 3.3 and Theorem 3.4 in Müller and Scarsini (2000), we can finally conclude $\widetilde{\mathbf{X}} \leq_{sm} \widetilde{\mathbf{Y}}$. ■

Definition 4.9 ***(Stop-loss order)*** *We say that $X \leq_{SL} Y$, in the stop-loss order sense, if $\mathbb{E}[v(X)] \leq \mathbb{E}[v(Y)]$ for all functions v with non-negative first and second derivatives, i.e. $v' \geq 0$ and $v'' \geq 0$ provided the expectations exist.*

Additional equivalent characterizations of stop-loss order can be found in Denuit et al. (2005). We remark that following Corollary 4.8, it is immediate to conclude that

$$S_{\widetilde{\mathbf{X}}} = \sum_{i=1}^{n} \widetilde{X}_i \leq_{SL} S_{\widetilde{\mathbf{Y}}} = \sum_{i=1}^{n} \widetilde{Y}_i.$$

In other words, stop-loss order of sums of the distorted random variables of the third kind is preserved. Furthermore, it follows that when $\mathbf{X} \leq_{sm} \mathbf{Y}$, we have $\mathbb{E}(S_{\widetilde{\mathbf{X}}}) \leq \mathbb{E}(S_{\widetilde{\mathbf{Y}}})$ and for any positive deductible d, we have $\mathbb{E}[(S_{\widetilde{\mathbf{X}}} - d)_+] \leq \mathbb{E}[(S_{\widetilde{\mathbf{Y}}} - d)_+]$.

5 Multivariate probability integral transform with distortion

In this section, we extend the results of Genest and Rivest (2001) to higher dimension and when we have distortion of the third kind. Let us briefly review the results in Genest and Rivest (2001). Without loss of generality, Genest and Rivest (2001) considered a random pair (X_1, X_2) with uniform margins on $[0,1]$ and associated copula denoted by $C(x_1, x_2)$. They showed that the probability of the event $C(X_1, X_2) \leq v$ has the following form:

$$K(v) = v + \int_v^1 \dot{C}\left(x_1, C_{x_1}^{-1}(v)\right) dx_1 = v - \lambda(v), \tag{28}$$

where $\dot{C}(x_1, x_2) = \partial C(x_1, x_2)/\partial x_1$ which is known to represent the conditional distribution function of X_2 given $X_1 = x_1$ and $C_{x_1}(x_2) = C(x_1, x_2)$.

As in Genest and Rivest (2001), without having to give up generality, we simplify the similar derivation to higher dimension by considering the random vector $\mathbf{X} = (X_1, \ldots, X_n)$ with uniform margins on $[0,1]$ and associated copula denoted by $C(X_1, \ldots, X_n)$. Define the event $A_k = \{C(X_1, \ldots, X_k, Y, 1, \ldots, 1) \leq v\}$ for integral values $k \geq 1$, denoting the $(k+1)$-th element of the vector by Y instead of X_{k+1} for notational simplicity. Denote the probability of the event A_{n-1} by $K_n(v) = \mathbb{P}(A_{n-1})$.

For $k = 1, \ldots, n-1$, define the event $E_k = \{C(X_1, \ldots, X_k, 1, \ldots, 1) \geq v\}$, the case where the observed values of the last $n-k$ terms are all equal to 1. Its complement will be denoted with a superscript c. For instance, in the case where $k = 1$, this is the same event as $E_1 = \{C(X_1, 1, \ldots, 1) \geq v\} = \{X_1 \geq v\}$. Furthermore, denote the set $e_k = \{(x_1, \ldots, x_k) |\ C(x_1, \ldots, x_k, 1 \ldots, 1) \geq v\}$ for $k = 1, \ldots, n-1$.

Lemma 5.1 *The probability of the event A_{n-1} can be expressed as*

$$K_n(v) = v + \sum_{k=1}^{n-1} \mathbb{E}\, I_{E_1 E_2 \cdots E_k} [\mathbb{E}\, I_{A_k} | X_1, X_2, \ldots, X_k]. \tag{29}$$

Proof. Since $C(x_1, \ldots, x_{n-1}, y) \leq \min(x_1, \ldots, x_{n-1}, y)$, $I_{E_1^c} I_{A_n} = I_{E_1^c}$, and for any $k = 1, \ldots, n-1$, $I_{E_k^c} I_{A_k} = I_{E_k^c}$, we have

$$\begin{aligned}
K_n(v) = \mathbb{P}(A_{n-1}) &= \mathbb{E}\, I_{\mathbb{E}_1^c} [\mathbb{E}\, I_{A_{n-1}} | X_1, \ldots, X_{n-1}] + \mathbb{E}[\mathbb{E}\, I_{\mathbb{E}_1} I_{A_{n-1}} | X_1, \ldots, X_{n-1}] \\
&= \mathbb{E}\, I_{E_1^c} + \mathbb{E}\, I_{E_1} [\mathbb{E}\, I_{A_{n-1}} | X_1, \ldots, X_{n-1}] \\
&= v + \mathbb{E}\, I_{E_1} [\mathbb{E}\, I_{A_{n-1}} I_{E_2^c} | X_1, \ldots, X_{n-1}] + \mathbb{E}\, I_{E_1} [\mathbb{E}\, I_{A_{n-1}} I_{E_2} | X_1, \ldots, X_{n-1}] \\
&= v + \mathbb{E}\, I_{E_1} I_{E_2^c} + \mathbb{E}\, I_{E_1} I_{E_2} [\mathbb{E}\, I_{A_{n-1}} | X_1, \ldots, X_{n-1}] \\
&= v + \mathbb{E}\, I_{E_1} I_{E_2^c} + \mathbb{E}\, I_{E_1} I_{E_2} I_{E_3^c} + \mathbb{E}\, I_{E_1} I_{E_2} I_{E_3} [\mathbb{E}\, I_{A_{n-1}} | X_1, \ldots, X_{n-1}] \\
&= v + \mathbb{P}(E_1 E_2^c) + \mathbb{P}(E_1 E_2 E_3^c) + \cdots + \mathbb{E}\, I_{E_1 E_2 \cdots E_n} [\mathbb{E}\, I_{A_{n-1}} | X_1, X_2, \cdots, X_{n-1}].
\end{aligned}$$

Note that we have the equivalent events

$$E_1 E_2 \cdots E_{k-1} E_k^c = E_1 E_2 \cdots E_{k-1} A_{k-1}, \qquad 1 < k \leq n,$$

from which the desired result in (29) follows. ■

Consider the partial derivatives of C. Denote the k-th partial derivative by

$$D_k C(x_1,\ldots,x_{n-1},y) = \frac{\partial}{\partial x_k} C(x_1,\ldots,x_{n-1},y),$$

and the first k partial derivative, for $k = 1,\ldots,n-1$ by

$$D_{1,\ldots,k} C(x_1,\ldots,x_{n-1},y) = \frac{\partial^k}{\partial u_1,\ldots,u_k} C(u_1,\ldots,u_{n-1},y)\Big|_{(x_1,\ldots,x_{n-1},y))}.$$

More generally, if we let $J \subset \{1,\ldots,n-1\}$ be such that $J = \{j_1,\ldots,j_k\}$, $j_i \neq j_l$, for $i \neq l$, then

$$D_J C(x_1,\ldots,x_{n-1},y) = \frac{\partial^k}{\partial u_{j_1},\ldots,u_{j_k}} C(u_1,\ldots,u_{n-1},y)\Big|_{(x_1,\ldots,x_{n-1},y))}.$$

Note that if $(x_1,\ldots,x_n) \in e_1 e_2 \cdots e_n$, it means that, given v, there is a solution y for the equation $C(x_1,\ldots,x_{n-1},y) = v$. Otherwise, for any $0 \leq y \leq 1$, we have $C(x_1,\ldots,x_{n-1},y) < v$, even if $v \leq x_i \leq 1$ for all $i = 1,\ldots,n-1$. For example, assuming $C(x_1,x_2,y) = x_1 x_2 y$, $v = \frac{1}{2}$, $v < x_1 = x_2 = \frac{2}{3}$, then $C(x_1,x_2,y) < C(x_1,x_2,1) = x_1 x_2 < \frac{1}{2}$. For simplicity, we use $\mathbf{X}_k$ to denote the vector $(X_1,\ldots,X_k)$ with $\mathbf{x}_k = (x_1,\ldots,x_k)$ for $k = 1,\ldots,n-1$.

Lemma 5.2 *For any $k = 1,\ldots,n-1$, we have*

$$\mathbb{E}\, I_{E_1 E_2 \cdots E_k}[\mathbb{E}\, I_{A_k} | \mathbf{X}_k] = \int \cdots \int_{e_1 e_2 \cdots e_k} D_{1,2,\cdots,k} C(\mathbf{x}_k, y_{\mathbf{x}_k,v}, 1,\ldots,1) d\mathbf{x}_k,$$

where $y_{\mathbf{x}_k,v} = \{y | C(\mathbf{x}_k, y, 1,\ldots,1) = v\} = \{y | C(x_1,\ldots,x_k,y,1,\ldots,1) = v\}$.

Proof. According to Theorem 2.27 in Schmitz (2003), we have

$$\begin{aligned}
\mathbb{E}\, I_{E_1 E_2 \cdots E_k}[\mathbb{E}\, I_{A_k} | \mathbf{X}_k] &= \int_{E_1 E_2 \cdots E_k} \mathbb{P}(Y \leq y_{\mathbf{x}_k,v} | \mathbf{X}_k = \mathbf{x}_k) d\mathbb{P} \\
&= \int \cdots \int_{e_1 e_2 \cdots e_k} D_{1,2,\cdots,k} C(\mathbf{x}_k, y_{\mathbf{x}_k,v}, 1,\ldots,1) d\mathbf{x}_k,
\end{aligned}$$

which proves the lemma. ∎

Therefore, by writing

$$\lambda_k(v) = -\int \cdots \int_{e_1 e_2 \cdots e_k} D_{1,2,\cdots,k} C(\mathbf{x}_k, y_{\mathbf{x}_k,v}, 1,\ldots,1) d\mathbf{x}_k,$$

we can express equation (29) as

$$K_n(v) = v - \sum_{k=1}^{n-1} \lambda_k(v). \tag{30}$$

It is straightforward to show that we recover formula (28) in the case where we have $n = 2$.

Example 5.3 ***(The case of independence)*** *Consider the case where C is the independent copula, i.e. $C(x_1,\ldots,x_k,y)=x_1\cdots x_k y$. For $n=2$, we have*

$$\lambda_1(v)=-\int_v^1 D_1C(x_1,y_{x_1,v})dx_1=-\int_v^1\frac{v}{x_1}dx_1=v\log(v).$$

In the case where $n=3$, because

$$e_1e_2^c=\left\{(x_1,x_2)|\ v\le x_1\le 1,\ 0\le x_2\le\frac{v}{x_1}\right\}\ \text{and}\ e_1e_2=\left\{(x_1,x_2)|\ v\le x_1\le 1,\ \frac{v}{x_1}\le x_2\le 1\right\},$$

we can derive the expression

$$\lambda_2(v)=-\int\int_{e_1e_2}D_{1,2}C(x_1,x_2,y_{(x_1,x_2),v})dx_1dx_2=-\int_v^1\int_{v/x_1}^1\frac{v}{x_1x_2}dx_1dx_2=-\frac{1}{2}v[\log(v)]^2.$$

Extending this to the case of n-dimension, it can be shown that

$$\lambda_k(v)=\frac{(-1)^{k+1}v}{k!}[\log(v)]^k=-\frac{v}{k!}[-\log(v)]^k,\qquad k=1,\ldots n-1. \tag{31}$$

In fact, we have

$$e_1e_2\cdots e_k=\left\{(x_1,\ldots x_k)|\ v\le x_1\le 1,\ \frac{v}{x_1}\le x_2\le 1,\cdots,\frac{v}{x_1\cdots x_{k-1}}\le x_k\le 1\right\},$$

so that it follows that

$$\lambda_k(v)=-\int_v^1\int_{\frac{v}{x_1}}^1\cdots\int_{\frac{v}{x_1\cdots x_{k-1}}}^1\frac{v}{x_1\cdots x_k}dx_1\cdots dx_k.$$

Now assuming that it holds for $k>1$, that is, $\lambda_k(v)=\frac{(-1)^{k+1}v}{k!}[\log(v)]^k$, we have

$$\begin{aligned}\lambda_{k+1}(v)&=-\int_v^1\int_{\frac{v}{x_1}}^1\cdots\int_{\frac{v}{x_1\cdots x_k}}^1\frac{v}{x_1\cdots x_{k+1}}dx_1\cdots dx_{k+1}\\&=-\int_v^1 dx_1\left(\int_{\frac{v}{x_1}}^1\cdots\int_{\frac{v}{x_1\cdots x_k}}^1\frac{v/x_1}{x_2\cdots x_{k+1}}dx_2\cdots dx_{k+1}\right)\\&=-\int_v^1\frac{(-1)^{k+1}(v/x_1)}{k!}[\log(v/x_1)]^k dx_1\\&=\frac{(-1)^{k+2}v}{(k+1)!}[\log(v)]^{k+1}=-\frac{v}{(k+1)!}[-\log(v)]^{k+1},\end{aligned}$$

and so by mathematical induction, equation (31) holds. ■

In extending these results to the case where we have distortion of the copula of the third kind as defined in Definition 3.6, we introduce the following notations:

$$y^*_{\mathbf{x}_k,v}=\{y|C_g(\mathbf{x}_k,y,1,\ldots,1)=v\}$$

and

$$C_{\mathbf{X}}(g(x_1), \ldots, g(x_k), g(y^*_{\mathbf{x}_k,v}), 1, \ldots, 1) = g(C_g(\mathbf{x}_k, y^*_{\mathbf{x}_k,v}, 1, \ldots, 1)) = g(v)$$

where $y_{(g(x_1),\ldots,g(x_k),g(v))} = g(y^*_{\mathbf{x}_k,v})$. Define

$$\lambda^g_k(v) = -\int \cdots \int_{e_1 e_2 \cdots e_k} D_{1,2,\cdots,k} C_g(\mathbf{x}_k, y^*_{\mathbf{x}_k,v}, 1, \ldots, 1) d\mathbf{x}_k. \tag{32}$$

We make the following proposition.

Proposition 5.4 *For any $k = 1, \cdots, n-1$, we have*

$$\begin{aligned}\lambda^g_k(v) &= \frac{\lambda_k(g(v))}{g'(v)} + \sum_{m=2}^{k} \sum_{\sum_{i=1}^m J_i = \{1,\ldots,k\}} f_m(v) \times \\ &\quad \int \cdots \int_{g(e_1 e_2 \cdots e_k)} D_{J_1} C_{\mathbf{X}}(\mathbf{z}_k, y^*_{\mathbf{z}_k,g(v)}, 1, \ldots, 1) \cdots D_{J_m} C_{\mathbf{X}}(\mathbf{z}_k, y^*_{\mathbf{z}_k,g(v)}, 1, \ldots, 1) d\mathbf{z}_k,\end{aligned}$$

where $g(e_1 e_2 \cdots e_k) = \{\mathbf{z}_k | z_1 = g(x_1), \cdots, z_k = g(x_k), \mathbf{x}_k \in e_1 e_2 \cdots e_k\}$ and $f_m(v)$ stands for $d^m g^{-1}(t)/dt^m$ evaluated at $t = g(v)$.

Proof. First, notice that

$$D_1 C_g(\mathbf{x}_k, y, 1, \ldots, 1) = g'(x_1) \frac{1}{g'(g^{-1}(C_{\mathbf{X}}(g(\mathbf{x}_k), y, 1, \ldots, 1)))} D_1 C_{\mathbf{X}}(g(\mathbf{x}_k), y, 1, \ldots, 1)$$

and

$$\begin{aligned} & D_{12} C_g(\mathbf{x}_k, y, 1, \ldots, 1) \\ = \; & g'(x_1) g'(x_2) \left[\frac{D_{12} C_{\mathbf{X}}(g(\mathbf{x}_k), y, 1, \ldots, 1)}{g'(g^{-1}(C_{\mathbf{X}}(g(\mathbf{x}_k), y, 1, \ldots, 1)))} - \frac{g''(g^{-1}(C_{\mathbf{X}}(g(\mathbf{x}_k), y, 1, \ldots, 1)))}{[g'(g^{-1}(C_{\mathbf{X}}(g(\mathbf{x}_k), y, 1, \ldots, 1)))]^3} \times \right. \\ & \left. D_1 C_{\mathbf{X}}(g(\mathbf{x}_k), y, 1, \ldots, 1) D_2 C_{\mathbf{X}}(g(\mathbf{x}_k), y, 1, \ldots, 1) \right]. \end{aligned}$$

By mathematical induction, it can be shown that for the case $k > 2$, we have

$$\begin{aligned} D_{1\cdots k} C_g(\mathbf{x}_k, y, 1, \ldots, 1) &= \prod_{i=1}^{k} g'(x_1) \cdots g'(x_k) \left[\sum_{m=1}^{k} \sum_{\sum_{i=1}^m J_i = \{1,\ldots,k\}} (g^{-1})^{(m-1)}(v) \times \right. \\ & \left. D_{J_1} C_{\mathbf{X}}(g(\mathbf{x}_k), y, 1, \ldots, 1) \cdots D_{J_m} C_{\mathbf{X}}(g(\mathbf{x}_k), y, 1, \ldots, 1) \right]. \end{aligned}$$

Making the change of variable $z = g(x)$, we have

$$\begin{aligned} & \int \cdots \int_{e_1 e_2 \cdots e_k} D_{1,2,\cdots,k} C_g(\mathbf{x}_k, y^*_{\mathbf{x}_k,v}, 1, \ldots, 1) d\mathbf{x}_k \\ = \; & \int \cdots \int_{g(e_1 e_2 \cdots e_k)} \frac{1}{g'(v)} D_{1,2,\cdots,k} C_{\mathbf{X}}(\mathbf{z}_k, y_{\mathbf{z_k},g(v)}, 1, \ldots, 1) d\mathbf{z}_k \\ & + \left\{ \sum_{m=2}^{k} \sum_{\sum_{i=1}^m J_i = \{1,\cdots,k\}} \int \cdots \int_{g(e_1 e_2 \cdots e_k)} (g^{-1})^{(m-1)}(v) D_{J_1} C_{\mathbf{X}}(\mathbf{z}_k, y_{\mathbf{z}_k,g(v)}, 1, \ldots, 1) \right. \\ & \left. \cdots D_{J_m} C_{\mathbf{X}}(\mathbf{z}_k, y_{\mathbf{z}_k,g(v)}, 1, \ldots, 1) d\mathbf{x}_k \right\}. \end{aligned}$$

The last term in the equation above is equal to the term on the right-hand side of the equation in the proposition. ■

The results of Proposition 5.4 is thus used to evaluate the following probability

$$K_n^g(v) = \mathbb{P}(C_g(X_1, \ldots, X_{n-1}, Y) \leq v) = v - \sum_{k=1}^{n-1} \lambda_k^g(v). \tag{33}$$

We note that in our definition of distortion of the third kind, we required g^{-1} to be absolutely monotonic to guarantee the legitimacy of the distorted copula C_g. However, in the examples following the definition, we also observed that this was not a necessary condition, which means that there are other distortion functions g that lead to a legitimate copula C_g. Proposition 5.4 only requires a distortion g leading to a legitimate copula C_g. To illustrate, consider example (3.11) where we distorted the independence copula leading to an Archimedean copula. Thus, the following example derives a formula for computing probabilities of the integral transform, $\mathbb{P}(C_\psi(u_1, \ldots, u_n) \leq v)$, for the class of Archimedean copulas. First, consider the following lemma.

Lemma 5.5 *Let ψ be an Archimedean generator. Then for any $k \geq 1$,*

$$(\psi^{-1}(t))^{(k)} = \frac{d^k \psi^{-1}(t)}{dt^k} = (-1)^k \sum_{r=1}^{k} a(k,r) f_r(t^*) e^{-rt}, \tag{34}$$

where $a(k,1) = 1$, $a(k,k) = 1$, and $a(k,r) = ra(k-1,r) + a(k-1,r-1)$, for $1 < r < k$, and $f_r(t^*)$ stands for $d^k g^{-1}(t^*)/dt^{*k}$, evaluated at $t^* = e^{-t}$.

Proof. Since $\psi^{-1}(t) = g^{-1}(e^{-t}) = g^{-1}(t^*)$, then $(\psi^{-1}(t))' = -(g^{-1}(t^*))' e^{-t}$ and

$$(\psi^{-1}(t))^{(2)} = (g^{-1}(t^*))^{(2)} e^{-2t} + (g^{-1}(t^*))' e^{-t}.$$

Assuming $(\psi^{-1}(t))^{(k-1)} = (-1)^{k-1} \sum_{r=1}^{k-1} a(k-1,r) f_r(t^*) e^{-rt}$ holds for $k > 2$, we have

$$\begin{aligned}
(\psi^{-1}(t))^{(k)} &= (-1)^{k-1} \left(f_{k-1}(t^*) e^{-(k-1)t} + \sum_{r=2}^{k-2} a(k,r) f_r(t^*) e^{-rt} + f_1(t^*) e^{-t} \right)' \\
&= (-1)^{k-1} (-f_k(t^*) e^{-kt} - (k-1) f_{k-1}(t^*) e^{-(k-1)t} + \sum_{r=2}^{k-2} a(k-1,r) f_{r+1}(t^*) e^{-rt}(-e^{-t}) \\
&\quad + (-r) a(k-1,r) f_r(t^*) e^{-rt} + f_2(t^*) e^{-t}(-e^{-t}) - f_1 e^{-t}) \\
&= (-1)^k \sum_{r=1}^{k} a(k,r) f_r(t^*) e^{-rt}.
\end{aligned}$$

Thus, it holds for all $k \geq 1$. ■

Example 5.6 ***(The class of Archimedean copulas)*** *Following example (3.11), we consider the n-dimensional Archimedean copula with*

$$C_g(x_1, \ldots, x_{n-1}, y) = \psi^{-1}\{\psi(x_1) + \cdots + \psi(x_{n-1}) + \psi(y)\}.$$

From formula (33) and Proposition 5.4, we find that for $k = 1$,

$$\lambda_1^g(v) = \frac{\lambda_1(g(v))}{g'(v)} = \frac{g(v)}{g'(v)} \log(g(v))$$

and for $k = 2$,

$$\lambda_2^g(v) = \frac{\lambda_2(g(v))}{g'(v)} - \frac{g''(v)}{(g'(v))^3} \int\int_{g(e_1e_2)} D_1 C_{\mathbf{X}}(z_1, z_2, y_{(z_1,z_2,g(v))}) D_2 C_{\mathbf{X}}(z_1, z_2, y_{(z_1,z_2,g(v))}) dz_1 dz_2.$$

Since

$$g(e_1e_2) = \left\{ (z_1, z_2) \Big| \; g(v) \le z_1 \le 1, \; \frac{g(v)}{z_1} \le z_2 \le 1 \right\},$$

and from the previous example, we then have

$$\begin{aligned} \lambda_2^g(v) &= \frac{\lambda_2(g(v))}{g'(v)} - \frac{g''(v)}{(g'(v))^3} \int_{g(v)}^1 \int_{g(v)/z_1}^1 \left[z_2 \frac{g(v)}{z_1 z_2} \right] \left[z_1 \frac{g(v)}{z_1 z_2} \right] dz_1 dz_2 \\ &= -\frac{1}{2} \frac{g(v)}{g'(v)} [\log(g(v))]^2 + \frac{1}{2} \frac{g''(v) g^2(v)}{(g'(v))^3} [\log(g(v))]^2. \end{aligned}$$

Extending this to the case $k > 2$, *we note that*

$$\begin{aligned} &\int \cdots \int_{g(e_1e_2\cdots e_k)} D_{J_1} C_{\mathbf{X}}(\mathbf{z}_k, y^*_{\mathbf{z}_k, g(v)}, 1, \ldots, 1) \cdots D_{J_m} C_{\mathbf{X}}(\mathbf{z}_k, y^*_{\mathbf{z}_k, g(v)}, 1, \ldots, 1) d\mathbf{z}_k \\ &= \int_{g(v)}^1 \int_{\frac{g(v)}{g(x_1)}}^1 \cdots \int_{\frac{g(v)}{g(x_1)\cdots g(x_{k-1})}}^1 \frac{g^m(v)}{z_1 \cdots z_k} d\mathbf{z}_k \\ &= \frac{(-1)^k g^m(v)}{k!} [\log(g(v))]^k. \end{aligned}$$

From Proposition 5.4, we have

$$\begin{aligned} \lambda_k^g(v) &= \frac{(-1)^k}{k!} [\log(g(v))]^k \sum_{m=1}^k \sum_{J_1+\cdots+J_m=\{1,\cdots,k\}} f_m(v) g^m(v) \\ &= \frac{(-1)^k}{k!} [\log(g(v))]^k \sum_{m=1}^k a(k,m) f_m(v) g^m(v), \end{aligned}$$

where we write

$$\sum_{m=1}^{k-1} \sum_{J_1+\cdots+J_m=\{1,\cdots,k-1\}} f_m(v) g^m(v) = \sum_{m=1}^{k-1} a(k-1,m) f_m(v) g^m(v).$$

Note that the index sets $\{J_1, \cdots, J_m, k\}, \{\{J_1, k\}, \cdots, J_m\}, \cdots, \{J_1, \cdots, \{J_m, k\}\}$, *for any* $J_1 + \cdots + J_m = \{1, \cdots, k-1\}$, *is equivalent to* $J_1 + \cdots + J_m = \{1, \cdots, k\}$ *such that we have*

$$\begin{aligned} &\sum_{m=1}^k \sum_{J_1+\cdots+J_m=\{1,\cdots,k\}} f_m(v) g^m(v) \\ &= \sum_{m=1}^{k-1} \sum_{J_1+\cdots+J_m=\{1,\cdots,k-1\}} m f_m(v) g^m(v) + \sum_{m=1}^{k-1} \sum_{J_1+\cdots+J_m+\{k\}=\{1,\cdots,k\}} f_m(v) g^{m+1}(v) \\ &= \sum_{m=1}^{k-1} m a(k-1,m) f_m(v) g^m(v) + \sum_{m=1}^{k-1} a(k-1,m-1) f_m(v) g^{m+1}(v) \\ &= \sum_{m=1}^k a(k,m) f_m(v) g^m(v). \end{aligned}$$

By mathematical induction, it follows that for any $k \geq 1$,

$$\lambda_k^g(v) = \frac{(-1)^k}{k!}[\log(g(v))]^k \sum_{m=1}^{k} a(k,m) f_m(v) g^m(v).$$

If we let $t = \psi(v)$, then $g(v) = e^{-t}$. Based on Lemma 5.5, we have

$$\lambda_k^g(v) = \frac{(-1)^k}{k!}[\psi(v)]^k[(\psi^{-1}(t))^{(k)}]. \tag{35}$$

It can easily be verified that formula (35) conforms well with the result in Section 5 of Genest and Rivest (2001).

6 Concluding remarks

This article extends the notion of probability distortion within the multivariate framework. Distortion in the univariate sense has been widely studied in the actuarial, insurance and financial literature, with a variety of applications: pricing of insurance and financial risks, quantifying risk measures, accommodating parameter uncertainty, to illustrate a few. These same set of applications can be conceptually extended to portfolios of correlated risks when applying distortion of copulas. In financial pricing, for example, it has been used to price Collaterizerd Debt Obligations (CDOs) where the payouts depend on an underlying portfolio of securities, see Crane and van der Hoek (2008). Wang (2007) applied a similar notion of distortion in pricing and measuring of multivariate risks, calling it exponential tilting. Risk analysts are constantly dealing with portfolios that involve multivariate risks; the notion of distortion of a copula can be a tool for enhancing their portfolio models to accommodate parameter uncertainty thereby accurately reflecting the magnitude of risks.

Our contribution to the literature involves carefully crafting the notion of distortion in the multivariate sense. When one distorts an existing copula, one must be careful to ensure that the resulting distortion will still lead to preserving the properties of a copula. Otherwise, the distortion can be mistakenly applied with dangerous implications. This is the primary reason why we examined three different approaches to the extension. First is the distortion of the first kind which distorts only the margins while preserving the original copula structure. Next is the distortion of the second kind whereby we apply the distortion on the margins while simultaneously altering the copula structure. Finally, in the distortion of the third kind, we synchronized the distortion of the copula and its respective margins. We further examined the notion of multivariate ordering of risks within these distortion frameworks. We primarily focused on the notion of supermodularity because this is the type of ordering commonly applied when several risks are concerned.

Finally, ancillary to the work completed in this article, we are happy to extend the formula developed by Genest and Rivest (2001) for computing the distribution of the probability integral transformation of a random vector and even further extend it to the case within the distortion framework. As far as the authors are concerned, these extension formulas to higher dimension has never appeared in the literature.

References

Nicole Bäuerle. Inequalities for stochastic models via supermodular orderings. *Stochastic Models*, 13:181–201, 1997.

Umberto Cherubini, Elisa Luciano, and Walter Vecchiato. *Copula Methods in Finance*. John Wiley & Sons, Ltd: West Sussex, England, 2004.

Glenis Crane and John van der Hoek. Using distortion of copulas to price synthetic CDOs. *Insurance: Mathematics and Economics*, 42:903–908, 2008.

Michel Denuit, Jan Dhaene, Marc Goovaerts, and Rob Kaas. *Actuarial Theory of Dependent Risks*. John Wiley & Sons, Ltd: West Sussex, England, 2005.

William Feller. *An Introduction to Probability Theory and its Applications, Vol. 2*. New York: John Wiley and Sons, Inc., 1971.

Edward W. Frees and Emiliano A. Valdez. Understanding relationships using copulas. *North American Actuarial Journal*, 2:1–25, 1998.

Christian Genest and Louis-Paul Rivest. On the multivariate probability integral transformation. *Statistics and Probability Letters*, 53:391–399, 2001.

Harry Joe. *Multivariate Models and Dependence Concepts*. Chapman and Hall: Boca Raton, Florida, 1997.

Masaaki Kijima. A multivariate estension of equilibrium pricing transforms: the multivariate esscher and wang transforms for pricing financial and insurance risks. *ASTIN Bulletin*, 36:269–283, 2006.

Stuart A. Klugman and Rahul Parsa. Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics*, 24:139–148, 1999.

Zinoviy M. Landsman and Emiliano A. Valdez. Tail conditional expectations for elliptical distributions. *North American Actuarial Journal*, 7:55–71, 2003.

Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press: Princeton, New Jersey, 2005.

Patricia M. Morillas. A method to obtain new copulas from a given one. *Metrika*, 61:169–184, 2005.

Alfred Müller and Marco Scarsini. Some remarks on the supermodular order. *Journal of Multivariate Analysis*, 73:107–119, 2000.

Roger B. Nelsen. *An Introduction to Copulas*. Springer-Verlag: New York, 2006.

A. Wayne Roberts and Dale E. Varberg. *Convex functions*. Academic Press: New York, 1973.

Volker Schmitz. Copulas and stochastic processes. Thesis, Institute of Statistics, Aachen University, 2003.

Abe Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institute de Statistique de L'Universite de Paris*, 8:229–231, 1959.

Shaun S. Wang. Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, 21:71–92, 1996.

Shaun S. Wang. Normalized exponential tilting: pricing and measuring multivariate risks. *North American Actuarial Journal*, 11:89–99, 2007.

David V. Widder. *The Laplace Transform*. Princeton University Press: New Jersey, 1946.

Menahem E. Yaari. The dual theory of choice under risk. *Econometrica*, 55:95–115, 1987.

Appendix: Wang transform is absolutely monotonic

In this appendix, we prove that the Wang transform is absolutely monotonic using mathematical induction. Define $s = g(t) = \Phi(\Phi^{-1}(t)+\gamma)$, for $\gamma \geq 0$, where Φ is the standard Normal distribution function. Then $g^{-1}(s) = \Phi(\Phi^{-1}(s) - \gamma)$ and if we let $s^* = \Phi^{-1}(s)$, we have the first and second order derivatives of g^{-1} :

$$(g^{-1}(s))' = \phi(s^* - \gamma)\frac{ds^*}{ds} = \frac{\phi(s^* - \gamma)}{\phi(s^*)} = e^{-\gamma^2/2}e^{\gamma s^*}$$

and

$$(g^{-1}(s))^{(2)} = \gamma e^{-\gamma^2/2}e^{\gamma s^*}\frac{ds^*}{ds} = \gamma e^{-\gamma^2/2}e^{\gamma s^*}\frac{1}{\phi(s^*)},$$

where ϕ is the density of a standard Normal. Clearly, both derivatives are non-negative. Applying Leibnitz rule together with the above results, for any $n > 2$, the $(n+1)$-th order derivatives of g^{-1} can be expressed as

$$(g^{-1}(s))^{(n+1)} = e^{-\gamma^2/2}\sum_{k=0}^{n}\binom{n}{k}(e^{\gamma s^*})^{(k)}(s^*)^{(n-k)}.$$

This implies that if $(s^*)^{(n)} \geq 0$ for any $n > 0$, then $(g^{-1}(s))^{(n)} \geq 0$ for all $n > 0$. Indeed, we have

$$\begin{aligned}
(s^*)^{(n+1)} &= \left[\frac{1}{\phi(s^*)}\right]^{(n)} = \sqrt{2\pi}(e^{s^{*2}/2})^{(n)} \\
&= \sqrt{2\pi}\sum_{j=0}^{n}\binom{n}{j}(e^{s^{*2}/2})^{(j)}(s^{*2}/2)^{(n-j)} \\
&= \sqrt{2\pi}\left(e^{s^{*2}/2}s^*\frac{ds^*}{ds}\right)^{(n-1)} + \sqrt{2\pi}\left(s^*\frac{ds^*}{ds}\right)^{(n-1)} \\
&\quad + \sqrt{2\pi}\sum_{j=1}^{n-1}\binom{n}{j}\left(e^{s^{*2}/2}s^*\frac{ds^*}{ds}\right)^{(j-1)}\left(s^*\frac{ds^*}{ds}\right)^{(n-j-1)}.
\end{aligned} \tag{36}$$

Because the highest order of derivatives of s^* in formula (36) is n, assuming $(s^*)^{(n)} \geq 0$, we have

$$(s^*)^{(n+1)} = (\Phi^{-1}(s))^{(n+1)} \geq 0.$$

By the process of inductive reasoning, we conclude that n-th order derivative of g^{-1} is non-negative for any $n > 0$. This proves that the Wang transform is absolutely monotonic.