Copulas, diagonals, and tail dependence

Fabrizio Durante\textsuperscript{*a}, Juan Fernández-Sánchez\textsuperscript{b}, Roberta Pappadà\textsuperscript{a,c}

\textsuperscript{a}School of Economics and Management
Free University of Bozen-Bolzano, Bolzano, Italy
\textsuperscript{b}Grupo de Investigación de Análisis Matemático
Universidad de Almería, Almería, Spain
\textsuperscript{c}Department of Statistical Sciences
University of Padua, Padua, Italy

Abstract
We present some known and novel aspects about bivariate copulas with prescribed diagonal section by highlighting their use in the description of the tail dependence. Moreover, we present the tail concentration function (which depends on the diagonal section of a copula) as a tool to give a description of tail dependence at finite scale. The tail concentration function is hence used to introduce a graphical tool that can help to distinguish different families of copulas in the copula test space. Moreover, it serves as a basis to determine the grouping structure of different financial time series by taking into account their pairwise tail behaviour.

Key words: Copulas, Diagonal section, Risk Management, Tail Dependence

1. Introduction

Since their introduction in the seminal work by A. Sklar [100, 101], copulas have been largely employed in the construction and estimation of multivariate stochastic models. As can be testified by a number of recent investigations and monographs devoted to the topic, copulas have enjoyed a great popularity in different applied sciences, especially when the major issue is to understand/quantify a risk coming from different sources. See, for instance, [9, 57, 58, 67, 72, 76, 94] and the (many) references therein.

Copulas are the functions that allow to aggregate individual risk factors (usually, expressed in terms of random variables) into one global risk output. Generally, such global risk is the multivariate probability distribution function coupling the individual (one-dimensional) risks by means of a copula, as outlined by Sklar’s recipe [100, Theorem 3]. In fact, it is known that, given a continuous multivariate random vector $\mathbf{X} = (X_1, \ldots, X_d)$ such that each $X_i$ is distributed according to a probability law $F_i$, the joint probability
distribution function $F$ of $X$ may be expressed as

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)).$$  \hspace{1cm} (1)

See also [14, 22] for Sklar’s theorem in a non-continuous setting.

In practical applications, however, the global risk is an object that is derivable from the multivariate probability distribution function, but may not directly coincide with it. In quantitative risk management, for instance, the global output can be a number, usually the Value-at-Risk (or any other suitable monetary risk measure) of the weighted portfolio [37, 89], or a vector, as in the case of vector-valued risk measures [12]. In environmental sciences, instead, the global risk is usually represented by a number, interpreted as the average return period before the occurrence of a dangerous event [88], or by a curve (the critical layer) in a multidimensional Euclidean space, as in the case of Kendall’s return period [47, 93, 95].

In all these situations, it is remarkable that the global risk is sensitive to the behaviour of the multivariate distribution function in some specific regions of its domain, usually corresponding to large quantile exceedances. Translated into the copula space, this implies that the behaviour of the copula in specific regions, usually neighbourhoods of the vertices of the domain $[0, 1]^d$, contain most of the information that are useful to estimate the risk. In Statistics, the latter regions are usually called tails of the distributions. Moreover, the specification of the dependence structure in the tails represents one of the main advantages of copula models. As stressed, for instance, by the Basel Committee on Banking Supervision\footnote{Developments in Modelling Risk Aggregation, October 2010, http://www.bis.org/publ/joint25.htm}: “The copula approach allows the practitioner to precisely specify the dependencies in the areas of the loss distributions that are crucial in determining the level of risk”.

The determination of copulas with a specific tail behaviour may hence allow to estimate correctly the region of the distribution that is most needed. As such, several investigations have been carried out during the years from different perspectives, ranging from extreme-value analysis [48, 69] to the concept of threshold copulas [55, 56].

In this work, we would like to review a specific (yet wide) method to construct (bivariate) copulas with given tail behaviour that is grounded on the knowledge of its diagonal section. In copula theory, methods of this type originated in [44, 83], even if similar studies have been conducted in the general theory of distribution function in [90, 91, 92]. The relevance of such constructions is mainly due to the fact that they are directly connected to a popular measure of tail behaviour, which is the tail dependence coefficient.

Interestingly, the idea of considering constructions of copulas (or, generally, aggregation functions) with given diagonal section originated in preliminary works about triangular norms [97, Chapter 5]. In particular, one question was whether the continuity of the diagonal section implies the continuity of the associated Archimedean triangular norm. A related problem in this literature has been to determine whether the knowledge of the diagonal section allows to determine uniquely the corresponding Archimedean triangular norms [2] (see also [39, 78, 79]). Such problems, which have their origin in the theoretical under-
standing of the behaviour of associative functions, have now found practical implications in copula estimation, as testified in [19, 51, 98].

The main goal of this paper is twofold. First, we review some basic facts about copulas with given diagonal sections and discuss possible related problems, showing how the use of recent analytical methods to deal with copulas may help in proving novel results. Secondly, we present some graphical tools about copulas with given diagonal section that can assist the decision maker to choose the relevant copula for the problem at hand, especially when, due to data scarcity, classical goodness-of-it techniques [40] may not be efficient.

The paper is organized as follows. Section 2 presents some known and complementary results about copulas with given diagonal section. Section 3 presents the tail concentration function as a tool to get a better idea about the tail behaviour of different copulas. Section 4, instead, introduces some novel tools to detect different tail behaviour. One tool proposes a tail-dependent 2D visualization of the copula test space that can be associated to a given dataset. The other one introduces a clustering procedures to grouping time series. Finally, Section 5 concludes.

2. Diagonal sections of copulas: basic properties with complements

Throughout the paper we adopt basic definitions and properties about copula theory, which can be found, for instance, in [35, 82].

Let \( \mathbb{I} := [0, 1] \). For any function \( F: \mathbb{I}^d \rightarrow \mathbb{R} \) we denote by \( \delta_F \) its diagonal section given by \( \delta_F(t) = F(t, t, \ldots, t) \) for all \( t \in \mathbb{I} \). The natural question is whether, given a suitable function \( \delta: \mathbb{I} \rightarrow \mathbb{I} \), we may obtain a \( d \)-dimensional copula whose diagonal coincides with \( \delta \). The full characterization of the problem is provided in [13, Proposition 5.1] and it is reproduced here.

**Theorem 2.1.** Let \( d \in \mathbb{N}, d \geq 2 \). The following statements hold:

(a) \( \delta: \mathbb{I} \rightarrow \mathbb{I} \) is the diagonal section of a \( d \)-dimensional copula \( C \).

(b) \( \delta \) satisfies the following conditions:

- (D1) \( \delta(1) = 1 \);
- (D2) \( \delta(t) \leq t \) for all \( t \in \mathbb{I} \);
- (D3) \( \delta \) is increasing;
- (D4) \( |\delta(v) - \delta(u)| \leq d|v - u| \) for all \( u, v \in \mathbb{I} \).

In the following, we mainly restrict our attention to the 2-dimensional case. We denote by \( C_\delta \) the class of copulas with diagonal section equal to \( \delta \) and by \( \mathcal{D}_2 \) the class of all possible diagonals, i.e. all functions \( \delta: \mathbb{I} \rightarrow \mathbb{I} \) satisfying (D1)–(D4) for \( d = 2 \). Notice that every \( \delta \in \mathcal{D}_2 \) admits derivative almost everywhere with respect to the Lebesgue measure. Furthermore (see, for example, [71]), \( \delta' \) is integrable and for any \( s, t \in \mathbb{I}, s < t \), \( \int_s^t \delta'(x)dx = \delta(t) - \delta(s) \).
In the 2-dimensional case, the problem of determining necessary and sufficient conditions for a function to be the diagonal section of a copula was explicitly mentioned and solved in [44, 83]. The solution is contained in the following result.

**Theorem 2.2.** Let $\delta \in \mathcal{D}_2$. Then the function $K_\delta : \mathbb{I}^2 \to \mathbb{I}$ given, for all $(u, v) \in \mathbb{I}^2$, by

$$K_\delta(u, v) = \min \left\{ u, v, \frac{\delta(u) + \delta(v)}{2} \right\}$$

is a copula in $C_\delta$, called diagonal copula.

Notice that, based on some ideas from [59], a diagonal copula has a nice probabilistic interpretation (which also provides a tool to prove that they are actually copulas). In fact, note that, for every $(x, y) \in \mathbb{I}^2$, we have

$$K_\delta(x, y) = \min \left\{ \frac{2x - \delta(x)}{2}, \delta(y) \right\} + \min \left\{ \frac{\delta(x)}{2}, 2y - \delta(y) \right\},$$

thus it is a convex combination of two bivariate distribution functions with the same copula, namely $M(u, v) = \min(u, v)$, but different marginals (see also [21]). This probabilistic interpretation provides also an immediate way for simulating a random sample from diagonal copulas. In the following, we set $\hat{\delta}(t) = 2t - \delta(t)$ for every $t \in \mathbb{I}$.

**Algorithm 2.1.**

1. Generate a random number $u \in \mathbb{I}$.
2. If $u \leq 0.5$, then
   (a) Generate a random number $s \in \mathbb{I}$.
   (b) Set $(u, v) = (\hat{\delta}^{-1}(s), \hat{\delta}^{-1}(s))$.
   Otherwise
   (a) Generate a random number $t \in \mathbb{I}$.
   (b) Set $(u, v) = (\delta^{-1}(t), \delta^{-1}(t))$.
3. Return $(u, v)$.

A random sample generated from the diagonal copula $K_\delta$ with $\delta(t) = t^2$ is illustrated in Figure 1. As it is evident from the chart, the support of such a copula is contained in the graph of two univariate functions. In fact, if $\delta \in \mathcal{D}_2$ is strictly increasing such that $\hat{\delta}(t) = 2t - \delta(t)$ is also strictly increasing, then $K_\delta$ concentrates the probability mass on the so-called hairpin set generated by $g = \delta^{-1} \circ \delta$. Specifically, the support of $K_\delta$ is contained in the graphs of the functions $g$ and $g^{-1}$. For more details, see [24, 83].

Another example of copulas in $C_\delta$ could be also derived from Bertino [4] (see also [41, 45]) and is reproduced here.

**Theorem 2.3.** For every diagonal $\delta \in \mathcal{D}_2$, let $\hat{\delta} : \mathbb{I} \to \mathbb{I}$ given by $\hat{\delta}(t) := t - \delta(t)$. Then the function $B_\delta : \mathbb{I}^2 \to \mathbb{I}$ defined by

$$B_\delta(u, v) := \min \{u, v\} - \min \{\hat{\delta}(t) : t \in [u \wedge v, u \vee v]\}$$

is a copula in $C_\delta$. 

4
The importance of Bertino copulas stems from the fact that they form a sharp pointwise

A random sample generated from the Bertino copula $B_\delta$ with $\delta(t) = t^2$ is illustrated in Figure 2. The sampling algorithm has been carried out via the conditional distribution method described, for instance, in [72]. As it is evident from the chart, the support of such a copula is contained in the two diagonals of $\mathbb{I}^2$. Copulas with this kind of support are also called X–copulas in [45, Example 2.9].

The importance of Bertino copulas stems from the fact that they form a sharp pointwise
lower bound in $\mathcal{C}_\delta$. The upper bound of the class, instead, may not be a copula, as summarized in the following result (for details, see [84, 64]).

**Theorem 2.4.** For every diagonal $\delta \in \mathcal{D}_2$ and for every $C \in \mathcal{C}_\delta$, we have

$$B_\delta \leq C \leq A_\delta,$$

(4)

where $A_\delta : \mathbb{I}^2 \to \mathbb{I}$ is the quasi-copula defined by

$$A_\delta(u, v) := \min \left\{ u, v, \max\{u, v\}, \max\{\hat{\delta}(t) : t \in [u \land v, u \lor v]\} \right\}.$$  

(5)

The lower bound of (4) is obviously sharp since it is a copula, while the upper bound may not be a copula and need not be sharp. In [104] conditions are given in order to ensure that $A_\delta = K_\delta$ and, hence, prove the sharpness of the upper bound under suitable assumptions on $\delta$. Notice that the diagonal copula is an upper bound in the class of symmetric copulas in $\mathcal{C}_\delta$ (see, e.g., [85, Theorem 2]).

**Remark 2.1.** Upper and lower bounds of (4) are also valid in the class of quasi-copulas having diagonal section equal to $\delta$. In such a case, they are obviously sharp.

Notice that both diagonal and Bertino copulas are symmetric and singular (for this concept, see, for instance, [23]). However, these properties are obviously not shared by all the elements in $\mathcal{C}_\delta$. In fact, the following facts hold:

- $\mathcal{C}_\delta$ is a rich set, in the sense that it contains an infinitude of elements up to the case when $\delta$ coincides with the identity function of $\mathbb{I}$, namely $\delta = \text{id}_I$ (see [21]). In fact, any convex combination of the diagonal and the Bertino copula with the same diagonal $\delta$ gives a copula with diagonal $\delta$. Moreover, from any copula with diagonal section different from $\text{id}_I$, a non-symmetric copula with the same diagonal may be constructed (see, e.g., [34, section 3.3]). Notice that other asymmetric copulas can be also obtained by the diagonal patchwork (see [28, 33, 85]).

- $\mathcal{C}_\delta$ contains absolutely continuous copulas if the set of all fixed points of $\delta$ has Lebesgue measure 0; moreover, if $\delta \neq \text{id}_I$, there exists at least a copula in $\mathcal{C}_\delta$ with a non-trivial absolutely continuous component. In general, given an absolutely continuous copula in $\mathcal{C}_\delta$, we can construct an infinitude of other absolutely continuous copulas in $\mathcal{C}_\delta$ via patchwork techniques (see [16, 34, 38]).

The problem of the construction of absolutely continuous copulas with given diagonal section has been particularly considered (see, for instance, [15, 27, 54]) since the presence of the copula density may be of particular interest when estimation has to be carried out via maximum likelihood techniques. However, most of the given examples do not have full support. The following result shows, instead, how constructions of copulas that are both absolutely continuous and have full support, i.e. the support is $\mathbb{I}^2$, are possible as well. We recall that the support of a copula $C$ is the complement of the union of all open subsets of $\mathbb{I}^d$ with $C$-measure zero. In other words, it is the smallest closed set on which the measure $\mu_C$ generated by $C$ is concentrated.
Theorem 2.5. Let $\delta \in \mathcal{D}_2$. The following statements are equivalent:

(a) $\delta$ is the diagonal of an absolutely continuous copula with full support;

(b) $\delta(t) < t$ for every $t \in (0, 1)$ and there is no interval $J \subset \mathbb{I}$ such that either $\delta' = 0$ on $J$ or $\delta' = 2$ on $J$.

Proof. Suppose that (a) holds. Then there is no $t_0 \in (0, 1)$ such that $\delta(t_0) = t_0$ since, otherwise, the corresponding copula would be an ordinal sum and, hence, cannot have full support. Moreover, if $\delta' = 0$ on an interval $J = (a, b) \subset \mathbb{I}$, it follows that $\delta$ is constant on $J$. Thus any copula $C$ with diagonal section $\delta$ is such that the $C$–measure, or $C$–volume of $[0, b]^2 \setminus [0, a]^2$, is equal to $V_C([0, b]^2 \setminus [0, a]^2) = 0$. Analogously, suppose that $\delta' = 2$ on $J = (a, b) \subset \mathbb{I}$ and $C$ is a copula with diagonal section $\delta$. It follows that the survival copula $\hat{C}$ associated with $C$, namely $\hat{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$ is such that $\delta' = 0$ on $(1 - a, 1 - b)$. Then, as above, $\hat{C}$ has no probability mass concentrated on $[0, 1 - a]^2 \setminus [0, 1 - b]^2$. It follows that $C$ has no probability mass concentrated on $[a, 1]^2 \setminus [b, 1]^2$. In both cases, $C$ cannot have full support.

Conversely, let $\delta \in \mathcal{D}_2$ and suppose that (b) holds. Let $C_{-1} := K_{\delta}$ be the associated diagonal copula which concentrates the probability mass on the graph of the functions $g$ and $g^{-1}$. Starting with $C_{-1}$, we can construct a sequence of copulas that may serve to show our assertion. Since $C_{-1}$ is symmetric, we can limit ourselves to define the following construction of copulas only in the triangle of the unit square with vertices $(0, 0)$, $(1, 0)$ and $(1, 1)$, since the extension to the remaining part of the domain $\mathbb{I}^2$ can be done by symmetry. Therefore, since $g(t) < t$ on $\mathbb{I}$, consider a countable sequence $R_n$ of rectangles that covers the set $\{(t, g(t)): t \in \mathbb{I}\}$ and such that two distinct rectangles have in common only points in their borders. For instance, for a fixed $a \in (0, 1)$, we consider $R_n = [g^{n+1}(a), g^n(a)] \times [g^{n+2}(a), g^{n+1}(a)]$ for every $n \in \mathbb{Z}$. Let $C_0$ be the copula obtained by rectangular patchwork techniques in the following way: $C_0$ coincides with $C_{-1}$ outside $\cup_n R_n$, while in each rectangle $R_n$ the probability mass is distributed according to the independence copula $\Pi_2(u, v) = uv$. The resulting $C_0$ is absolutely continuous and has diagonal section $\delta$, as can be easily deduced from the very definition of rectangular patchwork copulas [34].

Now, consider a dense set $\{x_k\}_{k \in \mathbb{N}} \subset \mathbb{I}$ that contains all the previous points of type $g^{n+1}(a)$ for every $n \in \mathbb{Z}$. Let $C_{0, k}$ be the copula that can be obtained as a rectangular patchwork from $C_0$ by spreading the probability mass on the rectangle $[g(x_k), x_k] \times [g^2(x_k), g(x_k)]$ according to the dependence copula. Obviously, the diagonal section of $C_{0, k}$ is equal to $\delta$. Moreover, $C_1 = \sum_{k \in \mathbb{N}} C_{0, k}$ is an absolutely continuous copula whose support is given the region $\{(x, y) \in \mathbb{I}^2: g^2(x) \leq y \leq g^{-2}(x)\}$. Analogously, by repeating the previous procedure, we can hence determine a copula $C_2$ that is absolutely continuous with diagonal section $\delta$ whose support is given by the region $\{(x, y) \in \mathbb{I}^2: g'(x) \leq y \leq g^{-1}(x)\}$. By iterating this procedure, we may define a sequence of absolutely copulas $(C_m)_{m \in \mathbb{N}}$ with diagonal section equal to $\delta$. Moreover, if $C = \sum_{m \in \mathbb{N}} C_m$, then the support of $C$ is $\mathbb{I}^2$. In fact, if $(x, y) \in \mathbb{I}^2$ with $y < x$, then there exists $m$ such that $y \in [g^m(x), g^{-m}(x)]$. Thus, $(x, y)$ belongs to the support of $C_m$ and, hence, the support of $C$ is $\mathbb{I}^2$. □
Diagonal and Bertino copulas are universal constructions of copulas with given diagonal section, since they provide a mapping that associated to each \( \delta \in \mathcal{P}_2 \) a copula in \( \mathcal{C}_\delta \). In the literature, a number of constructions are also known that work under additional assumptions on the diagonal section \( \delta \). Consider, for instance, the case of semilinear copulas and their extensions (see, e.g., [29, 42, 61, 62, 63]), copulas satisfying special functional equations [30, 75], copulas with diagonal and opposite diagonal sections [17], etc.

Moreover, recently, several methods have been also introduced in order to construct copulas with given diagonal starting from suitable aggregation functions (see, for instance, [33]) or fuzzy negations (see [1]). In particular, the latter method inspired the construction of copulas with given diagonal section proposed in [77], whose characterization (which was stated as open problem) is given below. We recall that a diagonal \( \delta \) is called simple (see, e.g. [85]) if \( \hat{\delta}(t) = t - \delta(t) \) is quasi–concave, i.e. for all \( \alpha \in \mathbb{I} \) and \( (u, v) \in \mathbb{I}^2 \),

\[
\hat{\delta}(\alpha u + (1 - \alpha)v) \geq \min\{\hat{\delta}(u), \hat{\delta}(v)\}.
\]

**Theorem 2.6.** Let \( \delta \) be a diagonal. Define the function \( C : \mathbb{I}^2 \to \mathbb{I} \) given by

\[
C(u, v) = \max\{0, \min\{u, v\} - \min\{\hat{\delta}(u), \hat{\delta}(v)\}\},
\]

(6)

Then \( C \) is a copula with diagonal \( \delta \) if, and only if, \( \delta \) is simple.

**Proof.** If \( \delta \) is simple, then, for every \( (u, v) \in \mathbb{I}^2 \), it follows that

\[
\min\{\hat{\delta}(u), \hat{\delta}(v)\} = \min\{\hat{\delta}(t) : t \in [u \land v, u \lor v]\}
\]

(see [85, Theorem 22 (ii)]). Thus the function \( C \) of (6) is a Bertino copula.

Conversely, suppose ab absurdo that \( \delta \) is not simple. Then, in view of the characterization given in [85, Theorem 22 (iii)], there exist \( z_1, z_2, z_3 \in \mathbb{I} \) such that \( z_1 < z_2 < z_3 \), and \( \hat{\delta}(z_1) = \hat{\delta}(z_3) > \hat{\delta}(z_2) \). In such a case, it holds that \( C(z_1, z_1) = C(z_1, z_3) = z_1 - \delta(z_1) \), while \( C(z_1, z_2) > z_1 - \delta(z_2) \), which is absurd since the horizontal section of a copula must be increasing. \( \square \)

At the end of this section, we discuss a probabilistic interpretation of the diagonal section. Consider the random variables \( U_1, \ldots, U_d \) defined on the same probability space \((\Omega, \mathcal{F}, \mathbb{P})\), having uniform distribution on \((0, 1)\) and \( C \) as their joint distribution function, then for every \( t \in \mathbb{I} \)

\[
\mathbb{P}(\max\{U_1, U_2, \ldots, U_d\} \leq t) = \mathbb{P}\left(\bigcap_{j=1}^{d}\{U_j \leq t\}\right) = C(t, t, \ldots, t).
\]

In other words, the largest order statistics of the vector \( U \) is only determined by the diagonal section of the copula \( C \). Moreover, in the case \( d = 2 \), it can be also proved that

\[
\mathbb{P}(\min\{U_1, U_2\} \leq t) = \mathbb{P}(U \leq t) + \mathbb{P}(V \leq t) - \mathbb{P}\left(\{U \leq t\} \cap \{V \leq t\}\right) = 2t - \delta_C(t).
\]
Thus, determining a copula \( C \) with a prescribed diagonal \( \delta \) is equivalent to determining a random vector \((U, V)\) such that \((U, V) \sim C\) and the marginal d.f.’s of the order statistics of \((U, V)\) are known. In its turn, by recourse to Sklar’s Theorem [100], it follows that the determination of copulas with prescribed diagonal section is equivalent to the following Fréchet problem in the class of distributions:

\[\text{Determine all the bivariate distribution functions having identically distributed univariate margins and prescribed distribution functions of order statistics.}\]

Starting with this interpretation, it is not surprising that some constructions of copulas with given diagonal sections (or with specified order statistics) are contained (in implicit form) in some early works about distribution functions of order statistics. See, for instance, [90, 91] or [59], where explicit constructions of copulas are derived from the literature.

3. Diagonal sections of copulas and tail dependence

The main importance in applications of the diagonal section of a copula is its role in determining the so-called tail dependence coefficients (shortly, TDC’s) of a bivariate copula. Such coefficients were suggested in [99], but their copula-based representation is made explicit in [60]. Here we recall that the quantile function associated with a distribution function \( F \), also called quasi–inverse of \( F \), is given by \( F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\} \).

**Definition 3.1.** Let \( X \) and \( Y \) be continuous random variables with distribution functions \( F_X \) and \( F_Y \), respectively. The upper tail dependence coefficient \( \lambda_U \) (shortly, UTDC) of \((X, Y)\) is defined by

\[
\lambda_U = \lim_{t \to 1^-} \mathbb{P} \left( Y > F_Y^{-1}(t) \mid X > F_X^{-1}(t) \right);
\]

and the lower tail dependence coefficient \( \lambda_L \) (shortly, LTDC) of \((X, Y)\) is defined by

\[
\lambda_L = \lim_{t \to 0^+} \mathbb{P} \left( Y \leq F_Y^{-1}(t) \mid X \leq F_X^{-1}(t) \right);
\]

provided that the above limits exist.

Therefore, the upper tail dependence coefficient indicate the asymptotic limit of the probability that one random variable exceeds a high quantile, given that the other variable exceeds a high quantile. Similar interpretation holds for the case of LTDC. As known (see, for instance, [82]), TDC’s only depend on the copula \( C \) of \((X, Y)\) in view of formulas:

\[
\lambda_L = \lim_{t \to 0^+} \frac{\delta_C(t)}{t} \quad \text{and} \quad \lambda_U = \lim_{t \to 1^-} \frac{1 - 2t + \delta_C(t)}{1 - t}.
\]

Because of their importance for the determination of the main feature of a copula, tail dependence coefficients are available for the popular families of copulas (see Table 1). In
Table 1: Tail dependence coefficients for popular families of copulas.

<table>
<thead>
<tr>
<th>Copula</th>
<th>$\lambda_L$</th>
<th>$\lambda_U$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Elliptical</td>
<td>$\geq 0$</td>
<td>$\geq 0$</td>
</tr>
<tr>
<td>Gumbel</td>
<td>0</td>
<td>$2 - 2^{1/\alpha}$</td>
</tr>
<tr>
<td>Clayton ($\alpha \geq 0$)</td>
<td>$2^{-1/\alpha}$</td>
<td>0</td>
</tr>
<tr>
<td>Frank</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Plackett</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

particular, they have been extensively studied for Archimedean copulas (see, e.g., [7, 8, 52, 68]). For extreme-value copulas, instead, LTDC is always 0 (up to the case of comonotone copula), while UTDC is equal to $2(1 - A(0.5))$, where $A$ is the Pickands dependence function (see, e.g., [48]).

Now, while tail dependence coefficients give an asymptotic approximation of the behaviour of the copula in the tail of the distribution, it might be also of interest to consider the case when the tail behaviour is considered at some (finite) points near the corners of the unit square. The distinction between tail dependence at asymptotic and sub-asymptotic levels has already been made several times in the literature, e.g. [10, 43, 73]. Its practical usefulness has been highlighted in [102].

An auxiliary function that may serve to visualize the tail dependence of a copula $C$ is the so-called tail concentration function (TCF, in short), defined as the function $q_C : (0, 1) \rightarrow \mathbb{I}$ given by

$$q_C(t) = \frac{\delta_C(t)}{t} \cdot 1_{(0,0.5]} + \frac{1 - 2t + \delta_C(t)}{1 - t} \cdot 1_{(0.5,1]}.$$  

Notice that $q_C(0.5) = (1 + \beta_C)/2$, where $\beta_C$ is the Blomqvist’s measure of association associated with $C$ (see [82]). The tail concentration function has been defined, for instance, in [105], while its estimation from empirical data has been presented in [86, 87].

For practical purposes, the tail concentration function can be more suited to assess the risk of joint extremes than its limits given by the UTDC and LTDC. In fact, when the speed of convergence of tail concentration function to the boundary 0 (or 1) is slow, this implies that the dependence in the finite tail can be significantly stronger than in the limit (compare with [73]).

The practical effect of considering the tail concentration function can be summarized in Figure 3. Here, we can note the different tail behaviour of several copulas sharing the same Blomqvist’s measure of association. The left figure displays the TCF plots for copulas with zero lower (upper) tail dependence coefficient, while the right figure considers copulas with non-zero lower (upper) tail dependence (see Table 1). In both figures the Blomqvist’s Beta is set to $\beta_C = 0.5$. As it can be noticed, the TCF of Gaussian copulas seems to converge to 0 (respectively, 1) slower than Frank and Plackett copulas. As regards the copulas with non-negative TDC’s, it seems that the convergence of the TCF to 0 (respectively, 1) is slower in the case of Clayton copulas (respectively, survival Clayton). Thus, Clayton copulas represent a natural choice for a conservative (from a risk manager
viewpoint) estimation of the tail of a joint distribution. Notice that the TCF’s of Gumbel
and Galambos copulas are very close each other.

From these plots, it should be evident the advantage of using the TCF to visualize tail
behaviour, instead of the related diagonal section. In particular, it is clear when there is
asymmetry in the TCF in the lower and upper part of its domain.

In practice, given a random sample \( \{(X_i, Y_i): i = 1, \ldots, n\} \) from the random pair
\((X, Y)\) with copula \(C\), the TCF can be approximated by the following procedure. First,
the marginal distributions are estimated by their empirical versions,

\[
F_n(t) = \frac{1}{n} \sum_{i=1}^{n} 1(X_i \leq t), \quad G_n(t) = \frac{1}{n} \sum_{i=1}^{n} 1(Y_i \leq t)
\]

Secondly, the copula \(C\) is estimated by the empirical copula \(C_n\), given for all \((u, v) \in \mathbb{R}^2\) by

\[
C_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} 1(U_i \leq u, V_i \leq v),
\]

where, for all \(i = 1, \ldots, n\), \(U_i = F_n(X_i)\) and \(V_i = G_n(Y_i)\). Thus, for any \(t > 0\), the
(empirical) tail concentration function is given by

\[
q_{\text{emp}}(t) = \frac{C_n(t, t)}{t} \cdot 1_{(0,0.5]}(t) + \frac{1 - 2t + C_n(t, t)}{1 - t} \cdot 1_{(0.5,1)}(t) \quad (7)
\]
Figure 4 presents the empirical estimation of the TCF for two time series considered in subsection 4.2.

Notice that $C_n$ depends on the sample size. In particular, $C_n(t, t)$ provides accurate estimation of the dependence structure when there is a sufficient number of points in $[0, t]^2$. Therefore, in order to allow more reliable estimate, the empirical tail concentration function is usually calculated on some interval $[\varepsilon, 1]$ for a suitable $\varepsilon > 0$ that is related to sample size.

**Remark 3.1.** Finally, it should be noted that other possible ways to detect finite tail dependence are available in the literature both in functional forms, e.g. tail copulas, threshold copulas (see, e.g., [55, 69] and references therein), and in numerical forms, e.g. conditional measures of association (see, e.g., [20, 25, 96]).

### 4. Graphical tools to detect tail dependence

Here we present two graphical tools that may help in the detection of tail dependence. They are based on the previously introduced TCF.

**4.1. Visualizing the tail-dependent copula-test space**

In [81], a graphical tool has been provided in order to advise the decision maker in the choice of the copula family that better fits given set of pairwise observations.

This method is based on two steps. First, one constructs the so-called *copula-test space*, i.e. the set of all possible families of copulas that are suitable for the data at hand (see [80]). Such a space is constructed, for instance, by taking into account some empirical measure of association calculated from the observations and the ability of a family of copulas to detect
such association. Then, a suitable distance between the empirical copula (as derived from data) and the (fitted) parametric families of copulas in the copula test space is introduced. The calculated distances are hence visualized in a Cartesian plane via principal coordinate analysis in such a way that one can assess visually which copula family is closer (in the given distance) to the empirical copula. In [81] the considered distance is a kind of $L^2$ distance calculated on a finite grid of the copula domain. Loosely speaking, it measures the mean squared error between the graph of the empirical copula and the graph of a fitted parametric copula.

However, as stressed also in [81], in some cases it would be also convenient to choose a metric that provides different weights in the tails. In fact, while a global distance is surely convenient to have an idea about the overall goodness-of-fit of the proposed model, in most practical problems connected with risk management, one is also interested whether the model captures some importance aspects of the tail dependence behaviour. In fact, as for instance, stressed in [53], it may happen that the central part of the distribution influences the estimation more than the tail part, a fact that could be not conservative from the viewpoint of a risk manager that wants to estimates some risk quantities derived from the model (like Value-at-Risk).

Therefore, here we introduce a modification of the graphical tool in [81] in order to detect which families of copulas are closer to the empirical copula in the tail dependence behaviour. Specifically, the main ingredient will be the introduction of a suitable dissimilarity measure based on the tail concentration function that can be used to visualize the distances among the copulas of the test space to the empirical copula.

To this end, let $(X_i, Y_i)_{i=1,\ldots,n}$ be a bivariate sample from an unknown copula; or, better, use rank transformation to obtain from any sample the pseudo-observations that are useful to identify the copula structure (see, e.g. [46]). Consider a set of $k$ copulas $C_1, C_2, \ldots, C_k$ belonging to different families that have been fitted to the available data. A dissimilarity between the empirical copula $C_n$ and the copula $C_i$, $i = 1, \ldots, k$ can be so defined:

$$d(C_n, C_i) = \int_0^1 (q_{\text{emp}}(t) - q_{C_i}(t))^2 dt,$$

where $q_{\text{emp}}$ is the TCF calculated via (7), while $q_{C_i}$ is the TCF associated with $C_i$. In other words, we consider a kind of $L^2$-type distance between the empirical TCF and the TCF of a copula $C_i$ fitted to the observations. Analogously, the dissimilarity between the $i$–th and the $j$–th copula is computed as

$$d(C_i, C_j) = \int_0^1 (q_{C_i}(t) - q_{C_j}(t))^2 dt$$

for $1 \leq i \neq j \leq k$. Notice that $d(C_i, C_j) = 0$ implies that $C_i$ and $C_j$ have the same diagonal section, but they may not coincide.

Obviously, dissimilarities (8) and (9) are both computed from a finite approximation of the TCF's at some points $t_1 < t_2 < \cdots < t_N$ in $I$ (eventually corresponding to the sample size). Nevertheless, even for small sample size, they seem to provide hints to distinguish
among different copulas. To this end, we have done a small simulation study in order to check whether the dissimilarity defined above are able to capture the tail behaviour of different copulas.

Specifically, we simulate $B = 500$ times bivariate observations of sample size $n \in \{250, 500\}$, respectively, from Clayton, Gumbel, and Gaussian copulas (with different values of Kendall’s $\tau$). Then, at each time, we fit a parametric copula family to the data (via inversion of Kendall’s $\tau$). Finally we calculate the dissimilarity between the empirical copula and the fitted copula via (8). The considered families of copulas have been Clayton (denoted by $C_1$), Gumbel ($C_2$), Frank ($C_3$), Gaussian ($C_4$), Plackett ($C_5$), Galambos ($C_6$).

The results obtained after the $B$ repetitions of this process are displayed in the form of box plots, for varying Kendall’s $\tau$ values $\tau \in \{0.25, 0.5, 0.75\}$ (see Figures 5 and 6).

The following considerations can be drawn:

- For the Clayton family ($C_1$), the tail properties of the simulated copula is identified in all situations, regardless of $\tau$-value and sample size. In fact, both for small and large sample sizes the box plots in panels (a) – (c) suggest that $\delta_{\text{emp},1}$ is minimal.

- For the Gumbel family ($C_2$), the simulation results suggest that the true copula is not always recognized as the best-fit copula. In particular, regardless of sample size, for $\tau \leq 0.5$ it seems that copulas $C_3, C_5, C_6$ exhibit small distances as well. For $\tau > 0.5$, the identification of the true data generating process seems more reliable.

- For the Gaussian family ($C_4$) with small sample size ($n = 250$), other copulas have similar behaviour as Gaussian copula for $\tau = 0.25, 0.5$, while for $\tau = 0.75$ the identification process seems to perform better. Moreover, for a larger sample size and $\tau \geq 0.5$, the Gaussian copula is unambiguously identified as true copula.

Given such preliminary results, let us illustrate the graphical procedure to find some possible copula candidates to describe the tail behaviour of bivariate observations with unknown dependence structure.

Let $(X_i, Y_i)_{i=1,\ldots,n}$ be a bivariate sample from an unknown copula. Let $\mathcal{C}^1, \ldots, \mathcal{C}^k$ be possible parametric families of copulas that can be suitable to describe the dependence in the given data (i.e. the copula test space). The procedure goes as follows.

1. For $i = 1, \ldots, k$ fit a copula $C_i$ from the family $\mathcal{C}^i$ by using classical methods (e.g. maximum likelihood estimation, inversion of Kendall’s $\tau$, etc.).

2. For $i = 1, \ldots, k$ calculate the dissimilarity between $C_i$ and the empirical copula $d_{\text{emp},i} := d(C_n, C_i)$ by using (8).

3. For the $k$ copulas $C_1, \ldots, C_k$, calculate the $K = k(k - 1)/2$ mutual dissimilarities $d_{i,j} := d(C_i, C_j)$ by using (9).

4. Consider the symmetric square matrix $D = (\sigma_{ij})$ (of dimension $k + 1$) defined as follows:

\[
\begin{align*}
\sigma_{1j} &= d_{\text{emp},j-1}, & j &= 2, \ldots, k + 1, \\
\sigma_{ij} &= d_{i-1,j-1}, & i, j &= 2, \ldots, k + 1, & i < j \\
\sigma_{ii} &= 0, & i &= 1, \ldots, k + 1.
\end{align*}
\]
Figure 5: Box plots resulting from the simulation study with $n = 250$, for bivariate observations of sample size $n$ from a Clayton ($C_1$), Gumbel ($C_2$) and Normal ($C_4$) copula with $\tau \in \{0.25, 0.5, 0.75\}$. 

(a) Clayton copula, $\tau = 0.25$  
(b) Clayton copula, $\tau = 0.5$  
(c) Clayton copula, $\tau = 0.75$  
(d) Gumbel copula, $\tau = 0.25$  
(e) Gumbel copula, $\tau = 0.5$  
(f) Gumbel copula, $\tau = 0.75$  
(g) Normal copula, $\tau = 0.25$  
(h) Normal copula, $\tau = 0.5$  
(i) Normal copula, $\tau = 0.75$
Figure 6: Box plots resulting from the simulation study with $n = 500$, for bivariate observations of sample size $n$ from a Clayton ($C_1$), Gumbel ($C_2$) and Normal ($C_4$) copula with $\tau \in \{0.25, 0.5, 0.75\}$. 

- (a) Clayton copula, $\tau = 0.25$
- (b) Clayton copula, $\tau = 0.5$
- (c) Clayton copula, $\tau = 0.75$
- (d) Gumbel copula, $\tau = 0.25$
- (e) Gumbel copula, $\tau = 0.5$
- (f) Gumbel copula, $\tau = 0.75$
- (g) Normal copula, $\tau = 0.25$
- (h) Normal copula, $\tau = 0.5$
- (i) Normal copula, $\tau = 0.75$
Such a $D$ is the dissimilarity matrix that describes the relation among $C_n$ (the empirical copula), $C_1, \ldots, C_k$.

5. Starting with the dissimilarity matrix $D$, perform a Multidimensional Scaling (MDS) in order to construct a configuration of points in $q$ dimensions, where the Euclidean distances (in the $q$–dimensional space) between the different copulas has to fit as closely as possible the dissimilarity information (for more details, see [65, 66]). As it is known from classical MDS, the final configuration is such that the distortion caused by a reduction in dimensionality is minimized by means of the so-called stress function (for more details, see [49]).

In practice, a suitable visualization is obtained when $q = 2$, although $q > 2$ could be preferred in some cases (depending on the value of the stress function).

As an illustration of the methodology, consider a set of bivariate observations generated from a known copula and suppose that we would like to check whether the described graphical tool is able to suggest some good candidate copula model for such data. To this end, consider as copula test space the one-parameter copula families $C_i, i = 1, \ldots, k$ mentioned above. Moreover, suppose that the random sample is of size $n = 250$, and is generated by Clayton, Gumbel, Frank, and Gaussian copulas, respectively, with Kendall’s tau $\tau = 0.5$. The results for four different situations are displayed in Figure 7.

Specifically, for each chart of Figure 7, we apply MDS on the matrix of dissimilarities $D$ defined as above, and check the stress function to see whether a 2D representation ($q = 2$) is feasible (in general a good representation should have a stress lower than 2.5%). Then, we plot the $k$ points $p_i = (x_i, y_i)$ corresponding to copula $C_i$ and $p_{emp} = (x_{emp}, y_{emp})$ corresponding to the empirical copula $C_n$ in a 2D graph. Notice that the fitting of parametric copulas has been done via inversion of Kendall’s $\tau$. As can be seen, the charts are often useful to identify the true data generating process.

Hence the graphical tool so obtained enables investigation of the goodness-of-fit by means of the relative distances between the empirical TCF and all TCF’s of the copulas of the test space at once. By including many other copula families with different characteristics, it is possible to have a 2D visual overview of the whole collection of copulas based on their tail features as expressed by the tail concentration functions. This can be used as a copula selection tool in practical fitting problems, when one wants to choose one or more copulas to model the dependence structure in the data, highlighting the information contained in the data.

4.2. Visualizing clusters of financial time series via tail dependence

An interesting way to visualize dependence between time series observations is to use clustering methods, which allow to group together time series that have similar patterns (for a review, see [70]). Such clustering procedures are of particular interest for financial time series, where the identification of different groups in a portfolio of asset returns can be useful for portfolio diversification (see, for instance, [3, 11, 36, 103]). In such a context, several methods have focused on Pearson correlation to describe similarities in time series [5, 74], however recent investigations have underlined the benefits of other (e.g. copula-based) measures of association. In particular, methods of this type have focused on the
Figure 7: Two-dimensional representation of copula test spaces with data generated by different copulas, as indicated in text.
notion of tail dependence coefficient [18, 32], conditional Spearman’s correlation [31], and ad-hoc measures related to the threshold copulas associated to a given copula [25]. Starting with these ideas, we propose here another clustering procedures for financial time series that is based on the TCF of the empirical copula associated with a pair of financial time series. This procedure goes as follows.

Let \((x_{it})_{t=1,\ldots,T}\) be a matrix of \(d\) financial time series \((i = 1, 2, \ldots, d)\) representing the returns of different assets and/or stock indices. Following [18, 31], we need to proceed in different steps.

- First, we fit a suitable copula-based time series model to the financial time series in order to focus the attention to the rank-invariant dependence between the variables of interest without considering marginal effects. Specifically, we assume that each univariate time series follows a specific ARMA-GARCH process with \(t\)-distributed innovations. The dependence among the time series is hence fully expressed by the knowledge of the copula coupling the different estimated probability integral transform variables \((u_{it})_{t=1,\ldots,T}\) (concentrated in \(\mathbb{R}^d\)) obtained after these fittings (see eq.(3) in [87] for the related procedures to extract such series).

- For each \(i,j \in \{1, \ldots, d\}, i < j\), we determine the empirical TCF \(q_{emp}^{ij}\) associated with the pair of observations \((u_{it}, u_{jt})_{t=1,\ldots,T}\) by means of (7). Then we calculate the quantity

\[
d_{ij} = \frac{T}{2} \sum_{k=1}^{T/2} \left( q_{emp}^{ij}(k/T) - 1 \right)^2,
\]

where \(k/T\) are the points of \([0, 0.5]\) in which the empirical TCF is defined. The value \(d_{ij}\) is a kind of distance between the empirical TCF and the TCF of the comonotone copula \(M(u,v) = \min\{u,v\}\) in one part of their domain corresponding to the lower tail of the copula. It provides a measure of dissimilarity between the estimated probability integral transform variables since it is equal to 0 when the variables are coupled by a comonotone copula (expressing maximal association), while it increases otherwise.

- For each \(i,j \in \{1, \ldots, d\}, i < j\), we hence define the dissimilarity matrix \(D = (d_{ij})\) that represents the tail association among all the pairs of estimated probability integral transform variables extracted from the original vector of time series. Such a matrix could be hence used to determine clusters among the \(d\) time series of financial returns by using the hierarchical agglomerative clustering techniques. Specifically, among all the agglomerative strategies we may apply the three most common criteria that differ in the computation of the distance between two groups: single linkage, complete linkage, average linkage (see, for instance, [49, 50]).

Summarizing, the procedure determines a dissimilarity matrix from the lower tail of the pairwise (empirical) copulas associated with the original time series. Then, such a matrix is used to determine subsets among the financial time series via classical tools. It follows
that, since the entries of the dissimilarity matrix are calculated as distances between the maximal TCF and the realized TCF in the interval \([0, 0.5]\) (i.e. in one part of the diagonal section corresponding to the lower tail of the joint distribution), the financial time series that are grouped together tend to comove when both are experiencing losses.

To illustrate our methodology, we consider daily returns of time series of Morgan Stanley Capital International (MSCI) Developed Markets indices designed to measure the equity market performance of developed markets. The Dataset includes the following markets: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom and the United States. We restrict to the time series of daily log–returns \((x_1^t, \ldots, x_d^t)\), \(d = 23, t = 1, \ldots, T\), in the period from June 4, 2002 to June 10, 2010 \((T=2093\ \text{observations}; \ \text{Source: Datastream})\). The same dataset was considered in [18, 32].

We preliminary fit ARMA-GARCH models for each series of returns with Student-t distributed errors to account for heavy tails. For all time series we then perform classical tests to check for the adequacy of the fit (for more details, see [18, 32]).

Given the estimated probability integral transform variables for each time series, we then compute the dissimilarity matrix, which is visualized in Figure 8.

As said, this matrix can be the input of several clustering algorithms. Here, for the sake of illustration, we restrict to hierarchical clustering techniques. Specifically, the complete linkage method is used to achieve more useful hierarchies than single or average linkage from a pragmatic point of view. Looking at the dendrogram produced by complete linkage scheme (see Figure 9) we find out that \(k = 4\) can be considered a good solution, where \(k\) denotes the number of clusters selected. Notice that the results can be mostly interpreted in terms of geographic proximity: the lower tail dependence tends to be higher within European markets, where the Scandinavian countries are grouped together; we can distinguish one North-American cluster while Pacific countries are divided in two separate clusters.

As explained in [18] (see also [26]) such an output may be used for automatic portfolio selection procedures in order to hedge the risk of a portfolio, especially during crisis periods. In fact, one may suppose to construct a portfolio with the same number of assets of the number of obtained clusters, i.e. by selecting one asset in each cluster in order to reduce the risk of joint losses (remember that financial time series in each group tends to comove in losses). Notice that the portfolio selection under cardinality constraints has been an important issue in portfolio management (see, for instance, [6] and references therein).

5. Conclusions

In this contribution we have presented some known and novel facts about copulas with a prescribed diagonal section, emphasizing some tools (especially, graphical) that may serve to enhance their possible application. Starting from the consideration that a single numerical coefficient cannot describe conveniently the tail dependence of a copula, we define the so-called tail concentration function (which depends on the diagonal section of a
Figure 8: Dissimilarity matrix of the MSCI World Index Data constituents according to the described method.

copula) in order to give a finer description of the tails. Hence the tail concentration function is used as a graphical tool to distinguish different families of copulas in a 2D configuration, in the same spirit of [81]. Moreover, it serves as a basis to determine the grouping structure of different financial time series in a portfolio, giving an alternative to related methods presented in [18, 31]. This work is also intended to serve as a warning against drawing simplistic conclusions from partial information about the dependence structure. In fact when looking at tail dependence, both the asymptotic and the finite behaviour should be taken into account for practical purposes.

Caveat emptor! In general, the identification of the dependence structure is a tricky task that cannot be easily solved. The methods here presented should be intended as tools that the decision maker may use, together with other formal methods (especially from extreme-value analysis) to understand and communicate the uncertainty about the tail behaviour.
Figure 9: Dendrogram of the MSCI World Index Data constituents according to complete linkage.

Acknowledgements

The first and third authors acknowledge the support of the Free University of Bozen- Bolzano, School of Economics and Management, via the project MODEX.

The first and second author have been supported by the Ministero de Ciencia e Innovación (Spain), under Research Project MTM2011-22394.

References


