Copula function's concentration set and its concentrated partition

Lujun Li, Yijun Wu, and Jingping Yang[*,†]

The research on the local correlation structure of copula function is an attractive topic. This paper investigates bivariate copula function's local correlation structure by defining its concentration set. The concentration set of a copula function is defined in $[0,1]^2$ with restrained Lebesgue measure such that the samples of the copula fall in the set with the largest probability. The method for finding the concentration set is provided and the properties of the concentration set are discussed. Based on the concentration set, a concentrated partition of $[0,1]^2$ for the copula function is introduced, and one measure for quantifying copula function's local correlation is defined by applying our concentrated partition. An empirical study is provided to support our idea of proposing the concentration set.

AMS 2000 subject classifications: Primary 60A10, 62H20; secondary 62P05.
Keywords and phrases: Copula function, Local correlation structure, Concentration set, Concentration measure.

1. INTRODUCTION

Copula function is a multi-dimensional distribution function with uniform $[0,1]$ margins. For a d-dimensional distribution function $H(x_1,\ldots,x_d)$ with marginal distributions F_i, $i = 1,\ldots,d$, Sklar's Theorem (Joe, 1997; Nelsen, 2006) states that there exists a copula function C such that

$$H(x_1,\ldots,x_d) = C(F_1(x_1),\ldots,F_d(x_d))$$

for $x_1,\ldots,x_d \in (-\infty,\infty)$, and if the marginal distributions F_i, $i = 1,\ldots,d$ are continuous, the copula function C is unique. Sklar's Theorem shows that the dependence structure in the distribution H can be fully captured by the copula function C. For detailed introduction about copula function, we refer to McNeil, Frey and Embrechts (2005) and Nelsen (2006). Now copula functions have been widely applied in finance and insurance, e.g., see Cherubini, Luciano and Vecchiato (2004).

*Corresponding author.
†Partially supported by the Key Program of National Natural Science Foundation of China (Grants No. 11131002) and the National Natural Science Foundation of China (Grants No. 11271033).

Due to the complexity of some copula functions, the approximation of copula function by some specific copula families with desired properties is an interesting topic. The approximation methods include *shuffle of min approximation* (Mikusiński, Sherwood and Taylor, 1992; Durante, Sarkoci and Sempi, 2009), *checkmin approximation* (Mikusiński and Taylor, 2010), *checkerboard approximation* (Li, Mikusiński and Taylor, 1998; Durrleman, Nikeghbali, and Roncalli, 2000), *Bernstein approximation* (Scancetta and Satchell, 2004) and *patched bivariate Fréchet approximation* (Zheng, Yang and Huang, 2011). In the above references, partition methods are applied to divide the probability space into some subspaces and then approximate the conditional copula on each subspace. More precisely, let $C(u,v)$ be a copula function and (U_1,U_2) be its sample. Given a positive integer $m \geq 2$, the first step of these approximation methods is to divide the unit square $[0,1]^2$ into subsets $\{I_{i,j} = (\frac{i}{m},\frac{i+1}{m}] \times (\frac{j}{m},\frac{j+1}{m}], \ 0 \leq i,j \leq m-1\}$, then the conditional distributions on the division $A_{i,j} = \{(U_1,U_2) \in I_{i,j}\}, 0 \leq i,j \leq m-1$ of the probability space are considered. However, this partition process does not consider the probability differences among the subsets. Note that for one copula function, its sample has a larger probability to fall in some subset of $[0,1]^2$ than other subsets, thus it will be meaningful to discuss how to divide $[0,1]^2$ into some subsets through considering the corresponding probabilities.

Inspired by the partition methods of copula approximation, in this paper we will discuss the local correlation structure of a copula function by defining its concentration set. More precisely, for any copula function C, it is known that there exists a measure μ_C on $[0,1]^2$ such that

$$\mu_C(A) = \iint_A dC(u_1,u_2)$$

for any measurable set $A \subseteq [0,1]^2$. The quantity $\mu_C(A)$ measures the probability that (U_1,U_2) falls in the set A, where (U_1,U_2) is a sample of copula function C. Given level $a \in [0,1]$, we consider two sets $A,B \in \{D \subseteq [0,1]^2 | \mu(D) \leq a\}$ with $\mu_C(A) < \mu_C(B)$, where μ is the Lebesgue measure. Thus the probability that (U_1,U_2) falls in the set B is larger than that of the set A. Taking the above ideas together, we define the concentration set $B^*(a)$ of level $a \in [0,1]$ by solving the following optimization problem:

$$(1.1) \qquad B^*(a) = \arg \max_{B \subseteq [0,1]^2,\, \mu(B) \leq a} \mu_C(B).$$

In other words, the concentration set $B^*(a)$ is a set in the family $\{B \subseteq [0,1]^2 \,|\, \mu(B) \le a\}$ with the largest probability $\max\{P((U_1, U_2) \in B) \,|\, B \subseteq [0,1]^2, \mu(B) \le a\}$. Actually, the concentration set is meaningful when we estimate or approximate a copula function. From the estimation viewpoint, by the definition of the concentration set we can expect that there are more samples in the concentration set, then the estimation for the subset would be relatively better due to a relatively larger sample size. From the approximation viewpoint, when we approximate the copula function by some specific copula functions, the approximation result would be more affected by the approximation error on the concentration set, due to its large probability. Therefore, the concentration set can help us get more insights about the local correlation structure of a copula function.

In order to show the importance of the concentration set, we will consider the conditional rank correlation measures on the concentration set, and the numerical results will provide detailed information about the local correlation structure of the copula function. On the other hand, we will propose a new partition method for copula functions, named concentrated partition, which divides $[0,1]^2$ into some concentration sets with different levels. The advantage of this partition method is that the probabilities of the subsets are ordered purposely. Based on our concentrated partition, we will define a measure named concentration measure for quantifying the concentration degree of a copula function. The concentration measure is an infinite-dimensional vector with zero and one as the minimum and maximum of each component. If the support of the copula has zero Lebesgue measure, the components in the concentration measure are all equal to one. And for the independent copula, the components in the concentration measure are all equal to zero. Thus the concentration measure can efficiently describe the concentration degree of copula functions.

This paper is organized as follows. In Section 2 we will discuss how to determine the concentration set. In Section 3 we will introduce a partition method by applying concentration sets, and a concentration measure based on the partition method will be defined. In Section 4, an empirical analysis on China Government Bond will be given to support our methodology. Conclusions are given in Section 5.

2. CONCENTRATION SET AND ITS PROPERTIES

To begin with the optimization problem (1.1), we notice that for any set $B \subseteq [0,1]^2$ satisfying $\mu(B) \le a$, it holds that

$$\mu_C(B \cup B') \ge \mu_C(B)$$

for any measurable $B' \subseteq [0,1]^2 \backslash B$ with $\mu(B') = a - \mu(B)$. Then $\mu(B \cup B') = a$ and $\mu_C(B \cup B') \ge \mu_C(B)$. It implies

that we can find one solution to (1.1) in the family $\{B \subseteq [0,1]^2 \,|\, \mu(B) = a\}$. Thus in the rest of this article, we will discuss the following optimization problem

$$(2.1) \qquad B^*(a) = \arg \max_{B \subseteq [0,1]^2, \, \mu(B)=a} \mu_C(B).$$

The solution to the optimization problem (2.1) may not be unique. For example, let us consider the independent copula $C(u_1, u_2) = u_1 u_2$. For any $B \subseteq [0,1]^2$, we have $\mu_C(B) = \mu(B)$. Thus for the independent copula, the solution to (2.1) is any subset $B \subseteq [0,1]^2$ with $\mu(B) = a$.

Since the solution may not be unique, we will focus on the solution family $\{B^*(a), a \in [0,1]\}$ satisfying $B^*(a) \subseteq B^*(b)$ for any $0 \le a \le b \le 1$, and we call this kind of solution family an *enlargement family*.

In the following discussion, we start to solve problem (2.1) by assuming that the density of the copula function exists, then we extend the results to the general case.

2.1 Concentration set when copula's density exists

Suppose that the density function of the copula $C(u_1, u_2)$ exists, denoted as $c(u_1, u_2)$. The objective function in the optimization problem (2.1) can be expressed as

$$B^*(a) = \arg \max_{B \subseteq [0,1]^2, \, \mu(B)=a} \iint_B c(u_1, u_2) du_1 du_2.$$

We begin to solve the optimization problem (2.1) via finding the concentration set defined by the copula density. Let

$$B(s) = \left\{ (u_1, u_2) \in [0,1]^2 \,|\, c(u_1, u_2) > s \right\},$$
$$D(s) = \left\{ (u_1, u_2) \in [0,1]^2 \,|\, c(u_1, u_2) = s \right\}.$$

Given $a \in [0,1]$, we define $s^*(a) = \sup\{s \,|\, \mu(B(s)) > a\}$. For simplicity, sometimes we write $s^*(a)$ as s^*. It is easy to verify that

$$\mu(B(s^*(a))) + \mu(D(s^*(a))) \ge a \ge \mu(B(s^*(a))).$$

Based on the above notations, we have the following results.

Theorem 2.1. *Let $A_{s^*(a)}(a) \subseteq D(s^*(a))$ be a set satisfying $\mu(A_{s^*(a)}(a)) = a - \mu(B(s^*(a)))$. Then the set $B^*(a)$ defined by*

$$B^*(a) = B(s^*(a)) \cup A_{s^*(a)}(a)$$

is one solution to the optimization problem (2.1).

Proof. Firstly we notice that for any $(u_1, u_2) \in B^*(a)$, $c(u_1, u_2) \ge s^*(a)$; and for any $(u_1, u_2) \in [0,1]^2 \backslash B^*(a)$, $c(u_1, u_2) \le s^*(a)$. Then for any $A \subseteq [0,1]^2$ with $\mu(A) = a$,

it holds that

$$\mu_C(B^*(a)) - \mu_C(A)$$
$$= \iint_{B^*(a)} c(u_1, u_2) du_1 du_2 - \iint_A c(u_1, u_2) du_1 du_2$$
$$= \iint_{B^*(a)\setminus A} c(u_1, u_2) du_1 du_2 - \iint_{A\setminus B^*(a)} c(u_1, u_2) du_1 du_2$$
$$\geq s^*(a) \cdot (\mu(B^*(a)\setminus A) - \mu(A\setminus B^*(a)))$$
$$= s^*(a) \cdot (\mu(B^*(a)) - \mu(A)) = 0.$$

Thus $B^*(a)$ is one solution to (2.1), and we have $\mu(B^*(a)) = \mu(A_{s^*(a)}(a)) + \mu(B(s^*(a))) = a$, so the proof is complete. $\square$

From the above theorem we can see that if $\mu(B(s^*(a))) = a$, then the set $B^*(a) = B(s^*(a))$ solves (2.1). On the other hand, if $\mu(B(s^*(a))) < a$, we know that $\mu(D(s^*(a))) \geq a - \mu(B(s^*(a)))$ in this case, so we can choose an adjustment subset $A_{s^*(a)}(a) \subseteq D(s^*(a))$ to guarantee that the Lebesgue measure of the union $B(s^*(a)) \cup A_{s^*(a)}(a)$ equals a. Actually, the adjustment subset is not unique. For consistency, we need to establish rules for choosing the adjustment set.

As we mentioned before, we hope that the solution family $\{B^*(a), a \in [0, 1]\}$ is an enlargement family. For this purpose, the adjustment sets $A_{s^*(a)}(a)$ can be chosen by some specific rules. Some rules are given in the following remark.

Remark 2.1. The adjustment sets $A_{s^*(a)}(a), a \in [0, 1]$ can be chosen according to one of the following rules:

(1) $A_{s^*(a)}(a) \subseteq D(s^*(a)) \cap \{(u_1, u_2) \in [0,1]^2 | u_1 \leq x, x \in [0,1]\}$;
(2) $A_{s^*(a)}(a) \subseteq D(s^*(a)) \cap \{(u_1, u_2) \in [0,1]^2 | u_2 \leq x, x \in [0,1]\}$;
(3) $A_{s^*(a)}(a) \subseteq D(s^*(a)) \cap \{(u_1, u_2) \in [0,1]^2 | u_1 + u_2 \leq 2x, x \in [0,1]\}$.

It is easy to verify that each of the rules guarantees that the concentration set family $\{B^*(a), a \in [0, 1]\}$ defined in Theorem 2.1 is an enlargement family.

In the next we give an example for using the above rules.

Example 2.1. For the independent copula $C(u_1, u_2) = u_1 u_2$, when we use the adjustment rule (1) in Remark 2.1, the solution to problem (2.1) is $B^*(a) = [0, a] \times [0, 1]$. For rule (3), the solution is $B^*(a) = \{(u_1, u_2) \in [0,1]^2 | u_1 + u_2 \leq \sqrt{2a}\}$ for $0 \leq a < 1/2$ and $B^*(a) = \{(u_1, u_2) \in [0,1]^2 | u_1 + u_2 \leq 2 - \sqrt{2 - 2a}\}$ for $1/2 \leq a \leq 1$.

Remark 2.2. Let $B^*(a)$ be the concentration set defined in Theorem 2.1 with some level $a \in [0, 1]$. Denote $b = \mu_C(B^*(a))$. If the density function $c(u_1, u_2)$ satisfies $c(u_1, u_2) > 0$ for all $(u_1, u_2) \in [0,1]^2$, then $B^*(a)$ is a solution of the dual problem

$$N^*(b) = \arg \min_{B \subseteq [0,1]^2, \, \mu_C(B)=b} \mu(B).$$

Because if there is a subset $N \subseteq [0, 1]^2$ satisfying $\mu_C(N) = b$ and $\mu(N) < a$, then for any $A \subseteq [0,1]^2 \setminus N$ with $\mu(A) = a -$ $\mu(N)$, the subset $B = N \cup A$ satisfies $\mu(B) = a$ and $\mu_C(B) > \mu_C(N) = b$, which is contradictory to the definition of $B^*(a)$ for solving the optimization problem (2.1).

2.2 Concentration set in the general case

In this subsection we will solve the optimization problem (2.1) for the general case.

By Lebesgue's decomposition theorem, the measure μ_C can be decomposed uniquely as

$$(2.2) \qquad \mu_C = \alpha \nu_c + (1 - \alpha) \nu_d,$$

where $\alpha \in [0, 1]$, ν_c and ν_d are probability measures on $[0, 1]^2$ satisfying that ν_c is absolutely continuous to the Lebesgue measure μ, and ν_d contains the discrete part and the singular part of μ_C. Referring to Halmos (1970, p. 134), there exists some $B_0 \subseteq [0, 1]^2$ satisfying

$$(2.3) \qquad \mu(B_0) = 0 \quad \text{and} \quad \nu_d(B_0) = 1.$$

Notice that B_0 may not be unique, but the difference between different versions of B_0 must have zero Lebesgue measure, and the choice of B_0 has no effect on the following results.

In fact, ν_c and ν_d are probability measures on $[0, 1]^2$, not necessarily generated by copula functions. Specifically, if $\alpha = 1$, the probability measure $\nu_c = \mu_C$ and it turns back to the case of Theorem 2.1. In the case $\alpha \in [0, 1)$, we can define $\tilde{B}^*(a)$ according to the methodology in Theorem 2.1 with respect to the density function $\tilde{c}(u_1, u_2) = d\nu_c/d\mu$. Although $\tilde{c}(u_1, u_2)$ is not necessarily a copula density, it is easy to check that the result of Theorem 2.1 also holds for general bivariate density functions, i.e., $\tilde{B}^*(a)$ is one solution to the following optimization problem

$$\tilde{B}^*(a) = \arg \max_{B \subseteq [0,1]^2, \, \mu(B) \leq a} \iint_B \tilde{c}(u_1, u_2) du_1 du_2.$$

Combining with the definition of B_0 in (2.3), we know that

$$(2.4) \qquad B^*(a) = \tilde{B}^*(a) \cup B_0$$

is one solution to the optimization problem (2.1). Moreover, the family $\{B^*(a), a \in [0, 1]\}$ defined above is an enlargement family as long as $\{\tilde{B}^*(a), a \in [0, 1]\}$ is an enlargement family.

2.3 The correlation measures in the concentration set

In this section, we consider conditional rank correlation measures in the concentration set.

For measuring the correlation between random variables, Kendall's tau and Spearman's rho (Kruskal, 1958) are the most commonly used rank correlation measures. Considering a random vector (X, Y) and its independent copy

(X_1, Y_1), Kendall's tau is defined as

$$\tau(X, Y) = P((X - X_1)(Y - Y_1) > 0)$$
$$- P((X - X_1)(Y - Y_1) < 0).$$

On the other hand, Spearman's rho is the linear correlation of the probability-transformed random variables, i.e.,

$$\rho(X, Y) = \text{corr}(F_X(X), F_Y(Y)).$$

Actually, let C be the copula function of continuous random vector (X, Y), then we have

$$\tau_C = \tau(X, Y) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1,$$

$$\rho_C = \rho(X, Y) = 12 \int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2) \, du_1 du_2.$$

As for the local and regional dependence measure, Holland and Wang (1987a,b) considered the dependence on the region of nonzero density function; Bjerve and Doksum (1993) defined a local correlation measure based on conditional mean and variance; Drouet-Mari and Kotz (2001) restricted Kendall's tau and Spearman's rho to an open neighborhood of a certain point in $[0, 1]^2$; Kolev, Anjos and Mendes (2006) generalized Spearman's rho to a conditional version. In this section, we will consider the conditional Kendall's tau and Spearman's rho in the concentration set.

For a copula function C, its concentration set $B^*(a)$ with level $a \in (0, 1]$ has been defined above. Next we denote $\tau_C^*(a)$ and $\rho_C^*(a)$ as the conditional Kendall's tau and Spearman's rho of U_1, U_2 on $\{(U_1, U_2) \in B^*(a)\}$ respectively, where (U_1, U_2) is a sample of copula C. In other words, let the conditional copula of random vector (U_1, U_2) under $\{(U_1, U_2) \in B^*(a)\}$ be denoted as $C_{B^*(a)}$, therefore we have

$$\tau_C^*(a) = 4 \int_0^1 \int_0^1 C_{B^*(a)}(u_1, u_2) dC_{B^*(a)}(u_1, u_2) - 1,$$

$$\rho_C^*(a) = 12 \int_0^1 \int_0^1 (C_{B^*(a)}(u_1, u_2) - u_1 u_2) \, du_1 du_2.$$

From this definition, we know that the conditional rank correlations on the concentration set equal the traditional global rank correlations when $a = 1$, i.e., $\tau_C^*(1) = \tau_C$ and $\rho_C^*(1) = \rho_C$. As a generalization of the traditional rank correlations, $\tau_C^*(a), \rho_C^*(a), a \in (0, 1]$ measure the local correlation of a copula function through functional viewpoint.

As an example, we consider Gaussian copula, student T-copula, Clayton copula and Gumbel copula. The conditional rank correlations on the concentration sets are shown in Figure 1. All of the four copulas have the same Kendall's tau, chosen as 0.3 in the example. Figure 1(a) shows that $\tau_C^*(a)$ of each copula is larger than 0.3 when $a < 1$, which means the
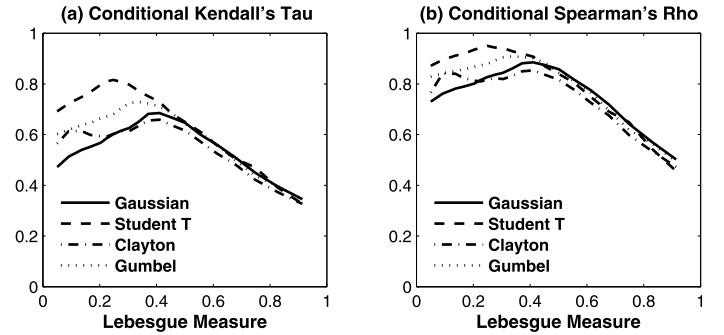

Figure 1. The curves $\tau_C^*(a)$ and $\rho_C^*(a), a \in (0, 1]$ of four different copulas. Kendall's tau of each copula equals 0.3. The correlation parameter of the Gaussian copula is 0.4540. The correlation parameter and the freedom of the student T-copula are 0.4540 and 3. The parameters of the Clayton and the Gumbel copula are 0.8571 and 1.4286 respectively.

concentration set is more likely to be around the diagonal line $u_1 = u_2$ in these copulas. The curve of $\tau_C^*(a)$ displays bimodal in the case of Clayton copula, while in other cases the curves are unimodal. An interesting fact is that the conditional Kendall's tau performs no obvious difference among the four copulas when $a > 0.4$, but the performance is quite different when $a < 0.4$, which means the four copulas have very different features in the high density area.

By checking the performance of the conditional rank correlation, we can obtain detailed information of the copula function itself. For example, for a small a the curve of conditional Kendall's tau of the student T-copula is above that of the Gaussian copula, which implies that the student T-copula has a larger conditional correlation in high density area.

Figure 1(b) shows the conditional Spearman's rho measure on the concentration set. The performance of $\rho_C^*(a)$ is similar to conditional Kendall's tau except for being larger. One fact worthy of mention is that the conditional Spearman's rho of student T-copula is close to 1 when $a \approx 0.22$, which means that the concentration set of level 0.22 is almost in the diagonal area.

3. CONCENTRATED PARTITION AND CONCENTRATION MEASURE

As mentioned in the introduction section, the regular partition method is applied in the approximation of copula functions. The partition process did not distinguish the probability of each subset. In this section, we will give one partition method by applying concentration sets with different levels, named concentrated partition, in which the probability of every subset is considered during the partition process. Following the concentrated partition, a new measure named concentration measure is proposed for measuring the concentration degree of copula functions.

3.1 Concentrated partition

For the copula function C and fixed $m \geq 2$, we will use concentration sets of copula C to divide the space $[0,1]^2$ into some subsets.

Given $\mathbf{a} = (a_1, \ldots, a_m)$ with $a_i > 0$ and $\sum_{i=1}^m a_i = 1$, for the copula function C we firstly find its concentration set $B_1^*(\mathbf{a}) \subseteq [0,1]^2$ with level a_1. Then we consider the subspace $[0,1]^2/B_1^*(\mathbf{a})$ and find another concentration subset $B_2^*(\mathbf{a})$ in the space $[0,1]^2/B_1^*(\mathbf{a})$ such that $\mu(B_2^*(\mathbf{a})) = a_2$. Sequently, we can obtain the sequence $B_n^*(\mathbf{a})$, $n = 3, \ldots, m$.

The above methodology can be summarized as the following optimization problem: finding disjoint subsets $\{B_1^*(\mathbf{a}), \ldots, B_m^*(\mathbf{a})\}$ of $[0,1]^2$ such that

$$
\begin{aligned}
B_i^*(\mathbf{a}) &= \arg\max \mu_C(B) \\
&\text{s.t. } B \subseteq [0,1]^2 \setminus \cup_{j=1}^{i-1} B_j^*(\mathbf{a}), \ \mu(B) = a_i,
\end{aligned}
$$

(3.1)

where we use the notation $\cup_{j=1}^0 B_j^*(\mathbf{a}) = \emptyset$.

Notice that $B_m^*(\mathbf{a}) = [0,1]^2 \setminus \cup_{j=1}^{m-1} B_j^*(\mathbf{a})$. The above partition divides the space $[0,1]^2$ into m disjoint subsets $B_1^*(\mathbf{a}), B_2^*(\mathbf{a}), \ldots, B_m^*(\mathbf{a})$ by using the concentration sets. Thus we call $B_1^*(\mathbf{a}), B_2^*(\mathbf{a}), \ldots, B_m^*(\mathbf{a})$ the copula C's concentrated m-partition of $[0,1]^2$ with level $\mathbf{a} = (a_1, \ldots, a_m)$. And $A_i = \{(U_1, U_2) \in B_i^*(\mathbf{a})\}, i = 1, \ldots, m$ is called the copula C's concentrated m-partition of probability space with level $\mathbf{a} = (a_1, \ldots, a_m)$. In the special case $a_i = 1/m, i = 1, \ldots, m$, we have $\mu(B_1^*(\mathbf{a})) = \mu(B_2^*(\mathbf{a})) = \cdots = \mu(B_m^*(\mathbf{a})) = 1/m$.

Next we will use the essential infimum and supremum of functions in our proof. The essential infimum of function f is defined as

$$
\operatorname{essinf} f = \sup\{b \mid \mu\{x \mid f(x) \leq b\} = 0\},
$$

and the essential supremum of function f is defined as

$$
\operatorname{esssup} f = \inf\{b \mid \mu\{x \mid f(x) \geq b\} = 0\}.
$$

Theorem 3.1. *Assume that the family $\{B^*(a), a \in [0,1]\}$ is an enlargement family solving the optimization problem (2.1). Then*

$$
\begin{aligned}
B_1^*(\mathbf{a}) &= B^*(a_1), \\
B_i^*(\mathbf{a}) &= B^*(a_1 + \cdots + a_i) \setminus B^*(a_1 + \cdots + a_{i-1}), \ 2 \leq i \leq m
\end{aligned}
$$

is one solution to the optimization problem (3.1). Moreover, let $A_i = \{(U_1, U_2) \in B_i^*(\mathbf{a})\}, i = 1, \ldots, m$, then

$$
P(A_1)/a_1 \geq P(A_2)/a_2 \geq \cdots \geq P(A_m)/a_m.
$$

Proof. Since the concentration sets $\{B^*(a), a \in [0,1]\}$ is an enlargement family, so the subsets $B_i^*(\mathbf{a})$, $i = 2, \ldots, m$ are well defined. By checking the definition of concentration set $B^*(a_1 + \cdots + a_i)$, $i = 1, \ldots, m$, it is obvious that $B_i^*(\mathbf{a})$, $i = 1, \ldots, m$ defined above solve the optimization problem (3.1).

To prove $P(A_i)/a_i$ is a decreasing series, we firstly consider the case that the copula density $c(u_1, u_2)$ exists. For $i = 1, \ldots, m-1$,

$$
\begin{aligned}
P(A_i)/a_i &\geq \operatorname{essinf}_{(u_1,u_2) \in B_i^*(\mathbf{a})} c(u_1, u_2) \\
&\geq \operatorname{esssup}_{(u_1,u_2) \in B_{i+1}^*(\mathbf{a})} c(u_1, u_2) \\
&\geq P(A_{i+1})/a_{i+1}.
\end{aligned}
$$

For the general case, using the similar argument and (2.4) we can obtain the conclusion. $\square$

Remark 3.1. Theorem 3.1 shows that the unit square $[0,1]^2$ can be divided into m subsets by using concentration sets. Note that $A_i = \{(U_1, U_2) \in B_i^*(\mathbf{a})\}, i = 1, \ldots, m$ are an ordered partition of the probability space in the case $a_1 = \cdots = a_m = 1/m$, in the sense that

$$
(3.2) \qquad P(A_1) \geq P(A_2) \geq \cdots \geq P(A_m)
$$

Hence the above ordered partition leads to an ordered copula decomposition as following

$$
(3.3) \quad
\begin{aligned}
C(u_1, u_2) &= \sum_{i=1}^n P(A_i) \cdot P(U_1 \leq u_1, U_2 \leq u_2 | A_i) \\
&= \sum_{i=1}^n P(A_i) \cdot C_i(F_i(u_1), G_i(u_2)),
\end{aligned}
$$

where for each i the function C_i is the conditional copula function on A_i, and F_i, G_i are the conditional marginal distributions on A_i.

The next corollary expresses the concentrated partition by using copula density functions.

Corollary 3.1. *(a) Suppose that the density c of the copula function C exists, and $\mu(D(s)) = 0$ for all $s \geq 0$. Then there exists $s_1 > s_2 > \cdots > s_{m-1} > 0$ such that*

$$
\begin{aligned}
B_1^*(\mathbf{a}) &= \{(u_1, u_2) \in [0,1]^2 | c(u_1, u_2) > s_1\}, \\
B_i^*(\mathbf{a}) &= \{(u_1, u_2) \in [0,1]^2 | c(u_1, u_2) \in (s_i, s_{i-1}]\}, 1 < i < m \\
B_m^*(\mathbf{a}) &= \{(u_1, u_2) \in [0,1]^2 | c(u_1, u_2) \leq s_{m-1}\}
\end{aligned}
$$

is copula C's concentrated m-partition of $[0,1]^2$ with level $\mathbf{a} = (a_1, \ldots, a_m)$.

(b) Suppose that for the copula function C the decomposition (2.2) holds. Let $\tilde{B}_1^*(\mathbf{a}), \ldots, \tilde{B}_m^*(\mathbf{a})$ be the corresponding optimization solution in Theorem 3.1 with respect to the density function $\tilde{c}(u_1, u_2) = d\nu_c/d\mu$, then

$$
\begin{aligned}
B_1^*(\mathbf{a}) &= \tilde{B}_1^*(\mathbf{a}) \cup B_0, \\
B_i^*(\mathbf{a}) &= \tilde{B}_i^*(\mathbf{a}) \setminus B_0, \ i = 2, \ldots, m
\end{aligned}
$$

is copula C's concentrated m-partition of $[0,1]^2$ with level $\mathbf{a} = (a_1, \ldots, a_m)$, where B_0 is defined in (2.3).

Proof. We only give the proof of the first part. The proof of the second part is simple and omitted.

If $\mu(D(s)) = 0$ for any s, $\mu(B(s))$ is a continuous decreasing function of s with $\mu(B(0)) = 1$ and $\mu(B(\infty)) = 0$.

Therefore for any $a \in [0, 1]$, we have $\mu(B(s^*(a))) = a$, in which $s^*(a) = \sup\{s \mid \mu(B(s)) > a\}$. Combining with the results in Theorem 3.1, we obtain the above expressions of $B_i^*(\mathbf{a})$. $\square$

3.2 Examples

In this subsection, we discuss our concentrated partition by focusing on Clayton copula, Gaussian copula and Fréchet copula.

Example 3.1 (Clayton Copula). *Archimedean copula* is defined as

$$C(u_1, u_2) = \phi^{-1}\left(\phi(u_1) + \phi(u_2)\right), \; u_1, u_2 \in [0, 1]$$

where the generator ϕ is a decreasing function from $[0, 1]$ to $[0, \infty]$, satisfying $\phi(0) = \infty$ and $\phi(1) = 0$. Clayton copula is a member of Archimedean family. The generator of Clayton copula is $\phi(t) = (t^{-\theta} - 1)/\theta$. Figure 2(a) shows the concentrated 3-partition of Clayton copula with level $\mathbf{a} = (1/3, 1/3, 1/3)$, in which the parameter of the Clayton copula $\theta =0.8571$, and its Kendall's tau equals 0.3. In this example, we have $P(A_1) = 0.4958$, $P(A_2) = 0.3418$, $P(A_3) = 0.1624$.

Clayton copula has a positive lower tail dependence coefficient, i.e., $\lim_{u\downarrow 0} C(u, u)/u > 0$, which means that there is a comovement near $(0, 0)$. Hence we can see an obvious asymmetry between the lower left and upper right corner in Figure 2(a).

Beyond the Archimedean copula family, Gaussian copula is also an important parametric copula family. It has the advantage that the Gaussian copula can illustrate the dependence structure only using the correlation coefficient, which is easy to calibrate and apply in practice.

Example 3.2 (Gaussian Copula). Figure 2(b) shows the concentrated 3-partition of Gaussian copula with level $\mathbf{a} = (1/3, 1/3, 1/3)$, where the correlation parameter $\rho = 0.4540$. We notice that the first concentration set with level $1/3$ is near the points $(0, 0)$ and $(1, 1)$. It is because that Gaussian copula with positive correlation parameter has a large density near these two points. In this case, we have $P(A_1) = 0.4811$, $P(A_2) = 0.3402$, $P(A_3) = 0.1787$. Figure 3 shows the concentrated partitions of Gaussian copula and student T-copula. From the concentrated partitions, we can see that tail dependence of student T-copula is higher than the Gaussian copula. In the view of the information content, the area near the left lower tail is valuable for calculating risk measures such as *Value-at-Risk* and *expected shortfall* (McNeil, Frey and Embrechts, 2005).

In the next example, we consider a copula which is not absolutely continuous to the Lebesgue measure.
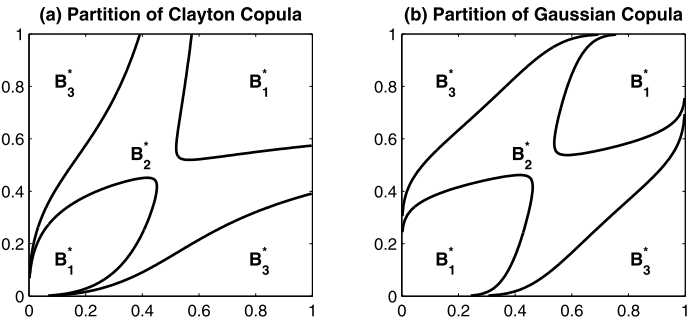
Figure 2. Concentrated partition of Clayton and Gaussian copula with level $\mathbf{a} = (1/3, 1/3, 1/3)$. The parameter of Clayton copula is 0.8571, and the correlation parameter of Gaussian copula is 0.4540. Kendall's tau of the two copulas both equal 0.3. According to Corollary 3.1, the edges of each part are contour lines of copula density functions.

Example 3.3 (Fréchet Copula). Fréchet copula has the following form

$$C(u_1, u_2) = \alpha \cdot \min(u_1, u_2) + \gamma \cdot \max(u_1 + u_2 - 1, 0) + (1 - \alpha - \gamma) \cdot u_1 u_2$$

for $(u_1, u_2) \in [0, 1]$, where $\alpha, \gamma \geq 0$ and $\alpha + \gamma \leq 1$. For the level $\mathbf{a} = (a_1, \ldots, a_m)$, $a_i > 0$ with $\sum_{i=1}^m a_i = 1$, we use the adjustment rule (1) in Remark 2.1. Let B_0 be the set defined in (2.3), so $B_0 = \{[0, 1]^2 \mid u_1 = u_2 \text{ or } u_1 + u_2 = 1\}$, then the concentrated partition of Fréchet copula is

$$B_1^*(\mathbf{a}) = ([0, a_1] \times [0, 1]) \cup B_0,$$
$$B_k^*(\mathbf{a}) = \left(\left[\sum_{i=1}^{k-1} a_i, \sum_{i=1}^k a_i\right] \times [0, 1]\right) \cap B_0^c, \; k = 2, \ldots, m.$$

Note that the set B_0 contains the support of the *comonotonic copula* $\min(u_1, u_2)$ and the *countermonotonic copula* $\max(u_1 + u_2 - 1, 0)$ in Fréchet copula. Then $P(A_1) = \alpha + \gamma + (1 - \alpha - \gamma)a_1$ and $P(A_k) = (1 - \alpha - \gamma)a_k$ for $k = 2, \ldots, m$. Recalling the decomposition in equation (3.3), we can compute the conditional margins as following. The marginal distributions can be expressed as

$$F_1(u) = P(U_1 \leq u | A_1)$$
$$= (\alpha u + \gamma u + (1 - \alpha - \gamma)\min(u, a_1))/P(A_1),$$
$$F_k(u) = P(U_1 \leq u | A_k)$$
$$= \min\left(\max\left(u - \sum_{i=1}^{k-1} a_i, 0\right)/a_k, 1\right)$$

for $k = 2, \ldots, m$. And $G_k(u) = P(U_2 \leq u | A_k) = u$ for all $k = 1, \ldots, m$. It is easy to verify that for $k \geq 2$, the conditional copula C_k on A_k is the independent copula $u_1 u_2$.
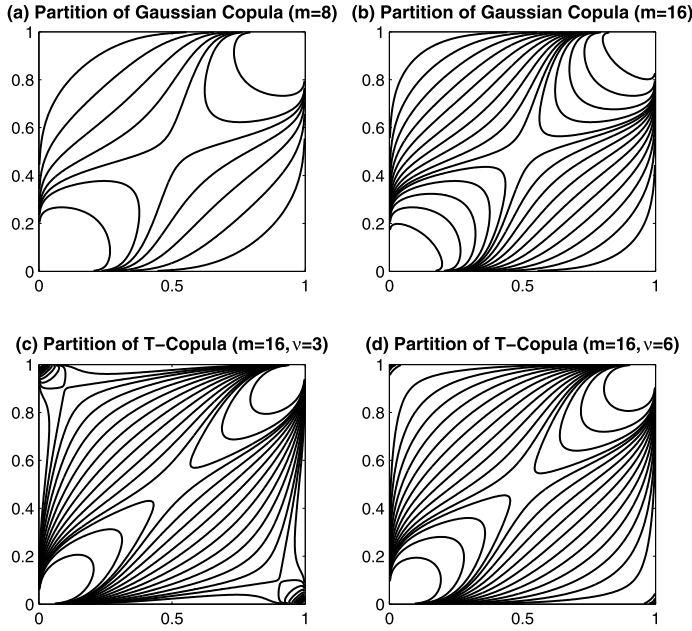
(a) Partition of Gaussian Copula (m=8) **(b) Partition of Gaussian Copula (m=16)**

(c) Partition of T−Copula (m=16, ν=3) **(d) Partition of T−Copula (m=16, ν=6)**

Figure 3. Concentrated partition of Gaussian copula and student T-copula with level $\mathbf{a} = (1/m, \ldots, 1/m)$. Kendall's tau of the copulas all equal 0.3. The freedom parameters of the T-copula in subplot (c) and (d) are 3 and 6 respectively.

For the conditional copula C_1, we have

$$
\begin{aligned}
C_1(u_1, u_2) = \big\{ &\gamma \max(F_1^{-1}(u_1) + u_2 - 1, 0) \\
&+ (1 - \alpha - \gamma) \min(F_1^{-1}(u_1), a_1) u_2 \\
&+ \alpha \min(F_1^{-1}(u_1), u_2) \big\} / P(A_1),
\end{aligned}
$$

where

$$
F_1^{-1}(x) = \begin{cases} P(A_1)x, & 0 \le x \le a_1/P(A_1); \\ a_1 + (1 - a_1)\frac{P(A_1)x - a_1}{P(A_1) - a_1}, & a_1/P(A_1) < x \le 1. \end{cases}
$$

Compared with the regular partition (e.g. Zheng, Yang and Huang, 2011), the concentrated partition is more efficient to describe the correlation structure. For instance, for the comonotonic copula $\min(u, v)$, the regular partition (e.g. Zheng, Yang and Huang, 2011) considers m^2 squares $\{I_i \times I_j, i, j = 1, \ldots, m\}$ with $I_i = [\frac{i-1}{m}, \frac{i}{m})$, and there are m subsets with positive probabilities. However, applying our concentrated partition method, only the first concentration set has positive probability, so all the correlation information is concentrated in the first partition subset.

3.3 Concentration measure of copula functions

Given positive integer $m > 1$, we set $\mathbf{a} = (1/m, \ldots, 1/m)$ in this subsection. Recall that for a copula function C, we can find its concentrated m-partition $\{B_i^*(\mathbf{a}), 1 \le i \le m\}$

in Theorem 3.1. Following the concentrated partition, we define

$$
\lambda_C(m) = \max_{2 \le i \le m} \big(\mu_C(B_{i-1}^*(\mathbf{a})) - \mu_C(B_i^*(\mathbf{a})) \big).
$$

From the definition of the concentration sets, we know that the solution $B_i^*(\mathbf{a})$ may not be unique, but $\lambda_C(m)$ is independent of this choice.

For any copula function C, we can calculate the sequence $\lambda_C(2), \lambda_C(3), \ldots$ Hence we call the vector

$$
\Gamma_C = (\lambda_C(2), \lambda_C(3), \ldots)
$$

the concentration measure of the copula function C. Note that this measure consists of infinite components, and for each $m \ge 2$,

$$
0 \le \lambda_C(m) \le 1.
$$

It is easy to check that $\Gamma_C = (1, 1, \ldots)$ when the Lebesgue measure of the support of C equals to zero, and $\Gamma_C = (0, 0, \ldots)$ when C is the independent copula. And if $\lambda_{C_1}(m) > \lambda_{C_2}(m)$ for each $m \ge 2$, then we can conclude that C_1 is more concentrated than C_2.

For some special copula functions, we will show the linearity of the measure Γ_C in the next theorem.

Theorem 3.2. *(I) If $\lambda_C(m) = 0$ for some $m \ge 2$, then C must be the independent copula $\Pi(u_1, u_2) = u_1 u_2, u_1, u_2 \in [0, 1]$ and $\Gamma_C = (0, 0, \ldots)$.*

(II) If copula C and C_1 satisfy $C = \gamma C_1 + (1 - \gamma)\Pi$ for some $\gamma \in [0, 1]$, we have $\Gamma_C = \gamma \Gamma_{C_1}$.

Proof. Given $m \ge 2$ and $\mathbf{a} = (1/m, \ldots, 1/m)$.

(I) Recall the decomposition in (2.2). Denote the concentrated partition with respect to C by $B_{m,i}^*$, $i = 1, \ldots, m$, and denote the concentrated partition with respect to the continuous part ν_c by $\tilde{B}_{m,i}^*$, $i = 1, \ldots, m$.

The fact $\lambda_C(m) = 0$ implies that for any $2 \le i \le m$,

$$
(3.4) \qquad \mu_C(B_{m,i-1}^*) = \mu_C(B_{m,i}^*).
$$

Next we will prove that the density of C exists almost surely. Then it will be proved that the density of C equals 1 almost surely.

Denote B_0 as the support of ν_d defined in (2.3). Corollary 3.1 says that

$$
B_{m,1}^* = \tilde{B}_{m,1}^* \cup B_0 \quad \text{and} \quad B_{m,i}^* = \tilde{B}_{m,i}^* \setminus B_0, \quad i = 2, \ldots, m.
$$

On the other hand, $\nu_c(\tilde{B}_{m,1}^*) \ge \nu_c(\tilde{B}_{m,2}^*)$ owing to the conclusion in (3.2). Combining with $\mu(B_0) = 0$, we have if $\alpha < 1$,

$$
\begin{aligned}
\mu_C(B_{m,1}^*) &= \mu_C(\tilde{B}_{m,1}^* \cup B_0) = \mu_C(B_0) + \mu_C(\tilde{B}_{m,1}^* \setminus B_0) \\
&= 1 - \alpha + \mu_C(\tilde{B}_{m,1}^* \setminus B_0) = 1 - \alpha + \alpha \cdot \nu_c(\tilde{B}_{m,1}^* \setminus B_0) \\
&= 1 - \alpha + \alpha \cdot \nu_c(\tilde{B}_{m,1}^*) > \alpha \cdot \nu_c(\tilde{B}_{m,1}^*) \ge \alpha \cdot \nu_c(\tilde{B}_{m,2}^*) \\
&= \alpha \cdot \nu_c(\tilde{B}_{m,2}^* \setminus B_0) = \mu_C(\tilde{B}_{m,2}^* \setminus B_0) = \mu_C(B_{m,2}^*),
\end{aligned}
$$

contradictory to (3.4). Therefore we conclude $\alpha = 1$, then the density function of copula C exists. Next we continue to prove that this density function $c(u_1, u_2)$ equals 1 almost surely. From Theorem 3.1, we know that for any $2 \leq i \leq m$,

$$\mu_C(B_{m,i-1}^*) \geq \text{essinf}_{(u_1,u_2) \in B_{m,i-1}^*} c(u_1, u_2)/m$$
$$\geq \text{esssup}_{(v_1,v_2) \in B_{m,i}^*} c(v_1, v_2)/m \geq \mu_C(B_{m,i}^*),$$

where the equalities hold if and only if

$$c(u_1, u_2) = \text{essinf}_{(w,z) \in B_{m,i-1}^*} c(w, z)$$
$$= \text{esssup}_{(w,z) \in B_{m,i}^*} c(w, z) = c(v_1, v_2)$$

almost surely for any $(u_1, u_2) \in B_{m,i-1}^*$ and any $(v_1, v_2) \in B_{m,i}^*$. It follows that $c(u_1, u_2) = 1$ almost surely in $[0,1]^2$. Therefore, the copula C is the independent copula, and $\Gamma_C = (0, 0, \dots)$.

(II) Denote the concentrated partition with respect to C_1 by $A_{m,i}^*$, $i = 1, \dots, m$. We will prove that $A_{m,i}^*$ is also a solution to (3.1) with respect to C.

Firstly we will prove that if $B^*(a)$ is a solution to (2.1) with respect to C_1 for $a \in [0,1]$, then it is also a solution to this problem with respect to C. In fact, for any $B \subseteq [0,1]^2$ with $\mu(B) = a$, we have $\mu_{C_1}(B^*(a)) \geq \mu_{C_1}(B)$ owing to the definition of $B^*(a)$, then

$$\mu_C(B^*(a)) = \gamma \mu_{C_1}(B^*(a)) + (1 - \gamma)\mu_\Pi(B^*(a))$$
$$\geq \gamma \mu_{C_1}(B) + (1 - \gamma)\mu_\Pi(B)$$
$$= \mu_C(B).$$

The above inequality implies that $B^*(a)$ is the solution to (2.1) with respect to C. According to Theorem 3.1, we know that $\{A_{m,i}^*, i = 1, \dots, m\}$ is also a solution to (3.1) with respect to C. Then we can get

$$\lambda_C(m) = \max_{2 \leq i \leq m} \left(\mu_C(A_{m,i-1}^*) - \mu_C(A_{m,i}^*) \right)$$
$$= \max_{2 \leq i \leq m} \left(\gamma \cdot \mu_{C_1}(A_{m,i-1}^*) - \gamma \cdot \mu_{C_1}(A_{m,i}^*) \right)$$
$$= \gamma \cdot \lambda_{C_1}(m).$$

Hence the proof is completed. □

Theorem 3.2 shows that the concentration measure Γ_C reaches its minimum if and only if C is an independent copula. Thus Γ_C can be regarded as a measure for the distance from the independent copula. When the concentration measure is large, the samples of the copula will concentrate in some small area.

Different from the traditional association measures, the concentration measure Γ_C is an infinite-dimensional vector. Through the components of the vector, we can obtain detailed information about the local correlation structure of copula functions.

Theorem 3.3. *Assume that the probability measure μ_C generated by copula C has the decomposition in (2.2).*

(I) We have $\lim_{m \to \infty} \lambda_C(m) = 1 - \alpha$. Furthermore, if $\lim_{m \to \infty} \lambda_C(m) = 0$, then C has density function.

(II) If $\Gamma_C = (\lambda, \lambda, \dots)$, then $\lambda = 1 - \alpha$ and ν_c in (2.2) is a probability measure generated by the independent copula.

Proof. Suppose that for $m \geq 2$, the concentrated partition for ν_c is denoted as $\tilde{B}_{m,i}^*, i = 1 \dots, m$, and the concentrated partition for μ_C is denoted as $B_{m,i}^*, i = 1 \dots, m$.

(I) Firstly we will prove that the limit of $\lambda_C(m)$ exists and equals $1 - \alpha$.

If $\alpha = 0$, we know that $\mu_C(B_{m,1}^*) = 1$. It follows that $\lambda_C(m) = 1$ for any m and we can get the conclusion. Now we assume that $\alpha \in (0, 1]$. Since ν_c is absolutely continuous to the Lebesgue measure, we have

$$0 \leq \lim_{m \to \infty} \max_{2 \leq i \leq m} \left(\nu_c(\tilde{B}_{m,i-1}^*) - \nu_c(\tilde{B}_{m,i}^*) \right)$$
$$\leq \lim_{m \to \infty} \max_{2 \leq i \leq m} \nu_c(\tilde{B}_{m,i-1}^*) = \lim_{m \to \infty} \nu_c(\tilde{B}_{m,1}^*) = 0.$$

Therefore, we obtain

$$\lim_{m \to \infty} \lambda_C(m)$$
$$= 1 - \alpha + \alpha \lim_{m \to \infty} \max_{2 \leq i \leq m} \left(\nu_c(\tilde{B}_{m,i-1}^*) - \nu_c(\tilde{B}_{m,i}^*) \right)$$
$$= 1 - \alpha.$$

If $\lim_{m \to \infty} \lambda_C(m) = 0$, then from the above equation we know that $\alpha = 1$ in (2.2), which implies that the density of copula C exists.

(II) Now we assume that $\lambda_C(m) = \lambda \in [0,1]$ for all $m \geq 2$, then $\lim_{m \to \infty} \lambda_C(m) = \lambda = 1 - \alpha$. On the other hand, $\lambda_C(m) = 1 - \alpha$ implies $\nu_c(\tilde{B}_{m,i}^*) = \nu_c(\tilde{B}_{m,i-1}^*)$ for any $2 \leq i \leq m$. By the same argument in the proof of Theorem 3.2, we conclude that the density $\tilde{c} = d\nu_c/d\mu$ satisfies $\tilde{c}(u, v) = 1$ on $[0,1]^2$ almost surely. Therefore ν_c is a probability measure generated by the independent copula. □

Next we will give some examples for a better understanding of the concentration measure Γ_C.

Example 3.4. For the Fréchet copula C defined in Example 3.3, $\Gamma_C = (\alpha + \gamma, \alpha + \gamma, \dots)$. We know that the Fréchet copula C is a mixture of the comonotonic copula, the countermonotonic copula and the independent copula, so $\lambda_C(m)$ is the sum of the weights of the comonotonic and countermonotonic parts. Thus $\lambda_C(m)$ measures the difference between the Fréchet copula and the independent copula.

Example 3.5. For Gaussian copula C_ρ^N with correlation $\rho \in [0, 1)$, we can calculate the measure $\lambda_{C_\rho^N}(m)$ numerically for different m, which is shown in Figure 4. For $\rho \in (-1, 0)$ we know that $\lambda_{C_\rho^N}(m) = \lambda_{C_{-\rho}^N}(m)$ from the symmetry of Gaussian copula density. From Figure 4 we can see that $\lambda_{C_\rho^N}(m)$ is positively correlated with ρ. It is because the
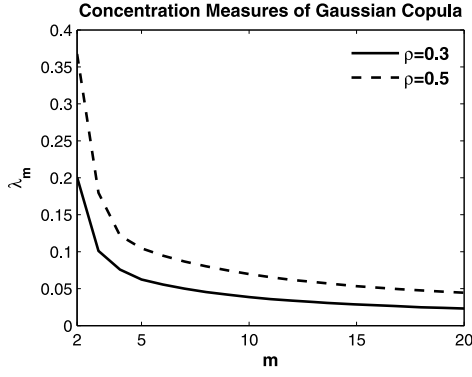
Figure 4. Concentration measure of Gaussian copula with different ρ and m.

difference between the Gaussian copula C_N^ρ and the independent copula becomes larger when ρ increases. For a fixed ρ, the measure $\lambda_C(m)$ becomes smaller when m increases, and tends to 0 as $m \to \infty$.

4. EMPIRICAL SUPPORT FOR THE CONCENTRATION SETS

In this section, we apply the concentrated partition to the copula of level and slope factors in Chinese government bond yield and give the empirical support for our methodology.

4.1 Data description and modeling

In this subsection we introduce the Nelson-Siegel model to determine the level factor l_t and the slope factor s_t in Chinese government bond yield data.

Nelson-Siegel (NS) model is widely used to describe the term structure of interest rates. This model is firstly introduced by Nelson and Siegel (1987) and supported by a number of papers such as Diebold and Li (2006) and Luo, Han and Zhang (2012). In this model the instantaneous forward rate $f(t, \tau)$ satisfies the following equation

$$f(t, \tau) = l_t + s_t e^{-\lambda\tau} + c_t \lambda\tau e^{-\lambda\tau},$$

where t is the current time, τ is the time to maturity and λ is the scale parameter. The three factors l_t, s_t and c_t are the level, slope and curvature factors of the yield curve, respectively. Therefore, the yield term structure is

$$(4.1) \quad y(t, \tau) = l_t + s_t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + c_t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right).$$

We use daily Chinese inter-bank treasury bond yields from January 1, 2006 to September 30, 2012 with 1689 trading days from the China bond website[1]. Following Diebold and Li (2006), we set $\lambda = 0.16$ as a pre-specified constant,

[1] http://www.chinabond.com.cn

then we use ordinary least squares (OLS) to estimate the level, slope and curvature factors for each day.

Next we focus on the daily difference of the level and slope factors, i.e., we study the historical copula function of $X_t = l_t - l_{t-1}$ and $Y_t = s_t - s_{t-1}$. Precisely, we estimate the marginal distributions of X_t and Y_t empirically, then we use Beta kernel function to estimate the copula density function.

Essentially, the pseudo-sample points are defined as

$$\hat{U}_t = \frac{\text{rank}\,(X_t)}{T + 1} = \frac{1}{T + 1} \sum_{k=1}^{T} 1_{X_k \leq X_t},$$

$$\hat{V}_t = \frac{\text{rank}\,(Y_t)}{T + 1} = \frac{1}{T + 1} \sum_{k=1}^{T} 1_{Y_k \leq Y_t},$$

where $T = 1{,}688$ is the length of samples (X_t, Y_t). The pseudo-samples are shown in Figure 5(a). Based on the pseudo samples $\hat{U}_t$ and $\hat{V}_t$, the Beta kernel estimation (Charpentier, Fermanian and Scaillet, 2007) for the density c is

$$\hat{c}(u_1, u_2) = \frac{1}{T} \sum_{t=1}^{T} K\left(\hat{U}_t; \frac{u_1}{h} + 1, \frac{1 - u_1}{h} + 1 \right)$$
$$\times K\left(\hat{V}_t; \frac{u_2}{h} + 1, \frac{1 - u_2}{h} + 1 \right),$$

where $(u_1, u_2) \in (0,1)^2$, h is the window width, and $K(\cdot; \alpha, \beta)$ is the Beta kernel function

$$K(x; \alpha, \beta) = \frac{x^{\alpha-1}(1 - x)^{\beta-1}}{B(\alpha, \beta)}, \ x \in (0,1), \ \alpha, \beta > 0,$$

in which the Beta function $B(\alpha, \beta) = \int_0^1 s^{\alpha-1}(1 - s)^{\beta-1} ds$. The Beta kernel function $K(\cdot; \alpha, \beta)$ is chosen to meet the natural property that the copula density is defined on the set $[0,1]^2$. Furthermore, it is free of boundary bias and can improve the accuracy of estimation. Figure 5(b) is the copula kernel density, and the window width h is chosen to be 0.031, which is optimized according to Chen (1999).

4.2 Concentrated partition of the copula between level and slope factors

In this subsection, we apply our concentrated partition to the above fitted copula, and the result of the empirical study shows our partition is more meaningful than the traditional regular partition.

From the pseudo-sample scatter plot in Figure 5(a), we find that the points gather at the corners $(0,0)$ and $(1,1)$, and are around the line $\{u_1 + u_2 = 1\}$, which is the support of the countermonotonic copula. Also in the estimated copula density in Figure 5(b), the density function is obviously higher in these regions.

Figure 5(c) shows the concentrated 3-partition, in which the Lebesgue measure of each part is 1/3. The concentrated

(a) Pseudo–Sample Plot

(b) Copula Kernel Density

(c) Concentrated Partition
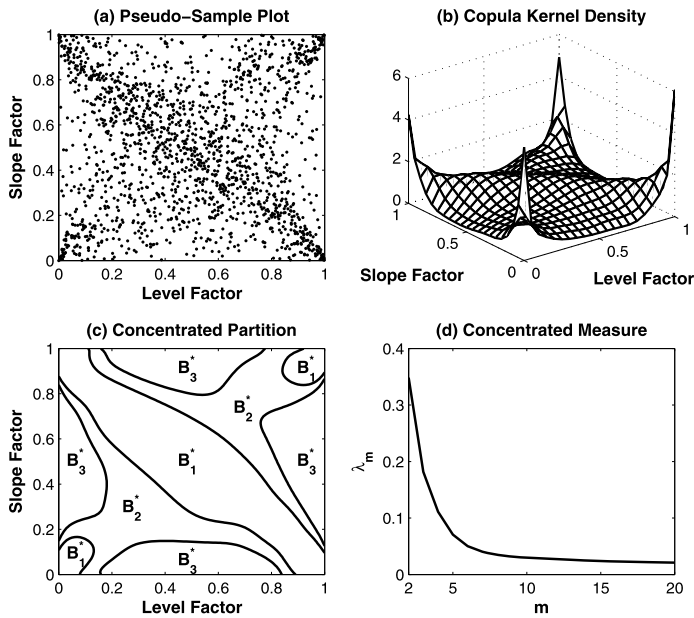
(d) Concentrated Measure

Figure 5. (a) Pseudo-samples of daily difference of the level and slope factors. (b) Copula density estimated by Beta kernel function. (c) Concentrated partition with level $\mathbf{a} = (1/3, 1/3, 1/3)$. (d) The concentration measure $\lambda_m(C)$.

partition finds out the most important set, which illustrates the main feature of the correlation between the two factors. The partition result shows that samples concentrate at the corners $(0,0)$ and $(1,1)$ as well as along the line $\{u_1 + u_2 = 1\}$.

Figure 5(d) shows the concentration measure $\lambda_C(m)$ for the fitted copula. When $m = 3$, the measure is about 0.3, which implies that the difference among the probabilities of the copula on the three subsets is fairly large. And when m is large, the partition becomes dense and $\lambda_C(m)$ is close to 0.

By investigating the details of the dependence structure between $X_t = \Delta l_t$ and $Y_t = \Delta s_t$, we can find that the daily change of short rate has a large kurtosis. According to equation (4.1), we have $\lim_{\tau \to 0} y(t, \tau) = l_t + s_t$. Hence for a small time-to-maturity τ, the daily change of $y(t, \tau)$ is more likely to be small if Δl_t and Δs_t are countermonotonic, and daily change of $y(t, \tau)$ is more likely to be large if Δl_t and Δs_t are comonotonic, so the daily change of short rate should have an obvious peak and fat tails. In fact, during January 2006 to September 2012, the kurtosis of the first difference of one-year rate is 14, while that of ten-year rate is 6. Therefore, looking into the copula function between Δl_t and Δs_t provides more detailed information of term structure.

5. CONCLUSION

In this paper we defined the concentration set of a copula function for discussing the local correlation structure of a copula function. Based on the concentration set, we also established a concentrated partition for copula function. The properties of the concentration set and the concentrated partition were discussed. Based on the concentrated partition, concentration measure were defined for measuring copula function's local correlation structure. At last, an empirical study was provided to support our idea of establishing the concept of the concentration set.

REFERENCES

BJERVE, S. and DOKSUM, K. (1993). Correlation curves: measures of association as function of covariates values. *Annals of Statistics* **21**, 890–902. MR1232524

CHARPENTIER, A., FERMANIAN, J. D. and SCAILLET, O. (2007). The estimation of copulas: theory and practice. *Copulas: from theory to application in finance*, 35–60. Risk Publications, London.

CHEN, S. X. (1999). Beta kernel estimators for density functions. *Computational Statistics and Data Analysis* **31**, 131–145. MR1718494

CHERUBINI, U., LUCIANO, E. and VECCHIATO, W. (2004). *Copula Methods in Finance*. John Wiley & Sons, Chichester. MR2250804

DIEBOLD, F. X. and LI, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics* **130**, 337–364. MR2211798

DROUET-MARI, D. and KOTZ, S. (2001). *Correlation and Dependence*. Imperial College Press, London. MR1835042

DURANTE, F., SARKOCI, P. and SEMPI, C. (2009). Shuffle of copulas. *Journal of Mathematical Analysis and Applications* **352**, 914–921. MR2501937

DURRLEMAN, V., NIKEGHBALI, A. and RONCALLI, T. (2000). Copulas approximation and new families. Working paper.

HALMOS, P. R. (1970). *Measure Theory*. Springer-Verlag, New York.

HOLLAND, P. W. and WANG, Y. J. (1987a). Regional dependence for continuous bivariate densities. *Communications in Statistics: Theory and Methods* **16**, 193–206. MR0885955

HOLLAND, P. W. and WANG, Y. J. (1987b). Dependence function for continuous bivariate densities. *Communications in Statistics: Theory and Methods* **16**, 863–876. MR0886560

JOE, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London. MR1462613

KOLEV, N., ANJOS, U. and MENDES, B. V. M. (2006). Copulas: a review and recent developments. *Stochastic Models* **22**, 617–660. MR2263860

KRUSKAL, W. H. (1958). Ordinal measures of association. *Journal of the American Statistical Association* **53**, 814–861. MR0100941

LI, X., MIKUSIŃSKI, P. and TAYLOR M. D. (1998). Strong approximations of copulas. *Journal of Mathematical Analysis and Applications* **225**, 608–623. MR1644300

LUO, X., HAN, H. and ZHANG J. E. (2012). Forecasting the term structure of Chinese Treasury yields. *Pacific-Basin Finance Journal* **20**, 639–659.

MCNEIL, A. J., FREY, R. and EMBRECHTS, P. (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, Princeton. MR2175089

MIKUSIŃSKI, P., SHERWOOD, H., TAYLOR, M. D. (1992). Shuffles of min. *Stochastica* **13**, 61–74. MR1197328

MIKUSIŃSKI, P. and TAYLOR, M. D. (2010). Some approximations of n-copulas. *Metrika* **72**, 385–414. MR2746583

Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd Edition. Springer, New York. MR2197664

Nelson, C. R. and Siegel, A. F. (1987). Parsimonious modeling of yield curve. *Journal of Business* **60**, 473–489.

Scancetta, A. and Satchell, S. (2004). The Bernstein copula and its applications to modelling and approximations of multivariate distributions. *Econometic Theory* **20**, 535–562. MR2061727

Zheng, Y., Yang, J. and Huang, J. Z. (2011). Approximation of bivariate copulas by patched bivariate Fréchet copulas. *Insurance: Mathematics and Economics* **48**, 246–256. MR2799306

Lujun Li
Department of Financial Mathematics
Peking University
Beijing, 100871
China
E-mail address: lujun.li@pku.edu.cn

Yijun Wu
Department of Financial Mathematics
Peking University
Beijing, 100871
China
E-mail address: wuyijun00@gmail.com

Jingping Yang
LMEQF, Department of Financial Mathematics
School of Mathematical Sciences
and Center for Statistical Sciences
Peking University
Beijing, 100871
China
E-mail address: yangjp@math.pku.edu.cn