J. R. Statist. Soc. B (2018)
80, *Part* 2, *pp.* 409–432

Semiparametric dynamic max-copula model for multivariate time series

Zifeng Zhao and Zhengjun Zhang
University of Wisconsin—Madison, USA

[Received May 2016. Final revision August 2017]

Summary. The paper presents a novel non-linear framework for the construction of flexible multivariate dependence structure (i.e. copulas) from existing copulas based on a straightforward 'pairwise max-'rule. The newly constructed max-copula has a closed form and has strong interpretability. Compared with the classical 'linear symmetric' mixture copula, the max-copula can be viewed as a 'non-linear asymmetric' framework. It is capable of modelling asymmetric dependence and joint tail behaviour while also offering good performance in non-extremal behaviour modelling. Max-copulas that are based on single-factor and block factor models are developed to offer parsimonious modelling for structured dependence, especially in high dimensional applications. Combined with semiparametric time series models, the max-copula can be used to develop flexible and accurate models for multivariate time series. A new semiparametric composite maximum likelihood method is proposed for parameter estimation, where the consistency and asymptotic normality of estimators are established. The flexibility of the max-copula and the accuracy of the proposed estimation procedure are illustrated through extensive numerical experiments. Real data applications in value-at-risk estimation and portfolio optimization for financial risk management demonstrate the max-copula's promising ability to capture accurately joint movements of high dimensional multivariate stock returns under both normal and crisis regimes of the financial market.

Keywords: Asymmetric dependence; Composite maximum likelihood; Copula construction; Market crisis; Mixture modelling; Tail dependence

1. Introduction

Modelling the multivariate joint behaviour of random variables is one of the most fundamental tasks in statistical modelling. The construction of multivariate distributions is technically difficult and most of the early multivariate modelling was restricted within the Gaussian or elliptical family. Thanks to Sklar's (1959) theorem, which states that multivariate dependence can be separated into a copula and individual marginal distributions, the 'time of the copula' has emerged for the construction of multivariate distributions. Various copulas have been proposed in the literature; see Joe (2014) and Nelsen (2006) for a summary. Copula-based models for multivariate distributions are widely used in a variety of applications; see Frees and Valdez (1998) in actuarial science and insurance, Cherubini *et al.* (2004) in finance and Genest and Favre (2007) in hydrology.

Because of its tractability, interpretability and flexibility in modelling non-extremal joint behaviour, the Gaussian copula is arguably the most widely used copula. Although Gaussian copulas perform well in many areas of applications, the financial market may turn out to be an exception. One of the most significant characteristics of financial data are their tail dependence,

Address for correspondence: Zhengjun Zhang, Department of Statistics, University of Wisconsin—Madison, 1300 University Avenue, Madison, WI 53706, USA.
E-mail: zjz@stat.wisc.edu

 1369–7412/18/80409

i.e., during a crisis, asset prices tend to move together. The failure of Gaussian copulas to capture this tail dependence in the pricing of collateralized debt obligations and related securities is considered to be one of the prominent causes of the recent financial crisis; for example, see Coval *et al.* (2009) and Salmon (2012) for more details.

The Gaussian copula's inability to model joint tail events inspires more research in the construction of copulas that can offer more sophisticated dependence structures. One direction is to exploit and extend the linear structure in Gaussian factor models by changing the distribution of latent factors from Gaussian to other distributions like *skewed t*, e.g. Hull and White (2004), Murray *et al.* (2013) and Oh and Patton (2017). This framework is capable of offering more sophisticated dependence like tail dependence and tail asymmetry. Also, it is particularly attractive in high dimensional applications thanks to the factor structure. Another promising direction is to use vine copulas, which build high dimensional copulas on the basis of a sequentially iterative pairwise construction of bivariate copulas; see Aas *et al.* (2009), Min and Czado (2010) and Almeida *et al.* (2012) for more details. Vine copulas can offer flexible dependence relationships and can be represented in graphs, which helps the modeller to understand the dependence structure visually. Combining the 'latent factor' idea with the pairwise construction idea of C-vine copulas, Krupskii and Joe (2013) proposed a factor copula model where, instead of imposing a Gaussian linear structure, bivariate copulas are used to specify the dependence between latent factors and observed variables. Theorem A1 in section 1 of the on-line supplementary material establishes a weak equivalence between the factor copula model based on the C-vine copula and the linear factor copula model under the additive model framework.

A much different yet important direction is to construct new sophisticated copulas based on existing copulas. The idea is that the new copula inherits various merits from its parents and thus offers more versatile dependence structures. In the literature, mixtures of distributions are a long existing technique for generating new distributions based on existing distributions. The mixture copula has a closed form cumulative distribution function (CDF) and is interpretable. Inspired by Cui and Zhang (2017) and the mixture technique, in this paper we propose a novel non-linear copula construction framework named the 'max-copula', which generates new copulas based on existing copulas via a straightforward pairwise max-function. The max-copula has a closed form CDF and has strong interpretability, especially in financial applications. Compared with mixture copulas, it can actively generate more flexible dependence structures, including asymmetric dependence structures. Moreover, because of its unique 'pairwise max-'characteristic, it can offer a better modelling of non-extremal behaviour while attaining an accurate modelling of tail dependence. Combined with semiparametric time series models, the max-copula has shown its advantages over the mixture copula: it can accurately capture asymmetric dependence and joint extremal movements of multivariate financial time series while simultaneously offering better modelling of non-extremal market behaviour.

The remainder of the paper is structured as follows. Section 2 presents the max-copula model, derives its quantile and tail dependence properties and discusses the selection of component copulas for the max-copula. Unique characteristics of max-copulas are emphasized. For high dimensional applications, single-factor and block factor max-copulas are developed. Section 3 describes the composite maximum likelihood estimation (CMLE) method under the semiparametric time series setting. Numerical experiments on the flexibility of max-copulas and the performance of CMLE are conducted in Section 4. In Section 5, we present empirical applications of max-copulas in the estimation of conditional value at risk, VaR, for a financial portfolio and in the construction of optimal portfolios based on 30 component stocks from the Dow Jones industrial average. Section 6 concludes with a discussion of potential extensions. All the proofs can be found in the on-line supplementary material.

The programs that were used to analyse the data and some simulated data can be obtained from

`http://wileyonlinelibrary.com/journal/rss-datasets`

2. Max-copula construction

2.1. Model specification

The idea of the max-copula is partially motivated by the mixture copula. Suppose that we have two copulas of the same dimension $d \geqslant 2$, say $\mathbf{C}_1$ and $\mathbf{C}_2$. For any $0 \leqslant c \leqslant 1$, the linear mixture $\mathbf{C} = c\mathbf{C}_1 + (1-c)\mathbf{C}_2$ is a new copula. The stochastic representation of $\mathbf{C}$ can be built as follows. Suppose that $\mathbf{U}_1 = (U_{11}, \ldots, U_{1d}) \sim \mathbf{C}_1$, $\mathbf{U}_2 = (U_{21}, \ldots, U_{2d}) \sim \mathbf{C}_2$, $X \sim \text{Bernoulli}(c)$ and $(\mathbf{U}_1, \mathbf{U}_2, X)$ are mutually independent; then $\mathbf{U} = (U_1, \ldots, U_d) = \max(\mathbf{U}_1^{1/X}, \mathbf{U}_2^{1/(1-X)}) \sim \mathbf{C}$, where $\max(\cdot)$ is a pairwise max-function, i.e. $U_i = \max(U_{i1}^{1/X}, U_{i2}^{1/(1-X)})$ (here we define $1/0 = \infty$). In what follows, we always assume that $(\mathbf{U}_1, \mathbf{U}_2, X)$ are mutually independent.

A closer examination reveals that the joint distribution of $\mathbf{U}$ is always a copula, as long as X is a random variable on [0, 1]. A general distribution on the interval [0,1] is $\text{beta}(a, b)$. Suppose that $X \sim \text{beta}(a, b)$; we obtain the following copula model:

$$\mathbf{U} = \max(\mathbf{U}_1^{1/X}, \mathbf{U}_2^{1/(1-X)}), \tag{1}$$

where $\mathbf{U}_1 \sim \mathbf{C}_1$, $\mathbf{U}_2 \sim \mathbf{C}_2$ and $X \sim \text{beta}(a, b)$. By an elementary argument, the copula $C_{\mathbf{U}}$ of $\mathbf{U}$ is

$$\mathbf{C}_{\mathbf{U}}(u_1, \ldots, u_d) = \mathbb{E}\{\mathbf{C}_1(u_1^X, \ldots, u_d^X)\, \mathbf{C}_2(u_1^{1-X}, \ldots, u_d^{1-X})\}. \tag{2}$$

A closed form solution of equation (2) is generally not available because of the expectation on X; however, it can be computed numerically via one-dimensional integration. In this paper, we consider the case when X is a Dirac mass on c, where $0 < c < 1$. Under this framework, the expectation can be dropped and we have

$$\mathbf{U} = \max(\mathbf{U}_1^{1/c}, \mathbf{U}_2^{1/(1-c)}),$$

i.e.

$$U_i = \max(U_{1i}^{1/c}, U_{2i}^{1/(1-c)}), \qquad \text{for } i = 1, 2, \ldots, d,$$

where $\mathbf{U}_1 \sim \mathbf{C}_1$, $\mathbf{U}_2 \sim \mathbf{C}_2$ and $\mathbf{C}_1$ and $\mathbf{C}_2$ are two existing copulas. Given u, $u^{1/c}$ is an increasing function of c and $u^{1/(1-c)}$ is a decreasing function of c. For each U_i, we have $P(U_i = U_{1i}^{1/c}) = P(U_{1i}^{1/c} > U_{2i}^{1/(1-c)}) = c$, so c can be viewed as the weight parameter that controls the relative strength of $\mathbf{C}_1$ and $\mathbf{C}_2$ in the max-copula $\mathbf{C}$, with a large c favouring $\mathbf{C}_1$ and small c favouring $\mathbf{C}_2$. The newly constructed $\mathbf{C}$ takes the form

$$\mathbf{C}(u_1, \ldots, u_d) = \mathbf{C}_1(u_1^c, \ldots, u_d^c)\, \mathbf{C}_2(u_1^{1-c}, \ldots, u_d^{1-c}), \tag{3}$$

and we call $\mathbf{C}$ a max-copula.

Remark 1. Both the mixture copula and the max-copula can be seen as limiting cases of the general copula (2) where $X \sim \text{beta}(a, b)$. Let $X \sim \text{beta}\{c/n, (1-c)/n\}$; we have $X \to_{\mathrm{d}} \text{Bernoulli}(c)$ as $n \to \infty$ and $X \to_{\mathrm{d}} \text{Dirac mass}(c)$ as $n \to 0$. Therefore, the general copula (2) can be seen as a 'bridge' between the max-copula and the mixture copula and is capable of generating more flexible dependence structures. Its properties are being investigated in a separate project.

Remark 2. We derive the max-copula (3) under the novel stochastic representation (1). We note that a more general form of equation (3) can be found in Liebscher (2008) and Durante and

Sempi (2015), which can be seen as a generalization of the Khoudraji device in Khoudraji (1995). To our best knowledge, the present paper provides the first thorough study of the max-copula's probabilistic properties and establishes novel semiparametric statistical inference procedures that make its real data application feasible.

Remark 3. An obvious generalization of max-copulas is to allow different cs for different (U_{1i}, U_{2i})s. This extension can produce non-exchangeable max-copulas even when its components $\mathbf{C}_1$ and $\mathbf{C}_2$ are exchangeable. Such an extension is not obvious in corresponding mixture copulas.

For mixture copulas, whether $\mathbf{U}$ behaves like $\mathbf{U}_1$ or $\mathbf{U}_2$ does not depend on $(\mathbf{U}_1, \mathbf{U}_2)$ itself, but on an independent Bernoulli random variable X, whereas for max-copulas it does depend on $(\mathbf{U}_1, \mathbf{U}_2)$. Following the pairwise max-rule, for each U_i in $\mathbf{U}$, we have $U_i = \max(U_{1i}^{1/c}, U_{2i}^{1/(1-c)})$, so $\mathbf{U}$ behaves like the more extreme of $\mathbf{U}_1^{1/c}$ and $\mathbf{U}_2^{1/(1-c)}$. This direct interaction between $\mathbf{U}_1$ and $\mathbf{U}_2$ is more realistic and has more meaningful interpretability in many areas. In financial applications, we can think of $\mathbf{U}$ as the risks for multiple stocks. Marginally speaking, for each stock U_i, there are two sources of risk coming from U_{1i} and U_{2i} with c controlling the relative weight and each stock taking the larger risk from the two sources. Jointly speaking, $\mathbf{U}_1$ and $\mathbf{U}_2$ represent the two sources of joint risks with different joint behaviour; for example $\mathbf{U}_1$ may follow a copula $\mathbf{C}_1$ with weak multivariate dependence, representing the joint risks in a 'normal state' market, whereas $\mathbf{U}_2$ can follow a copula $\mathbf{C}_2$ with strong joint tail dependence, representing the joint risks in a 'crisis state' market. Further details about the interpretation can be found in Section 2.3. As is seen later in real data applications, this unique characteristic of the max-copula helps it to capture the joint behaviour of multivariate financial time series accurately.

2.2. *Quantile dependence and tail dependence coefficient*

Suppose that the bivariate random vector $(U_1, U_2) \sim \mathbf{C}$, where $\mathbf{C}$ is the max-copula derived from $\mathbf{C}_1$ and $\mathbf{C}_2$, by equation (3). The quantile dependence $\lambda_{\max}^q$ between (U_1, U_2) takes the closed form

$$\lambda_{\max}^q = \begin{cases} P(U_1 \leqslant q | U_2 \leqslant q) = \dfrac{\mathbf{C}_1(q^c, q^c)\,\mathbf{C}_2(q^{1-c}, q^{1-c})}{q}, & q \in (0, 0.5], \\ P(U_1 > q | U_2 > q) = \dfrac{1 - 2q + \mathbf{C}_1(q^c, q^c)\,\mathbf{C}_2(q^{1-c}, q^{1-c})}{1-q}, & q \in (0.5, 1). \end{cases}$$

Sibuya (1959) introduced the concept of a tail dependence coefficient as a simple criterion to quantify the joint extreme behaviour of two random variables. Let (X_1, X_2) be a random vector; the quantity

$$\lambda^{\mathrm{U}} = \lim_{x \to x_F} P(X_2 > x | X_1 > x), \qquad x_F = \sup\{x : F(x) < 1\},$$

is called the upper tail dependence coefficient, provided that the limit exists (in this definition, X_1 and X_2 are required to have identical marginal distribution $F(x)$). The joint distribution of (X_1, X_2) is said to have upper tail dependence if $\lambda^{\mathrm{U}} > 0$ and upper tail independence if $\lambda^{\mathrm{U}} = 0$. Similarly, we can define the lower tail dependence coefficient λ^{L}. Theorem 1 states that the tail dependence coefficients of max-copulas have nice closed form expressions and behave differently from those of mixture copulas. In what follows, we denote the upper and lower tail dependence coefficients of $\mathbf{C}_1$ as λ_1^{U} and λ_1^{L}, and those of $\mathbf{C}_2$ as λ_2^{U} and λ_2^{L}.

Theorem 1. For the max-copula $\mathbf{C}_{\max}$ based on $\mathbf{C}_1$ and $\mathbf{C}_2$ with weight c, the upper and lower tail dependence coefficients are $\lambda_{\max}^{\mathrm{U}} = c\lambda_1^{\mathrm{U}} + (1-c)\lambda_2^{\mathrm{U}}$ and $\lambda_{\max}^{\mathrm{L}} = \lambda_1^{\mathrm{L}}\lambda_2^{\mathrm{L}}$. For the mixture copula $\mathbf{C}_{\text{mix}}$ based on $\mathbf{C}_1$ and $\mathbf{C}_2$ with weight c, the upper and lower tail dependence coefficients are $\lambda_{\text{mix}}^{\text{U}} = c\lambda_1^{\text{U}} + (1-c)\lambda_2^{\text{U}}$ and $\lambda_{\text{mix}}^{\text{L}} = c\lambda_1^{\text{L}} + (1-c)\lambda_2^{\text{L}}$.

By theorem 1, there is a clear difference between the tail behaviour of max-copulas and mixture copulas. If both component copulas $\mathbf{C}_1$ and $\mathbf{C}_2$ have symmetric tail dependence (i.e. $\lambda_1^{\text{L}} = \lambda_1^{\text{U}}$ and $\lambda_2^{\text{L}} = \lambda_2^{\text{U}}$), the mixture copula gives symmetric tail dependence, whereas the max-copula can clearly offer asymmetric tail dependence, which is often found to be more appealing in many applications. For example, a mixture of the two most widely used copulas, the Gaussian and the t-copula, gives symmetric tail dependence, whereas the corresponding max-copula offers asymmetric tail dependence with upper tail dependence and lower tail independence. On the basis of theorem 1, we can obtain the following corollary.

Corollary 1. Assume that both component copulas $\mathbf{C}_1$ and $\mathbf{C}_2$ are 'diagonally' symmetric (i.e. $\mathbf{C}_i(q,q) = 2q - 1 + \mathbf{C}_i(1-q, 1-q)$, for all $q \in (0,1)$, $i = 1,2$). By symmetry we have $\lambda_1^{\text{L}} = \lambda_1^{\text{U}} = \lambda_1$ and $\lambda_2^{\text{L}} = \lambda_2^{\text{U}} = \lambda_2$. If $\min(\lambda_1, \lambda_2) < 1$ and $\max(\lambda_1, \lambda_2) > 0$, then, for the upper quantile dependence $\lambda_{\max}^{q}$ and lower quantile dependence $\lambda_{\max}^{1-q}$ of $\mathbf{C}_{\max}$, there is always a $q^* > 0.5$ such that $\lambda_{\max}^{q} > \lambda_{\max}^{1-q}$, for all $q \in (q^*, 1)$. In contrast, for $\mathbf{C}_{\text{mix}}$, we have $\lambda_{\text{mix}}^{q} = \lambda_{\text{mix}}^{1-q}$, for all $q \in (0.5, 1)$.

Remark 4. In Nelsen (2006), a copula $\mathbf{C}$ is said to be 'radially' symmetric if $\mathbf{C}(u,v) = \bar{\mathbf{C}}(u,v)$ for all $(u,v) \in (0,1)^2$, where $\bar{\mathbf{C}}(u,v) = u + v - 1 + \mathbf{C}(1-u, 1-v)$ is the survival copula of $\mathbf{C}$. Diagonal symmetry in corollary 1 is a weaker condition than radial symmetry and all copulas in the elliptical copula family are radially symmetric.

Corollary 1 further demonstrates the fundamental difference between the max-copula and the mixture copula. The max-copula is an 'asymmetric system' that is capable of actively generating asymmetric dependence from symmetric component copulas, whereas the mixture copula is a 'symmetric system' that can only inherit asymmetric dependence from an asymmetric component copula.

An extreme example is when $\mathbf{C}_1 = \mathbf{C}_2 = \mathbf{C}$. For all copulas $\mathbf{C}$ we have $\mathbf{C}_{\text{mix}} = \mathbf{C}$, whereas for almost all widely used copulas $\mathbf{C}$ (except extreme value copulas which are max-stable) we have $\mathbf{C}_{\max} \neq \mathbf{C}$. Under the condition of corollary 1, the upper quantile dependence of $\mathbf{C}_{\max}$ is stronger than its corresponding lower quantile dependence, which is a desirable property, especially in financial risk management.

2.3. Choice of copulas and unique characteristics of the max-copula

Since the max-copula is based on component copulas $\mathbf{C}_1$ and $\mathbf{C}_2$, to construct a max-copula, we need to specify $\mathbf{C}_1$ and $\mathbf{C}_2$, which is a model selection problem. For bivariate max-copulas, a rule of thumb can be developed based on theorem 1: if we want the max-copula to have upper tail dependence, at least one of the copulas needs to be upper tail dependent, whereas, if we need lower tail dependence, both of the copulas are required to have lower tail dependence. Users can choose different copulas to suit different applications. In the elliptical copula family, the Gaussian copula has no upper and lower tail dependence; the t-copula has both upper and lower tail dependence. Meanwhile, in the Archimedean copula family, the Gumbel copula has upper tail dependence and the Clayton copula has lower tail dependence.

In general, the selection of max-copulas is a difficult topic, as is the selection of mixture copulas. A standard procedure in the literature is to use a likelihood-based information criterion (e.g. the Akaike information criterion or Bayesian information criterion) to select $\mathbf{C}_1$ and $\mathbf{C}_2$ from a pool of candidate copulas. There are various extensions of the Akaike information criterion and Bayesian information criterion for composite likelihood (which is later used for the estimation of max-copulas), e.g. Gao and Song (2010) and Varin *et al.* (2011). Another promising direction is to follow the work in Cai and Wang (2014), who proposed a penalized likelihood procedure via shrinkage operators for the selection of mixture copulas, which selects appropriate copulas and estimates related parameters simultaneously.

The primary purpose of this paper is to model multivariate financial time series, e.g. negative daily returns of multiple stocks. One of the most significant characteristics of the stock market is that stock returns tend to have greater dependence during a crisis and tend to behave more 'independently' otherwise. In other words, negative stock returns have asymmetric tail dependence with strong upper tail dependence and weak to no lower tail dependence; see Oh and Patton (2017) for an example. Thus, in general, we want to design a max-copula that has asymmetric tail behaviour with strong upper tail dependence.

As mentioned above, the Gaussian copula has no tail dependence and thus lacks the ability to capture the 'joint crash' property of the stock market. However, it performs well in capturing moderate scale stock returns, i.e. the Gaussian copula can be used to model the normal state stock market. In contrast, the Gumbel copula is capable of modelling joint upper tail dependence, whereas it may not perform well under moderate scale since it is an extreme value copula, i.e. the Gumbel copula can be used to model the crisis state stock market.

On the basis of the above observations, to capture better the multivariate dependence structure of the stock market, in this paper we choose $\mathbf{C}_1$ to be a Gaussian copula with correlation matrix Σ and $\mathbf{C}_2$ to be a Gumbel copula with parameter α. By theorem 1, the constructed max-copula $\mathbf{C}$ has a lower tail dependence coefficient of 0 and an upper tail dependence coefficient of $(1-c)(2-2^{1/\alpha})$. Since the Gaussian copula is a special case of the elliptical copula and the Gumbel copula is a special case of the Archimedean copula, the generalization of the max-copula based on the elliptical family and the Archimedean family easily follows.

If $c=1$, the max-copula $\mathbf{C}$ degenerates to a Gaussian copula $\mathbf{C}_1$ and there is no upper or lower tail dependence between the U_is. If $c=0$, the max-copula $\mathbf{C}$ degenerates to a Gumbel copula $\mathbf{C}_2$. When $0<c<1$, the Gaussian copula $\mathbf{C}_1$ helps to regulate the U_i's dependence structure under moderate scale and the Gumbel copula $\mathbf{C}_2$ helps to achieve upper tail dependence between the U_is. To help to understand the unique dependence structure of the max-copula better, Fig. 1 shows scatter plots of simulated data $\mathbf{U}=(U_1,U_2)$ from three sets of parameters (c,α,ρ) for a bivariate max-copula $\mathbf{C}(u_1,u_2)$. For a better visual illustration, the data have been transformed to be marginally normally distributed by using $\Phi^{-1}(U_i)$, where Φ is the cumulative distribution function of the standard normal distribution. In the graph, circles correspond to data where the Gumbel copula dominates (i.e. both points of $\mathbf{U}$ come from the Gumbel copula), triangles correspond to data where the Gaussian copula dominates (i.e. both points of $\mathbf{U}$ come from the Gaussian copula) and crosses correspond to the 'mixed' case where one point of $\mathbf{U}$ comes from the Gumbel copula and one from the Gaussian copula. The three plots demonstrate the case where the Gaussian copula has positive correlation ($\rho=0.5$) and the Gumbel copula has moderate upper tail dependence ($\alpha=2$). From left to right, c decreases from 0.8 to 0.2, i.e. the influence of the Gaussian copula is declining and that of the Gumbel copula is increasing. As is expected, the Gumbel copula dominantly regulates the upper tail area, whereas the Gaussian copula mainly affects the distribution in the non-extremal area.

The unique characteristic of the max-copula is well demonstrated in Fig. 1. Note here in the upper tail region, for all three cs, that $\mathbf{U}$ always come from the Gumbel copula because of its upper tail dependence, whereas, in the non-extremal region, the graph is a mix of three shapes with triangles or crosses taking the dominance, depending on the magnitude of c. Since, in comparison with the Gaussian copula, the Gumbel copula is not suitable for modelling non-extremal behaviour, the existence of the mixed cross points helps to achieve a better modelling

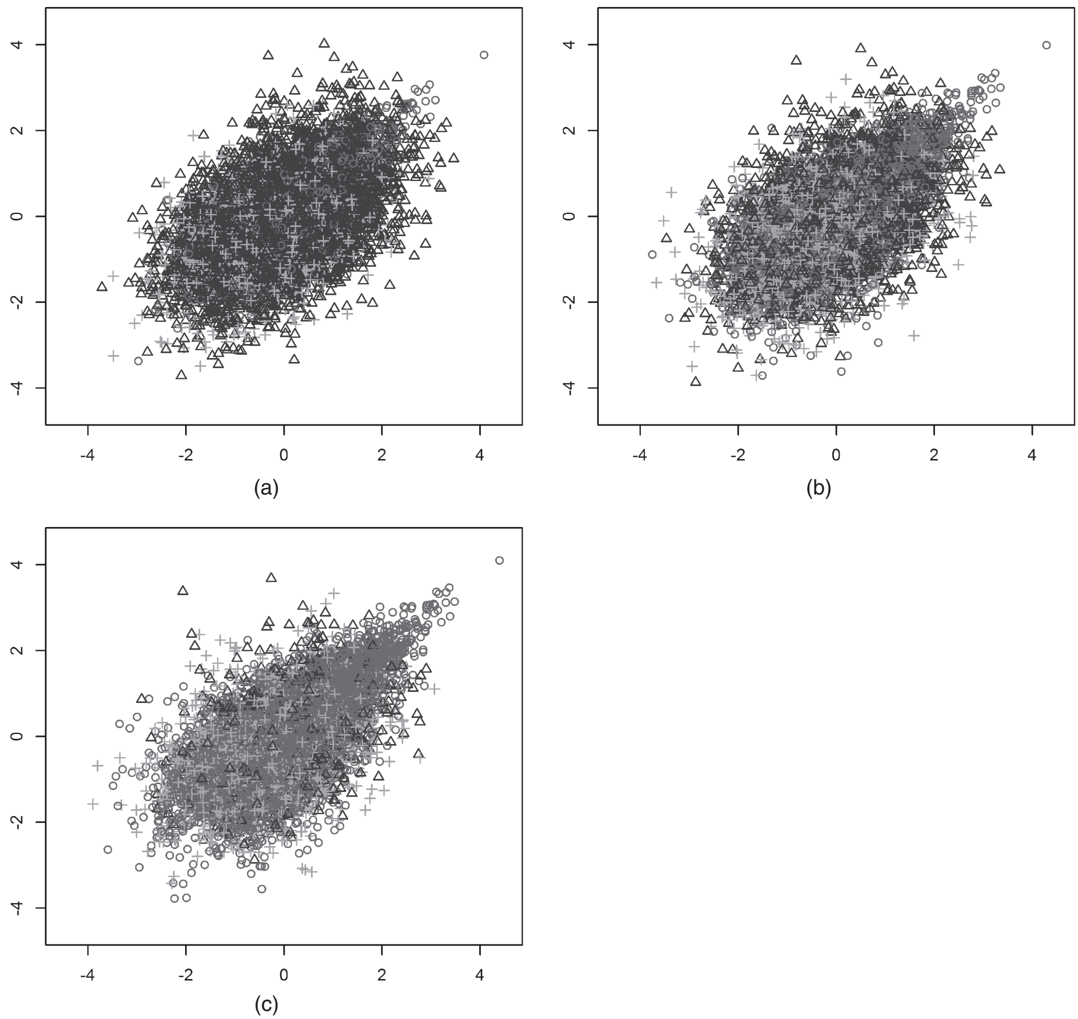

Fig. 1. Scatter plots of simulated two-dimensional max-copulas under three sets of parameters (c, α, ρ) (the marginals are transformed to the Gaussian scale for better illustration): (a) $(c, \alpha, \rho) = (0.8, 2, 0.5)$; (b) $(c, \alpha, \rho) = (0.5, 2, 0.5)$; (c) $(c, \alpha, \rho) = (0.2, 2, 0.5)$

of the non-extremal area by 'decreasing' the influence of the Gumbel copula. This mechanism takes effect especially in the case of multivariate max-copulas (i.e. $d \geqslant 3$), since multiple points of $\mathbf{U} = (U_1, \ldots, U_d)$ can come from the Gaussian copula; for example $(U_1, \ldots U_p)$ come from the Gaussian copula and $(U_{p+1}, \ldots, U_d)$ come from the Gumbel copula. This is a unique feature of the max-copula that is not shared by the mixture copula. With the unique feature, the max-copula achieves better modelling flexibility for the non-extremal behaviour while attaining good modelling for tail dependence. This feature is further demonstrated through simulation experiments in Section 4.1 and real data applications in Section 5.

2.4. Factor-structured max-copula

The number of parameters of an unstructured Gaussian copula's correlation matrix Σ increases quadratically as the dimension increases, which imposes a huge challenge on the estimation and inference of the model. To bypass this obstacle, we propose two factor-structured max-copulas, which offer flexible multivariate dependence modelling while remaining numerically tractable.

2.4.1. Single-factor max-copula

For low dimensional applications, we design a parsimonious max-copula by imposing a factor structure on the correlation matrix Σ of the Gaussian copula. For the single-factor max-copula, Σ takes a single-factor structure, where $\rho_{ij} = \beta_i \beta_j$ for all $1 \leqslant i, j \leqslant d$ and $|\beta_i| < 1$. For identification purposes, we assume that $\beta_1 > 0$. The Gumbel copula belongs to the one-parameter Archimedean copula family, which offers parsimonious modelling at the cost of assuming a rather restrictive exchangeable dependence structure among variables. Such an assumption is acceptable when the dimension is low, so we keep it unchanged in the single-factor max-copula. Thus, the parameters that are associated with the d-dimensional single-factor max-copula are $(c, \beta_1, \ldots, \beta_d, \alpha)$, which is of length $d + 2$ and is much smaller than $O(d^2)$.

2.4.2. Block factor max-copula

In practice, especially in high dimensional applications, situations where the multivariate observations come from several groups with similar characteristics are not uncommon. For example, in financial applications, stocks may come from different industrial sectors and it is expected that stocks from the same sector have common behaviour and are more closely related. To have a better modelling of such data, we propose the block factor max-copula. Assume that $\mathbf{U}$ consists of p groups and, for each group $i = 1, 2, \ldots, p$, it contains d_i group members. By a slight abuse of notation, we denote $\mathbf{U} = \cup_{i=1}^{p} (U_{i1}, U_{i2}, \ldots, U_{id_i})$, where $\mathbf{U}$ is of dimension $d = \Sigma_{i=1}^{p} d_i$.

For the Gaussian copula, we take advantage of the natural group structure by imposing a block factor structure on the correlation matrix Σ. Specifically, we assume that the Gaussian copula is implied by a multivariate normal distribution, denoted by $\mathbf{Z} = \cup_{i=1}^{p} (Z_{i1}, Z_{i2}, \ldots, Z_{id_i})$, which has the stochastic representation

$$Z_{ij} = \beta_i F_0 + \gamma_i F_i + \varepsilon_{ij},$$

where $i = 1, \ldots, p$, $j = 1, \ldots, d_i$, F_0 is the common factor across different groups, F_is are group-specific factors and ε_{ij}s are subject level noise. Also, all random variables are mutually independent and standard normal. Here F_0 introduces correlations across groups whereas F_is are responsible for group-specific correlations. The block factor structure requires $2p$ parameters instead of $O(d^2)$, which provides a much more parsimonious model.

We impose the group structure on the Gumbel copula by using the theory of hierarchical Archimedean copulas. Intuitively, hierarchical Archimedean copulas can be thought of as a block factor Archimedean copula. Here we extend the Gumbel copula to the one-level hierarchical Gumbel copula, which offers different within-group dependence and common between-group dependence. More formally, $\mathbf{U} = \cup_{i=1}^{p} (U_{i1}, U_{i2}, \ldots, U_{id_i})$ is said to follow a one-level hierarchical Gumbel copula if its cumulative distribution function can be written as

$$C(\mathbf{u}) = \psi_0 \left[\sum_{i=1}^{p} \psi_0^{-1} \{ C_i(u_{i1}, \ldots, u_{id_i}) \} \right],$$

where $\psi_0 = \exp(-x^{1/\alpha_0})$ is the Gumbel copula generator with parameter α_0, $C_i(\cdot)$s are Gumbel copulas with $C_i(u_{i1}, \ldots, u_{id_i}) = \psi_i \{ \Sigma_{j=1}^{d_i} \psi_i^{-1}(u_{ij}) \}$ and $\psi_i = \exp(-x^{1/\alpha_i})$s are the Gumbel copula generators with parameter α_is. For $C(\mathbf{u})$ to be a valid copula, we require $\alpha_0 \leqslant \min_{1 \leqslant i \leqslant p} \alpha_i$, i.e. the between-group dependence is weaker than the within-group dependence. Under the one-level hierarchical Gumbel copula, the within-group dependence parameter is α_i for the ith group and the between-group dependence parameter is α_0 for all different groups. For more details of hierarchical Archimedean copulas we refer readers to Joe (2014).

On the basis of the block factor Gaussian copula $\mathbf{C}_1$ and the hierarchical Gumbel copula $\mathbf{C}_2$, we now specify the block factor max-copula. Assume that $\mathbf{U}_1 = \cup_{i=1}^{p}(U_{i1,1}, U_{i2,1}, \ldots, U_{id_i,1})$ follows a block factor Gaussian copula $\mathbf{C}_1$ and $\mathbf{U}_2 = \cup_{i=1}^{p}(U_{i1,2}, U_{i2,2}, \ldots, U_{id_i,2})$ follows a one-level hierarchical Gumbel copula $\mathbf{C}_2$. We call the copula of

$$\mathbf{U} = \max(\mathbf{U}_1^{1/c}, \mathbf{U}_2^{1/(1-c)})$$

a block factor max-copula. The parameters that are associated with the d-dimensional multivariate copula are $(c, \beta_1, \ldots, \beta_p, \gamma_1, \ldots, \gamma_p, \alpha_0, \ldots, \alpha_p)$, which is of length $3p+2$, where p is the number of groups of $\mathbf{U}$. For variables within the same group i, i.e. $(U_{i1}, U_{i2}, \ldots, U_{id_i})$, the block factor max-copula reduces to a max-copula based on a Gaussian copula with exchangeable correlation $\rho = (\beta_i^2 + \gamma_i^2)/(1 + \beta_i^2 + \gamma_i^2)$ and a Gumbel copula with parameter α_i. For two variables from different groups i and j, e.g. (U_{i1}, U_{j1}), the block factor max-copula reduces to a bivariate max-copula based on a Gaussian copula with correlation

$$\rho = \frac{\beta_i \beta_j}{\surd(1 + \beta_i^2 + \gamma_i^2)\surd(1 + \beta_j^2 + \gamma_j^2)}$$

and a Gumbel copula with parameter α_0.

For parsimony, we use the same weight c for all p groups. Additional flexibility can be obtained by imposing group-specific weights. Here we use the factor Gaussian copula and the hierarchical Gumbel copula as component copulas for the block factor max-copula; the generalization to the factor elliptical copula and the hierarchical Archimedean copula follows readily.

2.5. Semiparametric dynamic max-copula model

The ultimate purpose of the max-copula is to model the joint behaviour of multivariate time series. In reality, it is almost impossible to observe an independent and identically distributed sequence of multivariate time series. Furthermore, the marginals by no means can behave like uniform random variables on [0,1]. To tackle these two problems, we follow the procedure in Chen and Fan (2006a) and propose a semiparametric dynamic max-copula (SDM) model. Let $\{\mathbf{Y}_t\}_{t=1}^{T}$ be a multivariate time series where $\mathbf{Y}_t$ is of dimension d and let $\mathcal{F}_{t-1}$ denote the information set at time $t-1$, i.e. the σ-field generated by $\{\mathbf{Y}_{t-1}, \mathbf{Y}_{t-2}, \ldots\}$. We specify the SDM model as follows.

2.5.1. *Parametric dynamic component*

We assume that the dynamics of multivariate time series happen in first- and second-order conditional moments and we further assume that the conditional mean and variance can be correctly parameterized up to a finite dimensional unknown parameter λ_0, i.e.

$$\mathbf{Y}_t = \mu_t(\lambda_0) + \boldsymbol{\eta}_t \surd H_t(\lambda_0),$$

where $\mu_t(\lambda_0) = (\mu_{t,1}(\lambda_0), \ldots, \mu_{t,d}(\lambda_0))' = \mathbb{E}(\mathbf{Y}_t|\mathcal{F}_{t-1})$ and $H_t(\lambda_0) = \operatorname{diag}\{h_{t,1}(\lambda_0), \ldots, h_{t,d}(\lambda_0)\}$, in which $h_{t,j}(\lambda_0) = \mathbb{E}\{(Y_{t,j} - \mu_{t,j}(\lambda_0))\}^2|\mathcal{F}_{t-1}]$, $j = 1, \ldots, d$.

2.5.2. *Non-parametric marginal component*

We do not impose any parametric assumption on the marginals of $\boldsymbol{\eta}_t = (\eta_{t1}, \ldots, \eta_{td})'$. Instead, we assume only that each marginal of η_{tj} is continuous and denote the true marginal distribution function as F_j^0.

2.5.3. Parametric max-copula component

According to Sklar (1959), there is a unique copula $\mathbf{C}^0$ such that $F^0(\boldsymbol{\eta}) = \mathbf{C}^0\{F_1^0(\eta_1), \ldots, F_d^0(\eta_d)\}$, where F^0 is the true joint distribution of $\boldsymbol{\eta}_t$. In an SDM model, we assume that $\mathbf{C}^0$ is a max-copula based on component copulas $\mathbf{C}_1$ and $\mathbf{C}_2$ and denote θ_0 as true parameters for the max-copula $\mathbf{C}^0$.

In what follows, $\mathbf{C}^0$ is taken to be a single-factor max-copula or a block factor max-copula depending on different applications, and we call the corresponding SDM model a single-factor SDM model or a block factor SDM model respectively.

3. Semiparametric composite maximum likelihood estimation

To estimate the SDM model fully, we need to estimate all three components of it. We largely follow the semiparametric copula estimation framework in Genest *et al.* (1995) and Chen and Fan (2006a). The multivariate time series that we observe are $\{\mathbf{Y}_t\}_{t=1}^T$ and the ultimate goal is to estimate the parameter θ_0 of the max-copula $\mathbf{C}^0$.

3.1. Estimation of the dynamic and marginal components

The parametric dynamic component and the non-parametric marginal component are treated as nuisance parameters and the estimation procedure is standard. After the parametric assumption on the dynamic component has been fixed, standard maximum likelihood estimation can be employed to estimate λ^0 by using the observations $\{\mathbf{Y}_t\}_{t=1}^T$. On the basis of estimated $\hat{\lambda}^0$, we can construct the fitted errors $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^T$. An empirical distribution function $\hat{F}_{Tj}$ is employed to estimate the non-parametric F_j^0, for $j = 1, \ldots, d$, where

$$\hat{F}_{Tj}(\cdot) = \frac{1}{T+1} \sum_{t=1}^{T} I(\hat{\eta}_{tj} \leqslant \cdot).$$

3.2. Estimation of the max-copula component

Using estimated $\hat{F}_T = (\hat{F}_{T1}, \ldots, \hat{F}_{Td})$, we can turn the fitted errors $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^{\mathrm{T}}$ into the transformed errors $\{\hat{\mathbf{U}}_t = \hat{F}_T(\hat{\boldsymbol{\eta}}_t)\}_{t=1}^T$, which can be seen as an approximation for the unobserved true copula processes $\{\mathbf{U}_t\}_{t=1}^T$ that drive $\{\mathbf{Y}_t\}_{t=1}^T$.

Genest *et al.* (1995) and Chen and Fan (2006a) used standard maximum likelihood estimation on $\{\hat{\mathbf{U}}_t\}_{t=1}^T$ to estimate the copula parameter θ_0. In our SDM model, because of the nature of max-copulas, the full log-likelihood function is analytically complicated and may also introduce numerical difficulty during optimization, especially when the dimension of the multivariate time series is high. To bypass the problem, we employ the CMLE method. To our best knowledge, this is the first time that CMLE has been applied to the semiparametric dynamic modelling of time series.

Consider a d-dimensional random vector $\mathbf{U} = (U_1, \ldots, U_d)$, with probability density function $f(\mathbf{u}; \theta)$ for some unknown parameter vector θ. Denote by $\mathcal{A} = \{\mathcal{A}_1, \ldots, \mathcal{A}_K\}$ a set of marginal events with associated likelihoods $L_k(\mathbf{u}; \theta) \propto f(\mathbf{u} \in \mathcal{A}_k; \theta)$. Following Lindsay (1988), a composite log-likelihood is the weighted sum $\mathrm{CL}(\mathbf{u}; \theta) = \Sigma_{k=1}^K w_k \log\{L_k(\mathbf{u}; \theta)\}$, where w_ks are non-negative weights to be chosen. For more details of CMLE, see Varin *et al.* (2011).

In this paper, we set $\mathcal{A}$ to be all pairwise combinations between elements of $\mathbf{U} = (U_1, \ldots, U_d)$, so $K = d(d-1)/2$, $\mathcal{A}_k = \{i, j\}$ for some $1 \leqslant i < j \leqslant d$ and $L_k(\mathbf{u}; \theta) = f_{ij}(u_i, u_j; \theta)$, where $f_{ij}(\cdot)$ denotes the pairwise probability density function between U_i and U_j. Also, we set $w_k = 1$ for all $k=1,\ldots,K$. As a result, the composite log-likelihood function based on one observation $\mathbf{u}=(u_1,\ldots,u_d)$ takes the form

$$\mathrm{CL}(\mathbf{u};\theta)=\sum_{k=1}^{K}\log\{L_k(\mathbf{u};\theta)\}=\sum_{i=1}^{d-1}\sum_{j=i+1}^{d}\log\{f_{ij}(u_i,u_j;\theta)\}.$$

3.2.1. *Pairwise likelihood function*

Suppose that $\mathbf{U}=(U_1,\ldots,U_d)$ follows the max-copula $\mathbf{C}(\mathbf{u};\theta)$. To employ CMLE, we first need to derive the pairwise copula density for each pair of (U_i,U_j). In what follows, we set $i=1$ and $j=2$, and others can be derived similarly. To be generic, in what follows, we assume that the correlation parameter of the Gaussian copula is ρ and the parameter of the Gumbel copula is α. Later, ρ and α can be replaced by the parameters in the single-factor or block factor max-copula. By construction, the pairwise CDF of (U_1,U_2) can be written as

$$\begin{aligned}C_{12}(u_1,u_2)&=C_{\mathrm{Gaussian}}(u_1^c,u_2^c,\rho)\,C_{\mathrm{Gumbel}}(u_1^{1-c},u_2^{1-c},\alpha)\\&=\Phi\{\Phi^{-1}(u_1^c),\Phi^{-1}(u_2^c),\rho\}\;\exp\left(-\left[\sum_{i=1}^{2}\{-\log(u_i^{1-c})\}^{\alpha}\right]^{1/\alpha}\right)\\&=\Phi\{\Phi^{-1}(u_1^c),\Phi^{-1}(u_2^c),\rho\}\;\exp\left(-(1-c)\left[\sum_{i=1}^{2}\{-\log(u_i)\}^{\alpha}\right]^{1/\alpha}\right).\end{aligned}$$

The pairwise probability density function can be obtained by taking the partial derivatives with respect to u_1 and u_2. The detailed expression can be found in section 2 of the on-line supplementary material.

3.2.2. *Composite likelihood method*

Suppose that we have a d-dimensional single-factor max-copula $\mathbf{C}$ with true parameter $\theta_0=(c_0,\beta_1^0,\beta_2^0,\ldots,\beta_d^0,\alpha_0)$. Following Section 3.2.1, the composite log-likelihood function based on the transformed errors $\{\hat{\mathbf{U}}_t\}_{t=1}^{T}$ can be written as

$$\mathrm{CL}_T(\theta)=\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{d-1}\sum_{j=i+1}^{d}l_{t,ij}(\hat{U}_{ti},\hat{U}_{tj};\theta)=\frac{1}{T}\sum_{t=1}^{T}l_t(\hat{\mathbf{U}}_t;\theta),$$

where $l_{t,ij}(\hat{U}_{ti},\hat{U}_{tj};\theta)=\log\{f_{ij}(\hat{U}_{ti},\hat{U}_{tj};\theta)\}=\log\{f_{ij}(\hat{U}_{ti},\hat{U}_{tj};c,\beta_i,\beta_j,\alpha)\}$ is the pairwise log-likelihood function between U_i and U_j based on the tth transformed error $\hat{\mathbf{U}}_t=(\hat{U}_{t1},\ldots,\hat{U}_{td})$ and $l_t(\hat{\mathbf{U}}_t;\theta)=\Sigma_{i=1}^{d-1}\Sigma_{j=i+1}^{d}l_{t,ij}(\hat{U}_{ti},\hat{U}_{tj};\theta)$. Denote $\hat{\theta}_T$ as the maximizer of the composite likelihood function $\mathrm{CL}_T(\theta)$ and call it the CMLE of θ.

For the block factor max-copula, we can further separate the composite likelihood function into the within-group likelihood and between-group likelihood. Suppose that there are p groups, each with group size d_i, and denote the tth transformed error as $\hat{\mathbf{U}}_t=\cup_{i=1}^{p}(\hat{U}_{t,i1},\ldots,\hat{U}_{t,id_i})$. The composite log-likelihood function can be written as

$$\mathrm{CL}_T(\theta)=\frac{1}{T}\sum_{t=1}^{T}\{\mathrm{CLWG}_t(\theta)+\mathrm{CLBG}_t(\theta)\},$$

where $\mathrm{CLWG}_t(\theta)$ stands for the within-group composite likelihood and $\mathrm{CLBG}_t(\theta)$ stands for the between-group composite likelihood. Following the above formulation, we have

$$\mathrm{CLWG}_t(\theta)=\sum_{k=1}^{p}\sum_{i=1}^{d_k-1}\sum_{j=i+1}^{d_k} l_{t,(ki,kj)}(\hat{U}_{t,ki},\hat{U}_{t,kj};\theta),$$

$$\mathrm{CLBG}_t(\theta)=\sum_{k=1}^{p-1}\sum_{g=k+1}^{p}\sum_{i=1}^{d_k}\sum_{j=1}^{d_g} l_{t,(ki,gj)}(\hat{U}_{t,ki},\hat{U}_{t,gj};\theta),$$

where $l_{t,(ki,kj)}(\hat{U}_{t,ki},\hat{U}_{t,kj};\theta)=\log\{f_{ki,kj}(\hat{U}_{t,ki},\hat{U}_{t,kj};\theta)\}=\log\{f_{ki,kj}(\hat{U}_{t,ki},\hat{U}_{t,kj};c,\beta_k,\gamma_k,\alpha_k)\}$ is the pairwise log-likelihood function between U_{ki} and U_{kj} that both come from the kth group, and $l_{t,(ki,gj)}(\hat{U}_{t,ki},\hat{U}_{t,gj};\theta)=\log\{f_{ki,gj}(\hat{U}_{t,ki},\hat{U}_{t,gj};\theta)\}=\log\{f_{ki,gj}(\hat{U}_{t,ki},\hat{U}_{t,gj};c,\beta_k,\gamma_k,\beta_g,\gamma_g,\alpha_0)\}$ is the pairwise log-likelihood function between U_{ki} and U_{gj} that come from the kth group and gth group respectively.

If we reindex $\hat{\mathbf{U}}_t$ as $\hat{\mathbf{U}}_t=(\hat{U}_{t1},\hat{U}_{t2},\ldots,\hat{U}_{td})$ with $d=\Sigma_{i=1}^{p}d_i$, it is easy to verify that we have

$$\mathrm{CL}_T(\theta)=\frac{1}{T}\sum_{t=1}^{T}\{\mathrm{CLWG}_t(\theta)+\mathrm{CLBG}_t(\theta)\}=\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{d-1}\sum_{j=i+1}^{d} l_{t,ij}(\hat{U}_{ti},\hat{U}_{tj};\theta)=\frac{1}{T}\sum_{t=1}^{T} l_t(\hat{\mathbf{U}}_t;\theta),$$

where $l_t(\hat{\mathbf{U}}_t;\theta)=\Sigma_{i=1}^{d-1}\Sigma_{j=i+1}^{d} l_{t,ij}(\hat{U}_{ti},\hat{U}_{tj};\theta)$. By reindexing, we unify the notation of $\mathrm{CL}_T(\theta)$ for the single-factor and block factor max-copulas.

As mentioned before, we set $w_k=1$ for all pairwise likelihoods. It is known that all weights lead to consistent estimation whereas some weights are more efficient. For the block factor max-copula, since pairwise likelihoods come from various sources (i.e. within group and between group), it is natural to set different weights for likelihoods from different sources, which may help to improve the estimation efficiency. However, since efficiency is not the main focus here and the simulation studies show satisfactory performance of the current weights, we leave the investigation of finding efficient weights as a future research question.

3.3. Asymptotic theory

Genest *et al.* (1995) and Chen and Fan (2006a,b), showed that, under the semiparametric setting, standard maximum likelihood estimation based on the transformed errors $\{\hat{\mathbf{U}}_t\}_{t=1}^{T}$ is consistent and asymptotically normal. It is also well known that with the classical independently and identically distributed setting, under some regularity conditions, both maximum likelihood estimation and CMLE are consistent and asymptotically normal, whereas CMLE is less efficient in terms of asymptotic covariance. On the basis of the above two observations, it is not surprising that CMLE is also consistent and asymptotically normal under the SDM model, as stated in theorems 2 and 3.

Theorem 2. Suppose that the observations $\{\mathbf{Y}_t\}_{t=1}^{T}$ and the composite likelihood function $l_t(\cdot)$ satisfy assumptions C and D in Chen and Fan (2006a); we have $\hat{\theta}_T\to_{\mathrm{p}}\theta_0$ as $T\to\infty$, i.e. $\hat{\theta}_T$ is consistent.

Before we state theorem 3, we first introduce some notation. Denote $l(u_1,\ldots,u_d;\theta)=\Sigma_{i=1}^{d-1}\Sigma_{j=i+1}^{d} l_{ij}(u_i,u_j;\theta)=\Sigma_{i=1}^{d-1}\Sigma_{j=i+1}^{d}\log\{f_{ij}(u_i,u_j;\theta)\}$, $l_\theta(u_1,\ldots,u_d;\theta)=\partial l(u_1,\ldots,u_d;\theta)/\partial\theta$, $l_j(u_1,\ldots,u_d;\theta)=\partial l(u_1,\ldots,u_d;\theta)/\partial u_j$, $l_{\theta\theta}(u_1,\ldots,u_d;\theta)=\partial^2 l(u_1,\ldots,u_d;\theta)/\partial\theta\,\partial\theta'$ and $l_{\theta j}(u_1,\ldots,u_d;\theta)=\partial^2 l(u_1,\ldots,u_d;\theta)/\partial u_j\,\partial\theta$, for $j=1,\ldots,d$. Also, denote $\{\mathbf{U}_t=(U_{t1},\ldots,U_{td})\}_{t=1}^{T}$ as the unobserved true copula processes that drive the observations $\{\mathbf{Y}_t\}_{t=1}^{T}$. Further denote

$$A_T^0=\frac{1}{T}\sum_{t=1}^{T}\{l_\theta(U_{t1},U_{t2},\ldots,U_{td};\theta_0)+\sum_{j=1}^{d}Q(U_{tj};\theta_0)\},$$

where $Q(U_{tj};\theta)\equiv E_{\theta_0}[l_{\theta j}(\mathbf{U}_s;\theta_0)\{I(U_{tj}\leqslant U_{sj})-U_{sj}\}|U_{tj}]$, $s\neq t$. Denote $B=-\mathbb{E}_{\theta_0}[l_{\theta\theta}(\mathbf{U}_t;\theta_0)]$ and $\Sigma=\mathrm{var}_{\theta_0}\{l_\theta(\mathbf{U}_t;\theta_0)+\Sigma_{j=1}^{d}Q(U_{tj};\theta_0)\}$.

Theorem 3. Let $\theta_0 \in \mathrm{int}(\Theta)$. Under assumptions D and N in Chen and Fan (2006a), we have

(a) $\hat{\theta}_T - \theta_0 = B^{-1} A_T^0 + o_p(T^{-1/2})$ and
(b) $(\hat{\theta}_T - \theta_0)\sqrt{T} \to_{\mathrm{d}} N(0, B^{-1}\Sigma B^{-1})$, i.e. $\hat{\theta}_T$ is asymptotically normal.

As mentioned in Chen and Fan (2006a), the additional term $Q(U_{tj};\theta_0)$ in A_T^0 is introduced by the estimation of the marginal distribution functions $F_j^0(\cdot)$, $j=1,\ldots,d$, and, if $F_j^0(\cdot)$ is completely known, $Q(U_{tj};\theta_0)$ will disappear. We note that CMLE implicitly imposes an incorrect working independence assumption, which may make it less efficient in terms of asymptotic covariance. However, the benefits of CMLE are that it bypasses the analytical difficulty of the full likelihood function and lowers the computational complexity and instability of the estimation procedures, especially in high dimension.

The classical plug-in estimator for the asymptotic covariance matrix is available. However, because of analytical complexity, a parametric bootstrap procedure is proposed to estimate the asymptotic variance for CMLE, which is consistent and gives good finite sample performance, as is shown in section 3.2 of the on-line supplementary material.

4. Simulation performance

4.1. Comparison between the max-copula and the mixture copula

In this section, we demonstrate the advantage of max-copulas over mixture copulas in the modelling of non-extremal behaviour. Specifically, we use a bivariate Gaussian copula as a 'surrogate' for the non-extremal joint behaviour of random vector (U_1, U_2). We compare the Kullback–Leibler (KL) distance of the max-copula and the mixture copula with the Gaussian copula, while keeping the upper tail dependence coefficients of the max-copula and the mixture copula the same. (All the KL distances that are calculated in this section are based on Monte Carlo integration.) We conduct experiments for two scenarios where the component copulas are either a Gaussian plus Gumbel copula or a Gaussian plus t copula. We demonstrate the simulation procedure for the Gaussian plus Gumbel copula in detail and report the results for both scenarios. Denote $\mathbf{C}_1(\rho)$ as a Gaussian copula and denote $\mathbf{C}_2(\alpha)$ as a Gumbel copula. Denote the max-copula based on $(c, \mathbf{C}_1(\rho), \mathbf{C}_2(\alpha))$ as $\mathbf{C}_{\max}(c,\rho,\alpha)$ and the corresponding mixture copula as $\mathbf{C}_{\mathrm{mix}}(c,\rho,\alpha)$. By theorem 1, $\mathbf{C}_{\max}(c,\rho,\alpha)$ and $\mathbf{C}_{\mathrm{mix}}(c,\rho,\alpha)$ have the same upper tail dependence coefficient. We conduct experiments for two different cases of settings.

In case 1, we calculate the KL distance of the max-copula $\mathbf{C}_{\max}(c,\rho,\alpha)$ and the mixture copula $\mathbf{C}_{\mathrm{mix}}(c,\rho,\alpha)$ to its component Gaussian copula $\mathbf{C}_1(\rho)$. We set $\rho=0.5$ and $\alpha=2$ since they are the typical parameters that are obtained in real data applications in Section 5, and we change c from 0.1 to 0.9. The result is summarized in Fig. 2(a). As can be seen, the max-copula is always closer to the component Gaussian copula $\mathbf{C}_1(\rho)$ in terms of KL distance. It confirms the observation in Fig. 1 that the unique pairwise max-rule of the max-copula helps to decrease the influence of the Gumbel copula in the non-extremal region and thus helps to offer better modelling of the non-extremal joint behaviour.

In case 2, we first set a target bivariate Gaussian copula $\mathbf{C}_{\mathrm{Gau}}(\rho_0)$. We then seek the max-copula $\mathbf{C}_{\max}(c,\rho,\alpha)$ and the mixture copula $\mathbf{C}_{\mathrm{mix}}(c,\rho,\alpha)$ that minimize their KL distance to the target $\mathbf{C}_{\mathrm{Gau}}(\rho_0)$ and also attain a given upper tail dependence coefficient λ^{U}. This is a constrained optimization problem where the parameters are (c,ρ,α) and the constraint is $(1-c)(2-2^{1/\alpha}) = \lambda^{\mathrm{U}}$. We fix $\lambda^{\mathrm{U}}=0.24$, the same as the estimated value in Section 5.1. We change the target Gaussian copula from $\rho_0=0.1$ to $\rho_0=0.9$ and report the minimized KL distance attained by the max-copula and the mixture copula in Fig. 2(b). The max-copula attains better ability

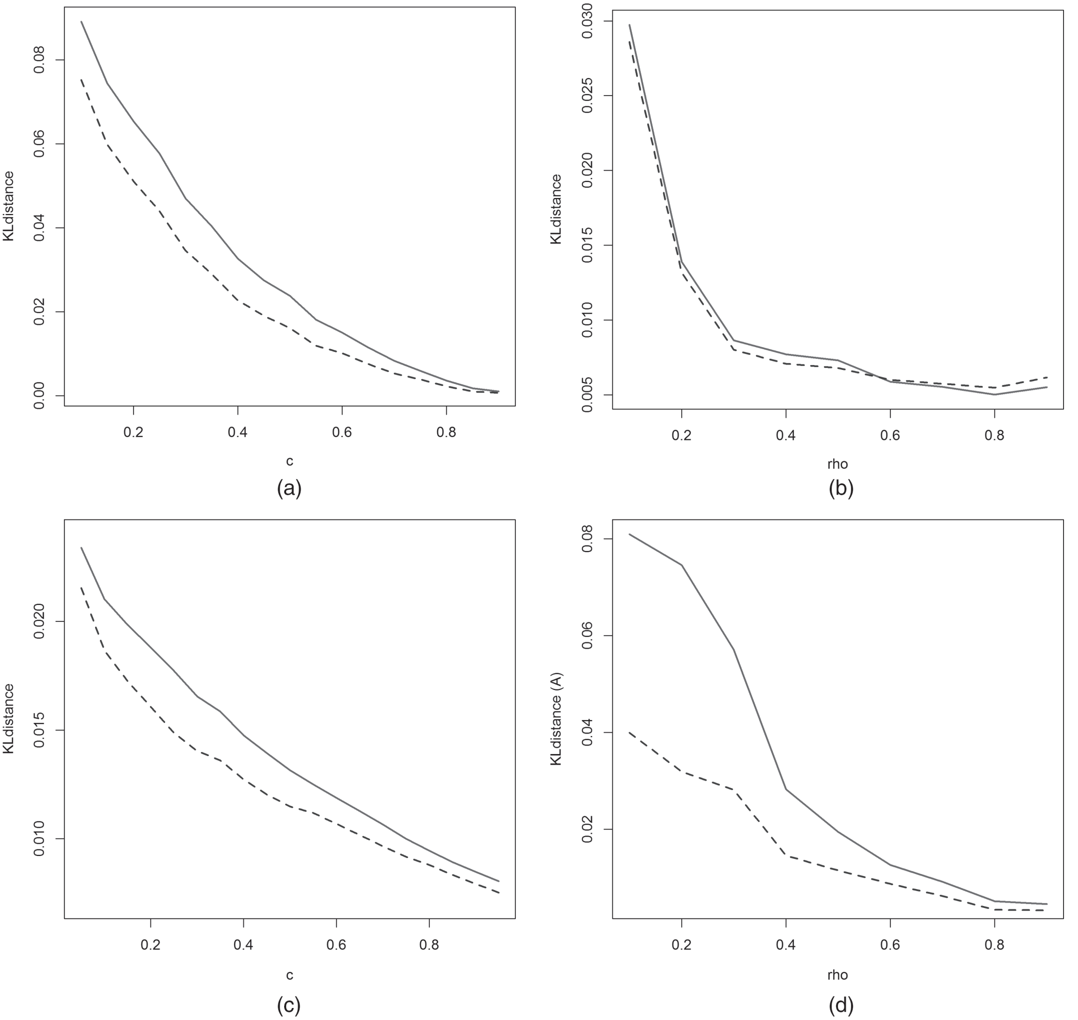

Fig. 2. KL distances of the max-copula and the mixture copula based on either (a), (b) a Gaussian copula plus Gumbel copula or (c), (d) a Gaussian copula plus t-copula to the prespecified Gaussian copulas (——, mixture copula; – – –, max-copula): (a) case 1, KL distance with respect to a component Gaussian copula; (b) case 2, minimized KL distance with respect to a target Gaussian copula; (c) case 1, KL distance with respect to a component Gaussian copula; (d) case 2, minimized KL distance with respect to a target Gaussian copula

to approximate the Gaussian copula in the range $0 < \rho_0 < 0.6$, which is the most commonly encountered range of Gaussian copulas in real data applications in Section 5. It indicates that, for Gaussian copulas with low-to-medium level correlation, the max-copula has a stronger approximation ability than the mixture copula, while attaining the same level of tail dependence.

We have also conducted experiments where $\mathbf{C}_1(\rho_1)$ is a Gaussian copula and $\mathbf{C}_2(\rho_2, \nu)$ is a t-copula. We fix $\nu = 4$ for the t-copula. For case 1, we set $\rho_1 = \rho_2 = 0.5$ and change c from 0.1 to 0.9. For case 2, we use the same target Gaussian copula $\mathbf{C}_{\text{Gau}}(\rho_0)$ and the same λ_{U}. Similar results are observed in Figs 2(c) and 2(d), which confirm the advantage of the max-copula in achieving good non-extremal behaviour modelling while attaining the desired tail dependence.

4.2. Performance of composite maximum likelihood estimation

In this section, we examine the performance of CMLE under single-factor and block factor SDM models. The data-generating process is as follows:

$$\begin{aligned} Y_{ti} &= \phi_0 + \phi_1 Y_{t-1,i} + \sigma_{ti}\eta_{ti}, \qquad t=1,2,\ldots,T, \quad i=1,\ldots,d, \\ \sigma_{ti}^2 &= \omega + \beta\sigma_{t-1,i}^2 + \alpha\sigma_{t-1,i}^2\eta_{t-1,i}^2, \\ \boldsymbol{\eta}_t &\equiv (\eta_{t1},\ldots,\eta_{td})' \overset{\text{IID}}{\sim} \mathbf{F}_\eta = \mathbf{C}(\Phi,\Phi,\ldots,\Phi), \end{aligned}$$

where Φ is the standard normal distribution function and $\mathbf{C}$ is the max-copula. Here we assume an AR(1)–generalized auto-regressive conditional heteroscedasticity GARCH(1,1) structure for the dynamic component of $\mathbf{Y}_t$. We set the parameters to be $(\phi_0,\phi_1,\omega,\beta,\alpha)=(0.01,0.05,0.05,0.85,0.10)$, which according to Oh and Patton (2013) broadly match the values of estimation from real world financial data.

4.2.1 Single-factor max-copula result

We conduct experiments on single-factor SDM models with $d=4$, where we set the true parameters $(c,\alpha,\beta_1,\beta_2,\beta_3,\beta_4)=(0.5,2,0.2,0.4,0.6,0.8)$. Under this setting, the Gaussian copula has correlations of 0.08, 0.12, 0.16, 0.24, 0.32, 0.48, and the Gumbel copula has an upper tail dependence coefficient of 0.59.

We simulate 500 data sets of sample size $T=(1000,2000,5000)$ and report the sample mean and sample standard deviation based on the 500 estimators. The result is summarized in Table 1. As can be seen, under the single-factor SDM model, CMLE is consistent, where both bias and variance of the estimators grow smaller as sample size T grows bigger.

4.2.2. Block factor max-copula result

In this section, we conduct numerical investigations on block factor SDM models. We assume $d=20$ and $p=4$, i.e. the observations are of dimension 20 with five subjects in each of the four groups. We set $c=0.5$, $(\beta_1,\ldots,\beta_4)=(1,1,1.2,1.2)$ and $(\gamma_1,\ldots,\gamma_4)=(0.8,0.8,1,1)$ for the block factor Gaussian copula, and set $(\alpha_0,\alpha_1,\ldots,\alpha_4)=(1.5,1.75,1.75,2,2)$ for the one-level hierarchical Gumbel copula. Under the current setting, the Gaussian copula has within-group correlations of 0.62 and 0.71 and between-group correlations of 0.38, 0.40 and 0.42; the Gumbel copula has within-group upper tail dependence coefficients of 0.51 and 0.59 and a common between-group upper tail dependence coefficient of 0.41.

We simulate 500 data sets of sample size $T=(1000,2000,5000)$ and report the sample mean and sample standard deviation based on the 500 estimators. The result is summarized in Table 2. As can be seen, CMLE is consistent where both bias and variance of the estimators become smaller as the sample size T grows bigger.

Table 1. Semiparametric CMLE for data generated by the single-factor SDM model with $d=4$ and $(c,\alpha,\beta_1,\beta_2,\beta_3,\beta_4)=(0.5,2,0.2,0.4,0.6,0.8)$†

T	c	α	β_1	β_2	β_3	β_4
1000	0.503 (0.080)	2.010 (0.148)	0.195 (0.105)	0.407 (0.072)	0.596 (0.075)	0.822 (0.087)
2000	0.494 (0.057)	1.992 (0.105)	0.193 (0.067)	0.403 (0.049)	0.601 (0.053)	0.813 (0.076)
5000	0.497 (0.037)	1.991 (0.075)	0.197 (0.043)	0.402 (0.033)	0.602 (0.031)	0.807 (0.040)

†The standard deviations of the estimators are in parentheses.

Table 2. Semiparametric CMLE for data generated by the block factor SDM model with $d=20, p=4$ and $(c, \beta_1, \beta_2, \beta_3, \beta_4, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4) = (0.5, 1, 1, 1.2, 1.2, 0.8, 0.8, 1, 1, 1.5, 1.75, 1.75, 2, 2)$†

T	c	β_1	β_2	β_3	β_4
1000	0.501 (0.092)	1.020 (0.110)	1.031 (0.100)	1.231 (0.122)	1.236 (0.125)
2000	0.492 (0.064)	1.018 (0.078)	1.008 (0.070)	1.216 (0.101)	1.214 (0.073)
5000	0.499 (0.039)	1.007 (0.042)	1.010 (0.048)	1.203 (0.051)	1.206 (0.050)
	γ_1	γ_2	γ_3	γ_4	α_0
1000	0.840 (0.133)	0.805 (0.159)	1.000 (0.158)	1.027 (0.144)	1.493 (0.084)
2000	0.821 (0.084)	0.827 (0.084)	1.002 (0.088)	1.006 (0.086)	1.500 (0.041)
5000	0.802 (0.056)	0.810 (0.054)	1.003 (0.056)	0.998 (0.045)	1.501 (0.028)
	α_1	α_2	α_3	α_4	
1000	1.742 (0.142)	1.745 (0.147)	2.011 (0.127)	1.992 (0.137)	
2000	1.736 (0.065)	1.737 (0.078)	2.000 (0.071)	1.998 (0.075)	
5000	1.746 (0.046)	1.741 (0.057)	1.999 (0.053)	1.996 (0.047)	

†The standard deviations of the estimators are in parentheses.

We have also conducted a simulation for a much larger block factor SDM model with $d=108$ and $p=9$, i.e. the observations are of dimension 108 with 12 subjects in each of the nine groups. Though there are 29 parameters in total, CMLE still performs well and provides decent accuracy when $T=500$. For brevity the result is provided in section 3.1 of the on-line supplementary material.

5. Real data application

In this section, we give two real data applications of the max-copula. The first is about the single-factor SDM model in the estimation of the conditional VaR for a financial portfolio, and the second is about the block factor SDM model in the construction of optimal portfolios based on 30 component stocks from the Dow Jones industrial average. For comparison, in each section, we fit the data with four different copulas (the max-copula, its component Gaussian copula, its component Gumbel copula and the corresponding mixture copula) and compare their performances.

5.1. Value-at-risk estimation for financial portfolios

Monitoring negative returns of a portfolio is essential in financial risk management. The most common practice is to use VaR, which is a certain extreme quantile (e.g. 0.95, 0.99 or 0.995) of the portfolio's negative return. Since a portfolio usually contains multiple constituents, to obtain an accurate estimation of its VaR, it is essential to have good modelling of the joint behaviour of the portfolio components. For a given portfolio with d-constituents $\mathbf{Y}_t = (Y_{t1}, \ldots, Y_{td})$ and weight $\mathbf{w}_t = (w_{t1}, \ldots, w_{td})$, the conditional daily VaR for day $t+1$ can be defined as a certain conditional quantile of $\Sigma_{i=1}^{d} w_{t+1,i} Y_{t+1,i}$ given the current information $\mathcal{F}_t$. For simplicity, here we consider only the case where the weight $\mathbf{w}_t$ is constant. Following McNeil and Frey (2000),

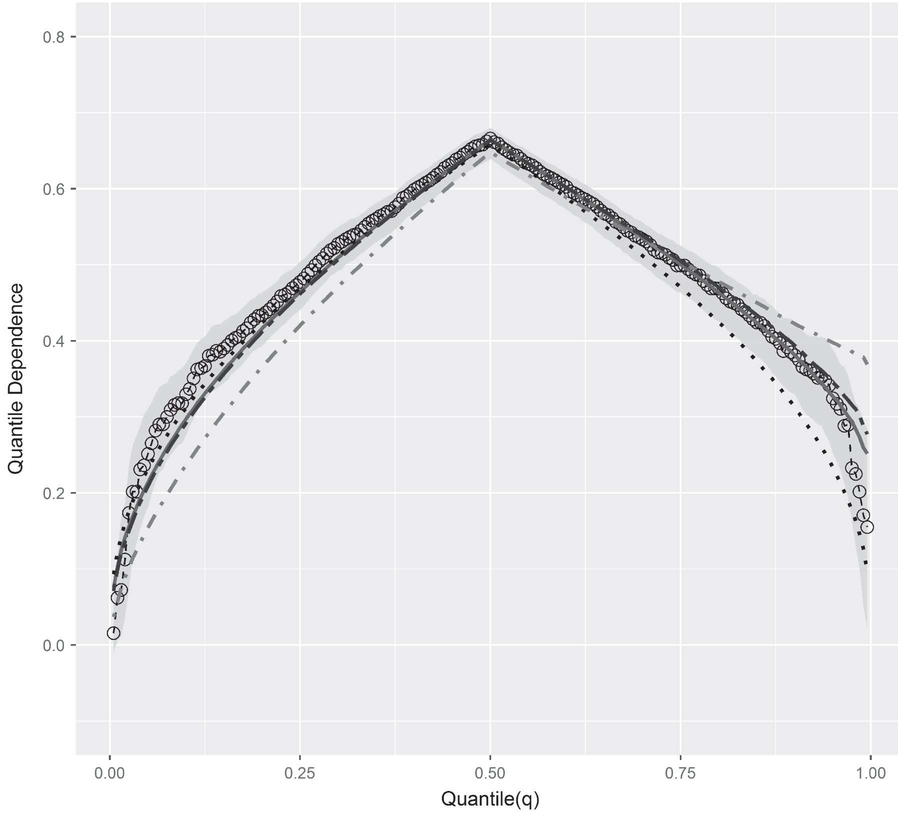

Fig. 3. Average of the sample quantile dependence functions of the three stocks and a bootstrap 95% (pointwise) confidence interval (■) for it: quantile dependence functions based on the estimated copulas max-copula (———), Gaussian copula (· · · · · ·), Gumbel copula (- — — —) and mixture copula (– —— –); ○, data

we propose a semiparametric copula framework for estimating the conditional VaR of a given portfolio based on the SDM model.

We consider the analysis of negative daily stock returns of Citigroup, General Electric and Pfizer, between September 1st, 1995, and August 31st, 2012, which consist of 4295 observations. The parametric dynamic component of each time series $\{Y_{ti}\}_{t=1}^{T}, i = 1, 2, 3,$ is set to be AR(1) + GARCH(1,1). For each pair (i, j) of the three stocks, we calculate the sample quantile dependence $\hat{\lambda}_{ij}^{q}$ based on the transformed errors $\{\hat{\mathbf{U}}_t\}_{t=1}^{4295}$ and plot the average sample dependence $(\hat{\lambda}_{12}^{q} + \hat{\lambda}_{13}^{q} + \hat{\lambda}_{23}^{q})/3$ in Fig. 3. A non-parametric bootstrap on the transformed errors $\{\hat{\mathbf{U}}_t\}_{t=1}^{4295}$ with $B = 1000$ replications is used to construct a pointwise 95% confidence interval for the sample quantile dependence estimates. As is clearly shown, the sample quantile dependence diminishes to 0 as q approaches 0 with 0 in the 95% confidence interval and stays positive as q comes near 1 with 0 outside the 95% confidence interval, which implies an asymmetric tail dependence structure.

For this low dimensional $(d = 3)$ application, we set the copula component of the SDM model to be the single-factor max-copula. We first fit the SDM model by using the whole data set and summarize the result in Table 3. As can be seen, the Gumbel copula plays a substantial role with $1 - \hat{c} = 0.42$. The estimated Gaussian copula has correlations of 0.27, 0.33 and 0.55, whereas

Table 3. Semiparametric CMLE for negative daily stock returns of Citigroup (β_1), General Electric (β_2) and Pfizer (β_3)†

c	α	β_1	β_2	β_3
0.58 (0.07)	1.98 (0.37)	0.67 (0.07)	0.82 (0.06)	0.41 (0.12)

†The bootstrapped standard deviations of estimators are in parentheses.

Table 4. Performance of conditional VaR based on the single-factor SDM model, with VaR-level 0.95, 0.99 and 0.995

Portfolio	*Expected VaR-level*	*Actual VaR-level*	*Expected violation*	*Actual violation*	*p-value*
$w_1 = 0.33$	0.950	0.952	164.75	158	0.63
$w_2 = 0.33$	0.990	0.991	32.95	28	0.43
$w_3 = 0.33$	0.995	0.995	16.48	17	0.81
$w_1 = 0.2$	0.950	0.952	164.75	158	0.63
$w_2 = 0.3$	0.990	0.992	32.95	27	0.34
$w_3 = 0.5$	0.995	0.997	16.48	10	0.14
$w_1 = 0.2$	0.950	0.951	164.75	161	0.81
$w_2 = 0.5$	0.990	0.992	32.95	27	0.34
$w_3 = 0.3$	0.995	0.995	16.48	18	0.71

the estimated Gumbel copula has an upper tail dependence coefficient of 0.58. Together with $\hat{c} = 0.58$, it implies that the three stocks have a strong upper tail dependence with $\hat{\lambda}^{\mathrm{U}} = 0.24$.

To test the performance of the single-factor SDM model in estimating the conditional VaR, we use the back-testing technique in McNeil and Frey (2000). Specifically, on each day t, we use a window of size 1000 (i.e. days $(t-999, \ldots, t)$) to estimate the SDM model. On the basis of the fitted model, we generate $B = 1000$ bootstrap samples $\{\mathbf{Y}_{t+1}^b = (Y_{t+1,1}^b, Y_{t+1,2}^b, Y_{t+1,3}^b)\}_{b=1}^B$ for day $t+1$ and estimate the conditional VaR by using the B bootstrapped portfolio returns $\{\Sigma_{i=1}^3 w_i Y_{t+1,i}^b\}_{b=1}^B$. A violation happens when the actual daily loss is over the qth quantile of the bootstrap sample, where q varies among 0.95, 0.99 and 0.995. Along with the single-factor max-copula, we also use the Gaussian copula, the Gumbel copula and the corresponding mixture copula for model comparison.

In total, we have a test sample of size $4295 - 1000 = 3295$ days. The back-testing is conducted for seven different weights $\mathbf{w}$. The expected and actual number of violations, as well as p-values of the binomial tests for three selected weights, are reported in Table 4 for brevity. More details can be found in section 4.1 of the supplementary material. As can be seen, the estimated VaR performs well, indicating that the single-factor SDM model captures the joint dynamics of the three stocks accurately. For brevity, the results for the other three copulas are not presented. In summary, the Gaussian copula tends to underestimate VaR and achieves 0.949, 0.989 and 0.994 for the mean VaR-level across the seven portfolios; the Gumbel copula tends to overestimate VaR and achieves 0.953, 0.991 and 0.996 for mean VaR-level; the mixture copula achieves a better result with mean VaR-level 0.951, 0.991 and 0.996. The numbers that are achieved by the max-copula are 0.952, 0.990 and 0.995, which are arguably the best among all the copulas, particularly at the extreme quantiles 0.99 and 0.995.

The reason why the max-copula outperforms the others is illustrated more clearly in Fig. 3, which plots the sample quantile dependence function along with those implied by the four fitted copulas. As can be seen, the Gumbel copula overestimates the upper tail dependence and also performs poorly in the non-extremal region. Although the Gaussian copula has a good fit for the non-extremal region, it suffers from an underestimation of the upper tail dependence. Both the mixture copula and the max-copula perform decently with the max-copula having a slightly better fit in the upper quantile and tail area.

We further conduct a formal goodness-of-fit test on all four copulas based on the Cramér–von Mises type test in Genest *et al.* (2009). The estimated p-value is 0.036 for the Gaussian copula, 0.00 for the Gumbel copula, 0.066 for the mixture copula and 0.072 for the max-copula, confirming the advantage of the max-copula.

5.2. Optimal portfolio construction on the Dow Jones industrial average

In this section, we extend the application of the SDM model from financial risk management to optimal portfolio construction as described in Harris and Mazibas (2013). The task here is to establish a weekly optimal portfolio construction framework for the 30 component stocks of Dow Jones industrial average index DJI30. (Compared with daily frequency portfolio optimization, weekly frequency portfolio optimization (i.e. we update the portfolio weight every trading week) is more realistic because of the transaction cost in real trading. Here a week refers to a trading week, which typically contains the five work days in a week.) DJI30 is a major US stock index which consists of 30 large publicly owned companies based in the USA. The 30 companies come from six industrial sectors, according to the standard industrial classification system. Detailed group information can be found in section 4.2 of the on-line supplementary material. It is known that stocks from the same industrial sector have common behaviour and closer relationships. Because of this natural group structure among the 30 stocks, the block factor SDM model is used for this high dimensional ($d = 30$) application. To capture the leverage effect in the conditional volatility of stock returns, we set the parametric dynamic component to be AR(1) plus Glosten–Jagannathan–Runkle–GARCH(1,1). More details about the Glosten–Jagannathan–Runkle–GARCH(1,1) process can be found in Glosten *et al.* (1993).

The initiative of this application is to demonstrate the block factor SDM model's ability to capture both normal and extreme joint movements of multiple stocks, which can be utilized for the construction of more risk rewarding portfolios. To serve this end, we choose the observation period from January 2nd, 2004, to December 30th, 2011, which consists of $T = 2015$ days and roughly 403 trading weeks. It covers the so-called Financial Crisis period from 2008 to 2009 and the post-crisis market rally from 2009 to 2011, which are later used as test samples to examine the performance of the proposed portfolio optimization framework under different market scenarios. Since Visa was not listed until 2008 and Verizon is the only stock in standard industrial classification 4, we remove these two stocks from the candidate pool for portfolio construction and consider the remaining 28 stocks from DJI30. In summary, the 28 stocks can be classified into five groups with group sizes of 8, 9, 3, 5 and 3.

We now describe the weekly optimal portfolio construction framework. Roughly speaking, in the proposed framework, we manage a portfolio that consists of the 28 component stocks in DJI30. At the end of each trading week, say day t, based on the estimated block factor SDM model, we seek to forecast an optimal portfolio weight $\hat{\mathbf{w}}_t = \{\hat{w}_i^t\}_{i=1}^{28}$ for the next trading week of days $(t+1, t+2, \ldots, t+5)$ such that a certain risk measure (e.g. variance, VaR or expected shortfall ES) of the constructed portfolio with the optimal weight $\hat{\mathbf{w}}_t$ is minimal among all possible portfolios. More formally, denote the portfolio weight as $\mathbf{w}=\{w_i\}_{i=1}^{28}$; we are solving the constrained optimization problem

$$\min_{\mathbf{w}} \Phi_t(\mathbf{w}), \qquad \mathbf{w} \geqslant 0, \quad \mathbf{w}'\mathbf{1}=1.$$

where $\Phi_t(\mathbf{w})$ denotes the selected risk measure of the portfolio with weight $\mathbf{w}$, which can be calculated on the basis of the estimated block factor SDM model on day t. The minimizer $\hat{\mathbf{w}}_t$ is the optimal portfolio weight.

In what follows, we set $\Phi_t(\mathbf{w})$ to be the ES of the cumulative portfolio returns in the next trading week of days $(t+1,\ldots,t+5)$. The reason that we choose ES instead of VaR is that ES is a convex function of portfolio weight $\mathbf{w}$, which makes the optimization an easier task. For a random variable Z, its $(1-\alpha)$-level ES is defined to be $\mathrm{ES}_{1-\alpha}=-(1/\alpha)[E(Z\mathbf{1}_{\{Z\leqslant z_\alpha\}})+z_\alpha\{\alpha-P(Z\leqslant z_\alpha)\}]$, where z_α denotes the α-quantile of Z. For more information on ES, readers are referred to McNeil *et al.* (2005). Conditioned on the estimated SDM model on day t, $\Phi_t(\mathbf{w})$ can be well approximated by the parametric bootstrap. Using the estimated SDM model, we generate $B=10000$ times bootstrap samples $\{\{\mathbf{Y}_{t+k}^b=(Y_{t+k,1}^b,\ldots,Y_{t+k,28}^b)\}_{k=1}^5\}_{b=1}^B$ for the returns of the 28 stocks on days $(t+1,\ldots,t+5)$ and approximate the ES $\Phi_t(\mathbf{w})$ by its empirical version based on the B bootstrap portfolio returns $\{\Sigma_{k=1}^5\Sigma_{i=1}^{28}w_iY_{t+k,i}^b\}_{b=1}^B$. For the level of ES, we use $1-\alpha=0.99$. The reason why we set $1-\alpha$ close to 1 is that we only want to avoid the extreme loss of the portfolio instead of all the risks. The intuition is that, if we try to eliminate all the risks, the constructed portfolio will also lack the ability to capture upward movements of the stock market.

As in Section 5.1, we use a 'back-testing' procedure to examine the performance of the weekly optimal portfolio constructed by the block factor SDM model. The basic steps are as follows. We initially estimate the block factor SDM model by using the first $s=1000$ observations. On the basis of the estimated model, we forecast the 1-week ahead out-of-sample optimal portfolio weight $\hat{\mathbf{w}}_s=\{\hat{w}_i^s\}_{i=1}^{28}$ for the next trading week of days $(s+1,s+2,\ldots,s+5)$ by using the optimal portfolio construction framework that was described above. The performance of the constructed optimal portfolio is $\hat{r}_s=\Sigma_{t=s+1}^{s+5}\Sigma_{i=1}^{28}\hat{w}_i^s r_{i,t}$, where $r_{i,t}$ is the actual log-return of the ith stock on day t. The estimation window is then rolled forward one trading week (5 days) and the optimal portfolio weight for the next week is generated. The last iteration uses days $\{T-1004,\ldots,T-5\}$ to generate optimal portfolio weight for the trading week of days $(T-4,T-3,\ldots,T)$. The starting date of the out-of-sample test is December 21st, 2007, and the ending date is December 30th, 2011, consisting of 203 weeks.

Besides the block factor max-copula, we also conduct experiments using the block factor Gaussian copula, the hierarchical Gumbel copula and the corresponding block factor mixture copula. The summary statistics of the portfolio returns in the out-of-sample test set for DJI30 and optimal portfolios constructed by each of the four copulas can be found in Table 5. As can be seen, compared with DJI30, all the copula-constructed portfolios deliver considerably higher annualized returns and lower risks, in terms of standard deviation, VaR and ES of the portfolio. Moreover, the portfolio that was constructed by the max-copula achieves the highest return and offers the best overall Sharpe ratio among all portfolios, followed by the mixture copula. We note that the Sharpe ratio is the most commonly used measure for risk-adjusted returns in the financial industry. For a given portfolio, its Sharpe ratio is defined to be $(r-r_\mathrm{f})/\sigma$, where r is its expected return, σ is its standard deviation and r_f is the risk-free rate. (Here, we set r_f to be the Federal funds rate, which was 0.00–0.25% during the test period. For simplicity we set $r_\mathrm{f}=0$.)

To demonstrate the comparison better, we plot the cumulative portfolio values of the four constructed portfolios and DJI30 throughout the test set in Fig. 4. As is clearly shown, the

Table 5. Summary statistics of the returns of different portfolios from December 21st, 2007, to December 30th, 2011†

Copula	*TR* (%)	*AR* (%)	*SD* (%)	*VaR95* (%)	*VaR99* (%)	*ES95* (%)	*ES99* (%)	*SR (overall)*	*SR (before)*	*SR (after)*
Max-copula	19.86	5.16	2.11	3.34	5.45	5.32	9.03	0.047	−0.169	0.165
Mixture	16.06	4.18	2.12	3.47	5.58	5.47	9.47	0.038	−0.179	0.155
Gaussian	12.87	3.35	2.19	3.19	5.90	5.63	9.58	0.029	−0.191	0.144
Gumbel	11.25	2.92	2.31	3.84	5.78	5.70	9.16	0.024	−0.143	0.113
DJI30	1.23	0.32	3.11	4.63	9.86	7.89	13.62	0.002	−0.254	0.142

†TR, total return; AR, annualized return; SR, Sharpe ratio.

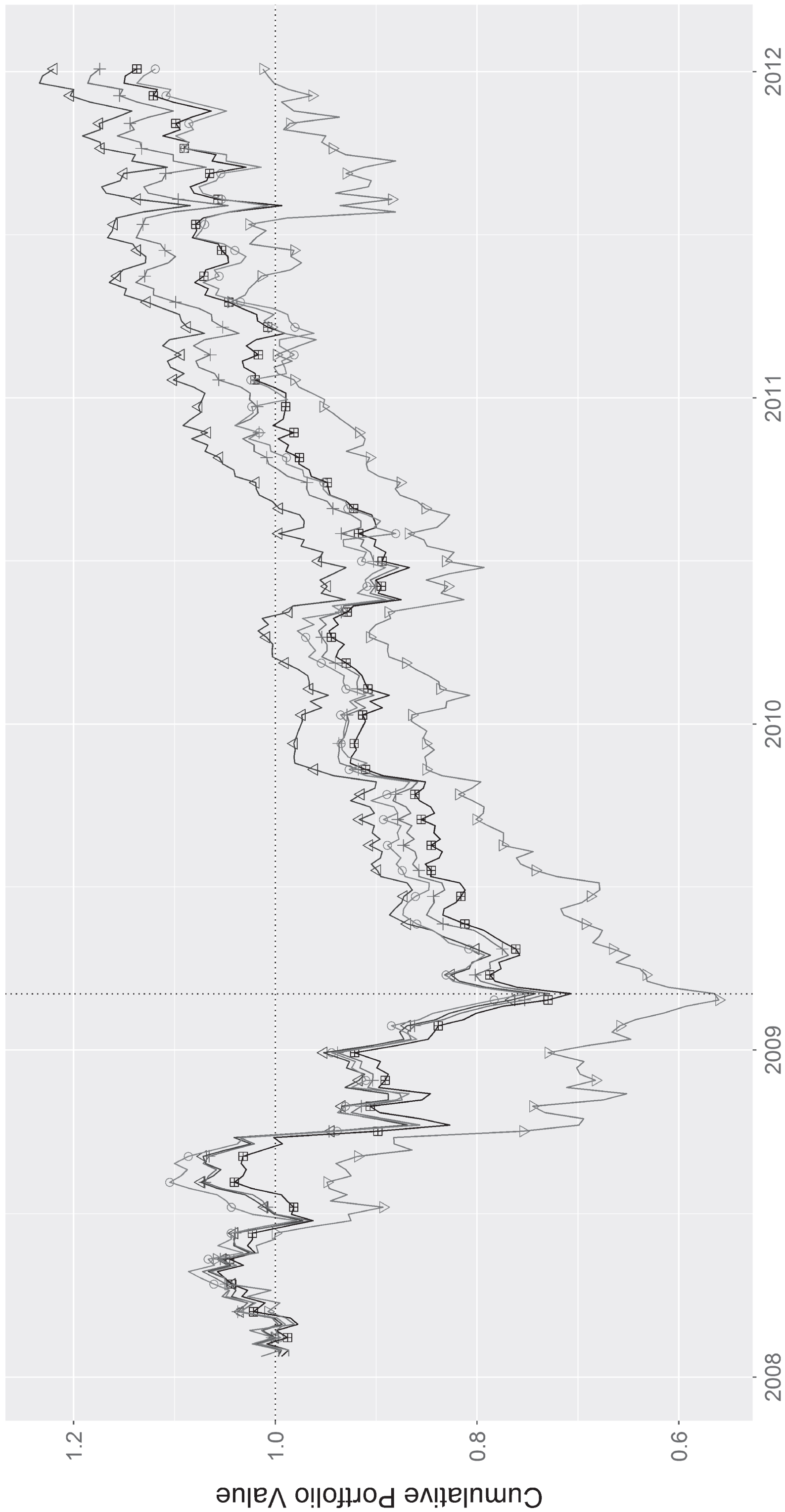

Fig. 4. Comparison of cumulative portfolio values of copula-based portfolios and DJI30 (∇) over the period December 21st, 2007–December 30th, 2011 (the cumulative portfolio value on day t is calculated by using $\exp(\Sigma_{s=1}^{t} \hat{r}_s)$, where $\hat{r}_s$ denotes the portfolio return on day s): ○, Gumbel copula; ⊞, Gaussian copula; +, mixture copula; △, max-copula

portfolio value that was constructed by the max-copula almost always stays on top. The vertical broken line marks the date (February 26th, 2009) when DJI30 reaches its minimum level, and we call it the minimal point. The market scenario before the minimal point can be seen as in the 'crisis' state and the point after the minimal point can be seen as in post-crisis 'normal or rally' state. To assess better the various copulas' performance under different market scenarios, separate analyses were performed on portfolio returns before and after the minimal point. We summarize the Sharpe ratio in Table 5 and more detailed information is in section 4.2 of the online supplementary material. As can be seen, the Gumbel copula does not perform well during the normal or rally state, the Gaussian copula does not perform well during the crisis state, and the max-copula offers the best balance of performance under both market scenarios (best in the normal or rally state and second best in the crisis state).

6. Conclusion

In this paper, we have proposed and studied the max-copula, which is a novel non-linear asymmetric copula-generating framework that constructs new flexible copulas on the basis of existing copulas through a pairwise max-function. The constructed max-copula enjoys tractable theoretical properties, such as closed form quantile and tail dependence functions. Moreover, it is capable of modelling asymmetric dependence and joint tail behaviour while offering good performance in non-extremal behaviour modelling. Max-copulas based on single-factor and block factor models were developed to offer parsimonious modelling for structured dependence, especially in high dimensional applications. Combined with semiparametric time series models, the framework proposed can further help to obtain flexible and accurate models for multivariate time series. The consistency and asymptotic normality of the CMLE proposed has been affirmed through extensive numerical experiments. The max-copula's ability to model multivariate financial time series has been demonstrated by the estimation of the conditional VaR for a financial portfolio and by weekly optimal portfolio constructions for the Dow Jones industrial average under various market scenarios.

Since the max-copula is relatively new, much work remains to be done. Its advantage and disadvantage relative to the mixture copula and the factor copula can further be explored. There are also several interesting extensions of the max-copula itself. A natural extension is to design a time varying max-copula. Instead of setting the weight parameter c to be constant, we can design an auto-regressive structure for c_t, such that c_t depends on $\{c_{t-i}, i>0\}$ and on whether $\mathbf{C}_1$ or $\mathbf{C}_2$ takes the lead on day $t-1$. Compared with constant c, a dynamic c_t may be more realistic in real data applications.

Acknowledgements

The authors thank the Joint Editor, Associate Editor and two referees for their valuable comments and suggestions. The work is partially supported by National Science Foundation grants NSF-DMS-1505367 and NSF-CMMI-1536978.

References

Aas, K., Czado, C., Frigessi, A. and Bakken, H. (2009) Pair-copula constructions of multiple dependence. *Insur. Math. Econ.*, **44**, 182–198.

Almeida, C., Czado, C. and Manner, H. (2012) Modeling high dimensional time-varying dependence using D-vine SCAR models. *Working Paper*. Technische Universität München, Munich. (Available from `http://arxiv.org/pdf/1202.2008v1.pdf`.)

Cai, Z. and Wang, X. (2014) Selection of mixed copula model via penalized likelihood. *J. Am. Statist. Ass.*, **109**, 788–801.
Chen, X. and Fan, Y. (2006a) Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *J. Econmetr.*, **135**, 125–154.
Chen, X. and Fan, Y. (2006) Estimation of copula-based semiparametric time series models. *J. Econmetr.*, **130**, 307–335.
Cherubini, U., Luciano, E. and Vecchiato, W. (2004) *Copula Methods in Finance*. New York: Wiley.
Coval, J., Jurek, J. and Stafford, E. (2009) The Economics of structured finance. *J. Econ. Perspect.*, **23**, 3–25.
Cui, Q. and Zhang, Z. (2017) Max-linear competing factor models. *J. Bus. Econ. Statist*., to be published.
Durante, F. and Sempi, C. (2015) *Principles of Copula Theory*. Boca Raton: CRC Press.
Frees, E. W. and Valdez, E. A. (1998) Understanding relationships using copulas. *Nth Am. Act. J.*, **2**, 1–25.
Gao, X. and Song, P. (2010) Composite likelihood Bayesian Information Criteria for model selection in high-dimensional data. *J. Am. Statist. Ass.*, **105**, 1531–1540.
Genest, C. and Favre, A. (2007) Everything you always wanted to know about copula modeling but were afraid to ask. *J. Hydrol. Engng*, **12**, 347–368.
Genest, C., Ghoudi, K. and Rivest, L. P. (1995) A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, **82**, 543–552.
Genest, C., Remillard, B. and Beaudoin, D. (2009) Goodness-of-fit tests for copulas: a review and a power study. *Insur. Math. Econ.*, **44**, 199–213.
Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. *J. Finan.*, **48**, 1779–1801.
Harris, R. and Mazibas, M. (2013) Dynamic hedge fund portfolio construction: a semi-parametric approach. *J. Bnkng Finan.*, **37**, 139–149.
Hull, J. and White, A. (2004) Valuation of a CDO and an nth to default CDS without Monte Carlo simulation. *J. Derivs*, **12**, 8–23.
Joe, H. (2014) *Dependence Modeling with Copulas*. Boca Raton: Chapman and Hall–CRC.
Khoudraji, A. (1995) Contributions a l'etude des copules et a la modelisation des valeurs extremes bivariees. *PhD Thesis.* Université Laval, Quebec.
Krupskii, P. and Joe, H. (2013) Factor copula models for multivariate data. *J. Multiv. Anal.*, **120**, 85–101.
Liebscher, E. (2008) Construction of asymmetric multivariate copulas. *J. Multiv. Anal.*, **99**, 2234–2250.
Lindsay, B. G. (1988) Composite likelihood methods. *Contemp. Math.*, **80**, 221–239.
McNeil, A. J. and Frey, R. (2000) Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *J. Empir. Finan.*, **7**, 271–300.
McNeil, A. J., Frey, R. and Embrechts, P. (2005) *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton: Princeton University Press.
Min, A. and Czado, C. (2010) Bayesian inference for multivariate copulas using pair-copula constructions. *J. Finan. Econmetr.*, **8**, 511–546.
Murray, J., Dunson, D., Lawrence, C. and Lucas, J. (2013) Bayesian Gaussian copula factor models for mixed data. *J. Am. Statist. Ass.*, **108**, 656–665.
Nelsen, R. B. (2006) *An Introduction to Copulas*. New York: Springer.
Oh, D. H. and Patton, A. (2013) Simulated method of moments estimation for copula-based multivariate models. *J. Am. Statist. Ass.*, **108**, 689–700.
Oh, D. H. and Patton, A. (2017) Modelling dependence in high dimensions with factor copulas. *J. Bus. Econ. Statist.*, **35**, 139–154.
Salmon, F. (2012) The formula that killed Wall Street. *Significance*, **9**, 16–20.
Sibuya, M. (1959) Bivariate extreme statistics. *Ann. Inst. Statist. Math.*, **11**, 195–210.
Sklar, A. (1959) Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, **8**, 229–231.
Varin, C., Reid, N. and Firth, D. (2011) An overview of composite likelihood methods. *Statist. Sin.*, **21**, 5–42.

Supporting information

Additional 'supporting information' may be found in the on-line version of this article:

'Supplementary material: Semi-parametric dynamic max-copula model for multivariate time series'.