Conditional Normal Extreme-Value Copulas

Pavel Krupskii\(^1\) and Marc G. Genton\(^2\)

June 23, 2020

Abstract

We propose a new class of extreme-value copulas which are extreme-value limits of conditional normal models. Conditional normal models are generalizations of conditional independence models, where the dependence among observed variables is modeled using one unobserved factor. Conditional on this factor, the distribution of these variables is given by the Gaussian copula. This structure allows one to build flexible and parsimonious models for data with complex dependence structures, such as data with spatial or temporal dependence. We study the extreme-value limits of these models and show some interesting special cases of the proposed class of copulas. We develop estimation methods for the proposed models and conduct a simulation study to assess the performance of these algorithms. Finally, we apply these copula models to analyze data on monthly wind maxima and stock return minima.

Key words: Factor copula; Residual dependence; Tail asymmetry; Tail dependence; Spatial dependence.

Short title: Conditional Normal Extreme-Value Copulas

\(^1\)University of Melbourne, Parkville 3010, Australia. E-mail: pavel.krupskiy@unimelb.edu.au
This research was supported by the Andrew Sisson fund.

\(^2\)Statistics Program, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. E-mail: marc.genton@kaust.edu.sa
This research was supported by KAUST.
1 Introduction

Copula models have become more popular in modeling non-Gaussian data. A copula function is a multivariate cumulative distribution function (cdf) with standard uniform $U(0,1)$ marginal cdfs. Sklar (1959) showed that for any $d$-variate continuous cdf, $F_{1,\ldots,d}(z_1,\ldots,z_d)$, with univariate marginal cdfs, $F_1,\ldots,F_d$, there exists a unique copula, $C_{1,\ldots,d}$, such that $F_{1,\ldots,d}(z_1,\ldots,z_d) = C_{1,\ldots,d}\{F_1(z_1),\ldots,F_d(z_d)\}$ for any $z_1,\ldots,z_d$. This implies that the copula can be used to construct a multivariate distribution with given marginal cdfs. This allows greater flexibility when modeling multivariate data.

Many existing copula families, however, are not suitable for modeling multivariate data because they cannot generate flexible dependence structures. For example, most Archimedean copulas have exchangeable dependence (McNeil et al., 2005) and the multivariate Student-$t$ copula is reflection symmetric; asymmetric versions of this copula have been proposed in the literature, but their parameter estimation can be computationally demanding (Yoshiba, 2018).

Vine copula models can be used to construct very flexible distributions; see, for example, Aas et al. (2009) and Kurowicka and Cooke (2006). These models require the estimation of $O(d^2)$ dependence parameters and might therefore be computationally demanding if $d$ is large. Truncated vines assume independence after conditioning on some variables and have $O(d)$ parameters (Brechmann et al., 2012). However, vine copula models lack interpretability and might be unable to capture some features of data. One example is data with spatial structure when dependence is weaker with larger distance, and this property is generally not satisfied by vine copula models.

Conditional independence models are another class of models in which observed variables are assumed to be independent, conditionally on unobserved (latent) factors. These
parsimonious and flexible models can be used for modeling data when all variables can be simultaneously affected by unobserved driving variables, such as financial stock returns; see Krupskii and Joe (2013, 2015) and Oh and Patton (2017). Extreme-value limits of these models can be used to analyze multivariate extremes data with a factor structure (Lee and Joe, 2018).

However, in many applications these latent factors may not explain all the dependence among the observed variables. To achieve greater flexibility, it is therefore useful to assume some residual dependence after conditioning on the unobserved factors. For example, Krupskii et al. (2018a) proposed a copula for modeling a non-Gaussian spatial process. In this model, the variables have a multivariate Gaussian distribution after conditioning on an unobserved latent variable which does not depend on a spatial location. Krupskii et al. (2018b) showed that the extreme-value limit of this copula is a flexible Hüsler-Reiss copula (Hüsler and Reiss, 1989) which is popular in different applications.

In this paper, we study extreme-value limits of conditional normal copula models proposed by Krupskii and Joe (2020). Let $U = (U_1, \ldots, U_d)^	op$, $U_j \sim U(0,1)$, $j = 1, \ldots, d$. Assume that $V_0 \sim U(0,1)$ is the unobserved (latent) variable and that $(U|V_0 = v_0) \sim C_N(\cdot; \Sigma)$ where $C_N(\cdot; \Sigma)$ is the Gaussian (Normal) copula with the correlation matrix $\Sigma$. Let $C_{j,0}$ be the copula linking $U_j$ and $V_0$ and let $C_{j|0}(u_j|v_0) = \partial C_{j,0}(u_j, v_0)/\partial v_0$ be the conditional copula cdf. Throughout this paper, we use small letters for the corresponding copula probability density functions (pdfs). We also assume that all bivariate copulas $C_{j,0}$ and their densities are strictly positive continuous functions on $(0,1)^2$.

The joint copula cdf and pdf for $U$ at $u = (u_1, \ldots, u_d)^	op$ are

$$C_U(u; \Sigma) = \int_0^1 C_N\{C_{1|0}(u_1|v_0), \ldots, C_{d|0}(u_d|v_0); \Sigma\}dv_0,$$

$$c_U(u; \Sigma) = \int_0^1 c_N\{C_{1|0}(u_1|v_0), \ldots, C_{d|0}(u_d|v_0); \Sigma\} \cdot \prod_{j=1}^d c_{j,0}(u_j, v_0)dv_0. \quad (1)$$
Note that in the general case the matrix $\Sigma$ can depend on the conditioning value $V_0 = v_0$.

These models form the general class of models where the variables are assumed to have a multivariate Gaussian distribution after conditioning on an unobserved latent factor. In this paper, we consider upper tail extreme-value limits of the copula model (1). The respective limiting extreme-value copula is defined as

$$C_{EV}^U(u_1, \ldots, u_d) = \exp[-V\{ -\ln(u_1), \ldots, -\ln(u_d) \}]$$

(2)

where $V(w_1, \ldots, w_d)$ is the exponent function (de Haan and Ferreira, 2006) and is defined as

$$V(w_1, \ldots, w_d) = \lim_{u \to 0} \frac{1}{u} \{ 1 - C_U(1 - uw_1, \ldots, 1 - uw_d) \}$$

assuming this limit exists.

We show that these extreme-value limiting copulas are computationally tractable and they can be used for modeling data with complex dependencies, such as multivariate extremes with dynamic dependence, or spatial extremes. Models with a multifactor structure can also be obtained when the multivariate Gaussian distribution has a factor correlation structure, i.e., when $C_N$ is a normal factor copula (Krupskii and Joe, 2013).

The rest of this paper is organized as follows. In Section 2 we review dependence properties of conditional normal copulas (1) and define the respective limiting extreme-value copulas. In Section 3 we consider some special cases with parsimonious dependence and provide more details on parameter estimation in Section 4. We then apply the proposed models to analyze monthly wind maxima and monthly stock returns minima data in Section 6. Section 7 concludes with a discussion.
2 Conditional Normal Copulas and Extreme-Value Limits

2.1 Dependence properties of conditional normal copulas

We now review some dependence properties of the copula $C_U$ in (1). We assume that the correlation matrix, $\Sigma$, is a constant and it does not depend on $v_0$. The properties below follow from the definition of the copula $C_U$:

1. If $\Sigma = I_d$, then $C_U$ simplifies to the one-factor copula cdf (Krupskii and Joe, 2013);

2. $C_U$ increases in concordance ordering as elements of $\Sigma$ increase (such that $\Sigma$ remains positive definite), with $C_{1,0}, \ldots, C_{d,0}$ fixed;

3. If $\Sigma = J_d$ (a $d \times d$ matrix of ones) and $C_{1,0} = C_{2,0} = \cdots = C_{d,0}$, then $C_U$ is the comonotonicity copula: $C_U(u) = \min(u)$. Note that in the general case with different linking copulas $C_{j,0}$ for $j = 1, \ldots, d$, $C_U$ is no longer the comonotonicity copula;

4. If all linking copulas $C_{j,0}$, $j = 1, \ldots, d$, are independence copulas, then $C_U$ is the normal copula with the correlation matrix $\Sigma$.

Tail properties of the copula $C_U$ depend on tail properties of the linking copulas, $C_{j,0}$, $j = 1, \ldots, d$, similar to the one-factor copula models. Therefore, $C_U$ has lower/upper tail dependence if the linking copulas $C_{j,0}$ have lower/upper tail dependence. We consider the lower tail as properties in the upper tail can be obtained in a similar way.

Proposition 1 (Krupskii and Joe, 2020) Assume that $\lim_{u \to 0} C_{j,0}(uw_j | uw_0) = b_{j,0}(w_j | w_0)$ and $\lim_{u \to 0} C_{0,j}(uw_0 | uw_j) = b_{0,j}(w_0 | w_j)$ where $b_{j,0}(\cdot | w_0)$ is a proper distribution function for any $w_0 > 0$ and $b_{0,j}(\cdot | w_j)$ is a proper distribution function for any $w_j > 0$, $j = 1, \ldots, d$. Then the lower tail function $b_{1:d}(w) := \lim_{u \to 0} C_U(uw)/u > 0$, where $w = (w_1, \ldots, w_d)^\top$. 

4
The assumptions of Proposition 1 are satisfied for many bivariate copulas with lower tail dependence, such as the reflected Gumbel, Clayton and other copula families; see Joe (2014) for more details on bivariate copula families and their dependence properties. The resulting copula $C_U$ has lower tail dependence in this case. At the same time, if $C_{j,0}$ has lower tail quadrant independence, then $C_U$ is a copula with no lower tail dependence as the next proposition shows.

**Proposition 2** *(Krupskii and Joe, 2020)* Let $\kappa_L$ be the lower tail order of $C_U$ *(Hua and Joe, 2011)*. Assume that $\Sigma$ is non-degenerate and $m_j^- u \leq C_{j,0}(u|v) \leq m_j^+ u$ for small enough $u > 0$ and some constants $m_j^-, m_j^+ > 0$, $j = 1, \ldots, d$. Then $\kappa_L = \kappa_\Sigma$ where $\kappa_\Sigma = \text{tr}(\Sigma^{-1})$ is the tail order of the normal copula with the correlation $\Sigma$.

Again, it is easy to check that the assumption of Proposition 2 is satisfied for many bivariate copulas with lower tail quadrant independence, including the Frank, Plackett and other copulas.

When using linking copulas $C_{j,0}$, $j = 1, \ldots, d$, with intermediate lower tail dependence, the copula $C_U$ also has intermediate lower tail dependence in many cases. For example, if $C_{j,0}$ are the Gaussian copulas with correlation parameters $\rho_j < 1$, then the copula $C_U$ is the Gaussian copula with the correlation matrix $\Sigma^*$ such that $(\Sigma^*)_{j,k} = \rho_j \rho_k + \sqrt{(1 - \rho_j^2)(1 - \rho_k^2)(\Sigma)_{j,k}} < 1$, $1 \leq j < k \leq d$. The next proposition shows that intermediate tail dependence for the copula $C_U$ can also be obtained when using extreme-value linking copulas $C_{j,0}$.

**Proposition 3** Assume that $\Sigma$ is non-degenerate and $C_{j,0}(u_1, u_2) = (u_1 u_2)^{A_j(\ln(u_2)/\ln(u_1 u_2))}$ where $A_j(\cdot) : [0, 1] \mapsto [1/2, 1]$ is a convex function such that $A_j(t) \geq \max(t, 1 - t)$, $j = 1, \ldots, d$. We also assume that $A_j(t)$ is a continuously differentiable function with $A_j'(t) > -1$
for $t \in (0, 0.5)$, $j = 1, \ldots, d$. Let the lower tail order of $C_{j,0}$ be $\kappa_j = 2A_j(1/2) > 1$. It implies that $C_U$ is a copula with intermediate lower tail dependence.

Proof: See Appendix A.1.

2.2 Extreme-value limits of conditional normal copulas

We now consider upper tail extreme-value limits of the copula model (1). Lower tail limits can be considered analogously. To get non-trivial limits with $C_N$ with non-degenerate $\Sigma$, we need to consider the case when the linking copulas $C_{j,0}$, $j = 1, \ldots, d$, have upper tail dependence because the copula $C_U$ has upper tail dependence in this case.

**Proposition 4** Assume that $b_{j|0}(w_j|w_0) = 1 - \lim_{u \to 0} C_{j|0}(1 - uw_j|1 - uw_0)$ and $b_{0|j}(w_0|w_j) = 1 - \lim_{u \to 0} C_{0|j}(1 - uw_0|1 - uw_j)$ where $b_{j|0}(\cdot|w_0)$ is a proper distribution function for any $w_0 > 0$ and $b_{0|j}(\cdot|w_j)$ is a proper distribution function for any $w_j > 0$. It follows that the exponent function of $C_{EV}^U$ is

$$V(w_1, \ldots, w_d) = \int_0^\infty \left[1 - C_N\{1 - b_{1|0}(w_1|w_0), \ldots, 1 - b_{d|0}(w_d|w_0); \Sigma}\right] dw_0, \ w_1, \ldots, w_d > 0.$$ 

Proof: See Appendix A.2.

With continuously differentiable tail functions $b_{j|0}(w_j|w_0)$, $j = 1, \ldots, d$, the exponent function in Proposition 4 is also a continuously differentiable function and therefore the resulting extreme-value copula is an absolutely continuous copula if $\Sigma$ is not a singular matrix. If $C_N$ is the comonotonicity copula (i.e., $\Sigma = J_d$), then the formula for $V(w_1, \ldots, w_d)$ simplifies to

$$V(w_1, \ldots, w_d) = \int_0^\infty \max_j \{b_{j|0}(w_j|w_0)\} dw_0.$$ 

One interesting class of limiting extreme-value copulas arises when $C_N$ is the comonotonicity copula and $c_{j,0}$, $j = 1, \ldots, d$, are continuous functions on $(0, 1]^2$. 

Proposition 5 Assume that $\Sigma = J_d$ is a matrix of ones and $c_{j,0}(u_j, u_0)$ are continuous functions on $(0, 1]^2$ and positive on $(0, 1)^2$, $j = 1, \ldots, d$. The exponent measure of $C_{U}^{EV}$ is

$$V(w) = V(w_1, \ldots, w_d) = \int_0^1 \max_j \{ w_j c_{j,0}(1, v_0) \} dv_0, \quad w_1, \ldots, w_d > 0.$$ 

Proof: See Appendix A.3.

Many copulas with tail quadrant independence satisfy the conditions of Proposition 5. It follows that the support of the corresponding limiting copulas is a subset of $[0, 1]^d$. For example, if $w_1$ is such that $w_1 \min_{0 < v_0 < 1} \{ c_{j,0}(1, v_0) \} > \max_{0 < v_0 < 1, k \neq 1} \{ w_k c_{k,0}(1, v_0) \}$, then $V(w_1, \ldots, w_d) = w_1$ and hence the density in the neighborhood of the corner $(1, 0, \ldots, 0)^T$ is zero. This class of copulas can therefore be used for modeling multivariate extremes with very strong dependence (close to comonotonic dependence).

Example 1: Let $d = 2$ and $C_{1,0}$ and $C_{2,0}$ are the Farlie-Gumbel-Morgenstern copulas with parameters $\theta_1 = 0.5$ and $\theta_2 = -0.5$, respectively. One can find that

$$V(w_1, w_2) = \begin{cases} \frac{9w_1^2 + 2w_1 w_2 + 9w_2^2}{8(w_1 + w_2)}, & \text{if } \frac{1}{3} < \frac{w_1}{w_2} < 3, \\ w_1, & \text{if } \frac{w_1}{w_2} > 3, \\ w_2, & \text{if } \frac{w_1}{w_2} < \frac{1}{3}. \end{cases}$$

This implies that the limiting copula density is positive only if $w_2^3 < u_1 < u_2^{1/3}$. It is seen that the exponent function is a continuously differentiable function with respect to $w_1$ and $w_2$ and hence the conditional limiting copula is also continuous. The copula density is a continuous function for $u_2^3 < u_1 < u_2^{1/3}$.

3 Parsimonious Dependence Structures

Different types of dependencies can be generated by the extreme-value limits of $C_{U}$, depending on the choice of the linking copulas, $C_{j,0}$, $j = 1, \ldots, d$, and the correlation matrix $\Sigma$. We will now consider some special cases resulting in parsimonious dependence structures for the extreme-value limit of the copula $C_{U}$ with $O(d)$ dependence parameters.
3.1 Copulas with spatial or temporal dependence

If $\Sigma$ is a spatial correlation matrix, then the limiting copula can be useful for modeling spatial extremes. Different covariance functions can be used to construct the covariance matrix with a spatial structure; see Gneiting et al. (2007) for an overview of covariance functions. Note that one should select the same linking copulas $C_{1,0} = C_{2,0} = \cdots = C_{d,0}$ to ensure that the $(j, k)$ margin converges to the comonotonicity copula when $\Sigma_{jk} \to 1$ for $1 \leq j < k \leq d$. Otherwise, one will get the limiting distribution as described in Proposition 5, which is not a comonotonicity copula in the general case.

Assuming the same linking copulas are used, one can control the rate of decay of tail dependence as a function of distance or time lag by selecting an appropriate spatial correlation matrix, as the next proposition shows.

**Proposition 6** Consider the $(1,2)$-margin of the copula $C^{EV}_U$ with $C_{1|0} = C_{2|0}$ defined in (2). Assume that $b_{1|0}(1|w) \sim w^{-\theta} \ell(w)$ as $w \to \infty$ for $\theta > 1$, where $\ell(w)$ is a slowly varying function. Under the assumptions of Proposition 4,

$$V(1, 1) = 1 + \left(\frac{1 - \rho}{\pi}\right)^{1/2} \int_{0}^{\infty} \phi[\Phi^{-1}\{b_{1|0}(1|w_0)\}]\,dw_0 + O((1 - \rho)^{3/2}),$$

as $\rho \to 1$, where $\Phi(\cdot)$, $\phi(\cdot)$ are the standard normal cdf and density, respectively, and $\rho = \Sigma_{1,2}$.

**Proof:** See Appendix A.4.

Assume that $\rho(d) = 1 - \eta - Cd^\alpha + o(d^\alpha)$ for some constants $C, \alpha > 0$, where $0 < \eta < 1$ is a nugget effect and $d$ is the distance between two locations. Proposition 6 implies that the upper tail dependence coefficient corresponding to the $(1,2)$-margin of the copula $C^{EV}_U$ is

$$\lambda_U(d) = 2 - V(1, 1) = 1 - \left(\frac{\eta + Cd^\alpha}{\pi}\right)^{1/2} \int_{0}^{\infty} \phi[\Phi^{-1}\{b_{1|0}(1|w_0)\}])\,dw_0 + o(d^{\alpha/2}).$$

A covariance function $\rho$ with a smaller $\alpha$ can therefore be selected to obtain a faster rate of decay of the tail dependence as a function of distance.
To model multivariate spatial extremes, different linking copulas (they can be from different parametric families or from the same parametric family but have different parameters) can be used for different variables together with the cross-covariance matrix $\Sigma$; see Genton and Kleiber (2015) for a review on cross-covariance models. For example, to model bivariate spatial extremes, one can select $C_{1,0} = C_{2,0} = \cdots = C_{m,0}$ for the first variable and $C_{m+1,0} = C_{m+2,0} = \cdots = C_{d,0}$ for the second variable, where $d = 2m$ and $m$ is the number of spatial locations.

To model spatial isotropy, one can select an isotropic correlation matrix $\Sigma$, and for non-stationary data, one can select a non-stationary correlation matrix $\Sigma$.

**Example 2:** The Hüsler-Reiss copula (Hüsler and Reiss, 1989) can be obtained as an extreme-value limit of the convolution of exponential and multivariate normal distributions (Krupskii et al., 2018b). In this model, the copula $C_U$ corresponds to the joint distribution of $W = (W_1, \ldots, W_d)^\top$, where $W_j = Z_j + \alpha_j V_0$, $Z = (Z_1, \ldots, Z_d)^\top$ has a multivariate normal distribution with $N(0, 1)$ marginals and the correlation matrix $\Sigma$ and where $Z$ is independent of $V_0$. Here, $C_{j,0}$ links $W_j = Z_j + \alpha_j V_0$ and $V_0 \sim \text{Exp}(1)$, and one can show that the tail function of this copula is $b_{j,0}(w_j|w_0) = \Phi \left( \alpha_j \ln \frac{w_j}{w_0} - \frac{1}{2\alpha_j} \right)$, $j = 1, \ldots, d$.

**Example 3:** An extremal $t$ copula (Demarta and McNeil, 2005) can be obtained as an extreme-value limit of the $t$ copula with $\nu > 0$ degrees of freedom (Nikoloulopoulos et al., 2009). In this model, the copula $C_U$ corresponds to the joint distribution of $W = (W_1, \ldots, W_d)^\top$, where $W_j = V_0 Z_j$, $Z = (Z_1, \ldots, Z_d)^\top$ has the multivariate normal distribution with $N(0, 1)$ marginals and the correlation matrix $\Sigma$ and where $Z$ is independent of $V_0$. Here, $C_{j,0}$ links $W_j = V_0 Z_j$ and $V_0 = (Y/\nu)^{-1/2}$ with $Y \sim \chi^2(\nu)$. 

9
3.2 Copulas with factor structure and dynamic dependence

The limiting extreme-value copula with a one factor structure can be obtained with \( \Sigma = I_d \); this class of models was studied by Lee and Joe (2018), who considered continuous copulas. Copulas with singular components can be obtained by the discontinuous functions \( b_{j|0}(w_j|w_0), j = 1, \ldots, d \). These copulas can be used to model the lifetime of system components that may fail simultaneously, with a positive probability. One example is Marshall-Olkin copulas with a factor structure (Krupskii and Genton, 2018). One application of such copulas is to model the times-to-default for components of a credit portfolio. A correlation matrix \( \Sigma \) with a one-factor structure can be used to obtain two-factor structure models, and bi-factor models can be obtained if \( \Sigma \) is a block-diagonal matrix. In all these cases, a one-dimensional integration is required to compute the copula density. Lee and Joe (2018) have provided some details about extreme-value copulas with two-factor structures; however, their approach requires the computation of two-dimensional integrals and is therefore not feasible in very high dimensions.

Dynamic dependence can be modeled by selecting a matrix \( \Sigma \) with a simple parsimonious structure with time-varying correlations and linking copulas \( C_{j,0}, j = 1, \ldots, d \), that do not change over time. The resulting copula in (1) can be used to model data with dynamic extremes, such as stock returns’ monthly maxima or minima, or time series extremes.

**Example 4:** Let \( t = 1, \ldots, T \), and we assume \( \Sigma = \Sigma(t) \) with time-varying correlations:

\[
(\Sigma_t)_{j,k} = \rho(t) \rho_j \rho_k, 1 \leq j \leq k \leq d,
\]

where \( 0 \leq \rho_1, \ldots, \rho_d \leq 1 \) and

\[
\rho(t) = \frac{1}{1 + \exp\{\eta(t)\}}, \quad \eta(t) = \beta_0 + \sum_{k=1}^{K} \beta_k V_k(t),
\]

where \( V_1(t), \ldots, V_K(t) \) are time-dependent external variables used to model nonstationarity.
in time series. Here, $\Sigma$ is a correlation matrix of $Z = (W_1, \ldots, W_d)^\top$ where

$$W_i = \{\rho(t)\}^{1/2}\rho_i Z_0 + \{1 - \rho(t)\rho_i^2\}^{1/2}Z_i, \quad i = 1, \ldots, d,$$

and $Z_0, Z_1, \ldots, Z_d \sim_{\text{i.i.d.}} N(0, 1)$. This correlation structure can be combined with linking copulas $C_{j,0}$ that have lower/upper tail dependence to model minima/maxima time series data. Krupskii and Joe (2020) used the conditional normal copula $C_U$ with correlation structure (3) and with BB1 copulas $C_{j,0}$ to model dynamic dependence in European bonds, credit default swaps, and US stock returns data.

4 Parameter Estimation

In this section, we show how the parameters of $C_{EV}^U$ in (2) can be estimated. We assume that $\theta_j$ is a vector of parameters for linking copulas $C_{j,0}$, $j = 1, \ldots, d$, and $\theta_\Sigma$ is a vector of parameters for the correlation matrix $\Sigma$.

4.1 Maximum likelihood inference for $C_{EV}^U$ with non-degenerate $\Sigma$

Here we show how to obtain parameter estimates for the copula $C_{EV}^U$ using the maximum likelihood approach. We assume that the conditions of Proposition 4 are satisfied. As likelihood inference is not feasible for extreme-value copulas in high dimensions, we will use the pairwise likelihood approach to estimate these parameters; see, for example, Lindsay (1998).

Let $\{u_i = (u_{i1}, \ldots, u_{id})^\top\}_{i=1}^n$ be a sample of size $n$ from $C_{EV}^U$. In applications, models of univariate data can be fitted first, and the probability integral transform can be applied to convert the original data to uniform $U(0,1)$ data. Alternatively, uniform scores data can be obtained using rank transforms (Genest et al., 1995).
The pairwise log-likelihood is
\[
\ell(u_1, \ldots, u_n; \theta) = \sum_{i=1}^{n} \sum_{1 \leq j < k \leq d} \ln c_{j,k}^{EV}(u_{j|i0}, u_{k|i0}; \theta_j, \theta_k, \theta_S), \quad \theta = (\theta_1^T, \ldots, \theta_d^T, \theta_S^T)^T.
\]

The copula pdf of the \((j, k)\)-margin is
\[
c_{j,k}^{EV}(u_j, u_k) = c_{j,k}^{EV}(u_{j|i0}, u_{k|i0}) \left\{ \frac{\partial V_{j,k}(w_j, w_k)}{\partial w_j} \cdot \frac{\partial V_{j,k}(w_j, w_k)}{\partial w_k} - \frac{\partial^2 V_{j,k}(w_j, w_k)}{\partial w_j \partial w_k} \right\},
\]
where \(w_j = -\ln u_j, w_k = -\ln u_k\), \(c_{j,k}^{EV}(u_j, u_k) = \exp\{-V_{j,k}(w_j, w_k)\}\) and
\[
V_{j,k}(w_j, w_k) = \int_0^\infty \left[ 1 - C_N\{1 - b_{j|i0}(w_j|w_0), 1 - b_{k|i0}(w_k|w_0); \rho_{jk}\}\right] dw_0, \quad \rho_{jk} = \Sigma_{j,k},
\]
\[
\frac{\partial V_{j,k}(w_j, w_k)}{\partial w_j} = \int_0^\infty C_N(1 - b_{k|i0}|1 - b_{j|i0}; \rho_{jk}) b_{j,0}(w_j, w_0) dw_0,
\]
\[
\frac{\partial V_{j,k}(w_j, w_k)}{\partial w_k} = \int_0^\infty C_N(1 - b_{j|i0}|1 - b_{k|i0}; \rho_{jk}) b_{k,0}(w_k, w_0) dw_0,
\]
\[
\frac{\partial^2 V_{j,k}(w_j, w_k)}{\partial w_k \partial w_j} = -\int_0^\infty c_N\{1 - b_{j|i0}(w_j|w_0), 1 - b_{k|i0}(w_k|w_0); \rho_{jk}\} b_{j,0}(w_j, w_0) b_{k,0}(w_k, w_0) dw_0,
\]
where \(b_{j,0}(w_j, w_0) = \partial b_{j|i0}(w_j|w_0)/\partial w_j, b_{k,0}(w_k, w_0) = \partial b_{k|i0}(w_k|w_0)/\partial w_k\).

Numerical integration is required to compute \(V_{j,k}(w_j, w_k)\) and its derivatives. The integrand is a slowly decaying function of \(w_0\), so some changes in variables are required to make the computation more efficient. More details about computing \(V_{j,k}(w_j, w_k)\) are provided in Appendix A.5. A quasi-Newton-Raphson method can then be used to estimate a vector of parameters \(\theta\) for the copula \(C_U^{EV}\).

### 4.2 Inference for \(C_U^{EV}\) with degenerate \(\Sigma\)

In this section, we assume that the conditions of Proposition 5 are satisfied for \(C_U^{EV}\). The pairwise likelihood approach is not feasible in this case, as the domain of the \((j, k)\) marginal copula depends on \(\theta_j\) and \(\theta_k\), and it can be very difficult to identify the parameter values \(\theta\) where the copula density is positive if the dimension \(d\) is large.
Instead, we can obtain consistent non-parametric estimates of the upper tail dependence coefficient $\lambda_{j,k} = \lambda_{j,k}(\theta_j, \theta_k) = 2 - V_{j,k}(1,1)$ for $1 \leq j < k \leq d$ (Ferreira, 2013) and define

$$\hat{\theta} = \text{argmin} \sum_{1 \leq j < k \leq d} \left\{ \hat{\lambda}_{j,k} - \lambda_{j,k}(\theta_j, \theta_k) \right\}^2,$$

where $\hat{\lambda}_{j,k}$ and $\lambda_{j,k}(\theta_j, \theta_k)$ are non-parametric and model-based estimates, respectively.

In Appendix A.6, we provide model details, including the explicit formula for $V_{j,k}(1,1)$ as a function of $\theta_j$ and $\theta_k$, for $C_{EV}^U$ with Clayton linking copulas $C_{j,0}, j = 1, \ldots, d$. Similarly, one can use different linking copulas with upper tail quadrant independence, such as the Frank copula.

Remark: This method can also be used to estimate the parameters of $C_{EV}^U$ with a non-degenerate $\Sigma$. The computation of $V_{j,k}(1,1)$ for different pairs $(j,k)$ is much easier with this family of copulas, and the respective estimates can be used as starting values for the quasi-Newton-Raphson algorithm discussed in the previous section. We provide more detail in the next section.

5 Monte Carlo Simulations

In this section, we conduct some simulation studies to check the performance of the parameter estimation methods discussed in Section 4.

5.1 Simulation 1

We simulate a data set from the extreme-value copula (2) with the reflected Clayton linking copulas that have upper tail dependence $C_{j,0}, j = 1, \ldots, d$ and $d = 10$. We assume that the vector of linking copula parameters is $\theta = (1, 1, 1, 2.5, 2.5, 2.5, 2.5, 1.5, 1.5, 1.5)^T$. We also assume that the correlation matrix has an autoregressive structure with $\Sigma_{j,k} = \rho^{|j-k|}$ and $\rho = 0.5$. To estimate the parameters $\theta$ and $\rho$, we first use the fast approach from Section 4.2
(method 1), and we use the respective estimates as starting values for the pairwise likelihood approach, as shown in Section 4.1 (method 2).

Table 1 shows the root mean squared errors (RMSE) of the parameter estimates $\hat{\theta}$ and $\hat{\rho}$, obtained by the two methods mentioned above for 500 samples of size $n = 200$ and $n = 1000$. The table shows that the fast method produces accurate estimates; the RMSEs of estimates of $\theta$ and $\rho$ are 5–10% and 20–25% higher using this method than with the pairwise likelihood approach. The estimates are less accurate for linking copulas with stronger dependence.

Table 1: RMSEs (multiplied by 100) of the parameter estimates, $\hat{\theta}$ and $\hat{\rho}$, obtained by the two methods. We used 500 samples of size $n = 200$ and $n = 1000$.

| sample size             | RMSE of $\theta$ | RMSE of $\hat{\rho}$ |
|-------------------------|-------------------|------------------------|
| $n = 200$, method 1     | 13 13 12 35 38 36 | 19 19 20 6.0           |
| $n = 200$, method 2     | 13 13 12 31 35 32 | 18 18 19 4.9           |
| $n = 1000$, method 1    | 5.9 5.2 5.3 17 16 15 | 8.0 8.1 8.7 2.5        |
| $n = 1000$, method 2    | 5.7 5.2 5.1 15 14 13 | 7.4 7.7 8.5 2.0        |

5.2 Simulation 2

We simulate a data set from the same extreme-value copula as in Simulation 1 but with $d = 20$, the vector of reflected Clayton linking copula parameters $\theta = (0.8, 0.8, 0.8, 0.8, 0.8, 1.2, 1.2, 1.2, 1.2, 1.6, 1.6, 1.6, 1.6, 1.6, 2, 2, 2, 2)^\top$, and $\rho = 0.7$. Again, we use methods 1 and 2 to estimate the copula parameters and compute the RMSEs based on these estimates; Table 2 shows the RMSEs of the estimates obtained using these two methods for 500 samples of sizes $n = 200$ and $n = 1000$.

Similar to the first simulation, the RMSEs of estimates of $\theta$ and $\rho$ are 5–15% and 25% higher with the fast method than with the pairwise likelihood approach. Again, the fast method is less accurate if the dependence is stronger. Both methods yield more accurate
estimates of the correlation parameter $\rho$ than those used in Simulation 1, so including more temporal or spatial data locations can improve estimates of the correlation matrix $\Sigma$.

Table 2: RMSEs (multiplied by 100) of the parameter estimates, $\hat{\theta}_i$, $i = 1, \ldots, 10$ (first line) and $i = 11, \ldots, 20$ (second line) and $\hat{\rho}$, obtained by the two methods. We used 500 samples of size $n = 200$ and $n = 1000$.

| sample size        | RMSE of $\theta$ | RMSE of $\hat{\rho}$ |
|--------------------|------------------|-----------------------|
| $n = 200$, method 1| 10 11 11 11 12 17 16 18 17 3.1  | 22 20 21 22 23 29 26 28 31 37 |
| $n = 200$, method 2| 10 11 10 10 11 16 15 16 16 2.5|
| $n = 1000$, method 1| 4.9 4.8 5.0 4.5 4.9 7.1 6.9 6.5 6.5 6.2 1.4| 9.1 9.8 9.8 10 9.6 13 13 12 13 12 |
| $n = 1000$, method 2| 5.1 4.7 4.8 4.6 4.9 6.8 6.9 6.4 6.2 6.0 1.1| 8.9 8.8 8.7 9.0 8.6 11 11 10 11 10 |

5.3 Simulation 3

We simulate a data set from an extreme-value copula (2) with reflected Clayton linking copulas. We use $d = 15$, and the vector of reflected Clayton linking copula parameters is $\theta = (2.5, 2.5, 2.5, 2.5, 2.5, 2.0, 2.0, 2.0, 2.0, 2.0, 1.5, 1.5, 1.5, 1.5, 1.5, 1.5)^T$. We select the correlation matrix $\Sigma$ with a block-diagonal structure that has the off-diagonal elements $\Sigma_{j,k} = \rho_{j,k}$ if $1 \leq j, k \leq 5$, $6 \leq j, k \leq 10$, $11 \leq j, k \leq 15$, and $\Sigma_{j,k} = 0$ otherwise. One can check that $(W_1^T, W_2^T, W_3^T)^T \sim N_{15}(0, \Sigma)$, where $W_g^T = (W_{g,1}, \ldots, W_{g,5})$ and

$$W_{g,i} = \rho_{g,i}Z_g + \left(1 - \rho_{g,i}^2\right)^{1/2}Z_{g,i}, \quad g = 1, 2, 3, \quad i = 1, \ldots, 5,$$

where $Z_1, Z_2, Z_3, Z_{1,1}, \ldots, Z_{3,5} \sim \text{i.i.d. } N(0,1)$. It implies that this correlation structure corresponds to the joint dependence of normal random variables from three independent groups, with the one-factor structure in each group. The respective copula (2) is an extreme-value limit of the bifactor copula model (Krupskii and Joe, 2015) with the reflected Clayton (nor-
mal) copulas linking the common (group-specific) factors, respectively, and the observed variables.

We use \( \rho_1 = (0.8, 0.6, 0.4, 0.2, 0.0)^\top \), \( \rho_2 = (0.4, 0.4, 0.4, 0.4, 0.4)^\top \) and \( \rho_3 = (0.0, 0.2, 0.4, 0.6, 0.8)^\top \), where \( \rho_g = (\rho_{g,1}, \ldots, \rho_{g,5})^\top \). Similar to the previous two simulations, we use methods 1 and 2 to estimate the parameters \( \theta, \rho_1, \rho_2 \) and \( \rho_3 \). Table 3 shows the RMSEs of the estimates obtained by the two methods for 500 samples of size \( m = 200 \) and \( n = 1000 \).

Table 3: RMSEs (multiplied by 100) of the parameter estimates, \( \hat{\theta} \) (first line) and \( \hat{\rho} = (\hat{\rho}_1^\top, \hat{\rho}_2^\top, \hat{\rho}_3^\top)^\top \) (second line), obtained by the two methods. We used 500 samples of size \( n = 200 \) and \( n = 1000 \).

| sample size | RMSE of \( \theta \) (top line) and \( \hat{\rho} \) (bottom line) |
|-------------|---------------------------------------------------------------|
| \( n = 200, \) method 1 | 30 33 32 32 33 27 24 26 26 25 18 18 18 19 17 |
|              | 16 15 14 18 21 22 21 21 21 19 18 15 15 16 17 |
| \( n = 200, \) method 2 | 29 32 32 32 33 26 23 25 25 25 18 19 18 18 18 |
|              | 15 14 14 17 19 22 19 23 20 20 15 15 12 14 15 |
| \( n = 1000, \) method 1 | 7.6 6.3 6.2 7.3 9.2 8.6 8.8 8.5 8.3 8.4 7.5 6.4 6.0 6.9 8.5 |
| \( n = 1000, \) method 2 | 14 13 13 14 13 10 11 11 10 11 7.9 7.7 8.0 7.6 7.9 |
|              | 6.6 5.4 5.2 6.4 7.5 7.5 7.3 7.3 7.0 7.1 6.2 5.8 5.3 5.8 6.9 |

The RMSEs of the estimates of \( \theta \) and \( \rho \) are 5% and 10–20% higher for the fast method compared to the pairwise likelihood approach. Estimates of \( \theta \) are less accurate if the dependence between the common factor and observed variables is stronger. Estimates of \( \rho \) are less accurate than estimates of \( \theta \) but the accuracy improves with a larger sample size.

5.4 Simulation 4

We simulate a data set from an extreme-value copula such that \( \Sigma \) is a matrix of ones, and we use the Clayton linking copulas with continuous densities on \((0, 1]^2\). We use \( d = 20 \), and the vector of Clayton linking copula parameters \( \theta = (0.8, 0.8, 0.8, 0.8, 0.8, 1.2, 1.2, 1.2, 1.2, 1.2, 1.2, 1.6, 1.6, 1.6, 1.6, 1.6, 2, 2, 2, 2, 2)^\top \).
We use method 1 to estimate the copula parameters; see Appendix A.6 for details. The model is not identifiable and we therefore fix the first parameter at 0.8. Table 4 shows RMSEs of the copula parameter estimates obtained by method 1 for 500 samples of size $n = 200$ and $n = 1000$.

Table 4: RMSEs (multiplied by 100) of the parameter estimates, $\hat{\theta}_i$, $i = 1, \ldots, 10$ (first line) and $i = 11, \ldots, 20$ (second line), obtained by method 1. We used 500 samples of size $n = 200$ and $n = 1000$.

| sample size | RMSE of $\hat{\theta}$ |
|-------------|-------------------------|
| $n = 200$   | 0.0 0.0 0.0 0.0 5.6 5.6 5.6 5.6 9.8 9.8 9.8 9.8 15 15 15 15 15 15 15 15 |
| $n = 1000$  | 0.0 0.0 0.0 0.0 2.3 2.3 2.3 2.3 4.6 4.6 4.6 4.6 6.8 6.8 6.8 6.8 6.8 6.8 6.8 6.8 |

Once the first parameter is fixed, the remaining parameters can be identified. The estimates are less accurate for linking copulas with stronger dependence but RMSEs are quite small even for a small sample size $n = 200$.

6 Empirical Studies

In this section, we apply the conditional normal extreme-value copulas to analyze two data sets. The first data set consists of the monthly maxima of daily average wind speed data, and the second data set is made of the monthly minima of daily stock log-returns.

6.1 Wind data

We consider monthly maxima of the average daily wind speed at a height of 10 meters above the ground recorded by 12 weather stations (Acme, Burneyville, Byars, Centrahoma, Chickasha, Ketchum Ranch, Lane, Ringling, Shawnee, Sulphur, Washington, and Waurika) in Oklahoma state, United States. The data are available at the mesonet.org website.
The stations are located within a small geographical region, the maximum distance between two stations being 194 km, and at an elevation between 181 and 430 m. We therefore assume spatial stationarity for these data. The average wind speed is higher in winter and spring due to thunderstorms, which produce gusty winds, and it is lower in summer when the weather is more settled. We therefore include three summer months (June, July and August) from 2000 to 2019, 60 months in total.

We use model (2) with the reflected Clayton copulas with the same parameter \( \theta \) and spatial covariance matrix \( \Sigma \) with \( \Sigma_{j,k} = (1 - \eta) \exp(-\gamma d_{j,k}^\alpha) \), where \( 0 < \eta < 1 \) is the nugget effect, \( \gamma > 0 \) and \( 0 < \alpha \leq 2 \) are parameters of the powered-exponential covariance function, and \( d_{j,k} \) is the distance (in km) between the \( j \)th and \( k \)th stations.

We transform the measurements taken at each station into uniform data using nonparametric ranks, and we estimate the model parameters using the pairwise likelihood approach. We get the following estimates:

\[
\hat{\theta} = 2.82, \quad \hat{\eta} = 0.49, \quad \hat{\gamma} = 0.063, \quad \hat{\alpha} = 2.
\]

Next, we compute the empirical and model-based estimates of Spearman’s correlation \( S_\rho \) and the upper tail dependence coefficient \( \lambda_U \) for different pairs of stations. For the model-based estimates, the formula for \( V_{j,k}(w_j, w_k) \) can be used to compute \( \lambda_U = 2 - V_{j,k}(1,1) \) for the \((j, k)\) pair of variables; see Appendix A.6. There is no explicit formula for \( S_\rho \), so we simulate a sample of size 10000 from the estimated model to compute the model-based estimates of correlations. The empirical and model-based estimates of \( S_\rho = S_\rho(d) \) and \( \lambda_U = \lambda_U(d) \) are plotted against distance \( d \) in Figure 1.

As shown in Figure 1, the conditional normal extreme-value copula fits the data well. The relationship between \( \lambda_U(d) \) and \( d \) is approximately linear for a small \( d \), which suggests that \((\eta + Cd^\alpha)^{1/2} = \eta^{1/2} + O(d) \) or \( \alpha = 2 \) under the assumptions of Proposition 6. The
estimate $\hat{\alpha} = 2$ is therefore in agreement with the result of this proposition.

### 6.2 Stock return data

We use monthly minima of daily stock log-returns from the S&P 500 index. We include stocks from three sectors: seven stocks from the industrial machinery sector with tickers \texttt{CMI}, \texttt{DOV}, \texttt{GWW}, \texttt{ITW}, \texttt{PH}, \texttt{SNA}, \texttt{SWK}; ten stocks from the electric utilities sector with tickers \texttt{AEP}, \texttt{ED}, \texttt{D}, \texttt{DUK}, \texttt{EIX}, \texttt{ETR}, \texttt{FE}, \texttt{PPL}, \texttt{PEG}, \texttt{SO}; and seven stocks from the regional banks sector with tickers \texttt{FITB}, \texttt{HBAN}, \texttt{KEY}, \texttt{MTB}, \texttt{PNC}, \texttt{RF}, \texttt{TFC}. The study period is 2000–2019, and we exclude 2007–2009, when the US subprime mortgage crisis severely affected the US market.

We use the model (2) with the reflected Clayton copulas with a vector of parameters $\theta$ and the covariance matrix $\Sigma$ with a block-diagonal structure with a vector of correlation
parameters $\rho$ and the off-diagonal elements $\Sigma_{j,k} = \rho_j \rho_k$ if a pair $(j, k)$ is from the same sector, and $\Sigma_{j,k} = 0$ otherwise. This corresponds to a model with the bifactor structure as in simulation 3 in Section 5. Bifactor models can handle data with factor structures, and can be suitable for modeling financial data with stock returns affected by global macroeconomic and sector-specific factors.

We transform the log-returns minima to uniform data using nonparametric ranks and fit the proposed model to the reflected data using the pairwise likelihood approach. The parameter estimates for the three groups are

$$\hat{\theta}_{1:7} = (1.07, 1.54, 0.99, 1.37, 1.31, 1.41, 1.06)^\top,$$
$$\hat{\rho}_{1:7} = (0.47, 0.14, 0.00, 0.76, 0.60, 0.26, 0.63)^\top,$$
$$\hat{\theta}_{8:17} = (0.61, 0.50, 0.59, 0.46, 0.51, 0.57, 0.51, 0.73, 0.57, 0.51)^\top,$$
$$\hat{\rho}_{8:17} = (0.93, 0.92, 0.92, 0.97, 0.89, 0.91, 0.80, 0.90, 0.88, 0.93)^\top,$$
$$\hat{\theta}_{18:24} = (1.06, 1.10, 1.13, 1.12, 1.45, 0.82, 1.08)^\top,$$
$$\hat{\rho}_{18:24} = (0.83, 0.90, 0.90, 0.75, 0.77, 0.95, 0.86)^\top.$$

The results indicate that the dependence between the observed variables and the common factor is stronger in the first and third sectors, while the group-specific factors are strongly correlated with the variables from the second and third groups. This implies that sectors 1 and 3 are strongly dependent on each other and that the residual dependence of these data is very strong, so a model with only one common factor is not suitable for the monthly minima of log-returns.

Similar to the wind data, we compute the empirical and model-based estimates of Spearman’s correlation $S_\rho$ and the upper tail dependence coefficient $\lambda_U$ for different pairs of
variables from the same sectors and from different sectors; see Figure 2.

Figure 2: Empirical and estimated correlations, $S_\rho(d)$ (first two rows, black and green points), and upper tail dependence coefficients, $\lambda_U(d)$ (last two rows, black and red points), for different pairs of stock returns from sector 1 (top left), sector 2 (top middle), sector 3 (top right), and sectors 1, 2 (bottom left), sectors 1, 3 (bottom middle), sectors 2, 3 (bottom right)
As shown in Figure 2, dependence within the sector, as measured by $S_{\rho}$ and $\lambda_U$, is stronger than dependence between the sectors, as expected. The dependence between sectors 1 and 3 is slightly weaker than the dependence within sector 1, and the dependence between sectors 1 and 2 and between sectors 2 and 3 is significantly weaker than the dependence within respective sectors, pointing to a weaker dependence between sectors 1 and 2, and sectors 2 and 3. The results also indicate that the proposed copula model with factor structure fits the data very well.

7 Discussion

In this paper, we considered a class of extreme-value copula models that are extreme-value limits of factor copula models with residual dependence modeled by a Gaussian copula. These are flexible models for multivariate extremes with complex dependence structures, such as spatial extremes or multivariate extremes with factor structures. Parsimonious dependence structures can be obtained with the appropriate bivariate linking copulas and the covariance matrix $\Sigma$ of the Gaussian copula. These models are computationally feasible in high dimensions, as only one-dimensional integration is required to compute the bivariate copula density.

We used reflected Clayton linking copulas in the empirical study, but different linking copulas can be used to increase the flexibility of the proposed class of models. Also, the underlying dependence structure is assumed to be known (e.g., the partition of financial log-returns into groups/sectors). However, the underlying dependence structure is not always known, so model selection procedures and goodness-of-fit tests are topics requiring future research. Another topic for further research is to propose efficient estimation methods for this class of models involving higher-dimensional marginals and some computationally tractable special cases.
Appendix

A.1 Proof of Proposition 3

With $\gamma_j = \frac{1}{k+1} A_j'(\frac{k}{k+1}) + A_j(\frac{k}{k+1})$ and $\eta_j(k) = (k+1)A_j(\frac{k}{k+1}) - k$, we find that $C_{j_{10}}(u|u^k) = \gamma_j u^{\eta_j(k)}$, $j = 1, \ldots, d$, and

$$C_U(u) = \int_0^\infty C_N\{C_{1_{10}}(u|u^k), \ldots, C_{d_{10}}(u|u^k)\} u^k \ln u \, dk$$

$$\leq \int_0^\infty C_N\{\gamma^* u^{\eta^*(k)}, \ldots, \gamma^* u^{\eta^*(k)}\} u^k \ln u \, dk \sim_{u \to 0} \int_0^\infty (\gamma^* u)^{\kappa \cdot \eta^*(k) + k} \ell(u) \, dk,$$

where $\ell(u)$ is a slowly varying function and $\gamma^* = \max_j \gamma_j$, $\eta^*(k) = \min_j \eta_j(k)$. It is seen that $\kappa_L \geq \min_{k \geq 0}\{\kappa \Sigma \cdot \eta^*(k) + k\} > 1$ because $\eta^*(k) \geq \max(0, 1-k)$ and $\eta^*(1) = \min_j \kappa_j - 1 > 0$.

Similarly, one can show that $\kappa_L \leq \min_{k \geq 0}\{\kappa \Sigma \cdot \eta^{**}(k) + k\} \leq \kappa \Sigma \cdot \eta^{**}(0) = \kappa \Sigma$, where $\eta^{**}(k) = \max_j \eta_j(k)$. This implies that $C_U$ has intermediate lower tail dependence. \hfill \Box

A.2 Proof of Proposition 4

Let $C_{j_{10}}(u_j|u_0) = 1 - C_{j_{10}}(u_j|1 - u_0)$, $j = 1, \ldots, d$. Similar to Proposition 1, we use Theorem 8.76 of Joe (2014):

$$V(w_1, \ldots, w_d) = \lim_{u \to 0} \frac{1}{u} \left[ 1 - \int_0^1 C_N\{1 - C_{1_{10}}(uw_1|w_0), \ldots, 1 - C_{d_{10}}(uw_d|w_0); \Sigma\} \, dw_0 \right]$$

$$= \lim_{u \to 0} \int_0^{1/u} \left[ 1 - C_N\{1 - C_{1_{10}}(uw_1|w_0), \ldots, 1 - C_{d_{10}}(uw_d|w_0); \Sigma\} \right] \, dw_0$$

$$= \int_0^\infty \left[ 1 - C_N\{1 - b_{1_{10}}(w_1|w_0), \ldots, 1 - b_{d_{10}}(w_d|w_0); \Sigma\} \right] \, dw_0. \hfill \Box$$

A.3 Proof of Proposition 5

We have $C_{j_{10}}(1 - uw_j|v_0) = 1 - uw_j c_{j,0}(1, v_0) + o(u)$, $j = 1, \ldots, d$, and therefore

$$V(w_1, \ldots, w_d) = \lim_{u \to 0} \frac{1}{u} \int_0^1 \left[ 1 - \min_j \{C_{j_{10}}(1 - uw_j|v_0)\} \right] \, dv_0$$

$$= \lim_{u \to 0} \frac{1}{u} \int_0^1 \left[ uw_j \max_j \{c_{j,0}(1, v_0) + o(1)\} \right] \, dv_0 = \int_0^1 \max_j \{wc_{j,0}(1, v_0)\} \, dv_0. \hfill \Box$$
A.4 Proof of Proposition 6

We use the following Lemma to prove this proposition.

Lemma 1: Let $C_N(u, v; \rho)$ be the normal copula with correlation $\rho$. If $\rho \to 1$, then

$$C_N(u, u; \rho) = u - \left(\frac{1-\rho}{\pi}\right)^{1/2} \phi\{\Phi^{-1}(u)\} + (1-\rho)^{3/2} \phi\{\Phi^{-1}(u)\} \{K_1(\rho) + K_2(\rho)\{\Phi^{-1}(u)\}^2\},$$

where $\max_{\rho} |K_i(\rho)| \leq K_0 < \infty$, $i = 1, 2$.

Proof of Lemma 1: Denote $C_N(u|v; \rho) = \partial C_N(u, v; \rho)/\partial v$ and $\delta = \sqrt{1-\rho}$. We find that

$$C_N(u, u; \rho) = 2 \int_0^u C_N(v|v; \rho) dv = 2 \int_0^u \Phi\left\{\frac{\delta}{\sqrt{2-\delta^2}} \Phi^{-1}(v)\right\} dv.$$

Let $h(\delta) = h(\delta, v) = \Phi\left\{\frac{\delta}{\sqrt{2-\delta^2}} \Phi^{-1}(v)\right\}$ and $g(\delta) = g(\delta, v) = \phi\left\{\frac{\delta}{\sqrt{2-\delta^2}} \Phi^{-1}(v)\right\}$. Using the Taylor expansion of $h(\delta)$ around $\delta = 0$ with a fixed $v$ yields:

$$h(\delta) = h(0) + h'(0)\delta + h''(\Upsilon)\delta^2, \quad 0 < \Upsilon < \delta,$$

where

$$h'(t) = \frac{2\Phi^{-1}(v)}{(2-t^2)^{1.5}} \cdot g(t), \quad h''(t) = -\frac{2t\Phi^{-1}(v)}{(2-t^2)^{3.5}} \cdot \left[2\{\Phi^{-1}(v)\}^2 - 3(2-t^2)\right] g(t).$$

It implies that

$$h(\delta) = 0.5 + 0.5\pi^{-1/2}\delta\Phi^{-1}(v) + w_1\delta^3\Phi^{-1}(v) - w_2\delta^5\{\Phi^{-1}(v)\}^3,$$

where $0 < w_1 < 6\phi(0)$ and $0 < w_2 < 4\phi(0)$ for any $0 < v < 1$ and $\delta \to 0$; hence,

$$C_N(u, u; \rho) = 2 \int_0^u h(\delta, v) dv = u + \pi^{-1/2}\delta I_1 + 2w_1\delta^3 I_1 - 2w_2\delta^5 I_2,$$

where

$$I_1 = \int_0^u \Phi^{-1}(v) dv = -\phi\{\Phi^{-1}(u)\}, \quad I_2 = \int_0^u \{\Phi^{-1}(v)\}^3 dv = -[2 + \{\Phi^{-1}(u)\}^2] \phi\{\Phi^{-1}(u)\}. $$
Finally,
\[
C_N(u, u; \rho) = u - \pi^{-1/2} \delta \phi \{ \Phi^{-1}(u) \} + \delta^3[K_1 + K_2(\Phi^{-1}(u))^2] \phi \{ \Phi^{-1}(u) \},
\]
where \( K_1 = -2w_1 + 4w_2 \), \( K_2 = 2w_2 \) and \(|K_1| < 24\phi(0)\), \( K_2 < 12\phi(0) \).

\[\square\]

**Proof of Proposition 6**: Note that \( \phi \{ \Phi^{-1}(u) \} \sim u\ell^*(u) \) as \( u \to 0 \), where \( \ell^*(u) \) is a slowly varying function. It implies that for any \( m \geq 0 \),
\[
\int_0^\infty [\Phi^{-1}(b_{1|0}(1|w_0))]^m \phi \{ \Phi^{-1}(b_{1|0}(1|w_0)) \} \, dw_0 < \infty.
\]

From Proposition 4 and Lemma 1, we find that, as \( \rho \to 1 \),
\[
V(1, 1) = \int_0^\infty [1 - C_N(1 - b_{1|0}(1|w_0), 1 - b_{1|0}(1|w_0); \rho)] \, dw_0 \\
= 2 - \int_0^\infty C_N(b_{1|0}(1|w_0), b_{1|0}(1|w_0); \rho) \, dw_0 \\
= 2 - \int_0^\infty b_{1|0}(1|w_0) \, dw_0 + \left( \frac{1 - \rho}{\pi} \right)^{1/2} \int_0^\infty \phi \{ \Phi^{-1}(b_{1|0}(1|w_0)) \} \, dw_0 + O((1 - \rho)^{3/2}) \\
= 1 + \left( \frac{1 - \rho}{\pi} \right)^{1/2} \int_0^\infty \phi \{ \Phi^{-1}(b_{1|0}(1|w_0)) \} \, dw_0 + O((1 - \rho)^{3/2}). \quad \square
\]

**A.5 Computation of \( V_{j,k}(w_j, w_k) \)**

Let \( I_{j,k}(w_0) = 1 - C_N(1 - b_{j|0}(w_j|w_0), 1 - b_{k|0}(w_k|w_0); \rho_{j,k}) \). We assume that, as \( w_0 \to \infty \),
\( b_{j|0}(w_j|w_0) = \ell_j(w_0)w_0^{-\phi_j} + o(w_0^{-\phi_j}) \) and \( b_{k|0}(w_k|w_0) = \ell_k(w_0)w_0^{-\phi_k} + o(w_0^{-\phi_k}) \), \( \phi_j, \phi_k > 1 \),
where \( \ell_j \) and \( \ell_k \) are slowly varying functions. It follows that \( I_{j,k}(w_0) = \ell_{j,k}(w_0)w_0^{-\phi_{j,k}} \) as \( w_0 \to \infty \) where \( \phi_{j,k} = \min(\phi_j, \phi_k) \) and \( \ell_{j,k} \) is a slowly varying function. The integrand
\( I_{j,k}(w_0) \) has a slow rate of decay in the tail and standard numerical integration methods
used to compute \( V_{j,k}(w_j, w_k) = \int_0^\infty I_{j,k}(w_0) \, dw_0 \) may not be efficient.

To make computations more efficient, we can write
\[
V_{j,k}(w_j, w_k) = \int_0^1 I_{j,k}(w_0) \, dw_0 + \int_1^\infty I_{j,k}(w_0) \, dw_0 = \int_0^1 I_{j,k}(w_0) \, dw_0 + \alpha \int_0^1 I_{j,k}(z_0^{-\alpha})z_0^{-\alpha-1} \, dz_0,
\]
where \( z_0 = \frac{w_0}{\alpha} \).
where the first integral has finite integration limits and bounded integrand. The second integrand can take large values if \( z_0 \) is close to zero. We therefore need to select the smallest \( \alpha > 0 \) such that \( I_{j,k}^*(z_0) = I_{j,k}(z_0^{-\alpha}z_0^{-\alpha-1}) < \infty \) for \( 0 \leq z_0 \leq 1 \). We have \( I_{j,k}^*(z_0) = \ell_{j,k}(z_0^{-\alpha})z_0^{(\phi_{j,k}-1)\alpha-1} \) as \( z_0 \to 0 \) and one can select \( \alpha = \alpha_{j,k} = \{\phi_{j,k} - 1\}^{-1} \). Now Gauss-Legendre quadrature (Stroud and Secrest, 1966) can be used to evaluate the two integrals and compute \( V_{j,k}(w_j, w_k) \).

The assumption about the tail behavior of \( b_{j|0} \) holds for many copulas with the upper tail dependence. For the reflected Clayton copula with parameter \( \theta_j \),

\[
b_{j|0}(w_j|w_0) = \left\{1 + (w_0/w_j)^{\theta_j}\right\}^{-1-1/\theta_j} = (w_0/w_j)^{-1-\theta_j} + o(w_0^{-1-\theta_j}), \quad \text{as } w_0 \to \infty,
\]

and therefore \( \alpha_{j,k} = 1/\min(\theta_j, \theta_k) \). For the Gumbel copula copula with parameter \( \theta_j \),

\[
b_{j|0}(w_j|w_0) = 1 - \left\{1 + (w_j/w_0)^{\theta_j}\right\}^{-1+1/\theta_j} = (1 - 1/\theta_j)(w_0/w_j)^{-\theta_j} + o(w_0^{-\theta_j}), \quad \text{as } w_0 \to \infty,
\]

and therefore \( \alpha_{i,k} = 1/\{\min(\theta_j, \theta_k) - 1\} \).

Similar ideas can be used to compute the derivatives of the exponent function. We found that the same transformation works very well in this case and that this change of variables greatly improves the accuracy of numerical integration and \( n_q = 35 \) quadrature points is sufficient to compute the density \( c_{j,k}^{EV} \) in most cases.

### A.6 Parameter estimation for \( C_{UV}^{EV} \) with Clayton linking copulas

Here we provide more details for \( C_{UV}^{EV} \) in Section 4.2 with Clayton linking copulas. Note that \( V_{j,k}(1, 1) = 1 \) if \( \theta_j = \theta_k \). Without loss of generality, we now assume that \( \theta_j > \theta_k \) for the \((j, k)\) margin. We have \( c(1, v; \theta) = (\theta + 1)v^\theta \) and

\[
V_{j,k}(1, 1) = \int_0^{v_{j,k}} c(1, v_0; \theta_k)dv_0 + \int_{v_{j,k}}^1 c(1, v_0; \theta_j)dv_0 = 1 - C(v_{j,k}|1; \theta_j) + C(v_{j,k}|1; \theta_k),
\]
where the conditional Clayton copula $C(v|1; \theta) = v^{\theta+1}$ and $v_{j,k} \in (0, 1)$ satisfies

$$c(1, v_{j,k}; \theta_j) = c(1, v_{j,k}; \theta_k) \Rightarrow v_{j,k} = \left( \frac{\theta_k + 1}{\theta_j + 1} \right)^{\frac{1}{\theta_j - \theta_k}},$$

and therefore

$$V_{j,k}(1, 1) = 1 + \frac{(\vartheta - 1)^{\vartheta-1}}{\vartheta^\vartheta}, \quad \vartheta = \frac{\theta_j + 1}{\theta_j - \theta_k}.$$

Similarly, one can show that the copula $C_{U}^{EV}$ and its lower-dimensional margins depend on $\vartheta(\theta_j, \theta_k) = (\theta_j + 1)/(\theta_j - \theta_k)$, or, equivalently, on $\vartheta^*(\theta_j, \theta_k) = \{\vartheta(\theta_j, \theta_k) - 1\}/\vartheta(\theta_j, \theta_k) = (\theta_k + 1)/(\theta_j + 1)$ for $1 \leq k < j \leq d$. The model is therefore non-identifiable and one can fix one parameter and estimate the remaining $d - 1$ parameters.

If the order of variables is ignored, the model is still non-identifiable even if one parameter is fixed.

**Example 5:** Assume that $d = 3$ and $\theta = (0.5, 1, 2)\top$. We generate a sample of size $N = 100$ from $C_{U}^{EV}$ assuming $\Sigma$ is a matrix of ones. We fix $\theta_2 = 1$ and find that the objective function (4) attains its minimum at $\hat{\theta} = (0.456, 1, 2.584)\top$ and $\tilde{\theta} = (1.747, 1, 0.116)\top$. We can see that

$$\frac{\hat{\theta}_1 + 1}{\hat{\theta}_2 + 1} = \frac{\hat{\theta}_2 + 1}{\hat{\theta}_1 + 1}, \quad \frac{\hat{\theta}_1 + 1}{\hat{\theta}_3 + 1} = \frac{\hat{\theta}_3 + 1}{\hat{\theta}_1 + 1}, \quad \frac{\hat{\theta}_2 + 1}{\hat{\theta}_3 + 1} = \frac{\hat{\theta}_3 + 1}{\hat{\theta}_2 + 1}.$$

To select the right solution, one can check bivariate scatter plots: if $\theta_j > \theta_k$ for the $(j, k)$ margin, the marginal density is zero around the corner $(0, 1)$ and the density is skewed to the lower right corner. Figure 3 shows scatter plots for the simulated data set. The plots indicate that $\theta_1 < \theta_2 < \theta_3$ and therefore $\hat{\theta}$ should be selected.

**References**

Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44, 182–198.
Figure 3: Bivariate scatter plots for the data set simulated from $C_{UV}^E$ with degenerate $\Sigma$ (matrix of ones) and Clayton linking copulas with $\theta = (0.5, 1, 2)^T$.

Brechmann, E.C., Czado, C., Aas, K., 2012. Truncated regular vines in high dimensions with applications to financial data. Canadian Journal of Statistics 40, 68–85.

Demarta, S., McNeil, A.J., 2005. The t copula and related copulas. Int. Stat. Rev. 73, 111–129.

Ferreira, M., 2013. Nonparametric estimation of the tail-dependence coefficient. REVSTAT – Statistical Journal 11, 1–16.

Genest, C., Ghoudi, K., Rivest, L.P., 1995. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika 82, 543–552.

Genton, M. G., Kleiber, W., 2015. Cross-covariance functions for multivariate geostatistics (with discussion). Statistical Science 30, 147–163.

Gneiting, T., Genton, M. G., Gutttorp, P., 2007. Geostatistical space-time models, stationarity, separability and full symmetry. In Finkenstaedt, B., Held, L. and Isham, V.(eds),
Statistics of Spatio-Temporal Systems, Chapman & Hall / CRC Press, Monograph in Statistics and Applied Probability, Boca Raton.

de Haan, L., Ferreira, A., 2006. Extreme Value Theory. Springer-Verlag, New York.

Hua, L., Joe, H., 2011. Tail order and intermediate tail dependence of multivariate copulas. Journal of Multivariate Analysis 102, 1454–1471.

Hüsler, J., Reiss, R. D., 1989. Maxima of normal random vectors: between independence and complete dependence. Statistics and Probability Letters 7, 283–286.

Joe, H., 2014. Dependence Modeling with Copulas. Chapman & Hall/CRC, Boca Raton, FL.

Krupskii, P., Genton, M. G., 2018. Linear factor copula models and their properties. Scandinavian Journal of Statistics 45(4), 861–878.

Krupskii, P., Huser, R., Genton, M. G., 2018a. Factor copula models for replicated spatial data. Journal of the American Statistical Association 521, 467–479.

Krupskii, P., Joe, H., 2013. Factor copula models for multivariate data. Journal of Multivariate Analysis 120, 85–101.

Krupskii, P., Joe, H., 2015. Structured factor copula models: theory, inference and computation. Journal of Multivariate Analysis 138, 53–73.

Krupskii, P., Joe, H., 2020. Flexible copula models with dynamic dependence and application to financial data. Journal of Econometrics and Statistics. Accepted.
Krupskii, P., Joe, H., Lee, D., Genton, M. G., 2018b. Extreme value limit of the convolution of exponential and multivariate normal distributions: Link to the Hüsler-Reiss distribution. Journal of Multivariate Analysis 163, 80–95.

Kurowicka, D., Cooke, R., 2006. Uncertainty Analysis with High Dimensional Dependence Modelling. Wiley Series in Probability and Statistics.

Lee, D., Joe, H., 2018. Multivariate extreme value copulas with factor and tree dependence structures. Extremes 21(1), 147–176.

Lindsay, B., 1998. Composite likelihood methods. Contemporary Mathematics 80, 220–239.

McNeil, A.J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management. Princeton University Press, Princeton, NJ.

Nikoloulopoulos, A. K., Joe, H., Li, H., 2009. Extreme value properties of multivariate t copulas. Extremes 12, 129–148.

Oh, D. H., Patton, A., 2017. Modeling dependence in high dimensions with factor copulas. Journal of Business and Economic Statistics 35, 139–154.

Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. Institute of Statistics of the University of Paris 8, 229–231.

Stroud, A., Secrest, D., 1966. Gaussian Quadrature Formulas. Prentice-Hall, Englewood Cliffs, NJ.

Yoshida, T., 2018. Maximum likelihood estimation of skew-t copulas with its applications to stock returns. Journal of Statistical Computation and Simulation 88(13), 2489–2506.