Astin Bulletin (2024), pp. 1–33
doi:10.1017/asb.2024.29

International Actuarial Association
Association Actuarielle Internationale

RESEARCH ARTICLE

A study of one-factor copula models from a tail dependence perspective

Nariankadu Shyamalkumar[1] and Siyang Tao[2]

[1]Department of Statistics and Actuarial Science, The University of Iowa, Iowa City, 52242, IA, USA
[2]Department of Mathematical Sciences, Ball State University, Muncie, 47306, IN, USA
Corresponding author: Siyang Tao; Email: tao@bsu.edu

Received: 16 January 2024; **Revised:** 14 July 2024; **Accepted:** 18 July 2024

Keywords: Tail dependence coefficient; tail dependence matrix; correlation polytope; BB1 family

Abstract

Modeling multivariate dependence in high dimensions is challenging, with popular solutions constructing multivariate copula as a composition of lower dimensional copulas. Pair-copula constructions do so by using bivariate linking copulas, but their parametrization, in size, being quadratic in the dimension, is not quite parsimonious. Besides, the number of regular vines grows super-exponentially with the dimension. One parsimonious solution is factor copulas, and in particular, the one-factor copula is touted for its simplicity – with the number of parameters linear in the dimension – while being able to cater to asymmetric non-linear dependence in the tails. In this paper, we add nuance to this claim from the point of view of a popular measure of multivariate tail dependence, the tail dependence matrix (TDM). We focus on the one-factor copula model with the linking copula belonging to the BB1 family, pointing out later the applicability of our results to a wider class of linking copulas. For this model, we derive tail dependence coefficients and study their basic properties as functions of the parameters of the linking copulas. Based on this, we study the representativeness of the class of TDMs supported by this model with respect to the class of all possible TDMs. We establish that since the parametrization is linear in the dimension, it is no surprise that the relative volume is zero for dimensions greater than three, and hence, by necessity, we present a novel manner of evaluating the representativeness that has a combinatorial flavor. We formulate the problem of finding the best representative one-factor BB1 model given a target TDM and suggest an implementation along with a simulation study of its performance across dimensions. Finally, we illustrate the results of the paper by modeling rainfall data, which is relevant in the context of weather-related insurance.

1. Introduction

Multivariate normal distribution, the classical model for incorporating dependence, while able to accommodate any valid correlation matrix, by its very nature, imposes normal marginals. Use of copulas, see Sklar (1959), Embrechts *et al*. (2011), Jaworski *et al*. (2010), Joe (2014), Nelsen (2006), and Salvadori *et al*. (2007), was driven by the desire to decouple modeling of the marginals and dependence. An early article, likely the foremost in the actuarial literature, exploring the use of copulas in actuarial applications is that of Frees and Valdez (1998), with the use of copulas now well established as the primary tool for dependence modeling – see Bassamboo *et al*. (2008) (credit risk), Chen *et al*. (2015) (mortality risks), Hua *et al*. (2017) (spatial dependence), Hsieh *et al*. (2021) (mortality risks), Kularatne *et al*. (2021) (general and life insurance), Li *et al*. (2021) (mortality risks), and Oh *et al*. (2021) (automobile insurance).

While Gaussian copula has found many successful uses, see Othus and Li (2010), Renard and Lang (2007), Song *et al*. (2009), and Van de Vyver and Van den Bergh (2018), a key feature of Gaussian dependence is that while it inherits the ability to model a wide range of central dependence from the multivariate normal distribution it also inherits tail independence. Li (2000) is a pioneering paper that

introduced the use of the Gaussian copula for modeling credit curves to incorporate default correlation between credits, with applications in the valuation of credit default swaps. While its simplicity made it wildly popular, its very popularity and simplicity, coupled with its inability to capture tail dependence, made its (ab)use far beyond its original intended purpose a contributor to the global financial crisis of 2008, see Donnelly and Embrechts (2010) and Salmon (2009). The importance of incorporating tail dependence in modeling has been noticed in many applications, see Blanchet and Davison (2011), Brechmann *et al*. (2012), Davis and Mikosch (2009), and Da Silva *et al.* (2014).

One requires a measure of tail dependence for modeling in the presence of tail dependence. The popular Pearson's correlation coefficient, a measure of central dependence, fails to capture tail dependence see Figure 7.4 of McNeil *et al*. (2015). Among the various bivariate tail dependence measures, the tail dependence coefficient is the most widely used measure in the literature. For a random vector (X_1, X_2), with copula $C(\cdot, \cdot)$, the lower tail dependence coefficient, $\chi(X_1, X_2)$, is defined as

$$\chi(X_1, X_2) := \lim_{u \downarrow 0} \frac{C(u, u)}{u},$$

provided that the limit exists, with the zero limit corresponding to tail independence. The tail dependence matrix (TDM), $(\chi(X_i, X_j))_{1 \le i,j \le d}$, is the array of tail dependence coefficients corresponding to a d-dimensional random vector $\mathbf{X}_d = (X_1, \ldots, X_d)$. The TDM serves as a d-dimensional measure of tail dependence, akin to the (linear) correlation matrix for central dependence. We denote the set of all d-dimensional TDMs by $\mathcal{T}_d$.

An extension of the Gaussian copula that allows for tail dependence is the t-copula, though imposing identical dependence in the lower and upper tails. The extension is in the same sense that the multivariate t-distribution is an extension of the multivariate normal distribution, with the latter lying at the $\nu = \infty$ boundary of the former, where ν is the degrees of freedom of a t-distribution. The multivariate t-distribution we refer to is the distribution of the following ratio, where X and Y are independent:

$$\frac{X}{\sqrt{Y/\nu}}, \quad \text{where } X \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \text{ and } Y \sim \chi^2_\nu.$$

The parameters of a t-copula comprise the correlation matrix of the underlying multivariate normal and the degrees of freedom of the chi-square distribution. Since a multivariate normal copula exists corresponding to every valid correlation matrix, a natural question is if the t-copula is similarly exhaustive with respect to $\mathcal{T}_d$. Interestingly, while the set of correlation matrices is neatly characterized as positive semidefinite matrices with a unit diagonal, the characterization of the set of TDMs is far more complex. As observed in Strokorb (2013) and Fiebig *et al*. (2017), $\mathcal{T}_d$ is a $\binom{d}{2}$-dimensional polytope, with explicit facet representations known only for $d \le 6$. Moreover, even determining if a given matrix belongs to $\mathcal{T}_d$ is NP-hard – see Shyamalkumar and Tao (2020) and Janßen *et al*. (2023). So, given the closed-form expression for the tail dependence coefficient of a t-copula, see equation (7.38) in McNeil *et al*. (2015) or equation (4) in Shyamalkumar and Tao (2022), it essentially follows that the t-copula does not cover $\mathcal{T}_d$. Surprisingly, in Shyamalkumar and Tao (2022), it it is shown that for $d = 6$, the relative volume of TDMs supported by t-copula is less than 1% – in this sense, thus providing only a sparse coverage of $\mathcal{T}_d$. Given that the t-copula is parameterized with $O(d^2)$ number of parameters, not to mention that it only represents a small part of $\mathcal{T}_d$, there is interest in the construction and study of models allowing for tail dependence with a more parsimonious parameterization, for example, with the number of parameters being linear in d.

A classical idea from the broader dependence modeling literature is to employ a linear factor structure, $X = L_{d\times p} V_{p\times 1} + \epsilon$, where V is a vector of factors with mean $\mathbf{0}$ and identity variance–covariance matrix, and ϵ, the idiosyncratic component, independent of V, has mean $\mathbf{0}$ and a diagonal variance-covariance matrix $\mathbf{\Lambda}$. Note that this results in the variance–covariance matrix $\Sigma_{d\times d}$ satisfying $\Sigma = LL' + \mathbf{\Lambda}$, and hence parametrized by $d(p+1) = O(d)$ number of parameters. The multivariate normal distribution with the variance-covariance matrix allowing such a decomposition admits such a linear factor structure with $V \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\epsilon \sim \mathcal{N}(\mathbf{0}, \Lambda)$. Similarly, the t-copula with such a correlation matrix

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

admits a near-linear factor structure – the non-linearity is with respect to the chi-square random variable in the denominator.

Such models, even with $p = 1$, find applications in finance and insurance. For example, in portfolio credit risk modeling, the common factor represents the systematic risk, and the reciprocal of the denominator represents the common shock the obligors face. The generality is achieved by relaxing the normality of the systematic and idiosyncratic risk and allowing for a more general class of common shocks. Bassamboo *et al*. (2008) models economies where the dependence amongst obligor defaults is primarily due to common shocks with light-tailed systematic and idiosyncratic risks, whereas Tang *et al*. (2019), Chen *et al*. (2023) allow heavy-tailed systematic and idiosyncratic risks, in particular, to show that during periods of market stress – like recessions – the exogenous common shocks can be dominated by the systematic and idiosyncratic risks.

A generalization of the above near-linear factor structure is the factor copula of Krupskii and Joe (2013), which retains the number of parameters to be $O(d)$. The factor copula model is a conditional independence model given p latent factors, which are taken, without loss of generality, to be standard uniform random variables. Specifically, such a factor copula is the joint distribution of marginally uniform random variable $(U_1, \ldots, U_d)$, specified by conditional independence given the latent factors V_i, $i = 1, \ldots, p$, and the conditional distributions of $(U_i, V_1, \ldots, V_p)$, $i = 1, \ldots, d$, by d many $(p+1)$-dimensional copulas. Clearly, this implies that the parametrization is of size $O(d)$. In many applications, there are natural candidates for factors; for example, in finance and insurance, these could represent systemic risk, market risk, sector risk, and common shock, among others. In such cases, the use of a factor copula is rather natural. But importantly, even otherwise, the factor structure alleviates the paucity of high-dimensional parametric copulas by allowing construction of them using lower-dimensional parametric copulas.

The simplest of factor copula models are the one-factor copula models. Theoretically, a single latent variable model is equivalent to the general p latent variables case, but they are quite distinct from the practical parametric modeling point of view. In the d-dimensional one-factor copula model, U_i, for $i = 1, \ldots, d$, are assumed to be conditionally independent given a latent random variable $V \sim U(0, 1)$. In particular, the one-factor copula can be expressed as

$$C(u_1, \ldots, u_d) = \int_0^1 \prod_{i=1}^{d} C_{i|0}(u_i|v)dv, \tag{1.1}$$

where $C_{i0}(u_i, v)$, for $i = 1, \ldots, d$, are bivariate copulas, and $C_{i|0}(u_i|v) := \partial C_{i0}(u_i, v)/\partial v$, for $i = 1, \ldots, d$, are the conditional distribution functions. Since $C_{i0}(u_i, v)$, for $i = 1, \ldots, d$, link the random variables U_i and V, they are referred to as the *linking copulas* of the one-factor copula models. The popularity of one-factor copulas partly derives from the fact that the collection of parametric bivariate copula families is substantially rich in contrast to those catering to higher dimensions.

We note that, more generally, vine copulas allow for the construction of copula models in higher dimensions, building up by linking together bivariate copulas using the regular vine specification, see Bedford and Cooke (2001, 2002). See Aas *et al*. (2009) See Czado and Nagler (2022), for a discussion of vine copulas in the insurance context, where an application to financial time series in the presence of tail dependence is discussed. One of the challenges of working with vine copula, see Morales-Napoles (2010), is that even after fixing the bivariate linking copula family, the choice of the regular vine tree structure is hard as there are, see Morales-Napoles (2010), as many as $d!2^{(d-2)(d-3)/2-1}$ of them. This is an issue that one-factor copula models avoid at the expense of generality. It is worth noting that the one-factor copula model is a particular case of a C-vine (canonical vine) copula model, where we have conditional independence in the trees below the first, see Figure 1; see Chapter 3 of Joe (2011).

One-factor copulas, and factor copulas in general, have gained importance in actuarial science, as is evident from their use across application areas. Hua *et al*. (2017) employs it for modeling thunderstorm loss data, with an innovative approach to model nonlinear high-dimensional spatial dependence while capturing micro-scale dependence and keeping model complexity low. In property insurance, modeling

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

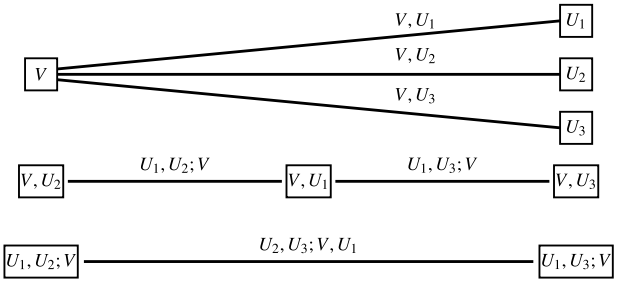


Figure 1. *A C-vine structure in four dimensions.*

hail loss, Gao (2022) introduces a spatial factor copula model to capture spatial dependencies between real estate properties and aspatial dependence induced by the common shock of experiencing the same storm. Chen *et al*. (2015) concerns multi-population mortality modeling, where the first stage of the model caters to the temporal dependence in a population's mortality index, using ARMA and GARCH, and the second stage caters to dependence between the standardized innovations across populations for a fixed time epoch. It is in this second stage that they employ a single-factor copula model. The copula model is that of Oh and Patton (2017), which offers a flexible dependence structure with the ability to model heterogeneity to exchangeability, and tail independence to symmetric or asymmetric tail dependence. In Hsieh *et al*. (2021), the interest is in mortality risk management of insurance policy pools, and it explores the use of factor copulas to model the dependence among the future lifetimes of insured individuals in an insurance policy pool. In particular, this is used to compute the hedge ratio for hedging mortality risk in a block of whole-life policies using a life settlement policy block as a hedge. Beyond mortality modeling, one-factor copula models find application in credit risk management, as demonstrated by Kolman (2014), who developed an alternative model for portfolio credit risk.

From the above, we see that the factor copula models have arisen naturally as parsimoniously parameterized models that accommodate nonlinear dependence and asymmetric tail dependence in high-dimensional data. Also, in a sense, even the simplest of them all, the one-factor copula, has found successful applications in insurance and finance. This prompted the current study to evaluate the range of tail dependence one-factor copula models can cater to with respect to the TDM as the measure of tail dependence. In other words, just the choice of a one-factor copula model implicitly determines a certain form of tail dependence?

Note that modeling with pair-copula constructions, in general, requires the choice of a parametric family for the linking copula. Since the analysis to answer the above-posed question is dependent on the linking copula family, we chose the family of copulas referred to as BB1 (Bivariate-Biparameter), see Joe and Hu (1996) and section 5.2 of Joe (1997). Our choice was driven by Joe (2014) which on page 220, after a comparative study of tail-dependent copulas, observes that *Because of the above properties, for the modeling with vine and factor copulas, it might be sufficient to consider the BB1, reflected BB1, and BB7 for the bivariate copula families with asymmetric lower and upper tail dependence.* Moreover, Joe (2011) on page 184 observes that *Because of the known concordance ordering only for BB1, the BB1 family might be preferable over BB4 and BB7 for getting asymmetric tail dependence with vine copulas.* Also, for actuarial interest, we note that Nikoloulopoulos *et al*. (2012) on page 15 states that *Sensitivity analysis, for extreme value inference of a portfolio, that uses more lower tail dependence than likelihood-based estimates, can be performed with BB1/BB7 copulas (but not t copulas) in the vine, and the forecasting can be substantially improved [sic].*

In the following, we briefly mention some studies where it has found use. Krupskii and Joe (2020) proposes a copula-based model for non-Gaussian data with time-varying dependence. Their model uses a static one-factor model with BB1 linking copula and incorporates time-varying dependence by modeling

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

the conditional distribution given the factor using the Gaussian copula with a time-varying variance-covariance matrix. In financial applications, the market indices are natural choices for the latent factor. In a stress-testing scenario for bond yields and CDS spreads, they find that the use of a tail-dependent copula like BB1 generates more distressed scenarios. Krupskii and Joe (2013) investigates the efficacy of one-factor and two-factor models alongside truncated regular-vine models with various choices for the linking copula. Their one-factor model using the BB1 as the linking copula, and their two-factor model using the BB1 copula for the first factor and the Frank copula for the second, with independence between the two factors, are among the best performers according to the AIC. Righi and Ceretta (2013) considers pair copula construction for modeling multivariate dependence between sector indices in the Brazilian market surrounding the subprime crisis period. They recommend a D-vine over a C-vine structure, and for each of the pair copulas, they considered Student's t, Frank, and many members of the BB and Joe families. They find that for some of the sector index pairs the BB1 was the preferred copula. De Luca *et al*. (2023), again in a pair copula construction setting, find a use for the BB1 copula for studying the influence of economic sectors on European stock returns employing novel tail dependence-based measures of influence. Nonfinancial references using BB1 copula include Lazoglou and Anagnostopoulou (2019) and Lan *et al*. (2021).

In Section 2, for the one-factor model with the BB1 linking copula family, we derive the tail dependence coefficients as a function of the linking copula parameters. Also, we establish some basic properties of this function as it is only implicitly defined. In Section 3, we explore the representativeness of the class of TDMs generated by the one-factor BB1 copula model. Since even with $O(d^2)$-parametrized t-copula leads to vanishing relative volume of the set of TDMs generated, with an $O(d)$-parameterization of our model, the relative volume is bound to be unsatisfactory. While we confirm this, we note that even with zero relative volume, the model could well cater to a wide variety of tail dependencies in an approximate sense. We present a novel manner of evaluating the representativeness of the set of TDMs supported by our model. While unsatisfactory in this new sense as well, it gives the modeler using a one-factor model a way to check the appropriateness of the model for his particular application. In Section 4, we formulate the problem of finding the best representative one-factor BB1 model given a target TDM and suggest an implementation along with a simulation study of its performance across dimensions. In Section 5, we model rainfall data using the one-factor BB1 model to illustrate the use of the results in our study. The analysis presented in this paper is centered around the lower tail of a one-factor copula model utilizing the BB1 family as the linking copula. While this focus is admittedly narrow, we explain in Section 6 that our findings have much broader implications.

Notation: Throughout this paper, all vectors and matrices are represented in boldface. The entry in the i-th row and j-th column of a matrix $\mathbf{A}$ is denoted interchangeably by a_{ij} or $a_{i,j}$. We use $\mathbf{I}$ to represent the identity matrix and $\mathbf{0}$ and $\mathbf{1/J}$ to represent vectors/matrices with all elements equal to 0 and 1, respectively. Since the TDMs are defined by their upper-diagonal elements, the vectorized form of a matrix in this paper refers to the vector consisting of its upper off-diagonal elements. The vectorized form of TDMs is used in the context of describing their geometry, for example, $\mathcal{T}_d$ as a convex polytope. We denote the indicator function of a set A by I_A. When working with matrices, say $\mathbf{A}_{m\times n} = (a_{ij})$, the Frobenius norm are denoted by $\|\cdot\|$, and is defined as $||\mathbf{A}|| := \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}$.

We focus on the set of TDMs generated by the one-factor model with BB1 family linking copulas. Let $\mathcal{T}_d'$ denote the set of all d-dimensional TDMs that can be supported by such copulas. We say a TDM is feasible if it is in $\mathcal{T}_d'$; otherwise, this TDM is infeasible.

2. Tail dependence: one-factor copula model with BB1 family

In the following, we derive the lower tail dependence coefficients for the one-factor copula model with the BB1 family for the linking copula. To this end, we define the BB1 family of copulas indexed by two

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

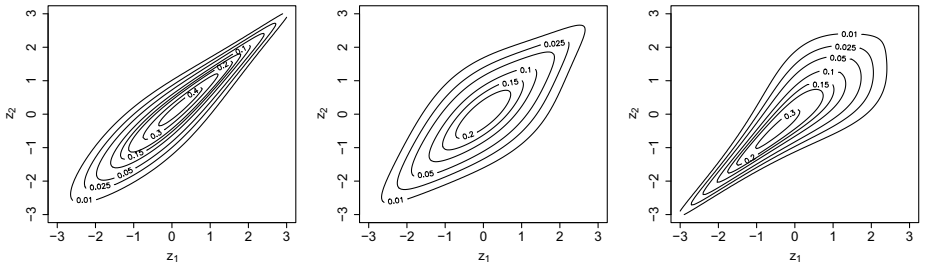


Figure 2. *Normalized contour plots for BB1 family with parameters* $(\theta, \delta) = (0.113, 3.80)$, $(0.585, 0.171)$, *and* $(2.63, 1.18)$. *The lower and upper tail dependence coefficients are* $(0.2, 0.8)$, $(0.5, 0.5)$, *and* $(0.8, 0.2)$, *respectively.*

parameters $\theta > 0$ and $\delta \geq 1$ as that given by

$$C(u, v; \theta, \delta) = \left\{1 + \left[\left(u^{-\theta} - 1\right)^{\delta} + \left(v^{-\theta} - 1\right)^{\delta}\right]^{\frac{1}{\delta}}\right\}^{-\frac{1}{\theta}}, \tag{2.1}$$

see Joe and Hu (1996) and section 5.2 of Joe (1997). This is the copula derived from the bivariate Weibull distribution of Lu and Bhattacharyya (1990). Also, the BB1 copula is referred to as the generalized Clayton copula, see page 259 of McNeil *et al*. (2015).

Remark 2.1. The limiting case of (2.1) as θ approaches zero is the Gumbel copula. It is noteworthy that as $\theta \to \infty$, the limiting copula is the comonotonicity copula $u \wedge v$, corresponding to $U = V$.

In Joe and Hu (1996), it is shown that for an Archimedean copula of the form

$$C(u, v) = \eta\left(\eta^{-1}(u) + \eta^{-1}(v)\right), \tag{2.2}$$

the lower tail dependence coefficient equals $2 \lim_{x\to\infty}\left[\eta'(2x)/\eta'(x)\right]$. Since the expression in (2.1) can be written in the form (2.2) with $\eta(x) = \left(1 + x^{1/\delta}\right)^{-1/\theta} \in (0, 1]$ and noting that $\eta^{-1}(x) = \left(x^{-\theta} - 1\right)^{\delta}$, the tail dependence coefficient for a BB1 copula equals

$$2 \lim_{x\to\infty} \frac{\eta'(2x)}{\eta'(x)} = 2 \lim_{x\to\infty} \left(\frac{1 + (2x)^{\frac{1}{\delta}}}{1 + x^{\frac{1}{\delta}}}\right)^{-\frac{1}{\theta}-1} 2^{\frac{1}{\delta}-1} = 2^{-\frac{1}{\theta\delta}}.$$

The normalized contour plots in Figure 2 display the ability of the BB1 family to model a variety of tail dependence.

Henceforth, we consider the case of a one-factor copula with the linking copulas belonging to the BB1 family with the parameters of the i-th linking copula denoted by $\theta_i (> 0)$, $\delta_i (\geq 1)$, for $i = 1, \ldots, d$. The contour plots in Figure 3 display the ability of the one-factor copula with BB1 family as the linking copula to model a variety of tail dependence. The tail dependence coefficients χ_{ij}, for $1 \leq i < j \leq d$, are given by

$$\chi_{ij} = \lim_{u\downarrow 0} \frac{1}{u} \int_0^1 C_{i|0}(u|v) C_{j|0}(u|v) dv = \lim_{u\downarrow 0} \int_0^{\frac{1}{u}} C_{i|0}(u|uw) C_{j|0}(u|uw) dw, \tag{2.3}$$

where

$$C_{i|0}(u|uw) = \left\{\left[\frac{u^{-\theta_i} - 1}{(uw)^{-\theta_i} - 1}\right]^{\delta_i} + 1\right\}^{\frac{1}{\delta_i}-1} \cdot \left\{(uw)^{\theta_i} + \left[\left[w^{\theta_i} - (uw)^{\theta_i}\right]^{\delta_i} + \left[1 - (uw)^{\theta_i}\right]^{\delta_i}\right]^{\frac{1}{\delta_i}}\right\}^{-\frac{1}{\theta_i}-1} \tag{2.4}$$

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

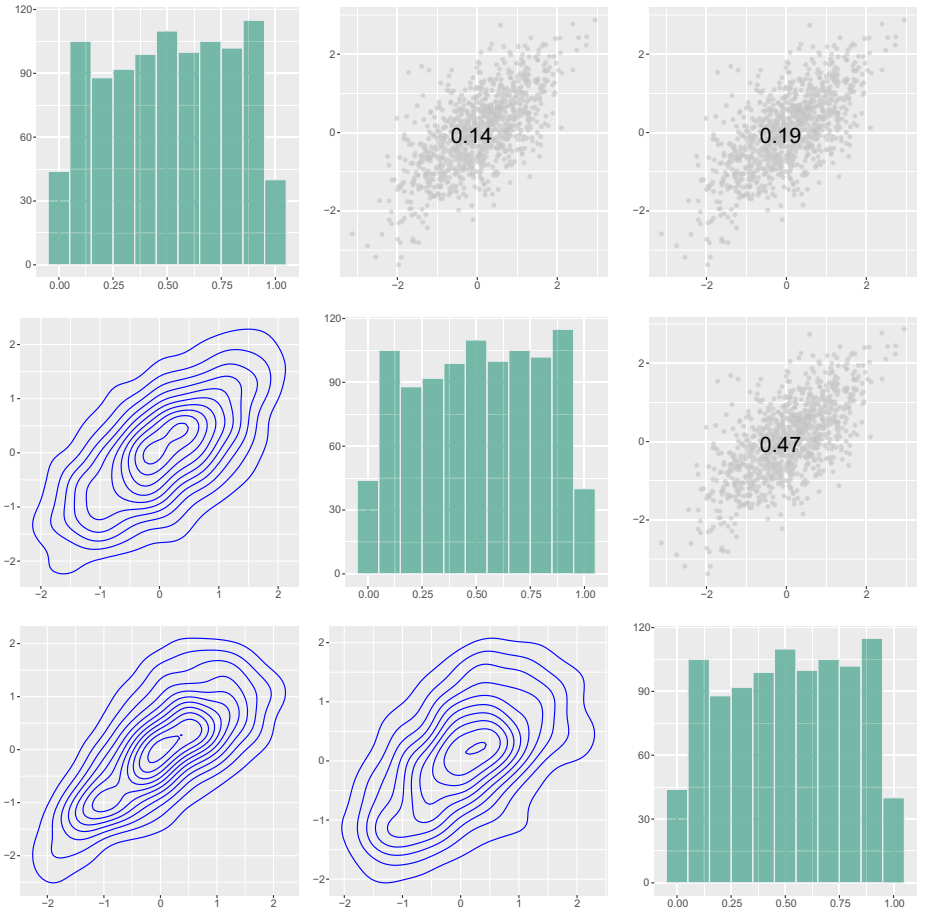


Figure 3. *Contour plot for one-factor copula with BB1 family as the linking copula with the same parameters as Figure 2.*

Theorem 2.1. *The tail dependence coefficients χ_{ij}, for $1 \le i < j \le d$, of the one-factor copula with the linking copula being the Family BB1 are given by*

$$\chi_{ij} = \int_0^\infty \left(1 + w^{\theta_i \delta_i}\right)^{-\frac{1}{\theta_i \delta_i} - 1} \left(1 + w^{\theta_j \delta_j}\right)^{-\frac{1}{\theta_j \delta_j} - 1} dw, \quad \theta_i, \theta_j \in (0, \infty),\ \delta_i, \delta_j \ge 1.$$

In the case that $\theta_i = \infty$, and $\theta_j \in (0, \infty)$, we have

$$\chi_{ij} = \int_0^1 (1 + w^{\theta_j \delta_j})^{-\frac{1}{\theta_j \delta_j} - 1} dw = 2^{-\frac{1}{\theta_j \delta_j}}.$$

Finally, if $\theta_i = \theta_j = \infty$, $\chi_{ij} = 1$.

Proof. The case when either or both θ's equal infinity, the result follows from the fact that the BB1 copula with $\theta = \infty$ is the comonotonic copula. Hence, in this case, the corresponding coordinate(s) equals the common factor, and χ_{ij} corresponds to that for the BB1 copula. For the case when $\theta_i, \theta_j \in (0, \infty)$, note that the limits of the conditional distributions $C_{i|0}(u|uw)$, for $i = 1, \ldots, d$, as u approaches 0, are given by

$$\lim_{u \downarrow 0} C_{i|0}(u|uw) = \left(1 + w^{\theta_i \delta_i}\right)^{\frac{1}{\delta_i} - 1 + \frac{1}{\delta_i}\left(-\frac{1}{\theta_i} - 1\right)} = \left(1 + w^{\theta_i \delta_i}\right)^{-\frac{1}{\theta_i \delta_i} - 1}, \quad i = 1, \ldots, d. \tag{2.5}$$

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

The result now follows by applying the dominated convergence theorem, as the integrand

$$w \mapsto C_{i|0}(u|uw)C_{j|0}(u|uw)I_{(0<uw<1)}$$

is bounded for all $u \in (0,1)$ by the integrable $g_i g_j$ of Lemma A.1 in the appendix. □

Note that the tail dependence coefficients χ_{ij}, can be expressed as $\chi_{ij} = \Psi(\theta_i\delta_i, \theta_j\delta_j)$, for $1 \le i < j \le d$, where for $0 < x, y < \infty$,

$$\begin{aligned}\Psi(x,y) &:= \int_0^\infty (1+w^x)^{-\frac{1}{x}-1}\,(1+w^y)^{-\frac{1}{y}-1}\,dw,\\ \Psi(x,\infty) &:= \Psi(\infty,x) := \Psi_1(x) := \int_0^1 (1+w^x)^{-\frac{1}{x}-1}dw = 2^{-\frac{1}{x}},\\ \Psi(\infty,\infty) &:= 1. \end{aligned} \tag{2.6}$$

Hence, the study of the geometric properties of the set of TDMs generated by the one-factor copula with the BB1 family for linking copulas, requires study of the symmetric function Ψ. Toward this, we denote the integrand above by ψ, that is

$$\psi(x,y,w) := (1+w^x)^{-\frac{1}{x}-1}\,(1+w^y)^{-\frac{1}{y}-1}, \qquad x,y>0,\ w \ge 0.$$

The proof of the following proposition is in the appendix.

Proposition 2.1. *Ψ is a symmetric strictly increasing function on $(0,\infty]^2$ with continuous partial derivatives.*

The following corollary conveniently lists some properties of Ψ that are needed for the results in the following sections, and the proof can be found in the appendix.

Corollary 2.1. *We have the following properties for Ψ:*

(i) The interval (0,1] is the the range of $\Psi(\cdot,\cdot)$.
(ii) If $\Psi(x,y) = \Psi(x',y') \in (0,1)$, then $sgn(x-x') + sgn(y-y') = 0$.

3. Representativeness of $\mathcal{T}_d'$ as a subset of $\mathcal{T}_d$

To study the geometric properties of $\mathcal{T}_d'$, we define the embedding function $\Psi_d : \mathbb{R}_+^d \to \mathcal{T}_d'$ by

$$(\Psi_d(\mathbf{x}))_{ij} := \begin{cases} \Psi(x_i,x_j), & 1 \le i \ne j \le d;\\ 1, & i=j, \end{cases}$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_d] \in \mathbb{R}_+^d$ is the vector of the parameters defining the tail dependence coefficients, that is $x_i := \theta_i\delta_i$, for $i = 1, \ldots, d$. Note that $\mathcal{T}_d'$ is the range of $\Psi_d(\mathbb{R}_+^d)$.

For $d = 2$, Corollary 2.1 readily shows that $\mathcal{T}_2' = (0,1] \subset \mathcal{T}_2 = [0,1]$. On the other hand, when $d \ge 3$, $\mathcal{T}_d'$ is a strict subset of $\mathcal{T}_d$. To see this, we consider MA(1) matrices with unit diagonal, that is, $\mathbf{T} = (t_{ij})$, where $t_{ij} = I_{i=j} + \alpha I_{|i-j|=1}$, for $1 \le i,j \le d$ and $\alpha \in (0, 1/2]$. Any such matrix is a member of $\mathcal{T}_d$, for $d \ge 2$ (see Section 4.1 of Embrechts *et al.*, 2016 or Proposition 6 of Shyamalkumar and Tao, 2020). However, its third-order leading principal sub-matrix of the form

$$\begin{bmatrix} 1 & \alpha & 0\\ \alpha & 1 & \alpha\\ 0 & \alpha & 1 \end{bmatrix}, \quad \alpha \in (0, 1/2].$$

is not feasible by Proposition 3.2 below. Therefore, $\mathcal{T}_d \setminus \mathcal{T}_d'$ in non-empty.

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

In the first subsection, we establish some topological and geometric characteristics of $\mathcal{T}_3'$. The second subsection introduces an imposed ordering and elaborates on geometric properties for $\mathcal{T}_d'$, $d \geq 4$.

3.1. Volume of $\mathcal{T}_d'$ relative to $\mathcal{T}_d$

As a starting point, we aim to determine the precise volume of the TDMs generated by the one-factor copula, using the Family BB1 as the linking copula in a three-dimensional setting. Our analysis begins with an exploration of a key property concerning the three-dimensional tail dependence coefficients, which is presented in the following proposition.

Proposition 3.1. *Let* $\chi_{12}, \chi_{13}, \chi_{23} \in (0, 1]$. *Then,*

$$\chi_{12} = \Psi(x, y), \quad \chi_{13} = \Psi(x, z), \quad \textit{and} \quad \chi_{23} = \Psi(y, z),$$

can have at most one solution. Furthermore, if $\{x, y, z\}$ *is the solution, then* $\chi_{12} \leq \chi_{13} \leq \chi_{23}$ *if an only if* $x \leq y \leq z$.

Proof. Suppose that $\{x, y, z\}$ and $\{x', y', z'\}$ are both solutions to the system of equations. Without loss of generality, we can assume that $x \geq x'$. By Corollary 2.1 (ii), we know that $y \leq y'$ and $z \leq z'$ from the first two equations. The third equation shows that $y = y'$ and $z = z'$. Hence, $x = x'$, and the two solutions are identical.

Now, let $\{x, y, z\}$ be the solution. The if part follows from monotonicity of Ψ from Proposition 2.1. For the only if part, let $\chi_{12} \leq \chi_{13} \leq \chi_{23}$. Again using monotonicity of Ψ, the first two equations along with $\chi_{12} \leq \chi_{13}$ indicate that $y \leq z$, and the last two equations similarly imply that $x \leq y$. Therefore, we have $x \leq y \leq z$. □

Before delving into properties of $\mathcal{T}_d'$ at higher dimensions, we further examine the representativeness of $\mathcal{T}_3'$. Proposition 3.1 establishes that $\mathcal{T}_3'$ displays invariance with respect to permutation of the coordinates. For expositional ease, we hence focus on the subset of the TDMs satisfying $\chi_{12} \leq \chi_{13} \leq \chi_{23}$. This subset, which we denote as $\mathcal{O}_3$, takes the form of a tetrahedron with vertices at (0,0,0), (0,0,1), (0, 1/2, 1/2), and (1,1,1). We further denote the collection of feasible points within $\mathcal{O}_3$ by $\mathcal{O}_3'$, that is $\mathcal{O}_3' := \mathcal{T}_3' \cap \mathcal{O}_3$. The volume of $\mathcal{O}_3$ (resp. $\mathcal{O}_3'$) is clearly one-sixth that of $\mathcal{T}_3$ (resp. $\mathcal{T}_3'$). Figure 4(a) displays the six cells of $\mathcal{T}_3$ belonging to its partition induced by the ordering of the coordinates. The following proposition gives a necessary and sufficient condition for membership in $\mathcal{T}_3'$, and using it derives its volume.

Proposition 3.2. *A matrix of the form*

$$\begin{bmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{bmatrix}, \quad 0 < a \leq b \leq c \leq 1,$$

is in $\mathcal{T}_3'$ *if and only if* $a \geq \Psi\left(\Psi_1^{-1}(b), \Psi_1^{-1}(c)\right)$. *Furthermore, the* $\mathcal{T}_3'$ *has a volume of about* 0.12882, *which is about 26% of the volume of* $\mathcal{T}_3$.

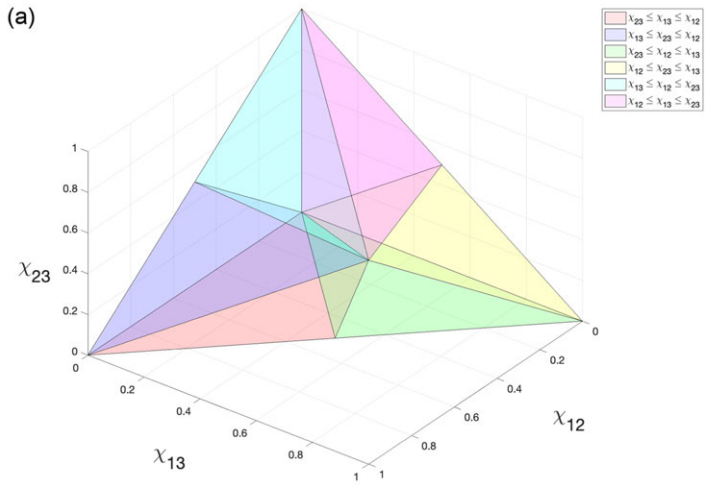
Proof. For the only if part, by Proposition 2.1, we establish that for any given $b = \Psi(x, z)$ and $c = \Psi(y, z)$, where $0 < b \leq c \leq 1$, the lower bounds for x and y are determined as $\Psi_1^{-1}(b)$ and $\Psi_1^{-1}(c)$, respectively. Consequently, the lower bound for $a = \Psi(x, y)$ is expressed as $\Psi\left(\Psi_1^{-1}(b), \Psi_1^{-1}(c)\right)$.

For the if part, let a, b, and c satisfy $\Psi(\Psi_1^{-1}(b), \Psi_1^{-1}(c)) \leq a \leq b \leq c \leq 1$. From Proposition 2.1, we have Ψ is a symmetric, strictly increasing, and continuous function on $(0, \infty]^2$. Since,

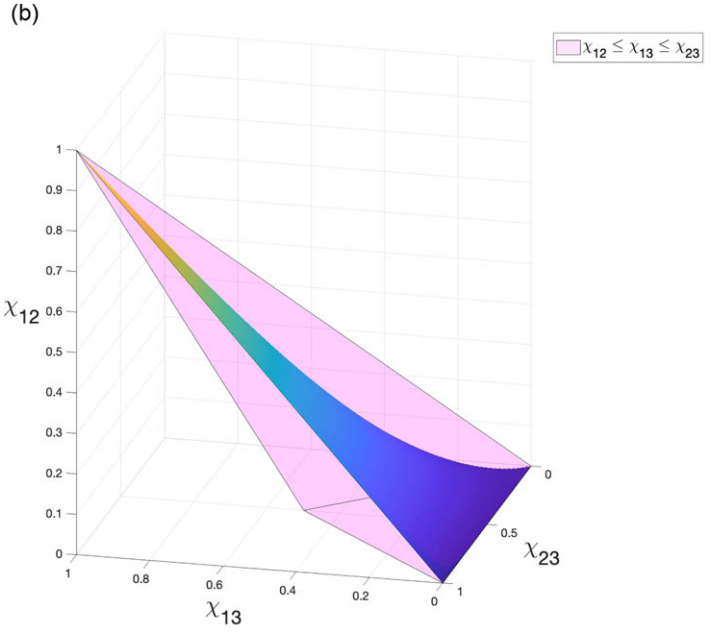
$$\Psi(\Psi_1^{-1}(b), 0) = 0, \text{ and } \Psi(\Psi_1^{-1}(b), \infty) = b,$$

using continuity of Ψ, we have the existence of a $z \geq \Psi_1^{-1}(c)$ satisfying $a = \Psi(\Psi_1^{-1}(b), z)$. This establishes that the parameter $(\Psi_1^{-1}(b), z, \infty)$ corresponds to the TDM with off-diagonal elements of a, b, and

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press



$\mathcal{O}_3$ (magenta) and the other cones of $\mathcal{T}_3$.



Lower bound on χ_{12} within $\mathcal{O}_3'$.

Figure 4. *$\mathcal{O}_3$ and the lower bound on χ_{12} within $\mathcal{O}_3'$.*

$\Psi(z, \infty)$. It is noteworthy that $\Psi(z, \infty) \geq \Psi(\Psi_1^{-1}(c), \infty) = c$. If $\Psi(z, \infty) = c$, our proof is complete; we now focus on the remaining case of $c < \Psi(z, \infty) \leq 1$. Consider functions $\eta_{a,b} : (0, \infty) \to (\Psi_1^{-1}(a), z)$ and $\eta_{b,b} : (0, \infty) \to (\Psi_1^{-1}(b), \infty)$, defined by

$$a = \Psi(\Psi_1^{-1}(b) + x, \eta_{a,b}(x)), \text{ and } b = \Psi(\Psi_1^{-1}(b) + x, \eta_{b,b}(x)),$$

respectively. Given that Ψ is strictly increasing and smooth, both $\eta_{a,b}$ and $\eta_{b,b}$ are continuous (see Lemma A.3 in the Appendix) and strictly decreasing. Defining $\phi(\cdot)$ by

$$\phi(x) := \Psi(\eta_{a,b}(x), \eta_{b,b}(x)), \quad x \in (0, \infty),$$

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

we note that

$$\lim_{x\downarrow 0}\phi(x)=\Psi(z,\infty)>c,\ \text{and}\ \lim_{x\uparrow\infty}\phi(x)=\Psi\left(\Psi_1^{-1}(a),\Psi_1^{-1}(b)\right)<a\le c.$$

The continuity of $\Psi(\cdot,\cdot)$, $\eta_{a,b}$ and $\eta_{b,b}$, imply the continuity of ϕ, and hence we have the existence of $z'>0$ satisfying $\phi(z')=\Psi(\eta_{a,b}(z'),\eta_{b,b}(z'))=c$. Consequently, the TDM with off-diagonal elements a, b, and c corresponds to the parameters $(\Psi_1^{-1}(b)+z',\eta_{a,b}(z'),\eta_{b,b}(z'))$.

Recalling that $\mathcal{O}_3'$ is given by

$$\{(\Psi(x,y),\Psi(x,z),\Psi(y,z)),\quad \text{for } 0<x\le y\le z\},$$

the curved surface of the lower bound on χ_{12} is plotted in Figure 4(b). The region above this curved surface represents $\mathcal{O}_3'$. Through integration across the curved surface and the $\chi_{13}-\chi_{23}$ plane, we can compute the volume below the surface as:

$$\int_0^1\int_0^c \Psi\left(\Psi_1^{-1}(b),\Psi_1^{-1}(c)\right)db\,dc\approx 0.145197.$$

As a result, the volume of $\mathcal{O}_3'$ is approximately 0.021470, and this result, when multiplied by six, yields the volume of $\mathcal{T}_3'$. □

Remark 3.1. We observe that $\mathcal{O}_3'$ is not a convex set. This is so as the midpoint of the line segment between the points (0,0,1) and (1/3, 1/2, 1/2) of $\mathcal{O}_3'$ (i.e. (1/6, 1/4, 3/4)) does not belong to $\mathcal{O}_3'$. This is verified by referring to Proposition 3.2 and noting that:

$$\Psi\left(\Psi_1^{-1}(3/4),\Psi_1^{-1}(1/4)\right)=0.23>1/6.$$

Furthermore, note that the non-convexity of $\mathcal{O}_3'$ implies the non-convexity of $\mathcal{T}_3'$.

Theorem 3.1. *$\mathcal{T}_d'$ has a non-zero volume in $\mathbb{R}_+^{\binom{d}{2}}$ if and only if $d\le 3$.*

Proof. Recall that the tail dependence coefficients of a d-dimensional one-factor copula with the linking copulas from a BB1 family are a function of the d-dimensional vector $(\theta_1\delta_1,\ldots,\theta_d\delta_d)\in(0,\infty)^d$. Defining,

$$\phi:(0,\infty)^d\longrightarrow[0,1]^{\binom{d}{2}}$$

$$(x_1,\ldots,x_d)\longmapsto(\Psi(x_1,x_2),\Psi(x_1,x_3),\ldots,\Psi(x_{d-1},x_d)),$$

we see that the interior of $\mathcal{T}_d'$ is contained in the range of ϕ. Note that $d<d(d-1)/2$, for $d\ge 4$, hence the domain of ϕ is in a lower dimensional Euclidean space than its range. Moreover, Proposition 2.1 implies that ϕ is in C^1. Hence, by using Sard's theorem (see Theorem A.1 of the Appendix), we see that the range of ϕ has zero Lebesgue measure in $[0,1]^{\binom{d}{2}}$. □

Remark 3.2. It is important to note that choosing a different linking copula from the BB families listed in Joe (1997) would lead to a different set of TDMs generated by the one-factor copula for $d=3$. In Table A1 in the Appendix, we provide additional information on related functions and the volume of the one-factor copula using two other BB families (BB4 and BB7) suggested in Section 8.6 of Joe (2011). To calculate the volume for the BB4 family, we employed a hit-and-run sampler, a method detailed in Kroese *et al*. (2013) and Shyamalkumar and Tao (2022). This approach allowed us to uniformly generate a sample of 100,000 points from $\mathcal{T}_3$, which we then assessed for their feasibility within the BB4 family. Importantly, it is worth noting that the tail dependence coefficient of the BB7 family adheres to the same mathematical form as that of the BB1 family. As a result, both families share identical volumes.

3.2. An imposed ordering

In the following, we show that the one-factor structure imposes a rather restrictive partial ordering on the tail dependence coefficients. While the discussion below is in the case of a BB1 linking copula,

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

the development suggests that the phenomenon is more general. We begin by showing that, in general, on the contrary, all possible orderings of the tail dependence coefficients are possible. In the following $\mathcal{I}_d := \{(1,2),(1,3),\ldots,(d-1,d)\}$, the set of all ordered pairs of positive integers less than or equal to d.

Theorem 3.2. *Given a linear ordering $\prec$ on $\mathcal{I}_d$, there exists a TDM, $\mathbf{T}_d = (t_{ij})$, such that the mapping $(i,j) \in \mathcal{I}_d \mapsto t_{ij}$ is strictly increasing in $\prec$.*

Proof. For $(i,j) \in \mathcal{I}_d$, let $0 \le \varepsilon_{ij} \le 1/d$ be such that $(i,j) \mapsto \varepsilon_{ij}$ is strictly decreasing in $\prec$, and define $\varepsilon_{ji} := \varepsilon_{ij}$. Defining $t_{ij} := 1/d - \varepsilon_{ij}$, for $1 \le i < j \le d$, we see that the mapping $(i,j) \mapsto t_{ij}$ is strictly increasing in $\prec$. In the following, we will show that $\mathbf{T}_d = (t_{ij})$ is a TDM. By Theorem 1 of Shyamalkumar and Tao (2020), it is equivalent to show that $\mathbf{T}_d/d$ is a Bernoulli Compatible Matrix (BCM). Also, by Corollary 3 therein, this is equivalent to showing the existence of events A_i, $i = 1,\ldots d$, on a probability space with identical probabilities of $1/d$ such that $t_{ij} = \Pr\left(A_i|A_j\right)$. Toward this, we construct disjoint sets A_{ij}, for $1 \le i < j \le d$, and B_i, $i = 1,\ldots d$, on a non-atomic probability space with probabilities

$$\Pr\left(A_{ij}\right) = \frac{1}{d}\left(\frac{1}{d} - \varepsilon_{ij}\right), \quad \text{and} \quad \Pr\left(B_i\right) = \frac{1}{d} - \sum_{j \ne i}\left(\frac{1}{d^2} - \frac{\varepsilon_{ij}}{d}\right).$$

This can be done as the sum of the probabilities of the above-mentioned sets are

$$\sum_{1 \le i < j \le d} \Pr\left(A_{ij}\right) + \sum_{i=1}^{d} \Pr\left(B_i\right) = \sum_{1 \le i < j \le d} \frac{1}{d}\left(\frac{1}{d} - \varepsilon_{ij}\right) + \sum_{i=1}^{d}\left[\frac{1}{d} - \sum_{j \ne i}\left(\frac{1}{d^2} - \frac{\varepsilon_{ij}}{d}\right)\right]$$

$$= 1 - \sum_{1 \le i < j \le d} \frac{1}{d}\left(\frac{1}{d} - \varepsilon_{ij}\right) \le 1.$$

With $A_{ji} := A_{ij}$, for $1 \le i < j \le d$, and

$$A_i := B_i \cup \left(\bigcup_{j \ne i} A_{ij}\right), \quad \text{for } i = 1,\ldots,d,$$

we see that $\Pr\left(A_i\right) = 1/d$, $A_i \cap A_j = A_{ij}$, and hence $t_{ij} = \Pr\left(A_i|A_j\right)$ satisfies the sought ordering. □

Remark 3.3. From the above, it is easily checked that for $\|\varepsilon\|_2 \le \sqrt{2}/d^2$, by Cauchy-Schwartz inequality we have $\|\varepsilon\|_1 \le 1/d$. The latter suffices for the validity of the above construction, and hence the above proof implies further that the TDM with off-diagonal elements equal to $1/d$ is in the interior of $\mathcal{T}_d$, and moreover using the fact that $\mathcal{T}_d$ is a convex polytope we have that the line segment between the extreme points origin and $\mathbf{1}$ of $\mathcal{T}_d$ lies in its interior.

In the following, we assume that the linking copula is the BB1 copula and define $\eta_i := \theta_i\delta_i$, $i = 1,\ldots,d$. In the case $d = 3$, we have the ordering of η_i's specifying the ordering of the tail dependence coefficients by the monotonicity of $\Psi(\cdot,\cdot)$. In particular, if $\eta_1 \le \eta_2 \le \eta_3$, then we have $\chi_{12} \le \chi_{13} \le \chi_{23}$. For $d \ge 3$, the lemma below shows that the ordering of η_is specifies a partial ordering on ordered pairs (i,j), for $1 \le i < j \le d$, denoted by $\prec_\eta$, and that the mapping $(i,j) \mapsto \chi_{ij}$ is non-decreasing with respect to this partial order. Figure 5(a) contains the Hasse diagram for the partial ordering with non-decreasing η_is for $d = 5$ and specifies the partial ordering for $d \le 5$.

For the following development, we require an arbitrarily chosen strict total order on $\mathcal{I}_d$. For simplicity, we use the reverse lexicographic ordering and denote it by $\prec_{rl}$. The notation used below does suppress the dependence on $\prec_{rl}$ as we note below that it has little impact on the main conclusion.

Proposition 3.3. *Given η_is, we define $(i,j) \prec_\eta (k,l)$ for $(i,j),(k,l) \in \mathcal{I}_d$ if $\min(\eta_i,\eta_j) \le \min(\eta_k,\eta_l)$ and $\max(\eta_i,\eta_j) \le \max(\eta_k,\eta_l)$, with ties broken using $\prec_{rl}$. The binary relation $\prec_\eta$ is a partial order, and moreover, the mapping $(i,j) \mapsto \chi_{ij}$ is non-decreasing with respect to $\prec_\eta$.*

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

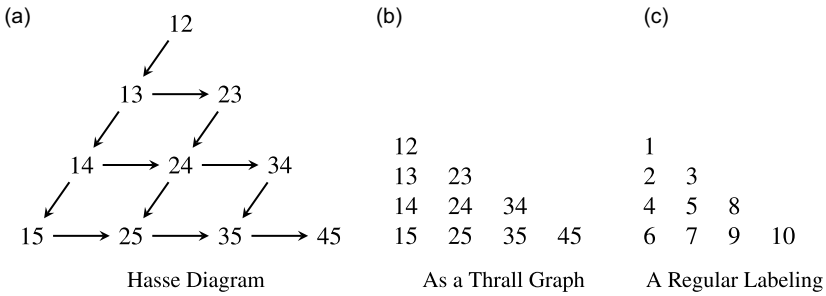


Figure 5. *Linear ordering of the poset: $d = 5$; increasing η_is.*

Proof. To show that $\prec_\eta$ is a partial order, we need to demonstrate the following properties for any $(i,j), (k,l), (u,v) \in \mathcal{I}_d$:

i. Reflexivity: $(i,j) \prec_\eta (i,j)$;
ii. Antisymmetry: If $(i,j) \prec_\eta (k,l)$ and $(k,l) \prec_\eta (i,j)$, then $(i,j) = (k,l)$;
iii. Transitivity: If $(i,j) \prec_\eta (k,l)$ and $(k,l) \prec_\eta (u,v)$, then $(i,j) \prec_\eta (u,v)$.

Part i. follows from the definition of $\prec_\eta$ and that $\prec_{rl}$ is a partial order. For part ii., we have that

$$\min(\eta_i, \eta_j) \le \min(\eta_k, \eta_l) \le \min(\eta_i, \eta_j) \quad \text{and} \quad \max(\eta_i, \eta_j) \le \max(\eta_k, \eta_l) \le \max(\eta_i, \eta_j).$$

Hence, $\min(\eta_i, \eta_j) = \min(\eta_k, \eta_l)$ and $\max(\eta_i, \eta_j) = \max(\eta_k, \eta_l)$. Since ties are broken (using $\prec_{rl}$), we have $(i,j) = (k,l)$. For part iii., the ordering implies that

$$\min(\eta_i, \eta_j) \le \min(\eta_k, \eta_l) \le \min(\eta_u, \eta_v) \quad \text{and} \quad \max(\eta_i, \eta_j) \le \max(\eta_k, \eta_l) \le \max(\eta_u, \eta_v).$$

Hence, since ties are broken using $\prec_{rl}$, we have $(i,j) \prec_\eta (u,v)$.

The non-decreasing nature of the mapping $(i,j) \mapsto \chi_{ij}$ with respect to $\prec_\eta$ follows from Proposition 2.1 as for $(i,j) \prec_\eta (k,l)$ we have that

$$\begin{aligned}\chi_{ij} = \Psi(\eta_i, \eta_j) &= \Psi(\min(\eta_i, \eta_j), \max(\eta_i, \eta_j)) \\ &\le \Psi(\min(\eta_k, \eta_l), \max(\eta_k, \eta_l)) = \Psi(\eta_k, \eta_l) = \chi_{kl}.\end{aligned}$$

□

On the other hand, given a set of tail dependent coefficients corresponding to a TDM in $\mathcal{T}_d'$, we can infer the ordering of the underlying η_is as demonstrated by the below proposition.

Proposition 3.4. *Let $\bar{\chi} := (\chi_{1,2}, \ldots, \chi_{d-1,d})$ be tail dependence coefficients corresponding to a member of $\mathcal{T}_d'$, with $d \ge 3$. Also, let $n_{\bar{\chi}}$ be the function defined on $\{1, \ldots, d\}$ by*

$$n_{\bar{\chi}}(i) = \sum_{m=1}^{d(d-1)/2} 2^{m-1} I_{(i_m = i \text{ or } j_m = i)},$$

where $(i_m, j_m) \in \mathcal{I}_d$, for $m = 1, \ldots, d(d-1)/2$, is defined by $\chi_{i_1,j_1} \le \cdots \le \chi_{i_{d(d-1)/2},j_{d(d-1)/2}}$, with the ties broken using $\prec_{rl}$. Then $n_{\bar{\chi}}$ is injective, and η_is underlying $\bar{\chi}$ are non-decreasing with respect to $<_\chi$, the strict total order on $\{1, \ldots, d\}$ under which $n_{\bar{\chi}}$ increasing.

Proof. To show that $n_{\bar{\chi}}$ is injective, assume on the contrary that $n_{\bar{\chi}}(i) = n_{\bar{\chi}}(j)$, for some $i < j$. Then,

$$\sum_{m=1}^{d(d-1)/2} 2^{m-1} I_{(i_m = i \text{ or } j_m = i)} = \sum_{m=1}^{d(d-1)/2} 2^{m-1} I_{(i_m = j \text{ or } j_m = j)},$$

implying that $I_{(i_m = i \text{ or } j_m = i)} = I_{(i_m = j \text{ or } j_m = j)}$, for $m = 1, \ldots, d(d-1)/2$. Since the latter implies that $i = j$, we have a contradiction. Thus, $n_{\bar{\chi}}$ is injective.

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

Next, we prove that the η_is underlying $\bar{\chi}$ are non-decreasing with respect to the total order $\prec_\chi$. For $i < j$ in the total order $\prec_\chi$, we have $n_{\bar{\chi}}(i) < n_{\bar{\chi}}(j)$, that is,

$$\sum_{m=1}^{d(d-1)/2} 2^{m-1} I_{(i_m=i \text{ or } j_m=i)} < \sum_{m=1}^{d(d-1)/2} 2^{m-1} I_{(i_m=j \text{ or } j_m=j)}.$$

Hence, $\arg\max_m (i_m = i \text{ or } j_m = i) \leq \arg\max_m (i_m = j \text{ or } j_m = j)$. Now, consider two cases. First, if $\arg\max_m (i_m = i \text{ or } j_m = i) < \arg\max_m (i_m = j \text{ or } j_m = j)$, then $\chi_{ki} \leq \chi_{k'i} \leq \chi_{kj}$, where χ_{kj} is the largest value in $\bar{\chi}$ that has j as an index, and $\chi_{k'i}$ is the largest value in $\bar{\chi}$ that has i as an index. By Proposition 2.1, this implies $\eta_i \leq \eta_j$. Second, if $\arg\max_m (i_m = i \text{ or } j_m = i) = \arg\max_m (i_m = j \text{ or } j_m = j) := l$, then $\arg\max_{m \neq l} (i_m = i \text{ or } j_m = i) < \arg\max_{m \neq l} (i_m = j \text{ or } j_m = j)$. This means that $\chi_{ki} \leq \chi_{k'i} \leq \chi_{kj}$, where χ_{kj} is the second largest value in $\bar{\chi}$ that has j as an index, and $\chi_{k'i}$ is the second largest value in $\bar{\chi}$ that has i as an index. Hence, by Proposition 2.1, we again have $\eta_i \leq \eta_j$. In both cases, we have shown that for $i < j$ in the total order $\prec_\chi$, $\eta_i \leq \eta_j$. Therefore, the η_is underlying $\bar{\chi}$ are non-decreasing with respect to the total order $\prec_\chi$. □

Proposition 3.4 implies that $\mathcal{T}_d'$ can be partitioned into $d!$ subsets with each corresponding to a permutation of $\{1, \ldots, d\}$, consistent with the ordering of the underlying η_is. Note that the use of $\prec_{rl}$ only is relevant for $\bar{\chi}$ with some tail dependence coefficients assuming the same value. By Proposition 3.3 we see that each of these subsets can be further partitioned corresponding to a permutation of $\mathcal{I}_d$, consistent with the ordering of the tail dependence coefficients, and by symmetry we see that there are $(d(d-1)/2)!/d!$ such subsets. In the case $d = 4$, we have 30 such cells of $\mathcal{T}_4$ corresponding to the identity permutation on $\{1, \ldots, d\}$, denoted by $\mathcal{O}_4$, corresponding to an increasing η_is. Interestingly, only two of them are nonempty if we ignore $\bar{\chi}$ with some tail dependence coefficients assuming the same value. These correspond to the permutations,

$$((1,2),(1,3),(2,3),(1,4),(2,4),(3,4)) \text{ and } ((1,2),(1,3),(1,4),(2,3),(2,4),(3,4)),$$

consistent, respectively, with

$$\chi_{12} < \chi_{13} < \chi_{23} < \chi_{14} < \chi_{24} < \chi_{34}, \text{ and } \chi_{12} < \chi_{13} < \chi_{14} < \chi_{23} < \chi_{24} < \chi_{34}.$$

This is so as these are the only permutations corresponding to linear orderings consistent with the partial order of Figure 5(a). In other words, of the 720 (non-empty interiors by Theorem 3.2 and Remark 3.3) cells of $\mathcal{T}_4$ determined by the ordering $\prec_\chi$, only a mere 48 ($2 \times 4!$) have a non-empty intersection of their interiors with $\mathcal{T}_4'$. The below theorem generalizes this observation for all $d \geq 3$. In Table A2 of the Appendix, the volumes of these cells of $\mathcal{O}_4$ calculated using polymake (see Gawrilow and Joswig, 2000) are given. It is worth noting that the nonempty cells 1 and 7 – polytopes – share a facet of dimension 5. The adjacency structure for all 30 polytopes of $\mathcal{O}_4$ is depicted in Figure A1 in the Appendix, where polytopes connected with lines indicate that they share a common facet of dimension 5.

Remark 3.4. For any point in $\mathcal{O}_4$, we can find the lower bound of its Euclidean distance to $\mathcal{O}_4'$, the feasible area under the strict total order $\prec_\chi$, by projecting it onto cells 1 or 7 of the partition $\mathcal{O}_4$. For example, the lower bound from the vertex $(1,0,0,0,0,1) \in \mathcal{O}_4$ to cells 1 and 7 is $\sqrt{0.8}$. The lower bound to the feasible area can be considered as a lower bound to the closest feasible matrix, which we will discuss in the next section.

For the proof of the theorem, we need the concept of a Young tableaux. A Young tableaux, see Section 5.1.4 of Knuth (1998), is an arrangement of $n_1 \geq n_2 \geq \ldots \geq n_{m-1} \geq n_m > 0$ distinct integers in an array of left-justified rows with the i-th row containing n_i elements, $i = 1, \ldots m$, such that the elements in each row are in an increasing order (left to right) and the elements in each column are increasing top to bottom. Below, find it convenient to refer to the underlying table structure as a Young graph, refer to the assignment of numbers to the graph as the labeling, and a labeling satisfying the above monotonicity property as a regular labeling. In Table 1a we depict a Young tableaux corresponding to the (4,2,1) partition of 7. In Table 1b, we depict a certain version of a shifted Young tableaux in which the i-th row

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

Table 1. *Young and Thrall tableaux*

(a) Young Tableaux				(b) Shifted Young Tableaux				(c) Thrall Tableaux		
1	2	3	5	1	2	3	5	1		
4	6				4	6		2	4	
7						7		3	6	7
								5		

Table 2. *Not all Young tableaux correspond to a Thrall Tableaux.*

(a) Young Tableaux				(b) Shifted Young Tableaux				(c) Thrall Tableaux		
1	3	4	5	1	3	4	5	1		
2	6				**2**	6		3	**2**	
7						7		4	6	7
								5		

begins at the i-th column, and in Table 1c, we depict the transpose of the shifted Young tableaux which we will refer to as a Thrall tableaux.

While there is an obvious equivalence between a regular labelings of a Thrall graph and its corresponding shifted Young graph, and such labelings yield a regular labeling of the corresponding Young graph, Table 2 shows that not all regular labelings of a Young graph yield a regular labeling of its corresponding Thrall graph. The following result from Thrall (1952) gives the number of regular labelings of a Thrall graph.

Theorem 3.3 (Thrall 1952). *The number of regular labelings of a Thrall graph associated with* $\tilde{n} = (n_1, \ldots, n_m)$*, a partition of* n*, with label set* $\{1, \ldots, n\}$ *is given by*

$$\frac{n!\Delta(\tilde{n})}{n_1!n_2!\cdots n_m!\nabla(n_1, \ldots, n_m)}, \tag{3.1}$$

where $n = n_1 + \cdots + n_m$*,*

$$\Delta(\tilde{n}) := \prod_{1\le i<j\le m} (n_i - n_j), \quad \textit{and} \quad \nabla(\tilde{n}) := \prod_{1\le i<j\le m} (n_i + n_j).$$

The number of Young tableaux of shape $(n_1, \ldots, n_m)$ is given by, see (34) on page 58 of of Knuth (1998),

$$\frac{n!\Delta(n_1 + m - 1, n_2 + m - 2, \ldots, n_m)}{(n_1 + m - 1)!(n_2 + m - 2)!\cdots n_m!}. \tag{3.2}$$

Interestingly, the formula in (3.2) was first derived from a group theoretic perspective in Frobenius (1900), and later a combinatorial proof was given in MacMahon (1909).

Theorem 3.4. *For* $d \ge 3$*, of the* $\binom{d}{2}!$ *subsets of* $\mathcal{T}_d$ *determined by* $<_\chi$*, the number with nonempty intersection of their interiors with* $\mathcal{T}_d'$ *is bounded above by*

$$N(d) := d!\binom{d}{2}!\frac{\prod_{i=1}^{d-2} i!}{\prod_{i=1}^{d-1}(2i-1)!}. \tag{3.3}$$

In particular, the corresponding proportion decays super-exponentially to zero as

$$\frac{d\prod_{i=1}^{d-1} i!}{\prod_{i=1}^{d-1}(2i-1)!} \le e^{-\Omega(d^2)}. \tag{3.4}$$

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

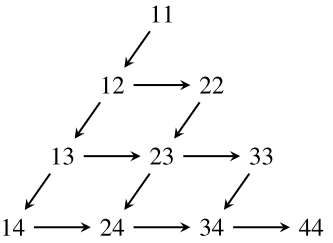


Figure 6. *Hasse Diagram – $n = 4$.*

Proof. Let us consider, without loss of generality, the interior of the subset of $\mathcal{T}_d$, which corresponds via $<_\chi$ to the identity permutation on $\{1, \ldots, d\}$. By Proposition 3.4, points in this interior belonging to $\mathcal{T}_d'$ have a strictly increasing underlying η_is. Hence, by Propositions 3.3 and 3.1, we have that any such point is strictly increasing in a strict linear order of $\mathcal{I}_d$ that is consistent with the partial order implied by a strictly increasing underlying η_is. This partial order for $d = 5$ is depicted in Figure 5(a). So the number of distinct strict total orders corresponding to such points is bounded above by strict linear orders on $\mathcal{I}_d$ consistent with the partial order implied by a strictly increasing underlying η_is. So to prove (3.3), by symmetry with respect to the permutation implied by $<_\chi$, it suffices to show that the latter equals $N(d)/d!$.

Considering the case $d = 5$, we see that the number of strict total orders satisfying the partial order of Figure 5(a) equals the number of regular labelings with $\{1, \ldots, 10\}$ of a Thrall graph such as that shown in Table 2(c). In general, it should be now clear that we seek the number of regular labelings of a Thrall graph associated with $\tilde{n}_d = (d-1, d-2, \ldots, 1)$, a partition of $n_d = d(d-1)/2$, with label set $\{1, \ldots, n_d\}$. The proof of (3.3) is completed by an application of Theorem 3.3, and by observing that

$$\Delta(\tilde{n}_d) = \prod_{i=1}^{d-2} i!, \quad \text{and} \quad \nabla(\tilde{n}_d) = \prod_{i=1}^{d-1} \frac{(2i-1)!}{i!}.$$

The super-exponential decay conveyed by (3.4) follows from Lemma A.5 in the Appendix. □

While Theorem 3.4 provides an upper bound, it is unlikely that it is exact. Consider a similar problem of the attainable orderings of $\{a_i a_j\}_{1 \le i \le j \le n}$ for n positive reals $a_1 > \ldots > a_n > 0$. For the case $n = 4$, the corresponding poset is denoted by the following Hasse diagram of Figure 6 (which is isomorphic to that in Figure 5(a)). But as discussed in Johnston (2014), the number of attainable linear orderings is 10 and not 12 (the total number of linear orderings consistent with the partial order of Figure 6). In this problem, while the *diagonal* is involved, the problem is seemingly simpler to analyze than the one we consider as instead of our $\Psi(\cdot, \cdot)$ it involves the simpler binary operator, the product. In the specific case of $d = 5$, all of the 12 possible linear orderings are attainable.

In Fiebig *et al*. (2017), they investigate the set of all $\{0, 1\}$ valued vertices of $\mathcal{T}_d$ as $\mathcal{T}_d$ being a convex set contains the convex hull of such vertices. In fact, Proposition 21 therein shows by such an analysis that the clique partition polytope is contained in $\mathcal{T}_d$. While the non-convexity of $\mathcal{T}'_d$ voids such a rationale, it is nevertheless interesting as a simple measure of the coverage of $\mathcal{T}_d$ by $\mathcal{T}_d'$. For this to be meaningful, for this discussion, we extend the class of BB1 copulas to include the lower tail-independent Gumbel copula to accommodate lower tail independence (i.e., tail dependence coefficient of zero). We note that if one of the linking copulae is Gumbel, then the corresponding coordinate is lower tail independent from the rest of the coordinates. This means that the set of all feasible $\{0, 1\}$ valued TDMs corresponds to a subset of coordinates being perfectly tail dependent with each other (tail dependence coefficient of 1; comonotonicity copula; $\theta = \infty$) and the rest of the coordinates being tail independent with the other $d - 1$ coordinates. In particular, the set of feasible $\{0, 1\}$ valued TDMs contains $2^d - d$ members out

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

Table 3. *The number of* $\{0, 1\}$*-valued and feasible vertices of* $\mathcal{T}_d$*, for* $3 \le d \le 6$*: although the number of* $\{0, 1\}$*-valued vertices in* $\mathcal{T}_d$ *(equal to the Bell number) increases super-exponentially with the dimension d, the number of feasible vertices approximately doubles with each increment in dimension.*

	d			
	3	4	5	6
# of vertices	5	15	214	28,895
# of $\{0, 1\}$-valued vertices	5	15	52	203
# of feasible vertices	5	12	27	58
Proportion of feasible vertices	100%	80%	13%	0.2%

of the dth Bell number of many such vertices of $\mathcal{T}_d$. Using the lower bound for the dth Bell number provided by Berend and Tassa (2010), we conclude that that the proportion of $\{0, 1\}$-valued vertices that are feasible is bounded above by

$$\left(\frac{2e \log d}{d}\right)^d \to 0, \quad d \to \infty$$

Table 3 presents the exact results for $d \le 6$.

4. Modeling with one-factor copula

From Theorem 3.1, we know that $\mathcal{T}_d'$ has zero volume for $d \ge 4$, and moreover, that feasible TDMs are confined to a relatively few sub-regions of $\mathcal{T}_d$. This, in particular, implies that the one-factor copula model's ability to handle arbitrary tail dependence is quite limited in higher dimensions, similar to the findings in the study by Shyamalkumar and Tao (2022) on t-copula. When an actuarial modeler uses the one-factor copula and relies on actuarial opinion to obtain a target TDM or estimates the TDM using data, there is a potential issue when the copula is unable to accommodate the target or estimated TDM. This is because expert opinion and estimates are inherently imprecise, and the modeler may need to carefully search for a closely approximating TDM within $\mathcal{T}_d'$. In this section, we present an approach for finding such a TDM.

Given an infeasible d-dimensional target TDM $\mathbf{T}_g = (t_{ij})$, we aim to find the closest feasible TDM in Frobenius norm. To achieve this, we formulate an optimization problem as the following:

$$\begin{aligned} &\text{minimize } \mathbf{h}_d(\mathbf{x}) := \tfrac{1}{2}||\Psi_d(\mathbf{x}) - \mathbf{T}_g||^2 = \sum_{1 \le i < j \le d} [\Psi(x_i, x_j) - t_{ij}]^2; \\ &\text{subject to } \mathbf{x} = [x_1, x_2, \ldots, x_d] \in \mathbb{R}_+^d. \end{aligned} \tag{4.1}$$

Here, $\Psi_d(\mathbf{x})$ is the TDM generated by the one-factor copula model with parameters $\mathbf{x}$, and $\Psi(x_i, x_j)$ denotes the (i,j)-th entry of $\Psi_d(\mathbf{x})$. The objective function $\mathbf{h}_d(\mathbf{x})$ measures the distance between the target TDM $\mathbf{T}_g$ and the TDM generated by the one-factor copula model with parameters $\mathbf{x}$. The optimization problem is subject to the non-negative constraints on $\mathbf{x}$ to ensure the feasibility of the resulting TDM.

Clearly, the definition of Ψ as a definite integral without a closed-form value rules out a closed-form solution to the above optimization problem. Hence, we must rely on numerical optimization methods to solve the optimization problem (4.1). The fmincon function of MATLAB (see MathWorks, 2021) finds local solutions to nonlinear non-convex problems using an interior-point algorithm. It is worth noting

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

that by providing the gradient,

$$\nabla \mathbf{h}_d(\mathbf{x}) = \left(\frac{\partial \mathbf{h}_d(\mathbf{x})}{\partial x_1}, \ldots, \frac{\partial \mathbf{h}_d(\mathbf{x})}{\partial x_d} \right),$$

where

$$\frac{\partial \mathbf{h}_d(\mathbf{x})}{\partial x_i} = \sum_{j \neq i} 2[\Psi(x_i, x_j) - t_{i,j}] \frac{\partial \Psi(x_i, x_j)}{\partial x_i},$$

we can significantly decrease the computational time. However, the built-in trust-region algorithm of the fmincon function performs worse; thus, we do not report its results.

To test the algorithm, we do not have a standard TDM test library available for $d \geq 7$. Therefore, we utilize two parametric families of TDMs presented in Shyamalkumar and Tao (2020). Firstly, we consider the two-sector matrices:

$$\begin{bmatrix} 1 & \alpha & \cdots & \alpha & \gamma & \gamma & \cdots & \gamma \\ \alpha & 1 & \cdots & \alpha & \gamma & \gamma & \cdots & \gamma \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha & \alpha & \cdots & 1 & \gamma & \gamma & \cdots & \gamma \\ \gamma & \gamma & \cdots & \gamma & 1 & \beta & \cdots & \beta \\ \gamma & \gamma & \cdots & \gamma & \beta & 1 & \cdots & \beta \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma & \gamma & \cdots & \gamma & \beta & \beta & \cdots & 1 \end{bmatrix} = \begin{bmatrix} (1-\alpha)\mathbf{I}_{d_1} + \alpha \mathbf{J}_{d_1} & \gamma \mathbf{J}_{d_1 \times d_2} \\ \gamma \mathbf{J}_{d_2 \times d_1} & (1-\beta)\mathbf{I}_{d_2} + \beta \mathbf{J}_{d_2} \end{bmatrix}. \tag{4.2}$$

where $(\alpha, \beta) \in [0, 1]^2$ and $\gamma \in [0, \gamma_u(\alpha, \beta)]$. The value of $\gamma_u(\alpha, \beta)$ can be computed using a linear programming formulation, as detailed in Shyamalkumar and Tao (2020). In this two-section case, it is natural to model it using a two-factor model. However, in our analysis, we utilize a one-factor copula to model this situation. Secondly, in order to test the algorithm, we consider the two-dependent matrices of the form:

$$\begin{bmatrix} 1 & \alpha & \beta & 0 & \cdots & 0 \\ \alpha & 1 & \alpha & \ddots & \ddots & \vdots \\ \beta & \alpha & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \alpha & \beta \\ \vdots & \ddots & \ddots & \alpha & 1 & \alpha \\ 0 & \cdots & 0 & \beta & \alpha & 1 \end{bmatrix}, \tag{4.3}$$

where the matrix of the form (4.3) is in $\mathcal{T}_d$ if and only if α and β satisfy the following constraints:

$$\begin{cases} \alpha, \beta \geq 0; \\ \alpha + 4\beta \leq 2; \\ 2\alpha - \beta \leq 1. \end{cases} \tag{4.4}$$

It is important to note that all of the matrices of the form (4.3) are not in $\mathcal{T}_d'$, except when $\alpha = \beta = 0$, by Lemma A.6 of the Appendix.

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

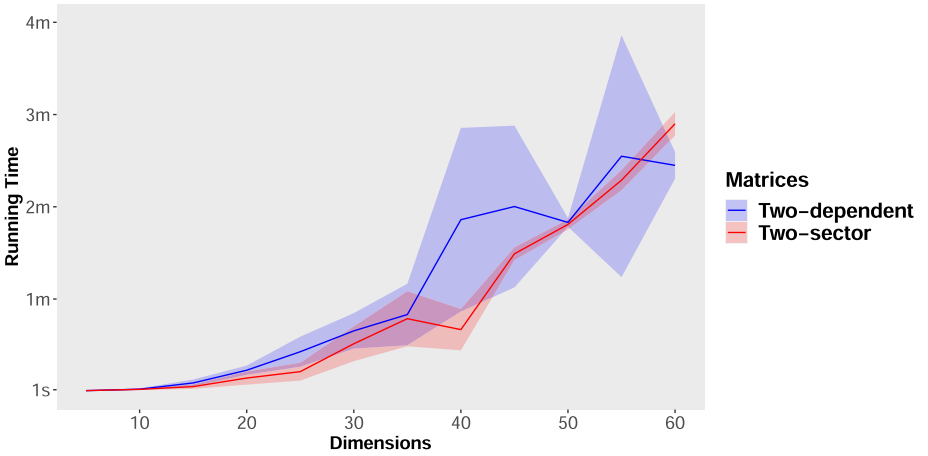


Figure 7. *Testing the performance of the interior-point algorithm: test instances were generated randomly by selecting 25 $d \times d$ matrices from two-sector matrices of the form (4.2) and two-dependent matrices of the form (4.3) for $d \in \{5, 10, \ldots, 60\}$.*

All the computations in this section were performed on a computer equipped with a single Intel Quad Core i7 2.7GHz processor and 16GB of RAM. To investigate the computational performance of the fmincon function in MATLAB for both the two-sector and two-dependence cases, we conducted experiments with various dimensions. For the two-sector case, we considered dimensions $d = 5, 10, \ldots, 60$ and randomly generated 25 pairs of $(\alpha, \beta) \in [0, 1]^2$ for each dimension. We calculated $\gamma_u(\alpha, \beta)$ for each pair and then randomly generated $\gamma \in [0, \gamma_u(\alpha, \beta)]$. Similarly, for the two-dependence case, we randomly generated 25 pairs (α, β) from show convex polytope (4.4) dimensions 5, 10, . . . , 60. For each dimension, we report the running times across all non-linear instances but in Figure 7.

As shown in the figure, the running times are under four minutes for all instances from both sets of TDMs, for dimensions up to 60. These experimental results show that the interior-point algorithm demonstrates good scaling for our problem. It is important to highlight, however, that the running times exhibit a non-linear, but less than cubic, growth.

5. An application to weather-related insurance

The global increase in both the frequency and severity of natural catastrophes, due to climate change or otherwise, and accompanying increased competition for food resources with the shrinking of arable land have contributed to the importance of crop insurance. Excessive rain and, in extreme cases, flooding or near-drought conditions are key risk events covered by crop insurance, as these can severely disrupt a farming business. From the point of view of an insurer or a reinsurer with a portfolio of crop insurance policies spanning a geographical area, an accurate assessment of the dependence structure of rainfall at different locations, especially in the tails, is a critical input to evaluate the geographical diversification or lack thereof, the latter affecting both pricing and capital considerations. In the following, we show how the methodology above can help develop a one-factor copula model that embodies the tail dependence exhibited in a rainfall dataset. We consider below the daily rainfall data from the European Climate Assessment & Dataset project (Klein Tank *et al*., 2002).[1] More precisely, we focus on a recent ten-year data (2011/1/1–2020/12/31) of 33 French cities with 1% of missing value.

[1]The data set can be downloaded at https://www.ecad.eu/dailydata/predefinedseries.php.

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

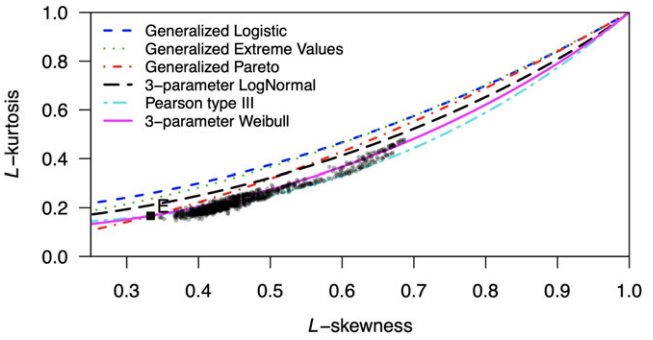


Figure 8. *The L-moments ratio diagrams of some common distributions.*

A random vector (X,Y) of rainfall in two cites on a given day, as suggested by Shimizu (1993), is modeled as described below:

$$\Pr(X=0, Y=0) = \lambda_0; \qquad \Pr(0 < X \le x, Y=0) = \lambda_1 F_1(x);$$

$$\Pr(X=0, 0 < Y \le y) = \lambda_2 F_2(y); \qquad \Pr(0 < X \le x, 0 < Y \le y) = \lambda_3 F(x,y),$$

where $x, y, \lambda_i > 0, i = 0, \ldots, 3$ satisfy $\sum_{i=0}^{3} \lambda_i = 1$, F_1 and F_2 are univariate continuous distribution functions supported on $(0, \infty)$, and F is a bivariate continuous joint distribution function supported on the positive quadrant. In the following, as we restrict our interest to the upper tail dependence of rainfall in each pair of cities, we only need to model the joint cumulative distribution function F.

Our approach is to model F by choosing a one-factor copula with parameters $(\tilde{\theta}, \tilde{\delta})$ and appropriate marginals F_X and F_Y. Following Serinaldi (2008), we use the L-moments ratio diagrams (see Hosking (1990)) for the purpose of choosing an appropriate distribution family for modeling the marginals. In Figure 8, the L-skewness versus L-kurtosis graph, we plot the $2 \times \binom{33}{2}$ empirical estimates from the marginals of truncated (to $(0, \infty)^2$) bivariate distributions along with the curves from some commonly used distribution families. Figure 8 suggests the three-parameter Weibull distribution as a good fit. To estimate the parameters, we use the fact that the Weibull distribution is a reverse Generalized Extreme Value distribution. This allows us to employ the L-moments-based estimators for the parameters of Generalized Extreme Value distributions listed in Table 2 of Hosking (1990). These estimators are conveniently made available on the R statistical environment by the packages lmom and lmomco (see Hosking, 2019 and Asquith, 2020).

Having estimated the marginals as described above, since we seek a model that captures well the tail dependence, we begin by estimating the tail dependence matrix. Among many nonparametric estimators for the tail dependence coefficient, the Capéraà-Fougères-Genest estimator of the tail dependence coefficient (see Capéraà *et al*., 1997) given by

$$\widehat{\chi}^{\mathrm{CFG}} = 2 - \prod_{i=1}^{n} \left(\frac{\log \min(u_{1,i}, u_{2,i})}{\log \max(u_{1,i}, u_{2,i})} \right)^{\frac{1}{2n}},$$

where n is the sample size, $u_{1,i} = \hat{F}_X(x_{1,i})$, and $u_{2,i} = \hat{F}_Y(x_{2,i})$, for $i = 1, \ldots, n$, shows the best performance in, for example Frahm *et al*. (2005). Using this estimator we compute the 33×33 estimated TDM, denoted by $\widehat{\mathbf{T}}_{33}^{\mathrm{CFG}}$, which interestingly is a valid TDM.

On the other hand, there are 236 of 5456, 3×3 principal submatrices that are determined to be not feasible TDMs by Proposition 3.2. It is noteworthy that 236 represents only about 4.3% of 5456, far smaller a proportion than the 26% relative volume of $\mathcal{T}_3'$ reported in Theorem 3.1. This, in particular, implies that $\widehat{\mathbf{T}}_{33}^{\mathrm{CFG}} \in \mathcal{T}_{33} \setminus \mathcal{T}_{33}'$, suggesting the use of the Frobenius norm closest one-factor feasible TDM, say $\widetilde{\mathbf{T}}_{33}$, to $\widehat{\mathbf{T}}_{33}^{\mathrm{CFG}}$ for estimating the parameters of a one-factor copula. As the upper tail-dependence

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

coefficient only determines $\theta\delta$, for our illustrative purposes we fix $\delta = 1.2$, but in practice, one can use constrained maximum likelihood or any other suitable statistical paradigms.

To illustrate the goodness of fit resulting from the above inference procedure, we consider two sets of three cities. The first set consisting of Bourges, Nice, and Saint-Girons is chosen as its corresponding principal submatrix of $\widehat{\mathbf{T}}_{33}^{\text{CFG}}$ is not in $\mathcal{T}_3'$; this is so as according to Proposition 3.2, infeasibility is implied by

$$\Psi\left(\Psi_1^{-1}(0.0820), \Psi_1^{-1}(0.2254)\right) = \Psi\,(0.2772, 0.4652) = 0.0325 > 0.0003.$$

This matrix, its counterpart in the feasible TDM $\widetilde{\mathbf{T}}_{33}$, and the corresponding one-factor copula parameters are given by

$$\begin{bmatrix} 1 & 0.0820 & 0.2254 \\ 0.0820 & 1 & 0.0003 \\ 0.2254 & 0.0003 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0.1011 & 0.2459 \\ 0.1011 & 1 & 0.0776 \\ 0.2459 & 0.0776 & 1 \end{bmatrix}, \text{ and } \begin{bmatrix} 1.0032 \\ 0.3502 \\ 0.6662 \end{bmatrix},$$

respectively. The above being quite similar suggests a good fit in the upper tails.

In Figure 9, the contour plots of the bivariate marginals from the fitted one-factor copula, with Weibull marginals, for these three cities being quite alike the contours of the kernel density estimate suggests a good overall fit. For this plot, we used the function sim1fact in the R package CopulaModel, see Joe (2014).

The second set of cities consisting of Bourges, Lyon, and Strasbourg was chosen as its corresponding principal submatrix of $\widehat{\mathbf{T}}_{33}^{\text{CFG}}$ is in $\mathcal{T}_3'$; this is so as according to Proposition 3.2, feasibility is implied by

$$\Psi\left(\Psi_1^{-1}(0.2620), \Psi_1^{-1}(0.2842)\right) = \Psi\,(0.5174, 0.5510) = 0.1168 < 0.2302.$$

This matrix, its counterpart in $\widetilde{\mathbf{T}}_{33}$, and the corresponding one-factor copula parameters are given by

$$\begin{bmatrix} 1 & 0.2842 & 0.2620 \\ 0.2842 & 1 & 0.2302 \\ 0.2620 & 0.2302 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0.2714 & 0.2772 \\ 0.2714 & 1 & 0.2267 \\ 0.2772 & 0.2267 & 1 \end{bmatrix}, \text{ and } \begin{bmatrix} 1.0032 \\ 0.7444 \\ 0.7636 \end{bmatrix},$$

respectively. Note that these matrices are nearly identical. Also, the contour plots of the bivariate marginals from the fitted one-factor copula for these three cities are similar to the contours of the kernel density estimate, as shown in Figure 10, suggesting a good overall fit. Relatively speaking, notice that we see a better alignment of contours in Figure 10 than in Figure 9, agreeing with the observations on the empirical TDM.

Among the 40,920, 4×4 principal submatrices, as expected by Theorem 3.1, none are feasible as their distance to the nearest feasible matrix using the optimization problem of (4.1) is positive. Propositions 3.3 and 3.4, by partitioning these matrices into cells depicted in Figure A1, give a qualitative description of proximity to feasibility. In particular, 5691 are located in cells 1 or 7. Additionally, 11,117 matrices are situated in cells adjacent to cells 1 or 7, namely cells 2, 3, 8, and 13. To be precise, what we mean by the qualitative statement is the stochastic ordering depicted in Figure 11, which plots the kernel density estimates of the log-squared distance of each member to feasibility categorized by their membership in the cells. These distances were computed using the optimization problem of (4.1).

Consider the following set of four cities: Strasbourg, Nantes, Troyes, and Caen. Their principal submatrix within $\widehat{\mathbf{T}}_{33}^{\text{CFG}}$ is not in the cells 1 or 7 of Table A2. This matrix, its counterpart in $\widetilde{\mathbf{T}}_{33}$, and the corresponding one-factor copula parameters are given by

$$\begin{bmatrix} 1 & 0.2553 & 0.3106 & 0.2761 \\ 0.2553 & 1 & 0.2691 & 0.3451 \\ 0.3106 & 0.2691 & 1 & 0.3587 \\ 0.2761 & 0.3451 & 0.3587 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0.2569 & 0.2801 & 0.2893 \\ 0.2569 & 1 & 0.3124 & 0.3230 \\ 0.2801 & 0.3124 & 1 & 0.3538 \\ 0.2893 & 0.3230 & 0.3538 & 1 \end{bmatrix}, \text{ and } \begin{bmatrix} 0.7636 \\ 0.8846 \\ 1.0225 \\ 1.0881 \end{bmatrix}.$$

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

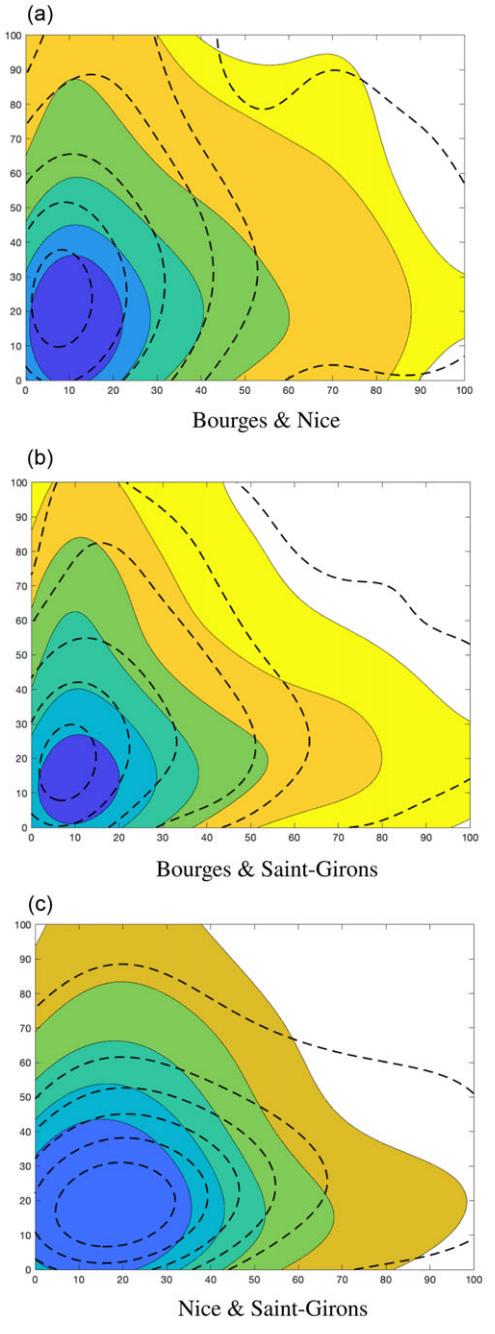


Figure 9. *Contours of the fitted kernel and one-factor-copula-based distribution (dotted lines).*

Proposition 3.4 indicates that $\mathcal{T}_4'$ can be divided into $4! = 24$ distinct subsets, each reflecting a permutation of the set $\{1, 2, 3, 4\}$, aligning with the sequence of the η_i values. Furthermore, according to Proposition 3.3, these subsets can be subdivided based on permutations of $\mathcal{I}_4$, arranged according to the tail dependence coefficients. Due to symmetry, this results in $6!/4! = 30$ such subdivisions. Notably, among these, only two subsets are nonempty. Located in cell 9 of Table A2 and Figure A1, the matrix's projection falls within cell 1, preserving the ordering of the underlying η_i's. Given the proximity of the TDM *projection*, this is not surprising.

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

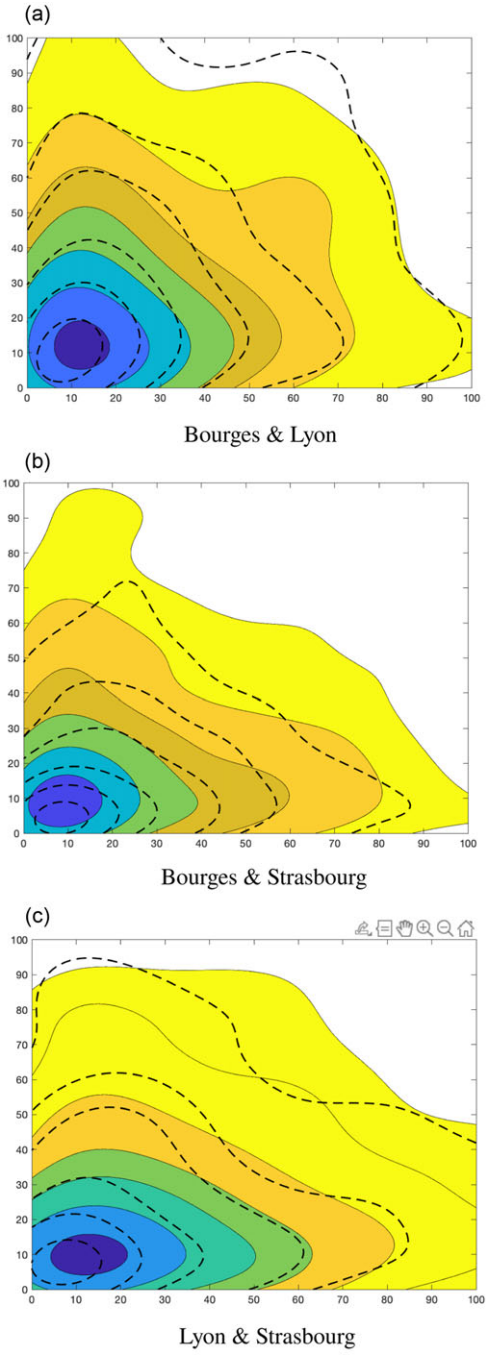


Figure 10. *Contours of the fitted kernel and one-factor-copula-based distribution (dotted lines).*

6. Broader implications of our analysis

To begin with, we note that the tail dependence coefficients of the following linking copulas share the same functional form as that of the BB1 family: Clayton copula (see, section 7.1 of McNeil *et al*., 2015 or Family B4 in section 5.1 of Joe, 1997), Galambos survival copula (see, section 4.9 of Joe, 2014),

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

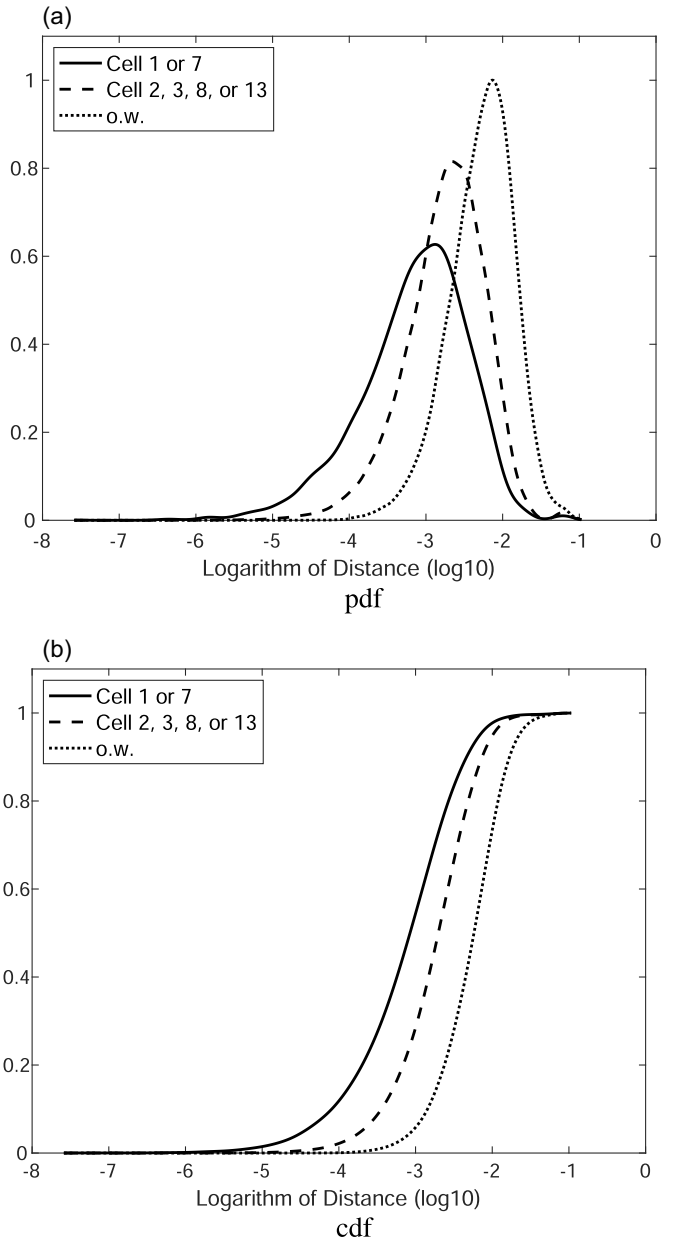


Figure 11. *Kernel density estimate plots for the three groups.*

BB4 survival copula (see, section 5.2 of Joe, 1997), and BB7 copula (see, Table A1 in the Appendix). Hence, our results in Section 3 carry over to the one-factor model with these linking copulas as well.

Suppose that the tail dependence coefficient is defined by $\Psi : \Theta^2 \to [0, 1]$, with the parameter space Θ being an interval of $\mathbb{R}$. In other words, we suppose that the i-th coordinate is associated with a parameter θ_i (possibly with parametrization, see FDG copulas below) such that the tail dependence coefficient between the i-th and the j-th coordinate equals $\Psi(\theta_i, \theta_j)$. Further, if Ψ is a symmetric, strictly monotone (not necessarily increasing) function with continuous derivatives, then the results outlined in Section 3

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

materially continue to hold. Only the computation of the volume and the non-convexity of $\mathcal{O}_3'$ referred to in Remark 3.1 use the specificity of the BB1 copula.

Examples of such linking copulas include the Gumbel survival copula (see, Section 7.1 of McNeil *et al*., 2015), B5 survival copula (see, Section 5.1 of Joe, 1997), BB1 survival copula, and BB3 survival copula (see, Section 5.2 of Joe, 1997). Interestingly, for all of these examples, the tail dependence coefficient between two coordinates has the form,

$$\Psi(x,y) := \int_0^\infty \left[1-\left(1+w^{-x}\right)^{\frac{1}{x}-1}\right]\left[1-\left(1+w^{-y}\right)^{\frac{1}{y}-1}\right] dw,$$

$$\text{with } \Psi(x,\infty) := \Psi(\infty,x) = 1-\int_0^1 (1+w^{-x})^{\frac{1}{x}-1} dw = 2-2^{\frac{1}{x}},$$

$$\text{and } \Psi(\infty,\infty) := 1, \qquad 1 \le x,y < \infty.$$

Note that the above Ψ is a symmetric strictly increasing function on $[1,\infty]^2$ with continuous partial derivatives in the interior.

FDG copula family – one-factor copula with Durante Generators – is a one-factor copula model with the Durante class of copulas, a tractable nonparametric class, as linking copulas. See Mazo *et al*. (2016) for the FDG copula and Durante (2006) for the Durante class.

The Durante class of copulas is given by

$$C(u,v) := (u \wedge v)\cdot f(u \vee v), \quad u,v \in [0,1]^2,$$

where $f:[0,1]\to[0,1]$, the generator, is a differentiable increasing function satisfying $f(1)=1$ with $u \mapsto f(u)/u$ a decreasing function. Note that the copula represents an exchangeable distribution, and by considering the conditional distribution of X given $Y \le v$, it is easy to see the need for the restrictions on f with differentiability added for the existence of the density. The lower and upper tail dependence coefficients are $f(0)$ and $1-f'(1)$, respectively.

It is interesting to note that the bivariate margins of an FDG copula are themselves members of the Durante class. Hence, if f_1 and f_2 are the generators of the first two coordinates, then their bivariate margin is a Durante class copula with generator

$$u \mapsto f_1(u)\cdot f_2(u) + u\int_u^1 f_1'(v) f_2'(v) \mathrm{d}v.$$

In particular, $f_1(0)\cdot f_2(0)$ and $(1-f'_1(1))(1-f'_2(1))$ are its lower and upper tail dependence coefficients, respectively. This implies that the ordering of the tail dependence coefficients of the linking copula imposes a partial ordering on the tail dependence coefficients of the FDG copula. Some parametric examples of generators are the Cuadras-Augé generator $u \mapsto u^{1-\theta}$ for $\theta \in [0,1]$, Fréchet generator $u \mapsto (1-\theta)u+\theta$, for $\theta \in [0,1]$ and the Durante-exponential generators

$$u \mapsto \exp\left(\frac{u^\theta - 1}{\theta}\right), \qquad \theta > 0.$$

Finally, since for any copula C, $C^*(u_1,u_2) := C(1-u_1,1-u_2)$ also forms a copula, and the lower tail dependence coefficient with respect to C corresponds to the upper tail dependence coefficient with respect to C^*, and vice versa, the applicability of our results is invariant to the focus being on the left or the right tail. Moreover, for copula families like the FDG copula, the results apply to both the tails of the copula.

7. Conclusion

In statistical modeling with reasonable complexity, it is common that explicit modeling choices sweep under the carpet implicit model assumptions, some of which may be unjustified. In this study, we

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

investigate if this happens in the context of modeling high-dimensional data that exhibits nonlinear tail-dependence when we use a one-factor copula model with a parametric family for the linking copulas. Apart from deriving basic tail dependence properties for this mode, we precisely describe the implicit ordering of the tail dependence coefficients, an implementation of a solution to the formulated optimization problem to help model data from a tail dependence perspective and conclude with an illustrative application.

We believe this paper's qualitative implications are broad; as a next step, an extension to the multifactor setting will be interesting as we continue to retain the linear-in-dimension parsimonious parameterization of the one-factor copula models.

Acknowledgments. We extend our sincere gratitude to the editor for choosing two excellent referees and to the referees whose detailed review led to a significantly improved article.

Competing interests. There are no competing interest to declare.

References

Aas, K., Czado, C., Frigessi, A. and Bakken, H. (2009) Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, **44** (2), 182–198.

Asquith, W. (2020) lmomco-package. URL https://rdrr.io/cran/lmomco/man/lmomco-package.html.

Bassamboo, A., Juneja, S. and Zeevi, A. (2008) Portfolio credit risk with extremal dependence: Asymptotic analysis and efficient simulation. *Operations Research*, **56** (3), 593–606.

Bedford, T. and Cooke, R.M. (2001) Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence*, **32**, 245–268.

Bedford, T. and Cooke, R.M. (2002) Vines–a new graphical model for dependent random variables. *The Annals of Statistics*, **30** (4), 1031–1068.

Berend, D. and Tassa, T. (2010) Improved bounds on Bell numbers and on moments of sums of random variables. *Probability and Mathematical Statistics*, **30** (2), 185–205.

Blanchet, J. and Davison, A.C. (2011) Spatial modeling of extreme snow depth. *Annals of Applied Statistics*, **5** (3), 1699–1725.

Brechmann, E.C., Czado, C. and Aas, K. (2012) Truncated regular vines in high dimensions with application to financial data. *Canadian Journal of Statistics*, **40** (1), 68–85.

Capéraà, P., Fougères, A.-L. and Genest, C. (1997) A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika*, **84** (3), 567–577.

Chen, H., MacMinn, R. and Sun,T. (2015) Multi-population mortality models: A factor copula approach. *Insurance: Mathematics and Economics*, **63**, 135–146.

Chen, S., Tong, Z. and Yang, Y. (2023) Portfolio losses driven by idiosyncratic risk. Mimeo, Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4626393.

Czado, C. and Nagler, T. (2022) Vine copula based modeling. *Annual Review of Statistics and Its Application*, **9**, 453–477.

Davis, R.A. and Mikosch, T. (2009) The extremogram: A correlogram for extreme events. *Bernoulli*, **15** (4), 977–1009.

De Luca, G., Ruscone, M.N. and Amati, V. (2023) The use of conditional copula for studying the influence of economic sectors. *Expert Systems with Applications*, 120582.

Donnelly, C. and Embrechts, P. (2010) The devil is in the tails: Actuarial mathematics and the subprime mortgage crisis. *ASTIN Bulletin: The Journal of the IAA*, **40** (1), 1–33.

Durante, F. (2006) A new class of symmetric bivariate copulas. *Journal of Nonparametric Statistics*, **18** (7-8), 499–510. doi: 10.1080/10485250701262242.

Embrechts, P., Frey, R. and McNeil, A. (2011) *Quantitative Risk Management*. Princeton, New Jersey: Princeton University Press.

Embrechts, P., Hofert, M. and Wang, R. (2016) Bernoulli and tail-dependence compatibility. *Annals of Applied Probability*, **26** (3), 1636–1658.

Fiebig, U.-R., Strokorb, K. and Schlather, M. (2017) The realization problem for tail correlation functions. *Extremes*, **20** (1), 121–168.

Frahm, G., Junker, M. and Schmidt, R. (2005) Estimating the tail-dependence coefficient: Properties and pitfalls. *Insurance: Mathematics and Economics*, **37** (1), 80–100.

Frees, E.W. and Valdez, E.A. (1998) Understanding relationships using copulas. *North American Actuarial Journal*, **2** (1), 1–25.

Frobenius, F.G. (1900) *Über die charactere der symetrischen gruppe, s'ber*. Akad. Wiss. Berlin, pp. 516–534.

Gao, L. (2022) Spatial models of insurance losses for claims management. Ph.D. Thesis, University of Wisconsin, Madison.

Gawrilow, E. and Joswig, M. (2000) Polymake: A framework for analyzing convex polytopes. In Polytopes–Combinatorics and Computation, pp. 43–73. Springer.

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

Hosking, J.R.M. (2019) lmom-package. URL https://rdrr.io/cran/lmom/man/lmom-package.html.
Hosking, J.R.M. (1990) L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society: Series B (Methodological)*, **52** (1), 105–124.
Hsieh, M.-H., Tsai, C.J. and Wang, J.L. (2021) Mortality risk management under the factor copula framework–with applications to insurance policy pools. *North American Actuarial Journal*, **25** (sup1), S119–S131.
Hua, L., Xia, M. and Basu, S. (2017) Factor copula approaches for assessing spatially dependent high-dimensional risks. *North American Actuarial Journal*, **21** (1), 147–160.
Janßen, A., Neblung, S. and Stoev, S. (2023) Tail-dependence, exceedance sets, and metric embeddings. In Extremes, pp. 1–39.
Jaworski, P., Durante, F., Hardle, W.K. and Rychlik, T. (2010) In *Copula Theory and its Applications*, Vol. 198. Springer.
Joe, H. (1997) *Multivariate Models and Dependence Concepts*, Monographs on Statistics and Applied Probability, Vol. 73. London: Chapman & Hall.
Joe, H. (2011) Tail dependence in vine copulae. *In Dependence Modeling*, pp 165–187. Hackensack, NJ: World Scientific Publishing.
Joe, H. (2014) *Dependence Modeling with Copulas*. Boca Raton, Florida: CRC Press.
Joe, H. and Hu, T. (1996) Multivariate distributions from mixtures of max-infinitely divisible distributions. *Journal of Multivariate Analysis*, **57**, 240–265.
Johnston, N. (2014) Counting the possible orderings of pairwise multiplication. URL http://www.njohnston.ca/2014/02/counting-the-possible-orderings-of-pairwise-multiplication/.
Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A.F.V., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L.V. and Petrovic, P. (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the european climate assessment. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, **22** (12), 1441–1453.
Knuth, D.E. (1998) Sorting and searching. In The Art of Computer Programming, Vol. 3.
Kolman, M. (2014) A one-factor copula-based model for credit portfolios. *Journal of Risk*, **17** (2), 93–132.
Kroese, D.P., Taimre, T. and Botev, Z.I. (2013) *Handbook of Monte Carlo Methods*, Vol. 706. Hoboken, New Jersey: John Wiley & Sons.
Krupskii, P. and Joe, H. (2013) Factor copula models for multivariate data. *Journal of Multivariate Analysis*, **12**, 85–101.
Krupskii, P. and Joe, H. (2020) Flexible copula models with dynamic dependence and application to financial data. *Econometrics and Statistics*, **16**, 148–167.
Kularatne, T.D., Li, J. and Pitt, D. (2021) On the use of archimedean copulas for insurance modelling. *Annals of Actuarial Science*, **15** (1), 57–81.
Lan, M., Shao, Y., Zhu, J., Lo, S. and Ng, S.T. (2021) A hybrid copula-fragility approach for investigating the impact of hazard dependence on a process facility's failure. *Process Safety and Environmental Protection*, **149**, 1017–1030.
Lazoglou, G. and Anagnostopoulou, C. (2019) Joint distribution of temperature and precipitation in the mediterranean, using the copula method. *Theoretical and Applied Climatology*, **135**, 1399–1411.
Li, D.X. (2000) On default correlation: A copula function approach. *The Journal of Fixed Income*, **9** (4), 43–54.
Li, J., Balasooriya, U. and Liu, J. (2021) Using hierarchical archimedean copulas for modelling mortality dependence and pricing mortality-linked securities. *Annals of Actuarial Science*, **15** (3), 505–518.
Lu, J.C. and Bhattacharyya, G.K. (1990) Some new constructions of bivariate weibull models. *Annals of the Institute of Statistical Mathematics*, **42**, 543–559.
MacMahon, P.A. (1909) Memoir on the theory of the partitions of numbers.–part iv. Philosophical Transactions of the Royal Society of London. Series A, **209** (441–458), 153–175.
MathWorks (2021) fmincon - nonlinear constrained optimization (optimization toolbox).
Mazo, G., Girard, S. and Forbes, F. (2016) A flexible and tractable class of one-factor copulas. *Statistics and Computing*, **26**, 965–979.
McNeil, A. J., Frey, R. and Embrechts, P. (2015) Quantitative Risk Management. Princeton Series in Finance. Princeton, NJ: Princeton University Press, revised edition. Concepts, techniques and tools.
Morales-Napoles, O. (2010) Counting vines. In Dependence Modeling: Vine Copula Handbook, pp. 189–218. World Scientific.
Nelsen, R.B. (2006) *An Introduction to Copulas*. New York, NY: Springer.
Nikoloulopoulos, A.K., Joe, H. and Li, H. (2012) Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis*, **56** (11), 3659–3673.
Oh, D.H. and Patton, A.J. (2017) Modeling dependence in high dimensions with factor copulas. *Journal of Business & Economic Statistics*, **35** (1), 139–154.
Oh, R., Ahn, J.Y. and Lee, W. (2021) On copula-based collective risk models: From elliptical copulas to vine copulas. *Scandinavian Actuarial Journal*, **2021** (1), 1–33.
Othus, M. and Li, Y. (2010) A gaussian copula model for multivariate survival data. *Statistics in Biosciences*, **2**, 154–179.
Da Silva, P.P., Rebelo, P.T. and Afonso, C. (2014) Tail dependence of financial stocks and cds markets – evidence using copula methods and simulation-based inference. *Economics*, **8** (1), 20140039.
Renard, B. and Lang, M. (2007) Use of a gaussian copula for multivariate extreme value analysis: Some case studies in hydrology. *Advances in Water Resources*, **30** (4), 897–912.

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

Righi, M.B. and Ceretta, P.S. (2013) Analyzing the dependence structure of various sectors in the brazilian market: A pair copula construction approach. *Economic Modelling*, **35**, 199–206.
Robbins, H. (1955) A remark on stirling's formula. *The American Mathematical Monthly*, **62** (1), 26–29.
Salmon, F. (2009) Recipe for disaster: The formula that killed wall street. *Wired Magazine*, **17** (3), 17–03.
Salvadori, G., De Michele, C., Kottegoda, N.T. and Rosso, R. (2007) *Extremes in Nature: An Approach Using Copulas*, Vol. 56. Dordrecht, The Netherlands: Springer Science & Business Media.
Serinaldi, F. (2008) Analysis of inter-gauge dependence by kendall's τ k, upper tail dependence coefficient, and 2-copulas with application to rainfall fields. *Stochastic Environmental Research and Risk Assessment*, **22** (6), 671–688.
Shimizu, K. (1993) A bivariate mixed lognormal distribution with an analysis of rainfall data. *Journal of Applied Meteorology and Climatology*, **32** (2), 161–171.
Shorack, G.R. (2017) *Probability for Statisticians*. Cham, Switzerland: Springer.
Shyamalkumar, N.D. and Tao, S. (2020) On tail dependence matrices. *Extremes*, **23** (2), 245–285.
Shyamalkumar, N.D. and Tao, S. (2022) t-copula from the viewpoint of tail dependence matrices. *Journal of Multivariate Analysis*, 105027.
Sklar, M. (1959) Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de statistique de l'Universite Paris*, **8**, 229–231.
Song, P.X.-K., Li, M. and Yuan, Y. (2009) Joint regression analysis of correlated data using gaussian copulas. *Biometrics*, **65** (1), 60–68.
Strokorb, K. (2013) Characterization and construction of max-stable processes. Ph.D. Thesis. URL http://hdl.handle.net/11858/00-1735-0000-0001-BB44-9.
Tang, Q., Tang, Z. and Yang, Y. (2019) Sharp asymptotics for large portfolio losses under extreme risks. *European Journal of Operational Research*, **276** (2), 710–722.
Thrall, R.M. (1952) A combinatorial problem. *Michigan Mathematical Journal*, **1** (1), 81–88.
Van de Vyver, H. and Van den Bergh, J. (2018) The gaussian copula model for the joint deficit index for droughts. *Journal of Hydrology*, **561**, 987–999.

A. Appendix

Lemma A.1. *$C_{i|0}(u|uw)$ is bounded above by the function g_i, for $i = 1, \ldots, d$, respectively, where*

$$g_i(w) := \begin{cases} 1, & 0 \le w \le 1; \\ 2^{\frac{1}{\delta_i}-1} w^{-(1+\theta_i)}, & 1 < w, \end{cases}$$

for $\theta_i \in (0, \infty)$.

Proof. Since, $C_{i|0}(u|uw)$ is a conditional distribution function, it is clearly bounded by 1. Hence, it suffice to consider the case $1 < w < (1/u)$. Since,

$$\frac{u^{-\theta_i} - 1}{(uw)^{-\theta_i} - 1} > 1,$$

the first term of (2.4) is bounded above by $2^{1/\delta_i - 1}$. Using the fact that $1 - (uw)^{\theta_i} \ge 0$, we can bound the second term of (2.4) above by $w^{-(1+\theta_i)}$, completing the proof. □

The following lemma is a version of Exercise 3.3 in Shorack (2017).

Lemma A.2. *Let $X(t, \cdot)$ be a measurable function on the measure space $(\Omega, \mathcal{A}, \mu)$ for each $t \in [a, b]$. Moreover, for a.e. ω, the partial derivative $\frac{\partial}{\partial t} X(t, \omega)$ exists for all t in the non-degenerate interval [a,b], and that*

$$\left| \frac{\partial}{\partial t} X(t, \omega) \right| \le Y(\omega), \quad \text{for all } t \in [a, b],$$

for an integrable Y. Then

$$\frac{d}{dt} \int_\Omega X(t, \omega) d\mu(\omega) = \int_\Omega \left[\frac{\partial}{\partial t} X(t, \omega) \right] d\mu(\omega).$$

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

Proof of Proposition 2.1. By splitting the integral representation of $\Psi(\cdot,\cdot)$ (see (2.6)) and change of variables on the integral part on $[1,\infty)$, we get

$$\Psi(x,y) = \int_0^1 (1+w^x)^{-\frac{1}{x}-1}(1+w^y)^{-\frac{1}{y}-1}\,dw + \int_0^1 w^{-2}\left(1+w^{-x}\right)^{-\frac{1}{x}-1}\left(1+w^{-y}\right)^{-\frac{1}{y}-1}dw$$
$$= \int_0^1 \left(1+w^{x+y}\right)(1+w^x)^{-1/x-1}(1+w^y)^{-1/y-1}\,dw.$$

This yields, for $(x,y)\in(0,\infty)^2$,

$$\Psi(x,y) < \int_0^1 (1+w^x)^{-1/x-1}(1+w^y)^{-1/y}\,dw < \Psi(x,\infty) < \Psi(\infty,\infty) = 1.$$

Hence, it suffices to show that Ψ has continuous partial derivatives which are positive on $(0,\infty)^2$. Toward this, we claim that,

$$\frac{\partial}{\partial x}\Psi(x,y) = \int_0^1 \left(1+w^{x+y}\right)(1+w^x)^{-\frac{1}{x}-1}(1+w^y)^{-\frac{1}{y}-1}\left[\frac{\ln(1+w^x)}{x^2}\right.$$
$$\left. - \left(1+\frac{x(1-w^y)}{1+w^{x+y}}\right)\frac{w^x\ln(w)}{x(1+w^x)}\right]dw. \quad \text{(A1)}$$

Note that the integrand is positive. Hence, it suffices to establish the above expression and its continuity. Toward this, note that the integrated is bounded by

$$2\left[x^{-2}\ln 2 + (1+x^{-1})w^x\ln(1/w)\right], \quad \text{(A2)}$$

and since this is monotone decreasing in x and integrable for $x>0$, we have by Lemma A.2 the validity of (A1) on $(0,\infty)^2$. Moreover, since the integrate in (A1) is continuous, the bound in (A2) and the dominated convergence theorem together imply the continuity of the partial derivatives of Ψ on $(0,\infty)^2$. □

Proof. (Corollary 2.1.) Part (i) directly follows from Proposition 2.1. For part (ii), the proof follows from observing that by Proposition 2.1, $x\geq x'$ (resp., $x<x'$) implies $y\leq y'$ (resp., $y>y'$). □

Lemma A.3. *Let* $0<a\leq b<1$ *and* $w\in(\Psi_1^{-1}(a),\infty]$ *such that* $a=\Psi(\Psi_1^{-1}(b),w)$*, and define the function* $\eta_{a,b}:(0,\infty)\to(\Psi_1^{-1}(a),w)$ *as* $a=\Psi(\Psi_1^{-1}(b)+x,\eta_{a,b}(x))$*. Then the function* $\eta_{a,b}$ *is continuous.*

Proof. Let $s>0$, $\epsilon>0$, and $u:=\eta_{a,b}(s/2)\vee(\Psi_1^{-1}(b)+s+\epsilon)$. According to Lemma A.4, with α increasing in its first argument and β decreasing, we have

$$\alpha(\Psi_1^{-1}(a),y)\leq\frac{\partial}{\partial y}\Psi(x,y)\leq\beta\left(\Psi_1^{-1}(a)\right):=M,\quad\forall x,y\in\left[\Psi_1^{-1}(a),u\right]^2.$$

The continuity of $\alpha(\Psi_1^{-1}(a),\cdot)$ implies that

$$m:=\min_{y\in[\Psi_1^{-1}(a),u]}\alpha(\Psi_1^{-1}(a),y)>0.$$

For $\delta:=\min\{s/2,m\epsilon/M\}$ and $t\in(s-\delta,s+\delta)$, we observe that

$$m|\eta_{a,b}(t)-\eta_{a,b}(s)|\leq|\Psi(\Psi_1^{-1}(b)+t,\eta_{a,b}(s))-\Psi(\Psi_1^{-1}(b)+t,\eta_{a,b}(t))|$$
$$=|\Psi(\Psi_1^{-1}(b)+t,\eta_{a,b}(s))-a|$$
$$=|\Psi(\Psi_1^{-1}(b)+t,\eta_{a,b}(s))-\Psi(\Psi_1^{-1}(b)+s,\eta_{a,b}(s))|\leq M|t-s|.$$

This implies $|\eta_{a,b}(t)-\eta_{a,b}(s)|\leq\frac{M}{m}|t-s|<\epsilon$, demonstrating the continuity of $\eta_{a,b}$ on $(0,\infty)$. □

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

Lemma A.4. *A lower and an upper bound for the partial derivative* $\frac{\partial}{\partial y}\Psi(x,y)$ *are given by functions* $\alpha(x,y)$ *and* $\beta(y)$*, where*

$$\alpha(x,y) := \frac{2^{-(1/x+1/y+3)}}{y(y+1)^2}, \quad \textit{and} \quad \beta(y) := 2\left[y^{-2}\ln 2 + (y+y^2)^{-1}\right], \textit{ for } y>0, \textit{ respectively.}$$

Proof. By (A1), we establish a lower bound of $\frac{\partial}{\partial y}\Psi(x,y)$ as

$$\frac{\partial}{\partial y}\Psi(x,y) \geq -\frac{2^{-(1/x+1/y+3)}}{y}\int_0^1 w^y \ln(w)dw = \alpha(x,y).$$

Similarly, according to equation (A2), the upper bound is calculated to be

$$\frac{\partial}{\partial y}\Psi(x,y) \leq \int_0^1 2\left[y^{-2}\ln 2 + (1+y^{-1})w^y \ln(1/w)\right]dw = \beta(y).$$

□

Theorem A.1 (Sard's Theorem). *Let* $f:\mathbb{R}^n \to \mathbb{R}^m$ *be* C^k*, where* $k \geq \max\{n-m+1, 1\}$*. Let* $M \subseteq \mathbb{R}^n$ *denote the critical set of f, which is the set of points* $x \in \mathbb{R}^n$ *at which the Jacobian matrix of f has rank* $< m$*. Then the image f(M) has Lebesgue measure 0 in* $\mathbb{R}^m$*.*

Lemma A.5. *For* $d \geq 3$*, we have*

$$\frac{d\prod_{i=1}^{d-1} i!}{\prod_{i=1}^{d-1}(2i-1)!} \leq e^{-\Omega(d^2)}. \tag{A3}$$

Proof. We use a looser version of the bound presented in Robbins (1955) that the following is valid for all positive integers n:

$$\sqrt{2\pi n}\cdot n^n e^{-n} < n! < \sqrt{2\pi n}\cdot n^n e^{-n+1}.$$

Then

$$\begin{aligned}\frac{d\prod_{i=1}^{d-1} i!}{\prod_{i=1}^{d-1}(2i-1)!} &\leq \frac{d\prod_{i=1}^{d-1}\sqrt{2\pi i}\cdot i^i e^{-i+1}}{\prod_{i=1}^{d-1}\sqrt{2\pi(2i-1)}\cdot(2i-1)^{2i-1}e^{-2i+1}} = d\prod_{i=1}^{d-1}\frac{e^i}{(2-1/i)^{i+1/2}(2i-1)^{i-1}} \\ &\leq \frac{de^{d(d-1)/2-1}}{1.5^{d(d-1)/2+(d-1)/2-1}(2e)^{d(d-3)/2-1}} = \frac{de^d}{3^{d(d-3)/2-1}1.5^{(3d-1)/2}} \\ &\leq d3^{-(d^2-5d)/2-1}.\end{aligned}$$

□

The following lemma follows directly from Proposition 3.2.

Lemma A.6. *If*

$$\begin{bmatrix} 1 & a & b \\ a & 1 & 0 \\ b & 0 & 1 \end{bmatrix} \tag{A4}$$

is in $\mathcal{T}_3'$*, then at least one of a and b must be equal to zero.*

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

Table A1. *Related functions and the approximate volume of the one-factor copula family with the linking copula being the BB4 and BB7 families.*

Family	BB4	BB7
$C(u,v;\theta_i,\delta_i)$, $i=1,2,3.$	$\left\{u^{-\theta_i}+v^{-\theta_i}-1-\left[(u^{-\theta_i}-1)^{-\delta_i}+(v^{-\theta_i}-1)^{-\delta_i}\right]^{-\frac{1}{\delta_i}}\right\}^{-\frac{1}{\theta_i}}$, $\theta_i\geq 0,\ \delta_i>0.$	$1-\left\{1-\left[\left(1-(1-u)^{\theta_i}\right)^{-\delta_i}+\left(1-(1-v)^{\theta_i}\right)^{-\delta_i}-1\right]^{-\frac{1}{\delta_i}}\right\}^{\frac{1}{\theta_i}}$, $\theta_i\geq 1,\ \delta_i>0.$
$C_{i\|0}(u\|uw)$, $i=1,2,3.$	$\left\{w^{\theta_i}+1-(uw)^{\theta_i}-\left[(w^{\theta_i}-(uw)^{\theta_i})^{-\delta_i}+(1-(uw)^{\theta_i})^{-\delta_i}\right]^{-\frac{1}{\delta_i}}\right\}^{-\frac{1}{\theta_i}-1}$ $\cdot\left\{1-\left[1+\left(\frac{u^{-\theta_i}-1}{(uw)^{-\theta_i}-1}\right)^{-\delta_i}\right]^{-\frac{1}{\delta_i}-1}\right\}$	$\left\{\left[\frac{1-(1-uw)^{\theta_i}}{1-(1-u)^{\theta_i}}\right]^{\delta_i}+1-(1-(1-uw)^{\theta_i})^{\delta_i}\right\}^{-\frac{1}{\delta_i}-1}$ $\cdot\left\{(1-uw)^{-\theta_i}-\left[\left(\frac{(1-uw)^{\theta_i}}{1-(1-u)^{\theta_i}}\right)^{\delta_i}+\left(\frac{(1-uw)^{\theta_i}}{1-(1-uw))^{\theta_i}}\right)^{\delta_i}-(1-uw)^{\theta_i\delta_i}\right]^{-\frac{1}{\delta_i}}\right\}^{\frac{1}{\theta_i}-1}$
$\lim_{u\downarrow 0} C_{i\|0}(u\|uw)$, $i=1,2,3.$	$\left[1+w^{\theta_i}-(1+w^{-\theta_i\delta_i})^{-\frac{1}{\delta_i}}\right]^{-\frac{1}{\theta_i}-1}\left[1-(1+w^{-\theta_i\delta_i})^{-\frac{1}{\delta_i}-1}\right]$	$(1+w^{\delta_i})^{-\frac{1}{\delta_i}-1}$
χ_{ij}, $1\leq i<j\leq 3.$	$\int_0^\infty\left[1+w^{\theta_i}-(1+w^{-\theta_i\delta_i})^{-\frac{1}{\delta_i}}\right]^{-\frac{1}{\theta_i}-1}\left[1-(1+w^{-\theta_i\delta_i})^{-\frac{1}{\delta_i}-1}\right]$ $\cdot\left[1+w^{\theta_j}-(1+w^{-\theta_j\delta_j})^{-\frac{1}{\delta_j}}\right]^{-\frac{1}{\theta_j}-1}\left[1-(1+w^{-\theta_j\delta_j})^{-\frac{1}{\delta_j}-1}\right]dw$	$\int_0^\infty\left(1+w^{\delta_i}\right)^{-\frac{1}{\delta_i}-1}\left(1+w^{\delta_j}\right)^{-\frac{1}{\delta_j}-1}dw$
Approx. Vol.	0.17057	0.12882

Table A2. *30 cells of the sub-partition $\mathcal{O}_4$. The reported value in the volume column is the volume of each cell multiplied by 207,360.*

Cell #	Ascending order of the indices						Vol.
1	12	13	23	14	24	34	31
2	13	12	23	14	24	34	24
3	12	23	13	14	24	34	27
4	13	23	12	14	24	34	22
5	23	12	13	14	24	34	17
6	23	13	12	14	24	34	17
7	12	13	14	23	24	34	72
8	13	12	14	23	24	34	54
9	12	23	14	13	24	34	32
10	13	23	14	12	24	34	28
11	23	12	14	13	24	34	18
12	23	13	14	12	24	34	18
13	12	14	13	23	24	34	81
14	13	14	12	23	24	34	81
15	12	14	23	13	24	34	43
16	13	14	23	12	24	34	47
17	23	14	12	13	24	34	18
18	23	14	13	12	24	34	18
19	14	12	13	23	24	34	56
20	14	13	12	23	24	34	76
21	14	12	23	13	24	34	28
22	14	13	23	12	24	34	41
23	14	23	12	13	24	34	24
24	14	23	13	12	24	34	27
25	13	23	14	24	12	34	48
26	23	13	14	24	12	34	36
27	13	14	23	24	12	34	72
28	23	14	13	24	12	34	36
29	14	13	23	24	12	34	63
30	14	23	13	24	12	34	45

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press

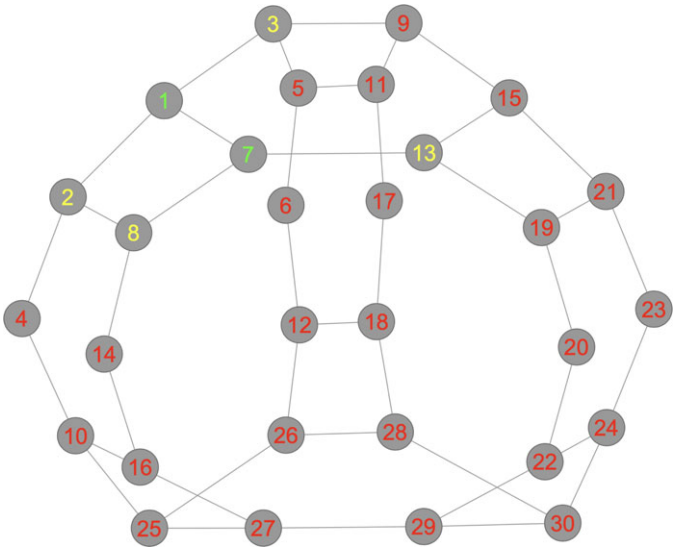


Figure A1. *Adjacency structure for the 30 cells of the sub-partition $\mathcal{O}_4$. Cells connected with edges share a common facet of dimension 5. The feasible points of $\mathcal{O}_4'$ are contained in cells 1 and 7.*

https://doi.org/10.1017/asb.2024.29 Published online by Cambridge University Press