CHAPTER 5

Micro Correlations and Tail Dependence

Roger M. Cooke[*,†,§], Carolyn Kousky[*,¶] and Harry Joe[‡]

[*]*Resources for the Future*
[§]*Cooke@Rff.org*
[†]*Department of Mathematics, Delft University of Technology*
[¶]*Kousky@Rff.org*
[‡]*Department of Statistics, University of British Columbia*
Harry.Joe@ubc.ca

An elementary though seemingly underappreciated finding shows that small global correlations are amplified by aggregation. We observe this behavior in flood damage claims in the US. We also observe that upper tail dependence seems to be amplified by aggregation in these data. We seek to understand this behavior. For sums of exponential variables which are conditionally independent given a gamma-distributed rate, we derive explicit expressions for upper tail dependence and prove that it goes to one as the number of summands goes to infinity, and that the lower tail dependence is zero. We also study sums of events under a latent variable model, where each event occurs if a uniform variable exceeds a threshold, and all uniform variables are conditionally independent given a "latent variable". We obtain a necessary and sufficient condition for strong asymptotic upper tail dependence as the number of summands goes to infinity. Curiously, the normal copula satisfies this condition, although it is not tail dependent via the usual definition. Thus, sums of events under the normal copula latent variable model have upper tail dependence increasing to 1. We also identify tail dependent-like behavior in finite sums of events with the latent variable model.

5.1 Introduction

Micro correlations will amplify the correlation of sums of globally correlated variables, and under certain circumstances they will also amplify tail dependence. This is of evident concern to risk managers, as it will compromise risk management based on diversification. The circumstances under which aggregation amplifies tail dependence are not well understood, and this chapter represents a first foray into the area of tail dependence amplification. We study latent variable models for sums of events, and L_1-symmetric variables. We obtain a condition that leads to upper tail dependence for two different sums of events. In the case of L_1-symmetric measures with gamma scale mixtures, we can prove that aggregation amplifies upper tail dependence.

In Section 5.2, we first discuss the issue of micro correlation and present loss data, draw from Kousky and Cooke,[8] where micro correlations amplify under aggregation. Section 5.3 shows results on tail dependence and aggregation, and Section 5.4 concludes with a discussion of further research.

5.2 Micro Correlations

Let $X_1, \ldots, X_N$ and $X_{N+1}, \ldots, X_{2N}$ be sets of random variables with the average variance σ^2 over the first N and second N random variables and average covariance γ within and between the two sets. The correlation of the sum of the first N and second N X's is:

$$\mathrm{corr}\left(\sum_{i=1}^{N} X_i, \sum_{i=N+1}^{2N} X_i\right) = \frac{N^2\gamma}{N\sigma^2 + N(N-1)\gamma} = \frac{N\gamma}{\sigma^2 + (N-1)\gamma}.$$

Evidently, if $\gamma > 0$ and $\sigma < \infty$, this goes to 1 as $N \to \infty$. Since $\sigma^2 > 0$, $\frac{\sigma^2}{N-1} \geq -\gamma$ which shows that for all N sufficiently large, $\gamma \geq 0$.

We can find micro correlations in many places once we start looking for them. We illustrate with two data sets: flood insurance claims data from the

US National Flood Insurance Program (NFIP) and data on crop insurance indemnities payments from the United States Department of Agriculture's Risk Management Agency. Both data sets are aggregated by county and year for the years 1980 to 2008. The data are in constant year 2000 dollars. Over this time period there has been substantial growth in exposure to flood risk, particularly in coastal counties. To remove the effect of growing exposure, we divide the claims per county per year by personal income per county per year available from the Bureau of Economic Accounts (BEA). Thus we study yearly flood claims per dollar income, per year per county. The crop loss claims are not exposure-adjusted, as an obvious proxy for exposure is not at hand, and exposure growth was less of a concern.

Suppose we randomly draw pairs of counties in the US and compute the correlation of their exposure-adjusted flood losses. Figure 5.1 shows the histogram of 500 such correlations. The average correlation is 0.04. A few counties have quite high correlations but the bulk is around zero. Indeed, based on the sampling distribution for the normal correlation coefficient, correlations less than 0.37 in absolute value would not be statistically distinguishable from zero at the 5% significance level. 91% of these correlations fall into that category.

Instead of looking at the correlations between two randomly chosen counties, consider summing 100 randomly chosen counties and correlating this with the sum of 100 distinct randomly chosen counties (i.e., sampling without

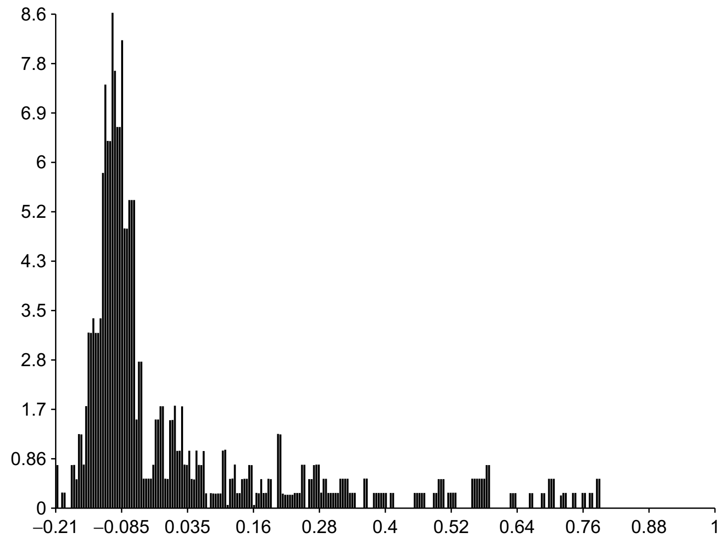

Figure 5.1. Histogram of 500 correlations of randomly paired US exposure-adjusted flood loss per county, 1980–2006. The average correlation is 0.04.

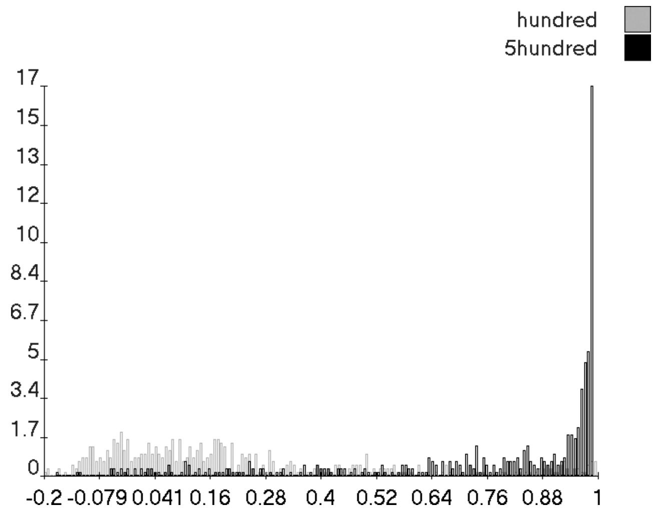

Figure 5.2. Similar to Fig. 5.1, but showing 500 correlations of random sums of 100 and 500.

replacement). If we repeat this 500 times, the centered histogram in Fig. 5.2 results; the average of 500 such correlations-of-100 is 0.23. The histogram at the upper extreme depicts 500 correlations-of-500; their average value is 0.71.

The flood damage per dollar exposure shows a lower correlation than the US crop losses in Fig. 5.3. The mean correlation is 0.13, and the mean of correlations-of-100 is 0.88.

It is interesting to compare the histograms of real loss distributions with a histogram in which each county is assigned an independent uniform variable. The histogram of 500 correlations of random pairs and correlations of random aggregations-of-500 are shown in Fig. 5.4.

5.3 Tail Dependence and Aggregation

In this section, we obtain some results on when aggregation amplifies tail dependence.

The definition of upper tail dependence is given below.

Definition 5.1 (Upper tail dependence). The upper tail dependence between random variables X and Y is

$$UTD(X,Y) = \lim_{q\to 1} \Pr(X > x_q | Y > y_q) \tag{5.1}$$

where $x_q = F_X^{-1}(q)$ and $y_q = F_Y^{-1}(q)$.

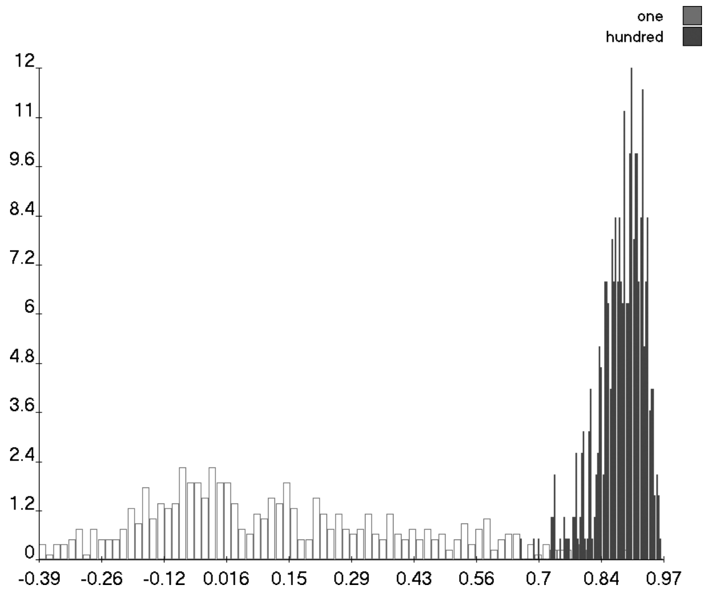

Figure 5.3. Histogram of 500 random correlations of US crop losses per county, 1980–2008, random pairs and random sums of 100.

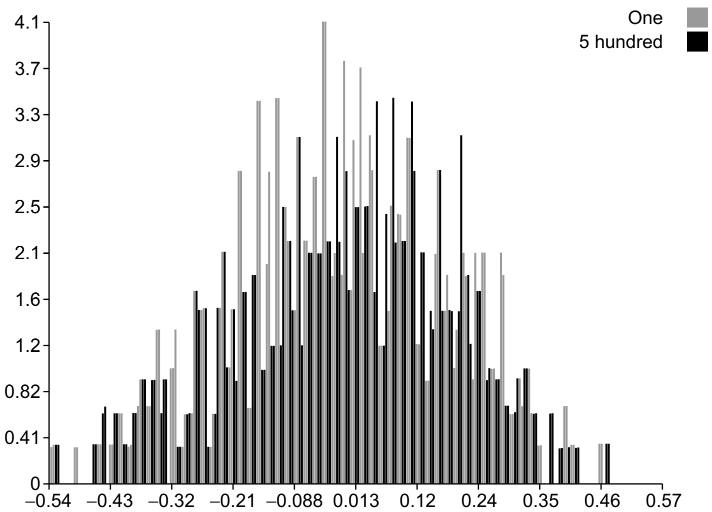

Figure 5.4. Histogram of 500 random correlations of independent uniforms assigned to each county, 1980–2008, random pairs and random sums of 500.

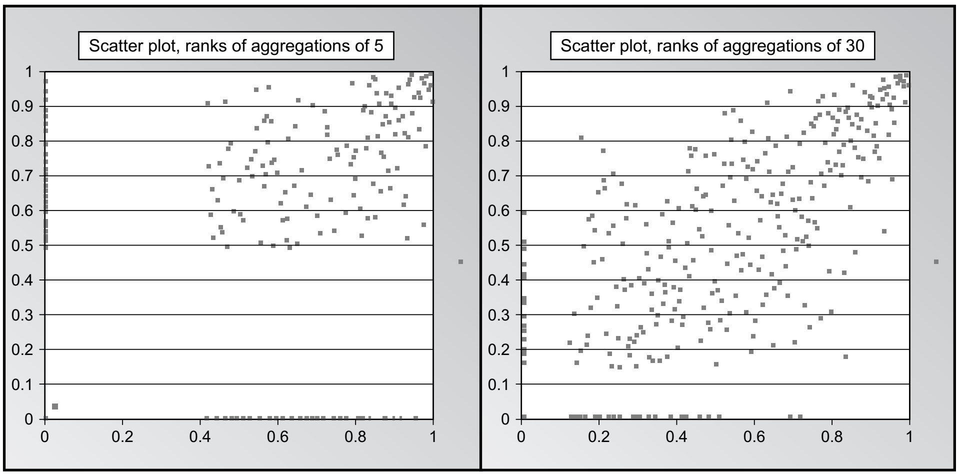

Figure 5.5. Percentile scatterplots of random aggregation of Florida county monthly flood losses. Left: two random aggregations of five counties; right: two random aggregations of 30 distinct counties.

Lower tail dependence is defined in a similar way in the lower quadrant: $LTD(X,Y) = \lim_{q\to 0} \Pr(X \leq x_q | Y \leq y_q)$. As is evident from the definition, tail dependence is a property of the copula. The normal copula has zero tail dependence for all correlation values in $(-1,1)$; see McNeil *et al.*[10]

A central question is whether tail dependence is also amplified by aggregation. In loss distributions we can see the amplification of tail dependence under aggregation. To see tail dependence, the yearly data are not sufficient. Figure 5.5 plots monthly flood loss data in the state of Florida from 1980 to 2008. We choose Florida because there are numerous counties with many non-zero losses in several months. There are two percentile scatterplots: that on the left shows two random aggregation of five counties while the plot on the right shows two random aggregations of 30 counties. Points on the axes correspond to months in which there were no losses in the corresponding aggregate variable. The plot suggests that the upper tail dependence is amplified by aggregation. We seek models to help understand why and when this happens.

5.3.1 *Latent variable models for tail dependence*

In simple latent variable models, a latent variable is an unobserved variable to which all observed variables are correlated, and conditional on which all observed variables are independent. Recognizing this structure as a C-vine with dependence confined to the first tree rooted at the latent variable, it is evident that this is the simplest of a wide class of models.

We first consider a finite class of events, where each event occurs when a physical variable exceeds some limit, and each physical variable is connected to a latent variable. For simplicity, let $U_1, \ldots, U_{2N}$ be uniform variables and suppose event E_i occurs if and only if $U_i > r$. Suppose further that the U_i are conditionally independent given a latent variable V, which is also uniform. To study such models, we require a copula joining U_i and V.

Specifically, $(U_i, V) \sim C(u, v)$ for all i and C is a bivariate copula. Let $C_{1|2}(u|v) = \partial C(u, v)/\partial v$ be the conditional distribution of U_i given $V = v$. We assume C has positive dependence in the sense of stochastic increasing, that is $\Pr(U_1 > u|V = v) = 1 - C_{1|2}(u|v)$ is strictly increasing in $v \in [0, 1]$ for all $0 < u < 1$. This condition is satisfied by all of the commonly used one-parameter families of copula when restricted to the region of positive dependence. For a fixed r in $(0, 1)$, let $Y_i = I(U_i > r)$ for the indicator of the extreme event E_i. Let $S_1(N) = Y_1 + \cdots + Y_N$ and let $S_2(N) = Y_{N+1} + \cdots + Y_{2N}$ be two aggregate numbers of extreme events in two sets. We study the (upper) tail dependence of S_1, S_2 under this simple latent variable model.

Let

$$p_r(v) = 1 - C_{1|2}(r|v), \quad q_r(v) = 1 - p_r(v), \quad 0 \le v \le 1. \tag{5.2}$$

For an integer k between 0 and N inclusive, and $j = 1$ or 2,

$$\begin{aligned} \Pr(S_j = k) &= \int_0^1 \Pr(S_j = k|V = v)\, dv \\ &= \int_0^1 \binom{N}{k} [p_r(v)]^k [q_r(v)]^{N-k}\, dv, \end{aligned}$$

and

$$\begin{aligned} &\Pr(S_1 = k_1, S_2 = k_2) \\ &\quad = \int_0^1 \Pr(S_1 = k_1|V = v) \Pr(S_2 = k_2|V = v)\, dv \\ &\quad = \int_0^1 \binom{N}{k_1} [p_r(v)]^{k_1} [q_r(v)]^{N-k_1} \binom{N}{k_2} [p_r(v)]^{k_2} [q_r(v)]^{N-k_2}\, dv. \end{aligned}$$

For a fraction $0 < \zeta < 1$, let $\lambda_U(r, \zeta, N) = \Pr(S_2 > N\zeta | S_1 > N\zeta)$. Then

$$\lambda_U(r, \zeta, N) = \frac{\int_0^1 \sum_{k_1 \ge N\zeta, k_2 \ge N\zeta} \binom{N}{k_1} [p_r(v)]^{k_1} [q_r(v)]^{N-k_1} \binom{N}{k_2} [p_r(v)]^{k_2} [q_r(v)]^{N-k_2}\, dv}{\int_0^1 \sum_{k \ge N\zeta} \binom{N}{k} [p_r(v)]^k [q_r(v)]^{N-k}\, dv}. \tag{5.3}$$

The analysis of (5.3) for large N is given next. Let Z be a standard normal random variable, with cumulative distribution function Φ. Let

$$g(v) = g(v; r, \zeta) = \frac{p_r(v) - \zeta}{\sqrt{p_r(v)q_r(v)}}.$$

By the normal approximation to binomial, for large N, (5.3) can be approximated by

$$\frac{\int_0^1 \{\Pr(Z > [N\zeta - Np_r(v)]/\sqrt{Np_r(v)q_r(v)}\,)\}^2\,dv}{\int_0^1 \Pr(Z > [N\zeta - Np_r(v)]/\sqrt{Np_r(v)q_r(v)}\,)\,dv} = \frac{\int_0^1 \Phi^2[N^{1/2}g(v)]\,dv}{\int_0^1 \Phi[N^{1/2}g(v)]\,dv}. \tag{5.4}$$

From the positive dependence assumption of stochastic increasing, $p_r(v)$ in (5.2) is increasing in v. Let

$$\begin{aligned} v_0 = v_0(r, \zeta) &= \sup\{v \in (0,1) : p_r(v) \le \zeta\} \\ &= \sup\{v \in (0,1) : g(v; r, \zeta) \le 0\}. \end{aligned}$$

Then (5.4) becomes

$$\frac{\int_0^{v_0} \Phi^2[N^{1/2}g(v)]\,dv + \int_{v_0}^1 \Phi^2[N^{1/2}g(v)]\,dv}{\int_0^{v_0} \Phi[N^{1/2}g(v)]\,dv + \int_{v_0}^1 \Phi[N^{1/2}g(v)]\,dv}. \tag{5.5}$$

If $0 \le v_0 < 1$, then

$$\lim_{N\to\infty} \int_0^{v_0} \Phi^j[N^{1/2}g(v)]\,dv = 0, \quad \text{and}$$

$$\lim_{N\to\infty} \int_{v_0}^1 \Phi^j[N^{1/2}g(v)]\,dv = 1 - v_0, \quad j = 1, 2.$$

Therefore, $\lambda_U(r, \zeta, N)$ in (5.3) goes to 1 as $N \to \infty$ if $0 \le v_0 < 1$, and

$$\lim_{N\to\infty} \lambda_U(r, \zeta, N) = 1 \quad \forall\, 0 < r < 1,\ \ 0 < \zeta < 1$$

if and only if $p_r(1) = \overline{C}_{1|2}(r|1) = 1$ for all $0 < r < 1$ or $C_{1|2}(u|1) = 0$ for all $0 < u < 1$.

If $p_r(1) = \overline{C}_{1|2}(r|1) < 1$, then $\lim_{N\to\infty} \lambda_U(r, \zeta, N) = 1$ only if ζ is small enough so that $0 < v_0 < 1$. If $v_0 = 1$ and $p_r(1) < \zeta$, then (5.5) is bounded above by $\max_{0\le v\le 1} \Phi[N^{1/2}g(v)]$ and this approaches 0 as $N \to \infty$.

For numerical computations, if the limit is 1, $\lambda_U(r, \zeta, N)$ is practical only if $\Pr(S_1 > \zeta N)$ is not too small and $v_0(r, \zeta)$ is not too close to 1; this means

ζ should not be too close to 1. For fixed ζ, $\Pr(S_1 > \zeta N)$ tends to get smaller as the (upper tail) dependence of (U_i, V) gets weaker.

The condition of $\overline{C}_{1|2}(r|1) = \Pr(U > r|V = 1) = 1$ for all $0 < r < 1$ or

$$C_{1|2}(u|1) = \Pr(U \leq u|V = 1) = 0 \quad \forall\, 0 < u < 1 \tag{5.6}$$

is an **upper tail dependence condition**. It is the same as $[U|V = v] \underset{p}{\rightarrow} 1$, as $v \uparrow 1$.

Equation (5.6) holds for all bivariate extreme value copulae, e.g., Gumbel and Galambos. The condition for (5.6) to hold for an Archimedean copula $C_\psi(u, v) = \psi(\psi^{-1}(u) + \psi^{-1}(v))$ is $\psi'(0) = -\infty$ and this is the same condition for the usual tail dependence (Theorem 3.12 in Joe[6]). Hence (5.6) fails to hold for the Frank copula. It also fails to hold for the Plackett copula but holds for the bivariate normal copula with positive correlation ρ. This means that (5.6) is not exactly the same as the usual tail dependence condition of $\lim_{v\uparrow 1} \overline{C}(v, v)/(1 - v)$ being positive because the bivariate normal copula does satisfy this. Some proofs of these cases are given in Appendix A.

Table 5.1 compares the conditional probability $\lambda_U(r, \zeta, N)$ for the Gumbel, bivariate normal and Frank copulae when $r = 0.9$, $\zeta = 0.7$, and the dependence parameters for the three copulae are chosen to get a rank correlation of 0.5.

The definition of tail dependence as limiting conditional probabilities of exceedence is not appropriate for finite sums of events. Nonetheless we can identify tail dependence-like behavior in finite sums of events. With Frank's copula, take the probability of the individual events as 0.1 and the correlation to the latent variable V as 0.9 (the parameter $\theta = 12.3$) which induces a correlation 0.36 between any two events. Figure 5.6 illustrates curious non-monotonic behavior in $P\{S_1 > i|S_2 > i\}$, for $N = 100$ and $i = 1, \ldots, 100$. This is caused by the interaction of two opposing "forces"; as i increases, $P\{S_1 > i\}$ goes down, while on the other hand, conditionalizing on $P\{S_2 > i\}$ drives the latent V up, which increases $P\{S_1 > i|S_2 > i\}$. The pattern with N fixed and ζ increasing is quite different from the pattern when ζ is fixed and N increasing.

5.3.2 *Sum of damages over extreme events*

Instead of the number of extreme events, consider the sum of losses or damages. The situation becomes more complex and the results depend strongly on the copula and the damage distributions. Figure 5.7 shows percentile scatterplots of events multiplied by independent damages, and where the

Table 5.1. Conditional probabilities $\Pr(S_2 > N\zeta | S_1 > N\zeta) = \lambda_U(r, \zeta, N)$ with $r = 0.9$, $\zeta = 0.7$, Spearman ρ_S = rank correlation = 0.5; leading to parameters $\theta = 1.54$ for the Gumbel, $\rho = 0.518$ for the bivariate normal (BVN), $\theta = 7.90$ for the Frank copulae respectively. Limit behavior depends on the comparison sign of $p_r(1) - \zeta$.

	$\lambda_U(r, \zeta, N)$		
N	Gumbel	BVN	Frank
10	0.604	0.264	0.144
20	0.687	0.411	0.067
30	0.733	0.528	0.034
40	0.763	0.620	0.019
50	0.785	0.695	0.010
60	0.801	0.755	0.006
70	0.815	0.804	0.003
80	0.826	0.845	0.002
90	0.835	0.877	0.001
100	0.843	0.903	0.001
	$p_r(v) = \overline{C}_{1\|2}(r\|v)$		
v	Gumbel	BVN	Frank
1	1.0	1.0	0.546
0.99999	0.994	0.861	0.546

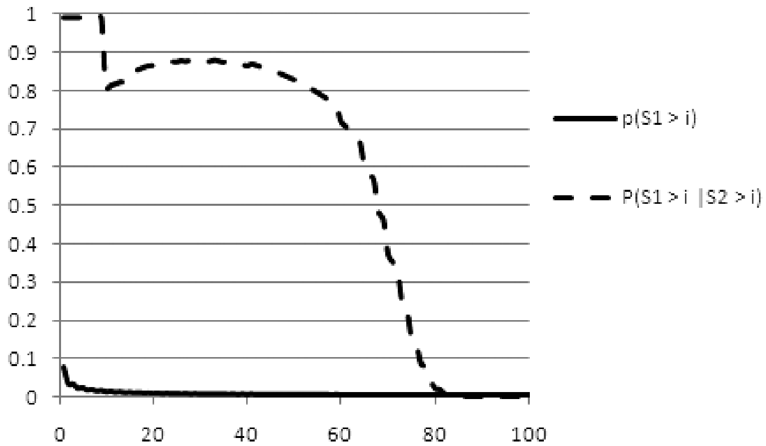

Figure 5.6. Tail dependent-like behavior of sums of events, probability of exceedence as function of i, for Frank's copula, $\theta = 12.3$, $N = 100$.

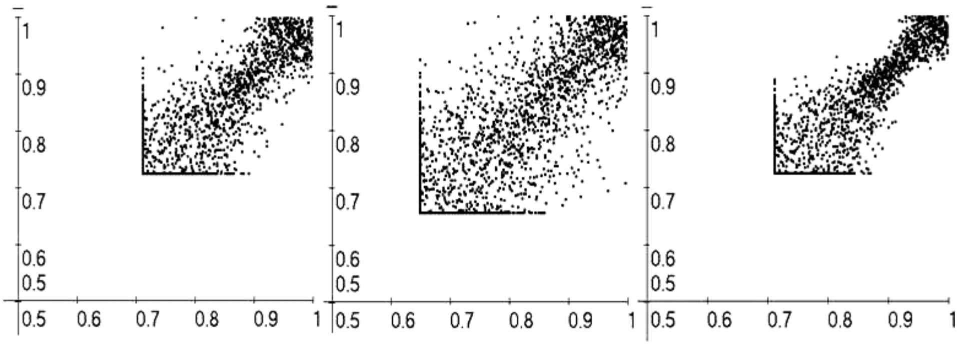

Figure 5.7. Tail dependent-like behavior of sums of events times damages. Left: Pareto 2 damages with Gumbel copula; middle: Pareto 2 damages with bivariate normal copula; right: exponential damages with Gumbel copula.

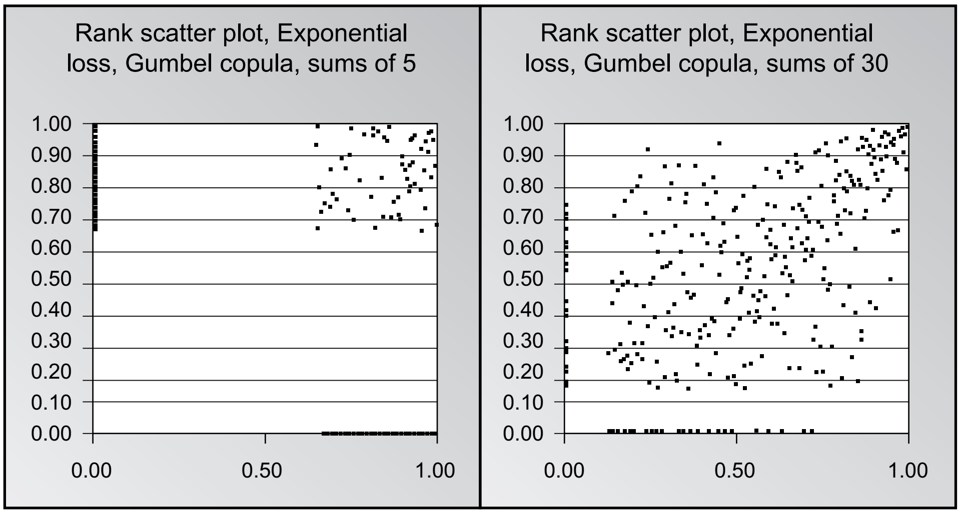

Figure 5.8. A model for Florida monthly flood damages, exponential damages linked to a latent variable with the Gumbel copula.

joining copulae are Gumbel and bivariate normal. Figure 5.8 shows exponentially distributed damages linked to a latent variable via the Gumbel copula, and parameters are chosen to resemble Fig. 5.5. This suggests that simple latent variable models may describe such loss phenomena satisfactorially.

Without considering sums of events, it is easy to construct simulations in which this amplification occurs. Figure 5.9 shows percentile plots of two normal variables X_1 and X_2 which each have rank correlation 0.1 to a latent variable V, and are conditionally independent given the latent variable. The

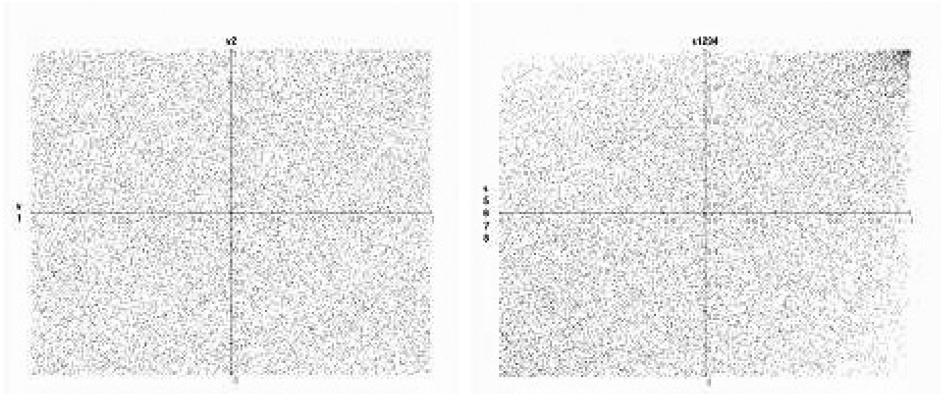

Figure 5.9. Percentile plots with tail dependence. Left: two normal variables rank-correlated 0.1 to a latent variable with Gumbel copula; right: distinct sums of 40 such variables, each similarly rank-correlated to the latent variable.

rank correlation is realized with the Gumbel copula, which has very weak tail dependence at that correlation value. This induces a very weak tail dependence between X_i and V. If we form sums of 40 such normal variables and consider the tail dependence of two such sums, we see in the right-hand plot of Fig. 5.9 that the tail dependence has become more pronounced.

Although tail dependence is a property of the copula, whether and to what degree tail dependence is amplified by aggregation depends on the marginal distributions. Figure 5.10 is similar to Fig. 5.9 except that the

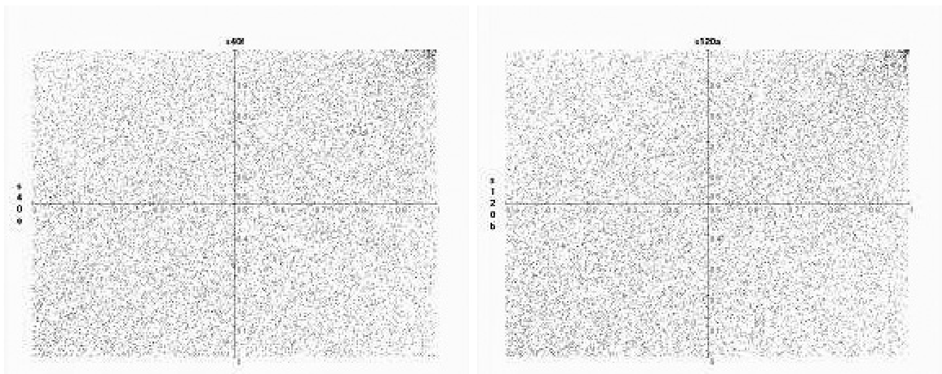

Figure 5.10. Percentile plots with tail dependence. Left: sums of 40 Pareto variables with survival function $(1/(1+x))$, each rank-correlated 0.1 to a latent variable with Gumbel copula; right: sums of 120 such variables.

variables are Pareto with survival function $S(x) = (1+x)^{-1}$. The amplification of tail dependence for 120 Pareto variables is much weaker than that for 40 normal variables. This Pareto distribution does not have a finite first moment (or, of course, correlation).

In certain cases we can prove some results for tail dependence. The following proposition, whose proof is in Appendix B, gives a lower bound for tail dependence of variables, which are tail dependent on a latent variable:

Proposition 5.1. *Suppose (U_1, V) and (U_2, V) are pairwise upper tail dependent with, respectively, coefficients $\lambda_1 > 0$ and $\lambda_2 > 0$, and (U_1, U_2) is conditionally independent given V. Also suppose that U_1 and U_2 are each stochastically increasing in V. Then (U_1, U_2) has an upper tail dependence coefficient that exceeds $\lambda_1\lambda_2$.*

5.3.3 L_1-symmetric measures

Results relating tail dependence to aggregation are difficult to obtain, since aggregation is not simply a question of the copula, but also of the marginal distributions. One case where analytic results are possible concerns the L_p-symmetric variables with $1/p \in \mathbb{N}$.

Recall the Gamma integral:

$$\int_0^\infty y^{\eta-1}e^{-\beta y}\,dy = \frac{\Gamma(\eta)}{\beta^\eta}; \quad \beta > 0, \ \ \eta > 0.$$

The Gamma(η, β) density with shape η and rate β is $f(y;\eta,\beta) = \beta^\eta y^{\eta-1}e^{-\beta y}/\Gamma(\eta)$, with mean η/β and variance η/β^2.

An atomless L_p-symmetric measure on $\mathbb{R}^n$ is one whose density at $(x_1,\ldots,x_N)$ depends only on the L_p norm $(\sum |x_i|^p]^{1/p}$. Berman[2] proved that L_p-symmetric measures on $\mathbb{R}$ can be uniquely represented as conditionally independent gamma transforms with shape $1/p$. For L_1 measures, we have conditionally independent exponentials given the failure rate. $(X_1,\ldots,X_N)$ have an L_1-symmetric distribution with Gamma(η,β) mixing distribution if, for any N, the N-dimensional marginal density is given by

$$f_N(x_1,\ldots,x_N) = \int \left\{\prod_{i=1}^N \lambda e^{-\lambda x_i}\right\} \beta^\eta\lambda^{\eta-1}e^{-\beta\lambda}d\lambda/\Gamma(\eta). \tag{5.7}$$

Setting $N = 1$ and integrating over λ, one finds the univariate density and survivor functions:

$$f_1(x) = \frac{\eta\beta^\eta}{(\beta+x)^{\eta+1}}; \quad 1 - F_1(x) = \left(\frac{\beta}{\beta+x}\right)^\eta, \tag{5.8}$$

which is the Pareto thick-tailed (leptokurtic) distribution with shape parameter η and scale parameter β. These multivariate distributions were first studied by Takahasi[11] and Harris.[5] Then unconditionally the joint survival function of $X_1, \ldots, X_N$ is

$$\begin{aligned}\Pr(X_1 > x_1, \ldots, X_N > x_N) &= \int_0^\infty \prod_{i=1}^N e^{-\lambda x_i} \frac{\lambda^{\eta-1}\beta^\eta}{\Gamma(\eta)} e^{-\beta\lambda} d\lambda \\ &= \frac{\beta^\eta}{[\beta + x_1 + \cdots + x_N]^\eta}. \end{aligned} \tag{5.9}$$

This is a special case of the multivariate Burr distribution of Takahasi,[11] with type II Pareto as a special case of Burr for the univariate margins. The multivariate Pareto distribution of Mardia[9] has type I Pareto margins rather than type II Pareto. From this distribution, Cook and Johnson[4] obtained the copula (replacing Pareto survival functions) as

$$C(u_1, \ldots, u_N; \eta) = [u_1^{-1/\eta} + \cdots + u_N^{-1/\eta} - (N-1)]^{-\eta}. \tag{5.10}$$

As an aside, Kimeldorf and Sampson[7] did the same thing but only for the bivariate case; Clayton[3] has the bivariate distribution as a gamma frailty model and through a derivation from a differential equation, but does not have the multivariate case. In this parametrization, dependence increases as η decreases. The copula (5.10) has lower tail dependence and the distribution (5.9) has upper tail dependence.

Consider the sum $S = X_1 + \cdots + X_N$, where $(X_1, \ldots, X_N)$ has density (5.7). Since $S|\Lambda = \lambda \sim \mathrm{Gamma}(n, \lambda)$,

$$\begin{aligned} f_S(r; N) &= \int_0^\infty \frac{1}{\Gamma(N)} \lambda^N r^{N-1} e^{-\lambda r} \cdot \frac{1}{\Gamma(\eta)} \lambda^{\eta-1} \beta^\eta e^{-\beta\lambda} d\lambda \\ &= \frac{r^{N-1}\Gamma(N+\eta)\beta^\eta}{\Gamma(N)\Gamma(\eta)(\beta+r)^{\eta+N}}. \end{aligned}$$

The sums have the same tail behavior as the one-dimensional margins. From (5.8), we obtain the mean of X. The variance, covariance and product moment correlation may be obtained from (5.7) with $N = 2$, giving:

$$\begin{aligned} \mu(X) &= \frac{\beta}{\eta - 1}; \quad \eta > 1 \\ \mathrm{Var}(X_1) &= \frac{\beta^2 \eta}{(\eta-1)^2(\eta-2)}; \quad \eta > 2 \end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(X_1, X_2) &= \frac{\beta^2}{(\eta-1)^2(\eta-2)}; \quad \eta > 2\\
\mathrm{corr}(X_1, X_2) &= \eta^{-1}; \quad \eta > 2\\
\mathrm{Var}(X_1 + \cdots + X_N) &= \mathrm{Var}(X_1)\{N + N(N-1)\eta^{-1}\}; \quad \eta > 2.
\end{aligned}$$

Note that the mean exists only if $\eta > 1$, and the variance, covariance and correlation require $\eta > 2$.

5.3.4 *Tail dependence for sums of L_1 measures*

Computations of tail dependence for sums of L_1 measures are tractable, and the same holds for L_p measures with $1/p \in \mathbb{N}$. If X is independent of Y then $UTD(X, Y) = 0$ but not conversely. Tail dependence is invariant under a monotone transformation of X and Y, hence it is a property of the copula joining X and Y.

Let $(X_1, \ldots, X_{2N})$ have density (5.7) with $2N$ replacing N.

The incomplete Gamma integral with positive integer parameter m is:

$$\frac{1}{\Gamma(m)} \int_y^\infty \lambda^m z^{m-1} e^{-\lambda z} = \sum_{k=0}^{m-1} \frac{(\lambda y)^i}{i!} e^{-\lambda y}, \quad y > 0.$$

Then

$$\begin{aligned}
\Pr\left(\sum_{i=1}^N X_i > r\right) &= \int_0^\infty \Pr\left(\sum_{i=1}^N X_i > r \,|\, \Lambda = \lambda\right) \frac{\lambda^{\eta-1}\beta^\eta}{\Gamma(\eta)} e^{-\beta\lambda} d\lambda\\
&= \int_0^\infty \sum_{k=0}^{N-1} \frac{(\lambda r)^k}{k!} e^{-\lambda r} \cdot \frac{\lambda^{\eta-1}\beta^\eta}{\Gamma(\eta)} e^{-\beta\lambda} d\lambda\\
&= \left(\frac{\beta}{\beta + r}\right)^\eta \left[\sum_{k=0}^{N-1} \frac{\Gamma(\eta+k)}{k!\,\Gamma(\eta)} \frac{r^k}{(\beta+r)^k}\right]. \qquad (5.11)
\end{aligned}$$

As $r \to \infty$, the bracketed term goes to

$$\left[\sum_{k=0}^{N-1} \frac{\Gamma(\eta+k)}{k!\,\Gamma(\eta)}\right].$$

Similarly,

$$\begin{aligned}\Pr&\left(\sum_{i=1}^{N} X_i > r \bigcap \sum_{i=N+1}^{2N} X_i > r\right)\\ &= \int_0^\infty e^{-2\lambda r}\left[\sum_{k=0}^{N-1}\frac{(\lambda r)^k}{k!}\right]^2 \cdot \frac{\lambda^{\eta-1}\beta^\eta}{\Gamma(\eta)} e^{-\beta\lambda} d\lambda\\ &= \left(\frac{\beta}{\beta+2r}\right)^\eta \sum_{k,j=0}^{N-1}\frac{\Gamma(k+j+\eta)}{k!j!\,\Gamma(\eta)}\left(\frac{r}{\beta+2r}\right)^{k+j}. \qquad (5.12)\end{aligned}$$

The tail dependence of sums of N L_1 variables is therefore the limiting ratio as $r \to \infty$ of (5.12) over (5.11):

$$\left(\frac{1}{2}\right)^\eta \frac{\sum_{k,j=0}^{N-1} 2^{-k-j}\frac{\Gamma(k+j+\eta)}{k!j!\,\Gamma(\eta)}}{\sum_{k=0}^{N-1}\frac{\Gamma(\eta+k)}{k!\,\Gamma(\eta)}}. \qquad (5.13)$$

Table 5.2 gives some values, comparing the number N of disjunct variables summed. We see that the tail dependence grows in N and decreases in the shape factor η. Also (5.13) converges to 1 as $N \to \infty$ for any $\eta > 0$ — a proof is given in Appendix C; the rate of convergence to 1 is slower for larger η.

Figure 5.11 shows rank scatterplots for sums of L_1 measures with shape $\eta = 3$. The first shows two variables, the second shows two sums of 10 variables, and the third shows two sums of 50 variables.

Table 5.2. Upper tail dependence for sums of N L_1 variables, the shape of the Gamma mixing distribution ranges from 1 to 5, 10, 15 and 20.

Shape	corr(X_1, X_2)	$N=1$	$N=3$	$N=5$	$N=10$	$N=50$
1		0.500	0.688	0.754	0.824	0.920
2		0.250	0.453	0.549	0.664	0.842
3	0.3333	0.125	0.289	0.388	0.523	0.767
4	0.2500	0.062	0.180	0.267	0.405	0.694
5	0.2000	0.031	0.109	0.180	0.307	0.624
10	0.1000	0.001	0.007	0.019	0.061	0.338
15	0.0667	3×10^{-5}	4×10^{-4}	0.002	0.009	0.160
20	0.0500	1×10^{-6}	2×10^{-5}	1×10^{-4}	0.001	0.066

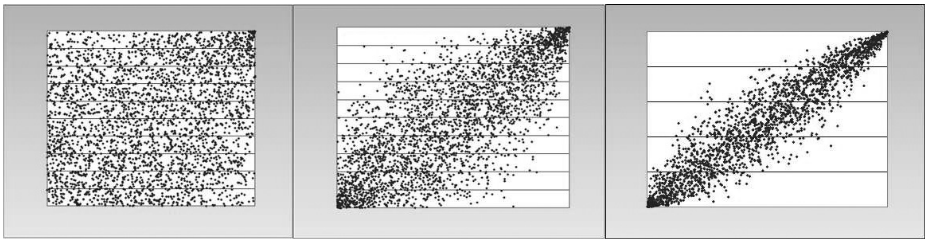

Figure 5.11. Percentile scatterplots for sums of L_1 variables, with shape of Gamma mixing distribution = 3. Left: 2 L_1 variables, rank correlation = 0.21; center: sums of 10 such variables, rank correlation = 0.77; right: sums of 50 such variables, rank correlation = 0.94.

5.3.5 *Lower tail dependence*

The multivariate Pareto model (5.9) does not have lower tail dependence, so it is not surprising that the aggregate losses S_1, S_2 do not have lower tail dependence. A derivation is given below, making use of the identity for the incomplete Gamma function with an integer shape parameter.

For $i = 1, 2$, let $S_i = S_i(N)$ denote the ith sum of N L_1 variables, as above. The marginal probability is

$$\Pr(S_i \le r) = \frac{\beta^\eta}{(r+\beta)^\eta} \sum_{k=N}^{\infty} \frac{r^k}{(r+\beta)^k} \frac{\Gamma(\eta+k)}{k!\,\Gamma(\eta)}$$

and

$$\begin{aligned}
&\Pr(S_1 \le r, S_2 \le r) \\
&\quad = \int_0^\infty \sum_{k=N}^{\infty} \frac{(\lambda r)^k}{k!} e^{-\lambda r} \cdot \sum_{j=N}^{\infty} \frac{(\lambda r)^j}{j!} e^{-\lambda r} \cdot \frac{\lambda^{\eta-1}\beta^\eta}{\Gamma(\eta)} e^{-\beta\lambda} d\lambda \\
&\quad = \frac{\beta^\eta}{(2r+\beta)^\eta} \sum_{k=N}^{\infty} \sum_{j=N}^{\infty} \frac{r^{k+j}}{(2r+\beta)^{k+j}} \frac{\Gamma(\eta+k+j)}{k!j!\,\Gamma(\eta)}.
\end{aligned}$$

Putting $z = r/\beta$ this becomes:

$$z^{2N} \frac{\Gamma(\eta+2N)}{N!N!\,\Gamma(\eta)} + O(z^{2N+1}), \quad r = \beta z \to 0,$$

and

$$\begin{aligned}
\Pr(S_1 \le r) &= \sum_{k=N}^{\infty} z^k (1+z)^{-\eta-k} \frac{\Gamma(\eta+k)}{k!\,\Gamma(\eta)} \\
&= z^N \frac{\Gamma(\eta+N)}{N!\,\Gamma(\eta)} + O(z^{N+1}), \quad y = \beta z \to 0.
\end{aligned}$$

The limit of the lower tail (for a fixed N) is:

$$\lambda_L = \lim_{r\to 0} \frac{\Pr(S_1 \le r, S_2 \le r)}{\Pr(S_1 \le r)} = \lim_{z\to 0} \frac{z^{2N} \frac{\Gamma(\eta+2N)}{N!N!\Gamma(\eta)} + O(z^{2N+1})}{z^N \frac{\Gamma(\eta+N)}{N!\Gamma(\eta)} + O(z^{N+1})}$$

$$= \lim_{z\to 0} \frac{z^N}{N!} \prod_{k=N}^{2N-1} (\eta + k) = 0.$$

5.4 Discussion

In this chapter, we have shown how some simple latent variable models lead to interesting results on tail dependence of aggregate losses. Further research consists of studying tail dependence on sums under more general dependence models, such as via vines. For example, $\sum_{i=1}^{N} X_1, \ldots, X_{2N}$ and $\sum_{i=N+1}^{2N} X_1, \ldots, X_{2N}$ are conjectured to have upper tail dependence of 1 as $N \to \infty$, if X_i have Pareto-like upper tails and their joint distribution has upper tail dependence. In analyzing data, tail dependence-like behavior is also of interest, as this behavior may obtain for more general classes of copulae.

Appendices

A. *Proofs for the tail dependence condition involving* $C_{1|2}(u|1)$

The conditional distributions $C_{1|2}$ for the common one-parameter copula families are given on pp. 146–147 of Joe.[6]

- For the Frank copula with parameter $\theta > 0$,
  $$\overline{C}_{1|2}(u|v) = [1 + e^{-\theta v} a(u)]^{-1}, \quad a(u) = (1 - e^{-\theta u})/(e^{-\theta u} - e^{-\theta}),$$
  so that $\overline{C}_{1|2}(u|1) = [1+e^{-\theta} a(u)]^{-1} = (e^{\theta} - e^{\theta u})/(e^{\theta} - 1) < 1$ for $0 < u < 1$.
- For the Plackett copula with parameter $\theta > 0$, $\overline{C}_{1|2}(u|1) = \theta(1-u)/[\theta(1-u)+u] < 1$ for $0 < u < 1$.
- For the bivariate normal copula with parameter $\rho > 0$, $\overline{C}_{1|2}(u|v) = 1 - \Phi([\Phi^{-1}(u) - \rho\Phi^{-1}(v)]/\sqrt{1-\rho^2}\,) \to 1 - \Phi(-\infty) = 1$ as $v \to 1$.
- For the Archimedean copula: with $C_\psi(u, v) = \psi(\psi^{-1}(u) + \psi^{-1}(v))$, where ψ is a Laplace transform,
  $$C_{1|2}(u|v) = \frac{\psi'(\psi^{-1}(u) + \psi^{-1}(v))}{\psi'(\psi^{-1}(v))}$$

so that

$$\lim_{v\to 1} C_{1|2}(u|v) = \lim_{s\to 0} \frac{\psi'(\psi^{-1}(u)+s)}{\psi'(s)} = \lim_{s\to 0} \frac{\psi'(\psi^{-1}(u))}{\psi'(s)}.$$

This is 0 if $\psi'(0) = -\infty$ and is in $(0,1)$ if $-\psi'(0) < \infty$.

- For the Extreme-value copula: Let $C(u,v) = e^{-A(-\log u, -\log v)}$, where $\max\{w_1, w_2\} \le A(w_1, w_2) \le w_1 + w_2$ and A is homogeneous of order 1. Let $A_2 = \partial A/\partial w_2$ which is homogeneous of order 0. Then $C_{1|2}(u|v) = C(u,v)A_2(-\log u, -\log v)\cdot v^{-1}$ so that $C_{1|2}(u|1) = uA_2(-\log u, 0) = 0$, assuming $A(w_1,w_2) \not\equiv w_1 + w_2$ and

$$A_2(w_1, 0) = \lim_{w_2\to 0} \frac{\partial A(w_1,w_2)}{\partial w_2} = \frac{\partial \lim_{w_2\to 0} A(w_1,w_2)}{\partial w_2} = \frac{\partial w_1}{\partial w_2} = 0.$$

It is easily shown directly that $A_2(w,0) = 0$ for the Gumbel and Galambos copulae with positive dependence. For the Gumbel copula, $A(w_1,w_2) = (w_1^\theta + w_2^\theta)^{1^\theta}$ (for $\theta > 1$), and for the Galambos copula, $A(w_1, w_2) = w_1 + w_2 - (w_1^{-\theta} + w_2^{-\theta})^{-1^\theta}$ (for $\theta > 0$)

B. *Proof of Proposition 5.1 and an example*

Proof. Since tail dependence is invariant under monotone increasing transforms, without loss of generality, we assume that U_1, U_2, V are uniform $(0,1)$ random variables. We need to show that $\lim_{u\uparrow 1} \Pr(U_2 > u|U_1 > u) \ge \lambda_1\lambda_2$.

Let $C_{U_1U_2V}(u_1,u_2,v)$ be the copula and joint distribution of U_1, U_2, V with margins $C_{U_1V}(u_1,v)$, $C_{U_2V}(u_2,v)$. Let $C_{12|V}, C_{1|V}, C_{2|V}$ be the partial derivatives with respect to v, and let $\overline{C}_{12|V}, \overline{C}_{1|V}, \overline{C}_{2|V}$ be the corresponding survival functions. Note that for $0 < u < 1$,

$$\begin{aligned}
\Pr(U_2 > u|U_1 > u) &\ge \Pr(U_2 > u, V > u|U_1 > u) \\
&= \frac{\Pr(U_2 > u, V > u, U_1 > u)}{1-u} = (1-u)^{-1}\int_u^1 \overline{C}_{12|V}(u,u|v)\,dv \\
&= (1-u)^{-1}\int_u^1 \overline{C}_{1|V}(u|v)\,\overline{C}_{2|V}(u|v)\,dv, \qquad (5.14)
\end{aligned}$$

where the last equality comes from conditional independence. The right-hand side of (5.14) is the same as

$$\mathbb{E}\left[\overline{C}_{1|V}(u|Z)\,\overline{C}_{2|V}(u|Z)\right], \qquad (5.15)$$

where Z is uniform on $[u, 1]$. With the stochastically increasing assumption, $\overline{C}_{1|V}(u|v)$ and $\overline{C}_{2|V}(u|v)$ are increasing in $v \in [u, 1)$. By positive dependence from Fréchet upper bound or co-monotonicity, the covariance of two increasing functions of a random variable is non-negative (if it exists), and hence (5.15) exceeds

$$\begin{aligned}
&\mathbb{E}[\overline{C}_{1|V}(u|Z)] \cdot \mathbb{E}[\overline{C}_{2|V}(u|Z)] \\
&\quad = (1-u)^{-1} \int_u^1 \overline{C}_{1|V}(u|v)\, dv \cdot (1-u)^{-1} \int_u^1 \overline{C}_{2|V}(u|v)\, dv \\
&\quad = \Pr(U_1 > u|V > u) \cdot \Pr(U_2 > u|V > u). \qquad (5.16)
\end{aligned}$$

Take the limit of (5.14) and (5.16) to get:

$$\begin{aligned}
&\lim_{u \uparrow 1} \Pr(U_2 > u|U_1 > u) \\
&\quad \geq \lim_{u \uparrow 1} \Pr(U_1 > u|V > u) \cdot \lim_{u \uparrow 1} \Pr(U_2 > u|V > u) = \lambda_1 \lambda_2 > 0. \qquad \square
\end{aligned}$$

Remark 5.1. Note that the stochastic increasing condition can be weakened to "$\Pr(U_i > u|V = v)$ is increasing in $v \in [u, 1)$ for all u near 1". Hence it is a weak condition that would be expected to hold if there is tail dependence. The stochastic increasing condition, as given in Proposition 5.1, usually holds in models with conditional independence given a latent variable, as shown in the example below.

Example 5.1. For the multivariate Pareto distribution (5.9) that derives from a Gamma mixture of exponentials, let (X_1, X_2) be such that $X_i|\Lambda = a$ are conditional exponential with mean a^{-1}, and $\Lambda \sim \text{Gamma}(\eta, \beta)$. Then with $U_1 = X_1$, $U_2 = X_2$, $V = \Lambda^{-1}$, U_1, U_2 are each stochastically increasing in V. From the copula (5.10), the bivariate upper tail dependence parameter of (X_1, X_2) is $2^{-\eta}$. We next obtain the common tail dependence parameter λ_1 for (X_i, V) for $i = 1, 2$ and show the inequality from the proposition. Because of scale invariance, we assume $\beta = 1$ for the following calculations. Let $G(z; \eta) = [\Gamma(\eta)]^{-1} \int_0^z y^{\eta-1} e^{-y}\, dy$ be the cumulative distribution function of the Gamma$(\eta, 1)$ random variable Λ. Then

$$\Pr(X_1 > x|\Lambda^{-1} > v) = \Pr(X_1 > x, \Lambda < v^{-1}) / \Pr(\Lambda^{-1} > v), \qquad (5.17)$$

$$\begin{aligned}
&\Pr(X_1 > x, \Lambda < v^{-1}) \\
&\quad = \Gamma^{-1}(\eta) \int_0^{v^{-1}} e^{-ax} a^{\eta-1} e^{-a}\, da = (1+x)^{-\eta} G(v^{-1}(1+x); \eta). \qquad (5.18)
\end{aligned}$$

X_1 has cumulative distribution function $F(x) = 1-(1+x)^{-\eta}$ $(x > 0)$ and inverse cumulative distribution function $F^{-1}(p) = (1-p)^{-1/\eta} - 1$ $(0 < p < 1)$. For z near 0, $G(z;\eta) \approx z^\eta/\Gamma(\eta+1)$. For $0 < u < 1$ that is close to 1, let $x = F^{-1}(u) = (1-u)^{-1/\eta} - 1$ and $v(u)$ be the u quantile of Λ^{-1}, so that $[v(u)]^{-1}$ is the lower $1-u$ quantile of Λ or $[v(u)]^{-1} \approx [(1-u)\Gamma(\eta+1)]^{1/\eta}$. Substitute into (5.17) and (5.18) to get:

$$\begin{aligned}
&\lim_{u\uparrow 1} \Pr(X_1 > F^{-1}(u) | \Lambda^{-1} > v(u)) \\
&\quad = \lim_{u\uparrow 1} \frac{(1-u)\, G(\Gamma^{1/\eta}(\eta+1)(1-u)^{1/\eta}(1-u)^{-1/\eta};\eta)}{1-u} = G(\Gamma^{1/\eta}(\eta+1);\eta).
\end{aligned}$$

To match Proposition 5.1, $\lambda_1 = \lambda_2 = G(\Gamma^{1/\eta}(\eta+1);\eta)$ and it can be shown numerically that

$$\lim_{u\uparrow 1} \Pr(X_2 > F^{-1}(u) | X_1 > F^{-1}(u)) = 2^{-\eta} \geq [G(\Gamma^{1/\eta}(\eta+1);\eta)]^2.$$

C. *Proof that* $\boldsymbol{\lambda_{U,\eta,N} \to 1}$ *as* $\boldsymbol{N \to \infty}$

Rewrite (5.13) as:

$$\lambda_{U,\eta,N} = 2^{-\eta} \frac{\sum_{k=0}^{N-1} \sum_{j=0}^{N-1} \frac{\Gamma(\eta+k+j)}{\Gamma(\eta)\, 2^{k+j} k! j!}}{\sum_{k=0}^{N-1} \frac{\Gamma(\eta+k)}{\Gamma(\eta)\, k!}}. \tag{5.19}$$

The numerator on the right-hand side of (5.19) can be written as

$$\sum_{\ell=0}^{2N-2} \frac{\Gamma(\eta+\ell)}{\Gamma(\eta)\, \ell!} A_{\ell,N}, \tag{5.20}$$

where

$$A_{\ell,N} = \sum_{0\leq k,j\leq N-1: k+j=\ell} \frac{\ell!}{2^{k+j} k! j!}.$$

For $0 \leq \ell \leq N-1$, then $A_{\ell,N} = 1$ from a binomial sum, and for $N \leq \ell \leq 2N-2$,

$$A_{\ell,N} = \sum_{i=\ell-N+1}^{N-1} \binom{\ell}{i} 2^{-\ell}.$$

It is shown in Lemma 5.2 below that $A_{\ell,N} \to 1$ as $N \to \infty$ for (approximately) fixed ℓ/N.

Next, (5.20) can be written as (with $k = \ell - N$ in second summation):

$$\sum_{k=0}^{N-1} \frac{\Gamma(\eta+k)}{\Gamma(\eta)\,k!} + \sum_{k=0}^{N-2} \frac{\Gamma(\eta+k+N)}{\Gamma(\eta)\,(k+N)!} A_{k+N,N} = D + \sum_{k=0}^{N-2} \frac{\Gamma(\eta+k+N)}{\Gamma(\eta)\,(k+N)!} A_{k+N,N}$$

where D is the denominator in (5.19). The proof is complete by showing that as $N \to \infty$,

$$D^{-1} \sum_{k=0}^{N-2} \frac{\Gamma(\eta+k+N)}{\Gamma(\eta)\,(k+N)!} A_{k+N,N} \to 2^\eta - 1$$

because then (5.19) goes to $2^{-\eta}[1 + (2^\eta - 1)] = 1$. This follows from the two lemmas below, together with the Lebesgue Dominated Convergence Theorem.

Lemma 5.1. *Let*

$$d_{\eta,k} = \frac{\Gamma(\eta+k)}{\Gamma(\eta)\,k!}, \quad k = 1, 2, \ldots.$$

As $N \to \infty$,

$$\frac{\sum_{k=0}^{N-2} \frac{\Gamma(\eta+k+N)}{\Gamma(\eta)\,(k+N)!}}{\sum_{k=0}^{N-1} \frac{\Gamma(\eta+k)}{\Gamma(\eta)\,k!}} = \frac{\sum_{k=0}^{N-2} d_{\eta,k+N}}{\sum_{k=0}^{N-1} d_{\eta,k}} \to 2^\eta - 1.$$

Proof. This is split into cases.

- $\eta = 1$: $d_{\eta,i} = 1$ for all i so the ratio is $1 = 2^1 - 1$.
- $\eta = 2$: $d_{\eta,k} = (k+1)$, $d_{\eta,k+N} = (k+N+1)$. Hence

$$\frac{\sum_{k=0}^{N-2} d_{\eta,k+N}}{\sum_{k=0}^{N-1} d_{\eta,k}} = \frac{\sum_{k=0}^{N-2}(k+N+1)}{\sum_{k=0}^{N-1}(k+1)} = \frac{3N(N-1)/2}{N(N+1)/2} \to 3 = 2^2 - 1.$$

- $\eta = 3$: $d_{\eta,k} = (k+2)(k+1)/2!$, $d_{\eta,k+N} = (k+N+2)(k+N+1)/2!$. Hence for large N,

$$\frac{\sum_{k=0}^{N-2} d_{\eta,k+N}}{\sum_{k=0}^{N-1} d_{\eta,k}} = \frac{\sum_{k=0}^{N-2}(k+N+2)(k+N+1)}{\sum_{k=0}^{N-1}(k+2)(k+1)}$$

$$\approx \frac{\int_0^N (x+N)^2\,dx}{\int_0^N x^2\,dx} = \frac{(2^3-1)N^3/3}{N^3/3} = 2^3 - 1.$$

- General $\eta > 0$: Since $\Gamma(\eta + i)/i!$ behaves like $i^{\eta-1}$ for large i (by applying Stirling's formula), then for large N,

$$\frac{\sum_{k=0}^{N-2} d_{\eta,k+N}}{\sum_{k=0}^{N-1} d_{\eta,k}} \approx \frac{\int_0^N (x+N)^{\eta-1}\,dx}{\int_0^N x^{\eta-1}\,dx} = \frac{(2^\eta - 1)N^\eta/\eta}{N^\eta/\eta} = 2^\eta - 1. \qquad \square$$

Lemma 5.2. *$A_{\ell_N,N} \to 1$ as $N \to \infty$ with $\ell_N/N \to a \in [1, 2)$.*

Proof. $A_{\ell,N} = \Pr(\ell - N + 1 \le Y \le N - 1)$, where $N \le \ell \le 2N - 2$ and $Y \sim \text{Binomial}(\ell, \frac{1}{2})$. By the normal approximation for large N and ℓ, this is approximately

$$\Pr\left(\frac{\ell - N + \frac{1}{2} - \frac{1}{2}\ell}{\frac{1}{2}\sqrt{\ell}} \le Z \le \frac{N - \frac{1}{2} - \frac{1}{2}\ell}{\frac{1}{2}\sqrt{\ell}}\right)$$
$$= \Phi\left(\frac{2N - 1 - \ell}{\sqrt{\ell}}\right) - \Phi\left(\frac{\ell - 2N + 1}{\sqrt{\ell}}\right)$$

where $Z \sim N(0, 1)$ and Φ is the standard normal cumulative distribution function. Let $\ell = \ell_N = [aN]$ where $1 \le a < 2$. Then, as $N \to \infty$,

$$\Phi\left(\frac{(2-a)N}{\sqrt{aN}}\right) - \Phi\left(\frac{-(2-a)N}{\sqrt{aN}}\right) \to 1. \qquad \square$$

References

1. Abramowitz M. and Stegun I.A. (1972). *Handbook of Mathematical Functions*, Dover, New York.
2. Berman S. (1980). Stationarity, isotropy and sphericity in l_p. *Probability Theory and Related Fields*, 54(1):21–23.
3. Clayton D.G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65:141–151.
4. Cook R.D. and Johnson M.E. (1981). A family of distributions for modelling non-elliptically symmetric multivariate data. *Journal of the Royal Statistical Society: Series B*, 43:210–218.
5. Harris C. (1968). The Pareto distribution as a queue service discipline. *Operations Research*, 16:307–316.
6. Joe H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
7. Kimeldorf G. and Sampson A.R. (1975). Uniform representations of bivariate distributions. *Communications in Statistics*, 4:617–627.
8. Kousky C. and Cooke R.M. (2009). Climate change and risk management: Challenges for insurance, adaptation and loss estimation. Discussion Paper RFF-DP-09-03, Resources for the Future, Washington DC.

9. Mardia K.V. (1962). Multivariate Pareto distributions. *Annals of Mathematical Statistics*, 33:1008–1015.
10. McNeil A.J., Frey R. and Embrechts P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press, New Jersey.
11. Takahasi K. (1965). Note on the multivariate Burr's distribution. *Annals of the Institute of Statistical Mathematics*, 17:257–260.