

The Grouped t-copula with an Application to Credit Risk

Stéphane Daul
Quantitative Risk
Management Methods
Swiss Re

Enrico De Giorgi
RiskLab
ETH Zürich

Filip Lindskog
RiskLab
ETH Zürich

Alexander McNeil
ETH Zürich

We describe a model that takes into account the tail dependence present in a large set of historical risk factor data using the modern concept of copulas. We extend the popular t-copula to obtain a new *grouped t-copula* which describes more accurately the dependence among risk factors of different classes. We explain how to estimate the parameters of the grouped t-copula and apply the method to a problem in credit risk management with a large number of risk factors. We measure the downside risk over one month for an internationally diversified credit portfolio and we observe that the new model gives different results to the t-copula and seems better able to capture the risk in a large set of risk factors.

Introduction

It has become increasingly popular to model vectors of risk factor log returns with so-called meta-t distributions, i.e. distributions with a t-copula and arbitrary marginal distributions. The reason for this is the ability of the t-copula to model the extremal dependence of the risk factors and also the ease with which the parameters of the t-copula can be estimated from data.

While the meta-t assumption makes sense for risk factors of similar type, e.g. foreign exchange rates, it may not accurately describe the dependence structure for a set of risk factor log returns where the risk factors are of very different type, for example a mixture of stock indices, foreign exchange rates and interest rates. Here the assumption of one global degrees of freedom parameter may be over-simplistic as seen from Table 1. In order to justify introducing a more flexible model it would be reasonable to require that the following two points are fulfilled:

- (i) the extended model should contain the meta-t model as a special case, and
- (ii) calibration of and simulation from the extended model should be no more difficult than in the meta-t model.

In this paper we propose a new, more flexible copula satisfying (i) and (ii) and we show how to estimate the parameters of the suggested model. Moreover, we apply the new copula to the problem of describing the risk of credit portfolios driven by a large set of risk factors and we show, by studying various risk measures, how the risk changes compared to models based on a Gaussian or t-copula.

Our investigation of this new copula was motivated by a real-life application in the area of credit risk modelling, where it was clear that the dependence structure of the risk factors driving default risk would be better captured by the new grouped t-copula than the standard t-copula. We present a slightly simplified version of the original motivating application in this paper.

Model

The t-copula

We begin by recalling the definition and properties of t-distributions and t-copulas. For more on copulas in general, see Joe (1997), Nelsen (1999) or Embrechts, McNeil, and Straumann (2002), and for more on the use of copulas in credit risk modelling, see Schubert and Schönbucher (2000) and Frey, McNeil, and Nyfeler (2001).

Let $\mathbf{Z} \sim \mathcal{N}_d(\mathbf{0}, \Sigma)$ and $R = \sqrt{\nu}/\sqrt{S}$, with $S \sim \chi_\nu^2$ (a Chi Square distribution with ν degrees of freedom), be independent. Then the \mathbb{R}^d -valued random vector \mathbf{Y} given by

$$\mathbf{Y} = R\mathbf{Z} \tag{1}$$

has a centered t-distribution with ν degrees of freedom. Note that for $\nu > 2$, $\text{Cov}(\mathbf{Y}) = \frac{\nu}{\nu-2}\mathbf{\Sigma}$. By Sklar's Theorem, the copula of \mathbf{Y} can be written as

$$C_{\nu, \boldsymbol{\rho}}^t(\mathbf{u}) = t_{\nu, \boldsymbol{\rho}}^d(t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_d)), \quad (2)$$

where $\rho_{ij} = \Sigma_{ij} / \sqrt{\Sigma_{ii}\Sigma_{jj}}$ for $i, j \in \{1, \dots, d\}$ and where $t_{\nu, \boldsymbol{\rho}}^d$ denotes the distribution function of $\sqrt{\nu}\mathbf{Z}/\sqrt{S}$, where $S \sim \chi_{\nu}^2$ and $\mathbf{Z} \sim \mathcal{N}_d(\mathbf{0}, \boldsymbol{\rho})$ are independent (i.e. the usual multivariate t-distribution function) and t_{ν} denotes the marginal distribution function of $t_{\nu, \boldsymbol{\rho}}^d$ (i.e. the usual univariate t-distribution function). In the bivariate case the copula expression can be written as

$$C_{\nu, \boldsymbol{\rho}}^t(u, v) = \int_{-\infty}^{t_{\nu}^{-1}(u)} \int_{-\infty}^{t_{\nu}^{-1}(v)} \frac{1}{2\pi(1-\rho_{12}^2)^{1/2}} \left\{ 1 + \frac{s^2 - 2\rho_{12}st + t^2}{\nu(1-\rho_{12}^2)} \right\}^{-(\nu+2)/2} ds dt. \quad (3)$$

Note that ρ_{12} is simply the usual linear correlation coefficient of the corresponding bivariate t_{ν} -distribution if $\nu > 2$. The density function of the t-copula is given by

$$c_{\nu, \boldsymbol{\rho}}^t(u_1, \dots, u_d) = \frac{1}{\sqrt{\det \boldsymbol{\rho}}} \frac{\Gamma(\frac{\nu+d}{2})\Gamma(\frac{\nu}{2})^{d-1} \prod_{k=1}^d \left(1 + \frac{y_k^2}{\nu}\right)^{\frac{\nu+1}{2}}}{\Gamma(\frac{\nu+1}{2})^d \left(1 + \frac{\mathbf{y}'\boldsymbol{\rho}^{-1}\mathbf{y}}{\nu}\right)^{\frac{\nu+d}{2}}}, \quad (4)$$

where $y_k = t_{\nu}^{-1}(u_k)$.

Let H_1, \dots, H_d be arbitrary continuous, strictly increasing distribution functions and let \mathbf{Y} be given by (1) with $\mathbf{\Sigma}$ a linear correlation matrix. Then

$$\mathbf{X} = (H_1^{-1}(t_{\nu}(Y_1)), \dots, H_d^{-1}(t_{\nu}(Y_d)))' \quad (5)$$

has a t_{ν} -copula and marginal distributions H_1, \dots, H_d . The distribution of \mathbf{X} is referred to as a meta-t distribution (see Fang, Fang, and Kotz 2002). Note that \mathbf{X} has a t-distribution if and only if H_1, \dots, H_d are univariate t_{ν} -distribution functions.

The coefficient of tail dependence expresses the limiting conditional probability of joint quantile exceedences (see Joe 1997). The t-copula has upper and lower tail dependence ($\bar{t}_{\nu+1}(x) = 1 - t_{\nu+1}(x)$):

$$\lambda = 2\bar{t}_{\nu+1} \left(\sqrt{\nu+1}\sqrt{1-\rho_{12}}/\sqrt{1+\rho_{12}} \right) > 0,$$

in contrast with the Gaussian copula which has $\lambda = 0$ (for a proof see Embrechts, McNeil, and Straumann 2002). From the above expression it is also seen that the coefficient of tail dependence is increasing in ρ_{12} and, as one would expect since a t-distribution converges to a normal distribution as ν tends to infinity, decreasing in ν . Furthermore, the coefficient of upper (lower) tail dependence tends to zero as the number of degrees of freedom tends to infinity for $\rho_{12} < 1$.

The estimation of the copula parameters are typically done as follows.

- (i) Kendall's tau is estimated for every pair of risk factor log returns. An estimate of the parameter $\boldsymbol{\rho}$ in (2) is obtained from the relation

$$\tau(X_i, X_j) = \tau(Z_i, Z_j) = \frac{2}{\pi} \arcsin(\rho_{ij}) \quad (6)$$

which holds for the distributions defined by (1) and (5). In fact it holds for any distribution with strictly increasing marginal distribution functions and a copula of an elliptical distribution which has a density (see Lindskog, McNeil, and Schmock (2003) for the more general version of this result), i.e. essentially any meta-elliptical distribution one would consider in applications. Note that in high-dimensional applications an estimate of $\boldsymbol{\rho}$ obtained from (6) may have to be modified to assure positive definiteness. This can be done by applying the so-called eigenvalue method, i.e. the negative eigenvalues are replaced by a small positive number (see Rousseeuw and Molenberghs 1993).

- (ii) Transforming each log return observation with its respective distribution function yields, under the meta-t assumption, a sample from a t-copula with known $\boldsymbol{\rho}$ -parameter. Finally, the degrees of freedom parameter ν is estimated by standard maximum likelihood estimation using (4).

In step (ii) we could use the empirical marginals or fitted distribution functions from a parametric family.

The grouped t-copula

Consider now the following model. Let $\mathbf{Z} \sim \mathcal{N}_J(\mathbf{0}, \boldsymbol{\rho})$, where $\boldsymbol{\rho}$ is an arbitrary linear correlation matrix, be independent of U , a random variable uniformly distributed on $(0, 1)$. Furthermore, let G_ν denote the distribution function of $\sqrt{\nu/\chi_\nu^2}$. Partition $\{1, \dots, J\}$ into m subsets of sizes s_1, \dots, s_m . Let $R_k = G_{\nu_k}^{-1}(U)$ for $k = 1, \dots, m$. If

$$\mathbf{Y} = (R_1 Z_1, \dots, R_1 Z_{s_1}, R_2 Z_{s_1+1}, \dots, R_2 Z_{s_1+s_2}, \dots, R_m Z_J)', \quad (7)$$

then the random vector $(Y_1, \dots, Y_{s_1})'$ has an s_1 -dimensional t-distribution with ν_1 degrees of freedom and, for $k = 1, \dots, m-1$, $(Y_{s_1+\dots+s_{k+1}}, \dots, Y_{s_1+\dots+s_{k+1}})'$ has an s_{k+1} -dimensional t-distribution with ν_{k+1} degrees of freedom (compare expressions (7) and (1)). Finally, let F_k denote the distribution function of Y_k and let H_1, \dots, H_J be some arbitrary continuous strictly increasing distribution functions. Then

$$\mathbf{X} = (H_1^{-1}(F_1(Y_1)), \dots, H_J^{-1}(F_J(Y_J)))'$$

is a generalisation of the meta-t model (5) which allows different subsets of the components to have different degrees of freedom parameters and its copula will be called the grouped t-copula. Note that if we take $m = 1$, $\nu_1 = \nu$, then \mathbf{X} has a meta-t distribution (and is equal in distribution to (5) if $d = J$ and $\boldsymbol{\Sigma} = \boldsymbol{\rho}$).

Simulation from the grouped t-copula is no more difficult than simulation from a t-copula.

- (i) Draw independently a random variate \mathbf{Z} from the J -dimensional normal distribution with zero mean, unit variances and linear correlation matrix $\boldsymbol{\rho}$, and a random variate U from the uniform distribution on $(0, 1)$.
- (ii) Obtain R_1, \dots, R_m by setting $R_k = G_{\nu_k}^{-1}(U)$ for $k = 1, \dots, m$. By (7) we obtain a random variate $(Y_1, \dots, Y_J)'$ from the grouped t-distribution.
- (iii) Finally,

$$(t_{\nu_1}(Y_1), \dots, t_{\nu_1}(Y_{s_1}), t_{\nu_2}(Y_{s_1+1}), \dots, t_{\nu_2}(Y_{s_1+s_2}), \dots, t_{\nu_m}(Y_J))'$$

is a random variate from the grouped t-copula.

Note that the grouped t-copula can be written down in a form similar to (3). However, we believe that the properties of the grouped t-copula is best understood from (7) and the above stochastic representation. Moreover, for calibration (see below) and simulation from the grouped t-copula there is no need for an explicit copula expression.

The calibration of this model is identical to that of the meta-t distribution except that the ML-estimation of the m degrees of freedom parameters has to be performed separately on each of the m risk factor subgroups. The key point is that the approximation

$$\tau(X_i, X_j) \approx \tau(Z_i, Z_j) = \frac{2}{\pi} \arcsin(\rho_{ij}) \quad (8)$$

is very accurate as seen in the Appendix. Again, the eigenvalue method may have to be applied to assure positive definiteness.

Application

We consider an internationally diversified credit portfolio with K counterparties. It is assumed that the systematic risk of each counterparty is adequately described by a set of risk factors, which are 92 country/industry equity indices (see Table 1); these are divided into 8 groups defined by country. We use the grouped t-copula to describe the dependence structure of the risk factors, and complete the model by specifying normally distributed marginals. The assumption of normally distributed marginals is reasonable for monthly returns and more sophisticated marginal assumptions would have little bearing on the results of this paper. We evaluate value-at-risk and expected shortfall for the one month portfolio loss distribution and we compare our results with value-at-risk and expected shortfall in the case that the risk factors follow a multivariate Gaussian distribution or a meta-t distribution with normal marginals.

The credit model

Consider a single counterparty k and the time horizon $T = 1$ month. Let I_k be the state variable for counterparty k at time horizon T . We consider only default events and not the impact of upgrades or downgrades on the credit quality. Therefore, we assume that I_k takes values in $\{0, 1\}$: the value 0 represents the default state, the value 1 is the non-default state. Let Y_k be a random variable with continuous distribution function $F_k(x) = \mathbb{P}[Y_k \leq x]$ and let $d_k \in \mathbb{R}$ such that $I_k = 0$ if and only if $Y_k \leq d_k$. The parameter d_k is called the *default threshold* and (Y_k, d_k) is the *latent variable model* for I_k , as described for example in Frey, McNeil, and Nyfeler (2001). As in KMV (see Crosbie

and Bohn 2002) Y_k represents the asset value monthly log return of counterparty k . We propose the following model for Y_k . For parameters \mathbf{a}_k and $\lambda_k \in [0, 1]$ let

$$Y_k = \sqrt{\lambda_k} \mathbf{a}'_k \boldsymbol{\Theta} + \sqrt{1 - \lambda_k} s_k \epsilon_k, \quad (9)$$

where $\boldsymbol{\Theta}$ is the vector of monthly risk factor log returns with a grouped t-copula and normally distributed marginals, $\mathbb{E}[\boldsymbol{\Theta}] = \mathbf{0}$ and $\epsilon_k \sim N(0, 1)$, independent of $\boldsymbol{\Theta}$. The parameter λ_k is the coefficient of determination for the systematic risk (how much of the variance can be explained by the risk factors) and $s_k^2 = \text{Var}(Y_k) = \mathbf{a}'_k \text{Cov}(\boldsymbol{\Theta}) \mathbf{a}_k$. Let π_k be the (unconditional) probability of default for counterparty k , i.e. $\pi_k = F_k(d_k)$. π_k is assumed to be given from some internal or external rating system or other procedures. The conditional probability of default for counterparty k given the risk factors $\boldsymbol{\Theta}$ can be written as

$$Q_k(\boldsymbol{\Theta}) = \mathbb{P}[Y_k \leq d_k | \boldsymbol{\Theta}] = \Phi \left(\frac{F_k^{-1}(\pi_k) - \sqrt{\lambda_k} \mathbf{a}'_k \boldsymbol{\Theta}}{\sqrt{1 - \lambda_k} \sqrt{\mathbf{a}'_k \text{Cov}(\boldsymbol{\Theta}) \mathbf{a}_k}} \right), \quad (10)$$

where Φ denotes the standard normal distribution function. Unlike in classical models such as KMV and CreditMetrics the distribution function of Y_k is unknown. Therefore, we work with the estimated conditional probability of default $\hat{Q}_k(\boldsymbol{\Theta})$ obtained by replacing $F_k^{-1}(\pi_k)$ in (10) by the empirical quantile estimate $\hat{F}_k^{-1}(\pi_k)$.

The default model described above is applied to each single counterparty in the credit portfolio with the additional assumption that counterparty defaults are conditionally independent given the risk factors $\boldsymbol{\Theta}$, i.e. the ϵ_k 's are independent. Conditional on a scenario $\boldsymbol{\Theta}$ counterparty defaults are simulated from independent Bernoulli distributions with parameters $\hat{Q}_k(\boldsymbol{\Theta})$. Let $I_k(\boldsymbol{\Theta}) \in \{0, 1\}$ be the conditional default indicator for counterparty k , let E_k be the corresponding exposure and let l_k be the loss given default. Then $L(\boldsymbol{\Theta}) = \sum_{k=1}^K I_k(\boldsymbol{\Theta}) l_k E_k$ gives the total credit loss under scenario $\boldsymbol{\Theta}$. Hence, the credit loss distribution is obtained by a three stage procedure:

- (i) Simulation of the monthly risk factor log returns $\boldsymbol{\Theta}$ from a grouped t-copula with normal marginals;
- (ii) For each counterparty k , simulation of the conditional default indicator $I_k(\boldsymbol{\Theta}) \in \{0, 1\}$ from a Bernoulli distribution with parameter $\hat{Q}_k(\boldsymbol{\Theta})$;
- (iii) Estimation of the credit loss distribution over a large set of scenarios for $\boldsymbol{\Theta}$, by integrating exposures and loss given default in the loss function $L(\boldsymbol{\Theta})$.

Data set and results

We calibrate the grouped t-copula and the normally distributed marginals using monthly risk factor log returns from 1992 to 2002 (hence 120 observations) obtained from DataStreamTM. Table 1 shows the estimated degrees of freedom parameters for various subsets of risk factors and the overall estimated degrees of freedom parameter. The difference between the various subset degrees of freedom parameters suggests that a grouped t-copula is more appropriate for describing the dependence structure.

Set	Number of risk factors	ν
AUS Indices	9	15
CAN Indices	14	24
CH Indices	4	19
FRA Indices	5	67
GER Indices	10	65
JPN Indices	15	14
UK Indices	15	17
US Indices	20	21
All	92	29

Table 1: Estimated degrees of freedom ν for various sets of risk factors. The country equities indices are for major industrial sectors.

The credit portfolio contains $K = 200$ counterparties with the same unconditional default probability $\pi = 1\%$. Each counterparty is assigned to a country so that there are 25 from each country. The weights \mathbf{a}_k and λ_k ($k = 1, \dots, 200$) are generated as follows. For λ_k we randomly choose values between 20% and 60%, which are common in credit modelling. Each counterparty is then described by two different risk factors (labelled i_1 and i_2) from the country to which it has been assigned, and the value of $a_k^{i_1}$ (and hence also that of $a_k^{i_2}$) are drawn from a uniform distribution on $(0, 1)$ such

that $a_k^{i_1} + a_k^{i_2} = 1$. Moreover, each counterparty has a total exposure of 1000 CHF and the loss given default is assumed to be uniformly distributed on $(0, 1)$.

We performed 500'000 simulations using a 92-dimensional Sobol sequence (see for example Press, Teukolsky, Vetterling, and Flannery 1992). The simulated default frequencies were all in the range 0.97 % - 1.03 % and the expected value of the portfolio loss distribution was estimated with less than 0.2% error.

We present the results by making a comparison between our new model incorporating the grouped t-copula (with degrees of freedom parameters shown in Table 1) and (1) a model with a t-copula (with 29 degrees of freedom) and (2) a model with a Gaussian copula (as in CreditMetrics and KMV); we take the Gaussian model as baseline and express the differences in the risk measures as percentages. In Table 2 we show various risk measures for the total credit loss distribution. Taking the tail dependence into account with the t-copula gives an increased assessment of the risk. When introducing the grouped t-copula, we get even larger risk measures; the 99%-shortfall is more than 10% larger than the normal case.

Measure	t ₂₉	grouped-t
Max. value	29.7%	41.4%
Std dev.	4.2%	5.3%
95% quantile	1.1%	1.7%
99% quantile	4.3%	6.0%
95% shortfall	3.4%	4.8%
99% shortfall	8.9%	10.8%

Table 2: Risk measures of the sample portfolio (described in text) using a t₂₉-copula or a grouped t-copula to model the dependence among the 92 risk factors. The values shown are the percentage deviations from those obtained with the normal copula.

Conclusion

We have described how to model large sets of risk factors of different classes using a grouped t-copula. This copula has the property that the random variables within each group have a t-copula with possibly different degrees of freedom parameters in the different groups, which gives a more flexible overall dependence structure more suitable for large sets of risk factors. Setting the degrees of freedom parameters to be equal yields the usual t-copula, and letting this common degrees of freedom parameter tend to infinity yields the Gaussian copula. Hence the grouped t-copula is an extension of the popular but less flexible Gaussian and t-copulas. When calibrated to our historical risk factor data set, we obtain a model for the risk factors which allows us to more accurately model the tail dependence present in the data. We then applied this model to the measure of credit risk for a portfolio driven by 92 risk factors. Although the estimated degrees of freedom parameters for the grouped t-copula look quite high it must be appreciated that these still give significantly increased risk measures compared to the Gaussian copula in a high dimensional application such as ours. This is particularly true in the setting of credit risk where joint extreme risk factors can lead to many more counterparty defaults.

References

- CROSBIE, P., AND J. BOHN (2002): “Modeling Default Risk,” *KMV*TM.
- EMBRECHTS, P., A. MCNEIL, AND D. STRAUMANN (2002): “Correlation and Dependence in Risk Management: Properties and Pitfalls,” in *Risk Management: Value at Risk and Beyond*, ed. by M. Dempster, pp. 176–223. Cambridge University Press, Cambridge.
- FANG, H.-B., K. FANG, AND S. KOTZ (2002): “The Meta-elliptical Distributions with Given Marginals,” *Journal of Multivariate Analysis*, 82(1), 1–16.
- FREY, R., A. MCNEIL, AND M. NYFELER (2001): “Copulas and credit models,” *RISK*, October, 111–114.
- JOE, H. (1997): *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
- LINDSKOG, F., A. MCNEIL, AND U. SCHMOCK (2003): “A note on Kendall’s tau for elliptical distributions,” in *Credit Risk. Measurement, Evaluation and Management*, ed. by G. Bol, G. Nakhaeizadeh, S. Rachev, T. Ridder, and K.-H. Vollmer, pp. 149–156. Physica-Verlag, Heidelberg.

NELSEN, R. (1999): *An Introduction to Copulas*. Springer, New York.

PRESS, W., S. TEUKOLSKY, W. VATTERLING, AND B. FLANNERY (1992): *Numerical Recipes in C*. Cambridge University Press, Cambridge.

ROUSSEEUW, P., AND G. MOLENBERGHS (1993): “Transformations of Non Positive Semidefinite Correlation Matrices,” *Communications in Statistics - Theory and Methods*, 22(4), 965–984.

SCHUBERT, D., AND P. SCHÖNBUCHER (2000): “Copula-Dependent Default Risk in Intensity Models,” University of Bonn, preprint, available at www.finasto.uni-bonn.de/schonbucher/.

Appendix

We show here that the approximation

$$\tau(X_i, X_j) \approx \tau(Z_i, Z_j) = \frac{2}{\pi} \arcsin(\rho_{ij}) \quad (11)$$

is very accurate. In the proposition below we calculate the left-hand side of (11). The discrepancy with the right-hand side is shown in Table 3. Note that the two last rows of Table 3 represent the worst case since the absolute value of the error in the approximation increases as $|\rho_{ij}|$ increases. The error increases nonlinearly in $|\rho_{ij}|$ (because of the arcsin function). This is clear from observing the values in the table: the increase in error from changing the correlation coefficient from 0.5 to 0.75 is very small but the increase in error from changing the correlation coefficient from 0.75 to 1 is substantially bigger. Note also that the error is symmetric in ρ_{ij} , i.e. the corresponding errors for negative correlations are obtained by just changing the sign of the values in Table 3. The accuracy of this approximation implies that the linear correlation matrix of the underlying multivariate normal random vector \mathbf{Z} can be estimated by using (6). We note that our linear correlation estimates were all less than 0.92 so the approximation was indeed very accurate.

The exact expression for $\tau(X_i, X_j)$ is given in the following proposition. Note that Kendall’s tau is invariant under strictly increasing marginal transformations of the underlying random vector.

$\bullet 10^{-3}$	ν_1									
$\rho = 0.5$	4	6	8	10	12	14	16	18	20	22
$\nu_2 = 8$	0.96	0.14	0.00	0.06	0.20	0.33	0.49	0.63	0.79	0.91
$\nu_2 = 14$	2.42	0.91	0.35	0.11	0.02	0.00	0.01	0.05	0.09	0.13
$\rho = 0.75$										
$\nu_2 = 8$	1.86	0.27	0.00	0.12	0.38	0.67	0.98	1.23	1.53	1.77
$\nu_2 = 14$	4.77	1.77	0.67	0.22	0.04	0.00	0.03	0.09	0.17	0.26
$\rho = 0.9$										
$\nu_2 = 8$	3.38	0.48	0.00	0.22	0.68	1.23	1.77	2.26	2.75	3.23
$\nu_2 = 14$	8.28	3.27	1.22	0.40	0.08	0.00	0.05	0.16	0.31	0.47
$\rho = 1$										
$\nu_2 = 8$	37.7	14.3	0.00	9.69	16.9	22.4	27.1	30.1	33.7	36.3
$\nu_2 = 14$	60.2	37.1	22.5	12.7	5.50	0.00	4.48	8.14	11.2	13.7

Table 3: MC-simulation of $\frac{2}{\pi} \arcsin(\rho) - \frac{2}{\pi} \mathbb{E}[\arcsin(\rho \psi(f_1(R), f_2(R), f_1(\tilde{R}), f_2(\tilde{R})))]$ for $\rho = 0.5$, $\rho = 0.75$, $\rho = 0.9$ and $\rho = 1$ based on independent samples of size 10000, where $f_1^2(R), f_1^2(\tilde{R}) \sim G_{\nu_1}$, $f_2^2(R), f_2^2(\tilde{R}) \sim G_{\nu_2}$ and the two comonotonic pairs $(f_1(R), f_2(R))$ and $(f_1(\tilde{R}), f_2(\tilde{R}))$ are independent. Note that the values should be multiplied by 10^{-3} .

Proposition. Let $\mathbf{Z} \sim \mathcal{N}_2(\mathbf{0}, \Sigma)$, let R be a strictly positive random variable and let $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$ be such that $f_1(x), f_2(x) > 0$ for $x > 0$. Then

$$\tau(f_1(R)Z_1, f_2(R)Z_2) = \frac{2}{\pi} \mathbb{E}(\arcsin(\rho_{12} \psi(f_1(R), f_2(R), f_1(\tilde{R}), f_2(\tilde{R}))))$$

where \tilde{R} is an independent copy of R , $\rho_{12} = \Sigma_{12} / \sqrt{\Sigma_{11}\Sigma_{22}}$ and

$$\psi(f_1(r), f_2(r), f_1(\tilde{r}), f_2(\tilde{r})) = \frac{f_1(r)f_2(r) + f_1(\tilde{r})f_2(\tilde{r})}{\sqrt{(f_1^2(r) + f_1^2(\tilde{r}))(f_2^2(r) + f_2^2(\tilde{r}))}}.$$

Proof. Let $\tilde{\mathbf{Z}}$ be an independent copy of \mathbf{Z} . For constants $r, \tilde{r} > 0$,

$$\begin{pmatrix} f_1(r)Z_1 \\ f_2(r)Z_2 \end{pmatrix} + \begin{pmatrix} f_1(\tilde{r})\tilde{Z}_1 \\ f_2(\tilde{r})\tilde{Z}_2 \end{pmatrix} = \begin{pmatrix} f_1(r) & 0 & f_1(\tilde{r}) & 0 \\ 0 & f_2(r) & 0 & f_2(\tilde{r}) \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ \tilde{Z}_1 \\ \tilde{Z}_2 \end{pmatrix},$$

where $(Z_1, Z_2, \tilde{Z}_1, \tilde{Z}_2)'$ is multivariate normally distributed. Hence

$$\begin{pmatrix} f_1(r)Z_1 \\ f_2(r)Z_2 \end{pmatrix} + \begin{pmatrix} f_1(\tilde{r})\tilde{Z}_1 \\ f_2(\tilde{r})\tilde{Z}_2 \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} \hat{Z}_1(r, \tilde{r}) \\ \hat{Z}_2(r, \tilde{r}) \end{pmatrix},$$

where $\hat{\mathbf{Z}}(r, \tilde{r}) \sim \mathcal{N}_2(\mathbf{0}, \boldsymbol{\Sigma}^{(r, \tilde{r})})$ with

$$\boldsymbol{\Sigma}^{(r, \tilde{r})} = \begin{pmatrix} \Sigma_{11}(f_1^2(r) + f_1^2(\tilde{r})) & \Sigma_{12}(f_1(r)f_2(r) + f_1(\tilde{r})f_2(\tilde{r})) \\ \Sigma_{12}(f_1(r)f_2(r) + f_1(\tilde{r})f_2(\tilde{r})) & \Sigma_{22}(f_2^2(r) + f_2^2(\tilde{r})) \end{pmatrix}.$$

Hence $\hat{\mathbf{Z}}(r, \tilde{r})$ has linear correlation coefficient $\rho_{12}\psi(f_1(r), f_2(r), f_1(\tilde{r}), f_2(\tilde{r}))$. We can now compute Kendall's tau for $(f_1(R)Z_1, f_2(R)Z_2)'$,

$$\begin{aligned} \tau(f_1(R)Z_1, f_2(R)Z_2) &= 2\mathbb{P}[(f_1(R)Z_1 - f_1(\tilde{R})\tilde{Z}_1)(f_2(R)Z_2 - f_2(\tilde{R})\tilde{Z}_2) > 0] - 1 \\ &= 2\mathbb{E}[\mathbb{P}[(f_1(R)Z_1 - f_1(\tilde{R})\tilde{Z}_1)(f_2(R)Z_2 - f_2(\tilde{R})\tilde{Z}_2) > 0 \mid (R, \tilde{R})]] - 1 \\ &= 2 \int \int \mathbb{P}[(f_1(r)Z_1 - f_1(\tilde{r})\tilde{Z}_1)(f_2(r)Z_2 - f_2(\tilde{r})\tilde{Z}_2) > 0] dF_R(r) dF_{\tilde{R}}(\tilde{r}) - 1 \\ &= 2 \int \int \mathbb{P}[\hat{Z}_1(r, \tilde{r})\hat{Z}_2(r, \tilde{r}) > 0] dF_R(r) dF_{\tilde{R}}(\tilde{r}) - 1 \\ &= 2 \int \int \left(\frac{1}{2} + \frac{1}{\pi} \arcsin(\rho_{12}\psi(f_1(r), f_2(r), f_1(\tilde{r}), f_2(\tilde{r}))) \right) dF_R(r) dF_{\tilde{R}}(\tilde{r}) - 1 \\ &= \frac{2}{\pi} \mathbb{E}[\arcsin(\rho_{12}\psi(f_1(R), f_2(R), f_1(\tilde{R}), f_2(\tilde{R})))]. \end{aligned}$$

□