High Dimensional Modelling and Simulation with Asymmetric Normal Mixtures *

Andreas Tsanakas†
Cass Business School, City University

Andrew Smith
Capital Markets Group, Deloitte

Abstract

A family of multivariate distributions, based on asymmetric normal mixtures, is introduced in order to model the dependence among insurance and financial risks. The model allows for straight-forward parameterisation via a correlation matrix and enables the modelling of radially asymmetric dependence structures, which are often of interest in risk management applications. Dependence is characterised by showing that increases in correlation values produce models which are ordered in the supermodular order sense. Explicit expressions for the Spearman and Kendall rank correlation coefficients are derived to enable calibration in a copula framework. The model is adapted to simulation in very high dimensions by using Kronecker products, enabling specification of a correlation matrix and an increase in computational speed.

Keywords: Dependence, copula, normal mixtures, Kronecker product, Monte-Carlo simulation.

*Part of this research was carried out while the corresponding author was employed by Lloyd's of London. The authors thanks Evangelia Desli, Peter Tavner and Henry Johnson for useful discussions and for giving them an interesting problem to solve. The views expressed in this paper are solely those of the authors and do not necessarily represent the views of Lloyd's.

†Corresponding author. Address: Cass Business School, 106 Bunhill Row, London EC1Y 8TZ, UK. Email: a(dot)tsanakas(dot)1(at)city(dot)ac(dot)uk

1 Introduction

The importance for risk management of modelling the dependence between multivariate insurance or financial risks has been well established. The effect of dependence on aggregate risk positions is discussed in Dhaene and Goovaerts (1996), Müller and Stoyan (2002), while model choices diverging from the multivariate normal paradigm are extensively discussed by Frees and Valdez (1998), Embrechts et al (2002), who propose the use of *copula* functions for modelling dependencies between risks. Extensive reviews of dependence models for insurance and financial risk management purposes can be found in McNeil et al (2005), Denuit et al (2005).

Copulas form a tool for constructing joint probability distributions, by separating the dependence structure from the marginal distributions of multivariate models. By a result known as *Sklar's Theorem* (e.g. Nelsen, 1999), for every n-dimensional random vector $\mathbf{X}$ with joint distribution function $F_{\mathbf{X}}$ and marginal distributions $F_i,\ i = 1, \dots, n$, there is a function $C : [0,1]^n \mapsto [0,1]$, called the copula of $\mathbf{X}$ such that

$$F_{\mathbf{X}}(x_1, \dots, x_n) = C\left(F_1(x_1), \dots, F_n(x_n)\right) \tag{1}$$

One way of constructing copulas is by considering a tractable family of multivariate distributions, such as the multivariate normal or t, and then obtaining the corresponding copula by transforming the marginal distributions to uniforms in $[0,1]$, i.e. by:

$$C(u_1, \dots, u_n) = F_{\mathbf{X}}(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)), \tag{2}$$

where $F_1^{-1}, \dots, F_n^{-1}$ are the (generalised) inverses of $F_1, \dots, F_n$. This technique is of particular interest when the random vector $\mathbf{X}$ is easy to simulate from; then simulation from an arbitrary random vector $\mathbf{Y}$ with marginal distributions $G_1, \dots, G_n$ and copula C is easily performed by $Y_1 \overset{d}{=} G_1^{-1}\left(F_1(X_1)\right), \dots, Y_n \overset{d}{=} G_n^{-1}\left(F_n(X_n)\right)$.

Through their shared copula, random vector $\mathbf{Y}$ thus inherits the dependence properties of $\mathbf{X}$. Hence it is important to specify what a desirable set of properties for $\mathbf{X}$ are. One possible list is:

1. Ability to simulate from $\mathbf{X}$ efficiently, potentially in high dimensions.
2. Specification via a correlation matrix, since pairwise correlations form one of the most popular (and interpretable) characterisations of dependence.
3. Flexibility to allow for asymmetry and/or tail dependence, as these are frequent features of insurance/financial data.
4. Characterisation of the resulting dependence structure via stochastic order relations, as this enables better understanding of the model and facilitates practical sensitivity testing (this point is discussed further in Section 3.1.

In this contribution a multivariate probability distribution is constructed that satisfies the properties listed above. The model emerges as an asymmetric (mean-variance) mixture of multivariate normal distributions. Mixtures of normal distributions, resulting from a randomisation of the covariance matrix, form a subcase of the more general class or elliptical distributions (Fang et al, 1987). Such distributions are easy to simulate and specify via a correlation matrix (or a generalisation thereof when second moments do not exist), as well as potentially displaying asymptotic tail dependence, as shown by Hult and Lindskog (2002). These models are however still radially symmetric, hence dependence for the left tails is the same as for the right ones. This can be addressed by considering normal mixtures where the mean is randomised as well as the covariance, as in Barndorff-Nielsen (1978), Demarta and McNeil (2005). However in these models the resulting correlation matrix is not easily specified from input parameters, which complicates parameterisation. The construction of asymmetric normal mixtures presented here addresses this problem.

The models discussed above all rely on the definition of a dispersion (correlation) matrix. Defining such a matrix can however become problematic in very high dimensions, e.g. of the order of 1000. For such cases the construction of correlation matrices via Kronecker products is proposed, and it is demonstrated how such a construction also improves computational efficiency.

Section 2 introduces the asymmetric mixed normal model proposed, along with its basic properties. In Section 3 the dependence properties are more closely examined, in particular with respect to stochastic orders, and formulas for rank correlations are derived. In Section 4, parameterisation issues are further discussed. Finally, in Section 5 the use of Kronecker products for the construction of very high dimensional dependence models is proposed.

2 The asymmetric mixed normal model

2.1 Definition and basic properties

The asymmetric mixed normal model proposed in this contribution was originally introduced by Smith (2002).

Definition 1. *An n-dimensional random vector $\mathbf{X}$ is defined as having a* multivariate asymmetric mixed normal distribution*, denoted by $\mathbf{X} \sim ANM_n(G, \mathbf{\Sigma}, \mathbf{u})$ whenever it satisfies*

$$\mathbf{X} \stackrel{d}{=} \gamma^{-1} \cdot (H-1) \cdot \mathbf{u} + \sqrt{H} \cdot \mathbf{L} \cdot \mathbf{Z}, \tag{3}$$

where,

- $\mathbf{Z}$ *is an n-vector of independent standard normal variables.*

- *H is a non-negative random variable, independent of $\mathbf{Z}$, following cumulative distribution G with mean 1 and standard deviation γ.*
- *$\mathbf{\Sigma} \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix and $\mathbf{L} \in \mathbb{R}^{n\times n}$ is such that $\mathbf{L} \cdot \mathbf{L}^T = \mathbf{\Sigma} - \mathbf{u} \cdot \mathbf{u}^T$.*
- *$\mathbf{u} \in \mathbb{R}^n$ is a 'non-centrality vector' such that $\mathbf{u}^T \cdot \mathbf{\Sigma}^{-1} \cdot \mathbf{u} \leq 1$.*

Note that when $\mathbf{u} = 0$, $\mathbf{X}$ reduces to a normal mixture, which is known to represent a large subclass of elliptical distributions (Fang et al., 1987). Conditional upon H, (3) represents a multivariate normal with mean vector $\gamma^{-1} \cdot (H-1) \cdot \mathbf{u}$ and covariance matrix $H \cdot \mathbf{\Sigma}$. Intuitively, the effect of (positive) non-centralities is that, conditional upon H, a high mean vector corresponds to high covariance and this is where the asymmetry originates.

The next few lemmas present stylised facts regarding the elementary properties of the multivariate asymmetric mixed normal distribution. The first two moments of $\mathbf{X}$ are given below[1].

Lemma 1. *Consider $\mathbf{X} \sim ANM_n(G, \mathbf{\Sigma}, \mathbf{u})$. It is:*

$$E(X_i) = 0, \quad i = 1, 2, \ldots, n. \tag{4}$$

$$Var(X_i) = \sigma_{ii}, \quad i = 1, 2, \ldots, n. \tag{5}$$

$$Cov(\mathbf{X}) = \mathbf{\Sigma}. \tag{6}$$

Proof. To begin with, denote

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_1 \\ \mathbf{L}_2 \\ \vdots \\ \mathbf{L}_n \end{pmatrix} \tag{7}$$

with $\mathbf{L}_i \in \mathbb{R}^{1\times n}$. From $\mathbf{L}\mathbf{L}^T = \mathbf{\Sigma} - \mathbf{u}\mathbf{u}^T$ it follows that $\mathbf{L}_i^T \mathbf{L}_i = \sum_{j=1}^n l_{ij}^2 = \sigma_{ii} - u_i^2$ and that $\mathbf{L}_i^T \mathbf{L}_j = \sigma_{ij} - u_i u_j$. It now is

$$\begin{aligned} E(X_i) &= E\left[E\left(\gamma^{-1}(H-1)u_i + \sqrt{H}\mathbf{L}_i\mathbf{Z}\,|H\right)\right] \\ &= E\left[\gamma^{-1}(H-1)u_i + \sqrt{H}E\left(\textstyle\sum_{j=1}^n l_{ij}Z_j\,|H\right)\right] \\ &= \gamma^{-1}E(H-1)u_i + E\left[\sqrt{H}\textstyle\sum_{j=1}^n l_{ij}E\left(Z_j\right)\right] = 0 \end{aligned} \tag{8}$$

[1] It is noted that for heavy tailed H the moments of $\mathbf{X}$ may not exist. However, similarly to the case of elliptical distributions, the discussion still holds, with the matrix $\mathbf{\Sigma}$ itself being finite.

For the variances it is

$$\begin{array}{rcl} Var(X_i) & = & E\left[E\left(X_i^2\,|H\right)\right] \\ & = & E\left[E\left(\gamma^{-2}(H-1)^2u_i^2 + H(\mathbf{L}_i\mathbf{Z})^2 + 2\gamma^{-1}(H-1))u_i\sqrt{H}\mathbf{L}_i\mathbf{Z}\,|H\right)\right] \\ & = & \gamma^{-2}E[(H-1)^2]u_i^2 + E\left[HE\left(\left(\sum_{j=1}^n l_{ij}Z_j\right)^2\right)\right] \\ & & +E\left[2\gamma^{-1}(H-1))u_i\sqrt{H}E\left(\sum_{j=1}^n l_ijZ_j\right)\right] \\ & = & u_i^2 + E\left(H\right)\sum_{j=1}^n l_{ij}^2 + 0 \\ & = & u_i^2 + \sigma_{ii} - u_i^2 = \sigma_{ii} \end{array} \tag{9}$$

Finally, for the covariance matrix of $\mathbf{X}$ we find

$$\begin{array}{rcl} Cov(\mathbf{X}) & = & E\left[E\left(\mathbf{X}\mathbf{X}^T\,|H\right)\right] \\ & = & E\left\{E\left[\left(\gamma^{-1}(H-1)\mathbf{u}+\sqrt{H}\mathbf{L}\mathbf{Z}\right)\cdot\left(\gamma^{-1}(H-1)\mathbf{u}^T+\sqrt{H}\mathbf{Z}^T\mathbf{L}^T\right)|H\right]\right\} \\ & = & \gamma^{-2}E\left[(H-1)^2\right]\mathbf{u}\mathbf{u}^T + E\left[HCov\left(\mathbf{L}\mathbf{Z}\right)\right] + 2\gamma^{-1}\mathbf{u}E\left[(H-1)\sqrt{H}E\left(\mathbf{L}\mathbf{Z}\right)^T\right] \\ & = & \mathbf{u}\mathbf{u}^T + E\left[H\left(\mathbf{\Sigma}-\mathbf{u}\mathbf{u}^T\right)\right] + 0 = \mathbf{\Sigma} \end{array} \tag{10}$$

□

Hence, the asymmetric mixed normal distribution as defined in (3) has zero mean and covariance matrix equal to the specified $\mathbf{\Sigma}$. A non-zero mean vector could of course be added, but this is not further considered here. Moreover, as it is the copula of $\mathbf{X}$ that is primarily of interest, we will sometimes consider standardised versions of $\mathbf{X}$, where $\mathbf{\Sigma}$ is a correlation matrix. In the next lemma it is shown that linear transformations of asymmetric normal mixtures are themselves asymmetric normal mixtures.

Lemma 2. *Consider* $\mathbf{X} \sim ANM_n(G, \mathbf{\Sigma}, \mathbf{u})$ *and* $\mathbf{A}$ *is a* $m \times n$ *dimensional matrix of rank* $m \leq n$. *Then* $\mathbf{AX} \sim ANM_m(G, \mathbf{A\Sigma A}^T, \mathbf{Au})$.

Proof. Consider:

$$\mathbf{AX} = \gamma^{-1}(H-1)\mathbf{Au} + \sqrt{H}\mathbf{ALZ} \tag{11}$$

It is sufficient that $(\mathbf{AL})(\mathbf{AL})^T = \mathbf{ALL}^T\mathbf{A}^T = \mathbf{A}\Sigma\mathbf{A}^T - (\mathbf{Au})(\mathbf{Au})^T$. □

Note that by letting each row of $\mathbf{A}$ contain at most one element equal to 1 and all other elements equal to 0, it follows from Lemma 2 that the marginals of an asymmetric mixed normal vector are themselves asymmetric mixed normal.

Lemma 3. *Consider* $\mathbf{X} \sim ANM_n(G, \mathbf{\Sigma}, \mathbf{u})$. *The joint distribution of* $\mathbf{X}$ *is given by*

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) = E\left[\Phi_n\left(a_1, a_2, \ldots, a_n; \mathbf{R}\right)\right], \tag{12}$$

where $\Phi_n(;\mathbf{R})$ *is an n-dimensional standard normal joint cdf with correlation matrix* $\mathbf{R} = \{r_{ij}\}$ *and*

$$\begin{aligned} a_i &= \frac{x_i - \gamma^{-1}(H-1)u_i}{\sqrt{H}\sqrt{\sigma_{ii} - u_i^2}} \\ r_{ij} &= \frac{\sigma_{ij} - u_i u_j}{\sqrt{(\sigma_{ii} - u_i^2)(\sigma_{jj} - u_j^2)}} \end{aligned} \tag{13}$$

Proof. For simplicity the proof is carried out for $n = 2$, proof for $n > 2$ being identical. It is:

$$\begin{aligned} &\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) \\ &= \mathbb{P}(\gamma^{-1}(H-1)u_1 + \sqrt{H}\textstyle\sum_{j=1}^2 l_{1j}Z_j \leq x_1, \gamma^{-1}(H-1)u_2 + \sqrt{H}\sum_{j=1}^2 l_{2j}Z_j \leq x_2 \\ &= E\left[\mathbb{P}\left(\textstyle\sum_{j=1}^2 l_{1j}Z_j \leq \frac{x_1 - \gamma^{-1}(H-1)u_1}{\sqrt{H}}, \sum_{j=1}^2 l_{2j}Z_j \leq \frac{x_2 - \gamma^{-1}(H-1)u_2}{\sqrt{H}} |H\right)\right] \end{aligned} \tag{14}$$

We know that

$$\begin{aligned} E[\textstyle\sum_{j=1}^2 l_{ij}Z_j] &= 0 \\ Var(\textstyle\sum_{j=1}^2 l_{ij}Z_j) &= \textstyle\sum_{j=1}^2 l_{ij}^2 = \sigma_{ii} - u_i^2 \\ Cov(\textstyle\sum_{j=1}^2 l_{1j}Z_j), \sum_{k=1}^2 l_{2k}Z_k)) &= \textstyle\sum_{j=1}^2\sum_{k=1}^2 l_{1j}l_{2k}Cov(Z_j, Z_k) \\ &= \textstyle\sum_{j=1}^2 l_{1j}l_{2j} = \sigma_{12} - u_1u_2 \end{aligned} \tag{15}$$

Consequently, since H is independent of $\mathbf{Z}$,

$$\begin{aligned} &\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) \\ &= E\left[\Phi_2\left(\frac{x_1 - \gamma^{-1}(H-1)u_1}{\sqrt{H}\sqrt{\sigma_{11} - u_1^2}}, \frac{x_2 - \gamma^{-1}(H-1)u_2}{\sqrt{H}\sqrt{\sigma_{22} - u_2^2}}; \frac{\sigma_{12} - u_1u_2}{\sqrt{(\sigma_{11} - u_1^2)(\sigma_{22} - u_2^2)}}\right)\right] \end{aligned} \tag{16}$$

□

Finally the restriction on the non-centrality vector is justified.

Lemma 4. *The condition* $\mathbf{u}^T \cdot \mathbf{\Sigma}^{-1} \cdot \mathbf{u} \leq 1$ *is sufficient for the decomposition* $\mathbf{L} \cdot \mathbf{L}^T = \mathbf{\Sigma} - \mathbf{u} \cdot \mathbf{u}^T$ *to exist.*

Proof. Let $\mathbf{\Sigma} = \mathbf{D} \cdot \mathbf{D}^T$, for $\mathbf{D} \in \mathbb{R}^{n \times n}$. Such $\mathbf{D}$ will always exist because $\mathbf{\Sigma}$ is positive definite and symmetric. Now consider matrix

$$\mathbf{L} = \mathbf{D} - \frac{\mathbf{u}\mathbf{u}^T(\mathbf{D}^{-1})^T}{1 + \sqrt{1 - \mathbf{u}^T\mathbf{\Sigma}^{-1}\mathbf{u}}} \tag{17}$$

It can be checked by direct calculation that $\mathbf{L}\mathbf{L}^T = \mathbf{\Sigma} - \mathbf{u}\mathbf{u}^T$. The existence of $\mathbf{L}$ is guaranteed by $\mathbf{u}^T\mathbf{\Sigma}^{-1}\mathbf{u} \leq 1$. □

2.2 Numerical illustration

In figures 1.-3. the dependence patterns induced by asymmetric normal mixtures are demonstrated. In figure 1. a scatter plot of 5000 samples from a standard bivariate normally distributed vector with $\sigma_{12} = 0.5$ is shown.

The corresponding sample ranks with could be viewed as a sample from the underlying copula are also shown.

In figure 2. the corresponding plots are given for a standard bivariate t distribution, with $\sigma_{12} = 0.5$ $d = 5$ degrees of freedom. This distribution can be constructed as an asymmetric normal mixture with $\mathbf{u} = \mathbf{0}$ and inverse Gamma distributed mixing variable $H \sim 1/Gamma\left(\frac{d}{2}, \frac{2}{d-2}\right)$. For this distribution it is $\gamma = \sigma(H) = \frac{1}{\sqrt{2d-2}}$. It can seen how the normal mixture introduces a higher dependence in the tails, while maintaining the radial symmetry of the normal distribution.

In figure 3. samples from an asymmetric generalisation of the t distribution are plotted. This is effected by taking H as before and letting $u_1 = u_2 = 0.7$. The effect of the non centrality vector u on the dependence structure is clearly seen, as positive dependence becomes concentrated in the top-right area of the distribution.

3 Dependence properties

3.1 Stochastic orders

Here the dependence properties of asymmetric normal mixtures are studied in some more depth. We start with the definition of the stochastic concordance and supermodular orders, which provides a much stronger characterisation of dependence than correlation. The presentation of standard results (with no regard to full generality of those results) is based upon Müller and Stoyan (2002).

Definition 2. *Consider random vectors* (X_1, X_2) *and* (Y_1, Y_2), *such that* $X_1 \stackrel{d}{=} Y_1$, $X_2 \stackrel{d}{=} Y_2$. *Then we say that* $\mathbf{X}$ *precedes* $\mathbf{Y}$ *in the concordance order and write* $\mathbf{X} \preceq_c \mathbf{Y}$, *if either of the following two equivalent conditions holds:*

i) $\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) \leq \mathbb{P}(Y_1 \leq x_1, Y_2 \leq x_2)$ *for all* x_1, x_2.

ii) $Cov(g(X_1), h(X_2)) \leq Cov(g(Y_1), h(Y_2))$ *for all increasing functions* g, h *such that the covariance exists.*

It is apparent from Definition 2 that concordance order is a property of the copulas of the random vectors $\mathbf{X}$ and $\mathbf{Y}$ and does not depend on the marginal distributions. The importance of concordance order in risk management is related to a result by Dhaene and Goovaerts (1996), which shows that among portfolios whose respective elements are equal in distribution, the more concordant portfolio is also the riskiest one in the stop-loss and convex order senses. It is a desirable property for a multivariate model that an increase in correlation makes the random vector more concordant, as this ensures the intuitively appealing property that higher correlations produce higher aggregate risk holds true. This is also practically relevant when sensitivity testing the implementation of a dependence model. If aggregate risk, defined as the sum of increasing functions of the elements of random vector $\mathbf{X}$, is measured by a risk measure that is consistent with the stop-loss order (see e.g. Denuit et al (2005) for a discussion of the relation between risk measures and stochastic orders), then an increase in input correlation should also yield a increase in the aggregate risk.

The above discussion generalises to the case of dimensions higher than 2, via the concept of supermodular order:

Definition 3. *Consider random n-vectors* $\mathbf{X}$ *and* $\mathbf{Y}$*, Then we say that* $\mathbf{X}$ *precedes* $\mathbf{Y}$ *in the supermodular order and write* $\mathbf{X} \preceq_{sm} \mathbf{Y}$*, if*

$$E[f(\mathbf{X})] \leq E[f(\mathbf{Y})], \tag{18}$$

for all supermodular functions f *such that the expectations exist.*

We need not assume assume equality of marginal distributions in the definition above, as such equality is actually a consequence of the supermodular order. Moreover it can be shown that the supermodular order generalises the concordance one, as formally stated in the next lemma (Müller and Stoyann, 2002, Th. 3.9.5).

Lemma 5. *Let* $\mathbf{X} \preceq_{sm} \mathbf{Y}$*. Then*

- $X_i \stackrel{d}{=} Y_i,\ i = 1, \ldots, n$
- $(X_i, X_j) \preceq_c (Y_i, Y_j),\ \forall i, j.$

The relationship between the supermodular order on random vectors and the stop-loss order on the sum of their elements is given by Theorem 8.3.3 in Müller and Stoyan (2002), essentially generalising Dhaene and Goovaerts (1996).

Is is now seen that an increase in the correlations between elements of an asymmetric mixed normal vector, makes the vector more dependent in the supermodular order sense.

Lemma 6. *Consider* $\mathbf{X} \sim ANM_n(G, \mathbf{\Sigma}, \mathbf{u})$ *and* $\mathbf{X}' \sim ANM_n(G, \mathbf{\Sigma}', \mathbf{u})$*, such that* $\sigma_{ii} = \sigma'_{ii}, \sigma_{ij} \leq \sigma'_{ij} \forall i, j$*. Then* $\mathbf{X} \preceq_{sm} \mathbf{X}'$*.*

Proof. We need to show that for all supermodular functions $E[f(\mathbf{X})] \leq E[f(\mathbf{X}')]$ holds. For $H, H' \sim G$ the mixing variables corresponding to $\mathbf{X}, \mathbf{X}'$, it then enough to show that $E[f(\mathbf{X})|H = h] \leq E[f(\mathbf{X}')|H' = h]$. We observe that $\mathbf{X}|H = h,\ \mathbf{X}'|H = h'$ are multivariate normal vectors with the identical marginal distributions and off-diagonal elements of the covariance matrix given by $c_{ij} = h(\sigma_{ij} - u_i u_j),\ c'_{ij} = h(\sigma'_{ij} - u_i u_j)$. Since $c_{ij} \leq c'_{ij}$ the Lemma is proved by Theorem 3.13.5 in Müller and Stoyan (2002). □

3.2 Rank correlations

While the asymmetric mixed normal model can be used to model quantities of interest in risk management, e.g. asset log-returns, it can also be used to model risk with arbitrary marginal distributions, by considering only the copula of an asymmetric mixed normal vector $\mathbf{X}$. Hence, beyond the Pearson correlation coefficient (and its generalisation via the matrix $\mathbf{\Sigma}$), expressions for the rank correlation coefficients by Spearman and Kendall are necessary. Rank correlations are of interest in the copula context because they are invariant to monotone transformations of the elements of $\mathbf{X}$.

Consider $\mathbf{X} \sim ANM_2(G, \mathbf{\Sigma}, \mathbf{u})$ and $\mathbf{X}', \mathbf{X}''$ independent copies of $\mathbf{X}$. The Spearman and Kendall rank correlation coefficients can then be defined via the equations (e.g. Nelsen, 1999):

$$\rho_s(X_1, X_2) = 12\mathbb{P}\left(X_1 \leq X_1', X_2 \leq X_2''\right) - 3 \tag{19}$$

and

$$\rho_\tau(X_1, X_2) = 4\mathbb{P}(X_1 \leq X_1', X_2 \leq X_2') - 1 \tag{20}$$

respectively.

Expressions for the rank correlation coefficients are given below.

Lemma 7. *For* $\mathbf{X} \sim ANM_2(G, \mathbf{\Sigma}, \mathbf{u})$ *Spearman's rank correlation coefficient is given by*

$$\rho_s(X_1, X_2) =$$
$$12E\left[\Phi_2\left(\frac{\gamma^{-1}u_1(H-H')}{\sqrt{(\sigma_{11}-u_1^2)(H+H')}}, \frac{\gamma^{-1}u_2(H-H'')}{\sqrt{(\sigma_{22}-u_2^2)(H+H'')}}; \frac{(\sigma_{12}-u_1u_2)H}{\sqrt{(\sigma_{11}-u_1^2)(\sigma_{22}-u_2^2)(H+H')(H+H'')}}\right)\right] - 4, \tag{21}$$

where $H, H', H'' \sim G$ *are independent copies of the mixing variable.*

Proof. Denote by H, H', H'' and $\mathbf{Z}, \mathbf{Z}', \mathbf{Z}''$ the mixing and normal variable corresponding to the independent pairs $\mathbf{X}, \mathbf{X}', \mathbf{X}'' \sim ANM_2(G, \mathbf{\Sigma}, \mathbf{u})$. Then we have:

$$\begin{aligned}
&\mathbb{P}\left(X_1 \leq X_1', X_2 \leq X_2''\right) = E\mathbb{P}\left(X_1 \leq X_1', X_2 \leq X_2'' | H, H', H''\right) = \\
&E\mathbb{P}\left(\begin{matrix} \gamma^{-1}(H-1)u_1 + \sqrt{H}(l_{11}Z_1 + l_{12}Z_2) \leq \\ \gamma^{-1}(H'-1)u_1 + \sqrt{H'}(l_{11}Z_1' + l_{12}Z_2'), \\ \gamma^{-1}(H-1)u_2 + \sqrt{H}(l_{21}Z_1 + l_{22}Z_2) \leq \\ \gamma^{-1}(H''-1)u_2 + \sqrt{H''}(l_{21}Z_1'' + l_{22}Z_2'') \end{matrix} \middle| H, H', H'' \right) = \\
&E\mathbb{P}\left(\begin{matrix} \sqrt{H}(l_{11}Z_1 + l_{12}Z_2) - \sqrt{H'}(l_{11}Z_1' + l_{12}Z_2') \leq \\ \gamma^{-1}(H'-H)u_1, \\ \sqrt{H}(l_{21}Z_1 + l_{22}Z_2) - \sqrt{H''}(l_{21}Z_1'' + l_{22}Z_2'') \leq \\ \gamma^{-1}(H''-H)u_2 \end{matrix} \middle| H, H', H'' \right)
\end{aligned} \tag{22}$$

For fixed H, H', H'', the joint distribution of

$$\begin{aligned} Y_1 &= \sqrt{H}(l_{11}Z_1 + l_{12}Z_2) - \sqrt{H'}(l_{11}Z_1' + l_{12}Z_2') \\ Y_2 &= \sqrt{H}(l_{21}Z_1 + l_{22}Z_2) - \sqrt{H''}(l_{21}Z_1'' + l_{22}Z_2'') \end{aligned} \tag{23}$$

is bivariate normal with mean vector $[0\ 0]^T$ and covariance matrix:

$$\begin{pmatrix} (\sigma_{11}-u_1^2)(H+H') & (\sigma_{12}-u_1u_2)H \\ (\sigma_{12}-u_1u_2)H & (\sigma_{22}-u_2^2)(H+H'') \end{pmatrix} \tag{24}$$

Hence

$$P\left(X_1 \le X_1', X_2 \le X_2''|H,H',H''\right) =$$
$$\Phi_2\left(\frac{\gamma^{-1}u_1(H-H')}{\sqrt{(\sigma_{11}-u_1^2)(H+H')}}, \frac{\gamma^{-1}u_2(H-H'')}{\sqrt{(\sigma_{22}-u_2^2)(H+H'')}}; \frac{(\sigma_{12}-u_1u_2)H}{\sqrt{(\sigma_{11}-u_1^2)(\sigma_{22}-u_2^2)(H+H')(H+H'')}}\right), \tag{25}$$

which by equation (19) completes the proof. ☐

Lemma 8. *For* $\mathbf{X} \sim ANM_2(G,\mathbf{\Sigma},\mathbf{u})$ *Kendall's rank correlation coefficient is given by*

$$\rho_\tau(X_1,X_2) =$$
$$4E\left[\Phi_2\left(\frac{\gamma^{-1}u_1(H-H')}{\sqrt{(\sigma_{11}-u_1^2)(H+H')}}, \frac{\gamma^{-1}u_2(H-H')}{\sqrt{(\sigma_{22}-u_2^2)(H+H')}}; \frac{(\sigma_{12}-u_1u_2)}{\sqrt{(\sigma_{11}-u_1^2)(\sigma_{22}-u_2^2)}}\right)\right]-1, \tag{26}$$

where $H, H' \sim G$ *are independent copies of the mixing variable.*

Proof. The proof is near identical to that of the previous lemma. Denote by H, H' and $\mathbf{Z}, \mathbf{Z}'$ the mixing and normal variable corresponding to the independent pairs $\mathbf{X}, \mathbf{X}' \sim ANM_2(G,\mathbf{\Sigma},\mathbf{u})$. Then we have:

$$\begin{aligned}
&\mathbb{P}\left(X_1 \le X_1', X_2 \le X_2'\right) = E\mathbb{P}\left(X_1 \le X_1', X_2 \le X_2'|H,H'\right) = \\
&E\mathbb{P}\left(\begin{array}{c} \gamma^{-1}(H-1)u_1 + \sqrt{H}(l_{11}Z_1+l_{12}Z_2) \le \\ \gamma^{-1}(H'-1)u_1 + \sqrt{H'}(l_{11}Z_1'+l_{12}Z_2'), \\ \gamma^{-1}(H-1)u_2 + \sqrt{H}(l_{21}Z_1+l_{22}Z_2) \le \\ \gamma^{-1}(H'-1)u_2 + \sqrt{H'}(l_{21}Z_1'+l_{22}Z_2') \end{array}\middle| H,H'\right) = \\
&E\mathbb{P}\left(\begin{array}{c} \sqrt{H}(l_{11}Z_1+l_{12}Z_2) - \sqrt{H'}(l_{11}Z_1'+l_{12}Z_2') \le \\ \gamma^{-1}(H'-H)u_1, \\ \sqrt{H}(l_{21}Z_1+l_{22}Z_2) - \sqrt{H'}(l_{21}Z_1'+l_{22}Z_2') \le \\ \gamma^{-1}(H'-H)u_2 \end{array}\middle| H,H'\right)
\end{aligned} \tag{27}$$

For fixed H, H', the joint distribution of

$$\begin{aligned} Y_1 &= \sqrt{H}(l_{11}Z_1+l_{12}Z_2) - \sqrt{H'}(l_{11}Z_1'+l_{12}Z_2') \\ Y_2 &= \sqrt{H}(l_{21}Z_1+l_{22}Z_2) - \sqrt{H'}(l_{21}Z_1'+l_{22}Z_2') \end{aligned} \tag{28}$$

is bivariate normal with mean vector $[0\ 0]^T$ and covariance matrix:

$$\begin{pmatrix} (\sigma_{11}-u_1^2)(H+H') & (\sigma_{12}-u_1u_2)(H+H') \\ (\sigma_{12}-u_1u_2)(H+H') & (\sigma_{22}-u_2^2)(H+H') \end{pmatrix} \tag{29}$$

Hence

$$P\left(X_1 \le X_1', X_2 \le X_2'|H,H'\right) =$$
$$\Phi_2\left(\frac{\gamma^{-1}u_1(H-H')}{\sqrt{(\sigma_{11}-u_1^2)(H+H')}}, \frac{\gamma^{-1}u_2(H-H')}{\sqrt{(\sigma_{22}-u_2^2)(H+H')}}; \frac{(\sigma_{12}-u_1u_2)}{\sqrt{(\sigma_{11}-u_1^2)(\sigma_{22}-u_2^2)}}\right), \tag{30}$$

which by equation (20) completes the proof. ☐

4 Parameterisation issues

4.1 Rank correlations

One can calculate Spearman's and Kendall's rank correlation for a particular choice of $G, \mathbf{\Sigma}, \mathbf{u}$ by equations (21) and (26) respectively, by numerical integration. There are however two practical issues that such a process will not address:

- A closed form formula for rank correlations may be more useful, e.g. for reasons of computational speed.
- When choosing parameters for a model where only the copula of an asymmetric normal mixture is of interest, one needs to work backwards from a specified set of rank correlation coefficients to the matrix $\mathbf{\Sigma}$ (say for a fixed non-centrality vector $\mathbf{u}$).

The approach taken here is to resolve these issues by deriving approximate formulas for the rank correlation coefficients. The approximation carried out by considering a discrete distribution G for the mixing variable H, defined on a finite number of points. This could be derived as an approximation to the generally continuous distribution G used in reality. The method is presented here only for the case of the Spearman rank correlation coefficient; the calculation for Kendall's rank correlation is very similar. Without loss of generality we assume that $\mathbf{\Sigma}$ is a correlation, rather then a covariance matrix.

Lemma 9. *Let $\mathbf{X} \sim ANM_2(G, \mathbf{\Sigma}, \mathbf{u})$, with $H \sim G$ such that $\mathbb{P}(H_i = h_i) = p_i$, $\sum_{j=1}^d p_j = 1$ for $h_1 < \cdots < h_d$ and $\sigma_{11} = \sigma_{22} = 1$. Then Spearman's rank correlation is given by:*

$$\rho_s(X_1, X_2) = 12 \left[\beta_0 + \sum_{m=1}^{\infty} \beta_m (\sigma_{12} - u_1 u_2)^m \right] - 3. \tag{31}$$

The coefficients β_m are given by

$$\begin{array}{rcl} \beta_0 & = & \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d p_i p_j p_k \Phi(a_1(h_i, h_j)) \Phi(a_1(h_i, h_k)) \\ \beta_m & = & \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \{ p_i p_j p_k \phi(a_1(h_i, h_j)) \phi(a_1(h_i, h_k)) \\ & & He_{m-1}(a_1(h_i, h_j)) He_{m-1}(a_2(h_i, h_k)) b(h_i, h_j, h_k)^m \}, \end{array} \tag{32}$$

where:

- $a_i(x, y) = \frac{\gamma^{-1} u_i (x-y)}{\sqrt{(1-u_i^2)(x+y)}}, \ i = 1, 2.$
- $b(x, y, z) = \frac{x}{\sqrt{(1-u_1^2)(1-u_2^2)(x+y)(x+z)}}$
- $He_k(x) = \sum_{i=0}^{[k/2]} \frac{k!}{i!(k-2i)!} (-1)^i 2^{-i} x^{k-2i}$ *are the Hermitian polynomials.*

- *Φ and ϕ are the standard normal cumulative distribution and density respectively.*

Proof. First note that for mixing variable H of the form considered here, equation (21) becomes

$$\rho_s(X_1, X_2) = 12\sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d p_i p_j p_k \;\Phi_2\left(a_1(h_i, h_j), a_2(h_i, h_k); (\sigma_{12} - u_1 u_2) b(h_i, h_j, h_k)\right) - 4. \tag{33}$$

Moreover, the bivariate standard normal distribution can be written as (Gupta, 1963)

$$\Phi_2(a_1, a_2; b) = \Phi(a_1)\Phi(a_2) + \phi(a_1)\phi(a_2)\sum_{m=1}^{\infty} \frac{1}{(m+1)!} He_m(a_1) He_m(a_2) b^{m+1} \tag{34}$$

The result follows directly from these two equations. □

Given Lemma 9, it is now possible to approximately calculate σ_{12} from $\rho_s = \rho_s(X_1, X_2)$ by performing a series reversion. In particular we can write:

$$\sigma_{12} \cong u_1 u_2 + \sum_{m=1}^{7} \delta_m \left(\frac{\rho_s + 3}{12} - \beta_0\right)^m, \tag{35}$$

where the coefficients δ_m of the reversed series can be calculated from those of the original series β_m by Abramowitz and Stegun (1972, p.16).

4.2 Choice of non-centrality vector

As discussed in Section 2, the non-centrality vector $\mathbf{u}$ is used in order to skew the dependence structure of a normal mixture. So choosing high elements of $\mathbf{u}$ will yield a very skew copula. The extent to which this can be carried out is nonetheless limited as very high values of in $\mathbf{u}$ would violate the constraint $\mathbf{u}^T \cdot \mathbf{\Sigma}^{-1} \cdot \mathbf{u} < 1$.

It is therefore of interest to ask: "given a correlation matrix $\mathbf{\Sigma}$, what is the largest $\mathbf{u}$ that one could use?". Answering this is our aim in this section. We note that using "the largest possible $\mathbf{u}$" is a decision by the modeler as to how the dependence structure should look like and has nothing to do in this context with statistical estimation of the $\mathbf{u}$ parameter vector. The respective choice of non-centrality vector is given in the following result:

Lemma 10. *The vector $\mathbf{u}$ for which the sum $\sum_{j=1}^n u_j^2$ is largest and $\mathbf{u}^T \cdot \mathbf{\Sigma}^{-1} \cdot \mathbf{u} < 1$, is proportional to the eigenvector of $\mathbf{\Sigma}$ corresponding to its largest eigenvalue.*

Proof. Let $\mathbf{S} = \mathbf{\Sigma}^{-1}$. As $\mathbf{S}$ is the inverse of a symmetric matrix, it is itself symmetric. Hence $\mathbf{u}^T \cdot S \cdot \mathbf{u} < 1 = \sum_{i=1}^n \sum_{j=1}^n s_{ij} u_i u_j$.

Consider now the following optimisation problem:

$$\max_{u_1,\ldots,u_n} \sum_{i=1}^{n} u_i^2, \qquad \text{such that:} \quad \sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij}u_iu_j = a, \tag{36}$$

where $a < 1$, e.g. $a = 0.99$. The corresponding Lagrangian is

$$\mathcal{L}(\mathbf{u}, \lambda) = \sum_{i=1}^{n} u_i^2 + \lambda \left(a - \sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij}u_iu_j \right) \tag{37}$$

It then is:

$$\begin{aligned}
\frac{\partial \mathcal{L}(\mathbf{u},\lambda)}{\partial u_k} &= 2u_k - \frac{\partial}{\partial u_k}\lambda \left(u_1 \sum_{j=1}^{n} s_{1j}u_j + \cdots + u_n \sum_{j=1}^{n} s_{nj}u_j \right) \\
&= 2u_k - \lambda \Big(u_1 \frac{\partial}{\partial u_k} \sum_{j=1}^{n} s_{1j}u_j + \cdots + \sum_{j=1}^{n} s_{kj}u_j + u_k \frac{\partial}{\partial u_k} \sum_{j=1}^{n} s_{kj}u_j \\
&\qquad + \ldots \frac{\partial}{\partial u_k} u_n \sum_{j=1}^{n} s_{nj}u_j \Big) \\
&= 2u_k - \lambda \left(\sum_{j=1}^{n} s_{kj}u_j + u_1 s_{1k} + \cdots + u_n s_{kn} \right) \\
&= 2u_k - 2\lambda \sum_{j=1}^{n} s_{kj}
\end{aligned} \tag{38}$$

Setting $\frac{\partial \mathcal{L}(\mathbf{u},\lambda)}{\partial u_k} = 0$ yields:

$$u_k = \lambda \sum_{j=1}^{n} s_{kj} \implies \mathbf{u} = \lambda \cdot \mathbf{S} \cdot \mathbf{u} \tag{39}$$

Hence the reciprocal of the Lagrange multiplier λ is an eigenvalue of $\mathbf{S}$ and the non-centrality vector $\mathbf{u}$ is the corresponding eigenvector. Consequently, λ is an eigenvalue of $\mathbf{\Sigma} = \mathbf{S}^{-1}$ with corresponding eigenvector $\mathbf{u}$.

Now, to determine which exactly of $\mathbf{\Sigma}$'s eigenvectors $\mathbf{u}$ corresponds to, consider

$$u_k = \lambda \sum_{j=1}^{n} s_{kj} \implies \sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n} \left\{ \lambda^2 \left(\sum_{j=1}^{n} s_{ij} \right)^2 \right\} \tag{40}$$

It is

$$\begin{aligned}
\left(\sum_{j=1}^{n} s_{ij} \right)^2 &= \sum_{j=1}^{n} \sum_{r=1}^{n} (s_{ir}u_r)(s_{ij}u_j) \\
&\overset{(39)}{=} \sum_{j=1}^{n} (s_{ij}u_j) \cdot \frac{1}{\lambda} u_i
\end{aligned} \tag{41}$$

Thus

$$\begin{aligned}
\sum_{i=1}^{n} u_i^2 &= \sum_{i=1}^{n} \lambda^2 \cdot \sum_{j=1}^{n} (s_{ij}u_j) \cdot \frac{1}{\lambda} u_i \\
&= \lambda \cdot \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij}u_ju_i \\
&= \lambda \cdot a
\end{aligned} \tag{42}$$

Therefore the largest value of the sum $\sum_{i=1}^{n} = u_i^2$ is achieved when $\mathbf{u}$ is an eigenvector of $\mathbf{\Sigma}$ corresponding to its largest eigenvalue. □

5 Kronecker products in high-dimensional simulation

5.1 Properties of the Kronecker product

Simulation algorithms based on the multivariate normal distribution, such as the asymmetric mixed normal model (3) are well suited for simulation in high dimensions, e.g. $n = 50$. However, a portfolio of insurance risks will sometimes be of much higher dimension. Consider for example the case of a large insurance company exposed to 50 lines of business, underwritten in 15 years over 3 territories. This immediately produces 2250 potentially dependent random variables. At such high dimension a number of problems occur:

- It becomes very difficult to specify a positive definite correlation matrix $\mathbf{\Sigma}$.
- The Cholesky decomposition algorithm used to factorise the matrix $\mathbf{\Sigma} - \mathbf{u} \cdot \mathbf{u}^T$ may fail because of numerical errors.
- Runtimes may become impracticably long, particularly for the matrix multiplication $\mathbf{L} \cdot \mathbf{Z}$.

A means to addressing these problems is to construct the matrix $\mathbf{\Sigma}$ using Kronecker products. Consider square matrices $\mathbf{A} = \{a_{ij}\} \in \mathbb{R}^{m \times m}, \mathbf{B} = \{b_{ij}\} \in \mathbb{R}^{n \times n}$. Then their Kronecker product $\mathbf{A} \otimes \mathbf{B}$ is a $mn \times mn$ matrix such that

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1m}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2m}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mm}\mathbf{B} \end{pmatrix} \tag{43}$$

Kronecker products have a number of useful properties (e.g. Van Loan, 2000) of which we note here:

1. $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$
2. $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$
3. $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$
4. If $\mathbf{A}, \mathbf{B}$ are positive definite matrices, then $\mathbf{A} \otimes \mathbf{B}$ is positive definite.
5. If $\mathbf{A} = \mathbf{M} \cdot \mathbf{M}^T$ and $\mathbf{B} = \mathbf{N} \cdot \mathbf{N}^T$, then $\mathbf{A} \otimes \mathbf{B} = (\mathbf{M} \otimes \mathbf{N}) \cdot (\mathbf{M} \otimes \mathbf{N})^T$.

Property 4. ensures that a positive definite correlation (covariance) matrix $\mathbf{\Sigma}$ can be constructed as the Kronecker product of 2 (or more) smaller correlation (covariance) matrices $\mathbf{A}, \mathbf{B}$. Property 5. breaks down the problem of decomposing a large correlation matrix to that of decomposing two smaller ones, thus reducing the potential for numerical error.

Consider now the matrix multiplication $(\mathbf{A}\otimes\mathbf{B})\cdot\mathbf{Z}$, where $\mathbf{A} \in \mathbb{R}^{m\times m}, \mathbf{B} \in \mathbb{R}^{n\times n}$ and $\mathbf{Z} \in \mathbb{R}^{mn\times 1}$. We can then write:

$$(\mathbf{A}\otimes\mathbf{B})\cdot\mathbf{Z} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1m}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2m}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mm}\mathbf{B} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \\ \vdots \\ \mathbf{Z}_m \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1j}\mathbf{B}\cdot Z_j \\ \sum_{j=1}^m a_{2j}\mathbf{B}\cdot Z_j \\ \vdots \\ \sum_{j=1}^m a_{mj}\mathbf{B}\cdot Z_j \end{pmatrix}, \tag{44}$$

where $\mathbf{Z}_j \in \mathbb{R}^{n\times 1},\ j = 1, \ldots, n.$. It is apparent that the matrix products $\mathbf{B} \cdot \mathbf{Z}_j$ are repeated in each block row of the matrix above. Hence they can be calculated in advance and reused as appropriate. Multiplication of an $mn \times mn$ matrix by an $mn \times 1$ vector generally requires $2(mn)^2$ elementary operations (additions and multiplications). However, if the $mn \times mn$ matrix can be represented by a Kronecker product as above, the computational workload drops to $2mn^2 + 2m^2n$ operations. If for example $m = n = 50$, this implies approximately a 25-fold reduction in the number of elementary operations required.

5.2 Kronecker products in the asymmetric mixed normal model

Here is shown how a high-dimensional version of the asymmetric mixed normal model can be constructed with the use of Kronecker products.

Lemma 11. *Consider* $\mathbf{X} \sim ANM_{mn}(G, \mathbf{\Sigma}, \mathbf{u})$*, such that*

- $\mathbf{\Sigma} = \mathbf{A} \otimes \mathbf{B}$ *for* $\mathbf{A} \in \mathbb{R}^{m\times m}, \mathbf{B} \in \mathbb{R}^{n\times n}$*, with* $\mathbf{A} = \mathbf{M}\mathbf{M}^T$*,* $\mathbf{B} = \mathbf{N}\mathbf{N}^T$
- $\mathbf{u} = \mathbf{v} \otimes \mathbf{w}$ *for* $\mathbf{v} \in \mathbb{R}^{m\times 1}, \mathbf{w} \in \mathbb{R}^{n\times 1}$

Then $\mathbf{X}$ *can be written as*

$$\mathbf{X} \overset{d}{=} \gamma^{-1}(H-1)(\mathbf{v}\otimes\mathbf{w}) + \sqrt{H}\cdot(\mathbf{M}\otimes\mathbf{N})\cdot\mathbf{Z} - \frac{\sqrt{H}}{1+\sqrt{1-(\mathbf{v}^T\mathbf{A}^{-1}\mathbf{v})\otimes(\mathbf{w}^T\mathbf{B}^{-1}\mathbf{w})}} \cdot \left((\mathbf{v}\mathbf{v}^T(\mathbf{M}^{-1})^T)\otimes(\mathbf{w}\mathbf{w}^T(\mathbf{N}^{-1})^T)\right)\cdot\mathbf{Z}, \tag{45}$$

where

- $\mathbf{Z}$ *is an* (mn)*-vector of independent standard normal variables.*
- H *is a non-negative random variable, independent of* $\mathbf{Z}$*, following cumulative distribution* G *with mean 1 and standard deviation* γ*.*

Proof. Follows directly from equations (3), (17) and the properties of the Kronecker product discussed in the previous section. □

Hence by constructing $\mathbf{\Sigma} = \mathbf{A} \otimes \mathbf{B}$, the problems of specifying a large positive definite correlation matrix and decomposing that matrix are addressed. Moreover, it can be seen from equation (45) that the matrix multiplication $\mathbf{L} \cdot \mathbf{Z}$ is broken down to a difference of two matrix multiplications, in each which the first factor can be expressed as a Kronecker product. Hence, the computation of $\mathbf{L} \cdot \mathbf{Z}$ can be substantially speeded up by using representation (45). We note that the preceding discussion easily generalises to the case of $\mathbf{\Sigma}$ expressed as a Kronecker product of more than two matrices.

Imposing a Kronecker-product structure on of $\mathbf{\Sigma}$ forms quite a strong assumption, so it is fair to ask whether such a specification makes sense. Consider the example of an insurance company that has exposures in m lines of business, n years, and r territories. Specify correlation matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, with dimensions $m \times m$, $n \times n$, $r \times r$ respectively. Interpret $\mathbf{A}$ as the correlation matrix between lines written in the same year and in the same territory, $\mathbf{B}$ as the correlation between the same line, written in the same territory over different years etc. Defining $\mathbf{\Sigma} = \mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C}$ produces an overall correlation matrix that is consistent with the above specification, with the Kronecker structure producing the cross-correlations between risks in different lines and different years or territories. As these cross-correlations emerge as products of correlation coefficients with modulus < 1, it is ensured that they are smaller than the corresponding correlations between lines within the same year and territory - hence a first reasonableness check is passed. There is of course no particular reason why the cross-correlations should have the prescribed form. Nonetheless, given that it would be very unlikely that an insurance company has enough data to statistically estimate a, say, 2000x2000 correlation matrix, this choice of correlation matrix structure seems to be an acceptable compromise.

References

Abramowitz, M. and Stegun, I. E. (1972). *Handbook of mathematical functions.* Dover, New York.

Barndorff-Nielsen O. E. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics* **5** 151-157.

Demarta, S. and McNeil, A. J. (2005). The t Copula and Related Copulas. *International Statistical Review* **73** (1), 111-129.

Denuit, M., Dhaene, J., Goovaerts, M.J., and Kaas, R. (2005). *Actuarial Theory for Dependent Risk: Measures, Orders and Models.* Wiley, New York

Dhaene, J., and Goovaerts (1996), M. J. Dependency of risks and stop-loss order. *ASTIN Bulletin* **26** (2), 201-212.

Embrechts, P., McNeil, A., and Straumann, D. (2002). Correlation and Dependence in Risk Management: Properties and Pitfalls. In *Risk Management: Value at Risk and Beyond,*, ed. M.A.H. Dempster, 176-223. Cambridge University Press.

Fang, K.T., Kotz, S., and Ng, K.W. (1987). *Symmetric Multivariate and Related Distributions.* Chapman & Hall, London.

Hult, H. and Lindskog, F. (2001). Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied Probability* **34** (3), 587-608.

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. *North American Actuarial Journal* **2** (1), 1-26.

Gupta, S. S. (1963). Probability integrals of multivariate normal and multivariate t. *The Annals of Mathematical Statistics* **34** 3, 792-828.

McNeil, A. J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press.

Müller, A. and D. Stoyan (2002). *Comparison Methods for Stochastic Models and Risks.* Wiley, Chichester.

Nelsen, R. B. (1999). *An Introduction to Copulas.* Springer, New York.

Smith, A. D. (2002). Dependent tails. *2002 GIRO Convention*, 8-11 October 2002, Disneyland Resort Paris.

van Loan, C. F. (2000). The ubiquitous Kronecker product. *Journal of Computational and Applied Mathematics,* **123** (1-2), 85-100.

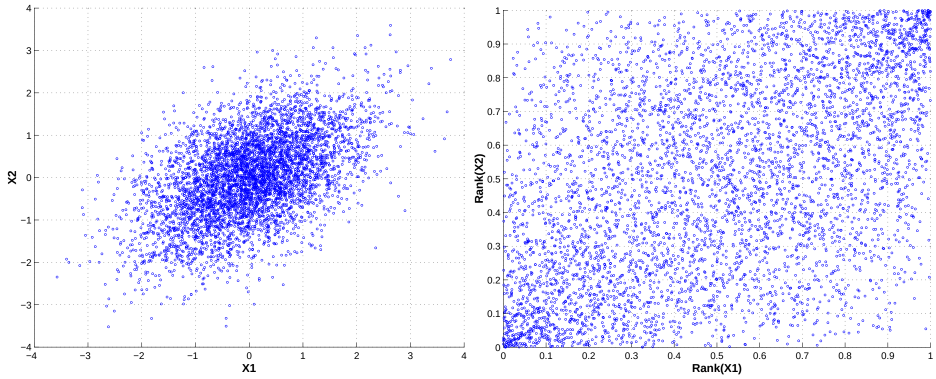

Figure 1: Samples and sample ranks from standard bivariate normal vector, $\sigma_{12} = 0.5$.

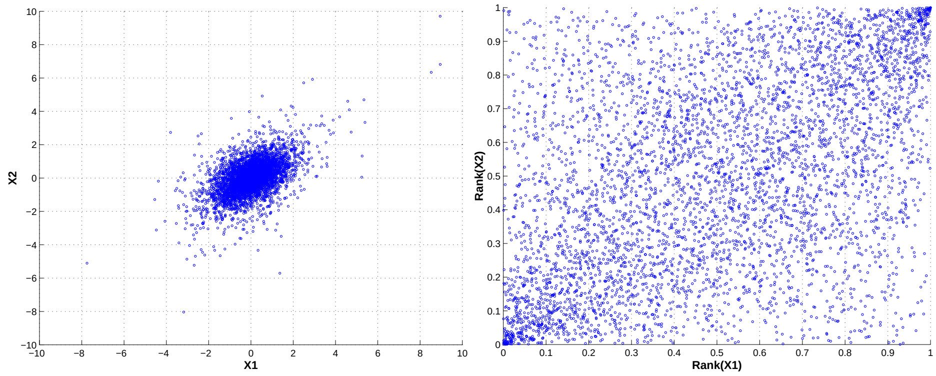

Figure 2: Samples and sample ranks from standard bivariate t vector, $\sigma_{12} = 0.5$, $\gamma = 8^{-0.5}(d = 5)$.

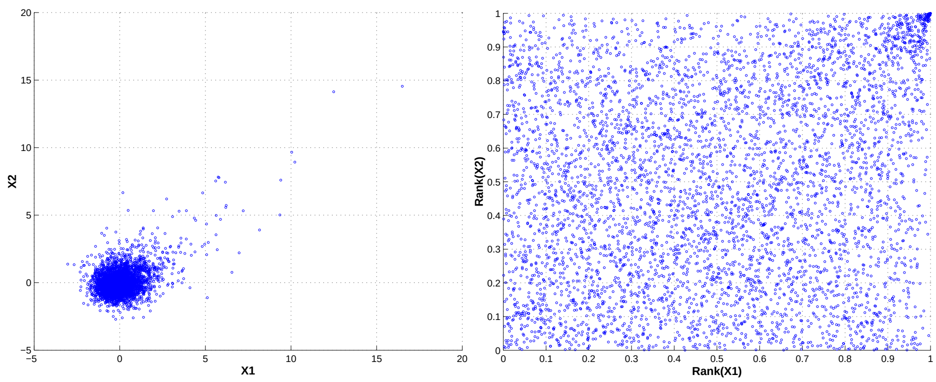

Figure 3: Samples and sample ranks from standard bivariate asymmetric t vector, $\sigma_{12} = 0.5$, $\gamma = 8^{-0.5}(d = 5)$, $u_1 = u_2 = 0.7$.