Proceedings of the 2005 Winter Simulation Conference
M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds.

FAST SIMULATION FOR MULTIFACTOR PORTFOLIO CREDIT RISK IN THE t-COPULA MODEL

Wanmo Kang

New Product Research
Moody's KMV
99 Church St.
New York, NY 10007, U.S.A.

Perwez Shahabuddin

Department of IE & OR
Columbia University
New York, NY 10027, U.S.A.

ABSTRACT

We present an importance sampling procedure for the estimation of multifactor portfolio credit risk for the t-copula model, i.e, the case where the risk factors have the multivariate t distribution. We use a version of the multivariate t that can be expressed as a ratio of a multivariate normal and a scaled chi-square random variable. The procedure consists of two steps. First, using the large deviations result for the Gaussian model in Glasserman, Kang, and Shahabuddin (2005a), we devise and apply a change of measure to the chi-square random variable. Then, conditional on the chi-square random variable, we apply the importance sampling procedure developed for the Gaussian copula model in Glasserman, Kang, Shahabuddin (2005b). We support our importance sampling procedure by numerical examples.

1 INTRODUCTION

A number of recent papers address better empirical fits of observed data by t-copula. See, for example, Mashal and Zeevi (2002) and Breymann, Dias, and Embrechts (2003). One reason for this is the asymptotic dependence property of t-copula (see Embrechts, Lidskog, and McNeal 2001) which captures the extreme co-movements of financial assets. The multifactor t-copula model of credit risk attempts to capture this in the credit risk setting. In this model the latent variables corresponding to obligors (e.g., normalized asset values of obligors) or equivalently, the risk factors, are assumed to have the multivariate t distribution. This is in contrast to the Gaussian copula model where these are assumed to have the multivariate Gaussian distribution.

Currently, there are no closed form analytical results for both the Gaussian copula and t-copula models. Lucas, Klaassen, Spreij, and Straetmans (2001) present approximations for the Gaussian copula case. Approximations for the t-copula case are presented in Lucas, Klaassen, Spreij, and Straetmans (2003), Kuhn (2004), and Kostadinov (2005). Their approach applies extreme value theory to the loss distributions conditional on the common factors. During this step, the idiosyncratic risks fade away by the strong law of large numbers. Then they focus on the remaining randomness, the common factors, which has a fixed size regardless of the number of obligors. This extreme value theoretic approach is useful for coming up with tail approximations of loss distributions. However in both the Gaussian copula and t-copula case, no bounds are provided for the approximation errors. Hence Monte Carlo simulation constitutes a viable alternative for the estimation of credit risk.

Credit default events of obligors are rare and thus the probability of large losses in a portfolio of credits is usually small. Naive simulation is known to be inefficient for estimation of small probabilities and importance sampling (IS) is widely used to increase simulation efficiency. Glasserman and Li (2003) and Glasserman, Kang, Shahabuddin (2005b) (henceforth GKSb) present *asymptotically optimal* importance sampling changes of measure for the estimation of credit risk in the single and multifactor Gaussian copula models, respectively. However, unlike the Gaussian copula model, a dependence structure based on the t-copula incurs a problem in devising importance sampling (IS) changes of measure. The student t-distribution is heavy-tailed, and hence the moment generating function does not exist. Hence the usual approaches for devising IS changes of measure are not applicable here.

Unlike the earlier approximation work mentioned above, we avoid a direct approach to this problem. In particular, we exploit the *conditional Gaussian* property of the multivariate t distribution used in a version of the multifactor t-copula model. This multivariate t can be represented as a multivariate Gaussian random vector divided by the square root of a univariate, scaled, chi-square random variable (see, e.g., Embrechts, Lidskog, and McNeal 2001). Then the risk factors are normally distributed conditional on the chi-square random variable. Thus, given a sample of the chi-square random variable, we can apply the fast simulation algorithms developed for the Gaussian copula model in GKSb. The large deviations result for the Gaussian model in Glasserman, Kang and Shahabuddin (2005a) (henceforth GKSa) and concepts related to *zero-variance changes of measure*, guide the IS for the chi-square random variable.

2 MULTIFACTOR PORTFOLIO CREDIT RISK MODELS

We consider the distribution of losses from default over a fixed horizon. We are interested in the estimation of the probability that the credit loss of a portfolio exceeds a given threshold. The default of each obligor is triggered if a latent variable associated with the obligor exceeds a threshold determined from its marginal default probability. The latent variables consist of a linear combination of factor variables that represent idiosyncratic risk and common risks to all obligors. We use the following notation:

m = the number of obligors to which the portfolio is exposed;
Y_k = default indicator (= 1 for default, = 0 otherwise) for the k-th obligor;
p_k = marginal probability that the k-th obligor defaults;
c_k = loss resulting from default of the k-th obligor;
L_m = $c_1 Y_1 + \cdots + c_m Y_m$ = total loss from defaults.

We are interested in the estimation of $\mathsf{P}(L_m > x)$ for a given threshold x when the event $\{L_m > x\}$ is rare. (For easy reference, we refer to the event $\{L_m > x\}$ as a *large loss* event.) The loss c_k may be assumed to be stochastic. However, in this paper, for simplicity we will assume the c_k to be deterministic and refer the reader to GKSa and GKSb for approaches for the stochastic case. For the dependence structure among Y_k's, we consider the two copula models mentioned before – the Gaussian copula and the t-copula.

2.1 The Gaussian Copula Model

Under the Gaussian copula, the dependence among the default indicators Y_k is given by the following. Let Φ be the cumulative distribution function (cdf) of a standard normal random variable. We set $Y_k = \mathbf{1}\{X_k > x_k\}$ where $X_1, X_2, \ldots$ are *correlated* standard normal random variables and $x_k := \Phi^{-1}(1 - p_k)$, so that $\mathsf{P}(Y_k = 1) = p_k$. The X_k's are referred to as latent variables. Correlations between these latent variables determine the dependence among the default indicators. In practice, these correlations are often derived from correlations in asset values or equity returns.

We consider the *multifactor Gaussian copula model with a finite number of types*. By types, we mean groups of homogeneous obligors in their dependence structure, which will be characterized formally below: There are d factors and t types of obligors. $\{\mathcal{I}_1^{(m)}, \ldots, \mathcal{I}_t^{(m)}\}$ is a partition of the set of obligors $\{1, \ldots, m\}$ into types. If $k \in \mathcal{I}_j^{(m)}$, then the k-th obligor is of type j and its latent variable is given by

$$X_k = \mathbf{a}_j^\top \mathbf{Z} + b_j \ \varepsilon_k \tag{1}$$

where $\mathbf{a}_j \in \mathbb{R}^d$ with $0 < \|\mathbf{a}_j\| < 1$, $b_j = \sqrt{1 - \mathbf{a}_j^\top \mathbf{a}_j}$, $\mathbf{Z}$ is a d dimensional standard normal random vector, and ε_k's are i.i.d., standard normal random variables, independent of the $\mathbf{Z}$. The $\mathbf{Z}$ represents systemic risk and ε_k represents idiosyncratic risk of the k-th obligor. $\mathbf{a}_j$ is the vector of factor loading coefficients (of the common factors) of obligors belonging to type j; b_j is the factor loading of the idiosyncratic risk factor. The $b_j = \sqrt{1 - \mathbf{a}_j^\top \mathbf{a}_j}$ ensures that the X_k's are $N(0,1)$'s. Let $n_j^{(m)} = |\mathcal{I}_j^{(m)}|$ denote the number of obligors of type j. Note that given Z, the probability of default of obligor k is

$$p_k(\mathbf{Z}) = \Phi\left(\frac{\mathbf{a}_j^\top \mathbf{Z} - \Phi^{-1}(1 - p_k)}{b_j}\right) . \tag{2}$$

2.2 The t-Copula Model

The t-copula model differs from the Gaussian copula in the sense that the latent variables have the multivariate t distribution, instead of the multivariate Gaussian distribution. One version of the multivariate t that possesses the property of extremal dependence is obtained by setting

$$X_k' = \sqrt{\frac{r}{V}} \cdot X_k = \sqrt{\frac{r}{V}} \cdot (\mathbf{a}_j^\top \mathbf{Z} + b_j \varepsilon_k)$$

where X_k, $\mathbf{Z}$, $\mathbf{a}_j$ and b_j are defined in exactly the same way as in (1), and $V \sim \chi_r^2$ (chi-square distribution with r degrees of freedom). Since X_k is $N(0,1)$, it is well known that X_k' has the univariate standard t distribution with r degrees of freedom. Then one sets $Y_k = \mathbf{1}\{X_k' > x_k\}$ where, as before, x_k is the default threshold. Let F_r be the cdf of a t distribution with r degrees of freedom. In this case we need $x_k = F_r^{-1}(1 - p_k)$ to ensure that $\mathsf{P}(Y_k = 1) = p_k$. Kuhn (2004) and Kostadinov (2005) also consider a t-copula model of this form.

3 LARGE DEVIATIONS AND FAST SIMULATION OF MULTIFACTOR GAUSSIAN COPULA MODEL

This section reviews the results and algorithms of GKSa and GKSb. Since the current work is based heavily on these papers, for the sake of completeness, we present a somewhat detailed review. For Gaussian copula model, GKSb propose an asymptotically optimal IS procedure. Their approach separates consideration of the credit exposures from the dependence mechanism and default probabilities. Define aggregated credit exposures,

$$C_j := \frac{1}{m}\sum_{k\in\mathcal{I}_j^{(m)}} c_k \quad \text{for } j=1,\ldots,t \text{ and}$$
$$C := \frac{1}{m}\sum_{k=1}^{m} c_k = \sum_{j=1}^{t} C_j \ .$$

GKSa introduce the concept of a *q-minimal index set*. These are sets of obligor types. We say that $\mathcal{J}$ is a q-minimal index set, $0<q<1$, if $\mathcal{J}\subset\{1,\cdots,t\}$ and

$$\max_{\mathcal{J}'\subset\mathcal{J},\mathcal{J}'\neq\mathcal{J}} \sum_{j\in\mathcal{J}'} C_j < qC \le \sum_{j\in\mathcal{J}} C_j \ . \tag{3}$$

The intuitive meaning of q-minimal index set is that $\mathcal{J}$ is one of the index sets sufficient for the portfolio loss to exceed the default threshold $x=qmC$ (note that mC is the maximum possible loss), if all obligors belonging to each index in $\mathcal{J}$ default, but this does not happen for any index set strictly included in $\mathcal{J}$. This characterization is important since, to achieve the optimal IS, it is crucial to change the probability measure on the common factors enough to increase the chance that $L_m > x$. But at the same time, the new probability measure has to be as close to the original measure as possible given that the default event has occurred.

For each type $j=1,\ldots,t$, define

$$d_j^{(m)} := \alpha_1^{(m)}\Phi^{-1}(1-\overline{p}_j)+\alpha_2^{(m)} b_j \Phi^{-1}(q)$$

and a halfspace

$$G_j^{(m)} := \left\{\mathbf{z}\in\mathbb{R}^d : \mathbf{a}_j^\top \mathbf{z} \ge d_j^{(m)}\right\} \tag{4}$$

where $\overline{p}_j = \max_{k\in\mathcal{I}_j^{(m)}} p_k$ is the maximum of default probabilities of obligors belonging to j-th type and $0\le\alpha_1^{(m)}<1$, $0\le\alpha_2^{(m)}<1$.

Define $\mathcal{M}_q$ as the family of all q-minimal index sets and

$$G_{\mathcal{J}}^{(m)} := \bigcap_{j\in\mathcal{J}} G_j^{(m)} \quad \text{for } \mathcal{J}\in\mathcal{M}_q \text{ and}$$
$$G_{\mathcal{M}_q}^{(m)} := \bigcup_{\mathcal{J}\in\mathcal{M}_q} G_{\mathcal{J}}^{(m)}.$$

Note that the condition $\|\mathbf{a}_j\|>0$ in Section 2.1 implies $\mathbf{a}_j\neq\mathbf{0}$. If all $\mathbf{a}_j\ge\mathbf{0}$, then $G_{\mathcal{J}}^{(m)}\neq\emptyset$ for any $\mathcal{J}\subset\{1,\cdots,t\}$. However, if some components of $\mathbf{a}_j$ are negative, these sets may be empty. Because we need to define the new IS distribution using these minimal index sets, a smaller $\mathcal{M}_q$ is desirable for efficient implementation. Hence, we introduce a *sufficient* subfamily of $\mathcal{M}_q$ which includes enough minimal index sets to define an efficient IS distribution. We denote it by $\mathcal{S}_q$. It satisfies (for all m):

Feasibility: For each $\mathcal{J}\in\mathcal{S}_q$, $G_{\mathcal{J}}^{(m)}\neq\emptyset$;
Covering property: $\bigcup_{\mathcal{J}\in\mathcal{S}_q} G_{\mathcal{J}}^{(m)} = G_{\mathcal{M}_q}^{(m)}$.

Note that the choice of $\mathcal{S}_q$ may not be unique, but the asymptotic efficiency of IS does not depend on this choice. For each $\mathcal{J}\in\mathcal{S}_q$, we define $\boldsymbol{\mu}_{\mathcal{J}}^{(m)}$ as the *unique* solution of the following *linearly constrained quadratic optimization* problem:

$$\boldsymbol{\mu}_{\mathcal{J}}^{(m)} := \operatorname{argmin}\left\{\|\mathbf{z}\| \ : \ \mathbf{z}\in G_{\mathcal{J}}^{(m)}\right\} \ . \tag{5}$$

The new importance sampling distribution for the common factors consists of a mixture of multivariate normal distributions with $\boldsymbol{\mu}_{\mathcal{J}}^{(m)}$, $\mathcal{J}\in\mathcal{S}_q$, as the mean vectors.

After sampling the common factors $\mathbf{Z}$, we apply IS to the idiosyncratic risks ε_k through changing the conditional marginal default probabilities from $p_k(\mathbf{Z})$ (see (2)) to exponentially twisted ones given by

$$p_{k,\theta}(\mathbf{Z}) = \frac{p_k(\mathbf{Z})e^{\theta c_k}}{1+p_k(\mathbf{Z})\left(e^{\theta c_k}-1\right)} \tag{6}$$

for some θ. The θ is chosen as

$$\theta_m(\mathbf{z}) := \operatorname*{argmin}_{\theta\ge 0}\{-\theta x + m\psi_m(\theta,\mathbf{z})\} \tag{7}$$

where $\psi_m(\theta,\mathbf{z})$ is the conditional cumulant generating function divided by m,

$$\psi_m(\theta,\mathbf{z}) := \frac{1}{m}\log \mathsf{E}\left[e^{\theta L_m}\middle|\mathbf{Z}=\mathbf{z}\right] \tag{8}$$
$$= \frac{1}{m}\sum_{k=1}^{m}\log\left(1+p_k(\mathbf{z})\left(e^{\theta c_k}-1\right)\right).$$

The IS procedure for the Gaussian copula model is summarized in Figure 1

To analyze the **MIS** algorithm one needs to focus on asymptotic regimes, where large losses rare: in this paper, we focus on the *small default probabilities* regime by imposing the following assumption in addition to those in Section 2.1.

Assumption SDP:

1. $0<c_k\le\overline{c}<\infty$ for $k=1,\ldots$.

Preliminaries: Determine $\mathcal{S}_q$ and $\boldsymbol{\mu}_{\mathcal{J}}^{(m)}$, $\mathcal{J} \in \mathcal{S}_q$, as discussed above. Call these $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K$. Let$\lambda_1, \ldots, \lambda_K$ denote the weights to be used in the mixture distribution. Choose the total number replications, n, and λ_i's so that $\lambda_i \cdot n$ is an integer for all i.

Main Loop: Repeat for replications $r = 1, \ldots, \lambda_i \cdot n$, and for $i = 1, \ldots, K$

1. Sample $\mathbf{Z}$ from $N(\boldsymbol{\mu}_i, \mathbf{I})$.
2. Find $\theta_m(\mathbf{Z})$ by solving (7).
3. Compute the twisted conditional default probabilities $p_{k,\theta_m(\mathbf{Z})}(\mathbf{Z})$, $k = 1, \ldots, m$ by (6) and generate Y_k, $k = 1, \ldots, m$.
4. Calculate $I_r^{(i)} = 1\{L_m > x\}$ $\times e^{-\theta_m(\mathbf{Z})L_m + m\psi_m(\theta_m(\mathbf{Z}),\mathbf{Z})}$ $\times \left(\sum_{i=1}^K \lambda_i \exp\left(\boldsymbol{\mu}_i{}^\top \mathbf{Z} - \frac{1}{2}\boldsymbol{\mu}_i{}^\top \boldsymbol{\mu}_i\right)\right)^{-1}$.

Return the estimate $\frac{1}{n}\sum_{i=1}^K \sum_{r=1}^{\lambda_i \cdot n} I_r^{(i)}$

Figure 1: Mixed Importance Sampling (MIS)

2. If the k-th obligor is of type j then its default probability is given by $p_k = p_j^{(m)} := \Phi(-s_j\sqrt{m})$ where $s_j > 0$. Hence the conditional default probability (given the factors $\mathbf{Z}$) of the same obligor is given by

$$p_k(\mathbf{Z}) = p_j^{(m)}(\mathbf{Z}) = \Phi\left(\frac{\mathbf{a}_j^\top \mathbf{Z} - s_j\sqrt{m}}{b_j}\right) .$$

3. For each type $j = 1, \ldots, t$,

$$C_j := \lim_{m\to\infty} \frac{1}{m} \sum_{k \in \mathcal{I}_j^{(m)}} c_k < \infty \quad \text{and}$$

$$C := \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^m c_k = \sum_{j=1}^t C_j .$$

4. The total loss from defaults and the portfolio default threshold are

$$L_m = \sum_{k=1}^m c_k Y_k^{(m)} \quad \text{and} \quad x \equiv x_m = q\sum_{k=1}^m c_k$$

where $Y_k^{(m)} = \mathbf{1}_{\left\{X_k > \Phi^{-1}\left(1 - p_j^{(m)}\right)\right\}}$ and $0 < q < 1$. Recall that $\sum_{k=1}^m c_k$ is the maximum possible loss and thus we are interested in the loss exceeding a fraction q of this. We impose a mild restriction on the possible values of q; q is not a value in the finite set, $\left\{\frac{1}{C}\sum_{j\in\mathcal{J}} C_j \ : \ \mathcal{J} \subset \{1, \cdots, t\}\right\}$.

We apply the original definition of q-minimal index set with these C_j and C. Define a halfspace

$$G_j = \left\{\mathbf{z} \in \mathbb{R}^d \ : \ \mathbf{a}_j^\top \mathbf{z} \geq s_j\right\} ,$$

and then for each $\mathcal{J} \in \mathcal{M}_q$, define

$$G_{\mathcal{J}} := \bigcap_{j\in\mathcal{J}} G_j .$$

Define $\boldsymbol{\gamma}_{\mathcal{J}}$ as the *unique* solution of the following *linearly constrained* problem:

$$\boldsymbol{\gamma}_{\mathcal{J}} = \begin{cases} \operatorname{argmin}\left\{\|\mathbf{z}\| \ : \ \mathbf{z} \in G_{\mathcal{J}}\right\} & \text{if } G_{\mathcal{J}} \neq \emptyset \\ (\infty, \ldots, \infty)^\top & \text{if } G_{\mathcal{J}} = \emptyset \end{cases} .$$

Define

$$\boldsymbol{\gamma}_* = \min\left\{\left\|\boldsymbol{\gamma}_{\mathcal{J}}\right\| : \mathcal{J} \in \mathcal{M}_q\right\}$$

breaking ties arbitrarily, if necessary. Note that $\boldsymbol{\gamma}_* = (\infty, \ldots, \infty)^\top$ and $\|\boldsymbol{\gamma}_*\| = \infty$ if $G_{\mathcal{J}} = \emptyset$ for all $\mathcal{J} \in \mathcal{M}_q$ by definition. The following large deviations result was proved in GKSa

Theorem 1 *In the multifactor Gaussian copula model with finite number of types, if assumption* **SDP** *is satisfied then*

$$\lim_{m\to\infty} \frac{1}{m} \log \mathsf{P}\left(L_m > x_m\right) = -\frac{1}{2}\left\|\boldsymbol{\gamma}_*\right\|^2 .$$

The sets G_j and $G_j^{(m)}$ (defined in (4)) are related, in the sense that under SDP, in the limit as $m \to \infty$, the halfspace $G_j^{(m)}/\sqrt{m}$ (a set where each element of $G_j^{(m)}$ is divided by $\sqrt{m}$) coincides with the halfspace G_j. This implies that under **SDP**, $\boldsymbol{\gamma}_{\mathcal{J}} = \lim_{m\to\infty} \frac{1}{\sqrt{m}}\boldsymbol{\mu}_{\mathcal{J}}^{(m)}$.

Denote the second moment of IS estimator as $M_2(x_m, \theta_m(\mathbf{Z}))$. If we show that the logarithmic limit of $M_2(x_m, \theta_m(\mathbf{Z}))$ is twice of the RHS constant in Theorem 1, then this is the fastest possible rate for any unbiased estimator because of Jensen's inequality. In the rare event simulation literature, estimators that achieve this are called *asymptotically optimal*. The asymptotic optimality can be interpreted as the following: there is a positive constant c (in fact, $c = \frac{1}{2}\|\boldsymbol{\gamma}_*\|^2$) for which $\mathsf{P}(L_m > x_m) = \exp(-c \cdot m + o(m))$ and $M_2(x_m, \theta_m(\mathbf{Z})) = \exp(-2c \cdot m + o(m))$. This means that the second moment of the estimator decreases at twice the exponential rate of the loss probability itself. For naive simulation, the second moment decreases at the rate $\exp(-c \cdot m + o(m))$.

The choices of $\alpha_1^{(m)}$ and $\alpha_2^{(m)}$ in (4) are important to achieve a large variance reduction. We limit ourselves to

$$\alpha_1^{(m)} = 1 - \epsilon_m \quad \text{and} \quad \alpha_2^{(m)} = 1 - \frac{1}{\sqrt{\log m}} \qquad (9)$$

where $\epsilon_m > 0$ is such that $\epsilon_m \to 0$ and $\epsilon_m \sqrt{m} \to \infty$ as $m \to \infty$. We use $m^{-1/3}$ in our experiments. GKSb proves that **MIS** with (9) is asymptotically optimal under **SDP**.

Theorem 2 *In the multifactor Gaussian copula model with finite number of types, suppose that assumption* **SDP** *holds and* $\mathcal{S}_q \neq \emptyset$*. If we apply* **MIS** *with (9) then*

$$\limsup_{m\to\infty} \frac{1}{m} \log M_2(x_m, \theta_m(\mathbf{Z})) \leq - \left\| \boldsymbol{\gamma}_* \right\|^2 .$$

Hence (using Theorem 1)

$$2 \lim_{m\to\infty} \frac{1}{m} \log \mathsf{P}\left(L_m > x_m\right) = \lim_{m\to\infty} \frac{1}{m} \log M_2(x_m, \theta_m(\mathbf{Z})),$$

and we have asymptotic optimality of the two-step IS estimator obtained by **MIS**.

For details of the analysis and another limiting parametric regime under which **MIS** is also asymptotically optimal, see GKSb.

3.1 Computational Issue in MIS: Approximate Importance Sampling on Common Factors

For instances with large number of types or large number of common factors, **MIS** may be computationally intractable. This maybe because either the size of $\mathcal{S}_q$ is too big (potentially it may have up to 2^t elements), or identifying $\mathcal{S}_q$ takes a long time. GKSb show that $\mathcal{S}_q$ can at most have t^d *distinct* elements. Also, they reformulated the problem of identifying these distinct elements as solving subset sum problems of negligible computing efforts. However, for large t and d, this could still be prohibitive. To reduce the computing burden, one needs to reduce t or d. To reduce the number of types t one may consider clustering ideas as in Hastie, Tibshirani, and Friedman (2001). GKSb focus on the reduction of the number of factors d.

Assume that initially the number of factors is D. Hence $\mathbf{a}_j \in \mathbb{R}^D$ for $j = 1, \ldots, t$. We want to reduce factor dimension from D to d ($< D$). GKSb suggest the use of *Principal Components Analysis* (PCA). By applying PCA (without mean adjusting) to $[\mathbf{a}_1 \cdots \mathbf{a}_t]^\top$, they choose the best subspace of $\mathbb{R}^D$ to explain the variations among factor loading vectors under the restriction of the subspace dimension being d. Then, with the projected factor loading vectors $\mathbf{a}'_j \in \mathbb{R}^d$ on this subspace (i.e., $\mathbf{a}'_j$ is the projection of $\mathbf{a}_j$ onto the subspace), they compute $\{\boldsymbol{\mu}'_{\mathcal{J}} : \mathcal{J} \in \mathcal{S}'_q\} \subset \mathbb{R}^d$ by solving the convex quadratic optimization (5) and subset sum problems (see GKSb for details of this subset sum problem) in $\mathbb{R}^d$ with t types. Here we use the notation $\mathcal{S}'_q$ to emphasize that these factor shifting mean vectors come from the approximation. One can also reduce the number of types by aggregating two types if their projected factor loading vectors are close to each other since the marginal default probabilities are allowed to vary within a type. Using the orthonormal basis on the subspace constructed by PCA, one can recover $\{\boldsymbol{\mu}_{\mathcal{J}} : \mathcal{J} \in \mathcal{S}'_q\} \subset \mathbb{R}^D$ corresponding to $\{\boldsymbol{\mu}'_{\mathcal{J}} : \mathcal{J} \in \mathcal{S}'_q\}$.

GKSb use $\{\boldsymbol{\mu}_{\mathcal{J}} : \mathcal{J} \in \mathcal{S}'_q\}$ to shift the common factors and to compute the likelihood ratios, but they use the *exact* factor loadings, $\{\mathbf{a}_j\}_{j=1}^T$, in the valuation of latent variables for each obligor. Hence this IS procedure using approximate mean shifting vectors is unbiased. One expects variance reductions because d of the most important factor loadings are considered.

4 FAST SIMULATION FOR THE t-COPULA MODEL

We now present the main contribution of this paper, i.e, an importance sampling algorithm for the t-copula model (see Section 2.2). In the discussion that follows we use a SDP type of regime where we parameterize $p_k = p_j^{(m)} = F_r(-s_j\sqrt{m})$ if the kth obligor is of type j (instead of $p_k = p_j^{(m)} = \Phi(-s_j\sqrt{m})$ that we used for the Gaussian copula model in the SDP regime). Assume that $V = v > 0$ is fixed. Then

$$X'_k = \sqrt{\frac{r}{v}} \cdot (\mathbf{a}_j^\top \mathbf{Z} + b_j \varepsilon_k) > x_k$$
$$\Leftrightarrow \quad \mathbf{a}_j^\top \mathbf{Z} + b_j \varepsilon_k > \sqrt{\frac{v}{r}} s_j \sqrt{m}.$$

That is, given $V = v$, we return to the Gaussian copula model with changed marginal default probability. We can then use **MIS**.

We now focus on the change of measure for the V. One can express $P(L_m > x_m) = E\left(P(L_m > x_m|V)\right)$. If $P(L_m > x_m|V = v)$ were computable easily for each v, then one could estimate $P(L_m > x_m)$, by first sampling V's and then computing $P(L_m > x_m|V)$ for each sampled V. A change of measure on the V that will yield *zero variance* in this estimation is given by

$$\frac{\mathsf{P}(L_m > x_m|V = v) f_{\chi_r^2}(v)}{\int_0^\infty \mathsf{P}(L_m > x_m|V = v) f_{\chi_r^2}(v) dv}. \qquad (10)$$

However since $P(L_m > x_m|V = v)$ is not easily computable, we will use an approximation to $P(L_m > x_m|V = v)$ as its surrogate, and then use that in (10) to come up with a change of measure on the V. We will use Theorem 1 to come up with such an approximation, since given $V = v$, we can use results from the Gaussian copula model.

Theorem 1 tells us that

$$\lim_{m\to\infty} \frac{1}{m} \log \mathsf{P}(L_m > x_m | V = v) = -\frac{1}{2}\|\boldsymbol{\gamma}_v\|^2$$

where $\boldsymbol{\gamma}_v$ is the optimal solution to the optimization problem based on the constraint sets

$$\left\{\mathbf{z} : \mathbf{a}_j^\top \mathbf{z} \geq s_j \sqrt{\frac{v}{r}}\right\}.$$

Denote the optimal solution to the $v = 1$ case by $\boldsymbol{\gamma}_*$. Then $\boldsymbol{\gamma}_v = \sqrt{v}\boldsymbol{\gamma}_*$ is the optimal solution to the $V = v$ case. Hence we get

$$\lim_{m\to\infty} \frac{1}{m} \log \mathsf{P}(L_m > x_m | V = v) = -\frac{1}{2}\|\boldsymbol{\gamma}_*\|^2 v.$$

We can re-write this as

$$\mathsf{P}(L_m > x_m | V = v) = \exp(-c_m \cdot v + o_v(m))$$

where

$$c_m = \frac{1}{2} \cdot m \cdot \|\boldsymbol{\gamma}_*\|^2. \qquad (11)$$

Note that $o_v(m)$ depends on v as well as m. This means that the m large enough to discard $o_v(m)$ may depend on v. Hence, for a fixed m, the conditional probability is not proportional to $\exp(-c_m \cdot v)$ for all values of v. However, we still use the approximation,

$$\mathsf{P}(L_m > x_m | V = v) \approx \text{constant} \times \exp(-c_m \cdot v) \qquad (12)$$

as a surrogate for $\mathsf{P}(L_m > x_m | V = v)$ in (10). This gives

$$e^{-c_m \cdot v} \cdot f_{\chi_r^2}(v)/M_m \qquad (13)$$

as the expression in (10) where

$$\begin{aligned} M_m &= \int_0^\infty e^{-c_m v} \cdot C v^{r/2-1} e^{-v/2} dv \\ &= \mathsf{E}[\exp(-c_m V)] = (2c_m + 1)^{-r/2}. \end{aligned}$$

We will use this as the candidate IS distribution for V.

A change of measure of the type given by (13) is referred to in the literature, as an "exponential tilting" or "exponential twisting" of the original measure by amount $-c_m$. One can easily show that exponentially tilting a chi-square distribution with r degrees of freedom by amount $-c_m$, yields a Gamma($r/2$, $(c_m + 1/2)^{-1}$), i.e., a gamma distribution with shape parameter $r/2$ and scale parameter $(c_m + 1/2)^{-1}$. Thus in the IS procedure, one will need to sample from this Gamma distribution. Let W denote a random variable with this new distribution. Note that one can sample W, by first sampling the original chi-square V, and then scaling by $(2c_m+1)^{-1}$, i.e., set $W = \frac{1}{2c_m+1}V$. We use the latter approach in our experiments. The associated likelihood ratio is given by

$$e^{c_m W} \cdot M_m. \qquad (14)$$

As mentioned before, after sampling V from the new distribution (i.e., sampling W), one can apply **MIS**. Although this can theoretically be done, this poses a computational problem. Since the conditional default probabilities given W change for every sample of W, we need to solve the optimization problems for finding the mean-shift vectors for each sample of W. To overcome this problem, we apply stratification to the W. For each stratum, we fix mean shiftings as the one computed for the midpoint of the interval on W. Note that since we use these mean-shifts only in the change of measure, this does *not* introduce bias in our estimates. Stratification also has the effect of reducing the variance of the likelihood ratio in (14).

The IS for the t-copula model is summarized in Figure 2.

5 NUMERICAL EXAMPLES

5.1 Exact IS for 30 Random Instances with the Small Number of Factors

We apply IS to an example based on the t-copula model. The description of (1) is as follows. 60% of the coefficients of the factor-loading vectors are non-zero. Each nonzero component is uniformly generated on $[-0.2, 1]$. Then they are scaled such that $\|\mathbf{a}_j\|$, $j = 1, \ldots, t$ are distributed uniformly on $[0.1, 0.7]$. The potential loss amount of each obligor is deterministic and chosen from a discrete uniform distribution on $\{1, 2, \ldots, 30\}$. The marginal default probabilities associated with the k-th obligor is generated by $0.0255 + 0.0245 \times \sin(16\pi k/m)$ such that it lies within $(0.1\%, 5\%)$. We also randomize the number of obligors in each type. We generate a uniform random number on $[0.4, 1]$ for each type. We divide this by the sum of the generated numbers for all types to compute proportions of the number of obligors of that type. We also make sure that the number of obligors in one type does not exceed 150% of that of any other type.

We test 30 randomly generated instances of 1000 obligors belonging to one of 25 types. We set the degree of freedom of the t-distribution to 5. (The effect of varying degrees of freedom is considered in Section 5.2.) The probabilities of the portfolio loss exceeding 30%, 40%, and 50% of the total credit exposure are estimated. These threshold values are chosen because they produce portfolio loss

Compute c_m: Apply the procedure for Gaussian copula to find $\boldsymbol{\gamma}_*$. Then compute c_m by (11).

Construct equiprobable strata: To construct s equiprobable strata, set $q_i = i/s$ for $i = 0, \ldots, s$. Find v_i, $1 \le i \le s$, such that $\mathsf{P}(V \le v_i) = \frac{q_{i-1}+q_i}{2}$, where V has the χ_r^2 distribution. Set $w_i = v_i/(2c_m + 1)$. (This is equivalent to finding w_i such that $\mathsf{P}(W \le w_i) = \frac{q_{i-1}+q_i}{2}$ where W is Gamma($r/2$, $(c_m + 1/2)^{-1}$)).

Main Loop: Repeat for each stratum $i = 1, \ldots, s$

1. By assuming $\sqrt{\frac{w_i}{r}}x_k$ as the new threshold of the Gaussian latent variable $\mathbf{a}_j^\top \mathbf{Z} + b_j \varepsilon_k$, determine $\mathcal{S}_q$ and compute the mean shifting vectors $\boldsymbol{\mu}_{\mathcal{J}}$ for $\mathcal{J} \in \mathcal{S}_q$.
2. Generate n_i samples of V from χ_r^2 distribution by the inverse transform method using n_i uniform samples on $[q_{i-1}, q_i)$. Then compute $W = V/(2c_m + 1)$. (This is equivalent to generating n_i samples of W from Gamma($r/2$, $(c_m + 1/2)^{-1}$) by the inverse transform method using n_i uniform samples on $[q_{i-1}, q_i)$.)
3. For each sampled W, change the threshold of Gaussian latent variable as $\sqrt{\frac{W}{r}}x_k$ and apply **MIS** with mean shifting vectors $\boldsymbol{\mu}_{\mathcal{J}}$ for $\mathcal{J} \in \mathcal{S}_q$.
4. For each sampled W, multiply the output from **MIS** with the likelihood ratio $e^{c_m W} \cdot M_m$.
5. Compute $\hat{\alpha}_i$ as the sample mean of the n_i outputs, and $\hat{\sigma}_i^2$ as the sample variance of the n_i outputs.

Estimator: Return $\frac{1}{s}\sum_{i=1}^{s}\hat{\alpha}_i$ as the estimator and $\frac{1}{s^2}\sum_{i=1}^{s}\hat{\sigma}_i^2/n_i$ as the sample variance of the estimator.

Figure 2: Importance Sampling for t-copula (**IS-T**)

probabilities within a range of $10^{-2.5}$ to 10^{-5}. We used the parametrization given by (9) for **MIS**. In this experiment, we use the exact minimal index sets, that is, we do not use approximate IS. Figure 3 depicts the pairs – portfolio loss probabilities and estimated variance reduction factors. The plot shows estimated variance reduction factors of more than 100 if the probability of the portfolio loss exceeding the threshold is less than 10^{-3}; and more than $10^{2.5}$ if it is less than 10^{-4}. Table 1 gives some details of the results.

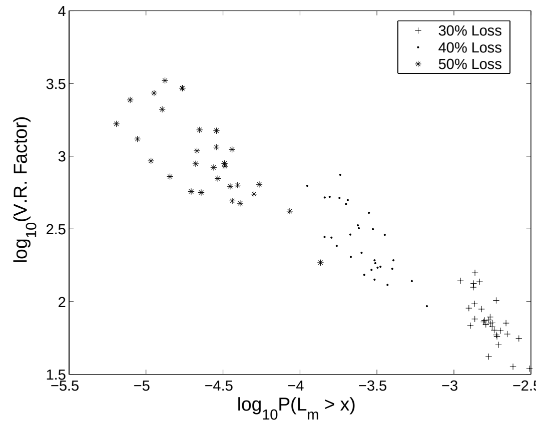

Figure 3: Plot of the Pairs – Estimates of Portfolio Loss Probabilities and Variance Reduction Factors – of 30 Random Instances, on Logarithmic Scales. For each Instance, the Triplet (+,·,∗) Represents the Numbers Corresponding to Portfolio Losses of More Than 30%, 40%, and 50%, Respectively, of the Total Credit Exposure

5.2 Exact IS for a Random Instance with Various Degrees of Freedom

To see the effect of degrees of freedom on the variance reductions, we fix one random instance (the first one in Table 1) and apply **IS-T** for various degrees of freedom for the t-distribution and various portfolio loss thresholds. We used 2,3,4,5, and 6 as the different degrees of freedom and 10%, 20%, 30%, 40%, 50%, and 60% as the different portfolio loss thresholds. Hence we tested the instance for 30 combinations of parameters. Figure 4 depicts line plots for each degree of freedom with respect to several portfolio loss thresholds. Except $r = 2$ case, we see a strong linear relationship between the loss probabilities and variance reductions. Since $r = 2$ case corresponds to the heaviest-tailed risk factors (which results in the largest loss probabilities among the tested degrees of freedom), the variance reductions seems not sufficiently large at the small portfolio loss thresholds. From this experiment, we see that the variance reductions are not affected much by the degrees of freedom since the variance reductions at portfolio loss thresholds corresponding to the same portfolio loss probability are quite similar.

5.3 Approximate IS for Structured Factor Models with Sparse Factor Loadings

This example is taken from Glasserman and Li (2003), but adapted to the t-copula case. There are 1000 obligors with probabilities of default given by

$$p_k = 0.01 \cdot (1 + \sin(16\pi k/m)), \quad k = 1, \ldots, 1000,$$

Table 1: Estimated Probabilities and Variance Reduction Factors at Three Loss Levels in the 25-type Model. The Degree of Freedom of χ^2 is 5. We Just Present 5 of the 30 Instances

Instance No.	Est. Prob.			Est. V.R.		
	q=30%	q=40%	q=50%	q=30%	q=40%	q=50%
1	0.0017	0.00031	3.5e-005	70	141	619
2	0.0014	0.00017	1.4e-005	76	241	724
3	0.0015	0.0002	1.7e-005	137	468	2936
4	0.0026	0.00053	8.6e-005	56	138	418
5	0.0013	0.00016	1.1e-005	133	525	2717

Table 2: The Ratio of Explained Squared Variations of Factor Loading Coefficients

	$(\alpha_R, \alpha_F, \alpha_G)$			
	(0.8,0.4,0.4)	(0.5,0.4,0.4)	(0.2,0.4,0.4)	(0.25, 0.15,0.05)
Single Dominating Factor in $\mathbb{R}^{21}$	79%	60%	25%	74%
Two Dominating Factors in $\mathbb{R}^{22}$	80%	64%	31%	77%

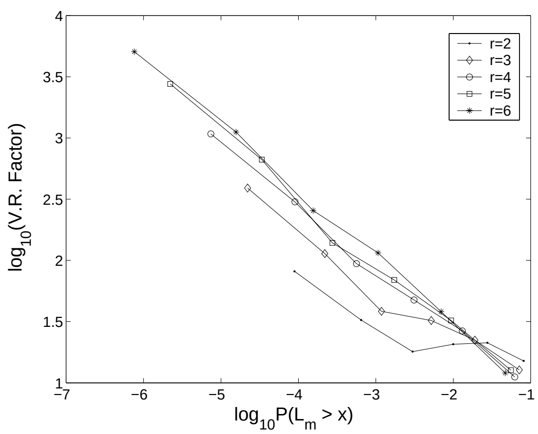

Figure 4: Plot of the Pairs – the Large Loss Probability of Portfolio and Estimate of Variance Reduction Factor – of Experiments in Logarithmic Scales

Table 3: Variance Reduction Factors in 21-factor Model, the Factor Loading Coefficients are (0.8, 0.4, 0.4)

Loss %	$P(L > x)$ by IS	Est. V.R.
10%	0.0268	18
30%	0.0081	49
50%	0.0029	54
70%	0.0008	144

Table 4: Variance Reduction Factors in 21-factor Model, the Factor Loading Coefficients are (0.5, 0.4, 0.4)

Loss %	$P(L > x)$ by IS	Est. V.R.
10%	0.0227	33
30%	0.0025	142
50%	0.0003	992
60%	0.0001	3088

Table 5: Variance Reduction Factors in 21-factor Model, the Factor Loading Coefficients are (0.2, 0.4, 0.4)

Loss %	$P(L > x)$ by IS	Est. V.R.
10%	0.0160	12
20%	0.0027	52
30%	0.0005	275
40%	0.00008	1863

Table 6: Variance Reduction Factors in 21-factor Model, he Factor Loading Coefficients are (0.25, 0.15, 0.05)

Loss %	$P(L > x)$ by IS	Est. V.R.
10%	0.0152	10
20%	0.0024	44
30%	0.0004	294
40%	0.00006	3281

Table 7: Variance Reduction Factors in 22-factor Model, the Factor Loading Coefficients are (0.8, 0.4, 0.4)

Loss %	$P(L > x)$ by IS	Est. V.R.
10%	0.0276	13
30%	0.0059	14
50%	0.0013	45
70%	0.0001	470

Table 8: Variance Reduction Factors in 22-factor Model, the Factor Loading Coefficients are (0.5, 0.4, 0.4)

Loss %	$P(L > x)$ by IS	Est. V.R.
10%	0.0210	22
20%	0.0052	29
30%	0.0015	157
40%	0.0005	225
50%	0.0001	760

Table 9: Variance Reduction Factors in 22-factor Model, the Factor Loading Coefficients are (0.2, 0.4, 0.4)

Loss %	$\mathsf{P}(L > x)$ by IS	Est. V.R.
10%	0.0163	12
20%	0.0024	39
30%	0.0004	321
40%	0.00006	2850

Table 10: Variance Reduction Factors in 22-factor Model, the Factor Loading Coefficients are (0.25, 0.15, 0.05)

Loss %	$\mathsf{P}(L > x)$ by IS	Est. V.R.
10%	0.0183	0.1
20%	0.0050	0.06
30%	0.0003	78
40%	0.00003	2867

and loss given default given by

$$c_k \;=\; 1+\frac{99}{999}(k-1), \quad k=1,\ldots,1000.$$

There are 100 types and 21 factors. The factor loading matrix A, the jth row of which is $\mathbf{a}_j$, is given by

$$A = \left(\begin{array}{c|ccc|c} & F & & & G \\ R & & \ddots & & \vdots \\ & & & F & G \end{array} \right), \quad G = \begin{pmatrix} c_G & & \\ & \ddots & \\ & & c_G \end{pmatrix}.$$

R is a column vector of 100 entries, all equal to a constant c_R; F is a column vector of 10 entries, all equal to a constant c_F; G is a 10×10 diagonal matrix with the diagonal elements set to a constant c_G. The 21-factor model thus has a single dominant market factor. Each set of 10 obligors (1-10, 11-20, $\cdots$) are of the same type. We consider different factor loadings: $(c_R, c_F, c_G) = (0.8, 0.4, 0.4)$ is associated with a large market factor case, $(c_R, c_F, c_G) = (0.5, 0.4, 0.4)$ with a medium market factor, and $(c_R, c_F, c_G) = (0.2, 0.4, 0.4)$ with a small market factor. We also consider $(c_R, c_F, c_G) = (0.25, 0.15, 0.05)$, suggested by Morokoff (2004).

In addition, we also consider a 22-factor model, with two dominant market factors. The 22-factor model differs from the 21-factor model only in R. R is now a 100×2 matrix with the first fifty entries of the first column and the last fifty entries of the second column equal to c_R; all other entries of R are zero.

To apply approximate IS, we compute singular value decompositions of A or more specifically, eigenvectors of $A^\top A$. We measured the effectiveness of these approximations by a sum of squared variations. Table 2 summarizes how much portion of squared variations of factor loading coefficients is explained by these approximations.

Approximate IS works well for both models. Tables 3 – 6 summarize the variance reduction estimates for the 21-factor model, and Tables 7 – 10 for the 22-factor model. We set the degrees of freedom of the χ^2 (and thus the t distribution) to 5 in these experiments.

ACKNOWLEDGMENTS

This work was partially supported by the National Science Foundation (U.S.A.) Grant DMI 03-00044. The work of the first author was performed while he was a graduate student at Columbia University.

REFERENCES

Breymann, W., A. Dias, and P. Embrechts. 2003. Dependence structure for multivariate high-frequency data in finance. *Quantitative Finance* 3: 1–14.

Demarta, S., and A. J. McNeil. 2004. The *t* copula and related copulas. Technical Report, ETH Zurich.

Embrechts, P., F. Lidskog, and A. J. McNeal. 2001. Modeling dependence with copulas and applications to risk management. Technical Report, Risklab, ETH Zurich.

Glasserman, P., W. Kang, and P. Shahabuddin. 2005a. Large deviations in multifactor portfolio credit risk. Technical Report, Graduate School of Business and IEOR Department, Columbia University.

Glasserman, P., W. Kang, and P. Shahabuddin. 2005b. Fast simulation of multifactor portfolio credit risk. Technical Report, Graduate School of Business and IEOR Department, Columbia University.

Glasserman, P., and J. Li. 2003. Importance sampling for portfolio credit risk. *Management Science*, to appear.

Hastie, T., R. Tibshirani, and J. Friedman. 2001. *The elements of statistical learning: Data mining, inference, and prediction*. New York: Springer-Verlag.

Kostadinov, K. 2005. Tail approximation for credit risk portfolios with heavy-tailed risk factors. Technical Report, The Munich University of Technology.

Kuhn, G. 2004. Tails of credit default portfolios. Technical Report, The Munich University of Technology.

Lucas, A., P. Klaassen, P. Spreij, and S. Straetmans. 2001. An analytical approach to credit risk in large corporate bond and loan portfolios. *Journal of Banking and Finance* 25: 1635–1664.

Lucas, A., P. Klaassen, P. Spreij, and S. Straetmans. 2003. Tail behavior of credit loss distributions for general latent factor models. *Applied Mathematical Finance* 10: 337–357.

Mashal, R., and A. Zeevi. 2002. Beyond correlation: Extreme co-movements between financial assets. Technical Report, Graduate School of Business, Columbia University, New York.

Morokoff, W.J. 2004. An importance sampling method for portfolios of credit risky assets. In *Proceedings of the 2004 Winter Simulation Conference*, ed. Ingalls, R.G., M.D. Rossetti, J.S. Smith, and B.A. Peters, 1668-1676. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

AUTHOR BIOGRAPHIES

WANMO KANG is a financial researcher at Moody's KMV. His main focus of interest is operations research and its application to finance. He received his Ph.D. in Operations Research from Columbia University. His e-mail address is <Wanmo.Kang@mkmv.com>.

PERWEZ SHAHABUDDIN is a Professor in the Industrial Engineering and Operations Research Department at Columbia University. From 1990 to 1995, he was a Research Staff Member at IBM T.J. Watson Research Center. He received a B.Tech. in Mechanical Engineering (1984) from the Indian Institute of Technology, Delhi, India, and an M.S. in Statistics (1987) and a Ph.D. in Operations Research (1990) from Stanford University. He serves as the Department Editor for Stochastic Models and Simulation at Management Science. His e-mail address is <perwez@ieor.columbia.edu>.