An Approximation for Credit Portfolio Losses

Rüdiger Frey

Universität Leipzig

Monika Popp

Universität Leipzig

Stefan Weber

Cornell University

April 27, 2007[*]

1 Introduction

Mixture models play an important role in the modeling of portfolio losses. In these models the risk of default of individual obligors (indexed by $i \in \{1, \ldots, m\}$) depends on an underlying set of common economic factors, denoted $\boldsymbol{\Psi}$. Given these factors, the losses due to default l_i of individual obligors are assumed to be stochastically independent. Dependence between different obligors stems only from dependence of the individual default probabilities on the set of factors. These models are used for both risk management of credit portfolios and valuation of multi-name credit derivatives. The current article investigates both issues.

The numerical evaluation of the portfolio loss distribution is usually based on the two-stage structure of mixture models. For instance, in order to sample from the loss distribution by standard Monte Carlo, one generates first a realization of the systematic factor variable $\boldsymbol{\Psi}$. In a second step one generates a sequence of independent variates $\hat{l}_i$, $1 \leq i \leq m$, according to the conditional distribution of $(l_i)_{1 \leq i \leq m}$ given $\boldsymbol{\Psi}$. Standard Monte Carlo can be quite slow, and so various numerical techniques for estimating the distribution of the total loss of a portfolio in mixture models have been developed. In this paper we focus on the the second stage, i.e. the conditional loss distribution given the underlying factors, and propose an alternative way for evaluating the conditional distribution of the total loss $L = \sum_{i=1}^{m} l_i$. Our approximation is based on a central limit theorem; error bounds can be derived from the Berry-Esseen-inequality. We compare the numerical performance of our method relative to the standard Vasicek-approximation and the true loss distribution obtained by standard Monte Carlo methods. It turns out that our suggested approximation technique often provides more accurate results than the Vasicek-approximation while being computationally less expensive

---
[*]Corresponding authors: Rüdiger Frey, Universität Leipzig, Mathematisches Institut, Fakultät für Mathematik und Informatik, 04081 Leipzig, Germany, Phone +49 341 97-32180, Fax +49 341 97-32197, email: `Ruediger.Frey@math.uni-leipzig.de`, web: `www.math.uni-leipzig.de/∼frey`; Stefan Weber, School of Operations Research and Industrial Engineering, Cornell University, 279 Rhodes Hall, Ithaca, NY 14853, USA, Phone (607) 254-4825, Fax (607) 255-9129, email: `sweber@orie.cornell.edu`, web: `people.orie.cornell.edu/∼sweber`.

than Monte Carlo algorithms. In particular, we use the loss distribution estimates for the calculation of CDO spreads and find that the accuracy is significantly improved if our method is used instead of the Vasicek approximation; this improvement comes with low additional numerical cost. Related work is surveyed in the interesting paper Glasserman and Ruiz-Mata [4]

The paper is structured as follows: Section 2 introduces the mixture model in which we work and provides an analysis of the approximation method. In Section 3 we present two methods for the estimation of portfolio loss probabilities in a Gaussian one-factor model and illustrate these techniques numerically for different portfolios.

2 Approximation and Error Bounds

We consider a portfolio of m obligors. The loss resulting from obligor $i \in \{1, 2, \ldots, m\}$ is modeled by a random variable which is denoted by $l_i \geq 0$. We are interested in the distribution of the portfolio loss $L^{(m)} = \sum_{i=1}^{m} l_i$. In particular, we provide methods to estimate the tail function of the loss distribution, i.e. the probability $P[L^{(m)} > x]$ of the event that $L^{(m)}$ exceeds a certain threshold x. This quantity is of interest for a variety of reasons: in credit risk management an estimate of the tail of the loss distribution can be used to compute risk measures such as Value at Risk or Expected Shortfall and hence economic and regulatory capital. Moreover, an efficient numerical procedure for computing the tail function of the portfolio loss distribution (under a risk-neutral measure) is useful also for the computation of CDO tranche spreads in factor copula models.

For the analysis we need the following condition on the structure of the model.

Assumption 2.1.

(i) For some $d < m$ there exists a d-dimensional random vector $\boldsymbol{\Psi} = (\Psi_1, ..., \Psi_d)'$ such that individual losses l_i are independent conditional on $\boldsymbol{\Psi}$.

(ii) The first three conditional moments of the random variables l_i, $i = 1, 2, \ldots, m$, are assumed to be finite. We introduce the following notation[1] for the conditional mean, and the centered conditional second and third moments:

$$
\begin{aligned}
\bar{l}_i(\boldsymbol{\psi}) &:= E[l_i | \boldsymbol{\Psi} = \boldsymbol{\psi}], \\
\sigma_i^2(\boldsymbol{\psi}) &:= \mathrm{Var}[l_i | \boldsymbol{\Psi} = \boldsymbol{\psi}], \\
\gamma_i(\boldsymbol{\psi}) &:= E[|l_i - \bar{l}_i(\boldsymbol{\psi})|^3 | \boldsymbol{\Psi} = \boldsymbol{\psi}], \qquad (\boldsymbol{\psi} \in \mathbb{R}^d).
\end{aligned}
$$

Letting H be the distribution function of the random vector $\boldsymbol{\Psi}$, we denote by $F_L^{(m)}(x) = P[L^{(m)} \leq x]$ and $F_{L|\boldsymbol{\Psi}}^{(m)}(x|\boldsymbol{\psi}) = P[L^{(m)} \leq x | \boldsymbol{\Psi} = \boldsymbol{\psi}]$ the unconditional resp. conditional distribution function of the total portfolio loss. The following proposition provides approximations

---

[1]The functions $\bar{l}_i$, σ_i^2 and γ_i are only $P^{\boldsymbol{\Psi}}$-almost surely well-defined. This technical detail will not cause any difficulties in the applications which we have in mind.

for $F^{(m)}$ and $F_{L|\Psi}^{(m)}$ together with error bounds which are derived from the central limit theorem and the Berry-Esseen inequality.

Proposition 2.2. *Suppose that Assumption 1 is satisfied. Letting Φ be the distribution function of the standard normal distribution, we consider the family of distribution functions*

$$G_{L|\Psi}^{(m)}(x|\boldsymbol{\psi}) := \Phi\left(\frac{x - \sum_{i=1}^{m} \bar{l}_i(\boldsymbol{\psi})}{\sqrt{\sum_{i=1}^{m} \sigma_i^2(\boldsymbol{\psi})}}\right).$$

Moreover, we set

$$G^{(m)}(x) := \int G_{L|\Psi}^{(m)}(x|\boldsymbol{\psi}) H(d\boldsymbol{\psi}).$$

Then there exists some constant A, independent of m, such that we have the following error estimate for the conditional and unconditional loss distribution and their approximations:

$$\sup_{x\geq 0} \left| F_{L|\Psi}^{(m)}(x|\boldsymbol{\psi}) - G_{L|\Psi}^{(m)}(x|\boldsymbol{\psi}) \right| \leq A \frac{\sum_{i=1}^{m} \gamma_i(\boldsymbol{\psi})}{\left(\sum_{i=1}^{m} \sigma_i^2(\boldsymbol{\psi})\right)^{3/2}} \quad and$$

$$\sup_{x\geq 0} \left| F^{(m)}(x) - G^{(m)}(x) \right| \leq A \int \frac{\sum_{i=1}^{m} |l_i - \bar{l}_i(\boldsymbol{\psi})|^3}{\left(\sum_{i=1}^{m} \sigma_i^2(\boldsymbol{\psi})\right)^{3/2}} H(d\boldsymbol{\psi}). \tag{1}$$

Remark 2.3. 1. In the sequel we will sometimes refer to $G^{(m)}(x)$ as *second-order approximation* of the portfolio loss distribution.
2. For typical portfolios the integral term on the right hand side of (1) becomes very small for m large. In particular, it will be shown in (3) below, that for a homogeneous portfolio the right hand side of (1) decays like $m^{-1/2}$. Bounds on the constant A are discussed in Remark 2.6 below.

The proof of the proposition is based on the following theorem which we quote from Petrov [8]:

Theorem 2.4 (Petrov, Theorem V.2.3). *Let $Z_1, ..., Z_m$ be independent random variables with $E[Z_i] = 0$ and $E[|Z_i|^3] < \infty$, $i = 1, ..., m$. Then there exists a constant A such that*

$$\sup_{x\in\mathbb{R}} \left| P\left[\frac{1}{\sqrt{\sum_{i=1}^{m} \tilde{\sigma}_i^2}} \sum_{i=1}^{m} Z_i < x\right] - \Phi(x) \right| \leq A C_m, \tag{2}$$

with

$$\tilde{\sigma}_i^2 = E\left[Z_i^2\right] \quad and \quad C_m = \frac{\sum_{i=1}^{m} E\left[|Z_i|^3\right]}{\left(\sum_{i=1}^{m} \tilde{\sigma}_i^2\right)^{3/2}}.$$

Proof of Proposition 2.2. Since the individual losses l_i, $i = 1, 2, \ldots, m$, are conditionally independent, Theorem 2.4 implies:

$$\sup_{x \geq 0} \left| P\left[\frac{\sum_{i=1}^{m}(l_i - \bar{l}_i(\boldsymbol{\psi}))}{\sqrt{\sum_{i=1}^{m} \sigma_i^2(\boldsymbol{\psi})}} < x \,\middle|\, \boldsymbol{\Psi} = \boldsymbol{\psi} \right] - \Phi(x) \right| \leq A \frac{\sum_{i=1}^{m} \gamma_i(\boldsymbol{\psi})}{\left(\sum_{i=1}^{m} \sigma_i^2(\boldsymbol{\psi})\right)^{3/2}}$$

which is equivalent to

$$\sup_{x \geq 0} \left| P\left[\sum_{i=1}^{m} l_i < x \,\middle|\, \boldsymbol{\Psi} = \boldsymbol{\psi} \right] - \Phi\left(\frac{x - \sum_{i=1}^{m} \bar{l}_i(\boldsymbol{\psi})}{\sqrt{\sum_{i=1}^{m} \sigma_i^2(\boldsymbol{\psi})}} \right) \right| \leq A \frac{\sum_{i=1}^{m} \gamma_i(\boldsymbol{\psi})}{\left(\sum_{i=1}^{m} \sigma_i^2(\boldsymbol{\psi})\right)^{3/2}},$$

and thus proves the approximation for the conditional distribution function. Taking expectations, Jensen's inequality yields the approximation for the unconditional distribution function:

$$\begin{aligned}
\left| F^{(m)}(x) - G^{(m)}(x) \right| &\leq \int |F_{L|\boldsymbol{\Psi}}^{(m)}(x|\boldsymbol{\psi}) - G_{L|\boldsymbol{\Psi}}^{(m)}(x|\boldsymbol{\psi})| \, H(d\boldsymbol{\psi}) \\
&\leq A \int \frac{\sum_{i=1}^{m} \gamma_i(\boldsymbol{\psi})}{(\sum_{i=1}^{m} \sigma_i^2(\boldsymbol{\psi}))^{3/2}} \, H(d\boldsymbol{\psi}).
\end{aligned}$$

$\square$

Homogeneous portfolios. Suppose that the individual losses l_i, $i = 1, \ldots, m$, are identically distributed given $\boldsymbol{\Psi}$, so that the portfolio is homogeneous. In that case the conditional moment functions are independent of i, $\bar{l}_i = \bar{l}$, $\gamma_i = \gamma$, and $\sigma_i = \sigma$, and we obtain the following simplifications for $G^{(m)}$ and the error bound.

Corollary 2.5. *Suppose that Assumption 1 holds. If the portfolio is moreover homogeneous, we obtain that*

$$G^{(m)}(x) = \int \Phi\left(\frac{x - m\,\bar{l}(\boldsymbol{\psi})}{\sqrt{m}\,\bar{\sigma}(\boldsymbol{\psi})} \right) H(d\boldsymbol{\psi}),$$

$$\sup_{x \geq 0} \left| F^{(m)}(x) - G^{(m)}(x) \right| \leq \frac{A}{\sqrt{m}} E\left[\frac{\bar{\gamma}(\boldsymbol{\psi})}{\bar{\sigma}^3(\boldsymbol{\psi})} \right]. \tag{3}$$

Remark 2.6. The optimal universal constants A in Proposition 2.2 and Corollary 2.5 are unknown, but lower and upper bounds can be provided. A lower bound for A is given by $\frac{3+\sqrt{10}}{6\sqrt{2\pi}}$, see Esseen [2]. Useful for applications are small upper bounds, since these can be used for the constant A in inequalities (1) and (3) which allows explicit calculations. Van Beek [1] gave the upper bound 0.7975. Using computational methods, Shiganov [9] obtained an upper bound of 0.7915 for the optimal constant in inequality (1) (the general case) and an upper bound of 0.7655 for the optimal constant in inequality (3) (the case of a homogeneous portfolio).

In practical applications it is often assumed that the individual losses are of the form

$$l_i = e_i \cdot \delta_i(\boldsymbol{\psi}) \cdot Y_i \qquad (i = 1, 2, \ldots, m)$$

where the positive constant e_i denotes the exposure due to obligor i, $\delta_i : \mathbb{R}^d \to (0, 1]$ is the corresponding percentage loss given default which is modeled as a deterministic function of the underlying factors, and the random Bernoulli variable Y_i represents the default indicator of obligor i ($Y_i = 1$ corresponds to default, $Y_i = 0$ to survival of firm i). The default indicators are assumed to be independent given the factors $\boldsymbol{\Psi}$, and the default probability of obligor i conditional on $\boldsymbol{\Psi} = \boldsymbol{\psi}$ is denoted by $p_i(\boldsymbol{\psi})$. If all firms have the same deterministic exposure $e_i = e$, and if both the conditional loss given default and the conditional default probabilities do not depend on i, i.e. $\delta_i(\boldsymbol{\psi}) = \delta(\boldsymbol{\psi})$ and $p_i(\boldsymbol{\psi}) = p(\boldsymbol{\psi})$ for all i, then the portfolio is homogeneous, and (3) can be expressed in terms of the constant e and the functions δ and p. We have

$$
\begin{aligned}
\bar{\gamma}(\boldsymbol{\psi}) &= (e\delta(\boldsymbol{\psi}))^3(1 - p(\boldsymbol{\psi}))p(\boldsymbol{\psi})\,(1 - 2 \cdot p(\boldsymbol{\psi})) \\
\bar{\sigma}^3(\boldsymbol{\psi}) &= (e\delta(\boldsymbol{\psi}))^3(p(\boldsymbol{\psi})\{1 - p(\boldsymbol{\psi})\})^{3/2}
\end{aligned}
$$

The error bound becomes

$$\sup_x \left| F^{(m)}(x) - G^{(m)}(x) \right| \leq \frac{0.7655}{\sqrt{m}} \cdot \int \frac{1 - 2 \cdot p(\boldsymbol{\psi})}{\sqrt{p(\boldsymbol{\psi})}\sqrt{1 - p(\boldsymbol{\psi})}} H(d\boldsymbol{\psi});$$

note that the right hand side of this expression depends only on the law of the random variable $Q := p(\boldsymbol{\psi})$. Some popular choices for the law of Q are discussed in Section 8.4 of McNeil, Frey & Embrechts [7].

3 Numerical Case Studies

In this section we test the numerical performance of the approximation proposed in the previous section. For this we compare the "true" distribution (computed by extensive standard Monte Carlo simulation) to the Vasicek approximation and the approximation which we propose in the current article. For the convenience of the reader we briefly describe the corresponding simulation algorithms in Section 3.1; numerical case studies are provided in Section 3.2; the application to CDO-tranches is discussed in Section 3.3.

The model. In our numerical analysis we will focus on a Gaussian one-factor Bernoulli mixture model. The underlying factor Ψ has a standard normal distribution, $\Psi \sim N(0, 1)$. We consider m obligors with unconditional default probabilities $\bar{p}_i$ and individual losses $l_i = e_i\delta_i(\cdot)Y_i$, $i = 1, ..., m$, where exposure e_i, loss given default $\delta_i(\cdot)$, and default indicators Y_i, $i = 1, ..., m$, are defined as before. The default indicators are constructed from the factors as follows. Let ϵ_i, $i = 1, 2, \ldots, m$, be independent standard normals which are independent of the factor Ψ. Setting

$$X_i = \sqrt{\rho} \cdot \Psi + \sqrt{1 - \rho} \cdot \epsilon_i, \qquad i = 1, 2, \ldots, m,$$

and $x_i = \Phi^{-1}(1 - \bar{p}_i)$, the default indicators

$$Y_i \;=\; 1_{\{X_i \geq x_i\}}, \qquad i = 1, 2, \ldots, m,$$

are conditionally independent given Ψ. The default probability of obligor i equals $\bar{p}_i$, and the corresponding conditional default probability $p_i(\psi)$ is given by

$$p_i(\psi) = P(Y_i = 1 | \Psi = \psi) = P(X_i \geq x_i | \Psi = \psi) = \Phi\left(\frac{\Phi^{-1}(\bar{p}_i) + \sqrt{\rho}\,\psi}{\sqrt{1 - \rho}}\right). \qquad (4)$$

3.1 The Algorithms

We now describe the algorithms for computing the conditional loss distribution used in the simulation study. In all simulations below the factor variable Ψ is not simulated; instead we apply numerical integration using the trapezoidal rule.

Vasicek method (first-order approximation). Vasicek's classical approximation is based on the law of large numbers. Under suitable conditions the average loss of a large portfolio can be approximated by the average conditional mean, i.e.

$$\lim_{m \to \infty} \frac{1}{m} L^{(m)} \;=\; \lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} \mathbf{E}\left[l_i | \Psi\right],$$

see Frey and McNeil [3]. Sufficient conditions are, e.g., that the exposures are bounded and that the sum on the right hand side converges.

A simple approximation of the conditional distribution of $L^{(m)}$ given $\Psi = \psi$ with $\psi \in \mathbb{R}$ is provided by the Dirac measure which is concentrated on the conditional mean $\sum_{i=1}^{m} \mathbf{E}\left[l_i | \Psi = \psi\right]$. The first-order approximation can thus be described as follows:

$$P[L^{(m)} > x | \Psi = \psi] \approx 1_{\left\{\sum_{i=1}^{m} e_i \delta_i(\psi) p_i(\psi) > x\right\}} = \begin{cases} 1 & \text{if } \sum_{i=1}^{m} e_i \delta_i(\psi) p_i(\psi) > x, \\ 0 & \text{otherwise} \end{cases}$$

The approximate unconditional probability $P[L^{(m)} > x]$ is obtained by integration over the factor Ψ.

Second-order Approximation. Proposition 2.2 in Section 2 leads to the following estimate for the conditional probability $P[L^{(m)} > x | \Psi = \psi]$ for fixed $\psi \in \mathbb{R}$:

Second-Order Algorithm:

1. Calculate $p_i(\psi) = \Phi\left(\frac{\Phi^{-1}(\bar{p}_i) + \sqrt{\rho}\,\psi}{\sqrt{1 - \rho}}\right)$.

2. Calculate the second order estimator for $P[L^{(m)} > x | \Psi = \psi]$:

$$
1 - \Phi \left(\frac{x - \sum_{i=1}^{m} e_i \, \delta_i(\psi) \, p_i(\psi)}{\sqrt{\sum_{i=1}^{m} e_i^2 \, \delta_i^2(\psi) p_i(\psi)(1 - p_i(\psi))}} \right).
$$

Again, the approximate unconditional probability $P[L^{(m)} > x]$ is obtained by integration over the factor Ψ.

Remark 3.1. A comparison of these algorithms explains our terminology "first-order" and "second-order": in the Vasicek method the conditional loss distribution is replaced by its mean and all randomness is due to fluctuations in Ψ. These constitute a first order effect, since they are of size $O(m)$. In contrast, in the second-order approximation also fluctuations in the conditional loss distribution – which are of order $O(\sqrt{m})$ – are taken into account via the normal approximation of the conditional loss distribution.

Both the first-order and second-order approximations rely on the evaluation of the Gaussian distribution function. This is computationally less demanding than Monte Carlo simulations. In the next section we provide numerical case studies and compare both precision and computational costs of the two methods and a Monte Carlo benchmark.

3.2 Numerical Results

We focus on two different applications. First we estimate probabilities of large portfolio losses which are important quantities for credit risk management. In this application, the probability measure P needs to be interpreted as the statistical measure. The second application is the calculation of CDO prices based on the portfolio loss distribution. In this situation, the probability measure P signifies a risk-neutral or pricing measure.

3.2.1 Numerical results for loss probabilities

We analyze the effect of three parameters, namely asset correlation ρ, default probabilities p_i, and portfolio size m.

The effect of the asset correlation ρ. The value of the parameter ρ determines the degree of dependency between different obligors: defaults are independent for $\rho = 0$; the larger ρ, the more dependent are the defaults. For varying ρ, we consider the first- and second-order approximation for portfolios of 200 obligors with identical annual default probabilities $p = 0.0112$. This value corresponds to BB-rated firms, see McNeil, Frey, and Embrechts [7].

Figure 1 displays the probabilities of exceeding a given loss amount as a function of this loss threshold level. For $\rho = 0$ the first-order approximation does not provide a reasonable estimate, whereas the second-order approximation still gives acceptable results. The larger ρ, the better are the first- and second-order approximations. The second-order approximation outperforms

the first-order approximation in most cases. This effect is most significant for small exceedance probabilities which correspond to larger threshold levels. The first-order approximation systematically understates the exceedance probabilities for large threshold levels, since it does not account for large values in the conditional loss distributions.

The impact of the parameter ρ itself on the accuracy of the approximations can be understood as follows. It is apparent from equation (4) that ρ governs the degree of dispersion of the random variable $Q_i := p_i(\Psi)$ around its mean $\bar{p}_i$: for small ρ the distribution of Q_i is very concentrated around $\bar{p}_i$, for larger values of ρ the distribution becomes more dispersed. This has implications for the overall unconditional loss distribution. If Q_i is very concentrated, almost all fluctuations of the loss distribution are due to fluctuations in the conditional loss distributions. If Q_i becomes more dispersed, the unconditional loss distribution is a mixture of the conditional loss distributions. The fluctuations of the conditional loss distributions become less important, while the influence of the factor distribution increases. This has the consequence that the accuracy of the approximation of the conditional loss distributions becomes less important (as long as it is unbiased), if we are interested in an approximation of the unconditional loss distribution.

The effect of $\bar{p}_i$. Again, we compare the algorithms for homogeneous portfolios of 200 obligors with identical exposures of $e_i = 1$. However, this time we keep $\rho = 0.054$ fixed and vary the value of the default probability. According to their rating classes, the individual default probabilities are taken from McNeil, Frey and Embrechts [7].

Rating class	$\bar{p}$
BB	0.0112
B	0.049
CCC	0.188

The results are displayed in Figure 2. The second-order approximation is better than the first-order approximation, in particular for low exceedance probabilities which correspond to large threshold levels. The first-order probability improves when the individual default probabilities are increased. It does still systematically understate the exceedance probabilities for large threshold levels, but the effect becomes less pronounced.

The effect of portfolio size m. For larger portfolios the difference between the first-order and the second-order approximation becomes less significant, although the second-order still gives slightly better results for low exceedance probabilities which correspond to larger threshold levels. This is illustrated in Figure 3 where we consider a portfolio of 2000 B-rated obligors. Since both the first-order and the second-order approximation converge to the true distribution as the size m of the portfolio, it is not surprising that their accuracy for a portfolio of 2000 obligors is extremely high.

Inhomogeneous portfolios. All numerical examples so far analyze the performance of the approximations for homogeneous portfolios. However, Figure 4 illustrates that the second-order approximation is also an adequate choice for inhomogeneous portfolios.

3.3 Application to synthetic CDO tranches

In the current section we apply the first-order and second-order approximations to collateralized debt obligations (CDO). Pricing tranches of CDOs requires a model for the cumulative loss process of a credit portfolio. A classical benchmark model which has been discussed in the literature is a Gaussian copula model which can equivalently be represented as a factor model; see for instance Section 9.7 of McNeil, Frey and Embrechts [7]. Here we focus on this particular model and investigate the accuracy of price estimates if we use the first- and second-order approximation to the portfolio loss distribution. Note however, that the second-order approximation for the conditional loss distribution can be applied to other factor copula models such as the double-t copula of Hull and White [5] or the NIG copula proposed by Kalemanova, Schmidt and Werner [6]. In the current context, the probability measure needs to be interpreted as a pricing or risk-neutral measure. In order to emphasize this fact we denote it by Q instead of P. All expectations are taken with respect to Q.

Theoretical background. A synthetic CDO tranche is based on a portfolio of m single-name credit default swaps on m different reference entities. The number of names m is typically equal to 125. The notional N of the CDO is the total exposure of the portfolio. A tranche is characterized by a maturity date T and lower and upper attachment points $0 \leq l \leq u \leq 1$ which are given as fractions of the notional of the CDO. The cumulative loss up to time t of the tranche $[l, u]$ is $L_t^{[l,u]} := (L_t - lN)^+ - (L_t - uN)^+$. Default and premium payments can conveniently be expressed in terms of the cumulative loss process.

At a time $\tau \leq T$ of the default of a name in the portfolio a default payment of size

$$\triangle L_\tau^{[l,u]} := L_\tau^{[l,u]} - L_{\tau-}^{[l,u]}$$

is made. Assuming that the short term interest rate is $(r(t))_{t \geq 0}$, the initial value of all default payments up to time T is given by

$$V_{[l,u]}^{\mathrm{def}} = E\left[\int_0^T \exp\left(-\int_0^t r(s)ds\right) dL_t^{[l,u]}\right].$$

To keep our analysis simple we assume that the interest rate is deterministic. Partial integration allows us to express the value of the default payments in terms of expectations of the loss process,

$$V_{[l,u]}^{\mathrm{def}} = \exp\left(-\int_0^T r(s)ds\right) E\left(L_T^{[l,u]}\right) + \int_0^T r(t)\exp\left(-\int_0^t r(s)ds\right) E\left(L_t^{[l,u]}\right) dt.$$

The premium payment leg consists of regular payments[2] at fixed future dates

$$t_1 < ... < t_N = T.$$

Given a spread x and setting $t_0 = 0$, the value of the regular premium payments equals

$$V_{[l,u]}^{\mathrm{prem}}(x) = x \sum_{n=1}^{N} (t_n - t_{n-1}) \exp\left(-\int_0^{t_n} r(s)ds\right) \left[(u-l)N - E\left(L_{t_n}^{[l,u]}\right)\right].$$

The fair tranche spread $x^{[l,u]}$ is then determined by equating the value of default and premium payments,

$$V_{[l,u]}^{\mathrm{def}} = V_{[l,u]}^{\mathrm{prem}}(x^{[l,u]}).$$

If we assume in addition that default can only occur at the dates $t_1 < ... < t_N$, then both sides of the equation can be expressed as functions of

$$E\left(L_t^{[l,u]}\right) = E\left((L_t - lN)^+ - (L_t - uN)^+\right), \qquad t = t_1, \ldots, t_N. \tag{5}$$

In the context of the Gaussian copula model which is specified below these expectations can be estimated on the basis of the first-order and second-order approximations.

The model. CDO pricing requires a dynamic model. However, as we have seen, the fair spread $x^{[l,u]}$ can be calculated, if the finite number of expectations in display (5) can be evaluated. This can be done, if the loss distributions L_t are specified for each date $t = t_1, \ldots, t_N$.

To be more precise, denote by τ_i the default time of firm i. We assume that defaults are independent conditional on a factor variable Ψ which is assumed to be standard normal. The risk-neutral conditional default probabilities at times t are given by

$$Q(\tau_i \le t | \Psi) = \Phi\left(\frac{\Phi^{-1}(F_i(t)) + \sqrt{\rho}\Psi}{\sqrt{1-\rho}}\right),$$

where $t \mapsto F_i(t) = Q(\tau_i \le t)$ is the distribution function of the default time τ_i. In a constant intensity framework, we have $F_i(t) = 1 - e^{-\lambda_i t}$ where λ_i is the risk-neutral default intensity for firm i. In this model, defaults can occur at any point in time, not only at the dates $t = t_1, ..., t_N$. For simplicity, however, we approximate τ_i by t_n whenever $\tau_i \in (t_{n-1}, t_n]$, $n = 1, \ldots, N$. Since the computation of CDO-spreads can then be reduced to evaluating the distribution of L_t for $t = t_1, ..., t_N$, we are faced with the same problem as in the evaluation of loss distributions.

Numerical results. For our numerical experiments we choose the following parameters: identical exposures $e_i = e = 1$, a constant percentage loss given default $\delta_i = \delta = 0.6$, maturity $T = 5$, identical default intensities $\lambda_i = \lambda = 0.007$, and $r = 0$. ρ represents the implied tranche

---

[2]In practice, there is moreover an accrued payment after default which is ignored for simplicity.

correlation. Implied tranche correlation[3] is chosen differently for every tranche to match market data (observed CDO-spreads) from August 4, 2004, see Hull and White [5].

The following table summarizes our calculations of the CDO-spreads. The fair tranche spread is given for each tranche and the corresponding value of implied tranche correlation ρ. The calculations of the fair spread is based on the formulas discussed above and requires a characterization of the loss distribution for every quarter. The true spread is obtained from Monte Carlo simulation. We compare this value to results which we obtain from the first-order and second-order approximation. Note, that the value for the equity tranche ($[0,3]$) corresponds to an upfront payment on the notional; the running spread is set to 5% by market convention. Levels for all other tranches are with no fixed running spread.

tranche	$[0,3]$ $\rho = 0.219$	$[3,6]$ $\rho = 0.042$	$[6,9]$ $\rho = 0.148$	$[9,12]$ $\rho = 0.223$	$[12,22]$ $\rho = 0.305$
first-order	30.66%	0.79%	0.53%	0.36%	0.18%
second-order	29.38%	1.51%	0.66%	0.42%	0.20%
true value	28.38%	1.55%	0.68%	0.42%	0.20%

We find that for all tranches the second-order approximation is a significant improvement compared to first-order approximation and even attains the true spread for the two most senior tranches. However, while first-order approximation gives poor results for the $[3,6]$ and the $[6,9]$ tranche, it performs better for the two most senior tranches. One can explain this by the fact that a loss of more than 10% of the total exposure is only possible for large Ψ, i.e. the losses in these tranches are driven by the factor risk and the approximation technique for the conditional loss distribution does not have much influence so that the results for all three methods are relatively close. This is in line with the findings from Section 3.2.1. Overall the second-order approximation outperforms the Vasicek approximation in the CDO setting.

3.3.1 Analysis of Computational Effort and Theoretical Error

In this section we analyse the computational effort of first- and second-order approximation. Furthermore, for the case of identically distributed losses we compare the actual numerical difference between the true distribution and the results obtained by second-order approximation with the theoretical error given by the Berry-Esseen inequality.

Computational Effort We compare the computing times for both first- and second-order approximation. In our implementation we obtain the following values for the calculation of $P[L^{(m)} > x]$ and of a CDO spread using the different methods:

computing time [seconds]	first-order	second-order
calculation of $P[L^{(m)} > x]$	0.088	0.090
calculation of CDO spread	145	149

---

[3]Implied tranche correlation is a convenient way to quote prices. It is the credit risk analogue of implied volatility in the equity world. Like the Black-Scholes model, the Gaussian copula model is merely a transformation tool which should, of course, not be interpreted as a realistic default model.

It is not surprising that first-order approximation is faster than second-order approximation, since the calculation of the first-order estimator is a bit simpler. However, if an efficient implementation for the normal distribution is used, the computation times for both techniques are really close. The time-consuming parts of the algorithm are integration over the underlying factor Ψ and the calculation of the estimator for all times $t_i, i = 1, ..., N$ (in the CDO case). Those steps have to be done for both algorithms. Note that in order to reach a similar level of precision standard Monte Carlo simulation would require a multiple of the times above.

Theoretical Error In the case of iid losses the theoretical error of second-order approximation is given by $\frac{A}{\sqrt{m}} E\left[\frac{\bar{\gamma}(\psi)}{\bar{\sigma}^3(\psi)}\right]$, see Corollary 2.5. According to Remark 2.6 we choose $A = 0.7655$ and want to compare the theoretical error to the actual numerical difference between second-order results and the true distribution. The calculation of the error is straightforward, we display it in the first column of the following table. In the second column we list the maximum difference between the second-order estimator $G^{(m)}(x)$ and the true distribution for various portfolios.

	theoretical error	max difference (computation)
200 CCC obligors	0.1944735	0.0130957
200 B obligors	0.5132462	0.0432077
200 BB obligors	1.2299930	0.1441425

We notice that the actual deviation from the true value is much smaller than the theoretical error. This stems from two reasons: first, the constant which we use in the Berry-Esseen theorem is not optimal, as discussed in Remark 2.6; second, the application of Jensen's inequality gives a very rough estimate in the proof of Proposition 2.2.

4   Conclusion

We have introduced a second-order approximation for estimating the distribution of portfolio losses. Compared to first-order approximation it provides a significant improvement in accuracy while it is easy to implement and much faster when compared to standard Monte Carlo. It is most useful for the estimation of small exceedance probabilities ($< 10\%$) for portfolios with less than 2000 obligors when asset correlations or default probabilities are low.

References

[1] Van Beek, P. (1972): "An application of Fourier methods to the problem of sharpening the Berry-Esseen inequality". *Z. Wahrscheinlichkeitstheorie verw. Geb.* 23:187-196.

[2] Esseen, C. G. (1956): "A moment inequality with an application to the central limit theorem". *Scand. Aktuarietidskr.* 39:160-170.

[3] Frey, R., and A. McNeil (2003): "Dependent defaults in models of portfolio credit risk". *Journal of Risk* 6(1):59–92.

[4] Glasserman, P. and Ruiz-Mata, J. (2006): 'Computing the credit loss distribution in the Gaussian copula model: a comparison of methods'. *The Journal of Credit Risk* 2(4), 33–66.

[5] Hull, J., and A. White (2004): "Valuation of a CDO and a nth to default CDS without Monte Carlo simulation". *Journal of Derivatives* 12, 8-23.

[6] Kalemanova, A., Schmid, B. and Werner, R. (2005), "The normal inverse gaussian distribution for synthetic CDO pricing", working paper, Risklab Germany.

[7] McNeil, A., R. Frey, P. Embrechts (2005): *Quantitative Risk Management.* Princeton University Press.

[8] Petrov, V. V. (1975): *Sums of Independent Random Variables.* Springer.

[9] Shiganov, I. S. (1986): "Refinement of the upper bound of the constant in the central limit theorem". In: *Problems of stability of stochastic models*, VINITI, Moscow, 109-115.

A   Appendix

All figures display tail probabilities, i.e. the probabilities of exceeding a given loss amount as a function of this loss threshold level. The x-axis represents the loss threshold in percent of total exposure; the y-axis represents the probability that the portfolio losses exceed the given threshold level. This probability is displayed on a logarithmic scale for greater clarity.
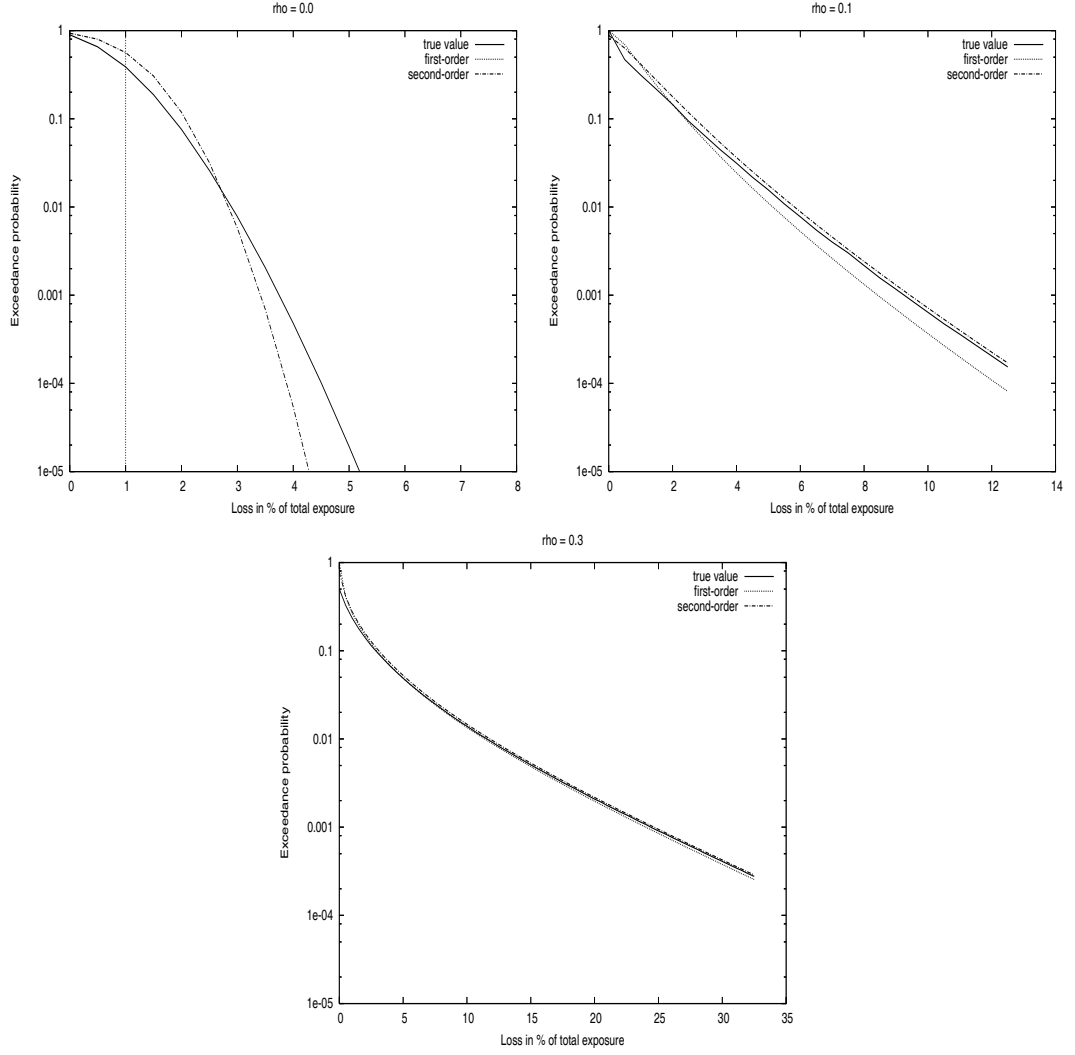
Figure 1: Impact of varying ρ.
For $\rho = 0$ the obligors are independent and the first-order approximation does not provide a reasonable estimate, whereas the second-order approximation still gives acceptable results. The larger ρ, the better are the first- and second-order approximations. The second-order approximation outperforms the first-order approximation in most cases. This effect is most significant for small exceedance probabilities (which correspond to larger threshold levels).
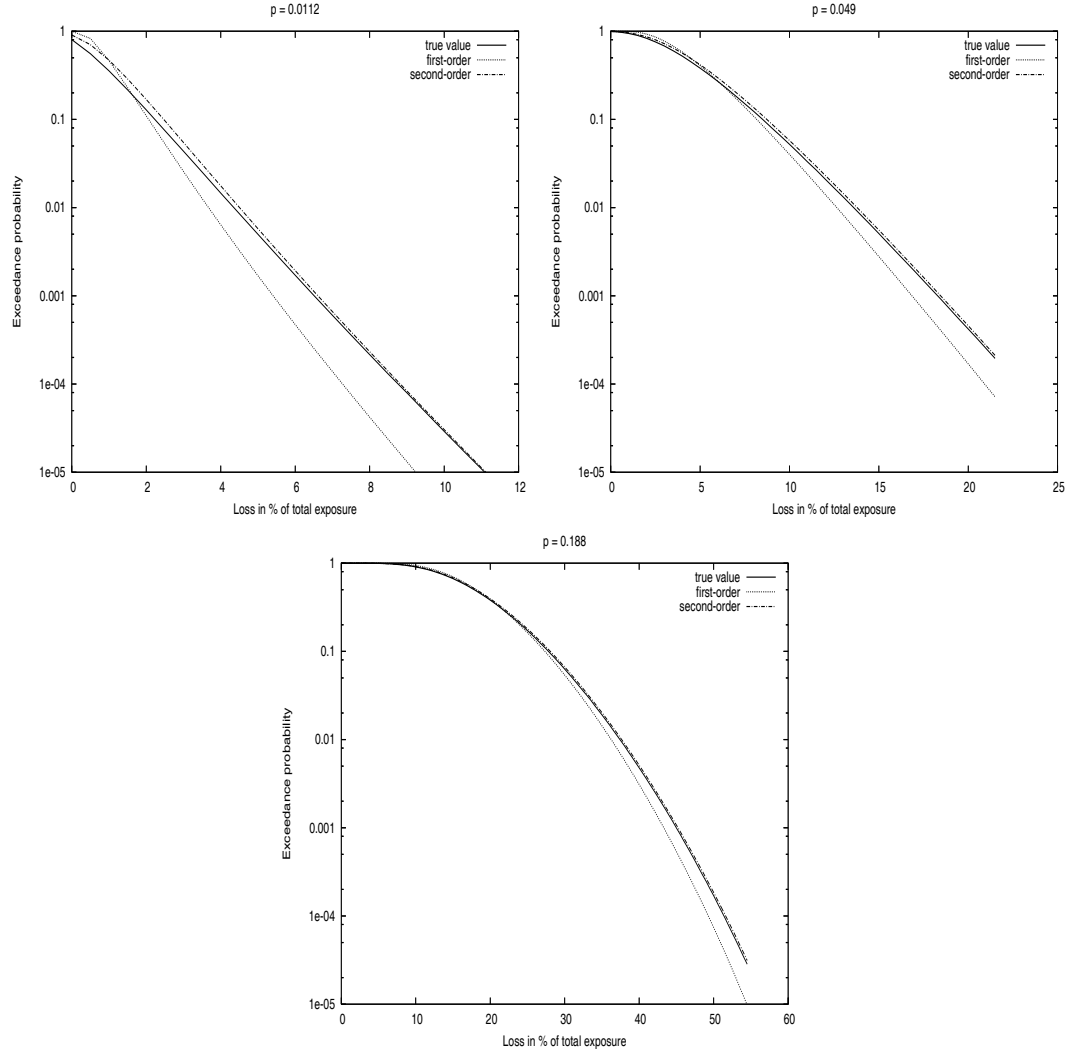
15

Figure 2: Impact of varying $\bar{p}$.
The second-order approximation is better than the first-order approximation, in particular for low exceedance probabilities (which correspond to large threshold levels). The first-order probability improves when the individual default probabilities are increased.
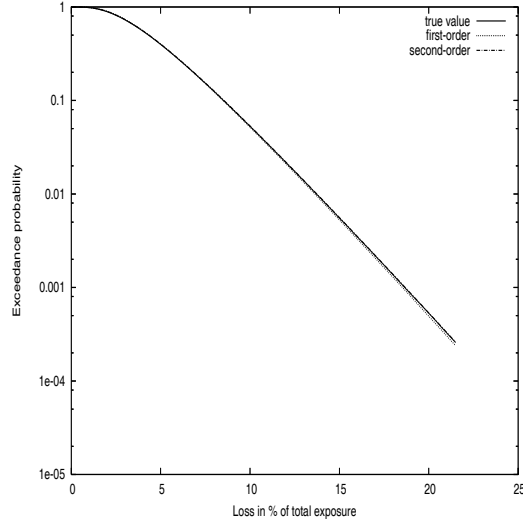
Figure 3: Impact of portfolio size.
For a portfolio of 2000 obligors with identical default probability of $\bar{p} = 0.049$ the results of both first- and second-order approximation nearly reach the true distribution. The second-order still gives slightly better results for low exceedance probabilities (which correspond to large threshold levels).
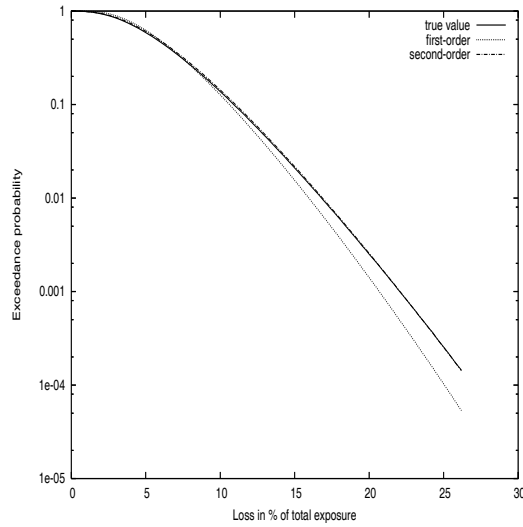

Figure 4: Impact of heterogeneity.
For an inhomogeneous portfolio of 40 BB-rated obligors with identical exposures $e_i = 5$, 60 B-rated obligors with identical exposures $e_i = 2$ and 100 CCC-rated obligors with identical exposures $e_i = 1$ the second-order approximation performs better than the first-order approximation. The effect is most significant for low exceedance probabilities (which correspond to large threshold levels).

17