Available online at www.sciencedirect.com
ScienceDirect
Journal of Banking & Finance 31 (2007) 1375–1398
Journal of BANKING & FINANCE
www.elsevier.com/locate/jbf

Correlation expansions for CDO pricing

Paul Glasserman *, Sira Suchintabandid *

Columbia Business School, 4R Uris Hall, New York, NY 10027, United States

Available online 19 December 2006

Abstract

This paper develops numerical approximations for pricing collateralized debt obligations (CDOs) and other portfolio credit derivatives in the multifactor Normal Copula model. A key aspect of pricing portfolio credit derivatives is capturing dependence between the defaults of the elements of the portfolio. But, compared with an independent-obligor model, pricing in a model with correlated defaults is more challenging. Our approach strikes a balance by reducing the problem of pricing in a model with correlated defaults to calculations involving only independent defaults. We develop approximations based on power series expansions in a parameter that scales the underlying correlations. These expansions express a CDO tranche price in a multifactor model as a series of prices in independent-obligor models, which are easy to compute. The approach builds on a classical approximation for multivariate Gaussian probabilities; we introduce an alternative representation that greatly reduces the number of terms required to evaluate the coefficients in the expansion. We also apply this method to the underlying problem of computing joint probabilities of multivariate normal random variables for which the correlation matrix has a factor structure.

JEL classification: G 13

Keywords: CDO's; Copulas; Factor models; Multifactor; Default risk

1. Introduction

The market for multi-name (portfolio) credit derivatives has been one of the fastest growing derivatives markets over the past decade. While credit default swaps have

* Corresponding authors.
E-mail address: pg20@columbia.edu (P. Glasserman).

0378-4266/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2006.10.018

facilitated trading in individual sources of credit risk, collateralized debt obligations (CDOs) and other multi-name credit derivatives have provided new mechanisms for the transfer of credit risk in an entire portfolio.

An important aspect of the valuation and risk management of multi-name credit derivatives is the modelling of dependence among sources of credit risk. On one hand, risk models that ignore this dependence and presuppose that obligors default independently are very tractable, but they omit a crucial feature of a portfolio view of credit risk. On the other hand, capturing the dependence among obligors comes at a price of increased complexity, both in the modelling and in the computation of the model's output, be it the price of a CDO, the value-at-risk of a portfolio, or the price of some other basket credit derivative.

Factor models of dependence – in which defaults become independent conditional on a set of underlying factors – lie at an intermediate level of complexity between independent-obligor models and models that allow arbitrary dependence. In particular, in a single-factor model, CDO prices can be computed efficiently using semi-analytical methods and numerical integration (see, for example, Andersen et al., 2003; Laurent and Gregory, 2003). For more general dependence structures for which no low-factor model provides an accurate fit, however, computing CDO prices (or other credit derivative prices) generally requires Monte Carlo simulation. This paper addresses models in which the number of underlying factors is substantially smaller than the number of names in the portfolio, but not so small as to allow use of the types of semi-analytic methods used in the single-factor models.

We work within the multifactor Normal Copula framework. We express a CDO tranche price in a dependent-obligor model as a series of easy-to-compute prices in independent-obligor models. To be more specific, we scale the obligors' correlation matrix and expand the desired tranche price as a power series of the scaling parameter. We then show that each term in this expansion can be expressed as a weighted finite sum of independent-obligor prices. Thus, approximating prices in a multifactor model reduces to a series of calls to an independent-obligor pricing routine. This method takes full advantage of the fact that independent-obligor prices are quick and easy to compute. We also show how the proposed approximation can be applied to compute quantities related to sensitivity analysis.

Our approach builds on Kibble's (1945) generalized tetrachoric series for multivariate normal random variables, in which a joint probability of correlated normal random variables is expanded as an infinite series of correlation coefficients. While various methods exist for special cases of the multivariate normal (e.g., bivariate (Ağca and Chance, 2003; Vasicek, 1998), trivariate (Kendall, 1941), equal correlation (Somerville, 1954)), Kibble's approximation is, in principle, general. However, Kibble's series in its original form contains a large number of terms and is computationally impractical. The approach we develop for approximating CDO prices also leads to an alternative and more efficient way for evaluating Kibble's series when the underlying correlation matrix has a factor structure. Indeed, although we focus on computing CDO prices, the same approach can be used to approximate the expectations of other functions of the loss distribution in the Normal Copula model.

The rest of this paper is organized as follows. Section 2 provides some background on CDOs. Section 3 reviews the Normal Copula model for credit risk. Section 4 describes the approximation regime for CDO prices for portfolios with weakly-correlated obligors. Section 5 describes how to extend the result from Section 3 to the case of strong correlation. In Section 6, we discuss the problem of sensitivity analysis. All technical details are

deferred to the appendices. The main result is stated formally in Appendix A and its proof is given in Appendix B.

2. Background on CDOs

2.1. Basic structure

Collateralized debt obligations (CDOs) have become an increasingly popular class of securities to emerge in the credit derivatives market. The CDO technology offers a way to redistribute the credit risk of a pool of debt instruments and create a family of securities with widely different risk profiles, ranging from very safe to highly speculative. For a comprehensive background on CDOs, see for example Hull and White (2004) or Schönbucher (2003). We hereafter describe the fundamental idea. In its most basic form, a CDO structure may be explained through Fig. 1.

A CDO structure is constructed based on a reference pool of assets, which can consist of a diversified group of debt instruments – let us consider them bonds for the purpose of our illustration. The number of obligors generally varies between 50 and 150. These bonds are subjected to credit risk; in other words, they may default – a situation that, depending on the recovery rate after default, can result in minor or major losses within the collateral pool.

CDO tranches offer the opportunity to buy and sell protection from certain a fraction of these losses. In Fig. 1, a CDO structure is created by "tranching out" credit losses on the reference pool into three types of securities: (1) an equity tranche, whose upper attachment point is 7% of the maximum loss in the reference portfolio; (2) a mezzanine tranche, whose attachment points are 7–25%; and (3) a senior tranche, whose lower attachment point is 25%.

The equity tranche absorbs the first 7% of the losses in the underlying pool. The holders of the equity tranche are paid a fixed rate specific to this tranche – 35% according to Fig. 1 – based on the outstanding nominal on the tranche. That is to say, the equity tranche holders initially are paid a return of 35% of the whole amount they invested (7% of the total bond principal). But suppose that, at time T_1, losses of 2% have been incurred (the grey stripe in Fig. 1), then at the subsequent coupon date the equity tranche holders will

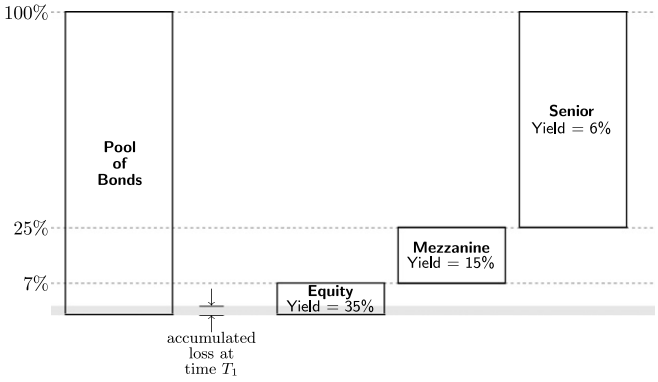

Fig. 1. Basic CDO structure.

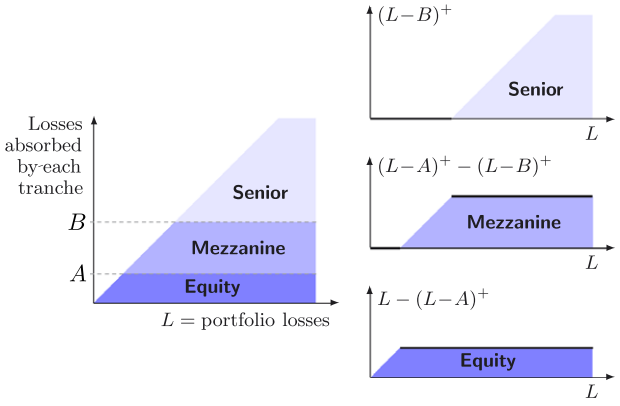

Fig. 2. Losses absorbed by each tranche.

be paid 35% of the remaining tranche notional (the unflooded portion in the equity tranche).

Similarly, the mezzanine tranche periodically receives coupon payments of 15% on the stochastically decaying tranche notional (initially set to $25\% - 7\% = 18\%$ of the total bond portfolio) until the losses reach 25% of the portfolio principal, at which point the contract is written down. The senior tranche is responsible for all losses in excess of the 25% absorbed by the equity and mezzanine tranches and, being the safest tranche, receives the lowest coupon rate among the three tranches. Fig. 2 displays the loss absorbed by each tranche as a function of the aggregated loss in the reference collateral pool.

The CDO structure in Fig. 1 is referred to as a *cash CDO*. There are natural extensions to this idea – liability structures with more (or fewer) than three tranches, combinations of fixed- and floating-coupon assets, and liability asset classes other than bonds. An increasingly popular variation of the CDO technology is the synthetic CDOs, which are very similar to cash CDOs, except that the bonds are replaced by credit default swaps; the arranger of a synthetic CDO passes the default risk on to the CDO tranches, thereby allowing investors to participate at different risk levels.

2.2. Portfolio loss approach in pricing CDOs

This paper focuses on pricing the CDO tranches using the so-called *portfolio loss approach*, in which all cashflows during the life of the CDO are expressed in terms of the *cumulative* losses in the reference portfolio at prespecified coupon dates. Thus, the price of a CDO tranche can be written as an expectation with respect to the distribution of the underlying portfolio loss process. The detailed explanation of this approach can be found, for example, in Andersen et al. (2003) and Laurent and Gregory (2003). We now briefly explain the approach.

From the payoff scheme described in the previous section, one can see that the value of a CDO tranche equals the default-free cashflows (calculated as a fixed coupon rate on the entire tranche notional) less the losses resulting from defaults in the reference pool. The fixed (default-free) leg is non-stochastic, and is valued like a stream of fixed cashflows. The main issue, therefore, is valuing the default payment leg.

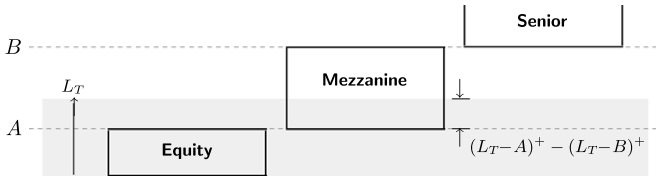

Fig. 3. Default losses absorbed by the mezzanine tranche at an arbitrary time T.

Let us decompose the default payment leg by considering the contribution from each coupon period. Suppose that up to the time of the ith coupon payment T_i, the reference pool has accumulated default losses of L_{T_i}. Then, by time T_i, the tranche notional will have depleted by $(L_{T_i} - A)^+ - (L_{T_i} - B)^+$, where A and B are the lower and upper attachment points, respectively (see Fig. 3). Hence, the expected loss to be experienced by holders of this tranche during the ith coupon period is $E(L_{T_i} - A)^+ - E(L_{T_i} - B)^+$ times the coupon rate. Discounting this amount to time zero and summing over all coupon periods during the life of the CDO gives the value of the default payment leg.

Thus, the problem of pricing a CDO tranche is reduced to that of evaluating expectations of the form $E(L - y)^+$, where L is the cumulative loss in the underlying portfolio at a specific coupon date, and y is an attachment point. In this paper, the snapshot of the loss portfolio L at a given coupon date is modelled within the normal copula framework, described in the following section. We then spend the rest of the paper discussing the computation of $E(L - y)^+$.

3. Normal copula model for credit risk

In this section, we describe the widely used Normal Copula model initiated by Li (2000), which has become an industry standard for pricing. We consider a CDO constructed from a reference portfolio of M obligors. As explained in the previous section, in order to value a CDO contract we only need the marginal distribution of the aggregated loss of the reference portfolio at coupon dates. Let L be the accumulated portfolio loss at a fixed coupon date. To specify the distribution of L, we introduce the following notation:

$$Y_i = \text{default indicator for } i\text{th obligor} = \begin{cases} 1 & \text{if } i\text{th obligor defaults,} \\ 0 & \text{otherwise;} \end{cases}$$

$p_i = P(Y_i = 1) = \text{marginal probability that } i\text{th obligor defaults;}$

$c_i = \text{loss from the default of } i\text{th obligor.}$

The marginal default probabilities p_i are assumed known (e.g., from credit spreads for the prices of credit default swaps). The losses given default c_i are assumed known and constant. Then, the portfolio loss at the given coupon date is

$$L = c_1 Y_1 + c_2 Y_2 + \cdots + c_M Y_M. \tag{3.1}$$

To model dependence among obligors we need to introduce dependence among the default indicators $Y_1, \ldots, Y_M$. In the normal copula model, dependence is introduced by letting

$$Y_i = 1\{X_i > v_i\}, \quad i = 1, 2, \ldots, M, \tag{3.2}$$

where $(X_1, X_2, \ldots, X_M)$ are correlated $\mathcal{N}(0,1)$ random variables. The "default boundary" v_i is chosen to match the default probability, that is

$$p_i = \overline{\Phi}(v_i). \tag{3.3}$$

where $\overline{\Phi}(x) = 1 - \Phi(x)$, with Φ the cumulative normal distribution.

While the correlation among the X_i may be specified by their correlation matrix, it is a common practice to introduce the correlation structure through a set of "factors" $Z_1, Z_2, \ldots, Z_d$ $(d \ll M)$,

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_M \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{M1} \end{bmatrix} Z_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{M2} \end{bmatrix} Z_2 + \cdots + \begin{bmatrix} a_{1d} \\ a_{2d} \\ a_{3d} \\ \vdots \\ a_{Md} \end{bmatrix} Z_d + \begin{bmatrix} b_1\epsilon_1 \\ b_2\epsilon_2 \\ b_3\epsilon_3 \\ \vdots \\ b_M\epsilon_M \end{bmatrix}. \tag{3.4}$$

Here, ϵ_i and $Z_1, \ldots, Z_d$ are independent $\mathcal{N}(0,1)$ random variables, a_{ij} and b_i are constants such that $a_{i1}^2 + \cdots + a_{id}^2 + b_i^2 = 1$ so that X_i is $\mathcal{N}(0,1)$. Each ϵ_i represents risk affecting only the ith obligor, while Z_j's are risk factors that affect more than one obligor. The factors Z_j are sometimes given economic interpretations (as industry or regional risk factors, for example). The matrix $\mathbf{A} = [a_{ij}]_{M\times d}$ is the *loading matrix*, and it completely determines the correlation matrix $\mathbf{C} = [\rho_{ik}]$,

$$\rho_{ik} = \sum_{j=1}^{d} a_{ij}a_{kj}. \tag{3.5}$$

In this case, $\mathbf{C}$ is said to be *of d-factor structure*.

We are interested in expectations of the form $E(L - y)^+$. This expectation may be viewed as the default payment leg of a single-period CDO tranche with lower attachment point y. As explained in the previous section, the price of a general CDO tranche (multiple coupon dates with both upper and lower attachment points) can be expressed as a linear combination of expectations of the form $E(L - y)^+$.

Our objective is to propose a deterministic method for approximating $E(L - y)^+$. In Section 4, we first present an approximation scheme for the case where the correlation ρ_{ij}'s are weak. In Section 5, we modify the method to overcome the case of strong correlations.

4. Approximation for weak correlation

4.1. Approximating $E(L - y)^+$

In this section, we decompose the tranche price $E(L - y)^+$ in a dependent-obligor model into an infinite sum of independent-obligor tranche prices. This expansion suggests a way to compute the tranche price.

We now describe our method. We begin by parameterizing $\mathbf{C}$, the correlation matrix of $X_1, \ldots, X_M$ in (3.4). For $t \in [0,1]$, let

$$\mathbf{C}_t = \begin{bmatrix} 1 & t\rho_{12} & t\rho_{13} & \cdots & t\rho_{1M} \\ t\rho_{21} & 1 & t\rho_{23} & & t\rho_{2M} \\ t\rho_{31} & t\rho_{32} & 1 & & t\rho_{3M} \\ \vdots & & & \ddots & \\ t\rho_{M1} & t\rho_{M2} & t\rho_{M3} & & 1 \end{bmatrix}. \tag{4.6}$$

Let E_t denote the expectation under which $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_t)$. Note that $t = 1$ corresponds to the original model, while $t = 0$ corresponds to an independent-obligor model. So, as t varies from 0 to 1, $E_t(L - y)^+$ is the progression of the tranche price from that of the independent-obligor case to the desired price of the given model. It can be shown (see Theorem 2 in Appendix A) that $E_t(L - y)^+$ is analytical in t. That is,

$$E_t(L - y)^+ = \delta_0 + \delta_1 t + \delta_2 \frac{t^2}{2!} + \cdots, \tag{4.7}$$

where $\delta_0, \delta_1, \ldots$ are real scalars. The key finding, as stated in Theorem 1 in Appendix A, is that each δ_i can be computed through a weighted finite sum of independent-obligor tranche prices. Thus, (4.7) reduces tranche pricing in a correlated model to simpler calculations in the independent-obligor case. We now give an outline of the result, leaving the technical details to the appendix.

According to Theorem 1, δ_n can be computed through a weighted finite sum of the form

$$\sum_J w_J \widetilde{E}_J(L - y)^+. \tag{4.8}$$

The number of terms in the summation and the corresponding set of weights w_J's depend on n and the number of factors d in the correlation structure. In each term, $\widetilde{E}_J(L - y)^+$ is the expectation under which all obligors of L are independent, with modified default probability $\tilde{p}_1^{(J)}, \ldots, \tilde{p}_M^{(J)}$. These probabilities are perturbations of the original $p_1, \ldots, p_M$. (See (A.26) for the perturbation formula.) The magnitude of perturbation is determined by a real variable s. As s approaches zero, the weighted sum (4.8) converges to δ_n,

$$\sum_J w_J \widetilde{E}_J(L - y)^+ \overset{s \downarrow 0}{\to} \delta_n. \tag{4.9}$$

This results in a method for computing the coefficient δ_n. In a perfect computer (no round-off error), we simply set s to be infinitesmally small, then compute the weighted sum (4.8) to get δ_n. In practice, we have found that the convergence (4.9) is quite fast, so a moderately small s yields a fairly accurate value for δ_n.

Note that since every coefficient in (4.7) can be expressed as a weighted sum of independent-obligor expectations $\widetilde{E}_J(L - y)^+$, the effect of (4.7) is that it expands the tranche price of a dependent-obligor model into a series of prices of independent-obligor models.

Now let us discuss the issue of computing time. To compute the nth-order approximate for $E_t(L - y)^+$, we must compute $\delta_0, \delta_1, \ldots, \delta_n$. Most of the computational cost is spent in computing expectations of the form $\widetilde{E}_J(L - y)^+$ in (4.9). In Appendix A, we show that the number of terms of the form $\widetilde{E}_J(L - y)^+$ involved in computing $\delta_0, \ldots, \delta_n$ is

$$\binom{n + 2d}{n},$$

Table 1
Computing time for the expansion (4.7)

	$d=3$	$d=5$	$d=7$	$d=10$
1st order	7	11	15	21
2nd order	28	66	120	231
3rd order	84	286	680	1771
4th order	210	1001	3060	10,626

One unit of computing time equals the amount of time it takes to compute an independent-model price.

where d is the number of factors in the correlation matrix. Table 1 shows the computing time for computing nth order approximant for some values of d.

The duration of one unit of computing time depends on the method and computing environment employed to compute independent-obligor tranche prices. In our numerical experiments, we use the recursive algorithm presented in Andersen et al. (2003) which builds the loss distribution recursively by adding one obligor per iteration. This recursion is particularly attractive for us because it gives the *exact* value of $\widetilde{E}_J(L-y)^+$ for any given independent-obligor loss portfolio L; this ensures that the summation (4.9) converges to the right limit when we send the perturbation parameter s to zero. Straightforward implementation of this recursion on MATLAB takes about 0.6 s for a loss portfolio with 200 obligors.

Since our approximation is based on the power series expansion (4.7) of a parameter that scales the correlation matrix, the magnitude of ρ_{ij}'s primarily determines the order n at which the expansion becomes accurate. The stronger the correlation among obligors, the more terms we need in the expansion. Therefore the approximation becomes impractical if the X_i are highly correlated. In the next section, we will discuss how to deal with heavier correlation.

The approximation is appropriate when the number of factors d is of moderate size. If d is small ($d \leqslant 2$, say), then numerical integration may be a better option. If $d=3$, then 210 units of time allow us to compute up to a 4th-order approximation, which can give good results. In the same amount of time, one can do 3-dimensional numerical integration with only $\sqrt[3]{210} \approx 6$ points per dimension, which would give poor results.

If d is too large, then the approximation becomes impractical because the number of terms in the summation (4.9) is not manageable. In this case, a remedy is to fit the correlation matrix $\mathbf{C}$ with a lower factor structure. A 10-factor structure may suffice for mimicking empirical correlation matrices.

To demonstrate the implementation of Theorem 1, we present a first example. A more realistic example will be given in the next section when we allow heavier correlation.

Consider a portfolio of $M=50$ names. The default probabilities are $p_i=0.02$ for all obligors. The exposure of obligor i is given by $c_i=i$, $i=1,\ldots,M$. The loading matrix $\mathbf{A}=[a_{ij}]_{50\times 5}$ is a sparse matrix whose nonzero elements are all equal to 0.2. The nonzero elements in the first column are $a_{1,1},\ldots,a_{12,1}$; in the 2nd column $a_{9,2},\ldots,a_{22,2}$; in the 3rd column $a_{19,3},\ldots,a_{32,3}$; in the 4th column $a_{29,4},\ldots,a_{42,4}$; and in the 5th column $a_{39,5},\ldots,a_{50,5}$. This factor structure results in sparse (and light) correlation matrix.

Suppose we want to approximate $E_t(L-y)^+$ at $y=200$ (this is about 16% the total exposure), using no more than 500 units of computing time. According to Table 1, this allows us to expand up to 3rd order. Then, the summation (4.9) will be over 286 different

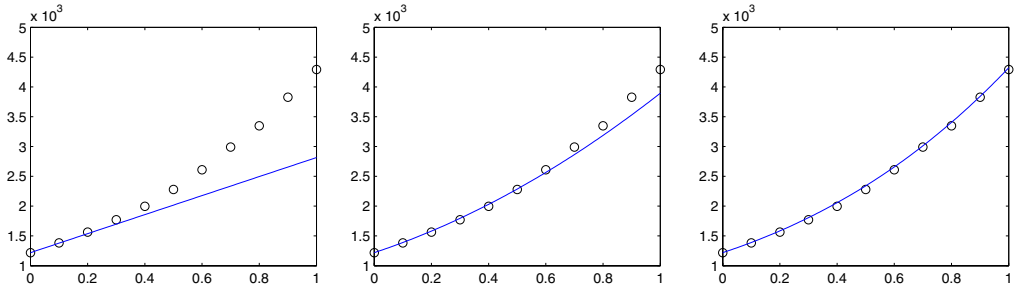

Fig. 4. Comparing first, second, and third order approximation.

values of $E_J(L-y)^+$, each of which is computed using independent-model pricing routine with perturbed probabilities $\tilde{p}_i^{(J)}$ from formula (A.26). In this example, we use perturbation parameter $s = 0.1$.[1] Then the δ_i are obtained from averaging these prices with different sets of weights. Their values are as follows.

δ_0	δ_1	δ_2	δ_3
.0012	.0016	.0022	.0026

Therefore, the 3rd-order approximation for $E_t(L-200)^+$ is

$$(1.2 + 1.6t + 2.2t^2/2! + 2.6t^3/3!) \times 10^{-3}.$$

Fig. 4 compares the first, second, and third order approximation. The true values are represented with "$\circ$". These are computed using Monte Carlo simulation. One can see that at $t = 1$, the 3rd-order approximation is quite accurate.

4.2. Approximating other expectations

It is worth noting that, while we have focused only on approximating $E_t(L-y)^+$, the same method can be used to approximate $E_t[f(L)]$, where f is an arbitrary function of the loss portfolio. We can show (see Theorem 2 in Appendix A) that for any $f: \mathbb{R} \to \mathbb{R}$, $E_t[f(L)]$ admits an infinite-series expansion in t. (We do not need to impose any regularity conditions on f because L has finite support.) That is,

$$E_t[f(L)] = \mu_0 + \mu_1 t + \mu_2 \frac{t^2}{2!} + \cdots,$$

where $\mu_0, \mu_1, \ldots$ are real constants. Analogous to (4.9), each μ_k can be calculated through a weighted finite sum, each of whose terms involves an expectation $\widetilde{E}_J[f(L)]$ under which all

[1] The parameter s may be taken arbitrarily small, subject only to the limits of machine precision. In our numerical examples, we have found that smaller values of s produce the same δ_i as $s = 0.1$, so we use this value consistently.

obligors of L are independent. For the formal statement of this result, see Theorem 2 in Appendix A.

This opens a wide range of application. For example, by choosing $f(L) = 1\{L > y\}$, Theorem 2 can be used to approximate the loss distribution and thus to calculate value-at-risk. This method can also be used to approximate joint probabilities of correlated normal random variables (see Proposition 2 in Appendix B).

5. Extension to strong correlation case

The approximation scheme described in the previous section performs well when the correlation among obligors is not too strong. This section shows how to deal with the case of heavy correlation.

The main modification lies in the parameterization scheme. Instead of parameterizing $\mathbf{C}$ so that $t = 0$ corresponds to an independent-obligor model as in (4.6), we now parameterize $\mathbf{C}$ so that $\mathbf{C}_0$ corresponds to a reference correlation matrix $\mathbf{R}$:

$$\mathbf{C}_t := (1-t)\mathbf{R} + t\mathbf{C}, \quad t \in [0,1], \tag{5.10}$$

where $\mathbf{R}$ has an r-factor structure. We will discuss the issue of selecting $\mathbf{R}$ later; at this point, we regard $\mathbf{R}$ as an r-factor approximation of $\mathbf{C}$. It is important that r should be significantly smaller than d – the number of factors in the original correlation matrix $\mathbf{C}$. (Remember, if d is already small, then there is little point in doing approximation, as numerical integration can be carried out efficiently.) In the following analysis, we assume that $r = 1$. The extension to $r > 1$ follows in an obvious manner.

Let $\mathbf{R}$ have a single-factor structure. That is, for some real scalars $\gamma_1, \ldots, \gamma_M$,

$$(i,j)\text{th element of } \mathbf{R} = \begin{cases} 1, & i = j \\ \gamma_i \gamma_j, & i \neq j \end{cases}. \tag{5.11}$$

Let E_t be the expectation under which the correlation of $X_1, \ldots, X_M$ is the $\mathbf{C}_t$ in (5.10). As t increases from 0 to 1, $E_t(L-y)^+$ progresses from the tranche price of a single-factor model to the desired price. We propose approximations of the form

$$E_t(L-y)^+ \approx \Delta_0 + \Delta_1 t + \Delta_2 \frac{t^2}{2!} + \cdots + \Delta_n \frac{t^n}{n!}. \tag{5.12}$$

Presently, we will specify the coefficients Δ_i's through a conditioning argument and an application of the result from the previous section.

First, we observe that if $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_t)$, where $\mathbf{C}_t$ is given by (5.10) and the reference correlation matrix $\mathbf{R}$ has a single-factor structure (5.11), then $\mathbf{X}$ can be decomposed as

$$X_i = \gamma_i Z + \sqrt{1-\gamma_i^2}\,\widetilde{X}_i. \tag{5.13}$$

Here, $Z \sim \mathcal{N}(0,1)$ and $\widetilde{\mathbf{X}} = (\widetilde{X}_1, \ldots, \widetilde{X}_M)$ is a multivariate normal vector, independent of Z, with zero mean and correlation structure

$$\begin{bmatrix} 1 & t\sigma_{12} & t\sigma_{13} & \dots & t\sigma_{1M} \\ t\sigma_{21} & 1 & t\sigma_{23} & & t\sigma_{2M} \\ t\sigma_{31} & t\sigma_{32} & 1 & & t\sigma_{3M} \\ \vdots & & & \ddots & \\ t\sigma_{M1} & t\sigma_{M2} & t\sigma_{M3} & & 1 \end{bmatrix}, \tag{5.14}$$

where $\sigma_{ik} = (\rho_{ik} - \gamma_i\gamma_k)/\sqrt{(1-\gamma_i^2)(1-\gamma_k^2)}$. We will refer to Z as the *market factor*, since it generates the market-wide correlation $\mathbf{R}$. In this connection, the σ_{ij}'s may be viewed as the *deviations* from the reference correlation matrix $\mathbf{R}$. Conditional on Z, L becomes a loss model with correlation matrix (5.14). It follows from Section 4 that $E_t[(L-y)^+|Z]$ admits an expansion in t.

$$E_t\left[(L-y)^+\middle|Z=z\right] = \delta_0(z) + \delta_1(z)t + \delta_2(z)\frac{t^2}{2!} + \cdots, \tag{5.15}$$

where the argument of $\delta_k(\cdot)$ signifies conditioning on $Z = z$. For given z, $\delta_k(z)$ can be computed using the method described in Section 4. Comparing (5.12) and (5.15) gives

$$\Delta_k = E[\delta_k(z)] = \int_{-\infty}^{\infty} \delta_k(z)\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}\,\mathrm{d}z. \tag{5.16}$$

Through numerical integration, $\Delta_0, \ldots, \Delta_n$ can be computed to give the nth order approximant for $E_t(L-y)^+$.

Note that computing $E(L-y)^+$ through other methods (Monte Carlo or numerical integration) involves some form of integrating in d-dimensional space, where d is the number of factors in the correlation matrix $\mathbf{C}$. The virtue of (5.16), therefore, is that it reduces a d-dimensional problem to one-layer integration.

The computational effort can be analyzed in units of time as in Section 4. In this case, however, (5.16) incurs extra cost for evaluating the integral. If the deviations σ_{ij} in (5.14) have a d'-factor structure, then the cost (in units of time) for computing the nth-order approximant in the heavy correlation case is

$$K \times \binom{n+2d'}{n}, \tag{5.17}$$

where K is the number of points used to evaluate the integral (5.16). Through examples, we found that $\delta_k(z)$ as a function of z is quite well-behaved and predictable. (See Fig. 2 in the following example for typical shapes of the curves $\delta_1(z)$ and $\delta_2(z)$.) So, $K = 10n$ typically gives good results.

The choice of the order of the approximation obviously depends on the rate at which the series (5.12) converges. Unlike the expansion (4.7) in the previous section, where $t = 0$ always corresponds to the independent-obligor case, we now have the freedom to choose $\mathbf{R}$ as the correlation structure at $t = 0$. Provided that $\mathbf{R}$ successfully captures the most important features of $\mathbf{C}$, the progression of $(t, E_t(L-y)^+)$ from $t = 0$ to $t = 1$ will nearly be a straight line. So linear and quadratic approximations generally give good results.

Whereas $\mathbf{R}$ supplies the market-wide correlation, the deviations σ_{ij} represent the excess correlation within and among groups of obligors. (It is not uncommon to find that

Table 2
Computing time of expansion (5.12)

	$d' = 3$	$d' = 4$	$d' = 6$	$d' = 8$
1st order	70	90	130	170
2nd order	560	900	1820	3060
3rd order	2520	4950	13,650	29,070

One unit of computing time equals the amount of time it takes to compute an independent-model price.

(5.14) has a block structure, such as that described in Gregory and Laurent (2004).) Therefore, d' – the number of factors of the deviations – admits the interpretation of being the number of sectors in the market. Or, d' can simply be the result of fitting a factor structure to the deviations through principal component analysis. In doing so, however, keep in mind that d' should be selected so that computation is economical. A reasonable range of d' is displayed in Table 2. The computing time is as given in (5.17) with $K = 10n$.

Let us now consider the time spent by other methods. In plain Monte Carlo with N replications, the computing time simply equals N units. To achieve reasonable precision, the required number of replications can easily be in the tens of thousands. So, according to Table 2, the 2nd-order (or even 3rd-order) approximation can be competitive for a wide range of d'. (This comparison will be made concrete in the numerical example below.)

Numerical integration consumes one unit of time per point. The number of points used in numerical integration increases exponentially with the number of factors d in the correlation structure of X_i's. Note that (5.13) implies that $d = d' + 1$. For example, using the same amount of time it takes to compute the 3rd-order approximation when $d' = 3$, numerical integration has about $\sqrt[4]{2520} \approx 7$ points per dimension; evidently, the curse of dimensionality is more pronounced in numerical integration than in approximation (5.12).

Before we give a numerical example, let us discuss briefly the means of selecting the reference correlation matrix $\mathbf{R}$. To ensure that the expansion (5.12) converges within the first few terms, the choice of $\mathbf{R}$ should approximate $\mathbf{C}$. Andersen et al. (2003) give an efficient algorithm for fitting, in a least square sense, a general correlation matrix with a low-factor structure. Alternatively, since the original covariance matrix $\mathbf{C}$ is often specified through the factor structure (3.5), a convenient way to choose the base single-factor correlation matrix $\mathbf{R}$ is to pick one of $Z_1, \ldots, Z_d$ as the market factor (the remaining Z's will determine the deviation σ_{ik}'s). This is appropriate especially when one of $Z_1, \ldots, Z_d$ carries heavier loadings than the others, and is therefore a clear choice for the market factor.

More generally, one can choose the reference correlation matrix $\mathbf{R}$ to be of r-factor structure, where $r > 1$. In this case, instead of having one market factor Z in the decomposition (5.13), we will have r market factors. Consequently, the expectation (5.16) will become an r-layer integral, and the effort in numerical integration will be exponential in r. So we should resort to this measure only when the single-factor structure fails to capture the nature of $\mathbf{C}$.

Example. We consider a portfolio of $M = 200$ names. For obligor $k = 1, \ldots, M$, the default probability and the loss given default are

$$p_k = .02\left(1 + \sin\frac{8\pi k}{M}\right);$$

$$c_k = \left\lceil \frac{8k}{M} \right\rceil^2.$$

The covariance matrix of the underlying normal random variables has a 7-factor structure with the following loading matrix.

$$\mathbf{A} = \begin{bmatrix} \mathbf{h} & \mathbf{g} & & & & \mathbf{e} & \\ & & \mathbf{g} & & & & \mathbf{e} \\ & & & \mathbf{g} & & \mathbf{e} & \\ & & & & \mathbf{g} & & \mathbf{e} \\ & & & & & \mathbf{e} & \\ & & & & & & \mathbf{e} \\ & & & & & \mathbf{e} & \\ & & & & & & \mathbf{e} \end{bmatrix}$$

where **h** is a column vector of 200 entries, all equal to 0.6; **g** is a column vector of 50 entries, all equal to 0.4; and **e** is a column vector of 25 entries, all equal to 0.3. We want to approximate $E(L-y)^+$ at $y = 700$ (this is about 7.7 $E[L]$ and about 14% of the maximum loss) using a 2nd-order approximation ($n = 2$).

We choose the reference single-factor structure to be the correlation matrix generated by the first column of **A**. Consequently, the residual correlations will have a 6-factor structure. Let $\delta_1(z)$ and $\delta_2(z)$ be the coefficients of the conditioned model. Fig. 5 shows

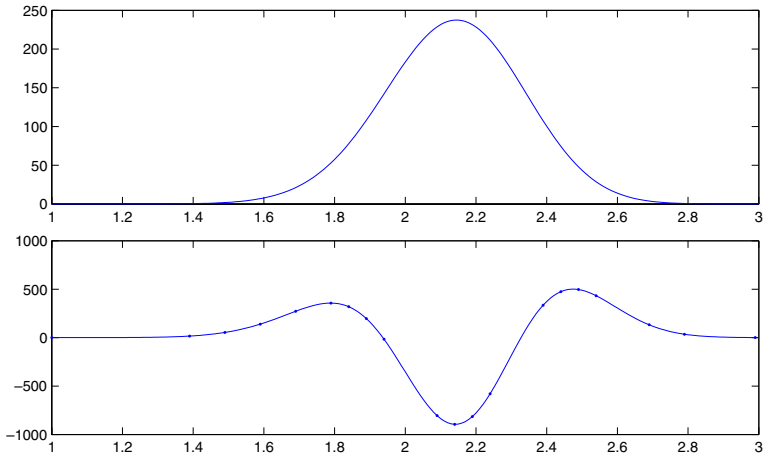

Fig. 5. $\delta_1(z)$ and $\delta_2(z)$ for $z \in [1,3]$.

$\delta_1(z)$ and $\delta_2(z)$ for a range of z. In this example, 20 values of z are enough to define the shape of these curves.

From several examples, we have found that the curve $\delta_1(z)$ is generally unimodal, $\delta_2(z)$ has three extrema, and $\delta_3(z)$ five extrema. Persumably, in general $\delta_k(z)$ has $2k-1$ extrema, though we have no theoretical result to substantiate this. This information can help save the number of points used in the integration (5.16) by putting them in strategic spots.

We then use numerical integration to compute the expectation (5.16). As a result, the first curve integrates to $\Delta_1 = 5.19$, and the second curve integrates to $\Delta_2 = 2.77$. The reference single-factor model is $\Delta_0 = E_0(L-700)^+ = 6.11$. Thus the 2nd-order approximation is $\Delta_0 + \Delta_1 + \frac{1}{2}\Delta_2 = 12.69$. Plain Monte Carlo with 10^6 replications gives 12.86 ± 0.07 with 95% confidence level.

Fig. 6 shows the true values of $E_t(L-y)^+$ at $t = 0, .2, .4, .6, .8, 1$. These are marked by "○". The 1st-order approximation is shown as the solid line, the 2nd-order approximation as the dashed line. Note that Δ_1 and Δ_2 represent the slope and the convexity at $t = 0$, respectively.

Now, let us compare this method with Monte Carlo simulation. Table 3 shows the computing time and outputs of each approach. We take the estimate in the last line as the representative for the true value. One can see that the variance of the Monte Carlo estimate is quite high. So, at 2000 replications, the resulting confidence interval carries little to no information. On the other hand, the 2nd-order approximation simply involves straightforward computation of 1820 independent-obligor prices, and gives a fairly accurate result.

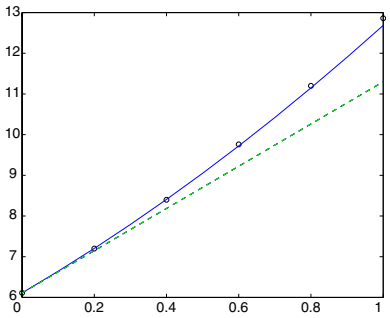

Fig. 6. First and second order approximation for $E_t(L-700)^+$, $t \in [0,1]$.

Table 3
Comparing computing time

	Computing time[a]	Output
Approximation (5.12)		
1st order	260	11.30
2nd order	1820	12.69
Plain Monte Carlo		
2000 replications	2000	11.28 ± 5.18[b]
10^7 replications	10^7	12.86 ± 0.07[b]

[a] One unit of computing time equals the amount of time it takes to compute an indep-obligor price.
[b] 95% confidence interval.

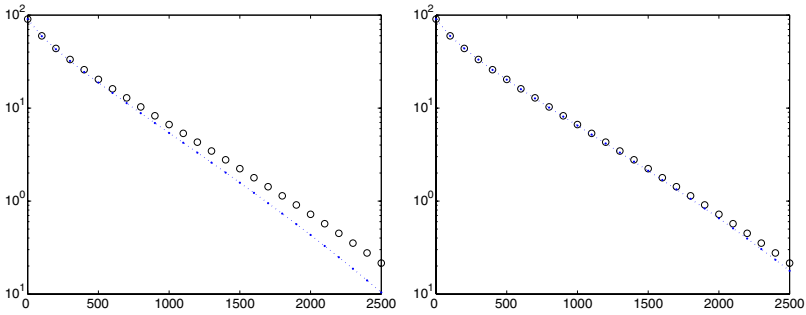

Fig. 7. First (left) and second (right) order approximation for $E_t(L-y)^+$ at $t=1$ and $0 \leqslant y \leqslant 2500$.

It should be noted that while simulation characterizes error in terms of confidence intervals, our approximation does not specify errors. A simple way to check the accuracy of the 2nd-order approximation is to compute the 3rd-order term and see that it is negligible (this, of course, comes with extra computational cost and does not provide a rigorous error bound).

Fig. 7 shows the approximation for a range of y. The true values are shown by "○". The dots on the left plot show the 1st-order approximation, and the right plot shows the 2nd-order approximation. One can see that the 1st-order approximation performs well for a wide range of y. This means that within this range of y, the progression of $E_t(L-y)^+$ with t is almost linear. As a rule of thumb, higher y requires higher order of approximation.

6. Sensitivity analysis

For purposes of hedging and risk management, it is often necessary to calculate the sensitivity of the tranche price $E(L-y)^+$ to default probabilities and correlation parameters. The results in the previous sections are quite useful in sensitivity analysis, as described in the following.

6.1. Sensitivity with respect to correlation

By definition, $\delta_1(\Delta_1)$ is the rate at which the price changes as the correlation structure deviates from an independent (single-factor) model. That is, $\delta_1(\Delta_1)$ represents the "directional sensitivity" of $E(L-y)^+$ with respect to the change of the correlation matrix in a given direction. The sensitivity with respect to an individual correlation coefficient, say ρ_{12}, may be obtained by setting $\sigma_{12}=1$ and all other $\sigma_{ij}=0$. Thus, our approach lends itself naturally to the calculation of these sensitivities. Gregory and Laurent (2004) provide an alternative approach to sensitivity analysis, in which the effect of the perturbation of correlation matrix is analyzed through characteristic functions.

6.2. Sensitivity with respect to default probability

If the obligors are independent, computing the sensitivity with respect to a default probability is easy (see also Andersen et al., 2003):

$$\frac{\partial}{\partial p_i}E(L-y)^+ = \frac{\partial}{\partial p_i}\left[(1-p_i)E(L^{\neg i}-y)^+ + p_i E(L^{\neg i}-y+c_i)^+\right]$$
$$= E(L^{\neg i}-y+c_i)^+ - E(L^{\neg i}-y)^+, \tag{6.18}$$

where $L^{\neg i}$ is the reduced portfolio obtained by removing the ith obligor. When obligors are dependent, however, computing sensitivity with respect to p_i is not as simple because the dependence forbids us to seperate p_i and split the expectation in the above manner. But according to the previous sections, a dependent-obligor model can be expressed as a weighted sum of independent-obligor models. Therefore, the problem of computing p-sensitivity for the dependent-obligor case reduces to that for the independent-obligor case. We now explain in more detail.

We first consider the case where the correlation is light. Taking the first $n+1$ terms in (4.7) and differentiating with respect to p_i yields the approximation

$$\frac{\partial}{\partial p_i}E_t(L-y)^+ \approx u_0 + u_1 t + \cdots + u_n\frac{t^n}{n!}, \tag{6.19}$$

where u_n is the derivative of δ_n with respect to p_i. Using (4.9), we have that u_n can be computed through the limit:

$$\sum_J w_J \frac{\partial}{\partial p_i}\widetilde{E}_J(L-y)^+ \overset{s\downarrow 0}{\to} u_n \tag{6.20}$$

Recall that each $\widetilde{E}_J(L-y)^+$ is an independent-obligor price with modified default probabilities $\tilde{p}_1^{(J)},\ldots,\tilde{p}_M^{(J)}$. So its p_i-sensitivity can be computed simply by the chain rule: first differentiate $\widetilde{E}_J(L-y)^+$ with respect to $\tilde{p}_i^{(J)}$ – which results in the expression (6.18) – then multiply by an extra factor $\partial\tilde{p}_i^{(J)}/\partial p_i$, whose analytical expression can be easily obtained by differentiating (A.26). Thus, computing u_n is no more difficult than computing δ_n.

Now we move on to the case where obligors are heavily correlated, with the covariance matrix (5.10). Differentiating (5.12) yields

$$\frac{\partial}{\partial p_i}E_t(L-y)^+ \approx U_0 + U_1 t + \cdots + U_n\frac{t^n}{n!}. \tag{6.21}$$

Through the same conditioning argument as in Section 4, we can show that

$$U_k = E[u_k(z)] = \int u_k(z)\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}\,\mathrm{d}z. \tag{6.22}$$

where $b_k(z)$ is the coefficient in the expansion of $(\partial/\partial p_i)E_t[(L-y)^+|Z=z]$, computed in the same way as (6.20). The integral can then be evaluated numerically.

Example. Consider the previous example and suppose we want to compute the 2nd-order approximant of $(\partial/\partial p_M)\,E(L-1000)^+$. Again, use first column as the reference correlation structure. The method (6.21, 6.22) gives $U_0 = 4.3$, $U_1 = 4.1$, and $U_2 = 2.3$. The 2nd-order approximation for $(\partial/\partial p_M)E_t(L-1000)^+$ is therefore $4.3 + 4.1t + 2.3t^2/2$. This is shown as the solid line in Fig. 8. The true values of the sensitivity at $t = 0, .2, .4, .6, .8, 1$ are marked with "$\circ$". These are computed by Monte Carlo simulation using the identity

$$\frac{\partial}{\partial p_i}E_{\mathcal{N}(\mathbf{0},\mathbf{C})}(L-y)^+ = E_{\mathcal{N}(v_i\mathbf{C}_i,\mathbf{C}-\mathbf{C}_i\mathbf{C}_i^\top)}\left[(L-y+c_i)^+ - (L-y)^+\right].$$

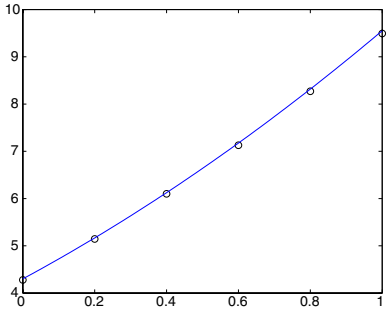

Fig. 8. Sensitivity at $t \in [0,1]$.

The subscript of the expectation denotes the distribution of the latent variable X_i's, and $\mathbf{C}_i$ denotes the ith column of $\mathbf{C}$.

7. Concluding remarks

We have developed a method for approximating CDO prices in the multifactor Normal Copula model. This method expresses the CDO price of a multifactor model as a series of prices of independent-obligor models, thereby exploiting the fact that independent-obligor prices are easy to compute. We also demonstrate how the proposed approximation can be used to compute quantities relating to sensitivity analysis; namely, sensitivity with respect to correlation structure and sensitivity with respect to default probabilities. Finally, we apply this method to compute joint probabilities of multivariate normal random variables.

Acknowledgements

This work is supported by NSF Grants DMI0300044 and DMS0410234.

Appendix A. Statement of main result

In stating the theorem, the following notations are useful. Let $H_n(x)$ denote the Hermite polynomial of degree n; that is, $H_0(x) = 1$, $H_1(x) = x$, and $H_{n+1}(x) = xH_n(x) - nH_{n-1}(x)$ for $n \geqslant 1$. Let $\varphi(x)$ denote the standard normal density function. Define

$$B_k(x) := \varphi(x)H_{k-1}(x), \quad k \geqslant 1. \tag{A.23}$$

Let L be the loss model (3.1, 3.2). Let E_t be the expectation under which X_i's have the correlation matrix (4.6). Assume that ρ_{ij}'s have a generalized d-factor structure:

$$\rho_{ik} = \sum_{j=1}^{d} \lambda_j a_{ij} a_{kj}, \tag{A.24}$$

where λ_j and a_{ij} $(1 \leqslant i \leqslant M,\ 1 \leqslant j \leqslant d)$ are real scalars. Note that (3.5) is a special case of (A.24) where the λ_j's are one. We write $\lambda_j = \lambda(j)$ and $a_{ij} = a(i,j)$ interchangeably. It is also

convenient to let a_{ij} and λ_j admit nonpositive index j by defining the convention $a(i,-j) = -a(i,j)$, $a(i,0) = 0$, $\lambda(-j) = \lambda(j)$, and $\lambda(0) = -2(\lambda_1 + \cdots + \lambda_d)$.

Define $D := \{0, \pm 1, \pm 2, \ldots, \pm d\}$. Let D^n denote the Cartesean product

$$D^n := \{(j_1, \ldots, j_n) : j_1, \ldots, j_n \in D\}. \tag{A.25}$$

For $J = (j_1, \ldots, j_n) \in D^n$, let $\lambda_J := \lambda(j_1), \ldots, \lambda(j_n)$. Let $p_1, \ldots, p_M$ be the default probabilities (3.3). Let $s_1, s_2, \ldots \in \mathbb{R}$. For each $J \in D^n$, let $\tilde{p}_i^{(J)}$ $(i = 1, \ldots, M)$ be a perturbation of p_i given by

$$\tilde{p}_i^{(J)} := p_i + \sum_{k=1}^{n} B_k(v_i) v_J(i,k). \tag{A.26}$$

The scalars $v_J(i,k)$ $(k = 1, \ldots, n)$ depend on $s_1, \ldots, s_n$ and are defined through the following recursion: for all $n \geqslant 2$ and for all $J = (j_1, \ldots, j_n) \in D^n$, $v_J(i,k) = 0$ whenever $k < 0$ or $k > n$ and

$$v_J(i,k) = v_{J'}(i,k) + s_n a(i, j_n) v_{J'}(i, k-1), \quad 1 \leqslant k \leqslant n, \tag{A.27}$$

where $J' = (j_1, \ldots, j_{n-1}) \in D^{n-1}$ is the truncation of J by deleting the nth (last) coordinate. The boundary conditions for the above recursion are: for $j \in D$, $v_J(i, 0) = 1$, $v_J(i,1) = s_1 a_{ij}$, and $v_J(i, k) = 0$ for all $k \neq 0,1$. It is easy to show that $v_J(i, k) \to 0$ as $s_1, \ldots, s_n \to 0$. So, for $s_1, \ldots, s_n$ sufficiently small, the $\tilde{p}_i^{(J)}$ are bonafide probabilities.

We are now ready to state the main theorem, which provides expressions for the coefficients $\delta_0, \delta_1, \ldots, \delta_N$ in an Nth order expansion of (4.7):

Theorem 1. *Let $\delta_0, \delta_1, \ldots$ be coefficients in the expansion (4.7). Fix $n \geqslant 1$. Let $s \in \mathbb{R}$. For each $J \in D^n$, let $\widetilde{E}_J(L-y)^+$ be the tranche price of an independent-obligor model (with the same c_i's) whose default probabilities are $\tilde{p}_1^{(J)}, \ldots, \tilde{p}_M^{(J)}$, calculated from (A.26, A.27) with $s_1 = \ldots = s_n = s$. For each $J \in D^n$, define the weight*

$$w_J := \frac{\lambda_J}{(2s^2)^n}. \tag{A.28}$$

Then,

$$\sum_{J \in D^n} w_J \widetilde{E}_J (L-y)^+ \to \delta_n \tag{A.29}$$

as $s \to 0$.

Let us briefly explain the role of the variable s and the reason for writing the δ_k's as limits. Although a more explicit expression for δ_k is available (it may be obtained by combining (B.41, B.40 and B.31)), it is very complicated and does not lend itself to computation. Therein lies the usefulness of Theorem 1; it provides a compact representation for δ_n as a limit in s. Since the limit can be evaluated by substituting a small number for s, computing δ_n is now feasible. This approach may be viewed as analogous to defining a generating function (in s) whose derivatives are the desired quantities and then approximating the derivatives by finite differences to simplify calculations. This analogy will be explicit in Proposition 1) in the Appendix.

To extend the theorem to cover $n = 0$, simply let $D^0 := \{\emptyset\}$ be a set containing one element – namely, the empty set. For $J = \emptyset \in D^0$, let $\tilde{p}_i^{(J)} := p_i$ for all i and let $\lambda_J := 1$.

It is important to note the following:

1. It appears that to compute the first $n+1$ coefficients $\delta_0, \delta_1, \ldots, \delta_n$ using (A.29), one needs to compute $\widetilde{E}_J(L-y)^+$ for all J in $D^0, D^1, \ldots, D^n$. It suffices, however, to compute $\widetilde{E}_J(L-y)^+$ only for $J \in D^n$. This is because for all k,
$$\left\{\widetilde{E}_J(L-y)^+\right\}_{J\in D^{k-1}} \subset \left\{\widetilde{E}_J(L-y)^+\right\}_{J\in D^k}.$$
To see this, let $J' = (j'_1, \ldots, j'_{k-1}) \in D^{k-1}$ and let $J = (j'_1, \ldots, j'_{k-1}, 0) \in D^k$. It is easy to see that $\tilde{p}_i^{(J')} = \tilde{p}_i^{(J)}$ for all i and, consequently, $\widetilde{E}_{J'}(L-y)^+ = \widetilde{E}_J(L-y)^+$.
2. The summation (A.29) seemingly contains $|D^n| = (2d+1)^n$ terms. Luckily, thanks to the symmetry of the formula (A.26), there are many repeated terms. If $J' = (j'_1, \ldots, j'_n)$ is a permutation of $J = (j_1, \ldots, j_n)$, then it is easy to show that $\tilde{p}_i^{(J')} = \tilde{p}_i^{(J)}$ and, consequently, $\widetilde{E}_{J'}(L-y)^+ = \widetilde{E}_J(L-y)^+$. Therefore, it suffices to compute $\widetilde{E}_J(L-y)^+$ only for $J \in \Lambda$, where Λ is a subset of D^n in which the coordinates are arranged monotonically:
$$\Lambda := \{(j_1, j_2, \ldots, j_n) \in D^n : j_1 \leqslant j_2 \leqslant \ldots \leqslant j_n\}.$$
So in effect, the summation (A.29) contains only $|\Lambda| = \binom{n+2d}{n}$ distinct values of $\widetilde{E}_J(L-y)^+$.
From these two points, we deduce that in order to obtain the nth-order approximant for $E_t(L-y)^+$, the number of terms of the form $\widetilde{E}_J(L-y)^+$ one needs to compute is $\binom{n+2d}{n}$.
The proof of Theorem 1 is given in Appendix B. In fact, we will prove the following theorem, which is a generalization of Theorem 1.

Theorem 2 (Generalization of Theorem 1). *Let* $f: \mathbb{R} \to \mathbb{R}$. *Then,* $E_t f(L)$ *is analytical in* t; *that is,* $E_t f(L)$ *admits the infinite expansion*:
$$E_t[f(L)] = \mu_0 + \mu_1 t + \mu_2 \frac{t^2}{2!} + \cdots.$$
Furthermore, where $w_J = \lambda_J/(2s^2)^n$ *and* $\widetilde{E}_J f(L)$ *is the expectation under which all obligors in* L *default independently with probabilities* $\tilde{p}_1^{(J)}, \ldots, \tilde{p}_M^{(J)}$ *calculated from* (A.26, A.27) *with* $s_1 = \ldots = s_n = s$, *we have*
$$\sum_{J\in D^n} w_J \widetilde{E}_J f(L) \to b_n.$$
as $s \to 0$.

Obviously, when $f(L) = (L-y)^+$, we retrieve Theorem 1. The comments from Theorem 1 apply to this theorem as well.

Appendix B. Proof of Theorem 2

We use the following notation throughout the proof. For a positive integer d, let $I_{[d]} := \{1, 2, \ldots, d\}$ and let $I_d^n = \{(j_1, \ldots, j_n) : j_1, \ldots, j_n \in I_d\}$.

The backbone of our result is Kibble's (1945) generalized tetrachoric series for multivariate normal joint probabilities, which we now recapitulate. Let $\mathbf{C}_t$ be the covariance matrix given by (4.6). Let P_t be a probability measure under which $(X_1,\ldots,X_M)\sim\mathcal{N}(\mathbf{0},\mathbf{C}_t)$. Kibble (1945) asserts that $P_t(X_1>v_1,\ldots,X_M>v_M)$ admits an infinite expansion:

$$P_t(X_1>v_1\ldots,X_M>v_M)=b_0+b_1t+b_2\frac{t^2}{2!}+\cdots, \tag{B.30}$$

and gives the expression for b_n as

$$b_n=\sum_{\substack{1\leqslant\kappa_1<\ell_1\leqslant M\\ \vdots\\ 1\leqslant\kappa_n<\ell_n\leqslant M}}\rho_{k_1\ell_1}\ldots\rho_{k_n\ell_n}B_{h_1}(v_1)\ldots B_{h_M}(v_M). \tag{B.31}$$

For each term in the summation, $h_i(1\leqslant i\leqslant M)$ is the number of times "i" occurs among the subscripts $k_1,\ell_1,\ldots,k_n,\ell_n$. The function $B_k(x)$ $(k=1,2,\ldots)$ are as defined in (A.23), and $B_0(x):=\overline{\Phi}(x)$.

Because the summation (B.31) contains a large number of terms, especially when M is large, it is not a practical way to compute b_n. Our result builds on Kibble's result by finding a compact representation for (B.31) (for a preview, fastforward to Proposition 1), which leads to a more cost-efficient way to compute b_n (see Proposition 2). We will see that after we have found an appropriate representation for (B.31), Theorem 2 will easily follow.

We begin by seeking a compact representation for the summation (B.31) when the ρ's have a single-factor structure. To do this, we will prove a more general statement below.

Lemma 1. *Let $\mathbf{C}^{(1)},\mathbf{C}^{(2)},\ldots,\mathbf{C}^{(n)}$ be covariance matrices of dimension $M\times M$, each having a single-factor structure. That is, for all $r=1,\ldots,n$,*

$$\mathbf{C}^{(r)}=\left[\rho_{k\ell}^{(r)}\right]_{M\times M},\quad \rho_{k\ell}^{(r)}=a_k^{(r)}a_\ell^{(r)}\quad(k\neq\ell).$$

Let $(\beta_r^{(i)}i=1,\ldots,M$ and $r=0,\ldots,n)$ be real constants. Consider

$$\sum_{\substack{1\leqslant\kappa_1<\ell_1\leqslant M\\ \vdots\\ 1\leqslant\kappa_n<\ell_n\leqslant M}}\rho_{k_1\ell_1}^{(1)}\ldots\rho_{k_n\ell_n}^{(n)}\beta_{h_1}^{(1)}\ldots\beta_{h_M}^{(M)}, \tag{B.32}$$

For each term in the summation, $h_i(1\leqslant i\leqslant M)$ is the number of times "i" occurs among the subscripts $k_1,\ell_1,\ldots,k_n,\ell_n$. Then, we have that the above summation equals

$$\frac{1}{2^n}\frac{\partial^{2n}}{\partial s_1^2\ldots\partial s_n^2}\bigg|_{\mathbf{s}=\mathbf{0}}G_1G_2\ldots G_M, \tag{B.33}$$

where G_i $(i=1,\ldots,M)$ is a polynomial in $s_1,\ldots,s_n$ defined as

$$G_i:=\sum_{T\subset I_n}\beta_{|T|}^{(i)}\prod_{r\in T}s_ra_i^{(r)}. \tag{B.34}$$

Note. If we read this lemma with $a_i^{(1)}=\ldots=a_i^{(n)}=a_i$ and $\beta_{ir}=B_r(v_i)$, then (B.33) is a compact representation for (B.31) in the single-factor case.

Proof. Note that $G_1,\ldots,G_M$ consist of 2^n terms each, all of whom distinct. Let G be the result of expanding the product $G_1G_2\ldots G_M$, so that G comprises 2^{nM} distinct terms. The operator (B.33) kills all terms of G except any term of the type $s_1^2s_2^2\ldots s_n^2\times$ constant, from which the operator extracts the constant. Therefore, to prove that (B.33) equals (B.32), it suffices to show that (A) if a term in G is of the form $s_1^2s_2^2\ldots s_n^2\times$ constant, then the constant matches a term in the summation (B.32); and (B) each term in the summation (B.32) multiplied by $s_1^2s_2^2\ldots s_n^2$ equals a term in G.

Let us first paint the picture of G. Let A be one of the 2^{nM} terms in G. Since G is the product of $G_1,\ldots,G_M$, it follows that $A=A_1A_2\ldots A_M$, where A_i is one of the 2^n terms in G_i. Evidently from (B.34), A_i is characterized by the following conditions: (I) A_i contains a factor $\beta_q^{(i)}$, where q equals the number of accompanying $a_i^{(\cdot)}$'s; (II) $a_i^{(r)}$ always appears alongside s_r in A_i, and vice versa; and (III) A_i contains no s_r^m with $m>1$. We now split our proof in two parts: verifying (A) and verifying (B).

Verify (*A*): Suppose that $A=A_1\ldots A_M$ contains exactly two s_1's. Because of (III), the two s_1's must come from two different A_i's – say, A_{k_1} and A_{ℓ_1} (assume $k_1<\ell_1$ without loss of generality). It follows from (II) that A must contain $a_{k_1}^{(1)},a_{\ell_1}^{(1)}$, and no other $a_\bullet^{(1)}$'s. By repeating this argument, we conclude that if A contains exactly two s_1's, exactly two s_2's, and so on, then for some $k_1<\ell_1$, $k_2<\ell_2,\ldots$,

$$A=s_1^2\ldots s_n^2\left(a_{k_1}^{(1)}a_{\ell_1}^{(1)}\ldots a_{k_n}^{(n)}a_{\ell_n}^{(n)}\right)\beta_{q_1}^{(1)}\ldots\beta_{q_M}^{(M)}$$

The factors $\beta_{q_i}^{(i)}$ appear as a consequence of (I). Obviously, the quantity in the parantheses equals $\rho_{k_1\ell_1}^{(1)}\ldots\rho_{k_n\ell_n}^{(n)}$. The number of times i occurs among $k_1,\ell_1,\ldots,k_n,\ell_n$ equals, as seen from above, the count of $a^{(\cdot)}$'s with subscript i; but this, according to (I), equals q_i. Therefore, $h_i=q_i$ and the proof of this part is complete.

Verify (*B*): Pick a term in the summation (B.32) and multiply by $s_1^2s_2^2\ldots s_n^2$:

$$\underbrace{s_1a_{k_1}^{(1)}}_{\text{Cluster 1}}\ \underbrace{s_1a_{\ell_1}^{(1)}}_{\text{Cluster 2}}\ \ldots\ \underbrace{s_na_{k_n}^{(n)}}_{\text{Cluster }2n-1}\ \underbrace{s_na_{k_n}^{(n)}}_{\text{Cluster }2n}\ \beta_{h_1}^{(1)}\ldots\beta_{h_M}^{(M)}=:A$$

the above product can be regrouped as $A=A_1A_2\ldots A_M$, where A_i collects $\beta_{h_i}^{(i)}$ and every cluster in which the subscript of a is i. One can easily see that A_i's satisfy (I), (II), and (III). Hence, A is a term in G. □

Now we move to the general d-factor case.

Proposition 1. *Let $v_1,\ldots,v_M$ be real numbers. Let P_t be a probability measure under which $X_1,\ldots,X_M$ are $\mathcal{N}(0,1)$ random variables with covariance matrix (4.6). Suppose that the ρ_{ij}'s have a generalized d-factor structure (A.24). Let b_0, b_1, $b_2,\ldots$ be the coefficients in the expansion (B.30). Then*

$$b_n=\frac{1}{2^n}\sum_{J\in I_d^n}\lambda_J\frac{\partial^{2n}}{\partial s_1^2\ldots\partial s_n^2}\bigg|_{\mathbf{s}=0}\tilde{p}_1^{(J)}\ldots\tilde{p}_M^{(J)},\tag{B.35}$$

where $\tilde{p}_i^{(J)}$'s are as given in (A.26, A.27).

Proof. From (A.27), it can be shown by induction that for $J=(j_1,\ldots,j_n)\in D^n$, $v_J(i,k)=\sum_{T\subset I_n,|T|=k}\prod_{r\in T}s_ra(i,j_r)$. Consequently, (A.26) can be written as

$$\tilde{p}_i^{(J)} := \sum_{T \subset I_n} B_{|T|}(v_i) \prod_{r \in T} s_r a(i, j_r).$$

where $B_0(x) := \overline{\Phi}(x)$ and $B_k(x)(k \geqslant 1)$ are as defined in (A.23). From the assumption that ρ's have a generalized d-factor structure (A.24), we can write $\rho_{k\ell}(1 \leqslant k \leqslant \ell \leqslant M)$ as

$$\rho_{k\ell} = \sum_{j=1}^{d} \lambda_j \rho_{k\ell}^{(j)}, \quad \text{where } \rho_{k\ell}^{(j)} = a_{kj} a_{\ell j}. \tag{B.36}$$

We then decompose $\rho_{k_1\ell_1}, \ldots, \rho_{k_n\ell_n}$ in (B.31) using (B.36). In decomposing $\rho_{k_1\ell_1}$ using (B.36), we use j_1 (instead of j) as the running index in the decomposition. Similarly, use j_2 as the running index for the decomposition of $\rho_{k_2\ell_2}$, and so on. Then, (B.31) becomes

$$\sum_{j_1,\ldots,j_n \in I_d} \lambda(j_1) \ldots \lambda(j_n) \sum_{\substack{1 \leqslant \kappa_1 \ell_1 \leqslant M \\ \vdots \\ 1 \leqslant \kappa_n \ell_n \leqslant M}} \rho_{k_1\ell_1}^{(j_1)} \ldots \rho_{k_n\ell_n}^{(j_n)} B_{h_1}(v_1) \ldots B_{h_M}(v_M).$$

To finish the proof, rewrite the inner summation using Lemma 1 with $a_i^{(r)} := a(i, j_r)$ and $\beta_r^{(i)} := B_r(v_i)$. □

The role of the following proposition is to rewrite the R.H.S. of (B.35) in limit form.

Proposition 2. *Same setup as Proposition 1. Then,*

$$b_n = \lim_{s \to 0} \sum_{J \in D^n} w_J \tilde{p}_1^{(J)} \ldots \tilde{p}_M^{(J)} \tag{B.37}$$

where w_J is given in (A.28), and $\tilde{p}_i^{(J)}$'s are perturbed probabilities calculated from (A.26, A.27) with $s_1 = \ldots = s_n = s$.

Proof. Let $g := \tilde{p}_1^{(J)} \ldots \tilde{p}_M^{(J)}$. Let us make the dependence of g on s_n and j_n explicit by writing $g = g(s_n, j_n)$. It can be seen from (A.27) that $g(-s_n, j_n) = g(s_n, -j_n)$ and $g(0, j_n) = g(s_n, 0)$. Using the finite difference formula,

$$\begin{aligned} \left.\frac{\mathrm{d}^2}{\mathrm{d}s^2}\right|_{s=0} g(s, j) &= \lim_{s \to 0} \frac{g(s, j) + g(-s, j) - 2g(0, j)}{s^2} \\ &= \lim_{s \to 0} \frac{g(s, j) + g(s, -j) - 2g(s, 0)}{s^2}. \end{aligned} \tag{B.38}$$

(To simplify notation, we have suppressed the subscript n in s_n and j_n.) Multiply (B.38) by $\lambda(j)$ and sum over $j \in I_d$:

$$\begin{aligned} \frac{1}{2} \sum_{j \in I_d} \lambda_j \left.\frac{\mathrm{d}^2}{\mathrm{d}s^2}\right|_{s=0} g(s, j) &= \lim_{s \to 0} \frac{1}{2s^2} \sum_{j=1}^{d} \lambda(j)[g(s, j) + g(s, -j) - 2g(s, 0)] \\ &= \lim_{s \to 0} \frac{1}{2s^2} \sum_{j \in D} \lambda(j) g(s, j), \end{aligned}$$

The last equality follows from the definition of $\lambda(j)$ when j is nonpositive (see page 23). Now write g as $g(s_1, \ldots, s_n; J)$ so that the dependence on $s_1, \ldots, s_n$ and J is made explicit. By repeating the above argument n times, we have

$$\frac{1}{2^n}\sum_{J\in I_d^n}\lambda_J\frac{\partial^{2n}g}{\partial s_1^2\ldots\partial s_n^2}\bigg|_{\mathbf{s}=0}=\lim_{s\to 0}\frac{1}{(2s^2)^n}\sum_{J\in D^n}\lambda_J g(s,\ldots,s;J)$$

The R.H.S. equals b_n according to Proposition 1). To complete the proof, note that $w_J=\lambda_J/(2s^2)^n$ and that $g(s,\ldots,s;\,J)$ is the product of the perturbed probabilities (A.26, A.27) with $s_1=\ldots=s_n=s$. □

We are now ready to prove Theorem 2 in Section 3.

Proof of Theorem 2. Lemma 2 implies that for every $S\subset I_M$, P_t (names $i\in S$ default) is analytical in t and

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}\bigg|_{t=0}P_t(\text{names } i\in S \text{ default})=\lim_{s\to 0}\sum_{J\in D^n}w_J\widetilde{P}_J(\text{names } i\in S \text{ default}). \tag{B.39}$$

Let E_S denote the event that obligors $i\in S$ default while the rest (obligors $\ell\in I_M\backslash S$) do not. Using the inclusion–exclusion formula from set theory, it is easy to argue that $P_t(E_S)$ is analytical in t and

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}\bigg|_{t=0}P_t(E_S)=\lim_{s\to 0}\sum_{J\in D^n}w_J\widetilde{P}_J(E_S) \tag{B.40}$$

for all $S\subset I_M$. To prove the theorem, we will first argue that $E_t f(L)$ has an infinite-series representation in t. For $S\subset I_M$, set $c(S):=\sum_{i\in S}c_i$. It is easy to see that

$$E_t f(L)=\sum_{S\subset I_M}f(c(S))P_t(E_S). \tag{B.41}$$

The summation is over all possible sets of defaulting obligors, and $f(c(S))$ is the payoff given E_S. Since $\mathrm{P}_t(E_S)$ is analytical in t, so is $E_t[f(L)]$. Finally, if we multiply (B.40) by $f(c(S))$ and sum over all $S\subset I_M$, we obtain

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}\bigg|_{t=0}E_t[f(L)]=\lim_{s\to 0}\sum_{J\in D^n}w_J\widetilde{E}_J[f(L)].$$

The proof is then complete. □

References

Ağca, S., Chance, D.M., 2003. Speed and accuracy comparison of bivariate normal distribution approximations for option pricing. Journal of Computational Finance 6 (4), 61–70.

Andersen, L., Sidenius, J., Basu, S., 2003. All your hedges in one basket. RISK (November), 67–72.

Gregory, J., Laurent, J.-P., 2004. In the core of correlation. RISK (October), 87–91.

Hull, J., White, A., 2004. Valuation of a CDO and an *n*th to default CDS without Monte Carlo simulation. The Journal of Derivatives 12 (2), 8–23, Winter.

Kendall, M.G., 1941. Proof of relations connected with the tetrachoric series and its generalization. Biometrika 32 (2), 196–198.

Kibble, W.F., 1945. An Extension of a Theorem of Mehler's on Hermite Polynomials. Proc. Cambridge Philos. Soc. 41, 12–15.

Laurent, J.-P., Gregory, J., 2003. Basket default swaps, CDOs and factor copulas. Journal of Risk 7 (4), 103–122.

Li, D., 2000. On default correlation: A copula approach. Journal of Fixed Income 9, 43–54.
Schönbucher, P.J., 2003. Credit Derivatives Pricing Models: Model, Pricing and Implementation. Wiley, Chichester, Hoboken, N.J.
Somerville, P.N., 1954. Some problems of optimum sampling. Biometrika 41 (3/4), 420–429.
Vasicek, O.A., 1998. A series expansion for the bivariate normal integral. Journal of Computational Finance 1 (4), 5–10.