APPROXIMATIONS AND CONTROL VARIATES FOR PRICING PORTFOLIO CREDIT DERIVATIVES

Zhiyong Chen

Bear, Stearns & Co., Inc.
383 Madison Avenue
New York, NY 10179, U.S.A.

Paul Glasserman

Columbia Business School
3022 Broadway, 101 Uris
Columbia University
New York, NY 10027, U.S.A.

ABSTRACT

Portfolio credit derivatives that depend on default correlation are increasingly widespread in the credit market. Valuing such products often entails Monte Carlo simulation. However, for large portfolios, plain Monte Carlo simulation can be slow. In this paper, we develop approximation methods for pricing collateralized debt obligation (CDO) tranches in the widely used factor copula approach. We also discuss using the approximations as control variates to improve the precision of Monte Carlo estimates. These approximation methods and control variate techniques could be applied to pricing other portfolio credit derivatives as well.

1 INTRODUCTION

A portfolio credit derivative is a contract tied to a portfolio of defaultable assets, such as corporate bonds and credit default swaps. Its payoff depends on the default experience of the underlying assets. One of the most popular types of portfolio credit products is the collateralized debt obligation (CDO). In a traditional CDO, the arranger redistributes credit losses on the underlying assets into several tranches (equity, mezzanine, and senior) and issues notes backed by each tranche. Tranches with lower seniority absorb losses before those with higher seniority and earn a higher coupon rate to compensate for the higher risk. See, e.g., Bruyere et al. (2006) and Schönbucher (2003) for general background on these types of contracts.

Valuing portfolio credit derivatives requires a model to describe the joint distribution of the default times of the underlying assets in the portfolio. We assume the risk-neutral default probability for each asset exists and is, in effect, chosen by the market. In practice, the marginal default distribution of the time to default is typically implied from related market information, such as spreads on corporate bonds or single-name default swaps (see, e.g., Duffie and Singleton (2003)). The dependence among the default times is usually specified by a copula model. The most widely used mechanism is the Gaussian copula model (Gupton, Finger, and Bhatia 1997; Li 2000). As pointed out in many articles (see, for example, Andersen, Sidenius, and Basu 2003), introducing a factor structure to the correlation matrices can yield significant improvements in computing speed. The one-factor Gaussian copula model has become the market standard for many purposes.

Monte Carlo simulation is usually required in computing prices or sensitivities of portfolio credit derivatives in multifactor models. However, calculation through plain Monte Carlo simulation can be slow. To address this difficulty, a few techniques have been introduced to accelerate computations for CDOs and other portfolio credit derivatives by approximating the conditional portfolio loss distribution. Andersen, Sidenius, and Basu (2003) approximate all losses as multiples of a basic loss unit, which allows the construction of the loss distribution through a recursive procedure. Hull and White (2004) develop a similar recursive method that divides potential losses into buckets, computes the probability of each bucket, and replaces the loss distribution in each bucket with its mean. Both methods are based on factor reduction and introducing some coarseness in loss levels.

In this paper, we develop an alternative approximation approach. Our method uses exact calculation of the distribution of the number of defaults, which is more tractable than the loss distribution. We then use two distributions to approximate the loss distribution conditional on the number of defaults, and thus construct an approximation of the portfolio loss distribution. In the special case of a homogeneous portfolio, the approximation distribution replicates the actual loss distribution. We give a detailed description of our methods, and analyze the cases where these methods are good and where they are not. We also show how to use the approximation methods to construct effective control variates that can be applied for variance reduction when the approximation itself is not sufficiently accurate to replace simulation.

The rest of this paper is organized as follows. Section 2 provides background on CDOs and the Gaussian copula model. Section 3 presents our approximation methods in the case of independent default times, and Section 4 extends our methods to the dependent default time case in the factor Gaussian copula model. Section 5 discusses a control variate technique for Monte Carlo simulation. Section 6 concludes the paper.

2 PROBLEM DESCRIPTION

This section gives a brief description of the credit derivatives we consider and reviews the popular Gaussian copula model with factor structure.

2.1 Collateralized Debt Obligations

We consider credit derivatives tied to a portfolio of N underlying assets subject to default, such as loans, bonds, and single-name credit default swaps. The value of N could vary from five to a few hundred. Each asset is associated with a default time τ_i and a recovery rate r_i, $i = 1, \ldots, N$. Let $\tau_i = \infty$ if the ith asset never defaults. For simplicity, we assume all recovery rates are constant and all asset notionals are 1. (In fact, differences in asset notionals could be absorbed into differences in recovery rates.) For convenience, we define $\tau = (\tau_1, \ldots, \tau_N)$.

In a CDO, credit losses on a pool of defaultable assets are tranched and passed to different investors. Lower seniority tranches act as cushions against losses in higher seniority tranches. In the usual structure of a CDO, a tranche absorbs losses from an attachment point S_l to a detachment point S_u. The cashflows of this tranche are as follows. At dates $0 < T_1 < \ldots < T_m \leq T$, where T is the maturity date, the tranche holder receives payment proportional to the notional principal left in the tranche. If the ith asset defaults, it causes a deterministic loss of $l_i = 1 - r_i$, called the loss given default (LGD). Order the assets so that $l_1 \leq \cdots \leq l_N$. For simplicity, we assume that the net default payments occur only at the coupon dates $T_1, \ldots, T_m$.

Let $L(t)$ be the cumulative loss on the collateral portfolio at time $t \leq T$, i.e., $L(t) = \sum_{i=1}^{N} l_i \mathbf{1}(\tau_i \leq t)$. The cumulative loss on the tranche at time t is

$$M(t) = (L(t) - S_l)^+ - (L(t) - S_u)^+,$$

where the "+" means the positive part of the value.

The discounted payoff of a CDO tranche can be written as the difference between a default payment leg and a premium payment leg,

$$V(\tau) = V_{def}(\tau) - V_{pre}(\tau).$$

To define the terms on right, let $D(t)$ be the discount factor for the interval from 0 to t. The discounted payoff of the default payment leg is

$$V_{def}(\tau) = \sum_{i=1}^{m} D(T_i) \left(M(T_i) - M(T_{i-1}) \right),$$

where $T_0 = 0$. Ignoring the accrual factor for payment days, the discounted payoff of the premium payment leg is

$$V_{pre}(\tau) = c \sum_{i=1}^{m} D(T_i)(S_u - S_l - M(T_i)),$$

where the constant c is the *spread* or the coupon rate of this tranche. Thus, the payoff of the CDO tranche can be priced as

$$
\begin{aligned}
V(\tau) &= (1+c) \sum_{i=1}^{m} D(T_i) M(T_i) - \sum_{i=1}^{m-1} D(T_{i+1}) M(T_i) \\
&\quad - c(S_u - S_l) \sum_{i=1}^{m} D(T_i).
\end{aligned} \tag{1}
$$

We observe that to compute $\mathbf{E}(V(\tau))$ we only need the first moment of the cumulative loss on the tranche at each date $T_1, \ldots, T_m$. Therefore, we are able to replace the problem of pricing a CDO tranche by a sequence of subproblems that estimate

$$\mathbf{E}(M(t)) = \mathbf{E}((L(t) - S_l)^+) - \mathbf{E}((L(t) - S_u)^+) \tag{2}$$

at $t = T_1, \ldots, T_m$. We will mainly discuss the subproblem in the rest of the paper. To calculate $\mathbf{E}(M(t))$, we approximate (or simulate) $L(t)$.

2.2 The Gaussian Copula Model

The expectations in (1) and (2) are taken in the risk-neutral measure, where the joint distribution of the default times, $f(t_1, \ldots, t_N)$, has not yet been specified. The marginal distribution of each τ_i is usually specified as follows:

$$F_i(t) = \mathbf{P}(\tau_i < t) \equiv 1 - \exp\left(-\int_0^t \lambda_i(s) \, ds \right),$$

where $\mathbf{P}$ is the risk-neutral measure and λ_i is the corresponding default hazard rate function. The function $F_i(t)$ is typically extracted from the market prices of credit default swaps or bonds, and is assumed known for all t.

The Gaussian copula (Gupton et al. 1997; Li 2000) is a widely used mechanism for specifying a joint distribution for the default times consistent with given marginals, and will provide a useful and illustrative example. In the Gaussian copula model, the dependence among $\tau_1, \ldots, \tau_N$

is determined by underlying jointly normal random variables $W_1, \ldots, W_N$. Each W_i has a standard normal distribution Φ, so $\Phi(W_i)$ is uniformly distributed on $(0,1)$ and $\tau_i = F_i^{-1}(\Phi(W_i))$ has distribution F_i. However, $W_1, \ldots, W_N$ are correlated, with covariance matrix Σ, and this introduces (and, indeed, completely characterizes) dependence among the default times $\tau_1, \ldots, \tau_N$. We will make the simplifying assumption that Σ has full rank so that no asset has its default time completely determined by those of the other assets.

2.3 Factor Models

In practice, the correlations determined by Σ are often specified through a factor model of the form

$$W_i = a_{i1}Z_1 + \cdots + a_{id}Z_d + b_i\varepsilon_i, \quad i = 1, \ldots, N,$$

or, in matrix-vector form,

$$\mathbf{W} = \mathbf{AZ} + \mathbf{B}\varepsilon, \tag{3}$$

in which

- $Z_1, \ldots, Z_d$, $d < N$, are systematic risk factors, normalized to be independent standard normals;
- ε_i, $i = 1, \ldots, N$, are idiosyncratic risks associated with the individuals assets, $N(0,1)$ distributed and independent of each other and of $Z_1, \ldots, Z_d$;
- $a_{i1}, \ldots, a_{id}$ are the factor loadings for the ith asset, and $\sum_{k=1}^{d} a_{ik}^2 \leq 1$;
- $b_i = \sqrt{1 - \sum_{k=1}^{d} a_{ik}^2}$, so that W_i is also $N(0,1)$ distributed.

The correlation matrix of $\mathbf{W}$ implied by (3) is $\mathbf{AA}^\top + \mathbf{B}^2$; so, for consistency, we require that this matrix equal Σ.

One of the most important features of factor models is that $W_1, \ldots, W_N$ become conditionally independent by fixing the factors $Z_1, \ldots, Z_d$.

3 APPROXIMATION METHODS WITH INDEPENDENT DEFAULT TIMES

The value of a CDO tranche is completely determined by the distribution of cumulative portfolio loss. This distribution is discrete with a finite number of possible values. However, the number of possibilities is still quite large and could grow exponentially when the number of assets increases. To find a good discretization of the loss distribution might be very challenging. However, it is much easier to get the distribution of the number of defaults in the portfolio. In particular, when all underlying names are independent, the probability of exactly n defaults can be determined recur-

sively (as in Andersen et al. 2003, Hull and White 2004, and Chen and Glasserman 2006).

In this section, instead of constructing the portfolio loss distribution directly, we consider approximating the loss distribution given the number of defaults in the pool. Through this approach, we develop approximation methods for valuing CDO tranches, and we present some numerical examples. We begin by considering the case of independent default times, a restriction in force throughout this section.

3.1 Notation and Basic Computation

We consider a single tranche CDO with N underlying names with loss $l_1 \leq \cdots \leq l_N$. We are interested in the problem of estimating

$$\mathbf{E}(M(t)) = \mathbf{E}((L(t) - S_l)^+) - \mathbf{E}((L(t) - S_u)^+)$$

for a given time t.

We use $p_1, \ldots, p_N$ to denote the default probabilities $F_1(t), \ldots, F_N(t)$, and sort them in ascending order $p_{(1)} \leq \ldots \leq p_{(N)}$. A simple and fast recursive algorithm allows calculation of the default probabilities. Let $Y_i(t) = I(\tau_i < t)$, i.e., the indicator that the ith asset defaults before t. Then $Y_i(t) = 1$ with probability p_i and is otherwise 0. To explain our algorithm, we introduce a Markov chain (indexed by i with t fixed)

$$X_i(t) = \sum_{j=1}^{i} Y_j(t), \quad i = 0, 1, \ldots, N,$$

that counts the number of defaults as we go through the number of underlying assets. That this process is indeed Markov follows from the independence of the default indicators Y_j. The chain is absorbed at time N into any of the states $\{0, 1, \ldots, N\}$. Define

$$P_k(n;t) = \mathbf{P}(X_k(t) = n);$$

this is the probability of n defaults in the first k assets before time t. Then the probability mass function of the loss distribution $P_N(n;t)$ can be obtained recursively following the rule

$$P_{k+1}(n;t) = p_{k+1}P_k(n-1;t) + (1 - p_{k+1})P_k(n;t),$$

with $P_0(0;t) = 1$ and let $P_k(n;t) = 0$ if $n > k$. This recursive method is a special case of the method introduced in Andersen et al. (2003), corresponding to all assets having the same LGD.

Suppose the number of defaults in the portfolio is n, which could be any value between 0 and N. Let $L(t|n)$ denote the total loss in the portfolio at time t given exactly

n credit defaults in the pool. Because $l_1 \leq l_2 \leq \cdots \leq l_N$,

$$\sum_{i=1}^{n} l_i \leq L(t|n) \leq \sum_{i=N-n+1}^{N} l_i.$$

When $n = 0$ or N, $L(t|n) = 0$ or $\sum_{i=1}^{N} l_i$ with probability 1. When $n = 1$,

$$L(t|n) = l_i, \quad \text{w.p.} \quad \frac{p_i}{1-p_i} \frac{\prod_{j=1}^{N}(1-p_j)}{P_N(1;t)}.$$

Therefore, $\mathbf{E}(M(t)|n)$ is easily computed when $n = 0$, 1, or N. Indeed, $\mathbf{E}(M(t)|0) = 0$ and $\mathbf{E}(M(t)|N) = S_u - S_l$. What we are interested in is the distribution of $L(t|n)$ when $1 < n < N$.

For convenience, we use n_l and n_u to denote the minimal and maximal values of n which make $S_l \leq L(t|n) \leq S_u$. It is not hard to see that

$$n_l = \min_{n}\{\sum_{i=1}^{n} l_{N-i} \geq S_l\}, \qquad n_u = \max_{n}\{\sum_{i=1}^{n} l_i \leq S_u\}.$$

We use $L_{\min}^n$, $L_{\max}^n$, $L_{\mod}^n$ to denote the minimum, maximum, and mode of $L(t|n)$, i.e.,

$$L_{\min}^n = \sum_{i=1}^{n} l_i, \qquad L_{\max}^n = \sum_{i=N-n+1}^{N} l_i, \qquad L_{\mod}^n \approx \sum_{i=N-n+1}^{N} l_{(i)},$$

where the last is an approximation of the real mode. Recall that $p_{(1)} \leq \cdots \leq p_{(N)}$, and this is the ordering of the $l_{(i)}$ used in the last expression.

In this section, we will use different distributions to approximate the distribution of $L(t|n)$. Once we have the approximate distribution, we are able to compute $M(t)$ analytically.

3.2 Homogeneous Approximation

The most straightforward approximation is the homogeneous approximation, i.e., we use a homogeneous portfolio, where each asset has LGD of $\bar{l} = \sum_{i=1}^{N} l_i/N$, to approximate the real portfolio. Under this approximation, $L_{\min}^n = L_{\mod}^n = L_{\max}^n = n\bar{l}$, and then the conditional loss on the tranche is deterministic,

$$M(t|n) = (n\bar{l} - S_l)^+ - (n\bar{l} - S_u)^+.$$

The expected tranche loss is given by

$$\mathbf{E}(M(t)) = \sum_{n=1}^{N-1} M(t|n)P_N(n;t) + (S_u - S_l)P_N(N;t).$$

The value of the CDO tranche in the homogeneous portfolio could then be calculated analytically, and is a good approximation of the desired tranche value in the real portfolio.

3.3 Approximation with Triangular Distribution

A simple way to incorporate additional information about the conditional loss distribution is to use a triangular distribution with endpoints $L_{\min}^n$ and $L_{\max}^n$ and mode $L_{\mod}^n$; i.e., the distribution with density

$$f_n(x) = \begin{cases} \frac{2(x-L_{\min}^n)}{(L_{\max}^n-L_{\min}^n)(L_{\mod}^n-L_{\min}^n)}, & L_{\min}^n \leq x \leq L_{\mod}^n; \\ \frac{2(L_{\max}^n-x)}{(L_{\max}^n-L_{\min}^n)(L_{\max}^n-L_{\mod}^n)}, & L_{\mod}^n < x \leq L_{\max}^n; \\ 0, & \text{otherwise.} \end{cases}$$

With this approximation, we are able to get the close form of $\mathbf{E}(M(t)|n)$. The possible expected values of $M(t)$ conditional on different n can be calculated by Proposition 1.

Proposition 1 *If $L(t|n)$ $(1 < n < N)$ has density f_n, then the value of $\mathbf{E}((L(t|n)-S)^+)$ for a given S is*

- $(L_{\min}^n + L_{\mod}^n + L_{\max}^n)/3 - S$, *if $L_{\min}^n \geq S$;*
- $\frac{(S-L_{\min}^n)^3}{3(L_{\max}^n-L_{\min}^n)(L_{\mod}^n-L_{\min}^n)} + (L_{\min}^n + L_{\mod}^n + L_{\max}^n)/3 - S$, *if $L_{\min}^n < S < L_{\mod}^n$;*
- $\frac{(L_{\max}^n-S_u)^3}{3(L_{\max}^n-L_{\min}^n)(L_{\max}^m-L_{\mod}^n)}$, *if $L_{\mod}^n \leq S < L_{\max}^n$;*
- 0, *if $L_{\max}^n \leq S$.*

Recall that $\mathbf{E}(M(t)|n) = \mathbf{E}((L(t|n) - S_l)^+) - \mathbf{E}((L(t|n) - S_u)^+)$, and thus it is straightforward to find a closed form expression for $\mathbf{E}(M(t)|n)$ according to the locations of $L_{\min}^n$, $L_{\mod}^n$, and $L_{\max}^n$. Furthermore, the expected value of $M(t)$ is given by

$$\mathbf{E}(M(t)) = \sum_{n=1}^{N-1} \mathbf{E}(M(t)|n)P_N(n;t) + (S_u - S_l)P_N(N;t).$$

3.4 Approximation with Binomial Distribution

The conditional loss distribution of $L(t|n)$ is discrete with possible values lying between $L_{\min}^n$ and $L_{\max}^n$ and mode $L_{\mod}^n$. We therefore consider a binomial approximation to $L(t|n)$. For the approximation, we suppose there are $K+1$ possible outcomes of $L(t|n)$ in $[L_{\min}^n, L_{\max}^n]$ and the success probability of q. A success means an increase of loss Δl. We use the binomial distribution with parameters q and K to approximate the conditional loss distribution, yielding the probability mass function is

$$\mathbf{P}(L(t|n) = L_{\min}^n + i\Delta l) = \frac{K!}{i!(K-i)!}q^k(1-q)^{K-i}.$$

A reasonable choice for Δl is the smallest difference between two LGDs. We then let $K = \lceil (L_{max}^n - L_{min}^n)/\Delta l \rceil$ and reset $\Delta l = (L_{max}^n - L_{min}^n)/K$ to make sure that L_{max}^n is also the maximal outcome in our binomial distribution. The mass function of the binomial distribution is maximized at $\lfloor (K+1)q \rfloor$, so we choose q as

$$q = \frac{(L_{mod}^n - L_{min}^n)}{(K+1)\Delta l} = \frac{n(L_{mod}^n - L_{min}^n)}{(K+1)(L_{max}^n - L_{min}^n)}.$$

Set $K = 0$ and $q = 0$ if $L_{max}^n = L_{min}^n$.

After selecting q and K, the conditional expected loss on the tranche at t is approximated by

$$\mathbf{E}(M(t)|n) = \sum_{i=0}^{K} \mathbf{E}(M(t)|L(t|n),n)\mathbf{P}(L(t|n) = L_{min}^n + i\Delta l)$$

and

$$\mathbf{E}(M(t)) = \sum_{i=1}^{N-1} \mathbf{E}(M(t)|n)P_N(n;t) + (S_u - S_l)P_N(N;t).$$

3.5 Remarks

We can expect that when the underlying portfolio is close to homogeneous, the three approximation methods should work very well. In fact, when the portfolio is homogenous, the triangular approximation and binomial approximation reduce to the homogeneous approximation, and all methods give the exact tranche value. However, if the underlying assets have quite different LGDs, these methods might have significant error.

The complexity of the computation in the triangular and binomial approximation methods is always $O(N^2)$.

3.6 Numerical Examples

We illustrate the approximation methods discussed so far with some numerical results. In these examples, we suppose a continuously compounded interest rate of $r = 5\%$. We will look at the $0-3\%$, $3\%-7\%$, $7\%-10\%$, $10\%-15\%$, and $15\%-30\%$ CDO tranches. Each CDO has a portfolio of 200 underlying assets and a maturity of 5 years. The coupon rate of each tranche is 3% and is paid quarterly. All Monte Carlo simulation results are based on 10^5 replications. These parameters will be used in subsequent sections as well.

In our test examples, the underlying assets consist of four groups. The assets in groups 1, 2, 3, and 4 have hazard rates of 0.02, 0.025, 0.03, and 0.04 respectively. In CDO C1, the assets in groups 1 to 4 have LGDs of 0.63, 0.64, 0.66 and 0.67, respectively. In CDO C2, the assets have LGDs of 0.5, 0.6, 0.7 and 0.8. The results are displayed in Tables 1 and 2.

Table 1: Numerical Results for CDO C1

Tranche	MC	Homog.	Triangular	Binomial
$0-3\%$	168.83 (0.05)	168.59	169.00	169.67
$3-7\%$	110.93 (0.11)	110.00	112.52	114.51
$7-10\%$	15.01 (0.06)	14.43	15.59	17.40
$10-15\%$	-4.21 (0.01)	-4.29	-4.14	-4.02
$15-30\%$	-15.82689 (0.00005)	-15.82691	-15.82685	-15.82701

Table 2: Numerical Results for CDO C2

Tranche	MC	Homog.	Triangular	Binomial
$0-3\%$	170.04 (0.05)	168.59	179.62	174.69
$3-7\%$	116.31 (0.11)	110.00	158.52	138.69
$7-10\%$	18.66 (0.07)	14.43	29.99	38.70
$10-15\%$	-3.57 (0.02)	-4.29	-1.72	3.56
$15-30\%$	-15.8265 (0.0002)	-15.8269	-15.8189	-15.7862

The computation of each of the three approximations is practically instantaneous. The approximation methods work well when the LGDs are close to each other (CDO C1), but not when the LGDs vary widely (CDO C2). The performance of the binomial approximation depends highly on the choice of K and q.

4 THE APPROXIMATION METHODS WITH DEPENDENT DEFAULT TIMES

We now return to the general case in which the default times are linked through the Gaussian copula model. In this setting, $\tau_1, \ldots, \tau_N$ are conditionally independent, given the common factors $Z_1, \ldots, Z_d$. Therefore, we are able to avoid some computation by applying the approximations after sampling the factors.

4.1 The Conditional Approximations

In a factor Gaussian copula model, the default times are given by $\tau_i = F_i^{-1}(\Phi(W_i))$, $i = 1,\ldots,N$, and the W_i admit the representation in (3). Conditioning on the factors yields

$$
\begin{aligned}
F_i^{\mathbf{Z}}(t) &\equiv \mathbf{P}(\tau_i \le t | \mathbf{Z}) \\
&= \mathbf{P}(F_i^{-1}(\Phi(W_i)) \le t | \mathbf{Z}) \\
&= \mathbf{P}\left(b_i \varepsilon_i \le F_i(t) - \sum_{k=1}^{d} a_{ik} Z_k\right) \\
&= \Phi\left(F_i(t) - \sum_{k=1}^{d} a_{ik} Z_k\right).
\end{aligned}
\tag{4}
$$

Conditional on the factors $\mathbf{Z}$, the underlying assets in the portfolio become independent, so we can apply the approximation methods discussed in Section 3, but with $F_i(t)$ replaced by $F_i^{\mathbf{Z}}(t)$.

4.2 Numerical Examples

Now we consider CDOs C3 and C4, based on modifying CDOs C1 and C2, respectively. The only difference is that the underlying assets in CDOs C3 and C4 are correlated through a one-factor Gaussian copula model. The factor loading for groups 1–4 in the portfolio of CDOs C3 and C4 are 0.2, 0.3, 0.4, and 0.5 respectively. We use a one-factor model for simplicity.

We combine the approximations with numerical integration over the factor in these examples; the computing time of the approximation method is about one percent of the time for Monte Carlo simulation. The results are displayed in Tables 3 and 4.

Table 3: Numerical Results for CDO C3

Tranche	MC	Homog.	Triangular	Binomial
$0 - 3\%$	144.8 (0.1)	144.7	144.7	144.6
$3 - 7\%$	90.0 (0.3)	89.2	90.5	91.5
$7 - 10\%$	29.6 (0.2)	28.9	29.7	30.7
$10 - 15\%$	17.7 (0.2)	17.0	17.9	9.2
$15 - 30\%$	-6.4 (0.1)	-7.0	-6.6	-15.5

We can see a similar pattern as in the independent case. The approximation methods work well only when the underlying assets have similar LGDs. The performance of the homogeneous approximation is much more stable than the other two methods. When the underlying portfolio is close to homogeneous, the triangular approximation is better than the homogeneous approximation, but is worse when the underlying portfolio is no longer "homogeneous". Both work much better than the binomial approximation.

Table 4: Numerical Results for CDO C4

Tranche	MC	Homog.	Triangular	Binomial
$0 - 3\%$	144.6 (0.1)	144.7	150.2	143.0
$3 - 7\%$	92.5 (0.3)	89.2	116.3	104.3
$7 - 10\%$	32.0 (0.2)	28.9	51.6	41.8
$10 - 15\%$	21.0 (0.2)	17.0	33.2	27.1
$15 - 30\%$	-3.6 (0.2)	-7.0	-2.7	-10.2

5 CONTROL VARIATE METHOD FOR MONTE CARLO SIMULATION

The approximation methods we discuss in previous sections work well only when the underlying portfolio is close to homogeneous. In more general cases, it would be difficult to determine the magnitude of the approximation error. Moreover, in multifactor models, simulation becomes necessary and applying approximation methods has no advantage (and could even be worse than plain Monte Carlo). Improving simulation accuracy becomes the key problem we need to solve.

One way to improve the simulation accuracy is to apply the control variate technique, one of the most effective variance reduction techniques. Recall that in the homogeneous portfolio case, we are able to calculate the tranche value analytically. Using a homogeneous portfolio, where each asset has LGD of $\bar{l} = \sum_{i=1}^{N} l_i / N$, to approximate the real portfolio, we are able to calculate the tranche value on the approximate portfolio. The value of the new tranche is then an attractive control variate for the desired tranche value.

Explicitly, we define the control variate as

$$
\begin{aligned}
\hat{V}(\tau) &= (1+c)\sum_{i=1}^{m} D(T_i)\hat{M}(T_i) - \sum_{i=1}^{m-1} D(T_{i+1})\hat{M}(T_i) \\
&\quad - c(S_u - S_l)\sum_{i=1}^{m} D(T_i),
\end{aligned}
$$

where $\hat{M}(t) = (n\bar{l} - S_l)^+ - (n\bar{l} - S_u)^+$, given that there are n defaults before t. In our control variate method, we use $V(\tau) - \hat{\beta}(\hat{V}(\tau) - \mathbf{E}(\hat{V}(\tau))$ as an unbiased estimator of the desired tranche value. The optimal coefficient β is given, as usual, by

$$
\beta = \frac{\mathbf{Cov}(V(\tau), \hat{V}(\tau))}{\mathbf{Var}(\hat{V}(\tau))};
$$

we use the sample estimate $\hat{\beta}$ calculated from the simulated replications. The exact value of $\mathbf{E}(\hat{V}(\tau))$ can be calculated in the independent case or in a one-factor model. In multifactor models, we replace it with its conditional expectation $\mathbf{E}(\hat{V}(\tau)|\mathbf{Z})$, since $\hat{V}(\tau) - \mathbf{E}(\hat{V}(\tau)|\mathbf{Z})$ has expectation zero.

We use CDOs C1, C2, C3, and C4 as examples to illustrate our control variate method. In all of the examples, the control variate method reduces the standard error significantly (especially in the correlated assets case), and has almost the same computing time as plain Monte Carlo simulation.

Table 5: Numerical Results for CDO C1

Tranche	Plain MC	Control Variate
$0-3\%$	168.83 (0.05)	168.832 (0.001)
$3-7\%$	110.9 (0.1)	110.884 (0.003)
$7-10\%$	15.01 (0.06)	14.972 (0.002)
$10-15\%$	-4.21 (0.01)	-4.2138 (0.0006)
$15-30\%$	-15.82689 (5E-5)	-15.826882 (8E-6)

Table 6: Numerical Results for CDO C2

Tranche	Plain MC	Control Variate
$0-3\%$	170.04 (0.05)	170.050 (0.008)
$3-7\%$	116.3 (0.1)	116.25 (0.02)
$7-10\%$	18.66 (0.07)	18.63 (0.01)
$10-15\%$	-3.57 (0.02)	-3.579 (0.005)
$15-30\%$	-15.8265 (0.0002)	-15.8264 (0.0001)

Table 7: Numerical Results for CDO C3

Tranche	MC Plain MC	Control Variate
$0-3\%$	144.8 (0.1)	144.650 (0.002)
$3-7\%$	90.0 (0.3)	89.621 (0.002)
$7-10\%$	29.5 (0.2)	29.331 (0.002)
$10-15\%$	17.7 (0.2)	17.553 (0.002)
$15-30\%$	-6.4 (0.1)	-6.627 (0.002)

Table 8: Numerical Results for CDO C4

Tranche	Plain MC	Control Variate
$0-3\%$	144.6 (0.1)	144.43 (0.01)
$3-7\%$	92.5 (0.3)	92.08 (0.02)
$7-10\%$	32.0 (0.2)	31.80 (0.01)
$10-15\%$	21.0 (0.2)	20.83 (0.02)
$15-30\%$	-3.6 (0.2)	-3.81 (0.02)

6 CONCLUDING REMARKS

In this paper, we have proposed approximation methods for pricing CDO tranches. These methods construct the loss distribution by approximating the conditional loss distribution given the number of defaults, and recursively calculating the exact distribution of the number of defaults. We tested three approximations to the conditional loss distribution;

the methods work well when the loss given default is nearly constant across assets in the portfolio, but not when these losses vary widely. Topics for further investigation include improving the specification of the conditional distributions and extending the methods to handle sensitivities.

We have also developed a variance reduction technique by using an approximate tranche value as a control variate. Our examples indicate that this method is fast and can significantly improve simulation precision.

A PROOF OF PROPOSITION 1

Proof When $L_{\max}^n \leq S$, we obviously have $\mathbf{E}((L(t|n) - S)^+) = 0$. When $L_{\min}^n \geq S$,

$$\mathbf{E}((L(t|n) - S)^+) = \mathbf{E}(L(t|n)) - S = \frac{L_{\min}^n + L_{\mathrm{mod}}^n + L_{\max}^n}{3} - S.$$

When $L_{\min}^n < S < L_{\mathrm{mod}}^n$,

$$\mathbf{E}((L(t|n) - S)^+) = \int_S^{L_{\max}^n} (x - S)f(x)dx$$
$$= \int_{L_{\min}^n}^{L_{\max}^n} (x - S)f(x)dx - \int_{L_{\mathrm{mod}}^n}^S (x - S)f(x)dx$$
$$= \frac{(S - L_{\min}^n)^3}{3(L_{\max}^n - L_{\min}^n)(L_{\mathrm{mod}}^n - L_{\min}^n)} + \frac{L_{\min}^n + L_{\mathrm{mod}}^n + L_{\max}^n}{3} - S.$$

When $L_{\mathrm{mod}}^n < S < L_{\max}^n$,

$$\mathbf{E}((L(t|n) - S)^+) = \int_S^{L_{\max}^n} (x - S)f(x)dx$$
$$= \frac{(L_{\max}^n - S_u)^3}{3(L_{\max}^n - L_{\min}^n)(L_{\max}^m - L_{\mathrm{mod}}^n)}.$$

∎

REFERENCES

Andersen, L., J. Sidenius, and S. Basu. 2003. All your hedges in one basket. *Risk* 16:67–72.

Bruyere, R., R. Cont, R. Copinot, L. Fery, C. Jaecl, T. Spitz, and G. Smart. 2006. *Credit Derivatives and Structured Credit: A Guide for Investors*. Chichester, England: Wiley.

Chen, Z., and P. Glasserman. 2006. Fast pricing of basket default swaps. To appear in *Operations Research*.

Duffie, D., and K. Singleton. 2003. *Credit Risk: Pricing, Measurement, and Management*. Princeton, New Jersey: Princeton University Press.

Gupton, G., C. Finger, and M. Bhatia. 1997. Creditmetrics Technical Document. Technical report, J.P. Morgan & Co., New York.

Hull, J., and A. White. 2004. Valuation of a CDO and an nth to default CDS without Monte Carlo simulation. *Journal of Derivatives* 12,2.

Li, D. 2000. On default correlation: A copula function approach. *Journal of Fixed Income* 9:43–54.

Schönbucher, P. 2003. *Credit Derivatives Pricing Models: Model, Pricing and Implementation*. Princeton, New Jersey: Princeton University Press.

AUTHOR BIOGRAPHIES

ZHIYONG CHEN is a Vice President in the Financial Analytics & Structured Transactions Group at Bear Stearns & Co., Inc. He obtained his Ph.D. degree in the Department of Industrial Engineering and Operations Research at Columbia University. His research focuses on credit risk modeling and Monte Carlo simulation methods.

PAUL GLASSERMAN is the Jack R. Anderson Professor at Columbia Business School. His research focuses on modeling and computational problems in derivatives pricing and risk management. His book, *Monte Carlo Methods in Financial Engineering* (Springer, 2004), received the 2005 Outstanding Simulation Publication Award and the 2006 Lanchester Prize.