Pricing Correlation-Dependent Derivatives Based on Exponential Approximations to the Hockey Stick Function*

Ian Iscoe[†] Ken Jackson[‡] Alex Kreinin[§] Xiaofang Ma[¶]

January 24, 2007

Abstract

Correlation-dependent derivatives, such as Asset-Backed Securities (ABS) and Collateralized Debt Obligations (CDO), have grown rapidly. Factor models in the conditional independence framework are widely used in practice to capture the correlated default events of the underlying obligors. An essential part of these models is the accurate and efficient evaluation of the expected loss of the specified tranche, conditional on a given value of a systematic factor (or values of a set of systematic factors). Unlike other papers that focus on how to evaluate the loss distribution of the underlying pool, in this paper we focus on the tranche loss function itself. It is approximated by a sum of

*This research was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

[†]Algorithmics Inc., 185 Spadina Avenue, Toronto, ON, M5T 2C6, Canada; Ian.Iscoe@algorithmics.com

[‡]Department of Computer Science, University of Toronto, 10 King's College Rd, Toronto, ON, M5S 3G4, Canada; krj@cs.utoronto.ca

[§]Algorithmics Inc., 185 Spadina Avenue, Toronto, ON, M5T 2C6, Canada; Alex.Kreinin@algorithmics.com

[¶]Department of Computer Science, University of Toronto, 10 King's College Rd, Toronto, ON, M5S 3G4, Canada; csxfma@cs.utoronto.ca

exponentials so that the conditional expectation can be evaluated in closed form without having to evaluate the pool loss distribution. As an example, we apply this approach to synthetic CDO pricing. Numerical results show that it is efficient.

1 Introduction

Correlation-dependent derivatives, such as Asset-Backed Securities (ABS) and Collateralized Debt Obligations (CDO), have grown rapidly. An ABS is a security based on a pool of non-mortgage assets. To create an ABS, a corporation creates a trust or a special purpose vehicle to which it sells the assets. While it is common to speak of the corporation as the issuer of the ABS, legally, it is the trust or the special purpose vehicle that sells securities to investors. An ABS can be structured into different tranches that have different credit ratings. A CDO is a security based on a pool of, generally non-mortgage, assets. Depending on the nature of the collateralized assets, a CDO may be called a collateralized loan obligation, a collateralized bond obligation or a synthetic CDO if it holds only loans, bonds or credit default swaps. Like an ABS, a CDO is usually structured into tranches offering investors various maturity and credit risk characteristics. Tranches are categorized as senior, mezzanine, and subordinated/equity, according to their degree of credit risk. If there are defaults or the CDO's collateral otherwise underperforms, scheduled payments to senior tranches take precedence over those of mezzanine tranches, and scheduled payments to mezzanine tranches take precedence over those to subordinated/equity tranches. Senior tranches typically have credit ratings of A to AAA, mezzanine tranches typically have ratings of B to BBB, while equity tranches are usually not rated. The ratings reflect both the credit quality of the underlying collateral as well as the amount of protection a given tranche is afforded by tranches that are subordinate to it.

Factor models in the conditional independence framework are widely used in practice to price these correlation-dependent derivatives so that analytic or semi-analytic formulas are available. An essential part of these models is the accurate and efficient evaluation of the expected loss of the specified tranche, conditional on a given value of a systematic factor (or correspondingly values of a set of systematic factors). To be specific, the problem is how to evaluate the conditional expectation of the piecewise linear payoff function of the loss Z

$$f(Z) = \min\left(u - \ell, (Z - \ell)^{+}\right), \tag{1}$$

where $x^{+} = \max(x, 0)$, $Z = \sum_{k=1}^{K} Z_k$, Z_k are conditionally mutually independent, but not necessarily identically distributed, nonnegative random variables in a conditional independence framework (see Section 2 for explanations), and ℓ and u are the attachment and the detachment points of the tranche, respectively, satisfying $u > \ell \geq 0$. Generally Z_k, for obligor k, is the product of the two components: a random variable directly related to its credit rating and a loss-given-default or mark-to-market related value. The payoff function, f, is also known as the stop-loss function in actuarial science [3], [16].

Note that the expectation of a function of a random variable depends on two factors: the distribution of the underlying random variable and the function itself. A standard approach to compute the expectation of a function of a random variable is to compute firstly the distribution of the underlying random variable, Z in our case, and then to compute the expectation of the given function, possibly using its special properties. Almost all research in finance [1], [7], [10], [15], [17] and in actuarial science [5], [19], to name a few, has focused on the first part due to the piecewise linearity of the payoff function.

In this paper, we propose a new approach to solving the problem. We approximate the non-smooth function f by a sum of exponentials over $[0, \infty)$. Based on this approximation, the conditional expectation can be computed from a series of simple expectations. Consequently, we do not need to compute the distribution of Z.

The remainder of this paper is organized as follows. The details of our approach outlined above are described in Section 2. As an example, this method is applied to synthetic CDO pricing in Section 3. The paper ends with some conclusions in Section 4, in which we summarize the advantages of our method over others, and indicate its scope of applicability.

2 Conditional expectation based on an exponential approximation

In the conditional independence framework, a central problem is how to evaluate the expectation

$$\mathbb{E}[f(Z)] = \int_M \mathbb{E}_M [f(Z)] \, \mathrm{d}\Phi(M),$$

where $\Phi(M)$ is the distribution of an auxiliary factor $\mathscr{M}$ (which can be a scalar or a vector),

$$\mathbb{E}_M [f(Z)] \equiv \mathbb{E} [f(Z) \,|\mathscr{M} = M]$$

and

$$Z = \sum_{k=1}^{K} Z_k, \tag{2}$$

where $Z_k \geq 0$ are mutually independent random variables, conditional on $\mathscr{M}$. It is obvious that Z is nonnegative. We denote by Ψ_M the distribution of Z conditional on $\mathscr{M} = M$, so that

$$\mathbb{E}_M [f(Z)] = \int_z f(z) \, \mathrm{d}\Psi_M(z). \tag{3}$$

Due to the piecewise linearity of the function f defined by (1), it is clear that once the distribution Ψ_M is obtained, the conditional expectation $\int_z f(z) \, \mathrm{d}\Psi_M(z)$ can be readily computed. Most research has focused on how to evaluate the conditional distribution of Z given the conditional distributions of Z_k. A fundamental result about a sum of independent random variables states that Z's distribution can be computed as the convolution of Z_k's distributions. Numerically, this idea is realized through forward and inverse fast Fourier transformations (FFT). A disadvantage of this approach is that it may be very slow when there are many obligors due to the number of convolutions to be calculated. For pools with special structures, it might be much slower than methods that are specially designed for those pools, such as recursive methods proposed by De Pril [5] and Panjer [19] and their extensions discussed in [3] and [16], and the one proposed by Jackson, Kreinin, and Ma [15] for portfolios where the Z_k sit on a properly chosen common lattice. To avoid computing too many convolutions, the target distribution can be approximated by parametric distributions matching the first few moments of the true distribution. For a large pool, a normal approximation is a natural choice as a consequence of the central limit theorem and due to its simplicity, although it may not capture some important properties, such as skewness and fat tails.

To capture these important properties for medium to large portfolios, compound approximations, such as the compound Poisson [13], improved compound Poisson [9], compound binomial and compound Bernoulli [20] distributions have been used. They have proved to be very successful, since they match not only the first few moments, but, most importantly, they match the tails much better than either normal or normal power distributions do. A key step in a method based on these compound approximations is the computation of convolutions by FFTs. As a result, the computational complexity of such an algorithm is superlinear in K, the number of terms in the sum (2).

As an alternative, in this paper, we propose an algorithm for which the computational complexity is linear in K. We focus on the stop-loss function f, instead of the distribution Ψ_M of Z. To emphasis the role of the attachment and the detachment points ℓ and u, we denote $f(x)$ by $f(x;\ell,u)$ and introduce an auxiliary function $h(x)$ defined on $[0,\infty)$: $h(x) = 1-x$ if $x \le 1$, 0 otherwise. Then we have

$$f(x) = f(x;\ell,u) = u\left[1 - h\left(\frac{x}{u}\right)\right] - \ell\left[1 - h\left(\frac{x}{\ell}\right)\right]. \tag{4}$$

In particular, if $\ell = 0$, we have

$$f(x;0,u) = \min\left(u, x^+\right) = \min\left(u, x\right) = u\left[1 - h\left(\frac{x}{u}\right)\right].$$

Note that $h(x)$ is independent of the constants ℓ and u. Therefore, if it can be approximated by a sum of exponentials over $[0,\infty)$, it is clear that $f(x;\ell,u)$ can be approximated by a sum of exponentials. Let

$$h(x) \approx \sum_{n=1}^{N} \omega_n \exp(\gamma_n x), \tag{5}$$

where ω_n and γ_n are complex numbers. Then from (4) we can see that $f(x;\ell,u)$ can be approximated by a sum of exponentials:

$$f(x;\ell,u) \approx u\left[1-\sum_{n=1}^{N}\omega_n \exp\left(\gamma_n \frac{x}{u}\right)\right] - \ell\left[1-\sum_{n=1}^{N}\omega_n \exp\left(\gamma_n \frac{x}{\ell}\right)\right]$$

$$\approx (u-\ell) - u\sum_{n=1}^{N}\omega_n \exp\left(\frac{\gamma_n}{u}x\right) + \ell\sum_{n=1}^{N}\omega_n \exp\left(\frac{\gamma_n}{\ell}x\right). \tag{6}$$

Based on this expression the conditional expectation $\mathbb{E}_M[f(Z)]$ defined in (3) can be computed as follows:

$$\begin{aligned}
&\mathbb{E}_M[f(Z)] \\
= \; & \int_z f(z)\,\mathrm{d}\Psi_M(z) \\
\approx \; & \int_z \left[(u-\ell) - u\sum_{n=1}^{N}\omega_n \exp\left(\frac{\gamma_n}{u}z\right) + \ell\sum_{n=1}^{N}\omega_n \exp\left(\frac{\gamma_n}{\ell}z\right)\right]\mathrm{d}\Psi_M(z) \\
= \; & (u-\ell) - u\sum_{n=1}^{N}\omega_n \int_z \exp\left(\frac{\gamma_n}{u}z\right)\mathrm{d}\Psi_M(z) \\
& +\ell\sum_{n=1}^{N}\omega_n \int_z \exp\left(\frac{\gamma_n}{\ell}z\right)\mathrm{d}\Psi_M(z) \\
= \; & (u-\ell) \\
& -u\sum_{n=1}^{N}\omega_n \int_{z_1,\ldots,z_K}\prod_{k=1}^{K}\exp\left(\frac{\gamma_n}{u}z_k\right)\mathrm{d}\Psi_{M,1}(z_1)\cdots\mathrm{d}\Psi_{M,K}(z_K) \\
& +\ell\sum_{n=1}^{N}\omega_n \int_{z_1,\ldots,z_K}\prod_{k=1}^{K}\exp\left(\frac{\gamma_n}{\ell}z_k\right)\mathrm{d}\Psi_{M,1}(z_1)\cdots\mathrm{d}\Psi_{M,K}(z_K) \\
= \; & (u-\ell) - u\sum_{n=1}^{N}\omega_n \prod_{k=1}^{K}\mathbb{E}_M\left[\exp\left(\frac{\gamma_n}{u}Z_k\right)\right] \\
& +\ell\sum_{n=1}^{N}\omega_n \prod_{k=1}^{K}\mathbb{E}_M\left[\exp\left(\frac{\gamma_n}{\ell}Z_k\right)\right],
\end{aligned} \tag{7}$$

where $\Psi_{M,k}$ is the conditional distribution of Z_k, $\mathbb{E}_M[\exp(cZ_k)]$ is the conditional expectation of $\exp(cZ_k)$, for $c = \frac{\gamma_n}{\ell}$ or $\frac{\gamma_n}{u}$, respectively. The last equality holds by noting that Z_k, thus cZ_k, are mutually independent conditional on a given value of $\mathscr{M}$. In this way we can see that, to compute the conditional expectation $\mathbb{E}_M[f(Z)]$, we only need to compute the conditional expectations $\mathbb{E}_M[\exp(cZ_k)]$ of individual obligors.

Since $h(z)$ is independent of the constants ℓ and u, for a given approximation accuracy the coefficients ω_n and γ_n for (5) need to be computed only once and the number of terms required can be determined a priori. As shown in a separate paper by the authors [12], the maximum absolute error in the approximation (5) is roughly proportional to $1/N$:

N	25	50	100	200	400
Max absolute error	6.4×10^{-3}	3.2×10^{-3}	1.6×10^{-3}	8×10^{-4}	4×10^{-4}

Table 1: The maximum absolute error in the approximation (5) for several values of N

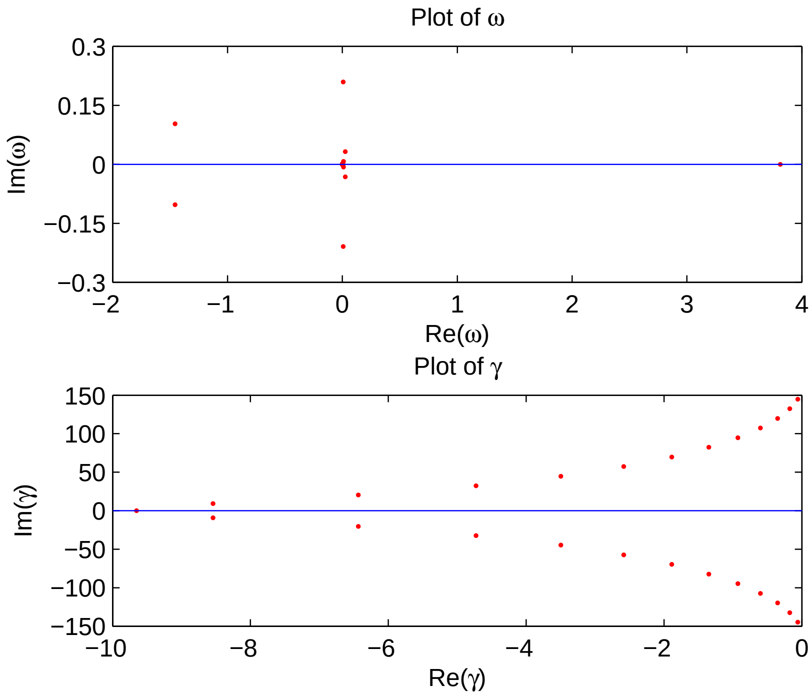

Figure 1: The exponents and the weights of the 25-term exponential approximation

As an example, the parameters γ_n and ω_n for a 25-term approximation are plotted in Figure 1. The top panel shows the values of ω_n; the bottom panel shows the values of γ_n. It is proved in [12] that, if γ_n is real, then ω_n is also real, and if γ_i and γ_j are a complex conjugate pair, then the corresponding ω_i and ω_j are also a complex conjugate pair, and vice versa. The data plotted in Figure 1 has this property. Exploiting this property, we can simplify the summations in (7) by noting that the sum of the complex conjugated i-th and j-th terms equals twice the real part of either one of these two terms. The bottom panel also shows that the real part of each γ_n is strictly negative. This property guarantees that the exponential approximation of (5) converges to zero as $x \to \infty$, and thus guarantees the existence of the conditional expectation $\mathbb{E}_M[\exp(cZ_k)]$. For a more complete discussion of the exponential approximation (5), see [12]. In particular, it is proved therein, that each γ_n has a nonpositive real part for any N.

3 Pricing synthetic CDOs

3.1 Overview of pricing methods

We illustrate our new method by applying it to synthetic CDO pricing. The underlying collateral of a synthetic CDO is a set of credit default swaps (CDSs). Factor models, such as the reduced-form model proposed by Laurent and Gregory [17] and the structural model proposed Vasicek [22] and Li [18] are widely used in practice to obtain analytic or semi-analytic formula to price synthetic CDOs efficiently. For a comparative analysis of different copula models, please see the paper by Burtschell, Gregory and Laurent [4].

Both exact and approximate methods for loss-distribution evaluation have been studied in [15]. Here we apply the exponential-approximation method to synthetic CDO pricing. An important advantage of our new approach is that it applies to more general models that incorporate dynamic interest rates, dynamic recovery rates, and other dynamic properties. However, as with many existing methods, our new method depends on the conditional independence framework.

3.2 The pricing equation and the Gaussian copula model

To illustrate our method, we use a simple one-factor Gaussian copula model. It is assumed that the interest rates are deterministic and the recovery rate corresponding to each underlying name is constant. Let $0 < t_1 < t_2 < \cdots < t_n = T$ be the set of premium dates, with T denoting the maturity of the CDO, and $d_1, d_2, \ldots, d_n$ be the set of corresponding discount factors. Suppose there are K names in the pool with recovery-adjusted notional values $LGD_1, LGD_2, \ldots, LGD_K$ in properly chosen units. Let $\mathscr{L}_i^P$ be the pool's cumulative losses up to time t_i and ℓ be the attachment point of a specified tranche of thickness S. An attachment point of a tranche is a threshold that determines whether some losses of the pool shall be absorbed by this tranche, *i.e.,* if the realized losses of the pool are less than ℓ, then the tranche will not suffer any loss, otherwise it will absorb an amount up to S. Accordingly, the detachment point of the tranche is $u = S + \ell$. Thus the loss absorbed by the specified tranche up to time t_i is $L_i = \min\left(S, (\mathscr{L}_i^P - \ell)^+\right)$. If we further assume the fair spread s for the tranche is a constant, then it can be calculated from the equation (see e.g., [6], [13])

$$s = \frac{\mathbb{E}\left[\sum_{i=1}^n (L_i - L_{i-1}) d_i\right]}{\mathbb{E}\left[\sum_{i=1}^n (S - L_i)(t_i - t_{i-1}) d_i\right]} = \frac{\sum_{i=1}^n \mathbb{E}\left[(L_i - L_{i-1}) d_i\right]}{\sum_{i=1}^n \mathbb{E}\left[(S - L_i)(t_i - t_{i-1}) d_i\right]}, \tag{8}$$

with $t_0 = 0$ and $\mathbb{E}[L_0] = 0$.

Since the discount factors d_i are deterministic, it follows from (8) that the problem of computing the fair spread s reduces to evaluating the expected cumulative losses $\mathbb{E}[L_i]$, $i = 1, 2, \ldots, n$. In order to compute this expectation, we have to specify the default processes for each name and the correlation structure of the default events. One-factor models were first introduced by Vasicek [22] to estimate the loan loss distribution and then generalized by Li [18], Gordy and Jones [8], Hull and White [10], Iscoe, Kreinin and Rosen [14], Laurent and Gregory [17], and Schönbucher [21], to name a few.

Let τ_k be the default time of name k. Assume the risk-neutral default probabilities $\pi_k(t) = \mathbb{P}(\tau_k < t)$, $k = 1, 2, \ldots, K$, that describe the default-time distributions of all K names are available, where τ_k and t take discrete values from $\{t_1, t_2, \ldots, t_n\}$. The dependence structure of the default times are determined in terms of their creditworthiness indices Y_k, which are defined by

$$Y_k = \beta_k X + \sigma_k \varepsilon_k, \qquad k = 1, 2, \ldots, K, \tag{9}$$

where X is the systematic risk factor, ε_k are idiosyncratic factors that are mutually independent and are also independent of X; β_k and σ_k are constants satisfying the relation $\beta_k^2 + \sigma_k^2 = 1$. These risk-neutral default probabilities and the creditworthiness indices are related by the threshold model

$$\pi_k(t) = \mathbb{P}\left(\tau_k < t\right) = \mathbb{P}\left(Y_k < H_k(t)\right), \tag{10}$$

where $H_k(t)$ is the default barrier of the k-th name at time t.

Thus the correlation structure of default events is captured by the systematic risk factor X. Conditional on a given value x of X, all default events are independent. If we further assume, as we do, that X and ε_k follow the standard normal distribution, then Y_k is a standard normal random variable and from (10) we have $H_k(t) = \Phi^{-1}\left(\pi_k(t)\right)$. Furthermore, the correlation between two different indices Y_i and Y_j is $\beta_i\beta_j$.

The conditional, risk-neutral default-time distribution is defined by

$$\pi_k(t;x) = \mathbb{P}\left(Y_k < H_k(t) | X = x\right). \tag{11}$$

Thus from (9) and (11) we have

$$\pi_k(t;x) = \Phi\left(\frac{H_k(t) - \beta_k x}{\sigma_k}\right). \tag{12}$$

The conditional and unconditional risk-neutral default-time probabilities at the premium date t_i are denoted by $\pi_k(i;x)$ and $\pi_k(i)$, respectively.

In this conditional independence framework, the expected cumulative tranche losses $\mathbb{E}\left[L_i\right]$ can be computed as

$$\mathbb{E}\left[L_i\right] = \int_{-\infty}^{\infty} \mathbb{E}_x\left[L_i\right] \mathrm{d}\Phi(x), \tag{13}$$

where $\mathbb{E}_x\left[L_i\right] = \mathbb{E}_x\left[\min\left(S, (\mathscr{L}_i^P - \ell)^+\right)\right]$ is the expectation of L_i conditional on the value of X being x, where $\mathscr{L}_i^P = \sum_{k=1}^K LGD_k \mathbf{1}_{\{Y_k < H_k(t_i)\}}$, and the indicators $\mathbf{1}_{\{Y_k < H_k(t_i)\}}$ are mutually independent conditional on X. Generally, the integration in (13) needs to be evaluated numerically using an appropriate quadrature rule:

$$\mathbb{E}\left[L_i\right] \approx \sum_{m=1}^{M} \mathrm{w}_m \mathbb{E}_{x_m}\left[\min\left(S, (\mathscr{L}_i^P - \ell)^+\right)\right]. \tag{14}$$

In the notation of Section 2, $Z = \mathscr{L}_i^P$, $Z_k = LGD_k \mathbf{1}_{\{Y_k < H_k(t_i)\}}$, and $\mathscr{M} = X$.

3.3 CDO pricing based on the exponential approximation

Notice from formula (14) that the fundamental problem in CDO pricing is how to evaluate the conditional expected loss $\mathbb{E}_{x_m}\left[\min\left(S,(\mathscr{L}_i^P-\ell)^+\right)\right]$ with a given value x_m of X. Since

$$\min\left(S,(\mathscr{L}_i^P-\ell)^+\right)=f(\mathscr{L}_i^P;\ell,\ell+S), \tag{15}$$

from (6) we see that

$$\begin{aligned}\min\left(S,(\mathscr{L}_i^P-\ell)^+\right)\approx&(\ell+S)\left[1-\sum_{n=1}^N\omega_n\exp\left(\gamma_n\frac{\mathscr{L}_i^P}{\ell+S}\right)\right]-\ell\left[1-\sum_{n=1}^N\omega_n\exp\left(\gamma_n\frac{\mathscr{L}_i^P}{\ell}\right)\right]\\=&S-(\ell+S)\sum_{n=1}^N\omega_n\exp\left(\frac{\gamma_n}{\ell+S}\mathscr{L}_i^P\right)+\ell\sum_{n=1}^N\omega_n\exp\left(\frac{\gamma_n}{\ell}\mathscr{L}_i^P\right).\end{aligned}$$

As a special case of (7) we have

$$\begin{aligned}\mathbb{E}_{x_m}\left[\min\left(S,(\mathscr{L}_i^P-\ell)^+\right)\right]\approx&S-(\ell+S)\sum_{n=1}^N\omega_n\mathbb{E}_{x_m}\left[\exp\left(\frac{\gamma_n}{\ell+S}\sum_{k=1}^K LGD_k\mathbf{1}_{\{Y_k<H_k(t_i)\}}\right)\right]\\&+\ell\sum_{n=1}^N\omega_n\mathbb{E}_{x_m}\left[\exp\left(\frac{\gamma_n}{\ell}\sum_{k=1}^K LGD_k\mathbf{1}_{\{Y_k<H_k(t_i)\}}\right)\right]\\=&S-(\ell+S)\sum_{n=1}^N\omega_n\prod_{k=1}^K\mathbb{E}_{x_m}\left[\exp\left(\frac{\gamma_n}{\ell+S}LGD_k\mathbf{1}_{\{Y_k<H_k(t_i)\}}\right)\right]\\&+\ell\sum_{n=1}^N\omega_n\prod_{k=1}^K\mathbb{E}_{x_m}\left[\exp\left(\frac{\gamma_n}{\ell}LGD_k\mathbf{1}_{\{Y_k<H_k(t_i)\}}\right)\right],\end{aligned} \tag{16}$$

where

$$\begin{aligned}\mathbb{E}_{x_m}\left[\exp\left(\frac{\gamma_n}{\ell+S}LGD_k\mathbf{1}_{\{Y_k<H_k(t_i)\}}\right)\right]&=\pi_k(i;x_m)\exp\left(\frac{\gamma_n}{\ell+S}LGD_k\right)+(1-\pi_k(i;x_m)),\\\mathbb{E}_{x_m}\left[\exp\left(\frac{\gamma_n}{\ell}LGD_k\mathbf{1}_{\{Y_k<H_k(t_i)\}}\right)\right]&=\pi_k(i;x_m)\exp\left(\frac{\gamma_n}{\ell}LGD_k\right)+(1-\pi_k(i;x_m)),\end{aligned}$$

since $\mathbf{1}_{\{Y_k<H_k(t_i)\}}=1$ with probability $\pi_k(i;x_m)$ and 0 with probability $1-\pi_k(i;x_m)$, and LGD_k is a constant.

3.4 Numerical results

In this section we present numerical results comparing the accuracy and the computational time for our new exponential-approximation method and the exact method, which we denote by JKM, proposed in [15]. The results presented below are based on a sample of 15 pools. For each pool, the number of names K is either 100, 200, or 400. The number of homogeneous groups in each pool is one of $1, 2, 4, 5$, or $K/10$, and all homogeneous groups in a given pool have an equal number of names. The notional values for each pool are summarized in Table 2. For example, the 200-name pool with local pool ID = 3 consists of four homogeneous groups with the notional values 50, 100, 150, and 200, respectively. For convenience, we also labeled each pool with a global pool ID. For each of the 100-name pools, the global and the local IDs coincide. For each of the 200- and 400-name pools, its global pool ID (GID) is its local pool ID plus 5 or 10, respectively. For example, a 200-name pool with local ID = 3 has GID = 8.

Local Pool ID	1	2	3	4	5
Notional values	100	50, 100	50, 100, 150, 200	20, 50, 100, 150, 200	$10, 20, \ldots, K$

Table 2: Selection of notional values of K-name pools

For each name, the risk-neutral cumulative default probabilities are one of two types, I and II, as defined in Table 3.

Type	1 yr.	2 yrs.	3 yrs.	4 yrs.	5 yrs.
I	0.0007	0.0030	0.0068	0.0119	0.0182
II	0.0044	0.0102	0.0175	0.0266	0.0372

Table 3: Risk-neutral cumulative default probabilities

The recovery rate is assumed to be 40% for all names. Thus the LGD of name k is $0.6N_k$. The maturity of a CDO deal is five years (*i.e.*, $T = 5$) and the premium dates are $t_i = i, i = 1, \ldots, 5$ years from today ($t_0 = 0$). The continuously compounded interest rates are $r_1 = 4.6\%, r_2 = 5\%, r_3 = 5.6\%, r_4 = 5.8\%$ and $r_5 = 6\%$. Thus the corresponding discount factors, defined by $d_i = \exp(-t_i r_i)$, are $0.9550, 0.9048, 0.8454, 0.7929$ and 0.7408, respectively. All CDO pools have five tranches that are determined by the attachment points (ℓ's) of the tranches. For this experiment, the five attachment points are: 0, 3%, 4%, 6.1% and 12.1%.

The constants β_k's lie in $[0.3, 0.5]$. In practice, the β_k's are known as tranche correlations and are taken as input to the model.

All methods for this experiment were coded in Matlab and the programs were run on a Pentium III 700 PC. The results are presented in Tables 4, 5 and 6.

The accuracy comparison results are presented in Table 4. The four numbers in each pair of brackets in the main part of the table are the spreads in basis points for the first four tranches of the corresponding pool. For example, $(2248.16, 927.59, 605.52, 248.31)$ are the spreads evaluated by the JKM method for the first four tranches of the 200-name homogeneous pool (with global pool ID, GID $= 6$). The entries under "25-term", "100-term", and "400-term" are the spreads evaluated using the exponential-approximation method with 25, 100 and 400 terms, respectively. From the table we can see that, as the number of terms increases, the accuracy of the spreads improves. To better illustrate the accuracy of our new approach, the relative errors in the spreads obtained using exponential approximations, with different numbers of terms, compared to the spreads computed by the exact JKM method are plotted in Figures 2 and 3.

The CPU times used by the JKM method and the exponential-approximation method using different numbers of terms for the test pools are presented in Tables 5 and 6, respectively. In Table 5 the numbers under "First tranche" and "First four tranches" are the times in seconds used by the exact JKM method to evaluate the spread for the first tranche and the spreads for the first four tranches of each pool, respectively. In Table 6 the numbers under "First tranche" and "First four tranches" are the times in seconds used by the exponential-approximation method using 25, 50, 100, 200 and 400 terms to evaluate the spread for the first tranche and the spreads for the first four tranches of each pool, respectively. Note that, for the exponential-approximation method, its CPU time is independent of the pool structure: its computational cost depends only on the number of names and the number of terms in the exponential approximation. It is interesting to note that, for a given pool, to evaluate any single tranche using the exponential-approximation method takes about as much time as to evaluate any other tranche. On the other hand, for the exact method, calculating the spread

GID	Exact	25-term	100-term	400-term
1	(2167.69, 925.62, 616.56, 255.67)	(2165.21, 930.60, 615.90, 255.66)	(2167.06, 926.63, 616.59, 255.65)	(2167.54, 925.88, 616.56, 255.67)
2	(2142.13, 945.03, 615.79, 257.43)	(2141.54, 943.15, 616.96, 257.50)	(2141.94, 944.66, 616.03, 257.43)	(2142.08, 944.94, 615.85, 257.42)
3	(2128.39, 941.00, 618.88, 258.75)	(2128.80, 940.35, 618.92, 258.89)	(2128.25, 941.19, 618.85,258.77)	(2128.35, 941.05, 618.86, 258.75)
4	(2097.58, 942.75, 622.30, 261.58)	(2097.24, 943.30, 622.47, 261.79)	(2097.46, 942.90, 622.26, 261.58)	(2097.55, 942.78, 622.29, 261.58)
5	(3069.39, 1165.62, 638.87, 154.37	(3069.45, 1165.84, 639.05 154.43)	(3069.29, 1165.80, 638.80, 154.37)	(3069.35, 1165.65, 638.88, 154.37)
6	(2248.16, 927.59, 605.52, 248.31)	(2246.74, 930.51, 605.30, 248.64)	(2247.83, 928.13, 605.50, 248.28)	(2248.07, 927.72, 605.52, 248.31)
7	(2237.60, 931.25, 606.10, 249.15)	(2236.79, 931.45, 606.74, 249.46)	(2237.36, 931.27, 606.22, 249.15)	(2237.54, 931.26, 606.12, 249.15)
8	(2229.45, 931.73, 607.47, 249.80)	(2229.10, 932.34, 607.63, 250.12)	(2229.31, 931.97, 607.44, 249.79)	(2229.41, 931.78, 607.47, 249.80)
9	(2212.52, 933.62, 609.33, 251.27)	(2212.51, 933.93, 609.54, 251.56)	(2212.45, 933.68, 609.32, 251.29)	(2212.50, 933.64, 609.33, 251.26)
10	(3350.42, 1171.60, 605.99, 127.05)	(3350.44, 1172.09, 606.28, 127.10)	(3350.37, 1171.63, 606.02, 127.05)	(3350.40, 1171.60, 606.00, 127.05)
11	(2291.12, 925.82, 600.30, 244.49)	(2290.57, 926.86, 600.67, 244.97)	(2290.93, 926.17, 600.24, 244.49)	(2291.07, 925.88, 600.30, 244.49)
12	(2285.92, 926.99, 600.81, 244.90)	(2285.72, 927.32, 601.22, 245.39)	(2285.80, 926.97, 600.84, 244.92)	(2285.89, 926.99, 600.81, 244.90)
13	(2281.84, 927.31, 601.32, 245.25)	(2281.82, 927.68, 601.66, 245.71)	(2281.77, 927.40, 601.32, 245.27)	(2281.82, 927.33, 601.32, 245.25)
14	(2273.15, 928.22, 602.32, 245.99)	(2273.27, 928.50, 602.63, 246.43)	(2273.12, 928.26, 602.34, 246.01)	(2273.14, 928.23, 602.32, 245.99)
15	(3427.70, 1157.67, 591.47, 122.31)	(3427.84, 1158.17, 591.77, 122.40)	(3427.70, 1157.70, 591.48, 122.31)	(3427.70, 1157.67, 591.47, 122.31)

Table 4: Accuracy comparison between the exact JKM method and the exponential-approximation method using 25, 100 and 400 terms

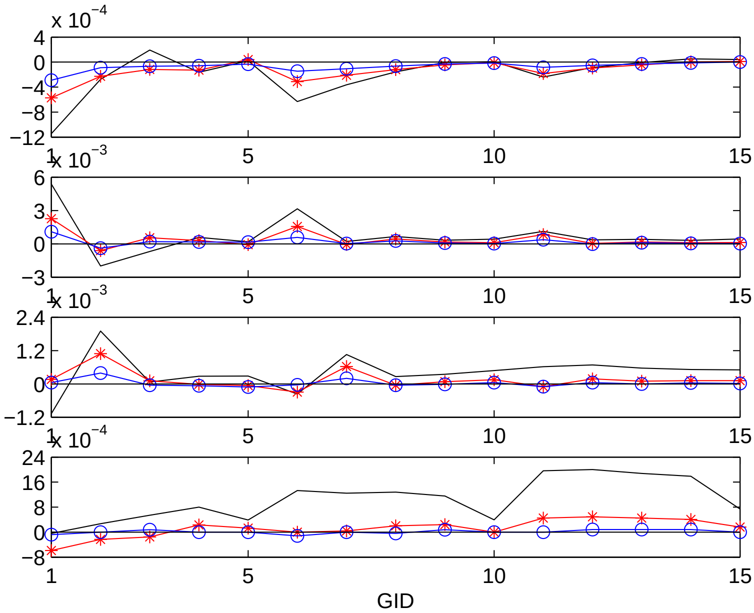

Figure 2: The graphs from top to bottom are the plots of the relative errors of the tranche spreads computed by our new method based on 25-, 50-, and 100-term exponential approximations compared to the exact spreads computed by the JKM method, for the tranches $[0\%, 3\%]$, $[3\%, 4\%]$, $[4\%, 6.1\%]$, and $[6.1\%, 12.1\%]$, respectively. The solid line is for the 25-term approximation. The line marked with small asterisks is for the 50-term approximation. The line marked with small circles is for the 100-term approximation.

for the j-th tranche takes as much time as calculating the spreads for the first j tranches.

4 Conclusions

A new method for pricing correlation-dependent derivatives has been proposed. The method is based on an exponential approximation to the "hockey stick" function. With this approximation, the evaluation of the conditional expectation of the stop-loss function of the credit portfolio can be computed by calculating a series of conditional expectations for individual obligors. In Section 3, we applied this method to synthetic CDO pricing where the correlation structure of the underlying obligors is specified through a simple one-factor Gaussian copula

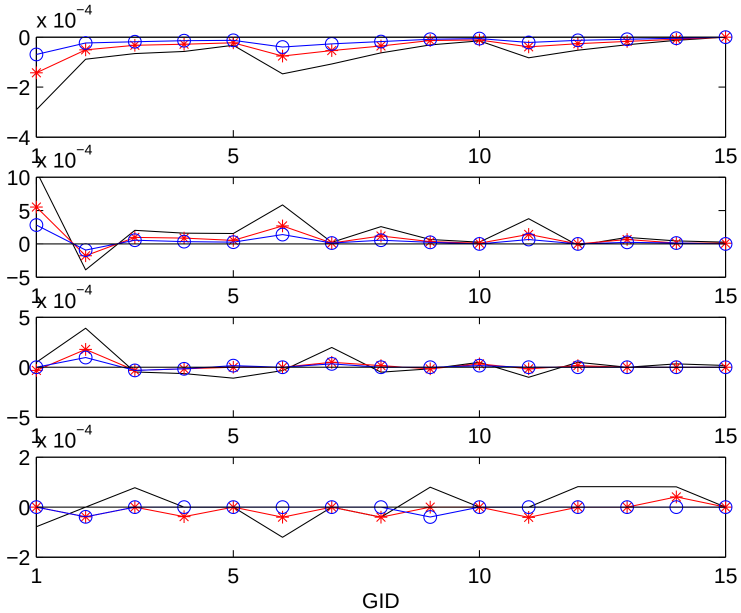

Figure 3: The graphs from top to bottom are the plots of the relative errors of the tranche spreads computed by our new method based on 100-, 200-, and 400-term exponential approximations compared to the exact spreads computed by the JKM method, for the tranches $[0\%, 3\%]$, $[3\%, 4\%]$, $[4\%, 6.1\%]$, and $[6.1\%, 12.1\%]$, respectively. The solid line is for the 100-term approximation. The line marked with small asterisks is for the 200-term approximation. The line marked with small circles is for the 400-term approximation.

model. This method could be applied to more general models provided that they belong to the conditional independence framework, such as the affine Markov chain model proposed by Hurd and Kuznetsov [11]. Also our new method should be applicable to a wide class of derivatives, not just those mentioned above. For example, it can be applied to the pricing of single tranche CDOs, and options on spreads of single tranche CDOs. From formula (7) we see that there are no restrictions on the distribution of Z_k. So this method has a wide range of applications. Compared to the saddlepoint approximation method used by Antonov, Mechkov, and Misirpashaev [2] and Yang, Hurd and Zhang [23], the main advantage of our new approach is that the coefficients can be computed in advance, whereas the saddlepoint method must compute some parameters dynamically.

GID	First tranche	First four tranches
1	0.39	0.46
2	0.44	0.48
3	0.52	0.57
4	0.57	0.70
5	0.81	0.85
6	0.53	0.71
7	0.58	0.76
8	0.67	0.88
9	0.76	1.26
10	1.41	2.32
11	0.86	1.41
12	0.95	1.56
13	1.06	1.86
14	1.32	3.38
15	4.50	12.31

Table 5: CPU time in seconds used by the JKM method to evaluate the first and the first four tranches of the test pools

References

[1] Leif Andersen, Jakob Sidenius, and Susanta Basu. All your hedges in one basket. *Risk*, pages 61–72, November 2003.

[2] Alexandre Antonov, Serguei Mechkov, and Timur Misirpashaev. Analytical techniques for synthetic CDOs and credit default risk measures. Available from http://www.defaultrisk.com/pp_crdrv_77.htm, May 2005.

[3] Robert Eric Beard, Teino Pentikäinen, and Erkki Pesonen. *Risk Theory: The Stochastic Basis of Insurance*. Monographs on Statistics and Applied Probability. Chapman and

	First tranche					First four tranches				
K/N	25	50	100	200	400	25	50	100	200	400
100	0.45	0.63	1.01	1.76	3.36	1.03	1.77	3.83	6.22	12
200	0.57	0.81	1.29	2.34	4.41	1.39	2.40	4.51	8.29	16.52
400	0.74	1.08	1.74	3.11	5.95	1.85	3.19	5.86	11.16	22.76

Table 6: CPU time in seconds used by the exponential-approximation method with different numbers of terms to evaluate the first and the first four tranches of the test pools

Hall, 3 edition, 1984.

[4] Xavier Burtschell, Jon Gregory, and Jean-Paul Laurent. A comparative analysis of CDO pricing models. Available from http://defaultrisk.com/pp_crdrv_71.htm, April 2005.

[5] Nelson De Pril. On the exact computation of the aggregate claims distribution in the individual life model. *ASTIN Bulletin*, 16(2):109–112, 1986.

[6] Ben De Prisco, Ian Iscoe, and Alex Kreinin. Loss in translation. *Risk*, pages 77–82, June 2005.

[7] Alain Debuysscher and Marco Szegö. The Fourier transform method – technical document. Working report, Moody's Investors Service, January 2003. Courtesy of Henry Tabe and Claudio Albanese.

[8] Michael Gordy and David Jones. Random tranches. *Risk*, 16(3):78–83, March 2003.

[9] Christian Hipp. Improved approximations for the aggregate claims distribution in the individual model. *ASTIN Bulletin*, 16(2):89–100, 1986.

[10] John Hull and Alan White. Valuation of a CDO and an n^{th} to default CDS without Monte Carlo simulation. *Journal of Derivatives*, (2):8–23, 2004.

[11] Tom Hurd and Alexey Kuznetsov. Affine Markov chain models of multifirm credit migration. Available from http://www.math.mcmaster.ca/tom/tom.html, 2004.

[12] Ian Iscoe, Ken Jackson, Alex Kreinin, and Xiaofang Ma. On exponential approximaton to the hockey-stick function. In preparation, January 2007.

[13] Ian Iscoe and Alex Kreinin. Valuation of synthetic CDOs. submitted to the Journal of Banking & Finance, 2005.

[14] Ian Iscoe, Alex Kreinin, and Dan Rosen. An integrated market and credit risk portfolio model. *Algo Research Quarterly*, 2(3):21–38, September 1999.

[15] Ken Jackson, Alex Kreinin, and Xiaofang Ma. Loss distribution evaluation for synthetic CDOs. In preparation, January 2007.

[16] Stuart A Klugman, Harry H Panjer, and Gordon E Willmot. *Loss Models From Data to Decisions*. John Wiley & Sons, Inc., 1998.

[17] Jean-Paul Laurent and Jon Gregory. Basket default swaps, CDOs and factor copulas. Available from http://www.maths.univ-evry.fr/mathfi/JPLaurent.pdf, September 2003.

[18] David X Li. On default correlation: A copula function approach. Technical Report 99-07, The RiskMetrics Group, 44 Wall St., New York, NY 10005, April 2000.

[19] Harry H Panjer. Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, 12(1):22–26, 1981.

[20] Susan M Pitts. A functional approach to approximations for the individual risk model. *ASTIN Bulletin*, 34(2):379–397, 2004.

[21] Philipp J Schönbucher. *Credit Derivatives Pricing Models*. Wiley Finance Series. John Wiley & Sons Canada Ltd., 2003.

[22] Oldrich Vasicek. Probability of loss distribution. Technical report, KMV, February 1987. Available from http://www.moodyskmv.com/research/.

[23] Jingping Yang, Tom Hurd, and Xuping Zhang. Saddlepoint approximation method for pricing CDOs. Available from http://www.math.mcmaster.ca/tom/tom.html, November 2005.