Analytical Pricing of CDOs in a Multi-factor Setting by a Moment Matching Approach

Antonio Castagna
Iason ltd.

Fabio Mercurio
Bloomberg LP and Iason ltd.

Paola Mosconi *
Banca IMI

First Version: January 6, 2012
This version: January 16, 2012

1 Introduction

The financial crisis started in the 2007 highlights the need for a more robust approach to accurately pricing and measuring the risk of investment positions held by financial institutions in credit derivatives such as collateralized debt obligations (CDOs). Many researchers devoted a great deal of efforts in analysing credit related products; see, amongst many others, Giesecke and Kim [7], Brigo et al. [2], Eckner [6], Duffie and Garleanu [5], Lopatin and Misirpashaev [10], Mortensen [13], Papageorgiou and Sircar [15].

Pricing of CDOs is typically performed by modelling the correlation between debtors. The approaches followed are of two types: a reduced form approach, relying on the specification of the default as a rare event (basically it is the first jump of a Poisson process), so that default correlation can be attained via some form of dependency of the debtors' default intensities; and a structural approach, based on the Merton's model in [12], assuming a dependency of the debtors' asset dynamics on one ore more common factors, thus obtaining a default correlation that can be seen as the result of a copula function. The latter approach is also the same adopted by the international regulation in the Basel II framework [14], for the purpose to compute the Credit VaR; this specific form of the approach, which underlies the Internal Ratings-Based (IRB) framework of the Pillar I of Basel II, is also known as the Asymptotic Single-Risk Factor (ASRF) paradigm and it hinges on the approximation in Vasicek [18], which consists in replacing the original portfolio loss distribution with an asymptotic one, whose VaR can be computed analytically. Basic hypothesis of this model include the homogeneity of the underlying portfolio and a common factor driving systematic risk. As such, the model does not consider concentration, sector and contagion effects, but it allows for an easy and fast computation of the

*P. M. gratefully acknowledges Giuseppe Pidoto for his prompt technical support at an early stage of the work. This paper expresses the views of its authors and does not represent the institutions where the authors are working or have worked in the past. Email: paola.mosconi@bancaimi.com.

credit portfolio VaR. The same hypothesis are often used to price also CDOs, as firstly shown in the work by Li [4].

In previous works by Bonollo, Mercurio and Mosconi [1] the basic ASRF hypothesis has been relaxed by introducing single name or imperfect granularity risk, due to the small size of the portfolio or to the presence of large exposures associated to single obligors; a sectoral concentration risk, due to imperfect diversification across sectoral factors; and finally a credit contagion risk taking into account the occurrence of default events triggered by inter-dependencies (legal, financial, business-oriented) among obligors. The model has then been generalized by Castagna, Mercurio and Mosconi [3] to a multi-scenario setting, allowing also for stochastic probabilities of default (PD) and recovery rates correlated to the PDs. Both works derive analytical approximations working very efficiently for the computation of the Credit VaR (*i.e.*: cumulated credit losses at high quantiles).

In what follows we try and apply the single-scenario version of the general model in Castagna, Mercurio and Mosconi [3] to the pricing of CDOs. We think it is important for financial institutions using a unified approach to both evaluate the Credit VaR and the risk of structured products they issue, and thus evaluate on a consistent and uniform basis the Economic Capital required to face unexpected credit losses, and the risk transferred out of the balance sheet via the securitisation activity. The framework should also be sophisticated and complex, so as to account for more realistic features of credit portfolios not considered under the simplified ASRF hypothesis. Clearly a more robust model is useful also to properly appreciate the risks embedded, on the asset side of the balance sheet, in the investments the institution have in credit linked products.

The approach we show in what follows is applied to a framework that matches these requirements, and at the same time it avoids to resort to cumbersome numerical procedure usually employed with more complex models, by retaining a closed-form feature that allows a quick and accurate pricing of CDO structures.

2 The Credit Portfolio Model

Recalling the model developed in [1] and [3], we consider a portfolio of loans where loans are associated to M distinct borrowers, each borrower having exactly one loan characterized by exposure EAD_i. We define the weight of a loan in the portfolio as $w_i = EAD_i / \sum_{i=1}^{M} EAD_i$. Each obligor is assigned a probability of default and a loss given default, the total number of them being in principle very large. In practice, however, on the basis of their creditworthiness, obligors are grouped in a certain number of rating classes, each featured by its own PD and LGD, so that all the obligors belonging to the same class share the same PD and LGD. The loss given default is described by means of a stochastic variable Q, whose independence of other sources of randomness is assumed.

In this framework, the portfolio loss L at a given time horizon[1] T can be written as

[1]Following [18], it turns out that the dependence of the portfolio loss on the time horizon T is totally encoded into the default probability (we assume a constant loss given default).

the sum of single obligor's losses:

$$L = \sum_{i=1}^{M} w_i \, L_i = \sum_{i=1}^{M} w_i \, Q_i \, \mathbb{I}_{\{D_i\}} \tag{1}$$

where the last equality has been written by virtue of the LGD independence of the other drivers of risk and the term D_i generically indicates the binary event of default associated to borrow i, to be specified in the following.

We start by describing the properties of the stochastic variable Q. We follow Gordy [8] and assume Q to be distributed according to[2] a Beta(α, β), where the parameters have been chosen in order to satisfy:

$$\alpha \equiv \mu$$
$$\beta \equiv 1 - \mu \, ,$$

with $\mu = \mathbb{E}(Q)$ the mean of the distribution. The standard deviation σ and the skewness γ_1 associated to Q, written in terms of μ, assume the following expressions:

$$\sigma = \sqrt{\frac{1}{2} \mu \, (1 - \mu)}$$
$$\gamma_1 = \frac{2\sqrt{2}}{3} \cdot \frac{1 - 2\mu}{\sqrt{\mu \, (1 - \mu)}} \, .$$

Next, we need to model the default event D_i appearing in (1). We extend the basic ASRF paradigm in order to include several risk factors (*sectors*) and *contagion* effects. Default occurs when a variable X_i describing the well-being of obligor i falls below a certain threshold ζ_i, where X_i is identified with the obligor's asset return and the threshold is related to the obligor's probability of default p_i at time T, i.e. $\zeta_i \equiv N^{-1}(p_i)$. Asset returns are assumed to follow a standard normal distribution:

$$X_i = r_i Y_i + \sqrt{1 - r_i^2} \, \xi_i \, , \tag{3}$$

where

$$Y_i = \sum_{k=1}^{N} \alpha_{ik} \, Z_k \qquad Z_k \sim \mathcal{N}(0, 1) \; i.i.d.$$

represents a systematic component synthetically describing the effects of several industry-geographic sectors (with factor loadings r_i) and

$$\xi_i \equiv \xi(\Gamma_i, \epsilon_i) = g_i \Gamma_i + \sqrt{1 - g_i^2} \, \epsilon_i$$

[2]General order moments are given by the recursive formula:

$$\mathbb{E}(Q^k) = \frac{\alpha + k - 1}{\alpha + \beta + k - 1} \, \mathbb{E}(Q^{k-1}) \, . \tag{2}$$

is an obligor specific term, which includes a purely idiosyncratic contribution proportional to $\epsilon_i \sim \mathcal{N}(0,1)$, *i.i.d.*, and a contagion[3] term $g_i\Gamma_i$. The unconditional correlation between pairs of distinct obligors is given by:

$$\rho_{ij} = r_i r_j \sum_{k=1}^{N} \alpha_{ik}\alpha_{jk} + \sqrt{1-r_i^2}\sqrt{1-r_j^2}\, g_i g_j \sum_{k=1}^{N} \gamma_{ik}\gamma_{jk} \,. \tag{4}$$

Summarizing, the portfolio loss (1) at time T can be recast as:

$$L = \sum_{i=1}^{M} w_i\, Q_i\, \mathbb{I}_{X_i \le N^{-1}(p_i)} \,, \tag{5}$$

where p_i are the probabilities of default at time T. As already mentioned in the Introduction, such expression is analytically tractable and allows to extract important information about the portfolio of loans in two main directions:

1. for high confidence level q, a second order Taylor expansion provides a very accurate estimate of the quantile $t_q(L)$. This quantity is fundamental in the calculation of the value at risk[4], entering the economic capital;

2. for lower confidence levels, such as those required in the structuring of CDO tranches[5], Taylor expansion breaks down, but a procedure based on the calculation of the exact moments of L and a *moment matching* technique allows to arrive at the tranches' price. This will be the topic at the hearth of this paper.

In Section 2.1 we simply quote the results about the high quantile expansion, referring to [1] and [3] for a detailed derivation and analysis. In Section 2.2 we provide a brief description of the *moment matching* (MM) method and analytical expressions of raw moments of L, up to order $k = 4$.

2.1 High Quantile Expansion: VaR

Let the variable $\overline{L}$ be defined as the limiting loss distribution in the one-factor Merton framework

$$\overline{L} = l(\overline{Y}) = \sum_{i=1}^{M} w_i\, \mu_i\, \hat{p}_i(\overline{Y}) \,, \tag{6}$$

---

[3] Contagion is modeled as follows: the whole portfolio is split into two clusters: the "I" (infecting) class of obligors and the "C" (infected) one, contagion factor loadings g_i being different from zero only for obligors belonging to the "C" class. The decomposition of the contagion effect onto different infected sectors C_k is summarized by the relation:

$$\Gamma_i = \sum_{k=1}^{N} \gamma_{ik}\, C_k \qquad C_k \sim \mathcal{N}(0,1)\ i.i.d.$$

[4] The regulatory value is $q = 99.9\%$.

[5] Typical values ranges from $q = 3\%$ to $q = 22\%$

where $\hat{p}_i(y)$ is the probability of default of borrower i, conditional on $\overline{Y} = y$

$$\hat{p}_i(y) = N\left[\frac{N^{-1}(p_i) - a_i y}{\sqrt{1 - a_i^2}}\right]. \tag{7}$$

(N indicates the cumulative normal distribution). The quantile of $\overline{L}$ at level q can be calculated analytically as

$$t_q(\overline{L}) = l(N^{-1}(1-q)).$$

Carrying out the Taylor expansion of $t_q(L)$ around $t_q(\overline{L})$, first-order contributions cancel out and the final result, up to the second order, assumes the form [16]

$$\Delta t_q \equiv t_q(L) - t_q(\overline{L}) = -\frac{1}{2\,l'(y)}\left[\nu'(y) - \nu(y)\left(\frac{l''(y)}{l'(y)} + y\right)\right]\Bigg|_{y=N^{-1}(1-q)} \tag{8}$$

The function $l(y)$ is defined as in (6), while $\nu(y) = \mathrm{var}[L|\overline{Y} = y]$ is the conditional variance of L on $\overline{Y} = y$. This function can be further decomposed in terms of its systematic and idiosyncratic components

$$\nu(y) = \nu_\infty(y) + \nu_{GA}(y), \tag{9}$$

where

$$\begin{aligned} \nu_\infty(y) &= \mathrm{var}[E(L|\{Z_k\})|\overline{Y} = y] = \qquad (10)\\ &= \sum_{i=1}^{M}\sum_{j=1}^{M} w_i w_j \mu_i \mu_j \left[N_2(N^{-1}[\hat{p}_i(y)], N^{-1}[\hat{p}_j(y)], \rho_{ij}^Y) - \hat{p}_i(y)\hat{p}_j(y)\right], \end{aligned}$$

$$\begin{aligned} \nu_{GA}(y) &= E[\mathrm{var}(L|\{Z_k\})|\overline{Y} = y] = \qquad (11)\\ &= \sum_{i=1}^{M} w_i^2 \left(\mu_i^2\left[\hat{p}_i(y) - N_2(N^{-1}[\hat{p}_i(y)], N^{-1}[\hat{p}_i(y)], \rho_{ii}^Y)\right] + \sigma_i^2 \hat{p}_i(y)\right). \end{aligned}$$

($N_2(.,.,.)$ is the bivariate normal cumulative distribution function). The first term, $\nu_\infty(y)$, accounts for the correction to the loss distribution due to the multi-factor setting, in the limit of an infinitely fine-grained portfolio. $\nu_{GA}(y)$ is the granularity adjustment term. The quantity ρ_{ij}^Y is the conditional correlation between pairs of distinct debtors. In the special case of homogeneous LGDs and default probabilities p_i, it becomes proportional to the Herfindahl-Hirschman index $HHI = \sum_{i=1}^{M} w_i^2$ (see [8]).

2.2 Moment Matching Method

The *Moment Matching* method (MM in the following) is based on the following idea: assume we have a random variable X whose terminal distributions is unknown but for which, under some conditions, we can compute the moments M_1, M_2,...,M_n. Then we can choose a known random variable Y, with the same n first moments, and replace X by Y to run the required computations. The Y proxy should not be an arbitrary choice, but be strictly related to the initial problem.

One of the most common example is the approximation of the sum of Lognormal variables, denoted by X and which is not lognormally distributed itself, by means of a Lognormal proxy Y. The MM approach has several relevant applications also in financial mathematics. For instance, it is very popular for Asian options, whose payoff is a function of the average price of some underlying S_t during the contract's life. In the Black-Scholes model the price is lognormally distributed at any time in the future, so one can argue that the sum of n prices, sampled at given observation dates, is approximately close to a Lognormal distribution. Turnbull and Wakemann [17] and Levy [9] are seminal research papers in this field.

In the specific case under consideration, the portfolio loss distribution L, given in eq. (5), is unknown but its moments of arbitrary order can be analytically computed. We therefore choose an approximating distribution L^*, whose distributional properties are known, and apply the MM technique. Depending on the number of parameters n which define L^*, we calibrate the first n exact raw moments M_n of L on the corresponding moments M_n^* of L^* and derive the values of the n parameters, by solving a system of n non-linear equations of the form

$$F(x) = 0$$

where $x = (x_1, \ldots, x_n)$ is the vector containing the n parameters of L^* and

$$F_i(x) = \frac{M_i - M_i^*(x)}{M_i}\,, \qquad i = 1, \ldots, n\,.$$

In Section 4, we will show two examples of approximating distributions L^* depending respectively on $n = 2$ and $n = 4$ parameters. Therefore, we complete this Section by quoting the analytical formulas of the first four raw moments of the exact distribution L.

First moment M_1

The first moment of the loss distribution (5) is simply given by the expected loss:

$$M_1 \equiv \mathbb{E}(L) = \sum_{i=1}^{M} w_i\,\mu_i\,p_i\,. \tag{12}$$

Second moment M_2

The second raw moment is calculated by starting from its very definition

$$M_2 \equiv \mathbb{E}(L^2) = \sum_{i,j=1}^{M} \mathbb{E}(L_i\,L_j)\,.$$

With some algebra we obtain

$$\begin{aligned} M_2 &= \sum_{i=1}^{M} w_i^2\,\mathbb{E}(Q_i^2)\,p_i + \sum_{i\neq j}^{M} w_i w_j \mu_i \mu_j\,N_2(N^{-1}[p_i], N^{-1}[p_j], \rho_{ij}) = \\ &= \sum_{i=1}^{M} w_i^2\left[\mathbb{E}(Q_i^2)\,p_i - \mu_i^2\,N_2(N^{-1}[p_i], N^{-1}[p_i], \rho_{ii})\right] + \\ &+ \sum_{i=1}^{M}\sum_{j=1}^{M} w_i w_j \mu_i \mu_j\left[N_2(N^{-1}[p_i], N^{-1}[p_j], \rho_{ij})\right]\,, \end{aligned} \tag{13}$$

where

$$\mathbb{E}(Q_i^2) = \mu_i^2 + \sigma_i^2\,.$$

The first contribution in (13) is the granularity adjustment which can be diversified away in the limit $M \to \infty$, while the second represents the systematic term, which cannot be made to vanish. ρ_{ij} defines the unconditional correlation between pairs of distinct obligors, eq. (4), with the understanding that ρ_{ii} is simply obtained from eq. (4), by replacing index j with index i.

Third moment

The third moment is defined by the following relation:

$$M_3 \equiv \mathbb{E}(L^3) = \sum_{i,j,k=1}^{M} \mathbb{E}(L_i\, L_j\, L_k)\,.$$

In explicit terms we get:

$$\begin{aligned} M_3 = & \sum_{i=1}^{M} w_i^3\, \mathbb{E}(Q_i^3)\, p_i + 3 \sum_{i \neq j}^{M} w_i^2 w_j\ \mathbb{E}(q_i^2)\, \mu_j\, N_2(N^{-1}[p_i], N^{-1}[p_j], \rho_{ij}) \\ & + \sum_{i \neq j \neq k}^{M} w_i w_j w_k\ \mu_i \mu_j \mu_k\, N_3(N^{-1}[p_i], N^{-1}[p_j], N^{-1}[p_k], \Sigma_{3\times 3}) \end{aligned} \tag{14}$$

where

$$\Sigma_{3\times 3} = \begin{pmatrix} 1 & \rho_{ij} & \rho_{ik} \\ & 1 & \rho_{jk} \\ & & 1 \end{pmatrix}$$

is the symmetric variance-covariance matrix and $\mathbb{E}(Q^3)$ can be expressed in terms of the mean value μ, the standard deviation σ and the skewness γ_1, of the generic stochastic variable Q associated to the LGD, as follows:

$$\mathbb{E}(Q^3) = \sigma^3 \gamma_1 + \mu\,(3\,\sigma^2 + \mu^2)\,.$$

Fourth moment

The fourth moment is given by:

$$M_4 \equiv \mathbb{E}(L^4) = \sum_{i,j,k,m=1}^{M} \mathbb{E}(L_i\, L_j\, L_k\, L_m)\,. \tag{15}$$

Introducing the compact notation

$$N_4(i,j,k,m) \equiv N_4(N^{-1}[p_i], N^{-1}[p_j], N^{-1}[p_k], N^{-1}[p_m], \Sigma_{4\times 4})\,,$$

where $\Sigma_{4\times 4}$ denotes the symmetric matrix encoding the correlation structure inside groups of four obligors

$$\Sigma_{4\times 4} = \begin{pmatrix} 1 & \rho_{ij} & \rho_{ik} & \rho_{im} \\ & 1 & \rho_{jk} & \rho_{jm} \\ & & 1 & \rho_{km} \\ & & & 1 \end{pmatrix},$$

eq. (15) assumes the explicit form:

$$\begin{aligned} M_4 &= \sum_{i=1}^{M} w_i^4 \, \mathbb{E}(Q_i^4) \, p_i + \\ &+ 2 \sum_{i \neq j}^{M} \left[2 \, w_i^3 w_j \, \mathbb{E}(Q_i^3) \mu_j + 3 \, w_i^2 w_j^2 \, \mathbb{E}(Q_i^2) \mathbb{E}(Q_j^2) \right] N_2(i,j) + \\ &+ 12 \sum_{i \neq j \neq k}^{M} w_i^2 w_j w_k \, \mathbb{E}(Q_i^2) \mu_j \mu_k \, N_3(i,j,k) + \\ &+ \sum_{i \neq j \neq k \neq m}^{M} w_i w_j w_k w_m \, \mu_i \mu_j \mu_k \mu_m \, N_4(i,j,k,m) \,. \end{aligned} \tag{16}$$

Given the recursive relation (2), the fourth raw moment of the LGD stochastic variable is equal to:

$$\mathbb{E}(Q^4) = \frac{\mu + 3}{4} \, \mathbb{E}(Q^3) \,.$$

3 Pricing CDO Tranches

Synthetic CDOs with maturity T_b are contracts obtained by putting together a collection of Credit Default Swaps (CDS) with the same maturity on different names and then tranching the loss associated to this pool, at two detachment points A and B such that $0 \leq A < B \leq 1$.

Each tranche is associated a so-called spread Spr_{AB} which represents the periodic premium paid by the protection buyer to the protection seller, in exchange for payments in case of loss. Given a set of maturities $T_i = T_1, \ldots, T_b$, the spread is given by the following formula (see *e.g.* Brigo et al. [2]):

$$\mathrm{Spr}_{AB} = \frac{\mathbb{E}\left[\sum_{i=1}^{b} D(0,T_i) \left(L_i^{AB} - L_{i-1}^{AB} \right) \right]}{\mathbb{E}\left[\sum_{i=1}^{b} D(0,T_i)(T_i - T_{i-1}) \left(1 - \frac{L_i^{AB} + L_{i-1}^{AB}}{2} \right) \right]}$$

which, assuming independence of interest rates from the other sources of risk, becomes:

$$\mathrm{Spr}_{AB} = \frac{\sum_{i=1}^{b} P(0,T_i) \left(\mathbb{E}[L_i^{AB}] - \mathbb{E}[L_{i-1}^{AB}] \right)}{\sum_{i=1}^{b} P(0,T_i)(T_i - T_{i-1}) \left(1 - \frac{\mathbb{E}[L_i^{AB}] + \mathbb{E}[L_{i-1}^{AB}]}{2} \right)} \,. \tag{17}$$

The expectations $\mathbb{E}[L_i^{AB}]$ are called Expected Tranche Losses (ETLs) and represents a sort of building blocks, useful in the calculation of the spread. In terms of the detachments points A and B they are given explicitly by:

$$\mathbb{E}[L_i^{AB}] = \frac{\mathbb{E}[(L_i - A)^+] - \mathbb{E}[(L_i - B)^+]}{B - A} \,. \tag{18}$$

For the equity tranche ($A = 0$) it is common practice to quote the upfront amount U_{0B} needed to make the contract fair when a running spread of 500bps is taken as a periodic spread in the premium leg.

Choosing a model to describe the portfolio loss L allows to determine the values of the ETLs and therefore the spread for each tranche. In the following we will specify two models and apply the pricing formulae to compute spreads for different tranches.

In the past, these formulae have been used in different ways and applied to the market of quoted instruments. For example, starting from market quotes, without specifying a portfolio model, Brigo et al. [2] have showed how to extract a term structure of ETLs in a model independent framework. Among the applications relying on the choice of a particular model, a special role has been played by the so called Gaussian Copula Large Homogeneous Portfolio approximation (GCLHP), which is based on the aforementioned ASRF approach, introduce by Vasicek [18] and further developed by JP Morgan [11]. The model is a simplified, single factor, no-contagion, version of that described in Section 2, where homogeneity of the portfolio implies in eq. (3) constant values of factor loadings $r_i = r$, probabilities of default $p_i = p$ and loss given default $\mu_i = \mu$. Furthermore, the large portfolio assumption allows to derive an analytical expression for the ETLs given by:

$$\mathbb{E}[L^{AB}] = \frac{N_2(N^{-1}(p), -N^{-1}(A), -\sqrt{1-r^2}) - N_2(N^{-1}(p), -N^{-1}(B), -\sqrt{1-r^2})}{B - A} \tag{19}$$

for unitary loss given default $\mu = 1$ and by:

$$\mathbb{E}[L_{LGD}^{A,B}] = \mathbb{E}[L^{\frac{A}{\mu}, \frac{B}{\mu}}] \tag{20}$$

for constant, but different from one, loss given default $\mu \neq 1$.

The GCLHP approach, due to its semi-analytical tractability, has been used in the past both as a standard tool to extract implied default correlations from market quotes and as an interpolator in order to quote off-market spreads. Implied correlations plays a role similar to that of the implied volatilities in the Black and Scholes framework. Two types of implied correlations have been identified: the so called *compound correlation* is obtained by fixing a tranche and inverting the pricing formulae (17) and (20), such that the correlation parameter r yield the market spread. Inversion, though working properly for the equity tranche 0-B, is problematic for mezzanine tranches A-B (because the spread is not monotonic in the correlation parameter).

In order to overcome the difficulties with *compound correlation*, another type has been introduced by JP Morgan [11], the so called *base correlation*. Such parameter, when inferred from market quotes, is extracted from the spreads associated to 0-A and the fictive tranches 0-B_1, 0-B_2...Therefore, it requires a bootstrapping procedure. In our case, spreads of the 0-B_1, 0-B_2... tranches can be easily obtained within our model. Therefore, we just need to invert the standard model pricing formulae (17) and (20). It turns out that *base correlation* is more like implied volatilities of equity options and, being monotonic in spread, is always invertible.

4 Approximating Distribution

In this Section we present the two and four parameter distributions that we have chosen in order to approximate the portfolio loss L.

The choice of the two parameter distribution has fallen on the Vasicek one, denoted by Vasicek(p, r). Historically, this distribution has played a crucial role in approximating the loss of a credit portfolio and has served as the basis for the development of the Gaussian Copula LHP model, which has been used as a standard model to price CDO tranches and to derive implied correlations from spreads quoted in the market.

Here we use it in a different way. We start from a portfolio of loans characterized by many dependencies (multi-factor environment, contagion, inhomogeneities in rating properties and exposure) and map it onto a Vasicek(p, r) such that the parameters p and r encode all of these dependencies. We use the resulting distribution to calculate CDO spreads for internal purposes.

Further, we use the Vasicek(p, r) as a building block to construct the four parameter distribution. We choose indeed a mixture of two Vasicek distributions with different values of the parameters.

In the following two Sections we summarize some stylized facts about the chosen distributions.

4.1 Two Parameter Distribution

We consider the approximating portfolio loss L^* to be distributed according to:

$$L^* \sim \text{Vasicek}(p, r)\,,$$

such that its cumulative distribution function (CDF) is given by:

$$F_{p,r}(x) = P(L^* \leq x) = N\left(\frac{\sqrt{1-r}\,N^{-1}(x) - t}{\sqrt{r}}\right)\,, \tag{21}$$

with $t \equiv N^{-1}(p)$. In order to calculate the moments of L^*, it turns out to be convenient to express the stochastic variable L^* as follows:

$$L^* = N\left(\frac{t - \sqrt{r}\,S}{\sqrt{1-r}}\right)\,, \tag{22}$$

where $S \sim \mathcal{N}(0, 1)$ is a standard normal variable. Moments of order n are therefore analytically computed by means of:

$$\begin{aligned} M_n^* \equiv \mathbb{E}[(L^*)^n] &= \mathbb{E}\left[N\left(\frac{t - \sqrt{r}\,S}{\sqrt{1-r}}\right)^n\right] = \\ &= \mathbb{P}[Y_1 \leq t, \ldots, Y_n \leq t] = \\ &= N_n(t, \ldots, t, \Sigma_{n \times n}) \end{aligned} \tag{23}$$

where $(Y_1, \ldots, Y_n)$ is a multivariate normal vector with $E[Y_i] = 0$, $\text{var}(Y_i) = 1$ and $\text{corr}(Y_i, Y_j) = r$, $i \neq j$. We conclude by quoting the analytical formula used to calculate the expected tranche loss:

$$\mathbb{E}[(L^* - x_0)^+] = N_2[N^{-1}(p), -N^{-1}(x_0), -\sqrt{1-r}]\,. \tag{24}$$

4.2 Four Parameter Distribution

We choose to approximate the unknown portfolio loss L with a variable L^* whose distribution depends on four parameters and whose CDF is a mixture of CDFs of two Vasicek distributions, given respectively by Vasicek(p_1, r) and Vasicek(p_2, r):

$$F_{a,p_1,r,p_2}(x) = a\, N\left[\frac{\sqrt{1-r}\, N^{-1}(x) - t_1}{\sqrt{r}}\right] + (1-a)\, N\left[\frac{\sqrt{1-r}\, N^{-1}(x) - t_2}{\sqrt{r}}\right] \qquad (25)$$

where $0 \leq a, r, p_1, p_2 \leq 1$.

The probability density function (PDF) is calculated by differentiating eq. (25) with respect to x, i.e.

$$f_{a,p_1,r,p_2}(x) = \frac{\partial}{\partial x} F_{a,p_1,r,p_2}(x) = a\, f_{p_1,r}(x) + (1-a)\, f_{p_2,r}(x)\,. \qquad (26)$$

and moments of order n are easily calculated by exploiting eq.s (23) and (26):

$$\begin{aligned} M_n^* &\equiv \mathbb{E}[(L^*)^n] = \int_{\mathbb{R}} x^n\, f_{a,p_1,r,p_2}(x)\, dx = \\ &= a \int_{\mathbb{R}} x^n\, f_{p_1,r}(x)\, dx + (1-a) \int_{\mathbb{R}} x^n\, f_{p_2,r}(x)\, dx = \\ &= a\, \mathbb{E}_1[(L^*)^n] + (1-a)\, \mathbb{E}_2[(L^*)^n] = \\ &= a\, \mathbb{E}_1\left[N\left(\frac{t_1 - \sqrt{r}\, S}{\sqrt{1-r}}\right)^n\right] + (1-a)\, \mathbb{E}_2\left[N\left(\frac{t_2 - \sqrt{r}\, S}{\sqrt{1-r}}\right)^n\right] = \\ &= a N_n(t_1, \ldots, t_1, \Sigma_{n\times n}) + (1-a) N_n(t_2, \ldots, t_2, \Sigma_{n\times n})\,, \end{aligned} \qquad (27)$$

where $\mathbb{E}_{1,2}$ denote expectations calculated respectively with densities $f_{p_1,r}$ and $f_{p_2,r}$ and

$$t_1 \equiv N^{-1}(p_1) \qquad t_2 \equiv N^{-1}(p_2)\,.$$

Further, the equivalent of eq. (24) for the mixture of two Vasicek distributions reads:

$$\begin{aligned} \mathbb{E}[(L^* - x_0)^+] = a\; & N_2[N^{-1}(p_1), -N^{-1}(x_0), -\sqrt{1-r}] \\ & + (1-a)\; N_2[N^{-1}(p_2), -N^{-1}(x_0), -\sqrt{1-r}]\,. \end{aligned} \qquad (28)$$

5 Implementation and Numerical Examples

In this Section we present an implementation of the portfolio model introduced in Section 2. Since a thorough analysis of the properties of high level quantiles and value at risk has been at the hearth of previous works by Bonollo, Mercurio and Mosconi [1] for the basic ASRF model and by Castagna, Mercurio and Mosconi [3] for a multi-scenarios setting of default probabilities and recovery rates, here we focus on applications to CDO pricing. In the following we present the results obtained, in terms of tranche spreads and implied correlations, for the two and four parameter distributions just introduced.

Before delving into the details of the analysis, we describe the main features of the portfolio under consideration. The portfolio consists of $M = 200$ names, $N = 4$ industry-geographic sectors and four rating classes, yielding a good rating quality $\mathbb{E}(L) \approx 1.51\%$ (at one year). The average one year probability of default is $p_{ave} \approx 3.71\%$ and it is assumed to increase at a constant rate of 2% per year. Loss given default is considered constant through time and its average value is fixed at $\mu_{ave} \approx 40\%$, yielding a constant recovery rate of $RR_{ave} \approx 60\%$. The portfolio is not homogeneous in its exposures and contagion effects are present (the number of "I" (infecting) firms being 50); however we have decided to avoid any particular name or sector concentration.

We are interested in pricing tranches covering losses between 0% and 3% (equity tranche) and losses covering 3%-6%, 6%-9%, 9%-12% (mezzanine tranches) and 12%-22%, 22%-100% (senior tranches). We consider maturities $T = 3, 5, 7, 10$ years. Discount factors which will be used in the pricing formulae are given by a standard discounting curve, summarized in Table 1.

t	$P(0, t)$
20/12/2011	1.000000
18/12/2012	0.983910
19/12/2013	0.973614
22/12/2014	0.957627
22/12/2015	0.937049
22/12/2016	0.912336
22/12/2017	0.885320
22/12/2018	0.857344
22/12/2019	0.829756
22/12/2020	0.802647
22/12/2021	0.775905

Table 1: Discount factors.

Exact moments M_n of the original portfolio loss distribution L are collected in tables appearing in the following sections.

5.1 Calibration Results

In this Section we collect the results of the calibration procedure performed in the two aforementioned cases, for different values of the maturities $T = 3, 5, 7, 10$ years.

For the two parameter distribution $L^* \sim \text{Vasicek}(p, r)$, the matching is performed on the the first two raw moments, yielding the results collected in Table 2.

	p	r
3y	0.0157	0.1144
5y	0.0164	0.1116
7y	0.0170	0.1100
10y	0.0181	0.1075

Table 2: Vasicek calibration.

The four parameter distribution corresponding to a mixture of two Vasiceks has been calibrated by using raw moments up to order four. The outcome of the calibration is showed in Table 3.

	a	p_1	p_2	r
3y	0.1764	0.0448	0.0095	0.0232
5y	0.1756	0.0459	0.0101	0.0224
7y	0.1615	0.0493	0.0108	0.0202
10y	0.1603	0.0521	0.0116	0.0176

Table 3: Mixture of Vasicek calibration.

It is interesting to compare the results obtained in the two cases. For example, we collect them in Table 4 for $T = 3$ year case. We show the values of the calibrated moments M_k^* along with the corresponding relative and absolute errors for the two choices of the approximating distribution L^*. It appears that in general the two distributions are practically equivalent as far as lower moments are considered, but a relative error ranging from 30%-40% is always present in the two parameter Vasicek case[6].

	Exact	Mixture			Vasicek		
k	M_k	M_k^*	*abs err*	*% err*	M_k^*	*abs err*	*% err*
1	0.01570	0.01573	2.69E-05	0.17%	0.01570	0	0.00%
2	0.00048	0.00048	3.19E-06	0.67%	0.00048	3.340E-08	0.01%
3	2.25E-05	2.25E-05	3.06E-08	0.14%	2.36E-05	1.101E-06	4.89%
4	1.24E-06	1.29E-06	4.94E-08	3.97%	1.72E-06	4.764E-07	38.32%

Table 4: Calibration results for $T = 3$ years.

5.2 CDO Tranches Pricing

In this Section we present the results pertaining to the calculation of the spreads associated to the given set of tranches 0%-3%, 3%-6%, 6%-9%, 9%-12%, 12%-22% and 22%-100%. We conclude our analysis with a discussion about the structure of implicit correlations implied by the model spreads.

First we focus on the Vasicek case. Table 5 collects the values of the expected tranche losses associated to the given set of tranches, across different maturities.

	ETL Vasicek					
	0-3%	**3-6%**	**6-9%**	**9-12%**	**12-22%**	**22-100%**
3y	0.4521	0.0580	0.0103	0.0022	2.16E-04	3.9E-07
5y	0.4697	0.0627	0.0111	0.0024	2.28E-04	3.9E-07
7y	0.4839	0.0673	0.0121	0.0026	2.47E-04	4.15E-07
10y	0.5092	0.0762	0.0140	0.0030	2.88E-04	4.74E-07

Table 5: Term structure of ETLs for the Vasicek distribution.

[6]Similar tables for $T = 5, 7, 10$ years are presented in the Appendix. The trend highlighted for the maturity $T = 3$ years is common to all of them.

This term structure of ETL, through the pricing formula (17) leads to the following values of the spreads:

Spread Vasicek					
0-3%	**3-6%**	**6-9%**	**9-12%**	**12-22%**	**22-100%**
5.00%	0.87%	0.15%	0.03%	0.00%	0.00%

Table 6: Spreads obtained for the Vasicek distribution model.

Following the convention of fixing the spread of the equity tranche at the value of 500bps, the upfront amounts to $U_{0\%3\%} = 0.217$.

An analogous analysis, carried out for the mixture of Vasicek distributions, leads to a very similar pattern.

	ETL Mixture					
	0-3%	**3-6%**	**6-9%**	**9-12%**	**12-22%**	**22-100%**
3y	0.4331	0.0821	0.0086	0.0004	3.06E-06	6.98E-13
5y	0.4498	0.0865	0.0095	0.0004	3.21E-06	5.65E-13
7y	0.4618	0.0930	0.0120	0.0005	3.61E-06	2.93E-13
10y	0.4843	0.1036	0.0146	0.0005	2.95E-06	5.66E-14

Table 7: Term structure of ETLs for the mixture of Vasicek distributions.

The ETL term structure is showed in Table 7 and leads to the following values of the spreads:

Spread Mixture					
0-3%	**3-6%**	**6-9%**	**9-12%**	**12-22%**	**22-100%**
5.00%	1.19%	0.15%	0.01%	0.00%	0.00%

Table 8: Spreads obtained for the mixture of Vasicek distributions model.

The upfront amount in this case is equal to $U_{0\%3\%} = 0.1862$.

We proceed further in the analysis by extracting the implied correlation. Being aware of the problems arising when trying to invert the pricing formulas in order to derive the *compound correlation* (see *e.g.* Brigo et al. [2]), we start from the *base* one. In order to infer it from the model spreads we need to invert the standard pricing formula of the Gaussian Copula LHP, by inputting the values of the spreads obtained in the Vasicek and mixture case, the desired tranches detachment points and the relevant values of the portfolio average probability of default (p_{ave} correctly computed at different time horizons), and recovery rate. In the *base correlation* case, the relevant tranches correspond to the following detachment points 0%-3%, 0%-6%, 0%-9%, 0%-12%, 0%-22% and 0%-100%. On quoted spreads, such tranches (except from the first one) are fictive and correlation must be retrieved through a bootstrap procedure (see for the details [11]). In our case, the derivation is simpler in that our models promptly give us the values of the (non-quoted) spreads. Moreover, on these fictive tranches spread is monotonic in correlation and the inversion leads to the outcome summarized in Table 9 and plotted in Figure 1. The correlation of the Gaussian Copula LHP model, which is constant, has been determined by inverting the fictive model tranche 0%-12%, where both the Vasicek model and the mixture of two Vasiceks share the same value of the spread.

	Base Correlation					
	0-3%	**0-6%**	**0-9%**	**0-12%**	**0-22%**	**0-100%**
Mixture	40.190%	10.410%	13.690%	18.540%	22.360%	22.360%
Vasicek	35.180%	10.190%	13.690%	18.540%	22.360%	22.360%
GC	18.540%	18.540%	18.540%	18.540%	18.540%	18.540%

Table 9: Base correlation.

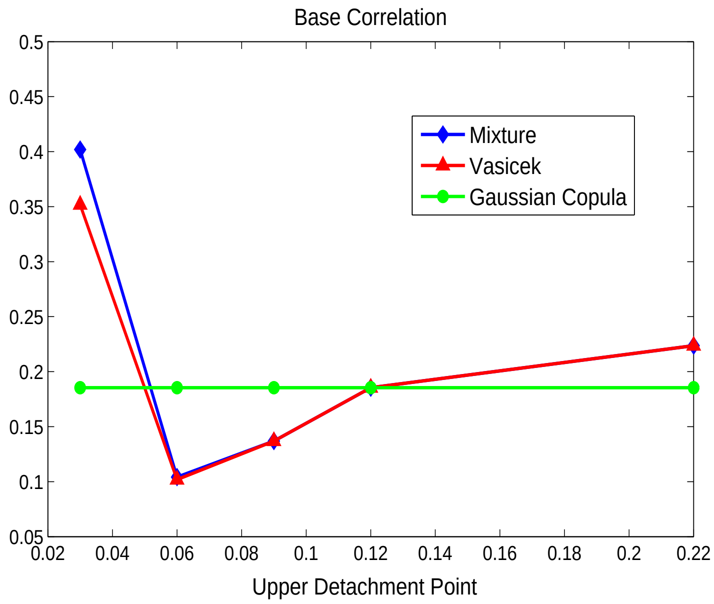

Figure 1: Base correlation.

Figure 1 is particularly interesting because it shows a *smile* behavior in the implied (*base*) correlation similar to what observed in the equity option market of implied volatilities and consistent with the existing literature (see *e.g.* [2], [13] and [11]). Also, in this case it is possible to infer that the smile behavior is due to the fatness of the tails in true portfolio loss distribution, which is not captured by a Gaussian Copula LHP approach, but that is implicit, by construction, in the two models we have proposed. We notice that the two approximating distributions we have proposed do not substantially differ. Indeed, the largest difference affects the fourth (and, only partially the third) moment of the distribution and, as far as spreads and correlations are concerned, this does not have a relevant impact.

As an illustrative example, we conclude by showing the results obtained in the case of the *compound correlation*. The outcome is summarized in Table 10.
Figure 2 shows a plot of the compound correlation. This pattern is consistent with patterns already encountered in literature (see *e.g.* Brigo et al. [2]) and, as already mentioned, is a reflection of the difficulty of extracting correlations on mezzanine tranches,

	Compound Correlation					
	0-3%	**3-6%**	**6-9%**	**9-12%**	**12-22%**	**22-100%**
Mixture	40.190%	52.730%	40.810%	33.870%	30.650%	22.360%
Vasicek	35.180%	40.630%	40.590%	41.170%	42.150%	22.360%
GC	18.540%	18.540%	18.540%	18.540%	18.540%	18.540%

Table 10: Compound correlation.

where spread is not a monotonic function of correlation. This is also related to the existence of multiple solutions to the inversion problem.

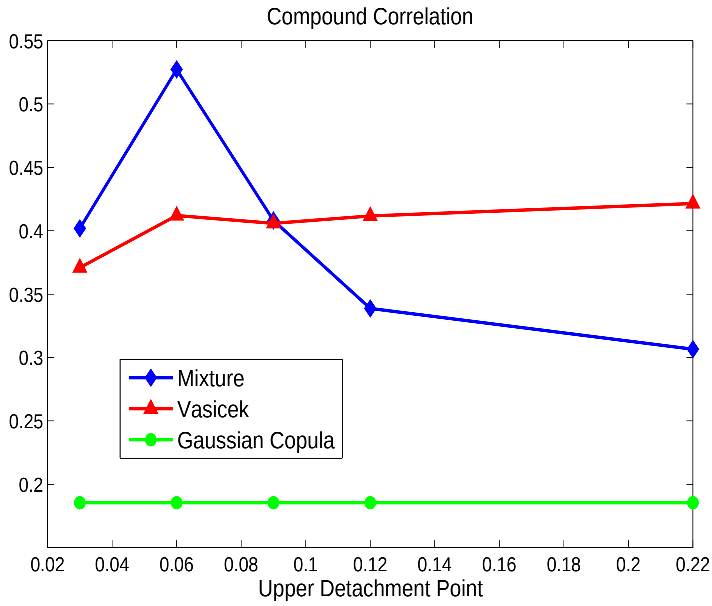

Figure 2: Compound correlation.

6 Conclusions

We have considered the credit portfolio model presented in [1] and [3], which describes an extension of the ASRF paradigm including several kinds of dependencies and inhomogeneities, and analyzed how to price CDOs backed by its loans. By means of a *moment matching* technique, we have mapped the unknown portfolio loss distribution onto two examples of known distributions and provided analytical formulas to determine the CDO tranches' spreads. The results obtained are consistent with the existing literature and, in terms of the implied correlation, a smile behavior is found, which can be explained by the fat-tail nature of the true loss distribution.

k	Exact M_k	Mixture M_k^*	*abs err*	*% err*	Vasicek M_k^*	*abs err*	*% err*
1	0.01638	0.01639	1.89E-06	0.01%	0.01640	1.541E-05	0.09%
2	0.00051	0.00051	1.27E-06	0.25%	0.00051	7.809E-07	0.15%
3	2.44E-05	2.39E-05	4.89E-07	2.00%	2.54E-05	9.845E-07	4.03%
4	1.37E-06	1.39E-06	1.77E-08	1.29%	1.84E-06	4.737E-07	34.62%

Table 11: Calibration results for $T = 5$ years.

k	Exact M_k	Mixture M_k^*	*abs err*	*% err*	Vasicek M_k^*	*abs err*	*% err*
1	0.01705	0.01702	2.88E-05	0.17%	0.01700	4.653E-05	0.27%
2	0.00054	0.00054	6.6E-07	0.12%	0.00054	2.585E-06	0.48%
3	2.65E-05	2.65E-05	1.29E-08	0.05%	2.72E-05	7.473E-07	2.82%
4	1.51E-06	1.59E-06	8.15E-08	5.41%	1.99E-06	4.812E-07	31.92%

Table 12: Calibration results for $T = 7$ years.

k	Exact M_k	Mixture M_k^*	*abs err*	*% err*	Vasicek M_k^*	*abs err*	*% err*
1	0.01809	0.01809	2.24E-06	0.01%	0.01810	1.009E-05	0.06%
2	0.00059	0.00060	4.61E-07	0.08%	0.00060	5.655E-07	0.10%
3	2.99E-05	3.00E-05	4.35E-08	0.15%	3.09E-05	9.324E-07	3.12%
4	1.75E-06	1.83E-06	8.77E-08	5.02%	2.28E-06	5.360E-07	30.70%

Table 13: Calibration results for $T = 10$ years.

Appendix

References

[1] M. Bonollo, F. Mercurio, and P. Mosconi. II second pillar: An analytical VaR with contagion and sectorial risks. *Working Paper, available at http://papers.ssrn.com*, 2009.

[2] D. Brigo, A. Pallavicini, and R. Torresetti. Calibration of cdo tranches with the dynamical generalized-poisson loss model. *RISK*, 20:70–75, 2006.

[3] A. Castagna, F. Mercurio, and P. Mosconi. Analytical credit var with stochastic probabilities of default and recoveries. *Wroking Paper, available at www.defaultrisk.com*, 2009.

[4] Li. D.i. On default correlation: A copula function approach. *Riskmetrics*, 2000.

[5] D. Duffie and N. Garleanu. Risk and valuation of collateralized debt obligations. *Financial Analysts Journal*, 57:41–59, 2001.

[6] A. Eckner. Computational techniques for basic affine models of portfolio credit risk. *J. Comput. Finance*, 15:63–97, 2009.

[7] K. Giesecke and B. Kim. Risk analysis of collateralized debt obligations, 2009.

[8] M.B. Gordy. A risk-factor model foundation for ratings-based bank capital rules. *Journal of Financial Intermediation*, 12:199–23, 2003.

[9] E. Levy. Pricing european average rate currency options. *Journal of International Money and Finance*, (14):474–491, 1992.

[10] A. Lopatin and T. Misirpashaev. Two-dimensional markovian model for dynamics of aggregate credit loss. *Adv. Econometrics*, 22:243–274, 2008.

[11] L. McGinty and R. Ahluwalia. Credit correlation: a guide. *Research Paper, JP Morgan*, 2004.

[12] R. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, (2):449–471, 1974.

[13] A Mortensen. Semi-analytical valuation of basket credit derivatives in intensity-based models. *Journal of Derivatives*, 13:8–26, 2006.

[14] Basel Commettee on Banking Supervision. Studies on credit risk concentration. *Basel Commettee Publication*, 2006.

[15] E. Papageorgiou and R. Sircar. Multiscale intensity models and name grouping for valuation of multi-name credit derivatives. *Applied Mathematical Finance*, 15:73–105, 2007.

[16] M. Pykhtin. Multi-factor adjustment. *Risk Magazine*, (3):85–90, 2004.

[17] S. Turnbull and L. Wakemann. A quick algorithm for pricing european average options. *Journal of Financial and Quantitative Analysis*, (26):377–389, 1992.

[18] O. Vasicek. Limiting loan loss probability distribution. *Working paper, KMV Corporation*, 1991.