

UNIT 5

MULTI NAME REDUCED FORM MODELS AND COPULAS

Damiano Brigo
www.damianobrigo.it

UNIT 5. Multi Name Reduced Form Models and Copulas

- Bottom up approach: from single defaults to the loss;
- Introducing dependence in defaults;
- Correlating intensities: Why doesn't it work;
- Correlating jump components: is linear correlation adequate?
- Copula functions.
- Factor Copulas.
- Pricing CDO's etc with factor copulas: Building the Loss distribution
  - Monte Carlo; Recurrence Approach ; Fast Fourier Transform
  - Probability Shifting ; Large Pool
- Implied correlation: Base and Compound.
- CDO Squared with Monte Carlo;
- Consistency with the whole correlation skew: the Implied copula of Hull and White

Notation

τ_i: default time of name i.

$\mathbb{Q}$: risk neutral probability measure.

$\mathbb{E}$, E: Risk Neutral expectation.

$\mathcal{F}_t$="info-on-default-free-markets-up-to-t";

$\mathcal{G}_t^i = \mathcal{F}_t \vee \sigma(\{\tau_i < u\}, u \le t)$, where $\sigma(\{\tau_i < u\}, u \le t) =$"info if default of name i occurred before t, and, if so, when exactly".

$\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\{$Number of defaulted names and total pool loss at any instant before $t)$.

$\mathbb{E}_t$: Risk neutral expectation, conditional on $\mathcal{G}_t$.

BOTTOM UP: Introducing dependence across single Defaults

Especially in the CDO payoffs, we have seen that the real underlying of the credit market is often the LOSS of the pool.

Models that consider directly the aggregate loss and worry later (if at all) about single name defaults are called TOP DOWN. We see these models later.

Models that consider single defaults, introduce dependence across them and then consider the aggregate loss are called BOTTOM UP.

These are the models we consider now. Mostly, they are based on copula functions.

BOTTOM UP. Credit Correlation

A first idea of correlation is given by the correlation between default indicators default correlation, given by (see for example Hull and White, 2000)

$$\rho = \frac{E[\mathbf{1}_{\{\tau_1 < T\}}\mathbf{1}_{\{\tau_2 < T\}}] - E[\mathbf{1}_{\{\tau_1 < T\}}]E[\mathbf{1}_{\{\tau_2 < T\}}]}{\mathrm{STD}[\mathbf{1}_{\{\tau_1 < T\}}]\mathrm{STD}[\mathbf{1}_{\{\tau_2 < T\}}]}$$

where τ_i is the default time for name i.

This is not a good definition for credit correlation: In fact **it is not possible to estimate it historically**. We would need the default data for identical companies, but it is obvious to see that this is not an easy task. Also, being the indicators non elliptically distributed in general, correlation would not be a good measure of dependence between them (more on this below).

Another possibility is to use **equity correlation as a proxy for credit correlation**, but this has **no formal justification**, except in particular structural models and under special conditions; also, typically **credit correlations are smaller than equity ones**.

BOTTOM UP: Introducing dependence across single Defaults

Simple single name reduced form models postulate that the default time τ be the first jump of a Poisson process with intensity $\lambda(t)$:

$$\text{Prob}(\tau \in [t, t+dt) | \tau > t, \text{market info up to } t) = \lambda(t)dt$$

where the "probability" dt factor λ is called **intensity** or **hazard rate**. It is also an **instantaneous credit spread** (more on this later). It is well known that in Poisson processes, transforming the 1st jump (default) time τ by its cumulated intensity $\Lambda(t) = \int_0^t \lambda(s)ds$ leads to an exponential random variable independent of any default-free quantity:

$$\Lambda(\tau) = \xi \sim \text{exponential}, \quad \mathcal{F} - \text{independent}.$$

If we assume λ to be strictly positive, we may define τ as

$$\tau = \Lambda^{-1}(\xi)$$

Introducing dependence in Defaults

Given names $1, 2, \ldots, n$, we may define dependency between the default times

$$\tau_1 = \Lambda_1^{-1}(\xi_1), \ldots, \tau_n = \Lambda_n^{-1}(\xi_n)$$

essentially in three ways.

1. Put dependency in (stochastic) intensities of the different names and keep the ξ of the different names independent;
2. Put dependency among the ξ of the different names and keep the (stochastic or trivially deterministic)intensities independent;
3. Put dependency both among (stochastic) intensities of the different names and among the ξ of the different names;

Introducing dependence in Defaults

1) Put dependency in (stochastic) intensities of the different names and keep the ξ of the different names independent;

In this case, one may induce dependence among the $\lambda_i(t)$ by taking diffusion dynamics for each of them and correlating the brownian motions.

$$d\lambda_i(t) = \mu_i(t, \lambda_i(t))dt + \sigma_i(t, \lambda_i(t))dW_i(t), \quad d\lambda_j(t) = \mu_j(t, \lambda_j(t))dt + \sigma_j(t, \lambda_j(t))dW_j(t),$$

$$dW_i dW_j = \rho_{i,j} dt, \quad \xi_i, \xi_j \text{independent}$$

Advantages: possible tractability; ease of implementation; default of one name does not affect the intensity of other names; The correlation can be estimated historically from time series of credit spreads;

Disadvantages: unrealistically low dependence across defaults $1_{\{\tau_i < T\}}, 1_{\{\tau_j < T\}}$. See e.g. Roncalli et al `http://gro.creditlyonnais.fr/content/wp/copula-intensity.pdf`

Introducing dependence in Defaults

2) Put dependency among the ξ of the different names and keep the (stochastic or trivially deterministic)intensities independent;

This is the framework that is currently used for correlation products in the market, especially for defining implied correlation.

Advantages: can take deterministic intensities, which makes life easier for the stripping of single name default probabilities; can reproduce sufficient levels of dependence across default times by putting dependence structures (called "copula functions") on the ξ's.

Disadvantages: no natural historical source for estimating the copula, often calibrated by means of dubious considerations; Default of one name affects the intensity of other names (realistic but untractable); Ignores credit spreads volatilities (large) and correlations.

Introducing dependence in Defaults

3) Put dependency among the ξ of the different names and among the stochastic intensities of different names (combine 1 and 2 above, very rarely used);

This is the most complicated framework.

Advantages: takes into account possible credit spread volatility, and can produce a sufficient amount of dependence among default times by putting dependence structures (called "copula functions") on the ξ's.

Disadvantages: no natural historical source for estimating the copula, often calibrated by means of dubious considerations; Default of one name affects the intensity of other names (realistic but untractable); Calculations are quite complicated, due to the presence of stochasticity both in the intensities and in the ξ's.

Introduction to Copula Functions

It is well known that linear correlation is not enough to express the dependence between two random variables in an efficient way in general.

Example: take X standard Gaussian and take $Y = X^3$. Y is a deterministic one-to-one transformation of X, so that the two variables give exactly the same information and should have maximum dependence. However, if we take the linear correlation between X and Y we easily get

$$(E(X^4) - E(X^3)E(X))/(\text{Std}(X^3)\text{Std}(X)) = 3/\sqrt{15} = \sqrt{3}/\sqrt{5} < 1$$

We get a dependence measure that is smaller than 1 (1 means maximum dependence).

So **correlation is not a good measure of dependence in this case**.

Introduction to Copula Functions

In standard financial models this problem with correlation as a dependence measure is usually absent because we are concerned with dependence between instantaneous Brownian shocks, which are Gaussian. **Correlation works well for Gaussian variables**.

In credit derivatives with intensity models we may find ourselves in the situation where we need to introduce dependence between the *exponential* components $\xi = \Lambda(\tau)$ of Poisson processes for different names.

This is usually done by means of Copula functions.

Introduction to Copula Functions

Given a random variable X, we may transform it in several ways through a deterministic function: $2X$, X^5, $\exp(X)$,.. A particularly interesting transformation function is the cumulative distribution function F_X of X. Let $U = F_X(X)$.

$$F_U(u) = \mathbb{Q}(U \leq u) = \mathbb{Q}(F_X(X) \leq u) = \mathbb{Q}(X \leq F_X^{-1}(u))$$

$$= \mathbb{Q}(X \leq z) = F_X(Z) = F_X(F_X^{-1}(u)) = u$$

However, the identity distribution function $F_U(u) = u$ is characteristic of a uniform random variable in $[0\ 1]$. This means that $U = F_X(X)$ is a **uniform distribution function**. Notice that since F_X is one to one, U contains the same information as X.

The idea then is to transform all random variables X by their F_X obtaining all uniform variables that contain the same information as the starting X. This way we rid ourselves of marginal distributions, obtaining only uniform rv's, *and can concentrate on introducing dependence directly for these standardized uniforms*.

Introduction to Copula Functions

Let $(U_1, ..., U_n)$ be a random vector with uniform margins and joint distribution $C(u_1, ..., u_n)$. $C(u_1, ..., u_n)$ is the *copula* of the random vector. It can be characterized by a number of properties that we do not repeat here.

An important result is given by **Sklar's theorem**: Let H be an n-dimensional distribution function with margins $F_1, ..., F_n$. Then there exists an n-copula C such that for all $\mathbf{x}$ in $\bar{\mathbb{R}}^n$,

$$H(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n)). \tag{18}$$

This result tells us that what we hinted at in the previous slide works. Indeed, one would intuitively write

$$\begin{aligned} H(x_1, ..., x_n) &= \mathbb{Q}(X_1 \leq x_1, \ldots, X_n \leq x_n) \\ &= \mathbb{Q}(F_1(X_1) \leq F_1(x_1), ..., F_n(X_n) \leq F_n(x_n)) \\ = \mathbb{Q}(U_1 \leq F_1(x_1), ..., U_n \leq F_n(x_n)) &= C(F_1(x_1), ..., F_n(x_n)) \end{aligned}$$

where C is the joint distribution function of uniforms $U_1, ..., U_n$.

Introduction to Copula Functions

Sklar's theorem: For any joint distribution function $H(x_1, ..., x_n)$ with margins $F_1, ..., F_n$ there exists a copula function $C(u_1, ..., u_n)$ (i.e. a joint distribution function on n uniforms) such that

$$H(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n))$$

Notice that C contains the pure dependence information. Notice the important point: Correlation between two variables is just a number, whereas a copula function between two variables is a two dimensional function.

We may also write

$$C(u_1, ..., u_n) = H(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)). \tag{19}$$

from which we see that we may use any known joint distribution function to define a copula

Introduction to Copula Functions

Consider again our example with X standard Gaussian and $Y = X^3$, $Z = X$. The copula between X and Z is the copula expressing maximum dependence (also correlation works in this case: $\text{Corr}(X, Z) = 1$).
This copula is the joint distribution of $U_1 = F_X(X)$ and $U_2 = F_Z(Z) = F_X(X) = U_1$,

$$\mathbb{Q}(U_1 < u_1, U_1 < u_2) = \mathbb{Q}(U_1 < \min(u_1, u_2)) = \min(u_1, u_2).$$

So this "min" copula corresponds to maximum dependence.
Now consider $Y = X^3$ and the copula between X and Y. Call $U_3 = F_Y(Y)$. This copula is

$$\mathbb{Q}(U_1 < u_1, U_3 < u_2) = \mathbb{Q}(F_X(X) < u_1, F_{X^3}(X^3) < u_2) =$$

$$\mathbb{Q}(F_X(X) < u_1, \mathbb{Q}(X^3 \le k)|_{k=X^3} < u_2) = \mathbb{Q}(F_X(X) < u_1, \mathbb{Q}(X \le k^{1/3})|_{k=X^3} < u_2) =$$
$$= \mathbb{Q}(F_X(X) < u_1, F_X(k^{1/3})|_{k=X^3} < u_2) = \mathbb{Q}(F_X(X) < u_1, F_X(X) < u_2),$$

the same as before. So with copulas also X and X^3 get maximum dependence.

Introduction to Copula Functions

This example actually has a more general version: if $g_1, \ldots, g_n$ are (say strictly increasing) one-to-one transformations, then the copula of some given $X_1, \ldots, X_n$ is the same as the copula for $g_1(X_1), \ldots, g_n(X_n)$ (not so for correlation). So **the copula is invariant for deterministic transformations that preserve the information**.

This tells us again that Copulas are really expressing the core of dependence.

In particular we can find some quantities (beside the standard linear correlation) expressing the level of *concordance* between two continuous random variables whose copula is C.

We now give two examples.

Introduction to Copula Functions

We present here the two most important measures of concordance. They provide the perhaps best alternatives to the linear correlation coefficient as a measure of dependence for pairs of non-gaussian (and non-elliptical) distributions, for which the linear correlation coefficient is inappropriate and often misleading.

- **Kendall's Tau** between two random variables X, Y

$$\tau(X,Y) = \mathbb{Q}\{(X - \tilde{X})(Y - \tilde{Y}) > 0\} - \mathbb{Q}\{(X - \tilde{X})(Y - \tilde{Y}) < 0\}$$

where $(\tilde{X},\tilde{Y})$ is i.i.d. as (X,Y). It can be proved that if (X,Y) is a couple of continuous random variables with copula C, then

$$\tau(X,Y) = 4 \int\int_{[0,1]^2} C(u,v) dC(u,v) - 1$$

Kendall's Tau for a pair of random variables (X,Y) is invariant under strictly increasing componentwise transformations.

Introduction to Copula Functions

- **Spearman's Rho** between two random variables X, Y

$$\rho_S(X,Y) = 3\mathbb{Q}\{(X-\tilde{X})(Y-Y') > 0\} - \mathbb{Q}\{(X-\tilde{X})(Y-Y') < 0\}$$

where (X,Y), (X',Y') and $(\tilde{X},\tilde{Y})$ are i.i.d. pairs. It can be proved that if (X,Y) is a couple of continuous random variables with copula C, then

$$\rho(X,Y) = 12 \int\int_{[0,1]^2} C(u,v)du\,dv - 3$$

Spearman's Rho for a pair of random variables (X,Y) is invariant under strictly increasing componentwise transformations.

Also, it can be proved that every copula C is bounded between tho functions C^+ and C^-, which are known as the Fréchet-Hoeffding bounds:

$$C(u_1,u_1,\dots,u_n)^- \le C(u_1,u_1,\dots,u_n) \le C(u_1,u_1,\dots,u_n)^+$$

Introduction to Copula Functions

$$C(u_1, u_1, \ldots, u_n)^- \leq C(u_1, u_1, \ldots, u_n) \leq C(u_1, u_1, \ldots, u_n)^+$$

where

$$C(u_1, u_1, \ldots, u_n)^- = \max(u_1 + u_2 + \ldots + u_n - 1, 0)$$

and

$$C(u_1, u_1, \ldots, u_n)^+ = \min(u_1, u_2, \ldots, u_n)$$

as in our example above.

While C^+ is a copula, C^- is a copula only in dimension 2.

We can define also an orthogonal copula $C^\perp$ corresponding to independent variables:

$$C(u_1, u_1, \ldots, u_n)^\perp = u_1 \cdot u_2 \cdot \ldots \cdot u_n$$

Introduction to Copula Functions

Consider now (X,Y) a pair of random variables with copula C, then

$$\begin{aligned} C = C^{+} &\quad \rightarrow \quad \tau_C = \rho_C = 1 \\ C = C^{\perp} &\quad \rightarrow \quad \tau_C = \rho_C = 0 \\ C = C^{-} &\quad \rightarrow \quad \tau_C = \rho_C = -1 \end{aligned}$$

Then C^{+} is the copula corresponding to the maximum dependence (correspondence) and C^{-} is the copula corresponding to the minimum dependence (correspondence). $C^{\perp}$ corresponds to perfect independence between two variables.

Before starting to introduce the most important families of copulas, let us define the concept of tail dependence.

Introduction to Copula Functions

The concept of tail dependence relates to the amount of dependence in the upper-right quadrant tail or lower-left-quadrant tail of a bivariate distribution. It is a concept that is relevant for the study of dependence between extreme values. Roughly speaking, it is the idea of "fat tails" for the dependence structure.

It turns out that tail dependence between two continuous random variables X and Y is a copula property and hence the amount of tail dependence is invariant under strictly increasing transformations of X and Y .

Definition: Let (X,Y) be a couple of continuous random variables with marginal distribution functions F and G. The coefficient of upper tail dependence of (X,Y) is

$$\lim_{u \uparrow 1} \mathbb{Q}\{Y > G^{-1}(u) | X > F^{-1}(u)\} = \lambda_U$$

provided that the limit $\lambda_U \in [0,1]$ exists. If $\lambda_U \in (0,1]$ X and Y are said to be asymptotically dependent in the upper tail; if $\lambda_U = 0$, X and Y are said to be asymptotically independent in the upper tail.

Introduction to Copula Functions

$$\lim_{u \uparrow 1} \mathbb{Q}\{Y > G^{-1}(u) | X > F^{-1}(u)\} = \lambda_U$$

Since $\mathbb{Q}\{Y > G^{-1}(u) | X > F^{-1}(u)\}$ can be rewritten as

$$\frac{1 - \mathbb{Q}\{X \leq F^{-1}(u)\} - \mathbb{Q}\{Y \leq G^{-1}(u)\} + \mathbb{Q}\{X \leq F^{-1}(u), Y \leq G^{-1}(u)\}}{1 - \mathbb{Q}\{X \leq F^{-1}(u)\}}$$

an alternative and equivalent definition (for continuous random variables), from which it is seen that the concept of tail dependence is indeed a copula property, is the following:

$$\lim_{u \uparrow 1} (1 - 2u + C(u,u))/(1-u) = \lambda_U$$

The concept of lower tail dependence can be defined in a similar way. If the limit

$$\lim_{u \downarrow 0} \mathbb{Q}\{Y \leq G^{-1}(u) | X \leq F^{-1}(u)\} = \lim_{u \downarrow 0} C(u,u)/u = \lambda_L$$

exists, then C has lower tail dependence if $\lambda_L \in (0, 1]$, and lower tail independence if $\lambda_L = 0$.

Some important copulas: Gaussian copulas

The canonical copula is the Gaussian (or Normal) copula, obtained using a multivariate normal distribution Φ_R^n as H:

$$C(u_1, ..., u_n) = \Phi_R^n(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n)) \tag{20}$$

where the Gaussian margins have all zero mean and unit variance, R is the correlation matrix and Φ^{-1} is the inverse of the usual standard normal cdf. Unfortunately this copula cannot be expressed in closed form. Indeed, in the 2-dimensional case we have:

$$C_R(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left\{ -\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)} \right\} ds\, dt, \tag{21}$$

ρ being the (only) correlation parameter in the matrix R. Notice that in case we are modelling the dependence among N names, the correlation matrix R in principle has $N(N-1)/2$ free parameters. Some properties:

- neither upper nor lower tail dependence
- $C(u, v) = C(v, u)$ i.e. *exchangeable copula*.

Some important copulas: Archimedean copulas

Archimedean copulas are an important class of copulas with a great quality: they can be expressed in closed form. In general Archimedean copulas arise from a particular function φ called the *generator* of the copula. In particular, if $\varphi : [0,1] \to [0,\infty)$ is a continuous, strictly decreasing function such that $\varphi(1) = 0$, then

$$C(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \tag{22}$$

is a copula if and only if φ is convex (in other words it must be $\varphi' < 0$ and $\varphi'' > 0$ under differentiability). We recall that $\varphi^{[-1]}$ is the *pseudo-inverse* of φ defined as:

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & 0 \le t \le \varphi(0) \\ 0 & t > \varphi(0) \end{cases}$$

If $\lim_{t\to 0} \varphi(t) = +\infty$ we say that φ is a *strict* generator and the copula is said to be a *strict* copula. According to the particular generator used, we have different families of copula functions. We give examples in dimension 2 but they can be easily generalized.

It happens that even if Archimedean copulas are known in closed form, they are difficult to simulate. On the contrary Gaussian copulas are not known in closed form but are easier to simulate. We will face later this issue of sampling from copulas.

Clayton family

Let us choose $\varphi(t) = (t^{-\theta} - 1)/\theta$ where $\theta \in [-1, \infty)\backslash\{0\}$. Then the Clayton family is:

$$C_\theta(u, v) = \max([u^{-\theta} + v^{-\theta} - 1], 0)^{-1/\theta}. \quad (23)$$

If $\theta > 0$ the copulas are strict and the copula expression simplifies to

$$C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}. \quad (24)$$

The Clayton copula has lower tail dependence for $\theta > 0$, and $C_{-1} = C^-$, $\lim_{\theta \to 0} C_\theta = C^\perp$ and $\lim_{\theta \to \infty} C_\theta = C^+$.

Frank family

Let us choose $\varphi(t) = -\ln \frac{e^{-\theta t}-1}{e^{-\theta}-1}$, where $\theta \in \mathbb{R}\backslash\{0\}$. This originates the Frank family

$$C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right). \qquad (25)$$

The Frank copulas are strict Archimedean copulas. Furthermore $\lim_{\theta\to-\infty} C_\theta = C^-$, $\lim_{\theta\to\infty} C_\theta = C^+$ and $\lim_{\theta\to 0} C_\theta = C^\perp$. The members of the Frank family are the only Archimedean copulas which satisfy the equation $C(u,v) = \breve{C}(u,v)$ (remark: only in two dimensions!!!). The members of the Frank family have no tail dependence (neither upper nor lower).

$\breve{C}(u,v)$ is the *survival copula* defined as

$$\mathbb{P}\big[X_1 > x_1, \ldots, X_n > x_n\big] = \breve{C}(\bar{F}_1(x_1), \ldots, \bar{F}_n(x_n))$$

where the $\bar{F}$'s are the margins survival functions. The survival copula is not linked to the copula in a simple way: It can be proved that in two dimensions the following relation holds:

$$\breve{C}(u,v) = u + v - 1 + C(1-u, 1-v).$$

Gumbel family

Let us choose $\varphi(t) = (-\ln t)^\theta$, where $\theta \geq 1$. This gives the Gumbel family

$$C_\theta(u, v) = \exp(-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}). \quad (26)$$

The Gumbel copulas are strict Archimedean copulas. Furthermore $\lim_{\theta\to\infty} C_\theta = C^+$, $C_1 = C^\perp$. Gumbel copulas describe **only positive dependence** between random variables; moreover they feature upper tail dependence.

Marshall Olkin Copula

This copula is outside the Archimedean family and is given by

$$C_{\alpha_1,\alpha_2}(u_1, u_2) = \min(u_1^{1-\alpha_1} u_2, u_1 u_2^{1-\alpha_2})$$

for two numbers $0 \leq \alpha_1, \alpha_2 \leq 1$. This copula has both an absolutely continuous and a singular component.

Its survival copula

$$\check{C}(u_1, u_2) = u_1 u_2 \min(u_1^{-\alpha_1}, u_2^{-\alpha_2})$$

links the first jump times of correlated Poisson processes coming from a Common Poisson Shock framework (more on this later with the GPCL loss model).

If $\alpha_1 = \alpha_2 = 0$ we have independence, otherwise if it is 1 we have C^+.

Some important copulas: t-Copulas

If the vector $\mathbf{X}$ of random variables has the stochastic representation $\mathbf{X} \sim \mu + \frac{\sqrt{\nu}}{\sqrt{S}}\mathbf{Z}$ where $\mu \in \mathbb{R}^n$, ν is a positive integer, $S \simeq \chi^2_\nu$ and $\mathbf{Z} \simeq \mathcal{N}_n(\mathbf{0}, \Sigma)$ are independent, then $\mathbf{X}$ has an n-variate t_ν-distribution with mean μ (for $\nu > 1$) and covariance matrix $\frac{\nu}{\nu-2}\Sigma$ (for $\nu > 2$). If $\nu \leq 2$ then Cov$(\mathbf{X})$ is not defined. In this case we just interpret Σ as being the shape parameter of the distribution of $\mathbf{X}$.

The copula of $\mathbf{X}$ defined above is can be written as

$$C^t_{\nu,R}(\mathbf{u}) = t^n_{\nu,R}(t^{-1}_\nu(u_1), \ldots, t^{-1}_\nu(u_n))$$

where $R_{ij} = \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$ for $i, j \in \{1, \ldots, n\}$ and where $t^n_{\nu,R}$ denotes the distribution function of $\sqrt{\nu}\mathbf{Y}/\sqrt{S}$ where $S \simeq \chi^2_\nu$ and $\mathbf{Y} \simeq \mathcal{N}_n(\mathbf{0}, R)$ are independent. Here t_ν denotes the (equal) margins of $t^n_{\nu,R}$, i.e. the distribution function of $\sqrt{\nu}Y_1/\sqrt{S}$.

Some important copulas: t-Copulas

In the bivariate case the copula expression can be written as

$$C^t_{\nu,R}(u,v) = \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(v)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} \left\{ 1 + \frac{s^2 - 2R_{12}st + t^2}{\nu(1-R_{12}^2)} \right\}^{-(\nu+2)/2} ds\, dt$$

Note that R_{12} is simply the usual linear correlation coefficient of the corresponding bivariate t_ν-distribution if $\nu > 2$. If (X_1,X_2) has a standard bivariate t-distribution with ν degrees of freedom and linear correlation matrix R, then $X_2|X_1 = x$ is t-distributed with $\nu + 1$ degrees of freedom and $\mathbb{E}(X2|X1 = x) = R_{12}x$, $\text{Var}(X_2|X1 = x) = \left(\frac{\nu+x^2}{\nu+1}\right)(1 - R_{12}^2)$.

Summary on copula properties

In the following table we collect the properties of the different copulas considered so far.

Copula	*Positive Dependence*	*Independence*	*Negative Dependence*	*Upper Tail Dependence*	*Lower Tail Dependence*
Clayton $\theta \in [-1, +\infty)$ $\theta \neq 0$	$C \to C^+$ $\theta \to +\infty$	$C \to C^\perp$ $\theta \to 0$	$C = C^-$ $\theta = -1$	no	only for $\theta > 0$
Frank $\theta \in \mathbb{R}\backslash\{0\}$	$C \to C^+$ $\theta \to +\infty$	$C \to C^\perp$ $\theta \to 0$	$C \to C^-$ $\theta \to -\infty$	no	no
Gumbel $\theta \in [1, +\infty)$	$C \to C^+$ $\theta \to +\infty$	$C = C^\perp$ $\theta = 1$	no negative dependence	yes	no
Gaussian $\rho \in (-1, 1)$	$C \to C^+$ $\rho \to +1$	$C = C^\perp$ $\rho = 0$	$C \to C^-$ $\rho \to -1$	no	no
t-Copula $R_{12} \in (-1, 1)$	$C \to C^+$ $R_{12} \to +1$ $\nu \to \infty$	$C = C^\perp$ $R_{12} = 0$ $\nu \to \infty$	$C \to C^-$ $R_{12} \to -1$ $\nu \to \infty$	yes	yes

Sampling from Copulas

The Gaussian copula has no closed form but is simple to simulate. Instead the Archimedean copulas have closed form but are difficult to sample, especially in high dim.

Here we present a general algorithm to sample out from copulas. The algorithm is based on general considerations. Let

$$C_k(u_1, \dots, u_k) = C(u_1, \dots, u_k, 1, \dots, 1), \quad k = 2, \dots, n-1$$

denote the $k-$dimensional margins of C, with $C_1(u_1) = u_1$ and $C_n(u_1, \dots, u_n) = C(u_1, \dots, u_n)$. Let $U_1, \dots, U_n$ have a joint distribution function C. Then the conditional distribution of U_k given the values of $U_1, \dots, U_{k-1}$, is given by

$$\begin{aligned} C_k(u_k|u_1, \dots, u_{k-1}) &= \mathbb{P}\{U_k \le u_k | U_1 = u_1, \dots, U_{k-1} = u_{k-1}\} \\ &= \frac{\partial^{k-1} C_k(u_1, \dots, u_k)}{\partial u_1 \dots \partial u_{k-1}} \Big/ \frac{\partial^{k-1} C_{k-1}(u_1, \dots, u_{k-1})}{\partial u_1 \dots \partial u_{k-1}} \end{aligned}$$

given that the numerator and the denominator exist and that the latter is $\neq 0$.

Sampling from Copulas

ARCHIMEDEAN COPULAS

- Simulate a random variable u_1 from $U(0,1)$.
- Simulate a random variable u_2 from $C_2(\cdot|u_1)$.

$\vdots$

- Simulate a random variable u_n from $C_n(\cdot|u_1, \ldots, u_{n-1})$.

GAUSSIAN COPULAS

- Find the Cholesky decomposition A of R where R is the correlation matrix
- Simulate n independent random variables $z_1, \ldots, z_n$ from $\mathcal{N}(0,1)$
- Set $\mathbf{x} = A\mathbf{z}$
- Set $u_i = \Phi(x_i)$, $i = 1, \ldots, n$
- $(u_1, \ldots, u_n)^T \sim C_R$.

Sampling from Copulas

t-COPULAS

- Find the Cholesky decomposition A of R
- Simulate n independent random variables $z_1, \ldots, z_n$ from $\mathcal{N}(0,1)$
- Simulate a random variate s from χ^2_ν independent of $z_1, \ldots, z_n$
- Set $\mathbf{y} = A\mathbf{z}$
- Set $x_i = \frac{\sqrt{\nu}}{\sqrt{s}} y_i$
- Set $u_i = t_\nu(x_i)$, $i = 1, \ldots, n$
- $(u_1, \ldots, u_n)^T \sim C^t_{\nu,R}$.

Sampling from Copulas

The main use of copula functions is to simulate dependent default times in a portfolio of defaultable names. We know that a single default is modelled by a Poisson process, in particular under positive intensity λ and cumulated intensity $\Lambda(t) = \int_0^t \lambda(s)ds$ we have

$$\tau = \Lambda^{-1}(\xi) = \Lambda^{-1}(-\ln(1-U))$$

where ξ is an exponential random variable that is $\mathcal{F}$-independent, and U is a uniform random variable that is $\mathcal{F}$-independent. Indeed, if we transform an exponential ξ with its distribution $1 - e^{-x}$ we have an uniform, $U = 1 - \exp(-\xi)$. Remember this one.

Sampling from Copulas

If we need to simulate a single default time τ we need to simulate first its intensity and then an independent uniform random variable; at this point we take minus the log of the uniform and apply the inverse of the cumulated intensity.

What in case of multiple dependent defaults?

$$\tau_1 = \Lambda_1^{-1}(-\ln(1 - U_1)), \tau_2 = \Lambda_2^{-1}(-\ln(1 - U_2)), ..., \tau_N = \Lambda_N^{-1}(-\ln(1 - U_N)).$$

Usually λ's are taken deterministic or independent of each other, and the dependency between the τ's is loaded into a copula on the U's.

One Factor Copula

In general practitioners use the Gaussian copula as a standard tool in pricing credit derivatives. But, as mentioned above, the number of free parameters in this case is $N(N-1)/2$ and it could be quite a problem when dealing with many names. Moreover the Gaussian copula has no analytical tractability, and this is an undesirable feature.

Still, the Gaussian (or t-) copula has the advantage to break complete dependence into sets of pairwise dependencies.

(One-)Factor copulas are a way to simplify the pricing process, since they reduce the dimensionality of the problem and also present more analytical tractability. This is the reason why they are attracting more and more interest and are becoming a standard when pricing CDOs and CDS index tranches.

One Factor Copula

One Factor copulas in the intensity/reduced form framework are particular Gaussian copulas on the $U = -\ln(1-\xi)$ coming from a factor structure.

However, it is possible to give a derivation of this copula inspired by the other (structural) framework, and by Merton's model in particular.

Recall that in Merton's model a company incurs in a default if the value of the firm falls below the debt value at the debt maturity.

Let us call V the gaussian log-firm value and K the log debt value: there is a default when $V < K$. Let us assume to have standardized V (zero mean and unit variance).

Prob to have a default is $\mathbb{Q}\{V < K\} = \Phi(K)$

Now let us imagine to deal with $i = 1, 2, \ldots, N$ different companies: Then each company has value V_i and debt K_i.

One Factor Copula

Following the work of Vasicek (Vasicek, 1987), we assume that each process V_i is driven by a common (systematic) factor and a specific (hydiosyncratic) factor.

$$V_i = \sqrt{\rho_i}\, M + \sqrt{1-\rho_i}\, Y_i$$

where Y_i, M are i.i.d as $N(0,1)$ so that $\text{corr}(V_i, V_j) = \sqrt{\rho_i \rho_j}$.

Now the N parameters ρ_i are sufficient to construct the whole correlation structure, and this is an improvement with respect to the standard Gaussian copula (N (N-1)/2). In particular if $\rho_i = \rho$ then $\text{corr}(V_i, V_j) = \rho$, unique correlation parameter. Important: conditional on the common M defaults are **independent** and default prob is:

$$\mathbb{Q}\{V_i < K_i | M\} \quad = \quad \mathbb{Q}\left(\sqrt{\rho_i}\, M + \sqrt{1-\rho_i}\, Y_i < K_i \middle| M\right) = \Phi\left(\frac{K_i - \sqrt{\rho_i}\, M}{\sqrt{1-\rho_i}}\right).$$

The value of the debt K_i is inversely computed as $K_i = \Phi^{-1}(p_i)$ where p_i is the default probability of each company i and is usually computed apart.

One Factor Copula

This is our inspiration, but we define now the copula directly on the U_i's leading to the default time in the intensity framework,

$$\tau_1 = \Lambda_1^{-1}(-\ln(1-U_1)), \tau_2 = \Lambda_2^{-1}(-\ln(1-U_2)), ..., \tau_N = \Lambda_N^{-1}(-\ln(1-U_N)).$$

We take $U_i = \Phi(X_i)$ where the X_i's are Gaussian variables defined as

$$X_i = \sqrt{\rho_i}\, M + \sqrt{1-\rho_i}\, Y_i, \qquad (27)$$

Y_i, M as before. Let us consider the Gaussian copula (20) in this framework:

$$C(u_1, \ldots, u_n) = \mathbb{Q}(X_1 < \Phi^{-1}(u_1), \ldots, X_n < \Phi^{-1}(u_n)). \qquad (28)$$

By iterated expectations, we can write the previous term as:

$$\mathbb{E}\left(\mathbb{Q}(X_1 < \Phi^{-1}(u_1), \ldots, X_n < \Phi^{-1}(u_n)|M)\right), \qquad (29)$$

One Factor Copula

$$\mathbb{E}\left(\mathbb{Q}(X_1 < \Phi^{-1}(u_1), \ldots, X_n < \Phi^{-1}(u_n)|M)\right), \tag{30}$$

By independence of X's, conditional on M, the probability is the product of probabilities

$$\mathbb{Q}_{|M}(X_i < \Phi^{-1}(u_i)) = \mathbb{Q}_{|M}(\sqrt{\rho_i}\, M + \sqrt{1-\rho_i}\, Y_i, < \Phi^{-1}(u_i)) = \Phi\left(\frac{\Phi^{-1}(u_i) - \sqrt{\rho_i} m}{\sqrt{1-\rho_i}}\right),$$

given that Y's are independent standard Gaussians, of which we take the external expectation:

$$C(u_1, \ldots, u_n) = \int \left(\prod_{i=1}^{n} \Phi\left(\frac{\Phi^{-1}(u_i) - \sqrt{\rho_i} m}{\sqrt{1-\rho_i}}\right)\right) \varphi(m)\, dm, \tag{31}$$

where $\varphi(v)$ is the usual standard Gaussian density (of M). **Thus, the n-dimensional copula is computed through a one dimensional integral.** In general also multi factor decompositions are possible, but this increases the complexity (see e.g. Laurent and Gregory, 2004 and Hull and White, 2004). We will see applications of this copula below.

Copulas and Multiname Credit Derivatives

How are copulas used in the credit market?

As we have seen before, one group of products that is attracting interest is the family of **multi-name products**. Since they deeply depend on the correlation between the underlying names, sometimes they are also called **correlation products**.

As basic examples of correlation products we have seen the n-th to Default, the Collateralized Debt Obligation (CDO), CDOs of CDOs, i.e. the so called CDO squared, and Leveraged Super Senior Tranches.

Copulas can help in understanding the meaning of credit correlation, how to extract values of this correlation from market data and how to use it in pricing engines.

Monte Carlo pricing with Copulas

To price multi-name credit products we need i) a pricing model and ii) an estimate of the value of the default dependence (copula) parameters to put in the model.

Let us skip for a moment the second issue (later) and let us suppose to have somehow estimated the copula parameters and concentrate on the pricing.

If we can write the payoff of the contract with respect to the series of default times, than it is natural to evaluate the contract by means of Monte Carlo simulations.

More precisely we can simulate a collection of default times $\tau_1, \ldots, \tau_N$ for the N components of the portfolio by means of copula and intensity, as seen before, and evaluate the realization of the payoff under each scenario.

Then we repeat it many times and average over scenarios, getting the price.

Monte Carlo pricing of a First to Default

For example we can think to price a First to Default (FtD) using this procedure.

As we have seen earlier, the (Running) FtD discounted payoff to the protection seller at time $t < T_a$ is $\Pi\text{RFtD}_{a,b}(t) :=$

$$D(t,\tau^1)(\tau^1 - T_{\beta(\tau^1)-1})R\mathbf{1}_{\{T_a<\tau^1<T_b\}} + \sum_{i=a+1}^{b} D(t,T_i)\alpha_i R\mathbf{1}_{\{\tau^1>T_i\}} - \mathbf{1}_{\{T_a<\tau^1\leq T_b\}}D(t,\tau^1)\ \text{LGD}_{i_1}$$

So we just simulate a series of default times and compute the payoff for that collection $\tau_1, \ldots, \tau_n$. Then we repeat this procedure simulating other sets of default times and have an estimate of $\text{FtD}_{a,b}(t) = \mathbb{E}_t\{\Pi\text{RFtD}_{a,b}(t)\}$.

Usually in the case of k-th to default baskets, the number of components is around 5 or 10, so the computational effort of a Monte carlo approach is not dramatic. The situation changes when considering CDO's and CDS's indices, where the number is around 100.

Monte Carlo pricing of a First to Default

If one wants to speed up the computation one can use the control variate variance reduction technique. Possible "analytical" payoffs (see earlier treatment of control variate on single name intensity models) to be used as control variate are

$$1_{\{\tau_1<T,\tau_2<T\}} + 1_{\{\tau_1<T,\tau_3<T\}} + \ldots + 1_{\{\tau_{N-1}<T,\tau_N<T\}}$$

where the expectation of each term, in case of Gaussian copulas, depends on pairwise correlations and can be computed analytically given approximations of the bivariate normal distribution function. Another possible payoff is the indicator

$$1_{\{\tau^1<T\}} = 1 - 1_{\{\tau_1\geq T,\tau_2\geq T,\ldots,\tau_N\geq T\}}$$

(without discount $D(0,\tau^1)$) whose expectation is the survival copula associated with the given copula model underlying default dependence. If this is known in closed form, we have an analytical value for this payoff.

CDO pricing

One of the key problems of CDO valuation is the estimation of the loss distribution of the underlying portfolio. This can be a problem when resorting to the commonly used Monte Carlo simulations, which can be time-consuming.

Factor models represent a useful and efficient framework to model the dependence structure: they can be used to derive portfolio loss distributions BOTTOM UP without numerical simulations. Instead, the solutions either have a closed-form or require low-dimensional numerical integration.

The complexity of computing the loss distribution with factor models is a function of the composition of the portfolio. Complexity is low for portfolios that comprise homogeneous credits, i.e. credits that have similar spread, recovery and correlation characteristics. Typical portfolios underlying CDOs, however, are not homogeneous and require advanced numerical techniques (such as Fast Fourier Transform) to determine loss distributions.

CDO pricing

A common approach to compute the fair value of a synthetic CDO is based on Monte Carlo simulations. The basic inputs in this approach include:

- The individual credit spread or default probability for each obligor
- The copula parameters in the chosen copula among the $U_i = -\ln(1 - \xi_i)$.

This basic MC setting suffers from two main drawbacks when considering CDO's:

- For an accurate estimate requires a large number of simulations. Time consuming.
- Estimation of the default correlation matrix. With N obligors, the $N \times N$ pairwise correlation matrix requires $N \cdot (N-1)/2$ estimates (we see later how to obtain them from market data) if we use a Gaussian copula parameterized by correlations.

CDO pricing: One-Factor approach

$$\tau_1 = \Gamma_1^{-1}(\xi_1), \ldots, \tau_N = \Gamma_N^{-1}(\xi_N).$$

We have already seen that in the one-factor approach, under deterministic intensity γ, each uniform $U = 1 - \exp(-\xi)$ can be decomposed into two components: A systematic factor M representing the market and a hydiosyncratic factor specific of that firm. More precisely, for each obligor i in the portfolio we have $1 - \exp(-\xi_i) = U_i = \Phi(X_i)$ with

$$X_i = \Phi^{-1}(1 - \exp(-\xi_i)) = \sqrt{\rho_i}\, M + \sqrt{1 - \rho_i}\, Y_i$$

where Y_i, M are i.i.d as $N(0, 1)$ and corr$(X_i, X_j) = \sqrt{\rho_i \rho_j}$.

Assume $\rho_i = \rho$ for each i, i.e. the correlation parameter is the same for each obligor. Then corr$(X_i, X_j) = \rho$ for each pair of names.

This is actually the market practice for which a single value of correlation is assumed across the portfolio when implying correlation from quoted CDO tranches for indices (later).

CDO pricing: One-Factor approach

We have, under deterministic intensity, recalling that $U = 1 - e^{-\xi}$, that the probability of the single default conditional on M with the factor copula is, as seen before:

$$\mathbb{Q}(\tau_i < T | M = m) = \mathbb{Q}(\Gamma_i(\tau_i) < \Gamma_i(T) | M = m) = \mathbb{Q}(\xi_i < \Gamma_i(T) | M = m) =$$

$$= \mathbb{Q}(U_i < 1 - \exp(-\Gamma_i(T)) | M = m) = \Phi\left(\frac{\Phi^{-1}(1 - \exp(-\Gamma_i(T))) - \sqrt{\rho_i} m}{\sqrt{1 - \rho_i}}\right)$$

CDO pricing: One-Factor approach

$$\mathbb{Q}(\tau_i < T | M = m) = \Phi\left(\frac{\Phi^{-1}(1 - \exp(-\Gamma_i(T))) - \sqrt{\rho_i}m}{\sqrt{1-\rho_i}}\right)$$

As noticed before, since conditional on M defaults are independent, joint default probabilities are just products of these objects, conditional on M. Then the only randomness left comes from M and can be integrated away with one last integration.

The final step is to determine the expected loss of the underlying portfolio in terms of these default times. Though factor models provide an appealing framework to perform this task the mathematical complexity and analytical tractability of the model itself will vary according to the specific properties of the underlying portfolio of credits.

CDO pricing: One-Factor approach

In particular the effectiveness of Factor models will depend on two main measures to assess the complexity of a portfolio of obligors:

- Portfolio Size: The number N of credits in the underlying portfolio
- Homogeneity: The uniformity of the underlying credits with respect to notional, recovery REC_i, default probability (see Γ_i) and correlation to common factors (ρ_i).

We are going to illustrate some methods to evaluate the Loss Distribution of the portfolio in the Factor model according to different specifics of the underlying portfolio itself.

CDO: One-Factor approach: Finite Size, Homogeneous Portfolio

We assume homogeneity in the portfolio in the sense that we consider **equal recovery rates**, **equal default probabilities** and **equal correlation** for the names in the basket.

Then, the probability of a single default in the portfolio, conditional on M, is

$$\mathbb{Q}(\tau_i < T|M = m) = \Phi\left(\frac{\Phi^{-1}(1-\exp(-\Gamma(T))) - \sqrt{\rho}m}{\sqrt{1-\rho}}\right)$$

and it is independent on the defaulting name i.

The main feature of the factor approach is that **conditionally to the common factor M the names are independent**. So we can easily compute the conditional probability to have n defaults among the N obligors with the Binomial distribution, given that now all names have the same statistics:

$$\mathbb{Q}\{n\,\text{defaults}|M = m\} = \binom{N}{n}\mathbb{Q}(\tau < T|M = m)^n\left(1 - \mathbb{Q}(\tau < T|M = m)\right)^{N-n}$$

CDO: One-Factor approach: Finite Size, Homogeneous Portfolio

$$\mathbb{Q}\{n\,\text{defaults}|M=m\} = \binom{N}{n}\mathbb{Q}(\tau<T|M=m)^n\,(1-\mathbb{Q}(\tau<T|M=m))^{N-n}$$

In order to compute the unconditional probability to have n defaults among the N obligors we just integrate the previous expression w.r. to the common factor M

$$\mathbb{Q}\{n\,\text{defaults}\} = \int_{-\infty}^{+\infty}\mathbb{Q}\{n\,\text{defaults}|M=m\}\varphi(m)dm$$

In general this integral has to be computed numerically: This is the reason why the factor approach is called **semi-analytical**. Indeed we have closed formulas for the number of defaults but we need to solve integrals numerically.

In the homogeneous portfolio framework, the computation of the associated loss distribution is straightforward. Given a constant recovery rate REC, for the underlying credits in the portfolio, the probability of having a percentage portfolio loss $L = n\cdot(1-\text{REC})$ given by n names defaulting in the portfolio is given by

$$\mathbb{Q}\{\text{Loss}=L\} = \mathbb{Q}\{n\,\text{defaults}\}$$

CDO: One-Factor approach: Infinite (Large) Size, Homogeneous Portfolio

If the numbers of reference entities in the portfolio becomes reasonably high, the formulation of the loss probability simplifies further.

If the portfolio is relatively large, conditional on M the fraction of credits of the underlying portfolio defaulting over a specific time horizon should be roughly equal to the common individual default probability of the underlying credits for that horizon (assumed equal across names due to homogeneity).

This is a consequence of a probabilistic result known as the Law of Large Numbers (which in turn can be derived as a consequence of the Central Limit Theorem). Let us see this in more detail.

CDO: JPMorgan Approach: The Large Pool Model

Let us introduce the definition of Clean Spread γ_{clean} for the pool, taking the itraxx as fundamental example:

$$\gamma_{clean} = \frac{R_T^{\text{i-Traxx}}}{1 - \text{Rec}} = \frac{R_T^{\text{i-Traxx}}}{\text{Lgd}}$$

($R_{0,T}^{\text{i-Traxx}}$ (bps) is the quoted premium rate (spread) of the CDS index for maturity T).

Justification of this formula has been given in the single name case.

If we denote by T the maturity of the contract, then the average cumulative default probability of the portfolio can be approximated by:

$$\text{PD}(T) = 1 - \exp\left(-\frac{\gamma_{clean}}{10000}T\right)$$

as in standard constant intensity model. The factor 10000 is due to expressing γ in bps.

CDO. JPMorgan Approach: The Large Pool Model

Conditional on the systemic factor M defaults are independent and the probability of one default is, as we have seen many times, the same for all names and given by:

$$\mathsf{PD}_1(M;\rho) = \Phi\left(\frac{\Phi^{-1}\left(\mathsf{PD}(T)\right) - \sqrt{\rho}M}{\sqrt{1-\rho}}\right)$$

Here the Large Pool assumption comes into play: we can assume the number of names going to infinity and check what happens to the default fraction in the pool.

CDO. JPMorgan Approach: The Large Pool Model

The default fraction or rate (DR) of the pool at a given time T is the number of defaulted names over the total number of names in the pool, and is what matters for a CDO payoff. We may write, conditional on M:

$$DR_T^N(M) = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\tau_i \le T | M\}}$$

But given that, conditional on M, $1_{\{\tau_i \le T|M\}}$ are i.i.d. with mean $\mathrm{PD}_1(M;\rho)$, the law of large numbers tells us that

$$DR_T^N(M) \to \mathrm{PD}_1(M;\rho) \quad \text{as} \quad N \to \infty$$

Since for equal recovery across names the loss is simply $(1 - \textsc{Rec})$ times the default fraction, the loss under infinite pool in our case is simply

$$\textsc{Loss}_T^{\infty}(M;\rho) = (1 - \textsc{Rec})\mathrm{PD}_1(M;\rho)$$

CDO. JPMorgan Approach: The Large Pool Model

If we consider a $[0, B]$ tranche we have that the **tranched loss** (conditional on M) is $\min(\mathsf{Loss}_T(M), B)$. JPMorgan computes the **expected tranched loss** by using the above result on the loss under large pool, so that, conditional on M, there is no need to take expectation, due to the simplification obtained through the law of large numbers:

$$\mathbb{E}[\mathsf{Loss}_{0,B}^{tr}](M,\rho) = \mathsf{Loss}_{0,B}^{tr,\infty}(M,\rho) := \frac{1}{B}\min\left(\mathsf{Loss}_T^{\infty}(M;\rho), B\right)$$

(notice that no expectation is acting really), and, unconditionally,

$$\mathbb{E}[\mathsf{Loss}_{0,B}^{tr}](\rho) = \mathsf{Loss}_{0,B}^{tr,\infty}(\rho) := \int \mathsf{Loss}_{0,B}^{tr,\infty}(m,\rho)\varphi(m)dm$$

CDO. JPMorgan Approach: The Large Pool Model

In particular, substituting all expressions and summing up, we find $\text{Loss}_{0,B}^{tr,\infty}(\rho) =$

$$= \boxed{\int \frac{1}{B} \min\left(\Phi\left(\frac{\Phi^{-1}\left(1 - \exp\left(\frac{-R_T^{\text{i-Traxx}}\ T}{\text{Lgd}\ 10000}\right)\right) - \sqrt{\rho} m}{\sqrt{1-\rho}} \right)(1 - \text{Rec}), B \right) \varphi(m) dm}$$

We can value this in closed form, obtaining

$$= \Phi(A_1) + \frac{\text{Lgd}}{B} \Phi_2 \left(-A_1, \Phi^{-1}\left(1 - \exp\left(\frac{-R_T^{\text{i-Traxx}}\ T}{\text{Lgd}\ 10000}\right)\right), -\sqrt{\rho} \right)$$

$$A_1 = \frac{1}{\sqrt{\rho}} \left[\Phi^{-1}\left(1 - \exp\left(\frac{-R_T^{\text{i-Traxx}}\ T}{\text{Lgd}\ 10000}\right)\right) - \sqrt{1-\rho}\ \Phi^{-1}(B/\text{Lgd}) \right]$$

and where $\Phi_2(\cdot, \cdot, c)$ is the bivariate standard normal cumulative distribution function with correlation c.

CDO. JPMorgan Approach: The Large Pool Model

For a generic tranche A, B, we have through the above formula

$$\mathsf{Loss}^{tr,\infty}_{A,B}(\rho_A, \rho_B) = \frac{1}{B-A}\left[B \ \mathsf{Loss}^{tr,\infty}_{0,B}(\rho_B) - A \ \mathsf{Loss}^{tr,\infty}_{0,A}(\rho_A)\right]$$

where we remember that we are computing everything at maturity T.

If we removed the infinite pool assumption we would need to put EXPECTED tranche losses in the payout.

Consistency of the model would impose $\rho_A = \rho_B$. We keep the possible inconsistency in view of implied correlation.

CDO. JPMorgan Approach: The Large Pool Model

Now, let us compute the outstanding tranche notional

$$\mathsf{ON}^{tr,\infty}_{A,B}(T,\rho_A,\rho_B) = \Big(1 - \mathsf{Loss}^{tr,\infty}_{A,B}(\rho_A,\rho_B)\Big)$$

Here too we do not need to take expectations since randomness has been ruled out by the law of large numbers in the basic loss. Since the payments are quarterly made, we can define a survival rate in the following way (T is expressed in years, annual compounding):

$$\mathsf{ON}^{tr,\infty}_{A,B}(T,\rho_A,\rho_B) = \left(1 + \frac{\mathsf{SurvivalRate}(\rho_A,\rho_B)}{4}\right)^{-4T}$$

CDO. JPMorgan Approach: The Large Pool Model

We can then extend, using this rate as pivot, the outstanding notional (as if it were a bond price) at all times t by

$$\mathsf{ON}^{tr,\infty}_{A,B}(t,\rho_A,\rho_B) = \left(1 + \frac{\mathsf{SurvivalRate}(\rho_A,\rho_B)}{4}\right)^{-4t}$$

Now assume, as JPMorgan does, that interest rates are null, or, equivalently, $P(0,t) = D(0,t) = 1$ for all t. Then we can write the tranche default leg price as

$$\mathsf{DefaultLeg}^{A,B}_{a,b}(0)(\rho_A,\rho_B) = \sum_{i=a+1}^{b} \left(\mathsf{ON}^{tr,\infty}_{A,B}(T_{i-1},\rho_A,\rho_B) - \mathsf{ON}^{tr,\infty}_{A,B}(T_i,\rho_A,\rho_B)\right)$$

and the tranche premium leg for unit spread as

$$\mathsf{PremiumLeg1}^{A,B}_{a,b}(0)(\rho_A,\rho_B) = \sum_{i=a+1}^{b} \alpha_i \mathsf{ON}^{tr,\infty}_{A,B}(T_i,\rho_A,\rho_B)$$

CDO. JPMorgan Approach: The Large Pool Model

In this setup the fair tranche spread $R_{0,T_b}^{A,B}$ paid in the premium leg that balances the default leg solves

$$R_{0,T_b}^{A,B}\ \mathsf{PremiumLeg1}_{0,b}^{A,B}(0)(\rho_A,\rho_B) = \mathsf{DefaultLeg}_{0,b}^{A,B}(0)(\rho_A,\rho_B)$$

from which

$$R_{0,T_b}^{A,B} = \frac{\mathsf{DefaultLeg}_{0,b}^{A,B}(0)(\rho_A,\rho_B)}{\mathsf{PremiumLeg1}_{0,b}^{A,B}(0)(\rho_A,\rho_B)}$$

This approach will be fundamental in defining **implied correlation**.

CDO. One-Factor approach: Large Homogeneous Portfolio

This approach has some limits:

- It can be used only for very large portfolios (typically basket of 100 names or more). This is the case for CDO's but not for k-th to default swaps
- The model cannot produce consistent probabilities for low number of defaults (zero or one). This is due to the fact that the model does not consider the absolute number of defaults but only the (continuous) fraction of the defaulted portfolio

CDO. One-Factor approach: Finite, Not Homogeneous Portfolio

This is the most typical case in the CDO world, but also the most difficult. As usual we need to compute the portfolio loss distribution. But here we have different recoveries REC_i (leading to different additive losses $\text{LGD}_i = 1 - \text{REC}_i$) and different default probabilities for the underlying names.

Find a way to compute the **conditional** on M portfolio loss distribution, so that we can obtain the **unconditional** distribution just integrating on the common factor.

There are three main approaches to compute the conditional distribution:

- Fast Fourier Transform Approach
- Recurrence Relation Approach
- Probability Shifting Approach

Loss Calculation: Factor Copula without MC. Fast Fourier Transform

We need to compute the $\text{Loss}(t)$ distribution at each time t, where

$$\text{Loss}(t) = \sum_{i=1}^{n} \text{Loss}_i \mathbf{1}_{\{\tau_i < t\}}$$

If the names are equally weighted in the portfolio we can substitute Loss_i with $\text{LGD}_i = 1 - \text{Rec}_i$. Computing the conditional version

$$\text{Loss}_m(t) = \sum_{i=1}^{n} \text{Loss}_i \mathbf{1}_{\{\tau_i < t | M = m\}} = \sum_{i=1}^{n} L_{i,m}(t)$$

where $L_{i,m} = \text{Loss}_i \mathbf{1}_{\{\tau_i < t | M = m\}}$ are independent variables. We know that the characteristic function φ of a sum of independent variables is given by the product of single characteristic functions (convolution product in the space of densities).

$$\varphi_{\text{Loss}_m}(u) = \mathbb{E}[e^{iu\text{Loss}_m}] = \mathbb{E}[e^{iu\sum_j L_{j,m}}] = \mathbb{E}[\prod_j e^{iuL_{j,m}}] = \prod_j \mathbb{E}[e^{iuL_{j,m}}] = \prod_j \varphi_{L_{j,m}}(u)$$

Loss Calculation: Factor Copula without MC. Fast Fourier Transform

$$\varphi_{\mathsf{Loss}_m}(u) = \prod_j \varphi_{L_{j,m}}(u) \iff \mathsf{DFT}(p_{\mathsf{Loss}_m}) = \mathsf{DFT}(p_{L_{1,m}}) \cdots \mathsf{DFT}(p_{L_{n,m}}),$$

since in general the characteristic function $\varphi_{\mathsf{Loss}^m}(u)$ of a (discrete) random variable is a (discrete) Fourier transform DFT of the density of the random variable Loss_m, and the same holds for single loss terms $\varphi_{L_{j,m}}(u)$. With the inverse discrete Fourier transform IDFT we can compute the loss density as

$$p_{\mathsf{Loss}_m} = \mathsf{IDFT}(\mathsf{DFT}(p_{L_{1,m}}) \cdots \mathsf{DFT}(p_{L_{n,m}}))$$

The Fast Fourier Transform is a particular method to efficiently compute the DFT.

The method is explained in great detail in Robertson, J. P., (1992). The Computation of Aggregate Loss Distributions

`http://www.defaultrisk.com/pp_related_01.htm`

Loss Calculation: Factor Copula without MC. Recurrence Relation

This approach is quite simple but it can be applied only when the recovery rates and the notionals are the same for each underlying name (see Hull and White, 2004).

The different realizations of the loss are just LGD times the number of defaults and we need the probability to have that specific number of defaults.

Here we do not assume default prob or correlation across pairs to be equal.

Since conditional on M the defaults are independent, it is possible to find a recurrence formula for the corresponding probabilities. The prob to have zero defaults at t is:

$$\pi_t(0|M) = \prod_{i=1}^{N} \mathbb{Q}(\tau_i > t|M)$$

We have seen earlier how to compute the conditional survival probability in a factor model

$$\mathbb{Q}(\tau_i > T|M = m) = 1 - \Phi\left(\frac{\Phi^{-1}(1 - \exp(-\Gamma_i(T))) - \sqrt{\rho_i}m}{\sqrt{1-\rho_i}}\right)$$

Loss Calculation: Factor Copula without MC. Recurrence Relation

Then it is easy to prove that the probability to have exactly one default is (either defaults name 1 and names $2, \ldots, N$ survive, or defaults name 2 and names $1, 3, \ldots, N$ survive, and these two events are disjoint, and so on):

$$\pi_t(1|M) = \pi_t(0|M) \sum_{i=1}^{N} \frac{1 - \mathbb{Q}(\tau_i > t|M)}{\mathbb{Q}(\tau_i > t|M)}$$

Applying the same reasoning and defining

$$w_i = \frac{1 - \mathbb{Q}(\tau_i > t|M)}{\mathbb{Q}(\tau_i > t|M)}$$

we have:

$$\pi_t(k|M) = \pi_t(0|M) \sum w_{p(1)} w_{p(2)} \ldots w_{p(k)}$$

where $\{p(1), p(2), \ldots, p(k)\}$ is a set of k numbers among $\{1, 2, \ldots, N\}$ and the summation is taken over all the possible $\frac{N!}{(N-k)!k!}$ combinations.

Loss Calculation: Factor Copula without MC. Recurrence Relation

Let us define $P_k = \sum w_{p(1)} w_{p(2)} \dots w_{p(k)}$ and $A_k = \sum_{i=1}^{N} w_i^k$; then, there is a Recurrence Formula saying

$$\begin{aligned} P_1 &= A_1 \\ 2P_2 &= A_1 P_1 - A_2 \\ 3P_3 &= A_1 P_2 - A_2 P_1 + A_3 \\ &\vdots \\ kP_k &= A_1 P_{k-1} - A_2 P_{k-2} + A_3 P_{k-3} - \dots + (-1)^{k+1} A_k \end{aligned}$$

the A's are easily computed based on conditional probabilities on M given earlier. The scheme then give the P's that, in turn, give the probability $\pi(k|M)$ of having k defaults and hence the loss if all the LGD and notionals are equal.

This method results to be computationally very fast, but has the limitation that it works only under strict requirements on the recoveries and the notionals.

Loss Calculation: Factor Copula without MC. Probability Shifting

A detailed description of this method can be found in Hull and White (2004) or in Andersen et al. (2003).

The idea is to divide the spectrum of possible losses into buckets of small size and to compute, conditional on M, the probability to have the loss inside one of these buckets.

The probability distribution is computed by recurrence: if we have the probability distribution of $n-1$ names, we can update the distribution when adding the n-th name.

In detail: Let us divide the possible loss range $[0, \text{Loss}^{MAX}]$ into buckets $\{0, b^0\}, \{b^0, b^1\}, \ldots, \{b^{K-1}, b^K\}$. We denote by $b_j = \{b^{j-1}, b^j\}$ the j-th bucket.

Then we indicate with p_j^n the probability that the loss is in the bucket b_j when n names are present. We assume that the loss is concentrated in the middle of the interval, so that if the loss is in the bucket j, we set $L_j = \frac{b^{j-1}+b^j}{2}$.

Loss Calculation: Factor Copula without MC. Probability Shifting

As usual we compute conditional on M (independence) and then integrate.

At the beginning we imagine to have a portfolio of zero names, so $p^0_j = 0$ for all j.

Now add a particular credit reference with default probability d_1, survival $1 - d_1$ and recovery REC_1. Then we update the probability distribution of the loss for a portfolio with one name. Obviously $p^1_0 = 1 - d_1$ and then there is a particular bucket containing $L_0 + \text{LGD}_1$ having probability d_1.

Loss Calculation: Factor Copula without MC. Probability Shifting

In general if we have all the series of the p_j^{n-1}'s with $n-1$ names considered, we can compute the updated p_j^n's when adding the $n-th$ name in the following way

$$p_{u(j)}^n = p_j^{n-1} \cdot d_n + p_{u(j)}^{n-1} \cdot (1-d_n)$$

where $u(j)$ indicates the bucket containing the loss of n names if the loss of $n-1$ names was in the bucket j.

This method has two main advantages: (i) Works well also for different notionals and recoveries, (ii) the width of the buckets is arbitrary.

Single Tranche CDO Pricing

Once we have the distribution of the portfolio loss, obtained by one of the methods above, we can compute the price of each tranche of a CDO.

As before (with the large pool model) we can extend tranched loss knowledge at time T_b to earlier loss knowledge by defining a survival rate for the expected outstanding tranche notional: Solve

$$1 - \mathbb{E}[\text{Loss}^{tr}_{A,B}(T)] = \left(1 + \frac{\text{SurvivalRate}}{4}\right)^{-4T}$$

in SurvivalRate and assume

$$1 - \mathbb{E}[\text{Loss}^{tr}_{A,B}(t)] = \left(1 + \frac{\text{SurvivalRate}}{4}\right)^{-4t}$$

Or else, to be more precise, we should apply the above methods for different maturities, to get the expected loss at different t's.

Single Tranche CDO Pricing

As we have already seen, a CDO tranche contract is made up by two legs of payments, the default leg

$$\text{PriceDefLeg}_{A,B}(0) = \int_0^T P(0,t)\mathbb{E}[d\text{Loss}^{tr}_{A,B}(t)]$$

and the premium leg

$$\text{PricePrLeg}(0) = R^{A,B}_{0,T}(0) \ \text{PricePrLeg1}(0),$$

$$\text{PricePrLeg1}(0) := \sum_{i=1}^{b} P(0,T_i)\alpha_i(1 - \mathbb{E}[\text{Loss}^{tr}_{A,B}(T_i)])$$

Still using the Factor copula, instead of resorting to one of the above approximations (large pool, probability shifting, recurrence relation) one can go for a Monte Carlo simulation, that makes no approximation on the loss evolution in time but computes it exactly. This can be numerically more intensive though.

Single Tranche CDO Pricing

We can compute the Tranche fair spread as

$$R_{0,T_b}^{A,B}(0) = \frac{\mathbb{E}[\int_0^{T_b} P(0,t) d\mathsf{Loss}_{A,B}^{tr}(t)]}{\mathbb{E}[\sum_{i=1}^{b} P(0,T_i)\alpha_i(1 - \mathsf{Loss}_{A,B}^{tr}(T_i))]}$$

If we use Monte Carlo, typically the numerator term inside the expectation has larger variance than the corresponding term in the denominator. The standard error can be computed only for the numerator assuming the denominator expectation to be exact. Otherwise, a conservative window built on standard errors for the numerator and the denominator can be built.

We typically use 200.000 paths without variance reduction techniques. In C++ it is a matter of few seconds.

The approximated methods can be helpful here. Particularly quick and precise is the probability shifting approach, although it gets slow if we apply it for different maturities to have a more precise evolution of the loss.

Implied Correlation

Up to now we have discussed how to compute the loss distribution starting from a given value for the Gaussian copula correlation, always neglecting how we can find this value from market quotes.

In market practice the default correlation is extracted implicitly from quotes of very liquid multi-name products such as the CDO index tranches. The correlation obtained in this way is called **implied correlation**.

Dealers provide quotes of the premium $R_{0,T_b}^{A,B}(0)$ paid for each tranche. Usually a Gaussian copula (standard or One Factor) is used for the pricing.

A single correlation value among all pairs of names in the portfolio is used in the Gaussian copula, that is the market quote **does not diversify across sector, country, etc, but assesses a sort of "average" correlation.**

Implied Correlation

In the tranche prices formulas write the losses as

$$\text{PriceDefLeg}_{A,B}(0) = \int_0^T P(0,t)d[B\ \mathbb{E}[\text{Loss}^{tr}_{0,B}(t)] - A\ \mathbb{E}[\text{Loss}^{tr}_{0,A}(t)]]/(B-A)]$$

$$= \underbrace{\frac{B}{B-A}}_{\beta}\text{PriceDefLeg}_{0,B}(0) - \underbrace{\frac{A}{B-A}}_{\alpha}\text{PriceDefLeg}_{0,A}(0)$$

and similarly the premium leg per unit spread

$$\text{PricePrLeg1}_{A,B}(0) = \sum_{i=1}^{b} P(0,T_i)\alpha_i(1-[B\ \mathbb{E}[\text{Loss}^{tr}_{0,B}(T_i)]-A\ \mathbb{E}[\text{Loss}^{tr}_{0,A}(T_i)]]/(B-A))$$

$$= \beta\text{PricePrLeg1}_{0,B}(0) - \alpha\text{PricePrLeg1}_{0,A}(0)$$

Implied Correlation

Now assume we put a Gaussian Factor copula with the same correlation parameter ρ_A in all loss calculations for the tranche $[0, A]$ and ρ_B for all loss calculations for $[0, B]$.

$$\text{PriceDefLeg}_{A,B}(0, \rho_A, \rho_B) = \beta \text{PriceDefLeg}_{0,B}(0, \rho_B) - \alpha \text{PriceDefLeg}_{0,A}(0, \rho_A)$$

$$\text{PricePrLeg1}_{A,B}(0, \rho_A, \rho_B) = \beta \text{PricePrLeg1}_{0,B}(0, \rho_B) - \alpha \text{PricePrLeg1}_{0,A}(0, \rho_A)$$

To obtain the implied correlation for a given maturity T_b, we insert the market quoted tranche spread $R^{A,B}_{0,T_b}(0)$ in the premium leg and find correlation parameters such that the two legs match:

$$\text{PriceDefLeg}_{A,B}(0, \rho_A, \rho_B) = R^{A,B\ \text{Mkt}}_{0,T_b}\ \text{PricePrLeg1}_{A,B}(0, \rho_A, \rho_B)$$

Implied Correlation

Consider the i-traxx tranches for example, to clarify the procedure. For a given maturity in 3y, 5y, 7y, 10y the market quotes

$$\text{Upfront}^{0,3\%\ \text{Mkt}} + 500bps \text{ running},$$

$$R^{3,6\%\ \text{Mkt}}, R^{6,9\%\ \text{Mkt}}, R^{9,12\%\ \text{Mkt}}, R^{12,22\%\ \text{Mkt}}$$

To obtain the implied correlation we proceed as follows.

Implied Correlation

First solve in $\rho_{3\%}$ for the equity tranche (this should be done with the upfront but here we assume we have converted it into an equivalent spread)

$$\text{PriceDefLeg}_{0,3\%}(0,\rho_3) = 500bps\ \text{PricePrLeg1}_{0,3}(0,\rho_3) + \text{Upfront}^{0,3\%\ \text{Mkt}}$$

The move on: now you have to choices: solve

$$\text{PriceDefLeg}_{3,6}(0,\rho_3,\rho_6) = R^{3,6\%\ \text{Mkt}}\ \ \text{PricePrLeg1}_{3,6}(0,\rho_3,\rho_6)$$

in ρ_6 (base correlation) or solve

$$\text{PriceDefLeg}_{3,6}(0,\bar\rho_{3,6},\bar\rho_{3,6}) = R^{3,6\%\ \text{Mkt}}\ \ \text{PricePrLeg1}_{3,6}(0,\bar\rho_{3,6},\bar\rho_{3,6})$$

in $\bar\rho_{3,6}$ (compound correlation).

Implied Correlation: Base VS Compound

$$\text{PriceDefLeg}_{3,6}(0, \rho_3, \rho_6) = R^{3,6\% \text{ Mkt}} \ \text{PricePrLeg1}_{3,6}(0, \rho_3, \rho_6)$$

solve in ρ_6 (base correlation), or solve

$$\text{PriceDefLeg}_{3,6}(0, \bar{\rho}_{3,6}, \bar{\rho}_{3,6}) = R^{3,6\% \text{ Mkt}} \ \text{PricePrLeg1}_{3,6}(0, \bar{\rho}_{3,6}, \bar{\rho}_{3,6})$$

in $\bar{\rho}_{3,6}$(compound correlation).

Compound correlation is more consistent at the level of single tranche: We value the whole payoff of the tranche premium and default legs with one single copula (model) with parameter $\bar{\rho}_{3,6}$.

Base correlation is inconsistent at the level of single tranche: we value different parts of the same payoff with different models, i.e. part of the payoff (involving $\text{Loss}_{0,3}$) is valued with a copula in ρ_3, while a different part (involving $\text{Loss}_{0,6}$) of **the same** payoff is valued with a copula in ρ_6.

Implied Correlation: Base VS Compound

The procedure goes on similarly:

$$\text{PriceDefLeg}_{6,9}(0, \rho_6, \rho_9) = R^{6,9\%\ \mathrm{Mkt}}\ \text{PricePrLeg1}_{6,9}(0, \rho_6, \rho_9)$$

solve in ρ_9 (base correlation), or solve

$$\text{PriceDefLeg}_{6,9}(0, \bar{\rho}_{6,9}, \bar{\rho}_{6,9}) = R^{6,9\%\ \mathrm{Mkt}}\ \text{PricePrLeg1}_{6,9}(0, \bar{\rho}_{6,9}, \bar{\rho}_{6,9})$$

in $\bar{\rho}_{6,9}$ (compound correlation).

And so on.

Implied Correlation: Model or simply Quoting Mechanism?

If the Gaussian copula assumptions were consistent with market tranche prices, there should be a unique Gaussian copula model consistent with the market.

In other terms all values $\rho_3, \rho_6, \rho_9, \ldots$ (base correlation) should be equal.

Or is we resort to compound correlation all values $\bar{\rho}_{0,3} = \rho_3, \bar{\rho}_{3,6}, \bar{\rho}_{6,9}\ldots$ should be equal to each other.

As we are going to see this does not happen: plotting base correlations

$$\rho_3, \rho_6, \rho_9, \ldots$$

we obtain a **correlation skew**, while when plotting compound correlations

$$\bar{\rho}_{0,3}, \bar{\rho}_{3,6}, \bar{\rho}_{6,9}\cdots$$

we obtain a **correlation smile**.

Implied (Compound) Correlation

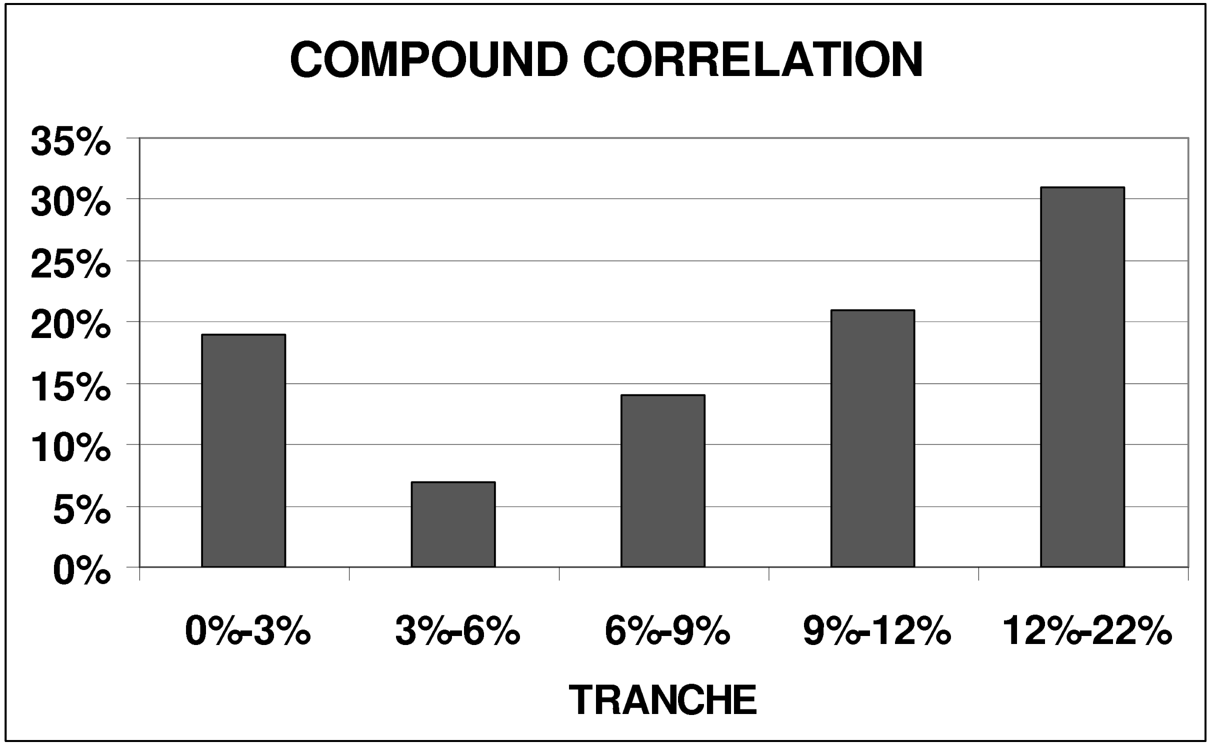


Figure 24: Example of compound correlation structure $\bar{\rho}$ for the DJ iTraxx.

Implied (Compound) Correlation

The value of correlation corresponding to the 1st mezzanine tranche (3%-6%) is lower than the one corresponding to the equity tranche, and after that the correlation grows.

The reason of this behavior must be sought for in the shape of the loss distribution leading to different investment considerations.

Consider the senior tranches: To reach the attachment points we need high losses, but this means that many defaults have to occur and this results in high correlations.

To obtain significant spreads (justifying the investment) we need high correlations.

On the contrary, if we think of the equity tranche $0 - 3\%$ it is clear that to obtain significant spreads we need low correlations: this tranche is impacted by every default, while a large correlation would imply a low probability of a single default (otherwise we would have a large probability to have many defaults with high correlation and in general this is not the case).

Implied (Base) Correlation

Now, let us suppose that we need to price a bespoke tranche (i.e. a non standard tranche, for example having different attachment points, e.g. 4%-7%).

Which value of correlation should we use? Is there a way to extract the correct value of correlation from the compound correlation smile?

The answer is negative: The **values of the implied compound correlations are peculiar** of the **PAIRS of standard attachment points**.

To overcome the problem of a correlation value linked to a pair of points a notion of correlation depending on a single point is better. We use then *base correlation*.

The idea is the following: We use market prices for quoted tranches and then we consider only equity tranches with the same lower attachment point. For example the DJ iTraxx tranches are transformed in 0%-3%, 0%-6%, 0%-9%, 0%-12%, 0%-22%.

Implied (Base) Correlation

Now, suppose we are interested in pricing the 4%-7% tranche: First we compute the base correlation curve, then we simply interpolate the curve to obtain the values ρ_4 and ρ_7 at 4% and 7%. Now we can price the considered tranche as a difference between the 0%-7% and the 0%-4% equity tranches.

Market contributors usually provide implied values both for the base and the compound correlation. Notice that by construction the compound correlation and the base correlation relative to the equity tranches are the same.

The term compound correlation derives from the fact that the loss for a mezzanine tranche is

$$\text{Loss}^{tr}_{A,B} = \frac{1}{B-A}\left[(\text{Loss}-A)^+ - (\text{Loss}-B)^+\right],$$

i.e. it is given by the composition (sum) of two (call) options on the loss.

Implied (Base) Correlation

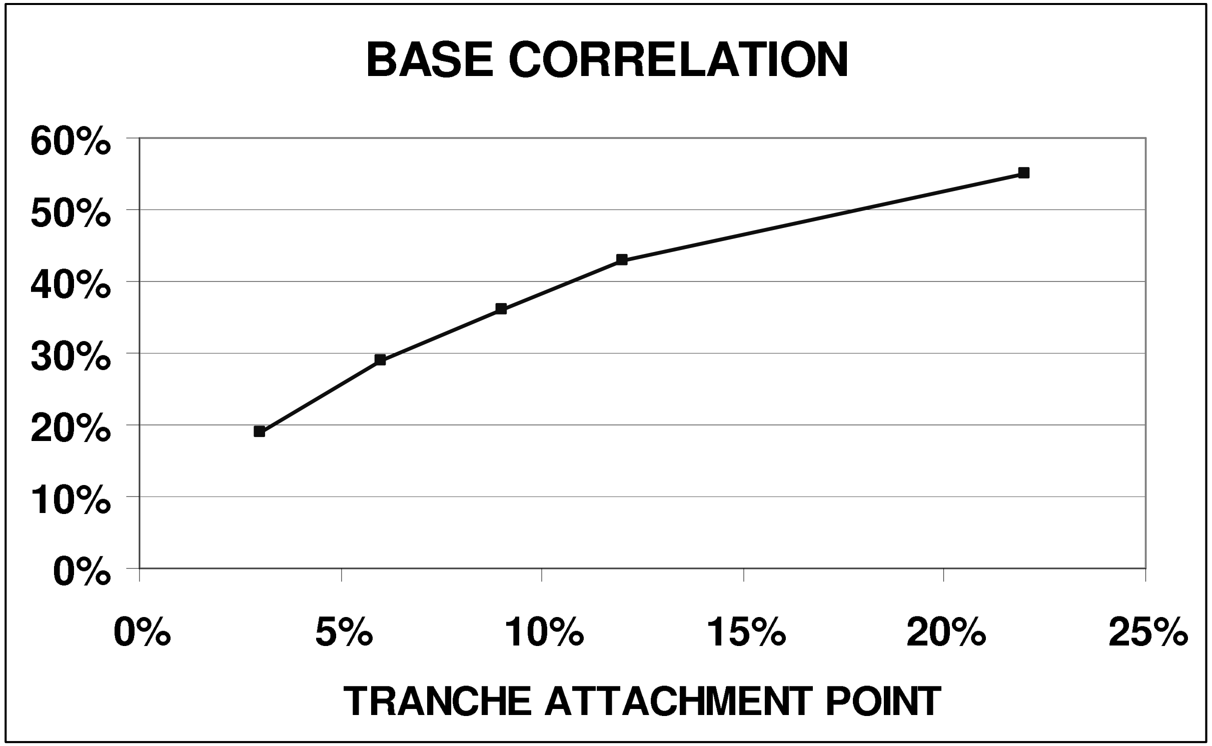


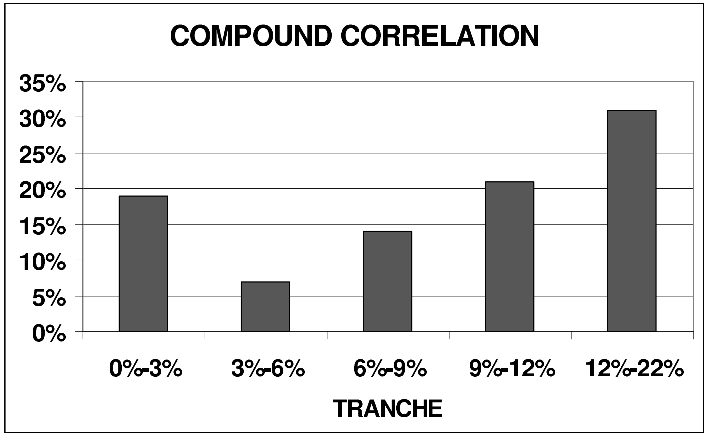
Figure 25: Example of base correlation structure for the DJ iTraxx.

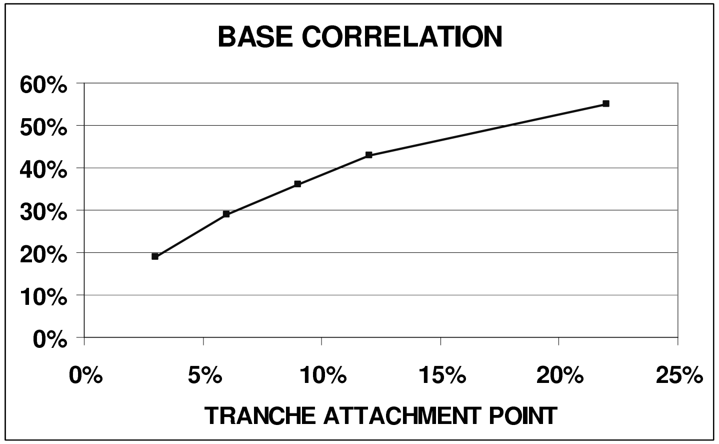
Implied (Base) Correlation

Notice two main differences:

First:The graph for the base correlation is a line while the graph for the compound is a collection of columns. This because the base correlation can be associated to a single (high) attachment point while the compound is associated to an entire tranche, so it would be incorrect to plot it as a line.

Second The values of base correlation are always increasing as a function on the attachment point, and are always greater than corresponding values of compound correlation.





Implied Correlation: Base VS Compound

$$\text{PriceDefLeg}_{3,6}(0, \rho_3, \rho_6) = R^{3,6\%\ \text{Mkt}}\ \text{PricePrLeg1}_{3,6}(0, \rho_3, \rho_6)$$

solve in ρ_6 (base correlation), or solve

$$\text{PriceDefLeg}_{3,6}(0, \bar{\rho}_{3,6}, \bar{\rho}_{3,6}) = R^{3,6\%\ \text{Mkt}}\ \text{PricePrLeg1}_{3,6}(0, \bar{\rho}_{3,6}, \bar{\rho}_{3,6})$$

in $\bar{\rho}_{3,6}$(compound correlation)

Compound correlation at times does not exist: the second equation may have no solution in some market conditions.

In these cases base correlation still exists: the fact that we (inconsistently) have two different ρ's in different parts of the payoff gives more freedom to the possible solutions of the first equation.

In these cases, however, the base correlation implies **negative expected tranche LOSS realizations** and is thus arbitrageable in principle. This happens when the base corr skew is very steep.

Implied Correlation: Base VS Compound

FEATURE ARTICLE 1

IMPLIED CORRELATION IN CDO TRANCHES:

A Paradigm to be Handled With Care

Stripping Correlations: MC vs Large Pool

We have seen how one strips implied correlations. However, in solving e.g.

$$\text{PriceDefLeg}_{3,6}(0, \rho_3, \rho_6) = R^{3,6\%\ \text{Mkt}}\ \text{PricePrLeg1}_{3,6}(0, \rho_3, \rho_6)$$

solve in ρ_6 (base correlation), or solve

$$\text{PriceDefLeg}_{3,6}(0, \bar{\rho}_{3,6}, \bar{\rho}_{3,6}) = R^{3,6\%\ \text{Mkt}}\ \text{PricePrLeg1}_{3,6}(0, \bar{\rho}_{3,6}, \bar{\rho}_{3,6})$$

in $\bar{\rho}_{3,6}$(compound correlation)

One can value the legs either by the large pool model assumption, by the Probability Shifting or Recurrence Relation method, or by full monte carlo valuation.

Stripping Correlations: MC vs Large Pool

Recall large pool model: all names have same default probabilities, recoveries, expected tranche loss is computed as tranched expected loss, zero interest rates etc.

The value of implied correlation changes considerably in the two cases

i) Monte carlo or Probability Shifting vs ii) large pool model (JPMorgan)

In the following we present two examples

Monte Carlo Approach

The simulation is performed using a standard Gaussian copula with a unique value of correlation. We simulate many a collection of default times and compute the expected values of the payment legs.

In this case we used the names of the DJ-iTraxx index: In particular we considered all the different spreads of the portfolio components (but a common recovery of 40%).

This method has a great disadvantage: It is quite heavy from a numerical viewpoint. In fact the stripping of correlations involves an iterative procedure for evaluating the tranche changing the correlation value: This means that at each round of the stripping (that is for each trial ρ) we need to run the Monte Carlo simulation, and this could be quite cumbersome.

Monte Carlo Approach

In the following Table we report an example of stripping of implied correlation starting from quoted tranches. In brackets we report the quoted correlation values.

Tranche	Spread	Base Correlation	Compound Correlation
0%-3%	24.125%	18.94%(19%)	18.94%(19%)
3%-6%	134	28.50%(29%)	4.83%(5%)
6%-9%	45	36.17%(37%)	13.32%(13%)
9%-12%	31	41.92%(43%)	21.98%(22%)
12%-22%	15.25	57.29%(55%)	30.44%(31%)

Table 21: Quoted DJ iTraxx tranche spreads on November 11th, 2004 by BNP Paribas. The maturity is March 20th, 2010 and the reference index spread is 37 bps. Both base correlation (third column) and compound correlation (forth column) are reported. The spread of the equity tranche is composed by a running spread of 500 bps plus an upfront premium: Here we report only this upfront part.

Monte Carlo Approach

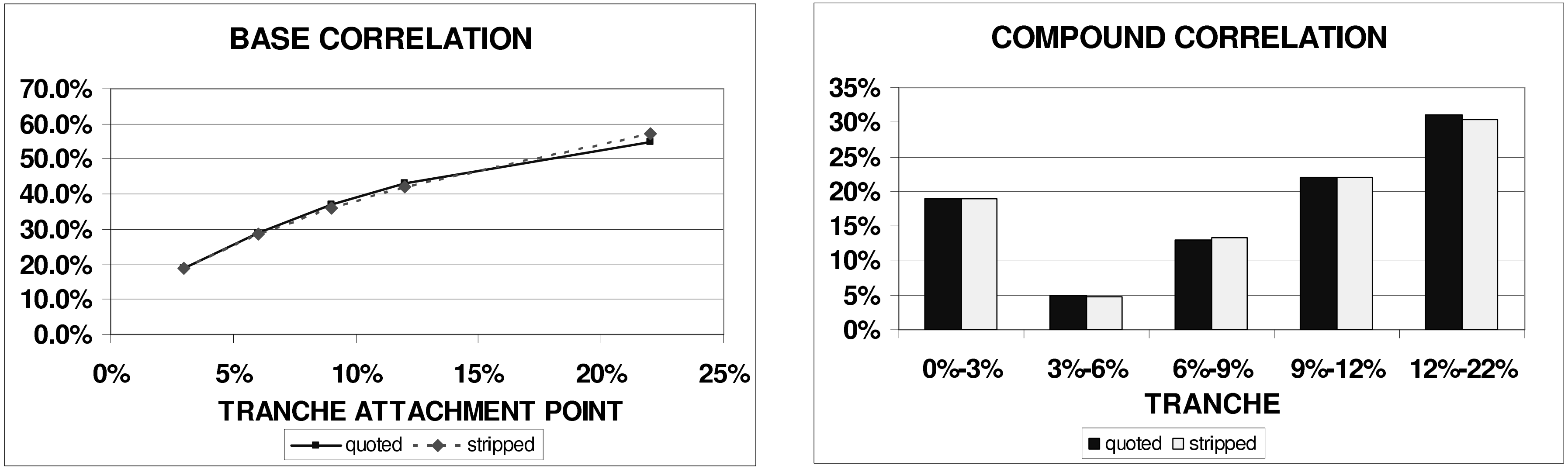


Figure 26: Comparison between quoted correlation curves and curves stripped with Monte Carlo. On the left we have the base correlation, on the right the compound correlation.

We used 100000 paths with no variance reduction technique, but 50.000 can often be sufficient.

Large Pool: JPMorgan Approach

The one-factor/large pool approach we have seen earlier is an alternative, and in particular JPMorgan has implemented and freely distributed a simplified version of it. For a detailed description of how JPM model works see also McGinty and Ahluwalia (2004). Initially to make implied correlation more objective JPMorgan used to assume interest rates all equal to 0, $P(t,T) = D(t,T) = 1$ for any t, T in premium and default legs.

Finally we remark the fact that they quote only base correlation: This is in line with the current practice to give importance only to base correlation (even if sometimes other dealers quote also the compound correlation).

One Factor Copula: JPMorgan Approach

We report the values of the base correlation stripped by JPM quotes of iTraxx tranches obtained implementing the Large Pool Model (in brackets we report the quoted values of correlation).

Tranche	Spread	Base Correlation Large Pool	Base Correlation MC
0%-3%	24.05%	25.9% (25.7%)	18.94%
3%-6%	134	35.5% (35.3%)	28.5 %
6%-9%	47	43.4% (43.2%)	36.17%
9%-12%	31.5	49.1% (48.7%)	41.92%
12%-22%	155	64.3% (63.9%)	57.29%

Table 22: Results of the stripping with JPMorgan approach.

Notice the difference with respect to the implied base correlations we have derived earlier and compared with Paribas quotes for example (and by most other entities), obtained with a more refined method (e.g. Monte Carlo) and without large pool assumptions.

One Factor Copula: JPMorgan Approach

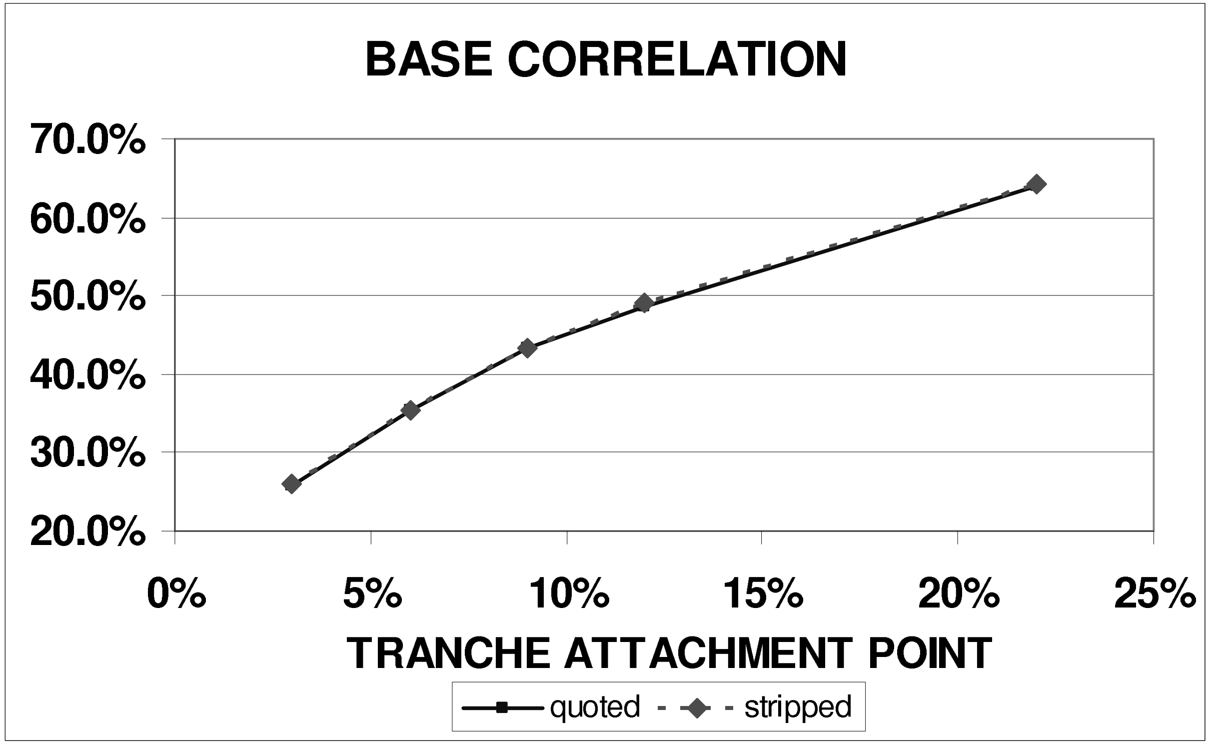


Figure 27: Comparison between quoted correlation curves and curves stripped with the JPMorgan approach.

One Factor Copula: JPMorgan Approach

The current debate is JPMorgan vs rest of the world. JPMorgan's method, with all its simplifications (especially the one identifying the conditional expected tranched loss with the conditional tranched expected loss), brings in a considerable difference with a more precise correlation stripping. The other entities use the more refined monte carlo or one factor approach. Strictly speaking then the refined method is more reliable.

JPMorgan answers to this by saying that they are not really providing the market with an estimation technique but merely with a quoting mechanism. And in order for the quoting to be uniquely determined, they do away with interest rates, different credit spreads etc. This way even two entities that disagree say on the zero curve for interest rates will agree on base correlation. Thus JPMorgan implied correlation is only based on inputs from the liquid credit market indices on which all parties agree

One Factor Copula: JPMorgan Approach

However, this answer is only partly satisfactory. What is indeed the point of giving a quoting mechanism if the quote cannot be interpreted in some intuitive way? And here the interpretation is that of a sort of "average" correlation parameter across names, countries, sectors etc. Now the more the model is precise, the more this parameter will resemble a correlation in the actual Gaussian copula.

Also, when pricing bespoke tranches with non-standard attachment points one directly refers to the base implied correlation curve, so that it has importance how this curve is built and whether one is being precise or not.

For example, if we interpolate large pool base correlations and then price a bespoke with a gaussian copula with those correlations but using Monte Carlo, we are not consistent.

CDO Squared with implied correlation

The best way to price the CDO squared (CDO^2) is the Monte Carlo simulation: In fact here no analytical or semi-analytical formulas are available. Moreover it is quite simple to take into account possible different characteristics of each underlying CDO.

Correlations are actually one of the most important issues in the pricing of the CDO^2. Because of the structure of the product, it is not difficult to imagine that correlation has a relevant role, since it strongly impacts the probability distribution of the losses.

CDO Squared using implied correlation

The problem is how to find the right correlations. One possibility is to proceed in the following way. Let us assume, for simplicity, that the underlying CDO tranches are all equal, for example with attachment points 5% and 7% (i.e. 2% of thickness).

Now let us consider one of the squared tranche, with attachment points a% and b%. Then the correlation corresponding to a% and b% is derived from the correlation of the underlying CDOs with attachment points $5\% + a\% \cdot 2\%$ and $5\% + b\% \cdot 2\%$.

Indeed, for the squared tranche to be affected, the loss in a basic CDO needs to arrive at 5%. To penetrate of $a\%$ in the squared loss, it is like to penetrate of $a\% * (7-5)$beyond 5% in the underlying average CDO. The maximum penetration in the squared tranche is when the whole $b\% * (7-5) = b\%2$ exceeds 5, leading to $5\% + b\% \cdot 2\%$.

$$[a\%, b\%] \longrightarrow [5\% + (7\% - 5\%)a, 5\% + b(7\% - 5\%)]$$

Squared CDO attach → basic CDO average attach

CDO Squared

$$[a\%, b\%] \quad \longrightarrow \quad [5\% + (7\% - 5\%)a, 5\% + b(7\% - 5\%)]$$

Squared CDO attach → basic CDO average attach

The correlations of these two attachment points can be directly interpolated from a CDS index, e.g. from i-Traxx tranche implied base correlations, that are the correct correlations to interpolate being based on a single attachment point.

These are not rigorous procedures and depend also on the diversification of the underlying portfolio. If this is not i-traxx or CDX we are in trouble.

CDO Squared

The ideal approach, however, would be calibrating a single copula with enough parameters (for example a mixture copula) to *all* the quoted index tranches available at the market at the same time. Then, with this single copula (and not with a Gaussian copula having each time a single different correlation parameter corresponding to a different tranche) we could price the CDO squared (and other credit derivatives) consistently with all available tranches correlation quotes.

This is under investigation.

CDO Squared

If we do not pursue this more rigorous approach, an assumption on the correlation has to be made, and stronger hypotheses can be needed. Anyway, once the correlation has been found, the pricing procedure is straightforward: put the correlation in the Gaussian copula and price the CDO squared tranche via Monte Carlo simulation, generating the joint $\tau_1 = \Gamma_1^{-1}(\xi_1), \ldots, \tau_N = \Gamma_N^{-1}(\xi_N)$, as seen many times in the lectures, by means of the calibrated Gaussian copula across the ξ's.

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach (Hull & White)

We give a summary of the method, trying to highlight its advantages and drawbacks.

A factor copula structure is assumed, similarly to the Gaussian 1 factor copula approach seen earlier, which we recall here:

$$\tau_1 = \Lambda_1^{-1}(-\ln(1-U_1)), \tau_2 = \Lambda_2^{-1}(-\ln(1-U_2)), \ldots, \tau_N = \Lambda_N^{-1}(-\ln(1-U_N)).$$

where we took $U_i = \Phi(X_i)$, with X_i's Gaussian variables defined as

$$X_i = \sqrt{\rho_i}\, M + \sqrt{1-\rho_i}\, Y_i,$$

with Y_i, M standard independent gaussian variables. We had

$$\mathbb{Q}(\tau_i < T|M) = \Phi\left(\frac{\Phi^{-1}(1-\exp(-\Gamma(T))) - \sqrt{\rho_i}M}{\sqrt{1-\rho_i}}\right)$$

with Γ the hazard function, assumed common for all single names (homogeneous pool)

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach

This time, instead:

- We do not model the copula explicitly, but
- we model default probabilities conditional on the systemic factor M of the copula.
- The copula will then be "hidden" inside these conditional probabilities.
- We assume a homogeneous model, in that the default probabilities of single names are all equal to each other.

Let us consider, for simplicity, survival (or equivalently default) probabilities that are associated to a constant-in-time hazard rate. We know that if we have a constant-in-time (possibly random) hazard rate λ for name i then the survival probability is

$$\mathbb{Q}(\tau_i > t) = \mathbb{E}[\exp(-\lambda t)]$$

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach

The implied copula approach postulates the following "scenario" distribution for the hazard rate λ conditional on the systemic factor M:

$$\lambda | M \sim \left\{ \begin{array}{ccc} \text{conditional hazard rate} & \text{Systemic scenario} & \text{Scenario probability} \\ \lambda_1 & M = m_1 & p_1 \\ \lambda_2 & M = m_2 & p_2 \\ \vdots & \vdots & \vdots \\ \lambda_s & M = m_s & p_s \end{array} \right.$$

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach

This way the default probability for a single name is, conditional on the systemic factor M,

$$\mathbb{Q}(\tau_i < t | M = m_j) = 1 - \exp(-\lambda_j t).$$

Compare with the Gaussian factor copula case:

$$\mathbb{Q}(\tau_i < T | M = m_j) = \Phi\left(\frac{\Phi^{-1}(1 - \exp(-\Gamma(T))) - \sqrt{\rho_i} m_j}{\sqrt{1 - \rho_i}}\right)$$

Unconditionally:

$$\mathbb{Q}(\tau_i < t) = \mathbb{E}[\mathbb{Q}(\tau_i < t | M)] = \sum_{j=1}^{s} p_j \mathbb{Q}(\tau_i < t | M = m_j) = \sum_{j=1}^{s} p_j (1 - \exp(-\lambda_j t))$$

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach

Conditional on M, all default times are independent, have the same hazard rate and their hazard rates are given by the above scenarios. Then to integrate against M we simply sum over all possible hazard rate scenarios multiplying by the scenario probability:

$$\text{TRANCHE}_{A,B}(0,R) = \sum_{j=1}^{s} p_j \text{TRANCHE}_{A,B}(0,R; \text{indep defaults across names with const haz rate } \lambda_j)$$

Each term on the right hand side can be easily computed according to the methods we have seen earlier (large pool model approximation, FFT, recurrence approach, Probability shifting approach, Monte Carlo...).

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach by Hull and White

Now we move to calibration: How can we calibrate say all the 5y DJ-i-Traxx tranches?

When the market gives us the tranche spread quote $R_{0,5y}^{A_i,B_i,\mathrm{MKT}}$ for two canonical attachments (e.g. $A = 3\%, B = 6\%$) we know that this should make the tranche net present value equal to zero. We might then try to solve

$$\sum_{j=1}^{s} p_j \mathrm{T}\text{RANCHE}_{A_i,B_i}(0, R_{0,5y}^{A_i,B_i,\mathrm{MKT}}; \text{indep defaults across names with const haz rate } \lambda_j) = 0$$

$$\text{for all } \ A_i, B_i.$$

In practice, we try and minimize the sum of squares of the above expression across the five canonical tranches:

$$\mathbf{p}^\star = \mathrm{argmin}_{p_1,\ldots,p_s} \sum_{i=1}^{5} \left[\sum_{j=1}^{s} p_j \mathrm{T}\text{RANCHE}_{A_i,B_i}(0, R_{0,5y}^{A_i,B_i,\mathrm{MKT}}; \text{ind defaults with const haz rate} \lambda_j) \right]^2$$

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach by Hull and White

$$\mathbf{p}^\star = \text{argmin}_{p_1,\ldots,p_s} \sum_{i=1}^{5} \left[\sum_{j=1}^{s} p_j \text{TRANCHE}_{A_i,B_i}(0, R_{0,5y}^{A_i,B_i,\text{MKT}}; \text{ind defaults with const haz rate } \lambda_j) \right]^2$$

Indeed, this is the approach suggested by Hull and White: calibrate the scenario probabilities p while pre-assigning the haz rate scenarios λ_j exogenously.

One may also calibrate the index itself besides the tranches, by adding the squared difference between the model implied index and the market index in the target function. The model implied index is easily computed by recalling that

$$\mathbb{Q}(\tau_i > t) = \sum_{j=1}^{s} p_j \exp(-\lambda_j t).$$

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach by Hull and White

One may naively think that since we have 6 quotes (5 tranches plus the index) then six parameters p could be enough to fit the market quotes, getting a low squared error out of the minimization. So one could set the number of scenarios $s = 6$.

Actually, this does not work. One may have to go up to $s = 30$ in order to be able to fit market data with a good precision.

We tried the model and checked that indeed, even with 27 points the error remains large, no matter how smartly one pre-selects the $\lambda_1, \ldots, \lambda_{27}$.

However, even with $s = 30$ one more problem surfaces. If we do not impose some further constraint or smoothing on the minimization problem, we are finding multiple solutions each time we change the initial guess.

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach by Hull and White

Hull and White propose to add to the target function a quantity that penalizes changes in convexity in the patterns of the scenario probabilities plotted against the default probabilities associated to each scenario:

$$\text{add } \ c \sum_{j=2}^{s-1} \frac{(p_{j+1} + p_{j-1} - 2p_j)^2}{0.5[\exp(-\lambda_{j-1} 5y) - \exp(-\lambda_{j+1} 5y)]} \ \text{ to the target function}$$

This term's numerators are second order differentials of the p's, and they penalize departures from convexity. Therefore, when we plot the graph $\{(i, p_i), i = 1, \ldots, s\}$ all kinds of humps are penalized. The constant c controls the degree of smoothing: too large a c privileges the smoothing over the fit, and we cannot recover exactly the market quotes. Too small a c fails to guarantee substantial uniqueness in the calibration solution. We found good results with $c = 1$.

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach by Hull and White

As a final remark, we mention a relationship between default probabilities and recovery rates that may be necessary to fit the market correlation skew in periods of turmoil (e.g July 2005).

Following results of an empirical study by Hamilton et al (2005) HW suggest to change recovery in each scenario by linking it to the conditional probability of default in that scenario:

$$\text{R}\textsc{ec}_j = 0.52 - 6.9(1 - \exp(-\lambda_j 5y)).$$

Then in computing the target function to be minimized we set recovery rates (and thus the related $\text{L}\textsc{gd} = 1 - \text{R}\textsc{ec}$ contributing to the loss function) in each scenario according to the above relationship. The final minimization problem reads:

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach by Hull and White

$$\mathbf{p}^\star = \mathrm{argmin}_{p_1,\ldots,p_s} \sum_{i=1}^{5} \left[\sum_{j=1}^{s} p_j \,\mathsf{TRANCHE}_{A_i,B_i}(0, R_{0,5y}^{A_i,B_i,\mathsf{MKT}}; \lambda = \lambda_j; \mathsf{REC} = \mathsf{REC}_j) \right]^2$$

$$+ \left[125 \left(\sum_{j=1}^{s} p_j R_{0,5y}^{\mathsf{MKT}} \sum_{k=1}^{b} \alpha_k P(0,T_k) e^{-\lambda_j T_k} - \sum_{j=1}^{s} p_j \mathsf{LGD}_j \sum_{k=1}^{b} P(0,T_k)(e^{-\lambda_j T_{k-1}} - e^{-\lambda_j T_k}) \right) \right]^2$$

$$+ c \sum_{j=2}^{s-1} \frac{(p_{j+1} + p_{j-1} - 2p_j)^2}{0.5[\exp(-\lambda_{j-1} 5y) - \exp(-\lambda_{j+1} 5y)]}$$

under the constraints $p_i > 0, \; p_1 + \ldots + p_{30} = 1.$

This problem can be effectively solved with the excel solver.

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach by Hull and White

Now we present a calibration example based on the same 5y i-Traxx data of November 11, 2004 we have seen before for the base/compound correlation examples.

This is a plot of the calibrated probabilities p^* against the default probabilities:

$$(1 - e^{-\lambda_1 5y}, p_1), (1 - e^{-\lambda_2 5y}, p_2), \ldots, (1 - e^{-\lambda_{30} 5y}, p_{30})$$

We take $c = 1$ in the smoothing part. The calibration error is negligible.

$1 - e^{-\lambda_i 5y}$	p_i
0.09%	19.78%
1.16%	17.97%
2.19%	16.04%
3.15%	13.82%
4.07%	11.26%
4.97%	8.46%
5.83%	5.64%
6.66%	3.14%
7.48%	1.29%
8.28%	0.28%
9.05%	0.00%
9.81%	0.00%
10.54%	0.00%
11.27%	0.00%
11.98%	0.01%

$1 - e^{-\lambda_i 5y}$	p_i
12.68%	0.03%
13.39%	0.12%
14.10%	0.25%
14.84%	0.34%
15.64%	0.37%
16.49%	0.30%
17.47%	0.17%
18.59%	0.05%
19.90%	0.00%
21.38%	0.00%
23.03%	0.00%
25.05%	0.06%
28.89%	0.16%
43.54%	0.24%
62.90%	0.23%

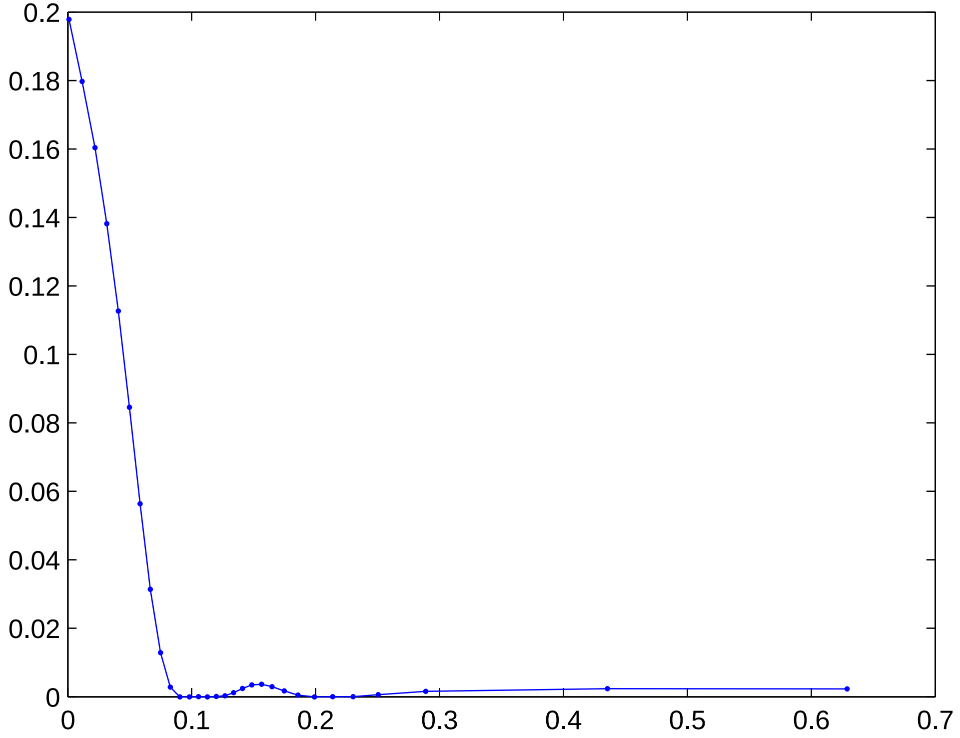


Figure 28: Implied copula calibrated parameters consistent with the whole correlation skew, plot of $(1 - e^{-\lambda_i 5y}, p_i)$.

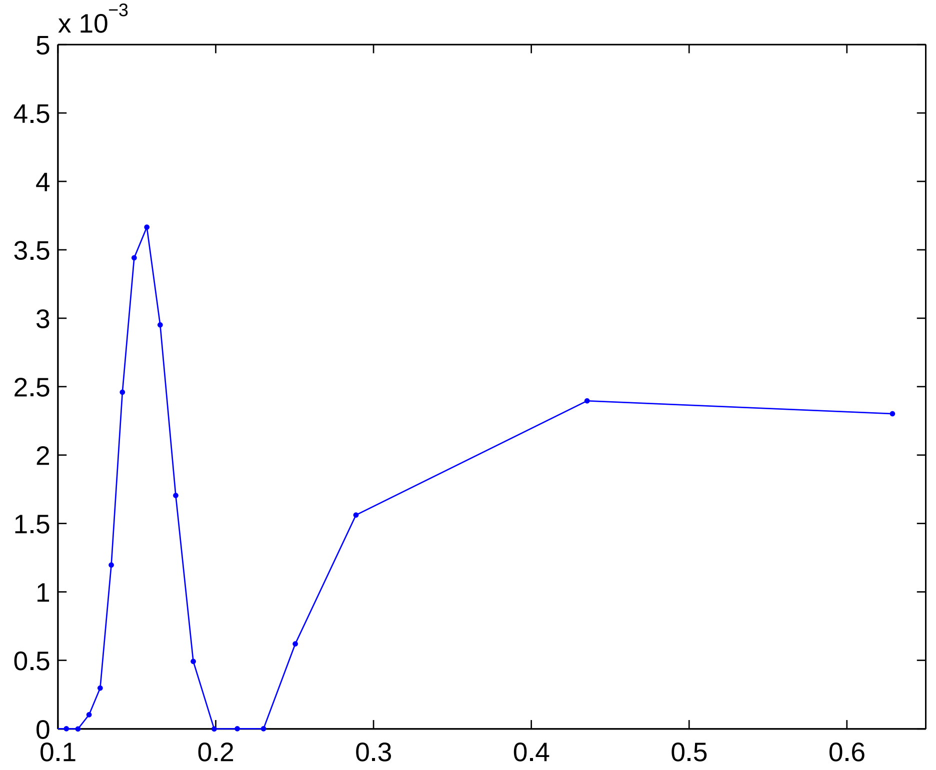


Figure 29: Implied copula calibrated parameters consistent with the whole correlation skew, plot of $(1 - e^{-\lambda_i 5y}, p_i)$ (zoom)

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach by Hull and White

Once the model is calibrated, any payoff can be priced with a weighted sum of prices each under a model assuming independent defaults and time-constant hazard rates homogeneous across all names.

$$\text{PriceIFC} = \sum_{j=1}^{s} p_j \ \text{Price(indep defaults}, \lambda = \lambda_j \text{ across all names } i = 1 \ldots N)$$

This is powerful. A drawback of this approach, like all other copula approaches, is that the core dependence across defaults is modelled as static, and there is no dependence dynamics. Pricing a tranche option would require in principle a dynamical model.

Also, there is no differentiation for single name probabilities, which are the same (although scenario-based) across names.

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach. Infinite pool.

Also with the implied factor copula one can resort to the infinite pool simplification, as we did for the gaussian factor copula. Conditional default indicators $1_{\{\tau_i < T | M = m_j\}}$ are i.i.d., so that their sample average as their number tends to infinity tends to the single common true mean:

$$DR^{j,N}(T) := \frac{1}{N}\sum_{i=1}^{N} 1_{\{\tau_i < T | M = m_j\}} \rightarrow \mathbb{E}1_{\{\tau_i < T | M = m_j\}} = \mathbb{Q}\{\tau_i < T | M = m_j\} = 1 - e^{-\lambda_j T}$$

when N tends to ∞.

Again, this way we avoid taking expectations, except the final one with respect to M. But conditional on $M = m_j$ all randomness has been ruled out by the law of large numbers and both the default rate and the loss are completely determined.

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach. Infinite pool.

$$DR(T)|M \quad \sim \quad \left\{ \begin{array}{ccc} \text{conditional default rate} & \text{Systemic scenario} & \text{Scenario probability} \\ DR^1(T) = 1 - \exp(-\lambda_1 T) & M = m_1 & p_1 \\ DR^2(T) = 1 - \exp(-\lambda_2 T) & M = m_2 & p_2 \\ \vdots & \vdots & \vdots \\ DR^s(T) = 1 - \exp(-\lambda_s T) & M = m_s & p_s \end{array} \right. \qquad (32)$$

The model is then used to calibrate the quoted index and tranches. For a preferred maturity T (typically 5 years) one sets the default rates DR to

$$DR^1(T) = 0,\ DR^2(T) = 1/125,\ DR^3(T) = 2/125,\ \ldots,\ DR^{125}(T) = 124/125$$

and inverts in the λ_j, solving

$$\lambda_j = -\ln(1 - DR^j(T))/T.$$

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach. Infinite pool.

Then one proceeds using

$$\text{Loss}_t|\{M = m_j\} = (1-\text{Rec}_j)(1-\exp(-\lambda_j t)),\ \ \text{Rec}_j = -0.1056289\ln(1-\exp(-\lambda_j)).$$

The recovery specification follows results of an empirical study we performed, see the first feature article by Torresetti, Brigo and Pallavicini (2006a) for more details and further references.

From the conditional loss $\text{Loss}_t|\{M = m_j\}$ one builds the payoff of tranches conditional on M, and then averages over M with a simple linear combination of the prices under each M scenario. See Torresetti, Brigo and Pallavicini (2006a) for more details and examples.

Calibrating all tranches consistently: The Implied Factor Copula (IFC) approach. Infinite pool.

FEATURE ARTICLE 2.

RISK NEUTRAL vs OBJECTIVE LOSS DISTRIBUTION AND CDO VALUATION.

References

- Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G. and Roncalli, T. (2000). Copulas for Finance. A reading Guide and Some Applications. *Working Paper* (available from `http://gro.creditlyonnais.fr`).
- Georges, P., Lamy, A-G., Nicolas, E., Quibel, G. and Roncalli, T., (2001). Multivariate survival modelling: a unified approach with copulas. *Working Paper* (available from `http://gro.creditlyonnais.fr`).
- Embrechts, P., Lindskog, F. and McNeil, A., (2001). Modelling Dependence with Copulas and Applications to Risk Management. *Working Paper*.
- Laurent, J.-P. and Gregory, J., (2003). Basket Default Swaps, CDO's and Factor Copulas. *Working Paper*.
- Cherubini, U., Luciano, E. and Vecchiato, W., (2004), *Copula Methods in Finance.* The Wiley Finance Series.
- Vasicek, O., (1987). Probability of loss on a loan portfolio. *Working Paper*, KMV Corporation.
- Hull, J. and White, A. (2004). Valuation of a CDO and an $\mathtt{n}^{\mathtt{th}}$ to Default CDS Without Monte Carlo Simulation, *Journal of Derivatives*, 12, 2.

- Nelsen, R. (1999), *An Introduction to Copulas*, Springer, New York.
- Li, D. X., (2000). On Default Correlation: A Copula Approach, *Journal of Fixed Income*, 9.
- Schönbucher, P. J., *Credit derivatives pricing models. Models, Pricing and Implementation*, (2003), Wiley.
- Sklar, A. (1959), *Fonctions de répartition á n dimensions et leurs marges*, Publications de l'Institut de Statistique de l'Université de Paris, 8.
- Fréchet, M. (1957): *Les tableaux de corrélation dont les marges et des bornes sont données* Annales de l'Université de Lyon, Sciences Mathématiques et Astronomie, 20.
- Brigo, D. and Alfonsi, A. (2004), *New families of Copulas based on periodic functions*, Updated version published in "Communications in Statistics: Theory and Methods", Vol 34, issue 7, 2005. Cermics report 2003-250.
- Hull, J. and White, A. (2000). Valuing Credit Default Swaps II: Modeling Default Correlations. *Journal of Derivatives*, Vol. 8, No. 3.
- Merton, R.C., (1974). On the pricing of Corporate Debt: The Risk structure of Interest Rates, *Journal of Finance*, 29.
- Laurent, J.-P. and Gregory, J., (2004). In the Core of Correlation, *Risk*, 17, 10.

- Robertson, J. P., (1992). The Computation of Aggregate Loss Distributions (available from www.defaultrisk.com).
- Melchiori, M. R., (2004). Credit Risk+ by Fast Fourier Transform (available from www.defaultrisk.com).
- BNP Paribas, (2004). DJ iTraxx: The main line for credit in Europe.
- Felsenheimer, J., Gisdakis, P., and Zaiser, M. (2004). DJ ITRAXX: Credit at its best! HVB Global Markets Research European Credit Strategy.
- McGinty, L., Ahluwalia, R., (2004). A Model for Base Correlation Calculation, JPM technical document.