Fortis Bank
Equity & Credit Derivatives Quantitative Research

The Correlation Smile Recovery

E. Vandenbrande, A. Vandendorpe, Y. Nesterov, P. Van Dooren

draft version : March 2, 2009

1 Introduction

Pricing and hedging exotic credit derivatives, even simple bespoke single tranche CDOs or First-To-Defaults, has been a challenge the past few years and has become even more difficult since the beginning of the credit crisis mid 2007. In order to fit the implied Index loss distribution in one consistent model [3], many methodologies have been proposed in the literature since the correlation crisis in 2005, ranging from simple mapping methods to more sophisticated models.

Mapping methods often fit the market but are difficult to use in order to price other payoffs than standard single tranches. Models that can price several payoffs, like CDO^2, have often difficulties in matching Index tranche quotes and single name CDS's in a consistent manner. The difficulty is to find a compromise between simplicity and robustness, which are necessary in order to price and hedge efficiently such products, and accuracy, i.e. ability to reproduce precisely daily quoted CDO index tranche prices. In an attempt to solve these issues, Andersen and Sidenius extend the Gaussian copula with random factor loadings and recovery rates [1]. They include a study on a correlation between recovery rates and the number of defaults. This feature is present in a natural way in the model we will present in this paper. Hull and White propose to find a probability distribution for the Market Factors in order to match the tranche quotes [7], [5] and [6]. Julien Turc introduces the concept of local correlation in [10]. In [3], Gregory and Laurent give an overview and a comparison of the several approaches.

However, none of these models can be easily extended to cope with a full rank correlation matrix. This feature is important in order to take into account the name by name correlation risk as viewed by the traders or the risk managers. Moreover, most of the preceding models have difficulties in preserving exactly the CDS spreads, the default probabilities and the expected recovery rate at default for all the obligors and all maturities.

In this paper, we propose a consistent pricing model for structured credit derivatives based on the market's standard Gaussian copula model that can perform all of these desirable features. The key idea is to use the gaussian copula with stochastic correlation and stochastic recovery rate. The two-dimensional probability distribution is the result of a calibration that tries to match quoted tranche prices. In order to obtain a smooth and stable probability surface, an entropy maximization regularization is used in the calibration process.

This paper is organized as follows. After some preliminary results in Section 2, the proposed model is introduced in Section 3. The optimization algorithm used for the model calibration is introduced in Section 4. Numerical examples are given in Section 5. Concluding remarks are given in Section 6.

2 Preliminary Results : The gaussian copula model

In this section we introduce the notation used in the rest of this paper as well as a brief reminder of the Gaussian copula model. We assume that we have a portfolio of N issuers, where

- $S_i(t)$ is the probability of survival of issuer i before time t,
- $Q_i(t)$ is the probability of default of issuer i before time t,
- τ_i is the default time of issuer i,
- RR_i is the expected recovery rate of issuer i,
- C is an $N \times N$ default correlation matrix.

The survival probability curve is generally stripped in order to match the market CDS quotes. Note that the expected recovery rate assumption has an influence on the survival probabilities. Under the standard Gaussian copula approach, the default dependence of the issuers is modelled by a Gaussian copula. The joint probability that $\tau_i < T$ for all i is

$$P(\tau_1 < T, \dots, \tau_N < T) = \Phi_C(\Phi^{-1}(Q_1(T)), \dots, \Phi^{-1}(Q_i(T)), \dots, \Phi^{-1}(Q_N(T))), \tag{2.1}$$

with

- $\Phi(\cdot)$, the standard normal distribution : $\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx$,
- $\Phi_C(\cdot)$, the multivariate standard normal distribution with correlation matrix C :

  $$\Phi(a_1, \dots, a_N) = \frac{1}{(2\pi)^{\frac{N}{2}} |C|^{\frac{1}{2}}} \int_{-\infty}^{a_1} \dots \int_{-\infty}^{a_N} e^{-\frac{1}{2} X^T C^{-1} X} dx_1 \dots dx_N,$$

  where $|C|$ denotes the determinant of C.

With this model, the recovery rate are most of the time assumed to be fixed (at 40% in case of the index).

It is well known that this model is unable to reproduce market tranche quotes. It gives rise to the well-known base correlation skew when trying to reproduce Market Tranche quotes with a homogeneous correlation matrix C. In the next section we will enhance this simple model to fit the tranche market in a better way.

3 The stochastic correlation model

Pricing credit structures boils down to choosing a default time copula and a recovery rate distribution. As it is impossible with the standard Gaussian copula model to recover market prices with a unique correlation matrix and fixed recovery rates we propose to put a probability distribution on the correlation matrices and recovery rates. The probability distribution will be calibrated from market prices using an entropy maximization entropy algorithm. In order to describe the model more formally, let us first define the *shifted scenarios*.

3.1 Shifted Scenario

In order to put a probability distribution on the correlations and the recovery rates, we first define a *stressed* scenario with correlation shift α and recovery rate shift β as follows:

- The shifted correlation matrix C_α is a convex combination of the original correlation matrix and an extreme (either uncorrelated or perfectly correlated) correlation matrix as follows:
  - If $-1 < \alpha \leq 0$, $C_\alpha = (1+\alpha)C - \alpha I_N$.
  - If $0 \leq \alpha < 1$, $C_\alpha = (1-\alpha)C + \alpha 1_N$.
- The shifted recovery rate for obligor i, denoted by $RR_{i,\beta}$, is computed similarly:
  - If $-1 < \beta \leq 0$, $RR_{i,\beta} = RR_i(1+\beta)$.

– If $0 \leq \beta < 1$, $RR_{i,\beta} = RR_i(1+\beta)M$, where M is a coefficient that depends on the portfolio and that is chosen such that the highest recovery rate cannot exceed 1. For instance, if the expected recovery rates of all the obligors of the portfolio are equal to 0.4, M will be equal to 1.25. For an arbitrary portfolio, one can choose the maximum possible recovery rate shift as follows:

$$M = \min_{i=1}^{N} \frac{1}{2RR_i}.$$

This ensures that any shifted recovery rate stays smaller than 1.

3.2 Assigning a Discrete Probability Distribution

In order to put a probability distribution on the shifted scenarios, we first need to discretize the universe of possible shifts. Let us fix the number of recovery rate shifts by k_{RR} and the number of correlation shifts by k_C. The total number of possible scenarios is equal to $k_{RR} \times k_C$. These are chosen in order to pave in a uniform way the square $[-1,1] \times [-1,1]$: Assume for instance that k_{RR} and k_C are odd integers. Then, $\forall\ 1 \leq i \leq k_C$, $\forall 1 \leq j \leq k_{RR}$,

$$\alpha_i = \frac{2i - k_C - 1}{k_C - 1}, \qquad \beta_j = \frac{2j - k_{RR} - 1}{k_{RR} - 1}.$$

The next step is to assign probabilities to each correlation and recovery rate scenario. Denote the probability associated to the scenario corresponding to the i-th correlation shift and the j-th recovery rate shift by $p_{i,j}$. We need to find the optimal probability matrix

$$P := \begin{bmatrix} p_{1,1} & \cdots & p_{1,k_{RR}} \\ \vdots & \ddots & \vdots \\ p_{k_C,1} & \cdots & p_{k_C,k_{RR}} \end{bmatrix}.$$

These market implied probabilities are chosen in order to match the Market tranche quotes as closely as possible. Denote by $MtM_{i,j,k}$ the Mark-to-Market corresponding the the $k-$th tranche of the scenario with the i-th correlation shift and the j-th recovery rate shift. Let us assume that there are n tranches to fit (typically, n equals 6). In order to allow more flexibility for the optimization part, let us associate a weight to each of the n tranches to fit, say w_i. When it is not possible to fit all the Market quotes perfectly, the weights can be used in the objective function of the optimization algorithm in order to allow to better fit the most important tranches. Assume that the probability matrix P is fixed. Using the properties of conditional expectations, the MtM of tranche i, denoted by $MtM_{P,i}$, is equal to

$$MtM_{P,i} = \sum_{k,l} p_{k,l} MtM_{k,l,i}. \tag{3.2}$$

The objective function, say $f(P)$, is the norm of the vector of Mark-to-Markets

$$f(P) := \left\| \begin{bmatrix} w_1 MtM_{P,1} \\ \vdots \\ w_n MtM_{P,n} \end{bmatrix} \right\|. \tag{3.3}$$

This function has to be minimized under the following constraints:

- All the probabilities have to be non-negative.

- The sum of probabilities must equal 1:

$$\sum_{i=1}^{k_C} \sum_{j=1}^{k_{RR}} p_{i,j} = 1. \tag{3.4}$$

- The expected recovery rate for each obligor must be preserved:

$$\sum_{i=1}^{k_C} \sum_{j=1}^{\frac{k_{RR}-1}{2}} \beta_j p_{i,j} + \sum_{i=1}^{k_C} \sum_{j=\frac{k_{RR}+1}{2}}^{k_{RR}} M \beta_j p_{i,j} = 0. \tag{3.5}$$

Remark 1 *Condition* (3.5) *ensures that the Mark-To-Markets of the CDS of each obligor of the portfolio are preserved. Indeed, the Mark-To-Market of a CDS is not affected by default correlation assumptions nor by recovery rate distribution as long as the expected value of the recovery rate conditional on default is preserved. Equation* (3.5) *ensures the latter.*

Note that all the preceding constraints are linear with respect to the probability coefficients $p_{i,j}$. In practice, the calibration procedure requires two steps :

- Compute all Mark-To-Markets, pvo1's and expected losses of the market tranches in every stressed scenario. We need to compute $n \times k_C \times k_{RR}$ tranche prices. This step is clearly time consuming, but needs to be performed only once a day.

- Use an optimization algorithm to minimize (3.3). This part of the calibration process is usually very fast (at most a few seconds).

Depending on the choice of the norm to minimize and additional smoothing conditions, several optimisation schemes are possible. If the objective function is the infinity norm of the Mark-to-Market vector, a simple choice would be to use the classical Simplex Algorithm. The problem is that one would obtain as a typical result only $n+2$ scenario with a positive probability, all the rest being zero. Moreover, a small change in the inputs (MtMs) can lead to a large change in the outputs (optimal probability distribution).

To cope with these problems, one can modify the optimisation scheme by adding an entropy term to the objective function (3.3):

$$f(P) := \left\| \begin{bmatrix} w_1 MtM_{P,1} \\ \vdots \\ w_n MtM_{P,n} \end{bmatrix} \right\| + \delta \sum_{i,j} p_{i,j} \ln p_{i,j}, \tag{3.6}$$

where δ is a smoothness parameter. The optimization problem becomes then strongly convex and can be solved efficiently. This stabilizes the solution. The detailed optimization algorithm is studied in Section 4.

4 Regularized solution in convex minimization

In this section, we propose a dual gradient algorithm that solves the optimization problem efficiently. For an introduction to convex optimization, see for instance [2, 9].

Let us write down the problem of our interest in the following form:

$$\min_{x \in Q} \; f(x). \tag{4.7}$$

where function $f(x)$ is convex and continuously differentiable on some closed convex set Q. Note that in general the set of optimal solutions of problem (4.7) can be unstable with respect to small perturbations of the objective function.

In order to treat this instability, we propose to use a *regularized* solution to this problem. For that, we need to choose an appropriate *prox-function* $d(x)$ of the set Q. This is a continuously differentiable non-negative function, which attains zero at some point $x_0 \in Q$. Moreover, it must be *strongly convex* on Q with some convexity parameter $\sigma > 0$:

$$d(y) \quad \geq \quad d(x) + \langle \nabla d(x), y - x \rangle + \tfrac{1}{2}\sigma \| y - x \|^2, \quad x, y \in Q. \tag{4.8}$$

Now we can define a trajectory of regularized optimal solutions to problem (4.7) as follows:

$$x^*_f(\delta) \quad = \quad \arg\min_{x \in Q} \left[\, f^{\delta}(x) \; \stackrel{\mathrm{def}}{=} \; f(x) + \delta \cdot d(x) \, \right], \tag{4.9}$$

where $\delta > 0$ is a tolerance parameter. In the remaining part of the paper, we justify a simple technique for approximating the points of this trajectory.

We assume that our variables belong to finite-dimensional real vector space R^n. For two (column) vectors x and y from R^n their *scalar product* is defined in the standard way:

$$\langle x, y \rangle \quad = \quad \sum_{i=1}^{n} x^{(i)} y^{(i)}.$$

If we fix for R^n some norm $\| \cdot \|$, then the *dual norm* is defined by

$$\| s \|_* \quad = \quad \max_{\|x\| = 1} \langle s, x \rangle.$$

Of course, the most important norm is the Euclidean one:

$$\| x \|_{(2)} \quad \stackrel{\mathrm{def}}{=} \quad \langle x, x \rangle^{1/2}.$$

Finally, we done by $\nabla f(x) \in R^n$ the *gradient* of function f at some point $x \in R^n$.

4.1 Dual gradient method

The stability of trajectory $x^*_f(\delta)$ is justified by the following fact.

Lemma 1 *Define the distance between two functions f_1 and f_2 in the following way:*

$$\Delta(f_1, f_2) \quad = \quad \max_{x \in Q} \| \nabla f_1(x) - \nabla f_2(x) \|_*. \tag{4.1}$$

Then, for any $\delta > 0$ we have

$$\| x^*_{f_1}(\delta) - x^*_{f_2}(\delta) \| \quad \leq \quad \tfrac{\Delta(f_1, f_2)}{\sigma \cdot \delta}. \tag{4.2}$$

Proof:
The first-order optimality conditions for function f_i^δ, $i = 1, 2$, can be written as follows:

$$\langle \nabla f_i(x^*_{f_i}(\delta)) + \delta \nabla d(x^*_{f_i}(\delta)), y - x^*_{f_i}(\delta)\rangle \;\equiv\; \langle \nabla f_i^\delta(x^*_{f_i}(\delta)), y - x^*_{f_i}(\delta)\rangle \;\geq\; 0 \tag{4.3}$$

for all $y \in Q$. Since function f_1^δ is strongly convex with parameter $\sigma\delta$, we have

$$\begin{aligned}
\sigma\delta \|x^*_{f_1}(\delta) - x^*_{f_2}(\delta)\|^2 &\leq \langle \nabla f_1^\delta(x^*_{f_1}(\delta)) - \nabla f_1^\delta(x^*_{f_2}(\delta)), x^*_{f_1}(\delta) - x^*_{f_2}(\delta)\rangle \\
&\overset{(4.3)}{\leq} \langle \nabla f_1^\delta(x^*_{f_2}(\delta)), x^*_{f_2}(\delta) - x^*_{f_1}(\delta)\rangle \\
&= \langle \nabla f_1(x^*_{f_2}(\delta)) + \delta \nabla d(x^*_{f_2}(\delta)), x^*_{f_2}(\delta) - x^*_{f_1}(\delta)\rangle \\
(\text{by (4.3) with } i = 2,\ y = x^*_{f_1}(\delta)) \quad &\leq \langle \nabla f_1(x^*_{f_2}(\delta)) - \nabla f_2(x^*_{f_2}(\delta)), x^*_{f_2}(\delta) - x^*_{f_1}(\delta)\rangle \\
&\leq \Delta(f_1, f_2) \cdot \|x^*_{f_2}(\delta) - x^*_{f_1}(\delta)\|.
\end{aligned}$$

□

Example 1 *Consider the case when $Q \subseteq \Delta_n \equiv \{x \in R^n_+ : \langle e_n, x\rangle = 1\}$. Then it is natural to define prox-function $d(x)$ using the* entropy function*:*

$$\eta(x) \;=\; \sum_{i=1}^n x^{(i)} \ln x^{(i)},$$

In this case, we can measure distances in Δ_n in l_1-norm:

$$\|h\| \;=\; \sum_{i=1}^n |h^{(i)}|, \quad h \in R^n.$$

The corresponding dual norm is then

$$\|s\|_* \;=\; \max_{1 \leq i \leq n} |s^{(i)}|, \quad s \in R^n.$$

Note that the entropy function is strongly convex with respect to l_1-norm with constant $\sigma = 1$. Hence, we can choose

$$d(x) \;=\; \omega_\eta(x_0, x),$$

where $x_0 = \arg\min_{x_0 \in Q} \eta(x)$, and $\omega_\eta(x_0, x)$ is the Bregman distance *between x_0 and x measured by function $\eta(\cdot)$:*

$$\omega_\eta(x_0, x) \;\overset{\text{def}}{=}\; \eta(x) - \eta(x_0) - \langle \nabla\eta(x_0), x - x_0\rangle. \tag{4.4}$$

Note that function $\omega_\eta(x, y)$ is strongly convex in y with the same parameter σ as $\eta(\cdot)$. In view of its definition, we have $\nabla d(x) = \nabla\eta(x) - \nabla\eta(x_0)$.

Consider now two quadratic functions:

$$f_A(x) \;=\; \tfrac{1}{2}\|Ax\|^2_{(2)}, \quad f_B(x) \;=\; \tfrac{1}{2}\|Bx\|^2_{(2)},$$

where $A, B \in R^{m\times n}$. Then $\nabla f_A(x) = A^T Ax$, and $\nabla f_B(x) = B^T Bx$. Therefore,

$$\begin{aligned}\Delta(f_A, f_B) &\leq \max_{x\in\Delta_n} \|(A^TA - B^TB)x\|_* = \max_{1\leq i\leq n} \|(A^TA - B^TB)e_i\|_* \\ &= \max_{1\leq i,j\leq n} |\langle a_i, a_j\rangle - \langle b_i, b_j\rangle|,\end{aligned}$$

where a_i and b_i are the columns of corresponding matrices. Denote $\epsilon_i = a_i - b_i$. Then

$$\begin{aligned}\langle a_i, a_j\rangle - \langle b_i, b_j\rangle &= \langle a_i, b_j + \epsilon_j\rangle - \langle a_i - \epsilon_i, b_j\rangle \\ &= \langle a_i, \epsilon_j\rangle + \langle b_j, \epsilon_i\rangle.\end{aligned}$$

Defining now the matrix norms

$$\|A\| = \max_{1\leq i\leq n} \|a_i\|, \quad \|A\|_d = \max_{1\leq i\leq n} \|a_i\|_*,$$

we obtain

$$\Delta(f_A, f_B) \leq (\|A\| + \|B\|) \cdot \|A - B\|_d.$$

□

In the sequel, we need several simple properties of Bregman distances.

(P1) For any linear function $l(x)$ we have $\omega_l(x, y) \equiv 0$.

(P2) For any convex function ϕ and coefficient $\beta > 0$ we have $\omega_{\beta\phi}(x, y) = \beta\omega_\phi(x, y)$.

(P3) For two convex functions ϕ_1 and ϕ_2 we have $\omega_{\phi_1+\phi_2}(x, y) = \omega_{\phi_1}(x, y) + \omega_{\phi_2}(x, y)$.

Finally, the Bregman distance can be used for estimating growth of convex function with respect to its minimal value.

Lemma 2 *Let function ϕ be convex and differentiable on Q. Then for any $x \in Q$ we have*

$$\phi(x) \geq \phi(x_*) + \omega_\phi(x_*, x), \tag{4.5}$$

where $x_* \in \operatorname{Arg}\min\limits_{y\in Q} \phi(y)$.

Proof:
Indeed, from the first-order optimality condition we know that

$$\langle\nabla\phi(x_*), x - x_*\rangle \geq 0, \quad x \in Q.$$

Therefore, for any $x \in Q$ we have

$$\phi(x) = \phi(x_*) + \langle\nabla\phi(x_*), x - x_*\rangle + \omega_\phi(x_*, x) \geq \phi(x_*) + \omega_\phi(x_*, x).$$

□

Let us show how to find a good feasible approximation to $x^*_f(\delta)$ by a simple (dual) gradient scheme. We assume that the objective function f has Lipschitz-continuous gradient:

$$\|\nabla f(x) - \nabla f(y)\|_* \quad \leq \quad L_f \|x - y\|, \quad x, y \in Q. \tag{4.6}$$

Let us describe first the *dual mapping* based on the prox-function $d(x)$. This mapping as applied to some vector g is defined as follows:

$$T_\gamma(g) \quad = \quad \arg\min_{y \in Q} [\, \langle g, y \rangle + \gamma d(y) \,]. \tag{4.7}$$

Lemma 3 *Let $\gamma \geq \frac{1}{\sigma} L_f$. Then for any $x \in Q$ we have*

$$f^\delta(T_{\gamma+\delta}(\nabla f(x) - \gamma \nabla d(x))) \quad \leq \quad \min_{y \in Q} [\, f(x) + \langle \nabla f(x), y - x \rangle + \gamma \omega_d(x, y) + \delta \cdot d(y) \,]. \tag{4.8}$$

Proof:
Let us fix an arbitrary $x \in Q$. Denote the objective function of the problem in the right-hand side of inequality (4.8) by $\psi_{\gamma,\delta}(y)$. Since d is strongly convex, we have

$$\omega_d(x, y) \quad \geq \quad \tfrac{1}{2}\sigma \|x - y\|^2.$$

Therefore, for $\gamma \geq \frac{1}{\sigma} L_f$, we have

$$\begin{aligned} \psi_{\gamma,\delta}(y) \quad &\geq \quad f(x) + \langle \nabla f(x), y - x \rangle + \tfrac{1}{2} L_f \|y - x\|^2 + \delta d(y) \\ &\geq \quad f(y) + \delta d(y) \;\equiv\; f^\delta(y). \end{aligned}$$

It remains to note that $T_{\gamma+\delta}(\nabla f(x) - \gamma \nabla d(x)) = \arg\min\limits_{y \in Q} \psi_{\gamma,\delta}(y)$. □

We will justify the dual gradient method by the technique of *estimate sequences*. In accordance to this approach, we need to update recursively the sequence of estimate functions

$$\psi_k(x) \quad = \quad \sum_{i=0}^{k-1} a_i [f(v_i) + \langle \nabla f(v_i), x - v_i \rangle] + (1 + \delta A_k) d(x),$$

where $\{a_i\}$ are some positive weights, $A_k = \sum\limits_{i=0}^{k-1} a_i$, and $\{v_i\} \subset Q$ are some auxiliary points which we will specify later. It is convenient to set $A_0 = 0$ and $\psi_0(x) = d(x)$.

Note that by construction we have

$$\psi_k(x) \quad \leq \quad A_k f^\delta(x) + d(x), \quad x \in Q. \tag{4.9}$$

Let us show how we can maintain another important condition:

$$\sum_{i=0}^{k-1} a_i f^\delta(y_i) \quad \leq \quad \psi^*_k \;\stackrel{\mathrm{def}}{=}\; \min_{x \in Q} \psi_k(x). \tag{4.10}$$

where $\{y_k\}$ is a minimization sequence for optimization problem in definition (4.9). Denote

$$v_k \quad = \quad \arg\min_{x \in Q} \psi_k(x), \quad k \geq 0.$$

In view of Lemma 2, for any $x \in Q$ and $k \geq 0$ we have

$$\begin{array}{rcl} \psi_k(x) & \geq & \psi_k^* + \omega_{\psi_k}(v_k, x) \stackrel{(P3),(P1)}{=} \psi_k^* + \omega_{(1+\delta A_k)d}(v_k, x) \\ & \stackrel{(P2)}{=} & \psi_k^* + (1 + \delta A_k)\omega_d(v_k, x). \end{array} \tag{4.11}$$

Note that condition (4.10) is valid for $k = 0$. Assume that it is valid for some $k \geq 0$. Then

$$\begin{array}{rcl} \psi_{k+1}^* & = & \min\limits_{x \in Q} \{ \ \psi_k(x) + a_k[f(v_k) + \langle \nabla f(v_k), x - v_k \rangle + \delta d(x) \ ] \ \} \\ & \stackrel{(4.11)}{\geq} & \min\limits_{x \in Q} \{ \ \psi_k^* + (1 + \delta A_k)\omega_d(v_k, x) + a_k[f(v_k) + \langle \nabla f(v_k), x - v_k \rangle + \delta d(x) \ ] \ \} \\ & \stackrel{(4.10)}{\geq} & \sum\limits_{i=0}^{k-1} a_i f^\delta(y_i) + a_k \min\limits_{x \in Q} [\ f(v_k) + \langle \nabla f(v_k), x - v_k \rangle + \gamma_k \omega_d(v_k, x) + \delta d(x) \ ], \end{array}$$

where $\gamma_k = (1 + \delta A_k)/a_k$. Hence, if we choose

$$\begin{array}{rcl} a_k & \leq & \sigma 1 + \delta \frac{A_k}{L_f}, \\ y_k & = & T_{\gamma_k + \delta}(\nabla f(v_k) - \gamma_k \nabla d(v_k)), \end{array}$$

then, in view of Lemma 3, condition (4.10) will be valid for iteration counter $k + 1$. If constant L_f is known, then we get the following method.

Set $\psi_0(x) = d(x)$, $A_0 = 0$, $\gamma = \frac{1}{\sigma} L_f$. For $k \geq 0$ iterate:

1. Compute $v_k = \arg \min\limits_{x \in Q} \psi_k(x)$.
2. Define $a_k = \sigma 1 + \delta \frac{A_k}{L_f}$, $A_{k+1} = A_k + a_k$. (4.12)
3. Compute $y_k = T_{\gamma + \delta}(\nabla f(v_k) - \gamma_k \nabla d(v_k))$.
4. Update $\psi_{k+1}(x) = \psi_k(x) + a_k[f(v_k) + \langle \nabla f(v_k), x - v_k \rangle + \delta d(x)]$.

By simple induction, it is easy to see that the coefficients a_k and A_k in this scheme admit a closed-form representation:

$$a_k \ = \ \frac{\sigma}{L_f} \left(1 + \sigma \cdot \frac{\delta}{L_f}\right)^k, \quad A_k \ = \ \frac{1}{\delta} \left[\left(1 + \sigma \cdot \frac{\delta}{L_f}\right)^k - 1 \right], \quad k \geq 0. \tag{4.13}$$

Hence, in view of inequalities (4.9), (4.10), the rate of convergence of this scheme is linear:

$$\sum_{i=0}^{k-1} a_i \left[f^\delta(y_i) - f^\delta(x_f^*(\delta)) \right] \ \leq \ d(x_f^*(\delta)). \tag{4.14}$$

The main drawback of the method (4.12) is the presence of usually unknown parameter L_f. However, it can be easily estimated by a well known very efficient backtracking strategy (see, for example, [8]).

4.2 Auxiliary problems

It remains to show how we can find the solution to the auxiliary problem (4.7). Let us assume that the feasible set Q has the following structure:

$$Q \quad = \quad \{x \in S : \; Ax = b\}, \tag{4.1}$$

where A is an $m \times n$-matrix, and the set S is *simple*. This means that we can solve *explicitly* the following minimization problem:

$$V(s) \quad = \quad \arg\min_{x \in S} [\langle s, x\rangle + d(x)]. \tag{4.2}$$

In this case, the problem (4.7) can be written in the *dual* form:

$$\begin{aligned}
\min_{x \in Q} [\langle g, x\rangle + \gamma d(x)] \quad &= \quad \min_{x \in S} \max_{y \in R^m} [\langle g, x\rangle + \gamma d(x) + \langle y, b - Ax\rangle] \\
&= \quad \max_{y \in R^m} \left\{ \langle b, y\rangle + \min_{x \in S} [\langle g - A^T y, x\rangle + \gamma d(x)] \right\} \\
&= \quad \max_{y \in R^m} \left\{ \langle b, y\rangle + \gamma \psi\left(\tfrac{1}{\gamma}(g - A^T y)\right) \right\},
\end{aligned}$$

where the *concave* function $\psi(s)$ is easily computable:

$$\psi(s) \quad = \quad \langle s, V(s)\rangle + d(V(s)).$$

Thus, our initial projection problem can be solved in the dual form:

$$\max_{y \in R^m} \left\{ \langle b, y\rangle + \gamma \psi\left(\tfrac{1}{\gamma}(g - A^T y)\right) \right\}. \tag{4.3}$$

In many important situations, the number of the linear constraints defining the set Q is small, so the problem (4.3) is simple. Note that the optimal solution $y^*_\gamma(g)$ to this problem can be used for computing the optimal solution of the problem (4.7):

$$T_\gamma(g) \quad = \quad V\left(\tfrac{1}{\gamma}(g - A^T y^*_\gamma(g)\right). \tag{4.4}$$

5 Practical Results

The Stochastic Correlation model has been calibrated using the 5-years tranche prices of the iTraxx index from March 2007 to September 2008. Clearly, one distinguishes two different regimes. Before August 2007, the probability distribution is mainly located in scenarii corresponding to very low correlations and around the expected recovery rate. The market seems to price only idiosyncratic risks corresponding to correlation shifts close to -1.

Note that a small probability bump already exists in March 2007, but it is almost not visible in Figure 1. As shown in Table 1, the calibrated model reproduces the market spreads very precisely. On August 2007, a systematic risk appears and gains in importance until today. This coincides with the subprime crisis that appeared in the summer 2007 in the credit market. Clearly, two probability peaks appear on Figure 2. The peak with a correlation shift at -1 corresponds to the idiosyncratic risk. The peak around a correlation shift of 1 and a recovery rate shift of -1 corresponds to the new important systemic risk initiated with the subprime crisis. Today, the model does not fit precisely

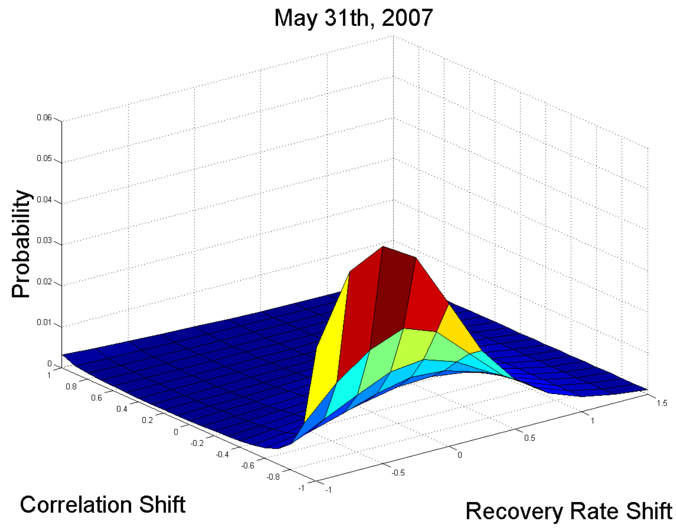

Figure 1: iTraxx Europe Implied Probability Distribution

Tranche Level	Market Spread	Obtained Spread	Error
0%-3%	6.25 %	6.243 %	-0.007 %
3%-6%	0.40 %	0.407 %	0.008 %
6%-9%	0.10 %	0.093 %	-0.009 %
9%-12%	0.05 %	0.052 %	0.007 %
12%-22%	0.02 %	0.029 %	0.010 %
22%-100%	0.01 %	0.006 %	0.001 %

Table 1: iTraxx Spread Calibration Error, May 31th, 2007

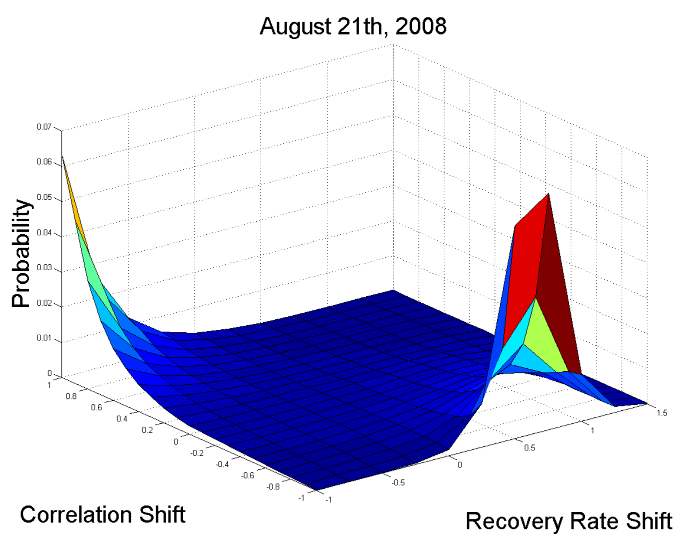

Figure 2: iTraxx Europe Implied Probability Distribution

the 12%-22% and 22%-100% tranches. Typically, the obtained fair spread of the former tranche is higher than the market tranche and inversely for the latter. A typical error of $12bp$ for the $12 - 22$ tranche and $-5bp$ for the $22 - 100$ tranche is observed today on the ITraxx Serie 10. Note however that under the present crisis situation in the Credit Market, today indices prices should be interpreted with great care.

6 Concluding Remarks

Compared with other models proposed in the literature, the proposed model has several advantages. As shown in Section 5, the proposed model is simple to understand and the obtained probability surface is easy to interpret. The calibration results are very satisfactory. Under normal market conditions previous to summer 2007, the fit is very good for all the tranches. Even in today's stressed situation, the fit remains globally satisfactory, except for the highest tranches.

Moreover, unlike the approach proposed in [5, 4], the CDS spreads, survival probabilities and expected recovery rates are all automatically preserved for all maturities and all the obligors. In other approaches, the CDS spreads are not easily preserved or preserved for one maturity only. An advantage compared to the more straightforward base correlation approach is that any payoff that depends on the loss distribution of the portfolio can be computed directly. Indeed, once the model is calibrated, simulating the defaults in a Monte-Carlo setting is straightforward. Note also that a natural mapping methodology is generated by the model. Indeed, once the probability distribution is found, it can be applied to any portfolio. Finally, thanks to the proposed optimisation scheme, the model parameters have proved to be particularly smooth and stable through time.

References

[1] L. Andersen and J Sidenius. Extensions to the gaussian copula: random recovery and random factor loadings. *Journal of Credit Risk*, 1(1), 2004.

[2] S. Boyd and L. Vandenberghe. *Convex Optimization*. 2004.

[3] X. Burtschell, J. Gregory, and J-P. Laurent. Beyond the gaussian copula: Stochastic and local correlation. 2005.

[4] J. Hull and A. White. Valuation of a cdo and an nth to default cds without monte carlo simulation. *Journal of Derivatives*, 12(2), 2004.

[5] J. Hull and A. White. The perfect copula. University of Toronto, 2005.

[6] J. Hull and A. White. Valuing credit derivatives using an implied copula approach. *Journal of Derivatives*, 14(2), 2006.

[7] J. Hull and A. White. An improved implied copula model and its application to the valuation of bespoke cdo tranches. University of Toronto, 2008.

[8] Y. Nesterov. Gradient methods for minimizing composite objective function. Technical Report 76, CORE Discussion Paper.

[9] Y. Nesterov. *Introductory Lectures on Convex Optimization, A Basic Course*. Kluwer Academic Publishers, 2003.

[10] J. Turc, P. Very, and D. Benhamou. Pricing cdos with a smile. Technical report, Societe Generale, 2005.