Implied Copula CDO Pricing Model: Entropy Approach

Alex Veremyev\textsuperscript{1}, Alex Nakonechnyi\textsuperscript{2}, Stan Uryasev\textsuperscript{3}, R. Tyrrell Rockafellar\textsuperscript{4}

RESEARCH REPORT \# 2009-1

Risk Management and Financial Engineering Lab
Department of Industrial and Systems Engineering
University of Florida, Gainesville, FL 32611

Current Version: June 12, 2009

Correspondence should be addressed to: Stan Uryasev

Abstract

A so-called “implied copula” CDO pricing model is considered for calibrating obligor hazard rates. To find the probability distribution of the hazard rates, Hull and White suggested minimizing the sum of deviations from no-arbitrage equations and a smoothing term. This paper proposes an alternative “entropy approach” to the implied copula model. The distribution is found by maximizing entropy with no-arbitrage constraints based on bid and ask prices of CDO tranches. To reduce the noise a new class of distributions is introduced, so-called “CCC distributions”. A case study shows that the entropy approach has a stable performance, while the Hull and White model is sensitive to the smoothing coefficient and the number of hazard rates on the grid. The MATLAB code used for conducting numerical experiments is provided.

Keywords: implied copula, maximum entropy, copula.

Introduction

\textsuperscript{1}University of Florida, ISE Department, P.O. Box 116595, 303 Weil Hall, Gainesville, FL 32611-6595; e-mail: averemyev@ufl.edu.

\textsuperscript{2}University of Florida, ISE Department, P.O. Box 116595, 303 Weil Hall, Gainesville, FL 32611-6595; e-mail: nakon@ufl.edu.

\textsuperscript{3}University of Florida, Risk Management and Financial Engineering Lab, ISE Department, P.O. Box 116595, 303 Weil Hall, Gainesville, FL 32611-6595; e-mail: uryasev@ufl.edu.

\textsuperscript{4}Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195-4350 email: rtr@math.washington.edu
The pricing of CDO contracts is a difficult quantitative problem faced by credit risk markets. The main issue is uncertainty about obligors default risk. This paper considers a so-called “implied copula” CDO pricing model for calibrating obligor hazard rates. The idea of this model is that, conditional on different market states, the obligors have different hazard rates. For example, if the market goes up then the obligor may have a lower risk of default (low hazard rate), or if the market goes down then it is more likely for the obligor to default during the contract period (high hazard rate).

To find the probability distribution of hazard rates, Hull and White (2006) suggested the so-called “implied copula” model. This is not a specific copula like Gaussian, Student-t, or double-t. It is called implied because it can be deduced from market quotes. The CDO tranche quotes are used for calibration. We considered the simplest version of the implied copula approach in which it is assumed that all companies being modeled have the same hazard rates and the same recovery rates (homogeneous case). Depending on the market scenario these hazard rates are different. To satisfy market quotes we need a search for probabilities to apply to individual hazard rates. The Hull and White (2006) model minimizes the sum of deviations from no-arbitrage equations and a smoothing term. The motivation in the deviation term comes from the equality between the mid-price of the CDO tranche and the expected payoff on this tranche (no-arbitrage constraint in risk-neutral setting). This equality may not be feasible for some CDO price quotes. The smoothing term is introduced to reduce the noise in the distribution. We observed, however, that the optimal solution is quite sensitive to the smoothing term coefficient.

This paper proposes an alternative “entropy approach” to the implied copula model. We found the distribution by maximizing the entropy with no-arbitrage constraints based on bid and ask prices of CDO tranches. In our numerical experiments these constraints were feasible and we did not need to introduce the deviation from no-arbitrage constraints. To reduce the noise in the distribution we introduced a new class of distributions, called “CCC distributions.” This is a wide class of distribution functions containing the normal, gamma, and the F distributions. By definition, for the continuous distributions in this class, the PDF is convex from the beginning to some point, then it is concave to some further point, and then it is again convex to the end. We called this class of distributions the CCC distributions (CCC is the abbreviation for convex/concave/convex). For discrete distributions we generalized this property to the points where the discrete distribution is defined. By a discrete distribution we mean a function which assigns some probability to each of finitely many hazard rates.

The paper presents a case study implementing Hull and White (2006) and our entropy approaches. We have demonstrated the approaches with the December, 2006 iTraxx tranche quotes. The data were obtained from the Arnsdorf and Halperin (2007) paper. We decided to use their data since it contained the bid and ask quotes. We also demonstrate results for the more recent data where the market was in unstable condition. To do the case study we used the Portfolio Safeguard (PSG) package (MATLAB Environment) by American Optimal Decisions (AOrDa.com). The case study shows that the entropy approach has a stable
performance, while the Hull and White model is sensitive to the smoothing coefficient and the number of hazard rates on the grid. We provide MATLAB code used for conducting numerical experiments. It can be found on The MathWorks web site (www.mathworks.com), in the file exchange- optimization area.

The paper proceeds as follows: Section 1 summarizes the implied copula model for hazard rates introduced by Hull and White (2006). Section 2 presents the entropy approach to the implied copula model. It provides the formal optimization problem statements and the heuristic algorithm for finding the implied probability distribution of hazard rates. Section 3 discusses the case study.

1. Conventional Copula and the Implied Copula

This section summarizes the implied copula approach proposed by Hull and White (2006). For a full model description a reader may refer to the Hull and White (2006) paper. A one-factor Gaussian copula model, first introduced by Li (2000), has became an industry standard. It models default intensities as a weighted sum of a market factor and an idiosyncratic term, a firm-dependent component. The model provides a correlation structure between default intensities of different obligors.

Define default intensities $X_i (1 \leq i \leq n)$ by:

$$X_i = \alpha_i V + \sqrt{1 - \alpha_i^2} W_i,$$  \hspace{1cm} (1)

where $V$ is a market factor and $W_i$ is an idiosyncratic term (firm-dependent component). Let $Q_i(t)$ be the cumulative distribution of (unconditional) time to default of company $i$ and let $F_i(t)$ be the cumulative distribution of $X_i$. Default intensity is then mapped to default time $\tau_i$ as $F_i(X_i) = Q_i(\tau_i)$.

A convenient way of defining $Q_i$ is through a company hazard rate. The latter has an interpretation of default intensity if the default is modeled as the first event in a non-homogeneous Poisson process. The hazard rate $\lambda_i(\tau)$ is related to $Q_i(t)$ in the following way:

$$\lambda_i(\tau) = \frac{1}{1 - Q_i(\tau)} \frac{dQ_i(\tau)}{d\tau}.$$  \hspace{1cm} (2)

Hazard rates are popular in credit risk applications due to ease of implementation, convenient analytic expressions and clear physical interpretation.

We define a grid $\lambda_1, ..., \lambda_I$ of possible hazard rates$^5$. In other words, we assume that in each scenario the hazard rate is constant and the same for all obligors. As in Hull and White (2008), we set the lowest hazard rate such that there is almost no chance to default ($\lambda_1 = 10^{-8}$), and the highest hazard rate such that almost all companies default immediately ($\lambda_I = 100$). The intermediate hazard rates are chosen so that the $\ln \lambda_k$ are

$^5$the hazard rate can be viewed as the severity of the credit environment over the life of the CDO.
equally spaced. We present results for the number of hazard rates on the grid from 100 to 1,000. We try to find out if the increasing number of scenarios of hazard rates leads to some limiting distribution. This property is expected from a “well defined” model where the precision increasingly improves the performance of the model. For a specific value of the market factor, defaults of each company or obligation are independent and described by their conditional hazard rates. These hazard rates are simultaneously higher or lower. Hull and White proposed a so-call “implied copula” model prescribing the same unconditional hazard rate to each company and then moved all hazard rates simultaneously (or, more precisely, proportionally) so that the collateral hazard takes on pre-defined values $\lambda_1, \ldots, \lambda_I$. The scenarios for hazard rate $\lambda_i$ have probabilities $p_i$.

To fit a probability distribution for hazard rates to the market we consider CDO (Collateralized Debt Obligation) price data. We use the 5-year quotes for iTraxx index tranches on December, 2006\(^7\). By sampling default scenarios corresponding to each level of $\lambda_i$, the net payoff\(^8\) of each tranche $j$ can be determined, conditional on the hazard rate scenario $\lambda_i$. Denote this payoff by $a_{ij}$. Note that this net payoff is calculated with the mid-quotes for the spreads for every tranche. Later, we will describe how the bid and ask quotes can be used in no-arbitrage consideration. A probability $p_i$ is assigned to $\{\lambda_i\}$ to form a probability distribution of hazard rates. No-arbitrage considerations in a risk-neutral setting assume that the expected net payoff of each CDO tranche is equal to zero\(^9\)

\[
\sum_{i=1}^{I} a_{ij}p_i = 0 \quad j = 1, \ldots, J .
\] (3)

The numerical experiments with the market data show that in some cases the constraint (3) is infeasible. In such cases we need to find a distribution approximately solving equation (3). Some criterion has to be defined to choose a distribution the closest to a feasible one. Hull and White (2006) proposed solving the following optimization problem to find a suitable probability distribution:

**Problem A**

\[
\min_p (D(p) + S(p))
\]

\(^6\)They also proposed an extension to this model where they assume a correlation between obligors (Hull and White [2006])

\(^7\)We obtained the data from Arnsdorf and Halperin (2007). We found that paper useful since it provides the bid and ask quotes.

\(^8\)the difference between expected present value of premium leg payments and default leg payments.

\(^9\)tranche payoffs (with both payment legs included) have to be zero under no-arbitrage assumptions. The tranche spread has to be established at such a level that the expected payoffs through the premium leg are precisely equal to the expected default losses, in other words, so that the premium leg has the same present value as the default leg.
subject to

probability distribution constraints

\[
\sum_{i=1}^{I} p_i = 1, \quad \text{subject to}
\]

\[
p_i \geq 0, \quad i = 1, \ldots, I.
\]  

(4)

(5)

where \( D(p) \) is a deviation term

\[
D(p) = \sum_{j=1}^{J} \left( \sum_{i=1}^{I} p_i a_{ij} \right)^2,
\]

(6)

and \( S(p) \) is a smoothing term

\[
S(p) = c \sum_{i=2}^{I-1} \left[ \frac{p_{i+1} + p_{i-1} - 2p_i}{0.5(d_{i+1} - d_{i-1})} \right]^2.
\]

(7)

The deviation term penalizes large deviations from zero of the net expected payoff of every tranche. The smoothing term enforces that every three consecutive points on the hazard rate distribution are approximately on the same line. The smoothing term is larger for larger differences from the straight line.

The smoothing term introduces distortion into resulting distribution, but a reasonable level of distortion may be better than a ragged distribution shape. The smoothing effect appears to decrease with the increase in the number of atoms in the distribution. Also, the coefficient \( c \) has to be chosen by trial and error.

Let us consider for instance the case presented in the paper of Hull and White (2006). We ran the case study to analyze results. First, we used the data (see the Figure 2) to simulate the expected cash flows on every tranche for different hazard rates. Then we solved the optimization Problem A. Figures 3 and 4 show graphs of optimal distributions obtained for different numbers of points and different values of the smoothing term coefficient \( c \).

It seems that if we find a “good” smoothing coefficient \( c \) for a particular number of atoms in the distributions, it will not work the same way if we change the number of atoms. Therefore, the smoothing coefficient \( c \) should be chosen individually for every number of points on the hazard rate grid. Moreover, we can not find a “reasonable” justification for why one smoothing coefficient is better than any other one.

The next section proposes an alternative approach which we call the “entropy approach” for finding the “best” probability distribution. With this approach we tried to exclude from the model arbitrary parameters, such as the smoothing coefficient.
2. Implied Copula: Entropy Approach

This section proposes an entropy approach to the implied copula model, discusses the reasons why and when such an approach is useful, and provides a heuristic algorithm to find the “best” probability distribution for hazard rates. Hull and White (2006) minimize the sum of squared deviations of tranche payoffs from “perfect fit” (6) and the smoothing term (7). In the recent Hull and White (2008) paper there is another approach to find a suitable probability distribution. The authors assume that the distribution is a log-t and calibrate its parameters to fit the market quotes. We propose an alternative maximum entropy principle and suggest finding the distribution in the class of CCC distributions (which will be described later).

The Maximum Entropy Principle (first introduced by Shannon, see also Golan (2002)) is popular in information theory. This principle is actively used in financial applications; see for instance Miller and Liu (2002), Chu and Satchell (2005), Mayer-Dautrich and Wagner (2007). The essence of the Maximum Entropy Principle is that with some given information about the distribution (specified through equations and constraints) we maximize the entropy and select the most “unknown” distribution. Therefore we are trying to find the most “unknown” distribution containing only available information about the distribution.

In this respect we want to point out that the information which is used in Hull and White model is not complete. The non-arbitrage equations are used for mid-spreads for tranches. That may be a reason why the constraints may not have a feasible solution. We argue that additional information is available in the bid and ask prices.

Instead of mid prices, we use bid and ask prices. Denote by \( a_{ij} \) and \( \bar{a}_{ij} \) the expected net payoff of tranche \( j \) conditional on hazard rate \( i \) for ask and bid prices, respectively. Then, the no-arbitrage constraints are as follows: for ask prices the expected net payoff \( (\sum_{i=1}^{I} a_{ij} p_i) \) of each CDO tranche is nonpositive and for bid prices \( (\sum_{i=1}^{I} \bar{a}_{ij} p_i) \) is nonnegative.

We maximize Shannon entropy \( H(p) = -\sum_{i=1}^{I} p_i \ln p_i \) subject to these no-arbitrage constraints. In other words, we propose to solve the following problem:

**Problem B**

\[
\begin{align*}
\min_p & \quad -H(p) \\
\text{subject to} & \\
\text{no-arbitrage constraints} & \\
\sum_{i=1}^{I} a_{ij} p_i & \leq 0, j = 1, \ldots, J , \quad (8) \\
\sum_{i=1}^{I} \bar{a}_{ij} p_i & \geq 0, j = 1, \ldots, J , \quad (9)
\end{align*}
\]
probability distribution constraints

\[ \sum_{i=1}^{I} p_i = 1, \quad \text{(10)} \]

\[ p_i \geq 0, \quad i = 1, \ldots, I. \quad \text{(11)} \]

We want to emphasize that the set of probability distributions satisfying constraints (8), (9) is bigger than for constraints (3). Therefore, constraints (8), (9) may have a feasible solution and we may not need to introduce the deviation term (6) to the objective.

Hull and White add a smoothing term to the objective function (7). As we indicated earlier in the previous section, the optimal solution is very sensitive to the choice of \( c \) in (7) and to the number of points \( I \) on the grid (see Figures 3 and 4). Furthermore, it seems that with an increasing number of points \( I \) in the optimization Problem A, the optimal solution does not stabilize. Some kind of stabilization can be seen for \( c = 10^{-5} \) at the right bottom graph in Figure 4). But, again, it is unclear why \( c = 10^{-5} \) should be used.

Here we want to quickly mention what we did next and how we came up to the conclusion to introduce the new class of functions. We solved Problem B for different numbers of points also. We found that the shape of the optimal solution eventually stabilized. Beyond the number of points being equal to 500, the optimal solutions have a similar shape. But we saw also that some “noise” was present in the optimal distribution. To cope with that we will later define the CCC class of discrete probability distributions.

We start with the general definition of CCC class of functions (not only distributions). We say that a function belongs to this class if it is convex on the left up to some point, then concave up to a further point, and then again convex on the right. CCC is the abbreviation for convex/concave/convex. Figure 1 shows an example of CCC function.

By definition a function \( f : \mathbb{R} \to \mathbb{R} \) is convex if for any \( x_1, x_2, \lambda : \lambda \in [0, 1] \) the following inequality holds:

\[ \lambda f(x_1) + (1 - \lambda) f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2) \]

Let \( x_3 = \lambda x_1 + (1 - \lambda)x_2; \) then \( \lambda(x_2 - x_1) = x_2 - x_3 \). Therefore, there is one to one correspondence between \( \lambda \) and \( x_3 \) and the convexity property can be rewritten as follows:

for any \( x_1, x_2, x_3 \) such that \( x_1 \leq x_3 \leq x_2 \) following inequality holds:

\[ (x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2) \geq (x_2 - x_1)f(x_3). \]

With this observation we can generalize the concavity/convexity property to any set \( X \subset \mathbb{R} \) not necessarily convex, closed, etc. We say that \( f : X \to \mathbb{R} \) is convex on \( X \) if for any \( x_1, x_2, x_3 \in X \):

\[ (x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2) \geq (x_2 - x_1)f(x_3) \]

Below is the formal definition of the CCC class of functions in general case.
**Definition (general case).** Let $f : X \to R$. Then $f(x)$ belongs to CCC class if and only if there exist $w_l, w_r \in R$ such that the following inequalities hold:

1. $w_l \leq w_r$,
2. $(x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2) \geq (x_2 - x_1)f(x_3)$, for all $x_1 \leq x_3 \leq x_2 \in (-\infty, w_l] \cap X$,
3. $(x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2) \leq (x_2 - x_1)f(x_3)$, for all $x_1 \leq x_3 \leq x_2 \in [w_l, w_r] \cap X$,
4. $(x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2) \geq (x_2 - x_1)f(x_3)$, for all $x_1 \leq x_3 \leq x_2 \in [w_r, +\infty) \cap X$.

First, we define the CCC class of continuous distributions.

**Definition (continuous case).** Let $f : R \to R$ be a continuous density function of some continuous distribution. Then $f(x)$ belongs to CCC class of distributions if $f(x)$ is a CCC function.

In our model we deal with discrete distributions. Let us define the CCC class of discrete distributions.

**Definition (discrete case).** Let $f : \{d_1, \ldots, d_I\} \to [0, 1]$ be a probability measure function on a sequence of points $d_1, \ldots, d_I : d_1 < d_2 < \cdots < d_I$, i.e. $\sum_{i=1}^I f(d_i) = 1$. Then $f(x)$ belongs to CCC class of discrete distributions if the probability measure $f$ belongs to the class of CCC functions.

Clearly, $f(x)$ belongs to the CCC class if and only if the inequalities 2-4 in the definition of the CCC class of functions hold for every three consecutive points $d_{i-1}, d_i, d_{i+1}$. In other words, the following proposition holds.
Proposition 1. Let \( f : \{d_1, \ldots, d_I\} \to [0, 1] \) be a probability measure function on a sequence of points \( d_1, \ldots, d_I : d_1 < d_2 < \cdots < d_I \), i.e \( \sum_{i=1}^I f(d_i) = 1 \). Then \( f \) belongs to the CCC class if and only if there exist indices \( w_l, w_r \) such that the following inequalities hold:

1. \( 1 \leq w_l \leq w_r \leq I \),
2. \((d_{i+1} - d_i)f(d_{i-1}) + (d_i - d_{i-1})f(d_{i+1}) \geq (d_{i-1} - d_{i+1})f(d_i), \) for all \( 1 < i < w_l \),
3. \((d_{i+1} - d_i)f(d_{i-1}) + (d_i - d_{i-1})f(d_{i+1}) \leq (d_{i-1} - d_{i+1})f(d_i), \) for all \( w_l < i < w_r \),
4. \((d_{i+1} - d_i)f(d_{i-1}) + (d_i - d_{i-1})f(d_{i+1}) \geq (d_{i-1} - d_{i+1})f(d_i), \) for all \( w_r < i < I \).

The proof is obvious and we leave it to the reader.

Later, we suppose that the distance between every two consecutive points \( d_i, d_{i+1} \) is the same. In this case Proposition 1 simplifies to:

Proposition 2. Let \( f : \{d_1, \ldots, d_I\} \to [0, 1] \) be a probability measure function \( \sum_{i=1}^I f(d_i) = 1 \) on a sequence of points \( d_1, \ldots, d_I : d_1 < d_2 < \cdots < d_I \), such that the distance between every two consecutive points \( d_i, d_{i+1} \) is the same. Then \( f(x) \) belongs to CCC class if and only if there exist \( d_{w_l}, d_{w_r} \) such that the following inequalities hold:

1. \( 1 \leq w_l < w_r \leq I \),
2. \( f(d_{i-1}) + f(d_{i+1}) \geq 2f(d_i), \) for all \( 1 < i < w_l \),
3. \( f(d_{i-1}) + f(d_{i+1}) \leq 2f(d_i), \) for all \( w_l < i < w_r \),
4. \( f(d_{i-1}) + f(d_{i+1}) \geq 2f(d_i), \) for all \( w_r < i < I \).

As was mentioned earlier our goal is to assign a probability \( p_i \) to every hazard rate \( \lambda_i \) to meet the market constraints. In this case, by a discrete distribution corresponding to a vector \( (p_1, \ldots, p_I) \) we mean a probability measure function \( f : \{\lambda_1, \ldots, \lambda_I\} \to [0, 1] \) such that \( f(\lambda_i) = p_i, \) \( i = 1, \ldots, I \).

We want to solve Problem B under the additional condition that the distribution belongs to the CCC class. The CCC class can be specified in the optimization problem by linear constraints. CCC constraints “regularize” the solution by reducing “noise” and they play the same role as the smoothing term \( S(p) \) in Problem A. In this way we can avoid arbitrariness in the choice of the smoothing coefficient \( c \). Also, the increasing number of points \( \lambda_i \) on hazard rate grid does not lead to additional noise in distribution.

The drawback is that there is no economic argument allowing us to claim that the real hazard rate distribution belongs to the CCC class. It is easy to imagine situations when some future event is expected to seriously affect the hazard rates, and investors are divided into two camps with very different expectations of hazard rates. Still, the advantage is that any “noise” is being effectively filtered out.
To solve Problem B in the CCC class of distributions we use Proposition 2 to introduce CCC constraints. We want to mention again that all $\ln(\lambda_i)$ are equally spaced on the interval $[\ln(10^{-8}), \ln(100)]$ in this setting.

The CCC constraints include constraints on the left slope, right slope and hump:

**Convexity of the left slope:**

\[
\frac{p_{i-1} + p_{i+1}}{2} \geq p_i, \quad i = 2, ..., w_l - 1 ,
\]

(12)

**Concavity of the hump:**

\[
\frac{p_{i-1} + p_{i+1}}{2} \leq p_i, \quad i = w_l + 1, ..., w_r - 1 ,
\]

(13)

**Convexity of the right slope:**

\[
\frac{p_{i-1} + p_{i+1}}{2} \geq p_i, \quad i = w_r + 1, ..., I - 1 ,
\]

(14)

The points $w_1, w_2$ may vary for different discrete distributions, therefore we incorporate them into the optimization problem as variables. By adding the CCC constraints to Problem B we have the following optimization problem:

**Problem C**

\[
\min_{w_l, w_r, p} -H(p)
\]

subject to

**no-arbitrage constraints**

\[
\sum_{i=1}^{I} a_{ij} p_i \leq 0 ,
\]

(15)

\[
\sum_{i=1}^{I} \bar{a}_{ij} p_i \geq 0 ,
\]

(16)

**CCC constraints:**

***constraints on inflection points***

\[
1 \leq w_l \leq w_r \leq I ,
\]

(17)

***convexity of the left slope***
\[
\frac{p_{i-1} + p_{i+1}}{2} \geq p_i, \ i = 2, \ldots, w_l - 1 ,
\] (18)

**concavity of the hump**

\[
\frac{p_{i-1} + p_{i+1}}{2} \leq p_i, \ i = w_l + 1, \ldots, w_r - 1 ,
\] (19)

**convexity of the right slope**

\[
\frac{p_{i-1} + p_{i+1}}{2} \geq p_i, \ i = w_r + 1, \ldots, I - 1 ,
\] (20)

**probability distribution constraints**

\[
\sum_{i=1}^{I} p_i = 1 ,
\] (21)

\[
p_i \geq 0, \ i = 1, \ldots, I .
\] (22)

Let us look at a subproblem of this problem. To formulate this problem suppose that we fixed \(w_l, w_r\).

**Problem** \(C(w_l, w_r)\)

\[
\min_p \ -H(p)
\]

**subject to**

**no-arbitrage constraints**

\[
\sum_{i=1}^{I} a_{ij} p_i \leq 0 ,
\] (23)

\[
\sum_{i=1}^{I} a_{ij} p_i \geq 0 ,
\] (24)

**CCC constraints:**

**convexity of the left slope**

\[
\frac{p_{i-1} + p_{i+1}}{2} \geq p_i, \ i = 2, \ldots, w_l - 1 ,
\] (25)

**concavity of the hump**
\[
\frac{p_{i-1} + p_{i+1}}{2} \leq p_i, \quad i = w_l + 1, \ldots, w_r - 1, \quad (26)
\]

*convexity of the right slope*

\[
\frac{p_{i-1} + p_{i+1}}{2} \geq p_i, \quad i = w_r + 1, \ldots, I - 1, \quad (27)
\]

*probability distribution constraints*

\[
\sum_{i=1}^{I} p_i = 1, \quad (28)
\]

\[
p_i \geq 0, \quad i = 1, \ldots, I. \quad (29)
\]

Clearly, to solve Problem C we need to solve Problem \(C(w_l, w_r)\) for all possible pairs of integers \(w_l, w_r\) such that \(1 \leq w_l \leq w_r \leq I\), and then choose the minimum among these solutions. Theoretically, the minimum should exist, but it may not be unique. In this case, we can pick any solution with this minimum. The number of subproblems (Problem \(C(w_l, w_r)\)) to solve is the order of \(n^2\). Recall that originally we proposed to solve Problem B, but since in our experiments its solutions have a noise, we suggested to find the solution to Problem B in the CCC class of functions (Problem C). We provide a heuristic algorithm for solving Problem C. We solve at first Problem B and then a sequence of Problem \(C(w_l, w_r)\)'s for different pairs of \((w_l, w_r)\). We do not prove that this algorithm provides an exact solution for Problem C.

Here is the formal description of algorithm. Explanations are provided after the formal description.

**Algorithm:**

**Step 0. Initial optimal solution.**

- Solve Problem B and denote its solution obtained for optimization problem by \(p^*\).
- Initialize \(w_l = w_r = \text{argmax}\{p^*_i : i = 1, \ldots, I\}^{10}, k = 0, H_0 = \infty\).

**Step 1. Solve Problem \(C(w_l, w_r)\)**

- Set \(k = k + 1, \text{exit\_flag} = 0\).
- Solve Problem \(C(w_l, w_r)\) and obtain the optimal solution \(p^*_k\) and \(H_k = H(p^*_k)\).

\(^{10}\)If the maximum is not unique, the algorithm should be performed for each point in the set \(\text{argmax}\{p^*_i : i = 1, \ldots, I\}\), and then the solution with the least objective value should be chosen.
Step 2. Shifting $w_r$ to the right

- If $w_r < I$ and $H_k \leq H_{k-1}$ then set $w_r = w_r + 1$, exit_flag = 1, and go to Step 1.

Step 3. Initialization of shifting $w_l$ to the left

- If $w_l > 1$ then set $w_l = w_l - 1$.
- If $w_l = 1$ then stop the algorithm, and $p_{k-1}^*$ is an approximation of the optimal solution.

Step 4. Solve Problem $C(w_l, w_r)$ (the same as Step 1)

- Set $k = k + 1$.
- Solve Problem $C(w_l, w_r)$ and obtain the optimal solution $p_k^*$ and $H_k = H(p_k^*)$.

Step 5. Shifting $w_l$ to the left

- If $w_l > 1$ and $H_k \leq H_{k-1}$ then set $w_l = w_l - 1$, exit_flag = 1, and go to Step 4.
- If exit_flag = 1, then go to Step 1.
- If ($w_l = 1$ or $H_k > H_{k-1}$) and exit_flag = 0, then stop the algorithm, and $p_{k-1}^*$ is an approximation of the optimal point.

The idea of this algorithm is that we step-by-step change inflection points $w_l, w_r$ and solve Problem $C(w_l, w_r)$. In Step 0 we solve Problem B and obtain an optimal solution $p^*$. Then we set $w_l = w_r = \text{argmax}\{p_i^* : i = 1, \ldots, I\}$. In other words, we find the maximum component of optimal vector $p^*$ and make $w_l, w_r$ equal to its index. In Step 1 we solve Problem $C(w_l, w_r)$ with these $w_l, w_r$ and obtain the optimal point and its objective value. Then, we shift $w_r$ to the right if it is possible, making $w_r = w_r + 1$. After that we go to Step 1 and again solve Problem $C(w_l, w_r)$ to obtain the optimal point and its objective value. Then we compare this objective value with the previous one obtained in Step 1 ($H_k$ and $H_{k-1}$). This procedure stops when the new objective value is greater then the previous one ($H_k > H_{k-1}$), or $w_r = I$. In Steps 3 to 5 we run the same procedure, but now we shift $w_l$ to the left. The procedure also stops when the new objective value is larger then the previous one ($H_k > H_{k-1}$), or $w_l = 1$. If during the steps 1 through 4 the less objective value is found by shifting $w_r$ or $w_l$, then these steps are needed to be performed again. In other words, we shift the points $w_r$ and $w_l$ to reach local optimality. Finally, the algorithm returns $p_{k-1}^*$ which is considered as an optimal point. We do not prove that this algorithm provides an optimal solution to Problem C. What we observe in our case study is that this algorithm works faster than solving Problem C and provides a reasonable solution.

3. Case study
We used Portfolio Safeguard (2008) in MATLAB environment to do the case study. We posted MATLAB files to run this case study on The MathWorks website (www.mathworks.com), in the file exchange-optimization area. The files can be used for both simulating the expected cash flow matrices using the tranche quotes and solving the optimization problem using these matrices. For the case study we have considered iTraxx index with different maturities.

First, we used 5-year iTraxx tranche quotes to simulate the expected cash flow matrices. For bid prices, mid prices and ask prices we simulated different matrices \((\bar{a}_{ij})_{i=1,\ldots,I}^{j=1,\ldots,J}\), \((a_{ij})_{i=1,\ldots,I}^{j=1,\ldots,J}\), \((\tilde{a}_{ij})_{i=1,\ldots,I}^{j=1,\ldots,J}\) for \(I=100, 200, \ldots, 1,000\). The number of tranches in the iTraxx index is six, so \(J = 6\).

For particular \(i, j\) we simulated the times to default of 125 companies in the iTraxx index and the corresponding tranche cash flows 10,000 times and then took the average. As we mentioned earlier, the time to default of each company is exponentially distributed with parameter \(\lambda_j\). For simulation we used the minimum hazard rate \(\lambda_1 = 10^{-8}\), the maximum hazard rate \(\lambda_I = 100\), and the distances between \(ln(\lambda_i)\) are equal. We assumed that the tranche payments are made quarterly, the recovery rate in case of default subject to to 40% and the annual risk free rate is 4%. The reader may refer to the Hull and White (2006) to find more details on the simulation procedure. This is a quite common technique and we do not focus on it.

We solved Problem A (Hull and White (2006)) for \(I=100, 300, 500\) and 1,000 points. It should be noticed that in their recent paper Hull and White tested another approach. They used optimization procedure to find a log-t distribution fitting the distributions of hazard rates. In this paper we compared our approach to the approach by Hull and White (2006). We used six different smoothing term coefficients in Problem A to compare results. The graphs are presented in Figures 3 and 4. The distribution functions in the graphs are not the actual solution vectors. We scaled them so that the areas under the graph are equal and the horizontal axis represents \(ln(\lambda)\). These graphs can be viewed as implied densities of hazard rate distributions. The results are quite sensitive to the parameters \(c\) and \(I\).

With our approach we ran the proposed heuristic algorithm described at the end of Section 2 for 100, 200, 300, 500, 800 and 1,000 points. The entropy maximization problem can be easily solved with PSG in MATLAB environment by calling PSG ‘riskprog’ optimization subroutine. We only need to put the matrix of constraints and ‘entropyr’ as parameters for this subroutine. Figure 5 shows six hazard rate distribution graphs for the six different values of \(I\) mentioned above and how the final distribution \(\tilde{p}_1\) differs from the intermediate \(\tilde{p}_0\) which is the optimal solution for Problem B. We found that imposing CCC function constraints has not changed significantly the shape of implied density functions. Some irregularities (which we call “noise”) were streamlined.

Figure 6 compares the final distributions \(\tilde{p}_1\) for different numbers of hazard rates \((I = 100, 300, 500, 800, 1,000)\). We conducted the case study on a laptop with processor Intel Core 2 CPU @2GHz. The optimization time for Problem B varied from 0.01 sec. for \(I = 100\) to 0.06 sec. for \(I = 1,000\), for Problem C it varied from 0.36 sec. for \(I = 100\) to 600 sec. for \(I = 1,000\).
100, 200, 300, 500, 800, and 1,000) on the grid. We also scaled them so that the areas below the graphs are equal. The last three graphs are almost identical, which seems quite natural. We did not observe the similar stability in the Hull and White (2006) even with a fixed smoothing coefficient $c$.

We applied our approach to the iTraxx tranche with different contract periods: 5 years, 7 years and 10 years. It should show whether the implied copula model can be used with the homogeneity assumption, i.e. that hazard rate of the company stays the same during the whole contract period. If this approach is reasonable, we should obtain a similar distribution of hazard rates for different contract periods. We simulated the matrices of expected cash flows using the prices from Figure 2 for $I=100$ for 5, 7 and 10-year contracts. Then, we used these matrices to solve Problem B and proposed heuristics. Figure 7 represents the solutions scaled the similar way and with $\ln(\lambda)$ on the horizontal axis. The graphs are quite similar and show little dependence of the length of the contract period.

We want to point out that the analyzed data were the market quotes for 5, 7, 10-year iTraxx on December 20, 2006. At that time the credit derivatives market was flourishing and expanding very fast. We also tested this model for the data taken for the recent times when the market was very unstable.

First, we used the data for the market quotes for the 5-year iTraxx on four different dates: 10/31/07, 12/31/07, 6/30/08 and 9/30/08. The data available to us contains only the closing prices. To get the bid and ask prices we used typical bid-ask spreads for that times varying from 2% to 7% depending on the tranche. Then, using the simulation technique described in the beginning of this section, we simulated expected cash flow matrices for the bid and ask prices with the number of hazard rate grid points $I = 100$. The implied density functions were obtained by solving Problem B. Figure 8 shows corresponding graphs. The graphs show the evolution of the hazard rate distribution function over the time. We want to mention that Problem C is infeasible with the assumed bid-ask spreads.

Second, we picked the two latest dates for which we have the price information for the market quotes for 5, 7, 10-year iTraxx. The expected cash flow matrices were simulated the same way. Figure 9 shows the hazard rate distribution functions for this case.

Finally, we want to mention that the obtained hazard rate distributions can be used for the pricing of various credit risk instruments.

References


Figure 2: Market quotes for 5, 7, 10-year iTraxx on December 20, 2006. Quotes for the 0 to 3% tranche are the percent of the principal that must be paid up front in addition to 500 basis points per year. Quotes for other tranches and the index are in basis points. The data was obtained from Arnsdorf and Halperin (2007).

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Low Strike</th>
<th>High Strike</th>
<th>Bid</th>
<th>Ask</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-Dec-11</td>
<td>0%</td>
<td>3%</td>
<td>11.75%</td>
<td>12.00%</td>
</tr>
<tr>
<td>20-Dec-11</td>
<td>3%</td>
<td>6%</td>
<td>53.75</td>
<td>55.25</td>
</tr>
<tr>
<td>20-Dec-11</td>
<td>6%</td>
<td>9%</td>
<td>14.00</td>
<td>15.50</td>
</tr>
<tr>
<td>20-Dec-11</td>
<td>9%</td>
<td>12%</td>
<td>5.75</td>
<td>6.75</td>
</tr>
<tr>
<td>20-Dec-11</td>
<td>12%</td>
<td>22%</td>
<td>2.13</td>
<td>2.88</td>
</tr>
<tr>
<td>20-Dec-11</td>
<td>22%</td>
<td>100%</td>
<td>0.80</td>
<td>1.30</td>
</tr>
<tr>
<td>20-Dec-11</td>
<td>0%</td>
<td>100%</td>
<td>24.75</td>
<td>25.25</td>
</tr>
<tr>
<td>20-Dec-13</td>
<td>0%</td>
<td>3%</td>
<td>26.88%</td>
<td>27.13%</td>
</tr>
<tr>
<td>20-Dec-13</td>
<td>3%</td>
<td>6%</td>
<td>130.00</td>
<td>132.00</td>
</tr>
<tr>
<td>20-Dec-13</td>
<td>6%</td>
<td>9%</td>
<td>36.75</td>
<td>38.25</td>
</tr>
<tr>
<td>20-Dec-13</td>
<td>9%</td>
<td>12%</td>
<td>16.25</td>
<td>18.00</td>
</tr>
<tr>
<td>20-Dec-13</td>
<td>12%</td>
<td>22%</td>
<td>5.50</td>
<td>6.50</td>
</tr>
<tr>
<td>20-Dec-13</td>
<td>22%</td>
<td>100%</td>
<td>2.40</td>
<td>2.90</td>
</tr>
<tr>
<td>20-Dec-13</td>
<td>0%</td>
<td>100%</td>
<td>33.50</td>
<td>34.50</td>
</tr>
<tr>
<td>20-Dec-16</td>
<td>0%</td>
<td>3%</td>
<td>41.88%</td>
<td>42.13%</td>
</tr>
<tr>
<td>20-Dec-16</td>
<td>3%</td>
<td>6%</td>
<td>348.00</td>
<td>353.00</td>
</tr>
<tr>
<td>20-Dec-16</td>
<td>6%</td>
<td>9%</td>
<td>93.00</td>
<td>95.00</td>
</tr>
<tr>
<td>20-Dec-16</td>
<td>9%</td>
<td>12%</td>
<td>40.00</td>
<td>42.00</td>
</tr>
<tr>
<td>20-Dec-16</td>
<td>12%</td>
<td>22%</td>
<td>13.25</td>
<td>14.25</td>
</tr>
<tr>
<td>20-Dec-16</td>
<td>22%</td>
<td>100%</td>
<td>4.35</td>
<td>4.85</td>
</tr>
<tr>
<td>20-Dec-16</td>
<td>0%</td>
<td>100%</td>
<td>44.50</td>
<td>45.50</td>
</tr>
</tbody>
</table>
Figure 3: Distributions of the collateral hazard rate, as implied in 5-year iTraxx tranche spreads. The prices from the table in Figure 3 are used. The implied copula of Hull and White approach is used. The distributions were found as solutions to Problem A for numbers of variables 100 and 300, and different smoothing term coefficients $c$. 

![Graphs for different values of $c$](image-url)
Figure 4: Distributions of the collateral hazard rate, as implied in 5-year iTraxx tranche spreads. The prices from the table in Figure 3 are used. The distributions were found as solutions to Problem A for numbers of variables 500 and 1,000, and different smoothing term coefficients c.
Figure 5: Distributions of the collateral hazard rate, as implied in 5-year iTraxx tranche spreads. The prices from the table in Figure 3 are used. The distributions were found as solutions to Problem $B$, and the proposed heuristic algorithm to find a solution in the CCC class for 100, 200, 300, 500, 800, 1,000 decision variables.
Figure 6: Distributions of the collateral hazard rate, as implied in 5-year iTraxx tranche spreads. The prices from the table in Figure 3 are used. The distributions were found in the CCC class using the proposed heuristic algorithm for 100, 200, 300, 500, 800, 1,000 decision variables.
Figure 7: Distributions of the collateral hazard rate, as implied in 5, 7 and 10-year iTraxx tranche spreads. The prices from the table in Figure 3 are used. The distributions were found as solutions to Problem B (upper chart) and in the CCC class (lower chart) using the proposed heuristic algorithm for 100 decision variables.
Figure 8: Distributions of the collateral hazard rate, as implied in 5-year iTraxx tranche spreads in different dates. The distributions were found as solutions to Problem B for 100 decision variables.
Figure 9: Distributions of the collateral hazard rate, as implied in 5, 7 and 10-year iTraxx tranche spreads in two different dates. The distributions were found as solutions to Problem B for 100 decision variables.