Advanced credit portfolio modeling and CDO pricing

Ernst Eberlein[1], Rüdiger Frey[2], and Ernst August von Hammerstein[1]

[1] Department of Mathematical Stochastics, University of Freiburg, eberlein@stochastik.uni-freiburg.de
[2] Department of Mathematics, University of Leipzig, ruediger.frey@math.uni-leipzig.de

1 Introduction

Credit risk represents by far the biggest risk in the activities of a traditional bank. In particular, during recession periods financial institutions loose enormous amounts as a consequence of bad loans and default events. Traditionally the risk arising from a loan contract could not be transferred and remained in the books of the lending institution until maturity. This has changed completely since the introduction of credit derivatives such as credit default swaps (CDSs) and collaterized debt obligations (CDOs) roughly fifteen years ago. The volume in trading these products at the exchanges and directly between individual parties (OTC) has increased enormously. This success is due to the fact that credit derivatives allow the transfer of credit risk to a larger community of investors. The risk profile of a bank can now be shaped according to specified limits, and concentrations of risk caused by geographic and industry sector factors can be reduced.

However, credit derivatives are complex products, and a sound risk-management methodology based on appropriate quantitative models is needed to judge and control the risks involved in a portfolio of such instruments. Quantitative approaches are particularly important in order to understand the risks involved in portfolio products such as CDOs. Here we need mathematical models which allow to derive the statistical distribution of portfolio losses. This distribution is influenced by the default probabilities of the individual instruments in the portfolio, and, more importantly, by the *joint behaviour* of the components of the portfolio. Therefore the probabilistic dependence structure of default events has to be modeled appropriately.

In this paper we use two different approaches for modeling dependence. To begin with, we extend the factor model approach of Vasiček [32, 33] by using more sophisticated distributions for the factors. Due to their greater

This version: 14.09.2007

flexibility these distributions have been successfully used in several areas of finance (see e.g. [9, 10, 11]). As shown in the present paper, this approach leads to a substantial improvement of performance in the pricing of synthetic CDO tranches. Moreover, in the last section we introduce a dynamic Markov chain model for the default state of a credit portfolio and discuss the pricing of CDO tranches for this model.

2 CDOs: Basic concepts and modeling approaches

A *collateralized debt obligation* (CDO) is a structured product based on an underlying portfolio of reference entities subject to credit risk, such as corporate bonds, mortgages, loans or credit derivatives. Although several types of CDOs are traded in the market which mainly differ in the content of the portfolio and the cash flows between counterparties, the basic structure is the same. The originator (usually a bank) sells the assets of the portfolio to a so-called *special purpose vehicle* (SPV), a company which is set up only for the purpose of carrying out the securitization and the necessary transactions. The SPV does not need capital itself, instead it issues notes to finance the acquisition of the assets. Each note belongs to a certain loss piece or *tranche* after the portfolio has been divided into a number of them. Consequently the portfolio is no longer regarded as an asset pool but as a collateral pool. The tranches have different seniorities; the first loss piece or *equity tranche* has the lowest, followed by *junior mezzanine*, *mezzanine*, *senior* and finally *super-senior* tranches. The interest payments the SPV has to make to the buyer of a CDO tranche are financed from the cash flow generated by the collateral pool. Therefore the performance or the default risk of the portfolio is taken over by the investors. Since all liabilities of the SPV as a tranche seller are funded by proceeds from the portfolio, CDOs can be regarded as a subclass of so-called asset-backed securities. If the assets consist mainly of bonds resp. loans, the CDO is also called collateralized bond obligation (CBO) resp. collateralized loan obligation (CLO). For a *synthetic CDO* which we shall discuss in more detail below, the portfolio contains only credit default swaps. The motivation to build a CDO is given by economic reasons:

- By selling the assets to the SPV, the originator removes them from his balance sheet and therefore he is able to reduce his regulatory capital. The capital which is set free can then be used for new business opportunities.
- The proceeds from the sale of the CDO tranches are typically higher than the initial value of the asset portfolio because the risk-return profile of the tranches is more attractive for investors. This is both the result from and the reason for slicing the portfolio into tranches and the implicit collation and rebalancing hereby. Arbitrage CDOs are mainly set up to exploit this difference.

In general, CDO contracts can be quite sophisticated because there are no regulations for the compilation of the reference portfolio and its tranching or

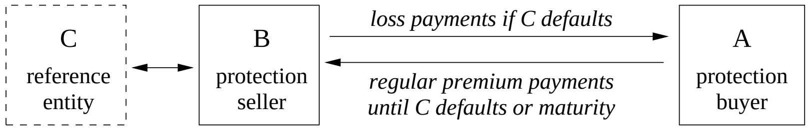

Fig. 1. Basic structure of a CDS

the payments to be made between the parties. The originator and the SPV can design the contract in a taylormade way, depending on the purposes they want to achieve. To avoid unnecessary complications, we concentrate in the following on synthetic CDOs which are based on a portfolio of credit default swaps.

2.1 Structure and payoffs of CDSs and synthetic CDOs

As mentioned before, the reference portfolio of a synthetic CDO consists entirely of *credit default swaps* (CDSs). These are insurance contracts protecting from losses caused by default of defaultable assets. The protection buyer A periodically pays a fixed premium to the protection seller B until a prespecified credit event occurs or the contract terminates. In turn, B makes a payment to A that covers his losses if the credit event has happened during the lifetime of the contract. Since there are many possibilities to specify the default event as well as the default payment, different types of CDSs are traded in the market, depending on the terms the counterparties have agreed on. The basic structure is shown in Figure 1. Throughout this article we will make the following assumptions: The reference entity of the CDS is a defaultable bond with nominal value L, and the credit event is the default of the bond issuer. If default has happened, B pays $(1-R)L$ to A where R denotes the recovery rate. On the other side A pays quarterly a fixed premium of $0.25 r_{CDS} L$ where r_{CDS} is the annualized fair CDS rate. To determine this rate explicitly, we fix some notation:

- r is the riskless interest rate, assumed to be constant over the lifetime $[0, T]$ of the CDS,
- $u(t)$ is the discounted value of all premiums paid up to time t when the annualized premium is standardized to 1,
- $G_1(t)$ is the distribution function of the default time T_1 with corresponding density $g_1(t)$ (its existence will be justified by the assumptions in subsequent sections).

The expected value of the discounted premiums (*premium leg*) can then be written as

$$PL(r_{CDS}) = r_{CDS}\, L \int_0^T u(t) g_1(t)\, \mathrm{d}t + r_{CDS}\, L\, u(T)(1 - G_1(T))\,.$$

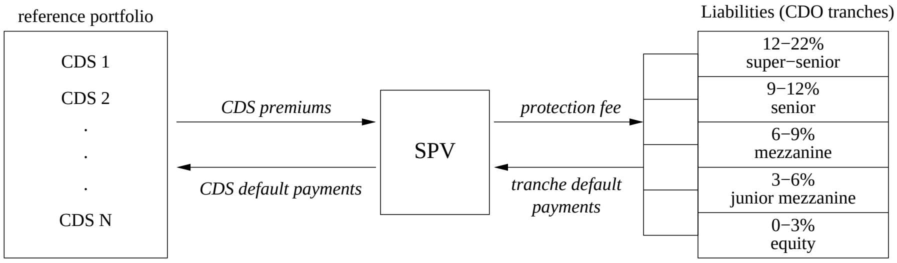

Fig. 2. Schematic representation of the payments in a synthetic CDO. The choice of the attachment points corresponds to DJ iTraxx Europe standard tranches.

The expected discounted default payment (*default leg*) is given by

$$D = (1 - R)L \int_0^T g_1(t)e^{-rt}\,\mathrm{d}t\,.$$

The no-arbitrage condition $PL(r_{CDS}) = D$ then implies

$$r_{CDS} = \frac{(1-R)\int_0^T g_1(t)e^{-rt}\,\mathrm{d}t}{\int_0^T u(t)g_1(t)\,\mathrm{d}t + u(T)(1 - G_1(T))} = \frac{D}{PL(1)}. \qquad (1)$$

To explain the structure and the cash flows of a synthetic CDO assume that its reference portfolio consists of N different CDSs with the same notional value L. We divide this portfolio in subsequent tranches. Each tranche covers a certain range of percentage losses of the total portfolio value NL defined by lower and upper *attachment points* $K_l, K_u \le 1$. The buyer of a tranche compensates as protection seller for all losses that exceed the amount of $K_l NL$ up to a maximum of $K_u NL$. On the other hand the SPV as protection buyer has to make quarterly payments of $0.25 r_c V_t$, where V_t is the notional value of the tranche at payment date t. Note that V_t starts with $NL(K_u - K_l)$ and is reduced by every default that hits the tranche. r_c is the fair tranche rate. See also Figure 2.

In recent years a new and simplified way of buying and selling CDO tranches has become very popular, the trading of *single index tranches*. For this purpose standardized portfolios and tranches are defined. Two counterparties can agree to buy and sell protection on an individual tranche and exchange the cash flows shown in the right half of Figure 2. The underlying CDS portfolio however is never physically created, it is merely a reference portfolio from which the cash flows are derived. So the left hand side of Figure 2 vanishes in this case, and the SPV is replaced by the protection buyer. The portfolios for the two most traded indices, the Dow Jones CDX NA IG and the Dow Jones iTraxx Europe, are composed of 125 investment grade US and European firms respectively. The index itself is nothing but the weighted credit default swap spread of the reference portfolio. In Sections 2.2 and 3.1 we shall derive the corresponding default probabilities. We will use market quotes for different iTraxx tranches and maturities to calibrate our models later in Sections 3.2 and 4.2.

In the following we denote the attachment points by $0 = K_0 < K_1 < \ldots < K_m \leq 1$ such that the lower and upper attachment points of tranche i are K_{i-1} and K_i respectively. Suppose for example that $(1-R)j = K_{i-1}N$ and $(1-R)k = K_i N$ for some $j < k,\ j,k \in \mathbb{N}$. Then the protection seller B of tranche i pays $(1-R)L$ if the $(j+1)^{\text{st}}$ reference entity in the portfolio defaults. For each of the following possible $k-j-1$ defaults the protection buyer receives the same amount from B. After the k^{th} default occurred the outstanding notional of the tranche is zero and the contract terminates. However, the losses will usually not match the attachment points. In general, some of them are divided up between subsequent tranches: If $\frac{(j-1)(1-R)}{N} < K_i < \frac{j(1-R)}{N}$ for some $j \in \mathbb{N}$, then tranche i bears a loss of $NL\big(K_i - \frac{(j-1)(1-R)}{N}\big)$ (and is exhausted thereafter) if the j^{th} default occurs. The overshoot is absorbed by the following tranche whose outstanding notional is reduced by $NL\big(\frac{j(1-R)}{N} - K_i\big)$. We use the following notation:

K_{i-1}, K_i	are the lower/upper attachment points of tranche i,
Z_t	is the relative amount of CDSs which have defaulted up to time t, expressed as a fraction of the total number N,
L_t^i	$= \min[(1-R)Z_t, K_i] - \min[(1-R)Z_t, K_{i-1}]$ is the loss of tranche i up to time t, expressed as a fraction of the total notional value NL,
r_i	is the fair spread rate of tranche i,
$0 = t_0$	$< \cdots < t_n$ are the payment dates of protection buyer and seller,
$\beta(t_0, t_k)$	is the discount factor for time t_k.

Remark 1. Under the assumption of a constant riskless interest rate r we would have $\beta(t_0, t_k) = e^{-rt_k}$. Since this assumption is too restrictive one uses zero coupon bond prices for discounting instead. Therefore $\beta(t_0, t_k)$ will denote the price at time t_0 of a zero coupon bond with maturity t_k.

The assumption that all CDSs have the same notional value may seem somewhat artificial, but it is fulfilled for CDOs on standardized portfolios like the Dow Jones CDX or the iTraxx Europe.

With this notation the premium as well as the default leg of tranche i can be expressed as

$$\begin{aligned} PL_i(r_i) &= \sum_{k=1}^{n} (t_k - t_{k-1}) \beta(t_0, t_k)\, r_i\, \mathrm{E}\big[\big(K_i - K_{i-1} - L_{t_k}^i\big) NL\big], \\ D_i &= \sum_{k=1}^{n} \beta(t_0, t_k) \mathrm{E}\big[\big(L_{t_k}^i - L_{t_{k-1}}^i\big) NL\big], \end{aligned} \tag{2}$$

where $\mathrm{E}[\cdot]$ denotes expectation. For the fair spread rate one obtains

$$r_i = \frac{\sum_{k=1}^{n} \beta(t_0, t_k)\big(\mathrm{E}\big[L^i_{t_k}\big] - \mathrm{E}\big[L^i_{t_{k-1}}\big]\big)}{\sum_{k=1}^{n} (t_k - t_{k-1})\beta(t_0, t_k)\big(K_i - K_{i-1} - \mathrm{E}\big[L^i_{t_k}\big]\big)} \,. \tag{3}$$

Remark 2. To get arbitrage-free prices, all expectations above have to be taken under a risk neutral probability measure, which is assumed implicitly. One should be aware that risk neutral probabilities cannot be estimated from historical default data.

Since payment dates and attachment points are specified in the CDO contract and discount factors can be obtained from the market, the remaining task is to develop a realistic portfolio model from which the risk neutral distribution of Z_t can be derived, i.e. we need to model the joint distribution of the default times $T_1, \dots, T_N$ of the reference entities.

2.2 Factor models with normal distibutions

To construct this joint distribution, the first step is to define the marginal distributions $Q_i(t) = P(T_i \le t)$. The standard approach, which was proposed in [21], is to assume that the default times T_i are exponentially distributed, that is, $Q_i(t) = 1 - e^{-\lambda_i t}$. The *default intensities* λ_i can be estimated from the clean spreads $r^i_{CDS}/(1-R)$ where r^i_{CDS} is the fair CDS spread of firm i which can be derived using formula (1). In fact, the relationship $\lambda_i \approx r^i_{CDS}/(1-R)$ is obtained directly from (1) by inserting the default density $g_1(t) = \lambda_i e^{-\lambda_i t}$ (see [22, section 9.3.3]).

As mentioned before, the CDX and iTraxx indices quote an average CDS spread for the whole portfolio in basis points (100bp = 1%), therefore the market convention is to set

$$\lambda_i \equiv \lambda_a = \frac{s_a}{(1-R)10000} \tag{4}$$

where s_a is the average CDX or iTraxx spread in basis points. This implies that all firms in the portfolio have the same default probability. One can criticize this assumption from a theoretical point of view, but it simplifies and fastens the calculation of the loss distribution considerably as we will see below. Since λ_a is obtained from data of derivative markets, it can be considered as a risk neutral parameter and therefore the $Q_i(t)$ can be considered as risk neutral probability distributions.

The second step to obtain the joint distribution of the default times is to impose a suitable coupling between the marginals. Since all firms are subject to the same economic environment and many of them are linked by direct business relations, the assumption of independence of defaults between different firms obviously is not realistic. The empirically observed occurrence of disproportionally many defaults in certain time periods also contradicts the independence assumption. Therefore the main task in credit portfolio modeling is to implement a realistic dependence structure which generates loss distributions that are consistent with market observations. The following approach goes back to [32] and was motivated by the Merton model [25].

For each CDS in the CDO portfolio we define a random variable X_i as follows:

$$X_i := \sqrt{\rho}\, M + \sqrt{1-\rho}\, Z_i, \qquad 0 \le \rho < 1, \quad i = 1, \dots, N, \tag{5}$$

where $M, Z_1, \dots, Z_N$ are independent and standard normally distributed. Obviously $X_i \sim N(0,1)$ and $\mathrm{Corr}(X_i, X_j) = \rho$, $i \neq j$. X_i can be interpreted as state variable for the firm that issued the bond which CDS i secures. The state is driven by two factors: the systematic factor M represents the macroeconomic environment to which all firms are exposed, whereas the idiosyncratic factor Z_i incorporates firm specific strengths or weaknesses.

To model the individual defaults, we define time-dependent thresholds by

$$d_i(t) := \Phi^{-1}(Q_i(t))$$

where $\Phi^{-1}(x)$ denotes the inverse of the standard normal distribution function or quantile function of $N(0,1)$. Observe that the $d_i(t)$ are increasing because so are Φ^{-1} and Q_i. Therefore we can define each default time T_i as the first time point at which the corresponding variable X_i is smaller than the threshold $d_i(t)$, that is

$$T_i := \inf\{t \ge 0 \,|\, X_i \le d_i(t)\}, \qquad i = 1, \dots, N. \tag{6}$$

This also ensures that the T_i have the desired distribution, because

$$P(T_i \le t) = P\big(X_i \le \Phi^{-1}(Q_i(t))\big) = P\big(\Phi(X_i) \le Q_i(t)\big) = Q_i(t),$$

where the last equation follows from the fact that the random variable $\Phi(X_i)$ is uniformly distributed on the interval $[0,1]$. Moreover, the leftmost equation shows that $T_i \stackrel{d}{=} Q_i^{-1}(\Phi(X_i))$, so the default times inherit the dependence structure of the X_i. Since the latter are not observable, but serve only as auxiliary variables to construct dependence, such models are termed 'latent variable' models. Note that by (4) we have $Q_i(t) \equiv Q(t)$ and thus $d_i(t) \equiv d(t)$, therefore we omit the index i in the following.

Remark 3. Instead of inducing dependence by latent variables that are linked by the factor equation (5), one can also define the dependence structure of the default times more directly by inserting the marginal distribution functions into an appropriately chosen *copula*. We do not discuss this approach here further, but give some references at the end of Section 2.3

To derive the loss distribution let A_k^t be the event that exactly k defaults have happened up to time t. From (6) and (5) we get

$$P(T_i < t\,|M) = P(X_i < d(t)\,|M) = \Phi\left(\frac{d(t) - \sqrt{\rho}M}{\sqrt{1-\rho}}\right).$$

Since the X_i are independent conditional on M, the conditional probability $P(A_k^t|M)$ equals the probability of a binomial distribution with parameters N and $p = P(T_i < t\,|M)$:

$$P(A_k^t|M) = \binom{N}{k} \Phi\Big(\frac{d(t)-\sqrt{\rho}M}{\sqrt{1-\rho}}\Big)^k \Big(1-\Phi\Big(\frac{d(t)-\sqrt{\rho}M}{\sqrt{1-\rho}}\Big)\Big)^{N-k}.$$

The probability that at time t the relative number of defaults Z_t does not exceed q is

$$\begin{aligned} F_{Z_t}(q) &= \sum_{k=0}^{[Nq]} P(A_k^t) \\ &= \int_{-\infty}^{\infty} \sum_{k=0}^{[Nq]} \binom{N}{k} \Phi\Big(\frac{d(t)-\sqrt{\rho}u}{\sqrt{1-\rho}}\Big)^k \Big(1-\Phi\Big(\frac{d(t)-\sqrt{\rho}u}{\sqrt{1-\rho}}\Big)\Big)^{N-k} \mathrm{d}P_M(u)\,. \end{aligned}$$

If the portfolio is very large, one can simplify F_{Z_t} further using the following approximation which was introduced in [33] and which is known as large homogeneous portfolio (LHP) approximation. Let $p_t(M) := \Phi\left(\frac{d(t)-\sqrt{\rho}M}{\sqrt{1-\rho}}\right)$ and G_{p_t} be the corresponding distribution function, then we can rewrite F_{Z_t} in the following way:

$$F_{Z_t}(q) = \int_0^1 \sum_{k=0}^{[Nq]} \binom{N}{k} s^k(1-s)^{N-k}\,\mathrm{d}G_{p_t}(s). \tag{7}$$

Applying the LHP approximation means that we have to determine the behaviour of the integrand for $N \to \infty$. For this purpose suppose that Y_i are independent and identically distributed (iid) Bernoulli variables with $P(Y_i = 1) = s = 1 - P(Y_i = 0)$. Then the strong law of large numbers states that $\bar{Y}_N = \frac{1}{N}\sum_{i=1}^N Y_i \to s$ almost surely which implies convergence of the distribution functions $F_{\bar{Y}_N}(x) \to \mathbb{1}_{[0,x]}(s)$ pointwise on $\mathbb{R} \setminus \{s\}$. For all $q \neq s$ we thus have

$$\sum_{k=0}^{[Nq]} \binom{N}{k} s^k(1-s)^{N-k} = P\left(\sum_{i=1}^N Y_i \le Nq\right) = P\big(\bar{Y}_N \le q\big) \underset{N\to\infty}{\longrightarrow} \mathbb{1}_{[0,q]}(s).$$

Since the sum on the left hand side is bounded by 1, by Lebesgue's theorem we get from (7)

$$\begin{aligned} F_{Z_t}(q) &\approx \int_0^1 \mathbb{1}_{[0,q]}(s)\,\mathrm{d}G_{p_t}(s) = G_{p_t}(q) = P\left(-\frac{\sqrt{1-\rho}\,\Phi^{-1}(q)-d(t)}{\sqrt{\rho}} \le M\right) \\ &= \Phi\left(\frac{\sqrt{1-\rho}\,\Phi^{-1}(q)-d(t)}{\sqrt{\rho}}\right) \end{aligned} \tag{8}$$

where in the last equation the symmetry relation $1 - \Phi(x) = \Phi(-x)$ has been used. This distribution is, together with the above assumptions, the current market standard for the calculation of CDO spreads according to equation (3). Since the relative portfolio loss up to time t is given by $(1-R)Z_t$, the expectations $\mathrm{E}\big[L^i_{t_k}\big]$ contained in (3) can be written as follows:

$$\mathrm{E}\big[L^i_{t_k}\big] = \int_{\frac{K_{i-1}}{1-R}\wedge 1}^{\frac{K_i}{1-R}\wedge 1} (1-R)\big(q-\tfrac{K_{i-1}}{1-R}\big)\mathrm{d}F_{Z_{t_k}}(q)+(K_i-K_{i-1})\big[1-F_{Z_{t_k}}\big(\tfrac{K_i}{1-R}\wedge 1\big)\big]. \tag{9}$$

2.3 Deficiencies and extensions

The pricing formula obtained from (3), (8) and (9) contains one unknown quantity: the correlation parameter ρ. This parameter has to be estimated before one can derive the fair rate of a CDO tranche. A priori it is not clear which data and which estimation procedure one could use to get ρ. In the Merton approach, defaults are driven by the evolution of the asset value of a firm. Consequently the dependence between defaults is derived from the dependence between asset values. The latter cannot be observed directly, therefore some practitioners have used equity correlations, which can be estimated from stock price data. A more direct and plausible alternative would be to infer correlations from historical default data, but since default is a rare event, this would require data sets over very long time periods which are usually not available.

With the development of a liquid market for single index tranches in the last years, a new source of correlation information has arisen: the *implied correlations* from index tranche prices. Similar to the determination of implied volatilities from option prices by inverting the Black–Scholes formula, one can invert the above pricing formula and solve numerically for the correlation parameter ρ which reproduces the quoted market price. This provides also a method to examine if the model and its assumptions are appropriate. If this is the case, the correlations derived from market prices of different tranches of the same index should coincide. However, in reality one observes a so-called *correlation smile*: the implied correlations of the equity and (super-)senior tranches are typically much higher than those of the mezzanine tranches. See Figure 3 for an example. The smile indicates that the classical model is not flexible enough to generate realistic dependence structures. This is only partly due to the simplifications made by using the LHP approach. The deeper reason for this phenomenon lies in the fact that the model with normal factors strongly underestimates the probabilities of joint defaults. This has led to severe mispricings and inadequate risk forecasts in the past. The problem became evident in the so-called *correlation crisis* in May 2005: the factor model based on normal distributions was unable to follow the movement of market quotes occuring in reaction to the downgrading of Ford and General Motors to non-investment grade.

A number of different approaches for dealing with this problem have been investigated. A rather intuitive extension to remedy the deficiencies of the

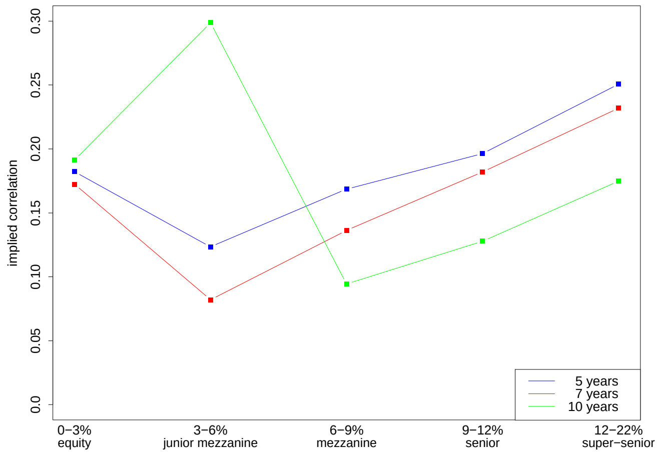

Fig. 3. Implied correlations calculated from the prices of DJ iTraxx Europe standard tranches at November 13, 2006, for different maturities T.

normal factor model which we shall exploit in Section 3, is to allow for factor distributions which are much more flexible than the standard normal ones. Different factor distributions do not only change the shape of F_{Z_t}, but also have a great influence on the so-called *factor copula* implicitly contained in the joint distribution of the latent variables. In fact, the replacement of the normal distribution leads to a fundamental modification of the dependence structure which becomes much more complex and can even exhibit *tail-dependence*. A necessary condition for the latter to hold is that the distribution of the systematic factor M is heavy tailed. This fact was proven in [24]. The first paper in which alternative factor distributions are used is [17] where both factors are assumed to follow a Student t-distribution with 5 degrees of freedom. In [19], Normal Inverse Gaussian distributions are applied for pricing synthetic CDOs, and in [1] several models based on Gamma, Inverse Gaussian, Variance Gamma, Normal Inverse Gaussian and Meixner distributions are presented. In the last paper the systematic and idiosyncratic factors are represented by the values of a suitably scaled and shifted Lévy process at times ρ and $1-\rho$.

Another way to extend the classical model is to implement *stochastic correlations* and *random factor loadings*. In the first approach which was developed in [15], the constant correlation parameter ρ in (5) is replaced by a random variable taking values in $[0,1]$. The cumulative default distribution can then be derived similarly as before, but one has to condition on both, the systematic factor and the correlation variable. The concept of random factor loadings was first published in [2]. There the X_i are defined by $X_i := m_i(M) + \sigma_i(M)Z_i$ with some deterministic functions m_i and σ_i. In the simplest case $X_i = m + (l\mathbb{1}_{\{M<e\}} + h\mathbb{1}_{\{M\geq e\}})M + \nu Z_i$ where $l, h, e \in \mathbb{R}$ are additional parameters and m, ν are constants chosen such that $\mathrm{E}[X_i] = 0$ and

$\text{Var}[X_i] = 1$. Further information and numerical details for the calibration of such models to market data can be found in [8].

As already mentioned in Remark 3, other approaches use copula models to define the dependence between the default times T_i. The concept of copulas was introduced in probability theory by Sklar [31]. A very useful and illustrative introduction to copulas and their application in risk management can be found in [22, chapter 5], for a thorough theoretical treatment, we refer to [26]. The first papers where copulas were used in credit risk models are [21] and [30]. A recent approach based on Archimedean copulas can be found in [5]. The pricing performance of models with Clayton and Marshall–Olkin copulas was investigated and compared with some other popular approaches in [7]. There the prices calculated from the Clayton copula model showed a slightly better fit to the market quotes, but they were still relatively close to those generated by the Gaussian model. The Marshall–Olkin copulas performed worse, since the deviations from market prices were greater than those of other models considered.

Alternatively, it is possible to come up with stochastic models for the dynamic evolution of the default state of the portfolio (instead of modeling just the distribution of the default times as seen from a given point in time t) and to look for dynamic models that can generate correlation skews. An example of this line of research is discussed in Section 4.

3 Calibration with advanced distributions

The factor distributions we implement to overcome the deficiencies mentioned above belong to the class of *generalized hyperbolic distributions* (GH) which was introduced in [4]. In the general case, their densities are given by

$$\begin{aligned} d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) &= a(\lambda,\alpha,\beta,\delta,\mu)\big(\delta^2+(x-\mu)^2\big)^{(\lambda-\frac{1}{2})/2} e^{\beta(x-\mu)} \\ &\quad \times K_{\lambda-\frac{1}{2}}\big(\alpha\sqrt{\delta^2+(x-\mu)^2}\big) \end{aligned} \tag{10}$$

with the norming constant

$$a(\lambda,\alpha,\beta,\delta,\mu) = \frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\,\alpha^{\lambda-\frac{1}{2}}\delta^\lambda K_\lambda(\delta\sqrt{\alpha^2-\beta^2}\,)}.$$

K_ν denotes the modified Bessel function of the third kind with index ν and $GH(\lambda,\alpha,\beta,\delta,\mu)$ the corresponding probability distribution. The influence of the parameters is as follows: $\alpha > 0$ determines the shape, $0 \le |\beta| < \alpha$ the skewness, $\mu \in \mathbb{R}$ is a location parameter and $\delta > 0$ serves for scaling. $\lambda \in \mathbb{R}$ characterizes certain subclasses and has considerable influence on the size of mass contained in the tails which can be seen from the asymptotic behaviour of the densities: $d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) \sim |x|^{\lambda-1}e^{-\alpha|x|+\beta x}$ for $|x| \to \infty$. See also Figure 4. Generalized hyperbolic distributions have already been shown to

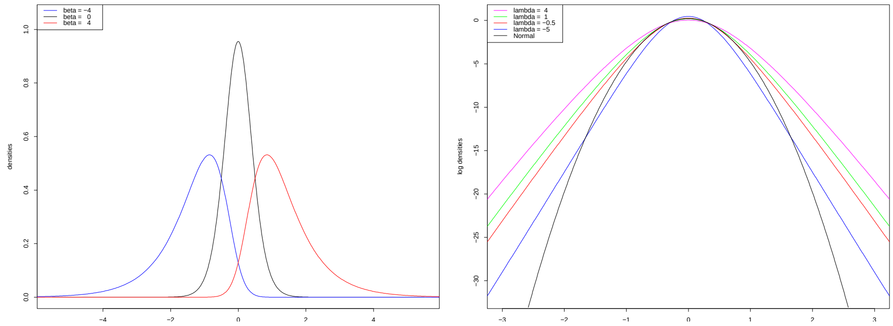

Fig. 4. Influence of the GH parameters β (left) and λ (right), where on the right hand side log densities are plotted.

be a very useful tool in various fields of mathematical finance. An overview over different applications can be found in [9]. Let us mention some special subclasses and limiting cases which are of particular interest and which we will use later to calibrate the iTraxx data:

For $\lambda = -0.5$ one obtains the subclass of *Normal Inverse Gaussian distributions* (NIG) with densities

$$d_{NIG(\alpha,\beta,\delta,\mu)}(x) = \frac{\alpha\delta}{\pi} \, \frac{K_1\big(\alpha\sqrt{\delta^2 + (x-\mu)^2}\big)}{\sqrt{\delta^2 + (x-\mu)^2}} \, e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)},$$

whereas $\lambda = 1$ characterizes the subclass of *hyperbolic distributions* (HYP) which was the first to be applied in finance in [10] and

$$d_{HYP(\alpha,\beta,\delta,\mu)}(x) = \frac{\sqrt{\alpha^2-\beta^2}}{2\alpha\delta K_1(\delta\sqrt{\alpha^2-\beta^2})} \, e^{-\alpha\sqrt{\delta^2+(x-\mu)^2}+\beta(x-\mu)}.$$

For positive λ, *Variance Gamma distributions* (VG), which were introduced in full generality in [23], can be obtained as weak limits of GH distributions. If $\lambda > 0$ and $\delta \to 0$, then the density (10) converges pointwise to

$$d_{VG(\lambda,\alpha,\beta,\mu)}(x) = \frac{(\alpha^2-\beta^2)^\lambda \, |x-\mu|^{\lambda-\frac{1}{2}}}{\sqrt{\pi}(2\alpha)^{\lambda-\frac{1}{2}}\Gamma(\lambda)} \, K_{\lambda-\frac{1}{2}}(\alpha|x-\mu|) \, e^{\beta(x-\mu)}.$$

However, if $\lambda < 0$ and $\alpha, \beta \to 0$, then (10) converges pointwise to the density of a scaled and shifted t-distribution with $f = -2\lambda$ degrees of freedom:

$$d_{t(\lambda,\delta,\mu)}(x) = \frac{\Gamma(-\lambda+\frac{1}{2})}{\delta\sqrt{\pi}\Gamma(-\lambda)} \left(1 + \frac{(x-\mu)^2}{\delta^2}\right)^{\lambda-\frac{1}{2}}, \text{ where } \Gamma(\lambda) = \int_0^\infty x^{\lambda-1}e^{-x}\mathrm{d}x.$$

For a detailed derivation of these limits and their characteristic functions, we refer to [11].

Both the skewness and especially the heavier tails increase significantly the probability of joint defaults in the factor model

$$X_i = \sqrt{\rho}\, M + \sqrt{1-\rho}\, Z_i. \tag{11}$$

In the following we assume that $M, Z_1, \dots, Z_N$ are independent as before, but $M \sim GH(\lambda_M, \alpha_M, \beta_M, \delta_M, \mu_M)$ and all Z_i are iid $\sim GH(\lambda_Z, \alpha_Z, \beta_Z, \delta_Z, \mu_Z)$ (including the above limiting cases). Thus the distribution functions of all X_i coincide. Denote the latter by F_X and the distribution functions of M and of the Z_i by F_M and F_Z, then one can derive the corresponding cumulative default distribution F_{Z_t} analogously as described in Section 2.2 and obtains

$$F_{Z_t}(q) \approx 1 - F_M\left(\frac{F_X^{-1}(Q(t)) - \sqrt{1-\rho}\, F_Z^{-1}(q)}{\sqrt{\rho}}\right). \tag{12}$$

Note that this expression cannot be simplified further as in equation (8) since the distribution of M is in general not symmetric.

Remark 4. As mentioned above, almost all densities of GH distributions possess exponentially decreasing tails, only the Student t limit distributions have a power tail. According to the results of [24], the joint distribution of the X_i will therefore show tail dependence if and only if the systematic factor M is Student t-distributed.

Further $GH(\lambda, \alpha, \beta, \delta, \mu) \xrightarrow{\mathcal{L}} N(\mu + \beta\sigma^2, \sigma^2)$ if $\alpha, \delta \to \infty$ and $\delta/\alpha \to \sigma^2$, so the normal factor model is included as a limit in our setting.

3.1 Factor scaling and calculation of quantiles

To preserve the role of ρ as a correlation parameter, we have to standardize the factor distributions such that they have zero mean and unit variance. In the general case of GH distributions we fix shape, skewness and tail behaviour by specifying α, β, λ and then calculate $\bar{\delta}$ and $\bar{\mu}$ that scale and shift the density appropriately. For this purpose we first solve the equation

$$1 = \mathrm{Var}[GH(\lambda, \alpha, \beta, \delta, \mu)] = \frac{\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} + \beta^2\frac{\delta^4}{\zeta^2}\left(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \frac{K_{\lambda+1}^2(\zeta)}{K_\lambda^2(\zeta)}\right)$$

with $\zeta := \delta\sqrt{\alpha^2 - \beta^2}$ numerically to obtain $\bar{\delta}$ and then choose $\bar{\mu}$ such that

$$0 = \mathrm{E}[GH(\lambda, \alpha, \beta, \bar{\delta}, \bar{\mu})] = \bar{\mu} + \frac{\beta\bar{\delta}^2}{\bar{\zeta}}\frac{K_{\lambda+1}(\bar{\zeta})}{K_\lambda(\bar{\zeta})}, \quad \bar{\zeta} = \bar{\delta}\sqrt{\alpha^2 - \beta^2}.$$

Since the Bessel functions $K_{n+1/2}$, $n \geq 0$, can be expressed explicitly in closed forms, the calculations simplify considerably for the NIG subclass. We have

$$\mathrm{Var}[NIG(\alpha, \beta, \delta, \mu)] = \frac{\delta\alpha^2}{(\alpha^2 - \beta^2)^{\frac{3}{2}}}\,, \quad \mathrm{E}[NIG(\alpha, \beta, \delta, \mu)] = \mu + \frac{\beta\delta}{\sqrt{\alpha^2 - \beta^2}}\,,$$

so the distribution can be standardized by choosing $\bar{\delta} = (\alpha^2 - \beta^2)^{\frac{3}{2}}/\alpha^2$ and $\bar{\mu} = -\beta(\alpha^2 - \beta^2)/\alpha^2$.
In the VG limiting case the variance is given by

$$\mathrm{Var}[\,VG(\lambda, \alpha, \beta, \mu)] = \frac{2\lambda}{\alpha^2 - \beta^2} + \frac{4\lambda\beta^2}{(\alpha^2 - \beta^2)^2} =: \sigma_{VG}^2,$$

so it would be tempting to use λ as a scaling parameter, but this would mean to change the tail behaviour which we want to keep fixed. Observing the fact that a VG distributed random variable X_{VG} equals in distribution the shifted sum of two Gamma variables, that is,

$$X_{VG} \stackrel{d}{=} \Gamma_{\lambda,\alpha-\beta} - \Gamma_{\lambda,\alpha+\beta} + \mu, \quad \text{where } d_{\Gamma_{\lambda,\sigma}}(x) = \frac{\sigma^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\sigma x} \mathbb{1}_{[0,\infty)}(x),$$

the correct scaling that preserves the shape is $\bar{\alpha} = \sigma_{VG}\,\alpha$, $\bar{\beta} = \sigma_{VG}\,\beta$. Then $\bar{\mu}$ has to fulfill

$$0 = \mathrm{E}[\,VG(\lambda, \bar{\alpha}, \bar{\beta}, \bar{\mu})] = \bar{\mu} + \frac{2\lambda\bar{\beta}}{\bar{\alpha}^2 - \bar{\beta}^2}.$$

The second moment of a Student t-distribution exists only if the number of degrees of freedom satisfies $f > 2$, so we have to impose the restriction $\lambda < -1$ in this case. Mean and variance are given by

$$\mathrm{Var}[t(\lambda, \delta, \mu)] = \frac{\delta^2}{-2\lambda - 2} \qquad \text{and} \quad \mathrm{E}[t(\lambda, \delta, \mu)] = \mu,$$

therefore one has to choose $\bar{\delta} = \sqrt{-2\lambda - 2}$ and $\bar{\mu} = 0$.

We thus have a minimum number of three free parameters in our generalized factor model, namely λ_M, λ_Z and ρ if both M and Z_i are t-distributed, up to a maximum number of seven $(\lambda_M, \alpha_M, \beta_M, \lambda_Z, \alpha_Z, \beta_Z, \rho)$ if both factors are GH or VG distributed. If we restrict M and Z_i to certain GH subclasses by fixing λ_M and λ_Z, five free parameters are remaining.

Having scaled the factor distributions, the remaining problem is to compute the quantiles $F_X^{-1}(Q(t))$ which enter the default distribution F_{Z_t} by equation (12). Since the class of GH distributions is in general not closed under convolutions, the distribution function F_X is not known explicitly. Therefore one central task of the project was to develop a fast and stable algorithm for the numerical calculation of the quantiles of X_i, because simulation techniques had to be ruled out from the very beginning for two reasons: The default probabilities $Q(t)$ are very small, so one would have to generate a very large data set to get reasonable quantile estimates, and the simulation would have to be restarted whenever at least one model parameter has been modified. Since the pricing formula is evaluated thousands of times with different parameters during calibration, this procedure would be too time-consuming. Further, the routine used to calibrate the models tries to find an extremal point by searching the direction of the steepest ascend within the parameter space in each optimization step. This can be done successfully only if the model prices depend exclusively on the parameters and not additionally on random effects. In the latter case the optimizer may behave erratically and will never reach an extremum.

We obtain the quantiles of X_i by Fourier inversion. Let $\hat{P}_X$, $\hat{P}_M$ and $\hat{P}_Z$ denote the characteristic functions of X_i, M and Z_i, then equation (11) and the independence of the factors yield

$$\hat{P}_X(t) = \hat{P}_M\big(\sqrt{\rho}\,t\big) \cdot \hat{P}_Z\big(\sqrt{1-\rho}\,t\big).$$

With the help of the inversion formula we get a quite accurate approximation of F_X from which the quantiles $F_X^{-1}(Q(t))$ can be derived. For all possible factor distributions mentioned above, the characteristic functions $\hat{P}_M$ and $\hat{P}_Z$ are well known; see [11] for a derivation and explicit formulas.

In contrast to this approach there are two special settings in which the quantiles of X_i can be calculated directly. The first one relies on the following convolution property of the NIG subclass,

$$NIG(\alpha, \beta, \delta_1, \mu_1) * NIG(\alpha, \beta, \delta_2, \mu_2) = NIG(\alpha, \beta, \delta_1 + \delta_2, \mu_1 + \mu_2),$$

and the fact that if $Y \sim NIG(\alpha, \beta, \delta, \mu)$, then $aY \sim NIG\big(\frac{\alpha}{|a|}, \frac{\beta}{a}, \delta|a|, \mu a\big)$. Thus if both M and Z_i are NIG distributed and the distribution parameters of the latter are defined by $\alpha_Z := \alpha_M\sqrt{1-\rho}/\sqrt{\rho}$ and $\beta_Z = \beta_M\sqrt{1-\rho}/\sqrt{\rho}$, then it follows together with equation (11) that $X_i \sim NIG\big(\frac{\alpha_M}{\sqrt{\rho}}, \frac{\beta_M}{\sqrt{\rho}}, \frac{\bar{\delta}_M}{\sqrt{\rho}}, \frac{\bar{\mu}_M}{\sqrt{\rho}}\big)$, where $\bar{\delta}_M$ and $\bar{\mu}_M$ are the parameters of the standardized distribution of M as described before.

In the VG limiting case the parameters α, β and μ behave as above under scaling, and the corresponding convolution property is

$$VG(\lambda_1, \alpha, \beta, \mu_1) * VG(\lambda_2, \alpha, \beta, \mu_2) = VG(\lambda_1 + \lambda_2, \alpha, \beta, \mu_1 + \mu_2).$$

Consequently if both factors are VG distributed and the free parameters of the idiosyncratic factor are chosen as follows, $\lambda_Z = \lambda_M(1-\rho)/\rho$, $\alpha_Z = \alpha_M$, $\beta_Z = \beta_M$, then $X_i \sim VG\big(\frac{\lambda_M}{\rho}, \frac{\bar{\alpha}_M}{\sqrt{\rho}}, \frac{\bar{\beta}_M}{\sqrt{\rho}}, \frac{\bar{\mu}_M}{\sqrt{\rho}}\big)$.

This stability under convolutions, together with the appropriate parameter choices for the idiosyncratic factor, was used in [19] and all models considered in [1]. We do not use this approach here because it reduces the number of free parameters and therefore the flexibility of the factor model. Moreover, in such a setting the distribution of the idiosyncratic factor is uniquely determined by the systematic factor, which contradicts the intuitive idea behind the factor model and lacks an economic interpretation.

3.2 Calibration results for the DJ iTraxx Europe

We calibrate our generalized factor model with market quotes of DJ iTraxx Europe standard tranches. As mentioned before, the iTraxx Europe index is based on a reference portfolio of 125 European investment grade firms and quotes its average credit spread which can be used to estimate the default intensity of all constituents according to equation (4). The diversification of the portfolio always remains the same. It contains CDSs of 10 firms from automotive industry, 30 consumers, 20 energy firms, 20 industrials, 20 TMTs (technology, media and telecommunication companies) and 25 financials. In each sector, the firms with the highest liquidity and volume of trade with respect to their defaultable assets (bonds and CDSs) are selected. The iTraxx portfolio is reviewed and updated quarterly. Not only companies that have defaulted in between are replaced by new ones, but also those which no longer fulfill the liquidity and trading demands. Of course, the recomposition affects future deals only. Once two counterparties have agreed to buy and sell protection on a certain iTraxx tranche, the current portfolio is kept fixed for them in order to determine the corresponding cash flows described in Section 2.1. The names and attachment points of the five iTraxx standard tranches are given in Figures 2 and 3. For each of them four contracts with different maturities (3, 5, 7 and 10 years) are available.

The settlement date of the sixth iTraxx series was December 20, 2006, so the 5, 7, and 10 year contracts mature on December 20, 2011 resp. 2013 and 2016. We consider the market prices of the latter on all standard tranches at November 13, 2006. For the mezzanine and senior tranches, these equal the annualized fair spreads r_i which can be obtained from equation (3) and are also termed *running spreads*. However, the market convention for pricing the equity tranche is somewhat different: In this case the protection buyer has to pay a certain percentage s_1 of the notional value $K_1 NL$ as an *up-front fee* at the starting time t_0 of the contract and a fixed spread of 500bp on the outstanding notional at $t_1, \dots, t_n$. Therefore the premium leg for the equity tranche is given by

$$PL_1(s_1) = s_1 K_1 NL + 0.05 \sum_{k=1}^{n} (t_k - t_{k-1}) \beta(t_0, t_k) \mathrm{E}\big[\big(K_1 - L^1_{t_k}\big) NL\big],$$

and the no-arbitrage condition $PL_1(s_1) = D_1$ then implies

$$s_1 = \frac{\sum_{k=1}^{n} \beta(t_0, t_k) \Big(\mathrm{E}\big[L^1_{t_k}\big] - \mathrm{E}\big[L^1_{t_{k-1}}\big] - 0.05(t_k - t_{k-1})\big(K_1 - \mathrm{E}\big[L^1_{t_k}\big]\big)\Big)}{K_1}. \quad (13)$$

Since the running spread is set to a constant of 500bp, the varying market price quoted for the equity tranche is the percentage s_1 defining the magnitude of the up-front fee.

We calibrate our generalized factor model by least squares optimization, that is, we first specify to which subclass of the GH family the distributions F_M and F_Z belong and then determine the correlation and distribution parameters numerically which minimize the sum of the squared differences between model and market prices over all tranches. Although our algorithm for computing the quantiles $F_X^{-1}(Q(t))$ allows us to combine factor distributions of different GH subclasses, we restrict both factors to the same subclass for simplicity reasons. Therefore in the following table and figures the expression VG, for example, denotes a factor model where M and the Z_i are variance gamma distributed. The model prices are calculated from equations (3) and (13), using the cumulative default distribution (12) resp. (8) for the normal factor model which serves as a benchmark. The recovery rate R which has a great influence on the expected losses $\mathrm{E}[L^i_{t_k}]$ according to equation (9) is always set to 40%; this is the common market assumption for the iTraxx portfolio.

One should observe that the prices of the equity tranches are usually given in percent, whereas the spreads of all other tranches are quoted in basis points. In order to use the same units for all tranches in the objective function to be minimized, the equity prices are transformed into basis points within the optimization algorithm. Thus they are much higher than the mezzanine and senior spreads and therefore react to parameter changes in a more sensitive way, which amounts to an increased weighting of the equity tranche in the calibration procedure. This is also desirable from an economical point of view since the costs for mispricing the equity tranche are typically greater than for all other tranches.

Remark 5. For the same reason, the normal factor model is usually calibrated by determining the implied correlation of the equity tranche first and then using this to calculate the fair spreads of the other tranches. This ensures that at least the equity price is matched perfectly. To provide a better comparison with our model, we give up this convention and also use least squares estimation in this case. Therefore the fit of the equity tranche is sometimes less accurate, but the distance between model and market prices is smaller for the higher tranches instead.

Our calibration results for the 5 and 7 year iTraxx tranches are shown in Figures 5 and 6. The normal benchmark model performs worst in all cases. The performance of the t model is comparable with the NIG and HYP models, whereas the VG model provides the best fit for both maturities. Since the t model is the only one exhibiting tail dependence (confer Remark 4) but does not outperform the NIG, HYP and VG models, one may conclude that this property is negligible in the presence of more flexible factor distributions. This may also be confirmed by the fact that all estimated GH parameters β_M and β_Z are different from zero which implies skewness of the factor distributions. Furthermore the parameter ρ is usually higher in the GH factor models than in the normal benchmark model, indicating that correlation is still of some importance, but has a different impact on the pricing formula because of the more complex dependence structure.

The VG model even has the potential to fit the market prices of all tranches and maturities simultaneously with high accuracy, which we shall show below. However, before that we want to point out that the calibration over different maturities requires some additional care to avoid inconsistencies when

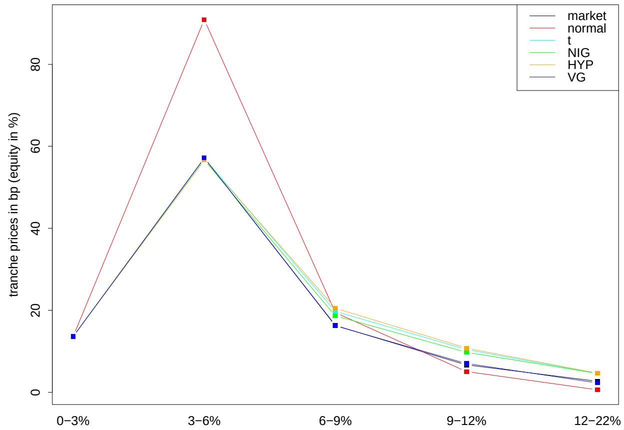

Fig. 5. Comparison of calibrated model prices and market prices of the 5 year iTraxx contracts.

calculating the default probabilities. As can be seen from Figure 7, the average iTraxx spreads s_a are increasing in maturity and by equation (4) so do the default intensities λ_a. This means that the estimated default probabilities $Q(t) = 1 - e^{-\lambda_a t}$ of a CDO with a longer lifetime are always greater than those of a CDO with a shorter maturity. While this can be neglected when concentrating on just one maturity, this fact has to be taken into account when considering iTraxx CDOs of different maturities together. Since the underlying portfolio is the same, the default probabilities should coincide during the common lifetime.

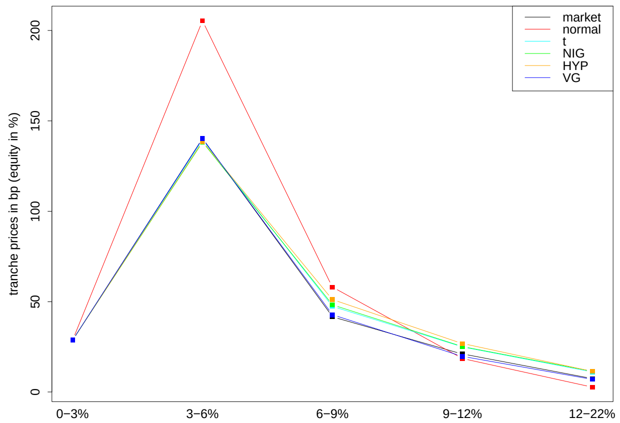

Fig. 6. Comparison of calibrated model prices and market prices of the 7 year iTraxx contracts.

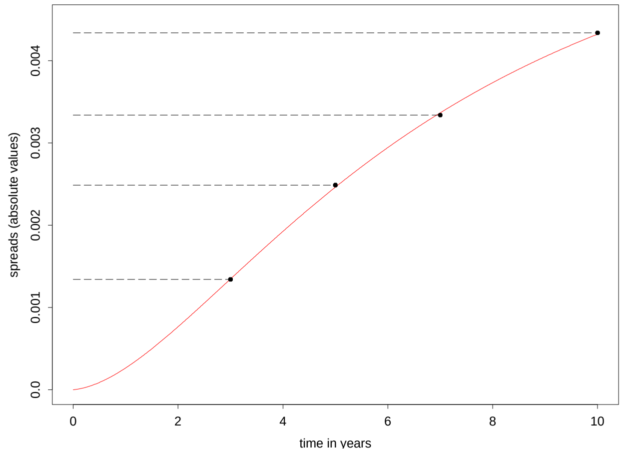

Fig. 7. Constant iTraxx spreads of November 13, 2006, and fitted Nelson–Siegel curve $\hat{r}_{NS}$ with parameters $\hat{\beta}_0 = 0.0072, \hat{\beta}_1 = -0.0072, \hat{\beta}_2 = -0.0069, \hat{\tau}_1 = 2.0950$.

To avoid these problems we now assume that the average spreads $s_a = s(t)$ are time-dependent and follow a Nelson–Siegel curve. This parametric family of functions has been introduced in [27] and has become very popular in interest rate theory for the modeling of yield curves where the task is the following: Let $\beta(0, t_k)$ denote today's price of a zero coupon bond with maturity t_k as before, then one has to find a function f (*instantaneous forward rates*) such that the model prices $\beta(0, t_k) = \exp\big(-\int_0^{t_k} f(t)\,\mathrm{d}t\big)$ approximate the market prices reasonably well for all maturities t_k. Since instantaneous forward rates cannot be observed directly in the market, one often uses an equivalent expression in terms of *spot rates*: $\beta(0, t_k) = \exp(-r(t_k)\,t_k)$, where the spot rate is given by $r(t_k) = \frac{1}{t_k}\int_0^{t_k} f(t)\,\mathrm{d}t$. Nelson and Siegel suggested to model the forward rates by

$$f_{NS(\beta_0,\beta_1,\beta_2,\tau_1)}(t) = \beta_0 + \beta_1 e^{-\frac{t}{\tau_1}} + \beta_2 \frac{t}{\tau_1}\, e^{-\frac{t}{\tau_1}}.$$

The corresponding spot rates are given by

$$r_{NS(\beta_0,\beta_1,\beta_2,\tau_1)}(t) = \beta_0 + (\beta_1 + \beta_2)\frac{\tau_1}{t}\Big(1 - e^{-\frac{t}{\tau_1}}\Big) - \beta_2 e^{-\frac{t}{\tau_1}}. \tag{14}$$

In order to obtain time-consistent default probabilities resp. intensities we replace s_a in equation (4) by a Nelson–Siegel spot rate curve (14) that has been fitted to the four quoted average iTraxx spreads, that is,

$$\lambda_a = \lambda(t) = \frac{\hat{r}_{NS}(t)}{(1-R)10000}, \tag{15}$$

and $Q(t) := 1 - e^{-\lambda(t)\,t}$. The Nelson–Siegel curve estimated from the iTraxx spreads of November 13, 2006, is shown in Figure 7. At first glance the differences between constant and time-varying spreads seem to be fairly large, but

Tranches	Market	VG	Market	VG	Market	VG
	5Y		7Y		10Y	
0–3%	13.60%	13.60%	28.71%	28.72%	42.67%	42.67%
3–6%	57.16bp	53.30bp	140.27bp	132.27bp	360.34bp	357.60bp
6–9%	16.31bp	17.19bp	41.64bp	41.83bp	105.08bp	111.17bp
9–12%	6.65bp	8.23bp	21.05bp	19.90bp	43.33bp	52.00bp
12–22%	2.67bp	3.05bp	7.43bp	7.34bp	13.52bp	18.97bp

Table 1. Results of the VG model calibration simultaneously over all maturities. Estimated parameters are as follows: $\lambda_M = 0.920$, $\alpha_M = 5.553$, $\beta_M = 1.157$, $\lambda_Z = 2.080$, $\alpha_Z = 2.306$, $\beta_Z = -0.753$, $\rho = 0.321$.

one should observe that these are the absolute values which have already been divided by 10000 and therefore range from 0 to 0.004338, so the differences in the default probabilities are almost negligible.

Under the additional assumption (15), we have calibrated a model with VG distributed factors to the tranche prices of all maturities simultaneously. The results are summarized in Table 1 and visualized in Figure 8. The fit is excellent. The maximal absolute pricing error is less than 9bp, and for the 5 and 7 year maturities the errors are, apart from the junior mezzanine

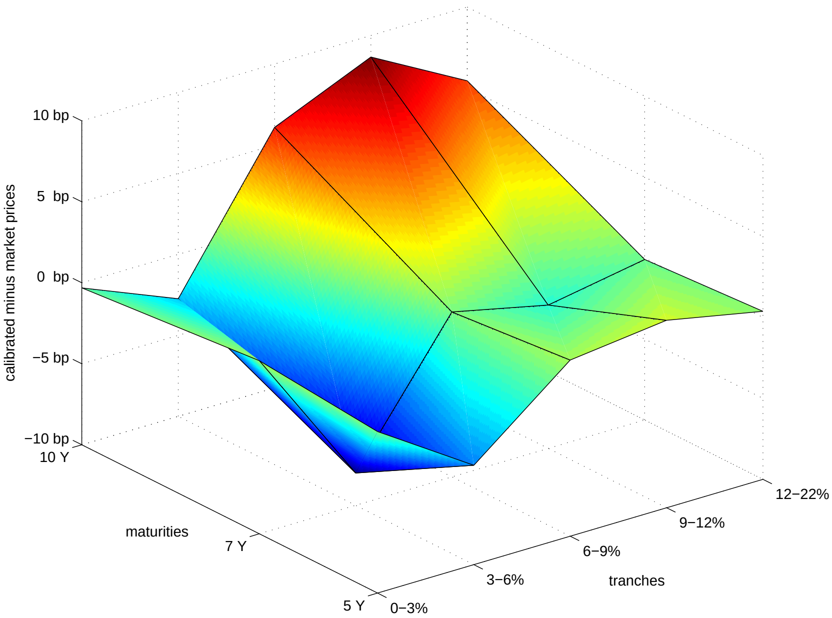

Fig. 8. Graphical representation of the differences between model and market prices obtained from the simultaneous VG calibration.

tranches, almost as small as in the previous calibrations. The junior mezzanine tranche is underpriced for all maturities, but it is difficult to say whether this is caused by model or by market imperfections. Nevertheless the overall pricing performance of the extended VG model is comparable or better than the performance of the models considered in [1, 7, 19], although the latter were only calibrated to tranche quotes of a single maturity.

Also note that this model admits a flat correlation structure not only over all tranches, but also over different maturities: all model prices contained in Table 1 were calculated using the same parameter ρ. Thus the correlation smiles shown in Figure 3 which in some sense question the factor equation (5) are completely eliminated. Therefore the intuitive idea of the factor approach is preserved, but one should keep in mind that in the case of GH distributed factors the dependence structure of the joint distribution of the X_i is more complex and cannot be described by correlation alone.

4 A dynamic Markov chain model

In this section we discuss an entirely different approach to explain observed CDO spreads, rooted more in the theory of stochastic processes. Our exposition summarizes results from [13].

4.1 The model

We begin with some notation. Given some probability space $(\Omega, \mathcal{F}, Q)$, Q the risk-neutral measure used for pricing, we define the *default indicator* of firm i at time t by $Y_{t,i} = \mathbb{1}_{\{T_i \leq t\}}$. Note that the default indicator process $Y_i = (Y_{t,i})_{t\geq 0}$ is a right continuous process which jumps from 0 to 1 at the default of firm i. The evolution of the default state of the portfolio is then described by the process $Y = (Y_{t,1}, \dots, Y_{t,N})_{t\geq 0}$; obviously, $Y_t \in S^Y := \{0,1\}^N$. We use the following notation for flipping the ith coordinate of a default state: given $y \in S^Y$ we define $y^i \in S^Y$ by

$$y_i^i := 1 - y_i \text{ and } y_j^i := y_j,\ j \in \{1, \dots, N\} \setminus \{i\}\,. \tag{16}$$

The *default history* (the internal filtration of the process Y) is denoted by $(\mathcal{H}_t)$, i.e. $\mathcal{H}_t = \sigma(Y_s : s \leq t)$. An $(\mathcal{H}_t)$-adapted process $(\lambda_{t,i})$ is called the *default intensity* of T_i (with respect to $(\mathcal{H}_t)$) if

$$Y_{t,i} - \int_0^{T_i \wedge t} \lambda_{s,i}\, \mathrm{d}s \text{ is an } (\mathcal{H}_t)\text{-martingale.}$$

Intuitively, $\lambda_{t,i}$ gives the instantaneous chance of default of a non-defaulted firm i given the default history up to time t. It is well-known that the default intensities determine the law of the marked point process (Y_t); see for instance [6] for a detailed account of the mathematics of marked point processes.

Modeling the dynamics of Y.

We assume that the default intensity of a non-defaulted firm i at time t is given by a function $\lambda_i(t, Y_t)$ of time and of the current default state Y_t. Hence the default intensity of a firm may change if there is a change in the default state of other firms in the portfolio; in this way dependence between default events can be modeled explicitly. Formally, we model the default indicator process by a time-inhomogeneous Markov chain with state space S^Y. The next assumption summarizes the mathematical properties of Y.

Assumption 1 (Markov family) *Consider bounded and measurable functions $\lambda_i : [0,\infty) \times S^Y \to \mathbb{R}_+$, $1 \le i \le N$. There is a family $Q_{(t,y)}$, $(t,y) \in [0,\infty) \times S^Y$, of probability measures on $(\Omega, \mathcal{F}, (\mathcal{H}_t))$ such that $Q_{(t,y)}(Y_t = y) = 1$ and such that $(Y_s)_{s \ge t}$ is a finite-state Markov chain with state space S^Y and transition rates $\lambda(s, y_1, y_2)$ given by*

$$\lambda(s, y_1, y_2) = \begin{cases} (1 - y_{1,i})\, \lambda_i(s, y_1), & \text{if } y_2 = y_1^i \text{ for some } i \in \{1, \dots, N\}, \\ 0 & \text{else.} \end{cases} \tag{17}$$

Relation (17) has the following interpretation: In t the chain can jump only to the set of neighbors of the current state Y_t that differ from Y_t by exactly one default; in particular there are no joint defaults. The probability that firm i defaults in the small time interval $[t, t+h)$ thus corresponds to the probability that the chain jumps to the neighboring state $(Y_t)^i$ in this time period. Since such a transition occurs with rate $\lambda_i(t, Y_t)$, it is intuitively obvious that $\lambda_i(t, Y_t)$ is the default intensity of firm i at time t; a formal argument is given in [13].

The *numerical treatment* of the model can be based on Monte Carlo simulation or on the Kolmogorov forward and backward equation for the transition probabilities; see again [13] for further information. An excellent introduction to continuous-time Markov chains is given in [28].

Modeling default intensities.

The default intensities $\lambda_i(t, Y_t)$ are crucial ingredients of the model. If the portfolio size N is large — such as in the pricing of typical synthetic CDO tranches — it is natural to assume that the portfolio has a homogeneous group structure. This assumption gives rise to intuitive parameterizations for the default intensities; moreover, the homogeneous group structure leads to a substantial reduction in the size of the state space of the model. Here we concentrate on the extreme case where the entire portfolio forms a single homogeneous group so that the processes Y_i are *exchangeable*; this simplifying assumption is made in most CDO pricing models; see also Section 2. Denote the number of defaulted firms at time t by

$$M_t := \sum_{i=1}^{N} Y_{t,i} \,.$$

As discussed in [13], in a homogeneous model default intensities are necessarily of the form

$$\lambda_i(t, Y_t) = h(t, M_t) \text{ for some } h : [0, \infty) \times \{0, \dots, N\} \to \mathbb{R}_+ . \qquad (18)$$

Note that the assumption that default intensities depend on Y_t via the number of defaulted firms M_t makes sense also from an economic viewpoint, as unusually many defaults might have a negative impact on the liquidity of credit markets or on the business climate in general. This point is discussed further in [12] and [16].

The simplest exchangeable model is the *linear counterparty risk* model. Here

$$h(t, l) = \lambda_0 + \lambda_1 l \,, \quad \lambda_0 > 0, \lambda_1 \geq 0, \, l \in \{0, \dots, N\} \,. \qquad (19)$$

The interpretation of (19) is straightforward: upon default of some firm the default intensity of the surviving firms increases by the constant amount λ_1 so that default dependence increases with λ_1; for $\lambda_1 = 0$ defaults are independent. Model (19) is the homogeneous version of the so-called looping-defaults model of [18].

The next model generalizes the linear counterparty risk model in two important ways: first, we introduce time-dependence and assume that a default event at time t increases the default intensity of surviving firms only if M_t exceeds some deterministic threshold $\mu(t)$ measuring the expected number of defaulted firms up to time t; second, we assume that on $\{l > \mu(t)\}$ the function $h(t, \cdot)$ is strictly *convex*. Convexity of h implies that large values of M_t lead to very high values of the default intensities, thus triggering a cascade of further defaults. This will be important in explaining properties of observed CDO prices below. The following specific model with the above features will be particularly useful:

$$h(t, l) = \lambda_0 + \frac{\lambda_1}{\lambda_2} \left(\exp\left(\lambda_2 \frac{(l - \mu(t))^+}{N} \right) - 1 \right), \quad \lambda_0 > 0, \, \lambda_1, \lambda_2 \geq 0 \,; \quad (20)$$

in the sequel we call (20) *convex counterparty risk* model. In (20) λ_0 is a level parameter that mainly influences credit quality. λ_1 gives the slope of $h(t, l)$ at $\mu(t)$; intuitively this parameter models the strength of default interaction for "normal" realisations of M_t. The parameter λ_2 controls the degree of convexity of h and hence the tendency of the model to generate default cascades; note that for $\lambda_2 \to 0$ (and $\mu(t) \equiv 0$) (20) reduces to the linear model (19).

The Markov property of M.

It is straightforward that for default intensities of the form (18) the process $M = (M_t)_{t \geq 0}$ is itself a Markov chain with generator given by

$$G^M_{[t]} f\,(l) = (N - l) h(t, l) \big(f(l + 1) - f(l) \big). \qquad (21)$$

In fact, since Assumption 1 excludes joint defaults, M_t can jump only to M_t+1. The intensity of such a transition is proportional to the number $N - M_t$ of surviving firms at time t as well as to their default intensity $h(t, M_t)$. This is important: since the portfolio loss satisfies $L_t = (1 - R)M_t/N$, the loss processes L^i of the individual tranches (and of course the overall portfolio loss) are given by functions of M_t, so that we may concentrate on the process M. As shown in [13] this considerably simplifies the numerical analysis of the model. Similar modeling ideas have independently been put forward in [3].

4.2 Analysis of CDO tranches

Next we turn to an analysis of synthetic CDO tranches in the context of the Markov chain model; in particular, we are interested in modeling the well-known implied correlation skew described in Section 2.3. Recall that according to equation (3), the computation of fair tranche spreads r_i boils down to evaluating the distribution of L_t^i — and hence the distribution of M_t — at the premium payment dates. The latter can be computed efficiently using the *Kolmogorov forward equations* or by simulation; see [13] for details.

The basic idea for generating correlation skews in the context of the convex counterparty risk model (20) is simple: by increasing λ_2 we can generate occasional large clusters of defaults without affecting the left tail of the distribution of L_t too much; in this way we can reproduce the spread of the mezzanine and senior CDO tranches in a way which is consistent with the observed spread of the equity tranche. In order to confirm this intuition we consider a numerical example with spread data from [17]. In Table 2 we give the CDO spreads if the convexity parameter λ_2 is varied; λ_0 and λ_1 were cali-

tranches	[0,3]	[3,6]	[6,9]	[9,12]	[12,22]	
market spreads	27.6%	168.0bp	70.0bp	43.0bp	20.0bp	
model spreads						$\sum$ abs. err.
$\lambda_2 = 0$	27.6%	223.1bp	114.5bp	61.1bp	16.9bp	120.8bp
$\lambda_2 = 5$	27.6%	194.2bp	95.7bp	54.9bp	23.3bp	67.1bp
$\lambda_2 = 8$	27.6%	172.1bp	80.0bp	46.7bp	23.7bp	21.5bp
$\lambda_2 = 8.54$	27.6%	168.0bp	77.1bp	45.1bp	23.5bp	12.7bp
$\lambda_2 = 10$	27.6%	156.9bp	69.4bp	40.7bp	22.7bp	16.7bp
state-dependent LGD $\delta_0 = 0.5$; $\delta_1 = 7.5$	27.6%	168.0bp	71.2bp	39.3bp	19.6bp	5.3bp

Table 2. CDO spreads in the convex counterparty risk model (20) for varying λ_2. λ_0 and λ_1 were calibrated to the index level of 42bp and the market quote for the equity tranche, assuming $\delta = 0.6$. For $\lambda_2 \in [8, 10]$ the qualitative properties of the model-generated CDO spreads resemble closely the behaviour of the market spreads; with state-dependent LGD the fit is almost perfect.

brated to the index level and the observed market quote of the equity tranche. The results show that for appropriate values of λ_2 the model can reproduce the qualitative behavior of the observed tranche spreads in a very satisfactory way. This observation is interesting, as it provides an explanation of correlation skews of CDOs in terms of the *dynamics* of the default indicator process. Similarly as in [2], the model fit can be improved further by considering a state-dependent loss given default of the form $\delta_t = \delta_0 + \delta_1 M_t$ with $\delta_0, \delta_1 > 0$; see again Table 2.

Comments.

Implied correlations for CDO tranches on the iTraxx Europe have changed substantially since August 2004. More importantly, the analysis presented in Table 2 presents only a "snapshot" of the CDO market at a single day. For these reasons in [13] the convex counterparty risk model (20) was recalibrated to 6 months of observed 5 year tranche spreads on the iTraxx Europe in the period 23.9.2005–03.03.2006. It turned out that the resulting parameters were quite stable over time.

In this paper we have calibrated a parametric version of the model (20) to observed CDO spreads. For an interesting nonparametric calibration procedure based on the Kolmogorov forward equation we refer to [3, 20, 29].

Dynamic Markov chain models are very useful tools for studying the practically relevant risk management of CDO tranches via dynamic hedging; we refer to the recent papers [14] and [20] for details.

References

1. H. Albrecher, S. A. Ladoucette, and W. Schoutens, *A generic one-factor Lévy model for pricing synthetic CDOs.* In: M. C. Fu, R. A. Jarrow, J.-Y. Yen, and R. J. Elliott (eds.), Advances in Mathematical Finance, Birkhäuser, Boston, 259–277 (2007)
2. L. Andersen, J. Sidenius, *Extensions to the Gaussian copula: random recovery and random factor loadings.* Journal of Credit Risk 1, 29–70 (2005)
3. M. Arnsdorf, I. Halperin, *BSLP: Markovian bivariate spread-loss model for portfolio credit derivatives.* Working paper, JP Morgan (2007)
4. O. E. Barndorff-Nielsen, *Exponentially decreasing distributions for the logarithm of particle size.* Proceedings of the Royal Society London, Series A 353, 401–419 (1977)
5. T. Berrada, D. Dupuis, E. Jacquier, N. Papageorgiou, and B. Rémillard, *Credit migration and basket derivatives pricing with copulas.* Journal of Computational Finance 10, 43–68 (2006)
6. P. Brémaud, *Point Processes and Queues: Martingale Dynamics.* Springer, New York (1981)
7. X. Burtschell, J. Gregory, and J.-P. Laurent, *A comparative analysis of CDO pricing models.* Working paper, BNP Paribas (2005)

8. X. Burtschell, J. Gregory, and J.-P. Laurent, *Beyond the Gaussian copula: stochastic and local correlation*. Journal of Credit Risk 3, 31–62 (2007)
9. E. Eberlein, *Application of generalized hyperbolic Lévy motions to finance*. In: O. E. Barndorff-Nielsen, T. Mikosch, and S. Resnick (eds.), Lévy processes: Theory and Applications, Birkhäuser, Boston, 319–337 (2001)
10. E. Eberlein, U. Keller, *Hyperbolic distributions in finance*. Bernoulli 1, 281–299 (1995)
11. E. Eberlein, E. A. von Hammerstein, *Generalized hyperbolic and inverse Gaussian distributions: limiting cases and approximation of processes*. In: R. C. Dalang, M. Dozzi, and F. Russo (eds.), Seminar on Stochastic Analysis, Random Fields and Applications IV, Progress in Probability 58, Birkhäuser, Basel, 221–264 (2004)
12. R. Frey, *A mean-field model for interacting defaults and counterparty risk*. Bulletin of the International Statistical Institute (2003)
13. R. Frey, J. Backhaus, *Credit derivatives in models with interacting default intensities: a Markovian approach*. Preprint, Universität Leipzig (2006), available from `www.math.uni-leipzig.de/~frey/publications-frey.html`.
14. R. Frey, J. Backhaus, *Dynamic hedging of synthetic CDO tranches with spread- and contagion risk*. Preprint, Universität Leipzig (2007), available from `www.math.uni-leipzig.de/~frey/publications-frey.html`.
15. J. Gregory, J.-P. Laurent, *In the core of correlation*. Risk, October, 87–91 (2004)
16. U. Horst, *Stochastic cascades, credit contagion and large portfolio losses*. Journal of Economic Behavior & Organization 63, 25–54 (2007)
17. J. Hull, A. White, *Valuation of a CDO and an n^{th} to default CDS without Monte Carlo simulation*. Journal of Derivatives 12, 8–23 (2004)
18. R. Jarrow, F. Yu, *Counterparty risk and the pricing of defaultable securities*. Journal of Finance 53, 2225–2243 (2001)
19. A. Kalemanova, B. Schmid, and R. Werner, *The normal inverse Gaussian distribution for synthetic CDO pricing*. Working paper, risklab germany (2005)
20. J.-P. Laurent, A. Cousin, and J. D. Fermanian, *Hedging default risk of CDOs in Markovian contagion models*. Working paper, ISFA Actuarial School, Université de Lyon (2007)
21. D. Li, *On default correlation: a copula function approach*. Journal of Fixed Income 9, 43–54 (2000)
22. A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management*. Princeton Series in Finance, Princeton University Press, Princeton (2005)
23. D. B. Madan, P. Carr, and E. C. Chang, *The variance gamma process and option pricing*. European Financial Review 2, 79–105 (1998)
24. Y. Malevergne, D. Sornette, *How to account for extreme co-movements between individual stocks and the market*. Journal of Risk 6, 71–116 (2004)
25. R. C. Merton, *On the pricing of corporate debt: the risk structure of interest rates*. Journal of Finance 29, 449–470 (1974)
26. R. B. Nelsen, *An Introduction to Copulas*. Lecture Notes in Statistics 139, Springer, New York (1999)
27. C. R. Nelson, A. F. Siegel, *Parsimonious modeling of yield curves*. Journal of Business 60, 473–489 (1986)
28. J. R. Norris, *Markov Chains*. Cambridge University Press, Cambridge (1997)
29. P. J. Schönbucher, *Portfolio loss and the term-structure of loss transition rates: a new methodology for the pricing of portfolio credit derivatives*. Working paper, ETH Zürich (2006)

30. P. J. Schönbucher, D. Schubert, *Copula-dependent default risk in intensity models*. Preprint, ETH Zürich and Universität Bonn (2001)
31. A. Sklar, *Fonctions de répartition à n dimensions et leurs marges*. Publications de l'Institute de Statistique de l'Université de Paris 8, 229–231 (1959)
32. O. A. Vasiček, *Probability of loss on loan portfolio*. Mimeo, KMV Corporation (1987)
33. O. A. Vasiček, *Limiting loan loss probability distribution*. Mimeo, KMV Corporation (1991)