Technische Universität München
– Center for Mathematical Sciences –
Chair of Mathematical Statistics

Time Series Models for Credit Default Swap Premia

Master's thesis

by

Márton Eifert

Supervisor: Prof. Dr. Claudia Klüppelberg

Submission date: April 5, 2013

Abstract

An idealized yet economically plausible relation between the “fair” risk-neutral premium rate of a credit default swap (CDS) and the risk-neutral default rate (intensity) of the underlying reference entity is established in this thesis. Therefore, central results on credit risk modelling within the intensity-based framework and on modelling time series by means of (Lévy-driven) continuous-time autoregressive moving-average (CARMA) processes are reviewed. In particular, the technique of recovering the noise process, i.e. isolating the background-driving Lévy process of an invertible CARMA process, is brought into focus.

Within the scope of an extensive practical model evaluation, historical premium rates of CDS on several European and North American reference issuers are used in order to determine the optimal CARMA polynomial degrees, to demonstrate the noise recovery and finally to identify the most suitable (parametric) distribution law of the approximated background-driving noise.

Zusammenfassung

In dieser Thesis wird ein idealisierter – dennoch ökonomisch plausibler – Zusammenhang zwischen der “fairen” risikoneutralen Prämie eines Credit Default Swaps (CDS) und der risikoneutralen Ausfallrate (Intensität) der dieser zugrundeliegenden Referenzobligation hergestellt. Dazu werden zentrale Erkenntnisse aus dem Bereich der Kreditrisikomodellierung mithilfe des intensitätsbasierten Ansatzes und der Zeitreihenmodellierung durch (Lévy-getriebene) zeitstetige, autoregressive Moving-Average-Prozesse (CARMA-Prozesse) vorgestellt. Insbesondere wird die Rückgewinnung des Hintergrundprozesses, d.h. der isolierten Darstellung des treibenden Lévy-Prozesses, näher beleuchtet.

Im Rahmen einer umfassenden, praktischen Auswertung werden historische Quotierungen von CDS-Prämien auf mehrere europäische und nordamerikanische Referenzschuldner herangezogen, um die für die Daten optimalen CARMA-Polynomgrade zu bestimmen, die Schätzung des Lévy-Prozesses zu veranschaulichen und schließlich die am besten passende (parametrische) Verteilung für die resultierenden Lévy-Inkremente zu identifizieren.

Acknowledgement

First and foremost, I would like to express my gratitude to my supervisor Prof. Dr. Claudia Klüppelberg, holder of the chair of Mathematical Statistics at the Technische Universität München, for her ever worthwhile advices, wide experience and extensive knowledge I could highly benefit from. It is especially thanks to her motivating nature that I have developed my taste for specific statistical topics such as time series analysis, extreme value theory and (financial) risk measurement. Not only after her deeply interesting lectures during my past studenthood but also after all kinds of face-to-face conversations, I may certainly contend that it is an honour and privilege to be taught by and have worked with her.

Emphasized mentioning and thanks is dedicated to Ph.D. Professor Helgi Tómasson, Faculty of Economics at the University of Iceland, for his very helpful notes and comments on using his `R`-package `ctarma` (publicly available online), as well as to Ph.D. Professor John P. Nolan, Department of Mathematics and Statistics at the American University, Washington D.C., whose software `STABLE` (publicly available as well) has also served very well.

Moreover, I thank my friend Alexander Müller M.Sc. for assisting me with thorough proof-reading, really constructive criticism and highly handy hints.

But most importantly, the steady backing of my parents, my brother and my friends earn my special appreciation—and more than ever my beloved girlfriend, who has patiently and tenderly helped me finalizing the past months of hard work.

Contents

List of Figures

List of Tables

Nomenclature

Global Abbreviations

càdlàg	*right continuous with left limits* (French: *continue à droite, limitée à gauche*)
CDS	*credit default swap*(*s*) (sg./pl.)
iid	*independent and identically distributed*
sde	*stochastic differential equation*

1 Introduction

$(\Omega, \mathcal{F}, \mathbb{F}, P)$	Filtered probability space endowed with the default-free market filtration $\mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}$ and probability measure P
$(\Omega, \mathcal{F}, \mathbb{F}, Q)$	Filtered probability space endowed with the default-free market filtration $\mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}$ and martingale measure Q
r	Risk-free short rate process $r := \{r_t\}_{t\geq 0}$ considered for a riskless investment (bank account)
$D(\cdot,\cdot)$	Risk-free discount factor given by $D(s,t) := \exp\{-\int_s^t r_u du\}$, $s,t \geq 0$
$\mathsf{E}[\cdot]\ (= \mathsf{E}_Q[\cdot])$	Expectation (with respect to the martingale measure Q)

2 Credit risk modelling

τ, $\mathbb{H}$	Default time τ and its natural filtration $\mathbb{H} := \{\mathcal{H}_t\}_{t\geq 0}$ (default filtration)
$(\Omega, \mathcal{G}, \mathbb{G}, Q)$	Filtered probability space endowed with the enlarged filtration $\mathbb{G} := \mathbb{F} \vee \mathbb{H}$ and martingale measure Q
F, $\bar{F}$	$\mathbb{F}$-conditional default probability $F := \{F_t\}_{t\geq 0}$ and survival probability $\bar{F} := \{\bar{F}_t\}_{t\geq 0}$ of the default time τ
Γ, γ	$\mathbb{F}$-hazard process $\Gamma := \{\Gamma_t\}_{t\geq 0}$ and $\mathbb{F}$-intensity process $\gamma := \{\gamma_t\}_{t\geq 0}$ of the default time τ
Π, $\tilde{\Pi}$	Risk-neutral price process of a defaultable ($\tilde{\Pi} := \{\tilde{\Pi}_t\}_{t\geq 0}$) and a default-free ($\Pi := \{\Pi_t\}_{t\geq 0}$) contingent claim
$\tilde{\Pi}^R_{\tilde{C}}$	Risk-neutral price process of a defaultable bond continuously paying the coupons $\tilde{C} := \{\tilde{C}_t\}_{t\geq 0}$ with recovery fraction $R \in [0,1]$ and unit notional amount $\tilde{N} := 1$
Π_C	Risk-neutral price process of a default-free bond continuously paying the coupons $C := \{C_t\}_{t\geq 0}$ with unit notional amount $N := 1$
Y, $\tilde{Y}$	Par yield of a defaultable ($\tilde{Y} := \{\tilde{Y}_t\}_{t\geq 0}$) and a default-free ($Y := \{Y_t\}_{t\geq 0}$) bond, continuously paying coupons
S^*	Par yield spread $S^* := \{S^*_t\}_{t\geq 0}$ of a defaultable continuously paying coupon bond over a default-free (otherwise equivalent) bond, given by $S^*_t := \tilde{Y}_t - Y_t$, $t \geq 0$
C^*	Par premium $C^* := \{C^*_t\}_{t\geq 0}$ of a (continuously paying) CDS

3 Continuous-time linear processes

L, W	(Two-sided) Lévy process $L := \{L_t\}_{t\in\mathbb{R}}$ and (two-sided) Brownian motion $W := \{W_t\}_{t\in\mathbb{R}}$

$\Delta L^{(h)}$ — Discretely sampled increments of the Lévy process L (of step size $h > 0$) given by $\Delta L_t^{(h)} := L_{th} - L_{(t-1)h}$, $t \in \mathbb{R}$

B^j, D^j — Discrete-time (B^j) and continuous-time (D^j) backshift operators of degree $j \in \mathbb{N}$

γ_y, ρ_y — Autocovariance function (ACVF) and autocorrelation function (ACF) of a (covariance) stationary process y

ϕ, ϑ — Autoregressive (ϕ) and moving-average (ϑ) polynomials of a discrete-time ARMA process

α, β — Autoregressive (α) and moving-average (β) polynomials of a continuous-time ARMA (CARMA) process

$\mathbf{A}$, $\mathbf{B}$ — Companion matrices of a CARMA process with characteristic polynomials α, β given by its state-space representation ($\mathbf{A}$) and of the inverted state equation ($\mathbf{B}$)

g, $g_{\mathbf{x}}$ — Real-valued kernel function of a stationary CARMA process and the vector-valued kernel function of its state vector process $\mathbf{x}$

4 Appropriate models for CDS premia

$\mathbb{T}$ — Finite time index set $\mathbb{T} := \{1, ..., T^*\}$ of discrete observations for some $T^* \in \mathbb{N}$

$\hat{y}$ — Discrete observations $\hat{y} := \{\hat{y}_t\}_{t \in \mathbb{T} \setminus \{1\}}$ of (first-order) log-differenced CDS premia given by $\hat{y}_t := \log C_t^* - \log C_{t-1}^*$, $t \in \mathbb{T} \setminus \{1\}$

ξ — Parameter vector $\xi := (\alpha_1, ..., \alpha_p, \beta_0, ..., \beta_q, \sigma)^\top$ of a CARMA(p,q) process; the admissible set is denoted by $\Xi \subset \mathbb{R}^{p+q+1} \times (0, \infty)$

θ — Parameter vector θ of the respective parametric distribution family with d parameters; the admissible set is denoted by $\Theta \subset \mathbb{R}^d$

ℓ — Likelihood function of the respective parameter vector ξ or θ

BIC	Bayesian information criterion (based on the log-likelihood $\log \ell$)
$AICC$	Akaike information criterion corrected for finite sample sizes (based on the log-likelihood $\log \ell$)

Preface

> *"[...] however, the proper test of a theory is not the realism of its assumptions but the acceptability of its implications [...]"* (Sharpe (1964))

Among the spectrum of various financial risks that a participant on capital markets—like a bank or insurance company—has to bear, *credit risk* is certainly one of the most important. It encompasses several hazards such as a company's partly or total failure or delay to comply promised payments contractually agreed with its counterparty. When risks like these eventuate, we call these events *default*, *bankruptcy* or *insolvency*. Actually, financial instruments traded on capital markets are always defaultable. For example, depending on the credit quality of (the issuer of) a corporate or governmental bond, the chance that the the total notional amount, including interests, is repaid to the investor at the bond's maturity may be higher or lower compared to other issuers—however, in most cases it may be doubted that the probability is really 100%.

Economists and mathematicians have studied the magnitude, intensity, probability as well as causes and impacts of defaults for more than 50 years. During this period, the development of theoretical fundamentals for describing and modelling the structure of credit risk always went hand in hand with the evolution and growth of global financial markets. With *credit default swaps* (*CDS*), the early 1990s have spawned probably the most influential, but also most quickly inflating credit derivatives. Primarily invented for protection against the default of a third party (like insurance contracts)—however, only available for over-the-counter trading—CDS rapidly evolved into mighty instruments for speculation during the recent 10–15 years as they are allowed to be bought (sold) in arbitrary amount without the necessity for the buyer (seller) to have any risky position in the reference entity (unlike classical insurance contracts). Regulators are only lagging behind with baffling, limiting and standardizing the capabilities of these instruments. Despite the loss in popularity that CDS suffered in recent years, as they are regarded as to be responsible for (or at least to have significantly fueled) the subprime mortgage crisis in the United States during 2004–2008, their influence in the current European sovereign debt crisis since 2010 has still been large, especially for hedging against (but also speculating on) rising or falling default risk of financial institutions and governments across Europe.

From a modelling point of view, CDS also exhibit the key advantages of being able to isolate credit risk—which in other financial securities is often connected with various different risk factors such as interest rate risk or liquidity risk: The premium of a CDS contract serves as a good quantifier of pure risk of the corresponding reference issuer's insolvency, because market participants raise (reduce) the premia they charge for newly issued contracts as soon as their expectations about his creditworthiness change into the negative (positive).

This thesis involves mathematical modelling aspects for describing the valuation of credit-risky financial instruments with certain contingent claim structures and illustrates the concepts especially by means of bonds and CDS. Furthermore and what is of greater interest, an introduction to continuous-time extensions to classical (i.e. discrete) linear time series modelling is given in order to provide useful and practical tools for describing the temporal patterns that are typical for CDS. The gap is then bridged by a very idealized framework established in this thesis in which the (risk-neutral) premium dynamics can be approximated with a very simple intensity-based expression.

This thesis is organized as follows: Chapter 1 gives a brief introduction into key concepts of risk-neutral valuation of contingent claims without default risk, in order to provide a basis for Chapter 2 in which the valuation concept is extended to account for credit risk in a certain manner by the hazard process (in general) or intensity-based approach (in particular). The theory described there consists largely of the results of Bielecki, Jeanblanc and Rutkowski arisen from many of their cooperative works (references given in the chapter) during the late 1990s through the early 2000s. Two credit risky instruments—bonds and CDS—are chosen representatives by means of which the theory is illustrated in a more concrete context.

Chapter 3 is a compact excursus into the world of modelling continuous-time linear time series, which will be used for applications in the practical part of this thesis. Its theory is mainly based on results of Brockwell and Davis, who are also authors of seminal discrete time series textbooks (references given in the chapter). They extend the theory of well-known and widespread discrete ARMA (autoregressive moving-average) models to the continuous time domain, resulting in the many-faceted class of (Lévy-driven) CARMA models. As a central aim of this thesis, methods of and conditions for recovering the background-driving Lévy process from (the sampled version of) a CARMA process are introduced.

After having presented the ideas of the Chapters 2 and 3, we will have a stochastic process observable and sufficiently appropriate for time series analysis on the one hand (the CDS premium process) and the necessary tools on the other hand (CARMA processes).

The practical part of this thesis, Chapter 4, describes the subsequent working stages that we were following during the phase of data analysis: from data acquisition and quality checks to finding and estimating appropriate models out of the CARMA class as well as recovering the background-driving noise process. This recovered noise process is put to further analysis in order to identify a suitable parametric distribution.

Chapter 5 concludes this thesis with an outlook to further topics and a brief summary.

Chapter 1.

Introduction

For introductory purposes, some key concepts of mathematical finance are briefly reviewed in order to establish a basis for the *risk-neutral valuation* of default-free claims. In Chapter 2, this will be extended to defaultable claims in a special modelling framework.

1.1. The default-free market

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space endowed with the filtration $\mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}$ satisfying the "usual conditions" that is, $\mathbb{F}$ is assumed to be complete and right-continuous. The probability measure P is referred to as the *physical measure*. In this thesis, $\mathbb{F}$ is representing the *default-free market information* which will become clear in the next chapter, when we add credit risk into the setting.

We define the (*default-free*) *money market account* modelled by a stochastic process $S^{(0)} := \{S_t^{(0)}\}_{t\geq 0}$ solving the stochastic differential equation (sde)

$$dS_t^{(0)} = S_t^{(0)} r_t dt,$$

where the positive and $\mathbb{F}$-adapted stochastic process $r := \{r_t\}_{t\geq 0}$ is called *short rate*. It is well-known that this sde has the solution

$$S_t^{(0)} = \exp\left\{\int_0^t r_s ds\right\}, \quad t \geq 0, \tag{1.1}$$

assuming $S_0^{(0)} = 1$ without loss of generality.

Suppose our market contains $N \geq 1$ different assets in addition to the money market account. Stating the price $S^{(i)}$ of some asset $i \in \{0, ..., N\}$ relatively to the price $S^{(j)}$ of another asset $j \in \{0, ..., N\}$ is a very general form of *discounting*. In that context, asset j is called *numéraire*, *reference asset* or simply *basis*. We will focus on the risk-free money market account as the numéraire of our choice ($j = 0$) for specific reasons clarified after the following definition.

Definition 1.1. (Discount factor and discounted price process) Let $S^{(i)} := \{S_t^{(i)}\}_{t \geq 0}$ denote the price process of asset $i \in \{0, ..., N\}$. Then:

(i) For some fixed $s \geq 0$, the process $\{D(s,t)\}_{t \geq 0}$ defined by

$$D(s,t) := \frac{S_s^{(0)}}{S_t^{(0)}} = \exp\left\{-\int_s^t r_u du\right\}, \quad t \geq 0,$$

is called *discount process* or *discount factor* from the viewpoint of time s.

(ii) The process $\bar{S}^{(i)} := \{\bar{S}_t^{(i)}\}_{t \geq 0}$ defined by

$$\bar{S}_t^{(i)} := \frac{S_t^{(i)}}{S_t^{(0)}} = D(0,t)\, S_t^{(i)}, \quad t \geq 0,$$

is called *discounted price process* of asset i.

Note that, in the special case $i = 0$, the discounted price process of the bank account is always $\bar{S}_t^{(0)} = D(0,t)\, S_t^{(0)} = S_0^{(0)} = 1$ for every $t \geq 0$. For $s \leq t$, the discount factor $0 < D(s,t) \leq D(t,t) = 1$ serves as the present value of one monetary unit at time s. For $s > t$, one speaks of *compounding* since $D(s,t) = D(t,s)^{-1} > 1$ is interpreted as the future worth of a risk-free investment of one monetary unit at time t.

Furthermore, observe that we have $D(s,t)\, D(t,T) = D(s,T)$ for every $0 \leq s \leq t \leq T$. We will frequently make use of these simple rules lateron.

Definition 1.2. (Risk-neutral pricing measure) Let Q denote a probability measure on $(\Omega, \mathcal{F})$, equivalent to P (in symbols $Q \sim P$), i.e. every P-nullset is a Q-nullset and vice versa. Then Q is called *martingale measure*, or *risk-neutral pricing measure* if for every asset $i \in \{0, ..., N\}$ the discounted price process $\bar{S}^{(i)}$ is an $\mathbb{F}$-martingale with respect to Q, i.e. $\mathsf{E}_Q[\bar{S}_t^{(i)} \mid \mathcal{F}_s] = \bar{S}_s^{(i)}$ for every $0 \leq s \leq t$ and every $i \in \{0, ..., N\}$.

Remark 1.3. It is not trivially guaranteed that an equivalent martingale measure Q always exists if we start modelling from the physical probability measure P. However, throughout the rest of this thesis, we take the existence of such a measure Q as our starting point, especially in order to ensure that the market is *arbitrage-free*.[1]

From now on, we turn to the risk-neutral world and work henceforth on the probability space $(\Omega, \mathcal{F}, \mathbb{F}, Q)$.

Notations 1.4. (Abbreviations)

(i) Whenever there is no risk of confusion, we omit the subscripted symbol Q when taking expectations $\mathsf{E}_Q[\cdot]$ and write $\mathsf{E}[\cdot]$ equivalently.

(ii) For properties holding Q-almost surely, we will write "Q-a.s." or simply "a.s.".

(iii) For measurability of some (real-valued) random variable X with respect to a σ-field $\mathcal{A}$ we use the notation $X \in \mathcal{A}$ for short.

1.2. Risk-neutral valuation in the default-free case

A debt claim arising from a payment X whose amount is possibly unknown a priori but depending on contractually agreed events (or realizations of stochastic processes adapted to the market filtration $\mathbb{F}$) is called *contingent claim*. Determining the fair price of such a contingent claim at some time $s \geq 0$ in a risk-neutral world means taking expectations of (functionals of) X conditional on the information $\mathcal{F}_s$ with respect to the pricing measure Q.

Using that each $\bar{S}^{(i)}$ is a martingale under Q, we obtain

$$S_s^{(i)} = \frac{\bar{S}_s^{(i)}}{D(0,s)} = \mathsf{E}\left[\frac{\bar{S}_t^{(i)}}{D(0,s)} \mid \mathcal{F}_s\right] = \mathsf{E}\left[\frac{D(0,t)}{D(0,s)} S_t^{(i)} \mid \mathcal{F}_s\right] = \mathsf{E}\left[D(s,t) S_t^{(i)} \mid \mathcal{F}_s\right] \tag{1.2}$$

for each $0 \leq s \leq t$ and every $i \in \{0, ..., N\}$. This *natural martingale pricing relation* (1.2) is summarized in the following result on the pricing of arbitrary default-free contingent claims:

[1] Formally, a *no-arbitrage condition* can be phrased as follows: There exists no such portfolio with present value zero today ($t = 0$), that has a non-negative value almost surely at a later time $t > 0$ and a positve value with probability greater than zero. This condition even ensures the *equivalence* of the absense of arbitrage and the uniqueness of a martingale measure. For a comprehensive introduction to arbitrage pricing theory in the context of mathematical finance, we refer to the seminal textbook of Föllmer and Schied (2011), Chapter 1.

Proposition 1.5. (Risk-neutral valuation of default-free contingent claims) *For some starting date $s \geq 0$ and a time interval $[s, s+T]$ of length $T > 0$, denote a contingent claim due at time $s+T$ by an $\mathcal{F}_{s+T}$-measurable integrable random variable X. Then, its risk-neutral price process $\Pi := \{\Pi_t\}_{t\geq 0}$ is given by*

$$\Pi_t = \mathsf{E}\left[D(t, s+T)\, X \mid \mathcal{F}_t\right] = \mathsf{E}\left[e^{-\int_t^{s+T} r_u du}\, X \mid \mathcal{F}_t\right], \quad t \geq 0. \tag{1.3}$$

Note that Π is indeed a martingale on $[s, s+T]$ with $\Pi_{s+T} = X$ and can therefore itself be considered as a discounted price process.

In the next chapter, we generalize this formula in a framework extended by credit risk. Both theory and applications of credit risk modelling are main subject of Chapter 2.

Chapter 2.

Credit risk modelling

For pricing securities within a complete, arbitrage-free market without default risk we need to calculate expectations of some discounted payoff structure with respect to the risk-neutral probability measure Q, conditional on the default-free market information $\mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}$, as was briefly shown in the introduction. The term "without default risk" particularly refers to the terminal payoff at a prespecified maturity date $s+T$ represented by a (from the creditor's point of view) $\mathcal{F}_{s+T}$-measurable, integrable random variable X, without taking into account the chance of a *default event* (or some more general *credit event*, for instance, the obligor's fail to meet some due payment.). For a defaultable financial instrument living on the finite time horizon $[s, s+T]$, the terminal payoff depends on whether the obligor has defaulted within this interval or not. In the simplest case, the value of the claim immediately reduces to zero as soon as default has occurred at some time before $s+T$. That is, the payoff should be of the form $X\mathbb{1}_{\{\tau>s+T\}}$, where τ denotes the random default time of the obligor. Additionally, recovery payments, only payed if and when default has occurred before or at $s+T$, may also be regarded, such that the complete payoff structure should be of the form

$$X\mathbb{1}_{\{\tau>s+T\}} + R\,\mathbb{1}_{\{\tau\leq s+T\}}$$

provided that default has not already occurred before s.

This contingency on credit risk requires additional modelling techniques introduced in the following such that we will be able to price defaultable contingent claims, i.e. contingent claims that might not be (fully) met due to some failure to pay on the obligor's side. To this end, there are generally two main paths we can follow: the *stuctural* or the *reduced-form* approach. Structural models make use of fundamental economic or company-specific business variables to define the default time τ endogenously as the firm value's first time to exceed a certain (deterministic or stochastic) threshold, called *default point*. The firm value is calculated as the present value of the company's assets, whereas the default point is structurally depending on the company's liabilities. The first and most famous models date back to Merton (1974) as well as Black and Cox (1976). In contrast to that, reduced-form models introduce a rather exogenous definition of τ; due the lack of easily comprehensible economic foundations as in the former approach, the latter is more sophisticated yet more involved. However, from the modelling point of view, the reduced-form approach brings along a wider range of possibilities. For this reason, this thesis is devoted to an extensive overview on this approach—while the structural approach will not be further treated.

Within the reduced-form framework, the predominantly and most often used terms of the *hazard process* and its associated *intensity* (or *hazard rate*) *process* are introduced and at the end, a risk-neutral valuation formula similar to (1.3) of Proposition 1.5 is established additionally accounting for defaultability.

Afterwards, the theoretical basis is explained by means of two popular examples of financial instruments with defaultable claims, in particular *bonds* and *credit default swaps* (*CDS*). In this context, important notions such as *yield* (*spread*) or *credit spread* are conceptualized in order to get the reader familiar with the key object that we are exploring in further detail in Chapter 4: the fair premium of a CDS.

2.1. The reduced-form approach

In this setting, the default event is introduced as the first jump of some indicator process and the required calculation methods for pricing, that is, taking expectations w.r.t. to an enlarged filtration $\mathbb{G}$, are derived from the properties of this counting process.[1]

2.1.1. Extension of the market filtration

Recall that the filtration $\mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}$ of the probability space $(\Omega, \mathcal{F}, \mathbb{F}, Q)$ previously introduced in Chapter 1 describes the flow of information over time that can be extracted from observable default-free market data such as the short rate process r. What cannot

[1] The following results are mainly based on the textbook of Bielecki and Rutkowski (2002), Chapters 4–6 and 8.

be extracted from $\mathbb{F}$ is the information whether (and if so, when) an obligor, e.g. a specific company or country, has defaulted. To this end, we first introduce an enlarged σ-field $\mathcal{F} \subseteq \mathcal{G}$ and an enlarged filtration $\mathbb{G} := \{\mathcal{G}_t\}_{t\geq 0}$ with $\mathcal{F}_t \subseteq \mathcal{G}_t$ for every $t \geq 0$, and then extend the probability space $(\Omega, \mathcal{F}, \mathbb{F}, Q)$ to $(\Omega, \mathcal{G}, \mathbb{G}, Q)$. Hence, valuation of credit-risky contingent claims especially means taking expectations of the discounted terminal payoff conditional on the extended filtration $\mathbb{G}$. To model the dynamics of default information, we fill the "gap" between $\mathbb{F}$ and $\mathbb{G}$ by means of the following terms.

Definition 2.1. (Default time and default indicator) Consider a non-negative random variable τ on $(\Omega, \mathcal{G})$ with properties $Q(\tau = 0) = 0$ and $Q(\tau > t) > 0$ for any $t > 0$. Then τ is called *default time* and the right-continuous increasing process $H := \{H_t\}_{t\geq 0}$ defined by $H_t := \mathbb{1}_{\{\tau \leq t\}}$, $t \geq 0$, is called the (*associated*) *default indicator process.*

Additionally, we set up the following notations to complete the extended framework.

Notations 2.2. (Default filtration and enlarged filtration)

(i) The *default filtration* $\mathbb{H} := \{\mathcal{H}_t\}_{t\geq 0}$ denotes the natural filtration of the indicator process H, that is $\mathcal{H}_t := \sigma(H_u : 0 \leq u \leq t) = \sigma(\{\tau \leq u\} : 0 \leq u \leq t)$ indicates whether (and if so, when) τ happened at some time before $t \geq 0$.

(ii) The filtrations $\mathbb{H}$ and $\mathbb{F}$ are now considered as "complementary" sub-filtrations of $\mathbb{G}$ in the sense that $\mathcal{G}_t := \mathcal{F}_t \vee \mathcal{H}_t := \sigma(\mathcal{F}_t \cup \mathcal{H}_t)$ for every $t \geq 0$ (we refer to this in symbols as $\mathbb{G} = \mathbb{F} \vee \mathbb{H}$).[2] All filtrations are assumed to satisfy the usual conditions.

Thus, all information available to the investor is now represented by $\mathbb{G}$, containing default-free ($\mathbb{F}$) and default information ($\mathbb{H}$).

Remark 2.3. Note that, by construction, $\mathbb{H}$ is the smallest possible filtration such that τ is a stopping time. Since $\mathbb{H} \subset \mathbb{G}$, τ is a $\mathbb{G}$-stopping time as well but not necessarily an $\mathbb{F}$-stopping time. We are not considering the case $\mathbb{H} \subset \mathbb{F}$ (i.e. $\mathbb{G} = \mathbb{F}$) where τ is an $\mathbb{F}$-stopping time by nature. For more details, see Bielecki and Rutkowski (2002), Chapters 5–6.

Furthermore, we will make use of the following simple finding without proof for verifications of later results.

[2] Of course, the term "complementary" shall not be confused with the classical set theoretical term "disjoint". Given $\mathbb{H}$ and $\mathbb{G}$, the equaility $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$ does not necessarily yield a unique filtration $\mathbb{F}$ (and we do not claim that) nor do $\mathbb{H}$ and $\mathbb{F}$ have to be independent.

Lemma 2.4. *For given $t \geq 0$ and any event $A \in \mathcal{G}_t$ one can find an event $B \in \mathcal{F}_t$ such that the equality $A \cap \{\tau > t\} = B \cap \{\tau > t\}$ holds true. Thus, for any random variable $X \in \mathcal{G}_t$ we find a random variable $Z \in \mathcal{F}_t$ such that*

$$\mathbb{1}_{\{\tau>t\}} X = \mathbb{1}_{\{\tau>t\}} Z, \quad t \geq 0. \tag{2.1}$$

Hence, every $\mathcal{G}_t$-measurable random variable concides with an $\mathcal{F}_t$-measurable random variable on $[0, \tau]$. Within the scope of this thesis, we focus on the probably most frequently applied class of reduced-form credit risk models, namely the hazard process approach.

2.1.2. Hazard process approach

First, let $F := \{F_t\}_{t\geq0}$ and $\bar{F} := \{\bar{F}_t\}_{t\geq0}$, defined for $t \geq 0$ by

$$F_t := \mathsf{E}\left[\mathbb{1}_{\{\tau\leq0\}} \mid \mathcal{F}_t\right] = Q(\tau \leq t \mid \mathcal{F}_t), \tag{2.2}$$

$$\bar{F}_t := 1 - F_t = Q(\tau > t \mid \mathcal{F}_t), \tag{2.3}$$

denote the ($\mathbb{F}$-*conditional*) *default* respectively *survival probability processes of* τ. Obviously, F ($\bar{F}$) is a bounded, non-negative submartingale (supermartingale). Let us consider the case $\bar{F}_t > 0$ for every $t \geq 0$ throughout the rest of this thesis. Now we are able to introduce the hazard process of τ.

Definition 2.5. (Hazard process) The process $\Gamma := \{\Gamma_t\}_{t\geq0}$ defined by $\Gamma_t = -\ln \bar{F}_t$, $t \geq 0$, is called $\mathbb{F}$-*hazard process of* τ.

One of the major goals in this *enlarged filtration framework* is now to derive calculation techniques for expectations $\mathsf{E}[Y \mid \mathcal{G}_t]$ in terms of expectations conditional on $\mathcal{F}_t$ involving functionals of τ and Γ. The following lemma yields an essential building block to this.

Lemma 2.6. (Key Lemma) *Let Y denote a $\mathcal{G}$-measurable random variable and let $X := \mathsf{E}[Y \mid \mathcal{G}_t]$. Then, for every $t \geq 0$, we have*

$$\mathbb{1}_{\{\tau>t\}} X = \mathbb{1}_{\{\tau>t\}} Z \tag{2.4}$$

with

$$Z := \frac{\mathsf{E}[\mathbb{1}_{\{\tau>t\}} Y \mid \mathcal{F}_t]}{Q(\tau > t \mid \mathcal{F}_t)} = e^{\Gamma_t} \mathsf{E}[\mathbb{1}_{\{\tau>t\}} Y \mid \mathcal{F}_t].$$

Proof. Fix some arbitrary $t \geq 0$ and use equation (2.1) from Remark 2.4. Taking expectations on both sides conditional on $\mathcal{F}_t$, $t \geq 0$ fixed, we obtain

$$\begin{aligned}\mathsf{E}\left[\mathbb{1}_{\{\tau>t\}}Y \mid \mathcal{F}_t\right] &= \mathsf{E}\left[\mathbb{1}_{\{\tau>t\}}\mathsf{E}\left[Y \mid \mathcal{G}_t\right] \mid \mathcal{F}_t\right], \quad \text{since } \mathcal{F}_t \subseteq \mathcal{G}_t, \\ &= \mathsf{E}\left[\mathbb{1}_{\{\tau>t\}}Z \mid \mathcal{F}_t\right], \quad \text{by definition of } X \text{ and Remark 2.4}, \\ &= Z\,\underbrace{\mathsf{E}\left[\mathbb{1}_{\{\tau>t\}} \mid \mathcal{F}_t\right]}_{Q(\tau>t|\mathcal{F}_t)} = Z\,e^{-\Gamma_t}\end{aligned}$$

almost surely for some $Z \in \mathcal{F}_t$. Solving for Z we get the desired result. □

As direct consequences of Lemma 2.6, we summarize some further results without proof, only mentioning that $\mathbb{1}_{\{\tau>t\}} = \mathbb{1}_{\{\tau>t\}}\mathbb{1}_{\{\tau>s\}}$ and $\mathbb{1}_{\{s<\tau\leq t\}} = H_s(1-H_t) = H_t - H_s$ for $0 \leq s < t$ are useful tricks for their verifications:

Corollary 2.7. *Let $0 \leq s < t$. Then,*

(i) for every $\mathcal{G}$-measurable random variable Y we have

$$\mathsf{E}\left[\mathbb{1}_{\{\tau>t\}}Y \mid \mathcal{G}_s\right] = \mathbb{1}_{\{\tau>s\}}e^{\Gamma_s}\mathsf{E}\left[\mathbb{1}_{\{\tau>t\}}Y \mid \mathcal{F}_s\right], \tag{2.5}$$

$$\textit{and } \mathsf{E}\left[\mathbb{1}_{\{s<\tau\leq t\}}Y \mid \mathcal{G}_s\right] = \mathbb{1}_{\{\tau>s\}}e^{\Gamma_s}\mathsf{E}\left[\mathbb{1}_{\{s<\tau\leq t\}}Y \mid \mathcal{F}_s\right], \tag{2.6}$$

(ii) for every $\mathcal{F}_t$-measurable random variable Y we have

$$\mathsf{E}\left[\mathbb{1}_{\{\tau>t\}}Y \mid \mathcal{G}_s\right] = \mathbb{1}_{\{\tau>s\}}\mathsf{E}\left[e^{\Gamma_s-\Gamma_t}Y \mid \mathcal{F}_s\right]. \tag{2.7}$$

In some valuation cases, the size of the contingent claim is unknown before default and revealed immediately after default has occured at τ. Therefore, the following lemma is also of particular interest.

Lemma 2.8. *Let $t \geq 0$. If $Z = \{Z_t\}_{t\geq 0}$ is an $\mathbb{F}$-predictable (bounded) process, then*

$$\mathsf{E}\left[Z_\tau \mid \mathcal{F}_t\right] = \mathsf{E}\left[\int_0^\infty Z_u dF_u \mid \mathcal{F}_t\right] \tag{2.8}$$

$$\textit{and } \mathsf{E}\left[Z_\tau \mid \mathcal{G}_t\right] = \mathbb{1}_{\{\tau\leq t\}}Z_\tau + \mathbb{1}_{\{\tau>t\}}e^{\Gamma_t}\mathsf{E}\left[\int_t^\infty Z_u dF_u \mid \mathcal{F}_t\right] \tag{2.9}$$

Proof. We prove the first equation (2.8) only for processes Z of the form $Z_t := \mathbb{1}_{(s,v]}(t)\epsilon_s$ for some elementary $\mathbb{F}$-predictable process $\epsilon = \{\epsilon_t\}_{t\geq 0}$. From Corollary 2.7 (ii), we have

$$\begin{aligned} \mathsf{E}\left[Z_\tau \mid \mathcal{F}_t\right] &= \mathsf{E}\left[\mathbb{1}_{(s,v]}(\tau)\epsilon_s \mid \mathcal{F}_t\right] \\ &= \mathsf{E}\left[\mathsf{E}\left[\mathbb{1}_{(s,u]}(\tau)\epsilon_s \mid \mathcal{F}_\infty\right] \mid \mathcal{F}_t\right] \\ &= \mathsf{E}\left[\epsilon_s Q(s < \tau \leq v \mid \mathcal{F}_\infty) \mid \mathcal{F}_t\right] \\ &= \mathsf{E}\left[\epsilon_s (F_v - F_s) \mid \mathcal{F}_t\right] \\ &= \mathsf{E}\left[\epsilon_s \int_s^v dF_u \mid \mathcal{F}_t\right] \\ &= \mathsf{E}\left[\int_0^\infty Z_u dF_u \mid \mathcal{F}_t\right] \end{aligned}$$

and, by applying the *Monotone Class Theorem*[3], this result extends to the even more general class of *càdlàg* $\mathbb{F}$-adapted processes, i.e. processes having right-continuous sample paths with left limits, as argued by Jeanblanc and Rutkowski (2000a,b,c).

The second equation (2.9) follows by applying the first equation (2.8) with $\tilde{Z}_\tau := \mathbb{1}_{\{\tau>t\}} Z_\tau$ combined with the Key Lemma 2.6:

$$\begin{aligned} \mathbb{1}_{\{\tau>t\}}\mathsf{E}\left[Z_\tau \mid \mathcal{G}_t\right] &= \mathbb{1}_{\{\tau>t\}} e^{\Gamma_t} \mathsf{E}\left[\tilde{Z}_\tau \mid \mathcal{F}_t\right] \\ &= \mathbb{1}_{\{\tau>t\}} e^{\Gamma_t} \mathsf{E}\left[\int_0^\infty \tilde{Z}_u dF_u \mid \mathcal{F}_t\right] \\ &= \mathbb{1}_{\{\tau>t\}} e^{\Gamma_t} \mathsf{E}\left[\int_t^\infty Z_u dF_u \mid \mathcal{F}_t\right]. \end{aligned}$$

On the other hand, $\mathbb{1}_{\{\tau\leq t\}}\mathsf{E}\left[Z_\tau \mid \mathcal{G}_t\right] = \mathbb{1}_{\{\tau\leq t\}} Z_\tau$ holds naturally true for any $\mathbb{F}$-predictable process Z. □

Remark 2.9. The proof of Lemma 2.8 extends to $\mathbb{G}$-predictable processes because any $\mathbb{G}$-predictable process coincides with an $\mathbb{F}$-predictable process on the set $[0, \tau]$. As mentioned in the proof, it even extends to the more general class of càdlàg $\mathbb{F}$-adapted processes, though it does not extend to the class of càdlàg $\mathbb{G}$-adapted processes (evident if we consider $Z_t := H_t$, $t \geq 0$, as example)!

[3] see for instance Protter (2004), Theorem 8 in Part I, Chapter 2

2.1.3. Intensity-based approach

From now on, we will follow the *intensity-based approach* which is a special case of the hazard process approach discussed in the previous section. We assume that the conditions of the following definition are satisfied throughout the rest of this thesis.

Definition 2.10. (Intensity process) Assume that there exists an $\mathbb{F}$-progressive process $\gamma := \{\gamma_t\}_{t\geq 0}$ with the following properties:

(i) γ is positive.

(ii) $\int_0^t \gamma_s ds < \infty$ a.s. for any $t \geq 0$.

(iii) $\bar{F}_t = Q(\tau > t \mid \mathcal{F}_t) = \exp\left\{-\int_0^t \gamma_s ds\right\}$ for any $t \geq 0$.

Then γ is called $\mathbb{F}$-*intensity process of* τ, or simply *intensity of* τ.

Comparing Definition 2.5 and Definition 2.10 (iii), one immediately sees that the hazard process Γ can be recovered by $\Gamma_t = \int_0^t \gamma_s ds$, $t \geq 0$.

One of the most practical intuitions behind the concept of an intensity is the following: From Definition 2.10 (iii), we can approximate the probability of default within an infinitesimally small time interval $(t, t+dt]$ by $Q(t < \tau \leq t+dt \mid \mathcal{G}_t) \approx \mathbb{1}_{\{\tau>t\}}\gamma_t dt$. That is, the probability of defaulting between t and $t + dt$ is (approximately) proportional to the length of the interval $(t, t + dt]$ with growth rate γ_t. Hence, γ is also often called the *default* or *hazard rate* in this credit risk modelling context. We will use all terms equivalently.

Next, we restate the results of Lemmata 2.6 and 2.8 by means of the intensity process γ:

Corollary 2.11. (Key Lemma, intensity-based approach) *Let $t \geq 0$. Then*

(i) For any $\mathcal{G}$-measurable random variable Y we have

$$\mathsf{E}\left[\mathbb{1}_{\{\tau>t\}}Y \mid \mathcal{G}_t\right] = \mathbb{1}_{\{\tau>t\}}e^{\int_0^t \gamma_s ds}\mathsf{E}\left[\mathbb{1}_{\{\tau>t\}}Y \mid \mathcal{F}_t\right],$$

(ii) If Z is an $\mathbb{F}$-predictable (bounded) process, then

$$\mathsf{E}\left[Z_\tau \mid \mathcal{F}_t\right] = \mathsf{E}\left[\int_0^\infty Z_u\gamma_u e^{-\int_0^u \gamma_s ds}du \mid \mathcal{F}_t\right],$$

$$\text{and } \mathsf{E}\left[Z_\tau \mid \mathcal{G}_t\right] = \mathbb{1}_{\{\tau\leq t\}}Z_\tau + \mathbb{1}_{\{\tau>t\}}\mathsf{E}\left[\int_t^\infty Z_u\gamma_u e^{-\int_t^u \gamma_s ds}du \mid \mathcal{F}_t\right].$$

Corollary 2.7 is similarly expressed in terms of γ using/replacing $e^{\Gamma_t}$by $e^{\int_0^t \gamma_s ds}$.

We conclude this section with another important relation between the $\mathbb{F}$-hazard or $\mathbb{F}$-intensity process and the so-called $\mathbb{H}$-compensator of the default time τ:

Definition 2.12. (Compensator of the default time τ) An $\mathbb{H}$-*compensator* $A := \{A_t\}_{t\geq 0}$ of the default time τ is an $\mathbb{H}$-predictable, right-continuous and increasing process with $A_0 = 0$ such that the process $M := \{M_t\}_{t\geq 0}$ given by $M_t := H_t - A_t$, $t \geq 0$, is an $\mathbb{H}$-martingale.

Proposition 2.13. (Uniqueness of the compensator) *The process $A := \{A_t\}_{t\geq 0}$ is the (unique) $\mathbb{H}$-compensator of the default time τ if and only if*

$$A_t = \Gamma_{\tau\wedge t} = \int_0^t \gamma_u(1 - H_u)du, \quad t \geq 0. \tag{2.10}$$

Moreover, the process $M := \{M_t\}_{t\geq 0}$ given by $M_t := H_t - A_t = H_t - \Gamma_{\tau\wedge t}$, $t \geq 0$, is even a $\mathbb{G}$-martingale.

Proof. First, let A be defined as in equation (2.10). Then it is obvious that $A_0 = \Gamma_0 = 0$ and that A is right-continuous and increasing, since it adopts the same properties from the hazard process Γ. Secondly, $t \mapsto t \wedge \tau$ is a continuous, $\mathbb{H}$-adapted process which implies $\mathbb{H}$-predictability of A.

If we show the last assertion, i.e. that $H - A$ is a $\mathbb{G}$-martingale, of course, this implicitly verifies that $H - A$ is an $\mathbb{H}$-martingale as well and that A is the desired compensator.

Therefore, let $0 \leq s < t$ and define $Z_u := \Gamma_{u\wedge t} - \Gamma_{u\wedge s}$ for $u \geq 0$. Then, using Corollaries 2.7

and 2.11, and the fact that $Z_\tau = (\Gamma_{\tau\wedge t} - \Gamma_{\tau\wedge s}) = 0$ on $\{\tau \leq s\}$, this yields

$$\begin{aligned}
&\mathsf{E}\left[(H_t - H_s) - (A_t - A_s) \mid \mathcal{G}_s\right] \\
&\quad = \mathsf{E}\left[H_t - H_s \mid \mathcal{G}_s\right] - \mathsf{E}\left[Z_\tau \mid \mathcal{G}_s\right] \\
&\quad = \mathbb{1}_{\{\tau>s\}} e^{\Gamma_s} \mathsf{E}\left[\mathbb{1}_{\{s<\tau\leq t\}} \mid \mathcal{F}_s\right] - \mathbb{1}_{\{\tau\leq s\}} Z_\tau - \mathbb{1}_{\{\tau>s\}} e^{\Gamma_s} \mathsf{E}\left[\int_s^\infty Z_u \gamma_u e^{-\Gamma_u} du \mid \mathcal{F}_s\right] \\
&\quad = \mathbb{1}_{\{\tau>s\}} e^{\Gamma_s} \mathsf{E}\left[e^{-\Gamma_s} - e^{-\Gamma_t} - \int_s^t (\Gamma_u - \Gamma_s)\gamma_u e^{-\Gamma_u} du - (\Gamma_t - \Gamma_s) \int_t^\infty \gamma_u e^{-\Gamma_u} du \mid \mathcal{F}_s\right] \\
&\quad = \mathbb{1}_{\{\tau>s\}} e^{\Gamma_s} \mathsf{E}\left[e^{-\Gamma_s} - e^{-\Gamma_t} - \int_s^t \Gamma_u dF_u + \Gamma_s \int_s^t dF_u - (\Gamma_t - \Gamma_s) \int_t^\infty dF_u \mid \mathcal{F}_s\right] \\
&\quad = \mathbb{1}_{\{\tau>s\}} e^{\Gamma_s} \mathsf{E}\left[e^{-\Gamma_s} - e^{-\Gamma_t} - \int_s^t \Gamma_u dF_u + \Gamma_s(F_t - F_s) - (\Gamma_t - \Gamma_s)(1 - F_t) \mid \mathcal{F}_s\right] \\
&\quad = \mathbb{1}_{\{\tau>s\}} e^{\Gamma_s} \mathsf{E}\left[e^{-\Gamma_s} - e^{-\Gamma_t} - \int_s^t \Gamma_u dF_u - \Gamma_s e^{-\Gamma_s} + \Gamma_t e^{-\Gamma_t} \mid \mathcal{F}_s\right]
\end{aligned}$$

One can observe, that the terms within the above expectation indeed sum up to zero, by applying the product rule to

$$\begin{aligned}
\int_s^t \Gamma_u dF_u &= \Gamma_t F_t - \Gamma_s F_s - \int_s^t F_u d\Gamma_u \\
&= \Gamma_t(1 - e^{-\Gamma_t}) - \Gamma_s(1 - e^{-\Gamma_s}) - \underbrace{\int_s^t d\Gamma_u}_{\Gamma_t - \Gamma_s} + \underbrace{\int_s^t e^{-\Gamma_u} d\Gamma_u}_{e^{-\Gamma_s} - e^{-\Gamma_t}} \\
&= \Gamma_t e^{-\Gamma_t} - \Gamma_s e^{-\Gamma_s} + e^{-\Gamma_s} - e^{-\Gamma_t}.
\end{aligned}$$

The opposite direction follows immediately by the uniqueness of the *Doob-Meyer decomposition* which states that the supermartingale H has the unique represenation $H = M + A$ with a local $\mathbb{H}$-martingale M and *the* (indeed unique) $\mathbb{H}$-compensator A of τ.[4] □

Notations 2.14. In this section, we have adopted the notation of the $\mathbb{F}$-hazard ($\mathbb{F}$-intensity) process in terms of the Greek letter Γ (γ). This is consistent with the textbook of Bielecki and Rutkowski (2002), which this section is mainly based on. It is, however, important to mention that the authors therein accurately differ between several hypotheses concerning the constellation of the filtrations $\mathbb{G}$, $\mathbb{F}$ and $\mathbb{H}$, among others. In our context, we adopt the most frequently assumed situation of $\mathbb{F} \subsetneq \mathbb{G}$ such that τ is not an $\mathbb{F}$-stopping time in general as well as uniqueness γ which is related to the process F being continuous and increasing.

[4] see, for example, Protter (2004), Part III, Chapter 3

For more details on this and further general cases to see the difference between the $\mathbb{F}$-hazard process denoted as above by Γ and the $(\mathbb{F}, \mathbb{G})$-*martingale hazard process* denoted by Λ, we refer the interested reader to the papers of Jeanblanc and Rutkowski (2000a,b,c), for instance.

Before turning to applications of credit risk modelling in the intensity-based framework, we give a short yet central result as an analogy to formula (1.3) of Proposition 1.5.

2.1.4. Risk-neutral valuation in the defaultable case

Similar to the results on risk-neutral valuation in the default-free case (Proposition 1.5 in Chapter 1) and the techniques derived in the reduced-form approach (Sections 2.1.2 and 2.1.3), we end up here at a risk-neutral pricing formula for contingent claims now additionally involving default risk. Note that our market filtration with respect to which we take conditional expectations is now $\mathbb{G} = \mathbb{F} \vee \mathbb{H}$.

Theorem 2.15. (Risk-neutral valuation of defaultable contingent claims, intensity-based approach) *For some starting date $s \geq 0$ and a time interval $[s, s+T]$ of length $T > 0$, denote a defaultable contingent claim due at time $s+T$ by $Y := X\,\mathbb{1}_{\{\tau > s+T\}}$ for some $\mathcal{F}_{s+T}$-measurable integrable random variable X. Then, its risk-neutral price process $\tilde{\Pi} := \{\tilde{\Pi}_t\}_{t\geq 0}$ is given by*

$$\tilde{\Pi}_t := \mathsf{E}\left[D(t, s+T)\, Y \mid \mathcal{G}_t\right] = \mathbb{1}_{\{\tau > t\}} \mathsf{E}\left[e^{-\int_t^{s+T} r_u + \gamma_u du} X \mid \mathcal{F}_t\right], \quad t \geq 0. \tag{2.11}$$

Proof. Applying Corollary 2.11 (i) immediately verifies equation (2.11). □

As one might observe, the present value of Y, discounted with $D(t, s+T) = \exp\left\{\int_t^{s+T} r_u du\right\}$ at the short rate $r = \{r_t\}_{t\geq 0}$, is equivalent to the present value of X, discounted at a higher rate consisting of $r + \gamma$. We will see what this higher "premium" means in the context of the next section, when we turn our attention to applications of this formula to (defaultable) contingent claims specified by the contractual structure of some important financial instruments.

2.2. Financial instruments subject to default risk

In this section, we apply the techniques from the previously introduced intensity-based framework on pricing credit risky securities and derivatives. In doing so, we focus on *bonds* and so-called *credit default swaps* (*CDS*) which are also commonly understood as insurance contracts against the default risk of a reference obligation. Especially in the first decade of the 21st century, they gained an essential and meaningful role in the credit derivatives market and became a central topic to financial researchers and mathematicians as well.

For this section, recall the reduced-form framework within the probability space $(\Omega, \mathcal{G}, \mathbb{G}, Q)$ endowed with the filtration $\mathbb{G} = \mathbb{F} \vee \mathbb{H}$ and risk-neutral pricing measure Q. Furthermore, assume that an $\mathbb{F}$-intensity process γ of the default time τ exists (uniquely).

2.2.1. Defaultable bonds

Bonds are debt securities between two parties: The *bond holder* (*investor* or *buyer*) and the *bond issuer* (*obligor* or *seller*) both of which can be corporations, banks or governments. The holder invests a certain amount of money to the bond issuer today who in turn promises a future a priori fixed payment, called *notional* (or *nominal, principal*, *face value, par value*), at a fixed maturity date. Moreover, the bond issuer commits himself to pay periodically a constant (*fixed*) or a time-varying (*floating*) fraction of the notional, called *coupon*. Beyond the standard contractual items of bonds, such as the size of the face value and coupons as well as maturity and payment dates, we do not discuss more complex attributes in further detail. We only mention that—within a standardized framework—there are far more possibilities of structuring bonds such as exchangeability to shares of the issuer's or a third party's common stock (*convertible* vs. *exchangeable* bonds) and classification by liquidation priority/credit rating (*senior* vs. *subordinated* bonds, *investment-grade* vs. *high-yield* bonds) etc. which we will neglect within the scope of this thesis.

We find bonds amongst the most popular financial instruments not only in terms of simplicity of their properties but also in terms of historical meaning and development of the markets they are traded on. As they are standardized securities, buying and selling takes primarily place at auctions but also secondarily on over-the-counter markets (bonds can be resold during their lives).

The goal is now to price bonds, that is, to calculate their present value which equals the amount of money the buyer has to invest today in order to get back the notional amount after the remaining time to maturity. On the other hand, concerning credit risk, the investor is also facing the possibility that the obligor might fail to meet the agreed redemption of the notional at maturity or to pay his regular coupon obligations. We assume the credit-worthiness of the issuer to be already fully reflected in the price of the bond, regardless of what its credit rating might be.

Before stating some formal results, think of the possibility that the obligor fails to repay the face value up to a certain (but not the total) amount. That is, he manages to pay back a particular fraction of the agreed notional amount, called *recovery of face value* (RFV).[5] For different modelling purposes but also with regards to different prudential and bankruptcy court regulations, one can assume the recovery fraction to denote the current value of an otherwise equivalent but default-free bond (*recovery of treasury*, RTV)[6] or the current (or instantaneous "pre-default") market value of the considered bond itself (*recovery of market value*, RMV)[7]. The distinction between these aspects is not subject of our considerations here; we will only focus on the RFV assumption as it is also the implicitly assumed in CDS. For a good theoretical overview, we refer the interested reader to Duffie and Singleton (1999) or Uhrig-Homburg (2002).

Assume a defaultable bond, issued at time $s \geq 0$ ("today"), committing the obligor to pay a floating coupon, denoted by an $\mathbb{F}$-predictable non-negative process $\tilde{C} = \{\tilde{C}_t\}_{t\geq 0}$, continuously during the life $[s, s+T]$ of the bond and to repay the notional $\tilde{N}$ to the investor at maturity $s+T$. Without loss of generality, we consider unit notional amounts $\tilde{N} = 1$. Furthermore, let τ denote the obligor's random default time at which he stops every payment, including coupons if $\tau \in [s, s+T]$; this is referred to as *credit event*. Instead of losing the full notional, assume that the investor receives at least the recovery fraction $R \in [0, 1]$ of the unit notional amount. Of course, R is unknown before τ in general. However, since the determination of recovery claims on a defaulted bond by prudential authorities might even extend over several years after the credit event, implicit or explicit data for recovery rates are difficult to find. Therefore we suggest that R does neither depend on the short rate $\{r_t\}_{t\geq 0}$ nor on the default time (hence on its hazard rate $\{\gamma_t\}_{t\geq 0}$) nor on time at all. Instead, we claim the (albeit simplifying) assumption of a constant recovery rate.[8]

---

[5] as first described by Brennan and Schwartz (1980) and Duffee (1998)

[6] as first described by Jarrow and Turnbull (1995)

[7] as first described by Duffie and Singleton (1999)

[8] More concretely, for theorists as well as practitioners, it is convenient to classifiy credit-risky instruments into different seniorities $\mathcal{S} \in \{\text{secured, senior-unsecured, subordinated}, ...\}$ and assume fixed levels $R = R(\mathcal{S})$, decreasing in $\mathcal{S}$, of recovery at default for all bonds within the same seniority. Recent works that relax these assumptions and model R as another $\mathcal{F}_\tau$-measurable random variable besides the short and

Then, the discounted cashflows $\mathrm{dcf}_{\tilde{C}}^{R}(s, s+T)$ of the bond are summarized as follows:

$$\begin{aligned}\mathrm{dcf}_{\tilde{C}}^{R}(s, s+T) := \int_s^{s+T} & D(s,u)\,\tilde{C}_u\,\mathbb{1}_{\{\tau>u\}}du \\ & + \; D(s,s+T)\,\mathbb{1}_{\{\tau>s+T\}} \\ & + \; D(s,\tau)\,R\,\mathbb{1}_{\{\tau\le s+T\}} \end{aligned} \tag{2.12}$$

where $D(\cdot,\cdot)$ denotes the discount factor from Definition 1.1 with respect to the short rate $r = \{r_t\}_{t\ge 0}$.

By applying the reduced-form techniques provided in the previous section, especially Theorem 2.15, we obtain the following risk-neutral pricing result within the intensity-based framework:

Proposition 2.16. (Risk-neutral price of a defaultable bond, intensity-based) *Denote by $\gamma = \{\gamma_t\}_{t\ge 0}$ the hazard rate of the bond issuer's default time τ. Then, in the intensity-based framework, the price of the above defaultable coupon bond with unit notional and given constant recovery $R \in [0,1]$, denoted by $\tilde{\Pi}_{\tilde{C}}^{R}(s, s+T)$, is the conditional expectation of the discounted future cashflows $\mathrm{dcf}_{\tilde{C}}^{R}(s, s+T)$ from equation (2.12), i.e.*

$$\begin{aligned}\tilde{\Pi}_{\tilde{C}}^{R}(s, s+T) &= \mathsf{E}\left[\mathrm{dcf}_{\tilde{C}}^{R}(s, s+T) \mid \mathcal{G}_s\right] \\ &= \mathbb{1}_{\{\tau>s\}} \int_s^{s+T} \mathsf{E}\left[\tilde{C}_u e^{-\int_s^u r_t+\gamma_t dt} \mid \mathcal{F}_s\right] du \\ &\quad + \; \mathbb{1}_{\{\tau>s\}}\,\mathsf{E}\left[e^{-\int_s^{s+T} r_t+\gamma_t dt} \mid \mathcal{F}_s\right] \\ &\quad + \; \mathbb{1}_{\{\tau>s\}}\,R \int_s^{s+T} \mathsf{E}\left[e^{-\int_s^u r_t+\gamma_t dt}\gamma_u \mid \mathcal{F}_s\right] du. \end{aligned} \tag{2.13}$$

For the bond price of a default-free issuer (equivalently, of a default-free bond) denoted by Π_C with corresponding coupon process $C := \{C_t\}_{t\ge 0}$ (and unit notional $N = 1$), one

hazard rates (with possible dependence structures) have been done by Pan and Singleton (2008) and Schläfer and Uhrig-Homburg (2009), to name but a few.

similarly obtains that[9]

$$\begin{aligned}\Pi_C(s,s+T) &= \mathsf{E}\left[\int_s^{s+T} D(s,u)\,C_u\,du \mid \mathcal{G}_s\right] \\ &\quad + \mathsf{E}\left[D(s,s+T) \mid \mathcal{G}_s\right] \\ &= \int_s^{s+T} \mathsf{E}\left[C_u e^{-\int_s^u r_t dt} \mid \mathcal{F}_s\right] du \\ &\quad + \mathsf{E}\left[e^{-\int_s^{s+T} r_t dt} \mid \mathcal{F}_s\right].\end{aligned}$$

Remark 2.17. It might be worth indicating that the price Π_C of a default-free bond can be rewritten in terms of the price of a defaultable bond with same coupon process $C := \{C\}_{t\geq 0}$ and recovery $R = 1$, i.e. $\Pi_C = \tilde{\Pi}_C^1$ which means explicitly that

$$\begin{aligned}\Pi_C(s,s+T) &= \mathsf{E}\left[\int_s^{s+T} D(s,u)\,C_u\,du \mid \mathcal{G}_s\right] \\ &\quad + \mathsf{E}\left[D(s,s+T) \mid \mathcal{G}_s\right] && (2.14) \\ = \tilde{\Pi}_C^1(s,s+T) &= \mathsf{E}\left[\int_s^{s+T} D(s,u)\,C_u\,\mathbb{1}_{\{\tau>u\}}\,du \mid \mathcal{G}_s\right] \\ &\quad + \mathsf{E}\left[D(s,s+T)\,\mathbb{1}_{\{\tau>s+T\}} \mid \mathcal{G}_s\right] && (2.15) \\ &\quad + \mathsf{E}\left[D(s,\tau)\,\mathbb{1}_{\{\tau\leq s+T\}} \mid \mathcal{G}_s\right]. && (2.16)\end{aligned}$$

Note that these two bonds arise from two different obligors (with different creditworthinesses), one that might default at a random time τ and one other that will never default (his default time is virtually equal to infinity)—but their prices are identical: The idea behind this modification is to decompose the above terms of the default-free bond price formula (in particular, the present value of the redemption payment, line (2.14)) into the present values of redemption at maturity $s+T$ (this corresponds to no default of the defaultable issuer, line (2.15)) and of redemption at τ (this corresponds to default of the defaultable issuer, line (2.16)). This "complification" of the simple default-free bond price formula will be useful in the next section, when we try to reproduce a CDS contract synthetically by selling a defaultable bond and buying a default-free bond in order to obtain a fair insurance premium against default.

[9] This is done by considering $\gamma_t := 0 \Rightarrow \Gamma_t = 0 \Rightarrow \bar{F}_t = Q(\tau > t \mid \mathcal{F}_t) = e^{-0} = 1$ for every $t \geq 0$, i.e. the issuer pays everything back and can almost surely never default.

Remark 2.18. It should be stressed at this point that assuming a continuous payment schedule (just as continuous discounting) is an *idealized* but very useful consideration for both illustrating the applied techniques and approximating reality, where actually fixed (discrete, e.g. monthly or quarterly) payment periods and discrete (or linear) discounting are conventional. Because of these benefits, we remain in this idealized setup.

Next, we are interested in quantifying the credit quality of a defaultable bond in terms of a premium relative to a default-free bond. Therefore, as a measure, we want to use its so-called *par yield spread*. We mainly follow the ideas of LEVY (2009) and CECCHETTI and DI CESARE (2012), who formalize the concepts assuming constant short and hazard rates which we claim to be, however, also valid in our more general context.

Definition 2.19. (Par yield) Let $\tilde{\Pi}^R_{\tilde{C}}$ denote the price of a defaultable coupon bond associated with its coupon process $\tilde{C} = \{\tilde{C}_t\}_{t\geq 0}$ and recovery rate R. Then, the (continuously payed) *par yield* is defined as the coupon $\tilde{Y} := \{\tilde{Y}_t\}_{t\geq 0}$ for which the bond price is equal to its par value, i.e. the *par condition*

$$\tilde{N} = 1 = \tilde{\Pi}^R_{\tilde{Y}}(s, s+T) \tag{2.17}$$

is satisfied. Similarly, $Y := \{Y_t\}_{t\geq 0}$ denotes the par yield of a default-free bond with par condition $\Pi_Y(s, s+T) = 1 = N$. In the intensity-based framework, we obtain more concretely that[10]

$$\begin{aligned} \tilde{Y}_t &= r_t + (1-R)\gamma_t, \quad t \geq 0, & (2.18)\\ Y_t &= r_t, \qquad\qquad\quad\;\; t \geq 0, & (2.19) \end{aligned}$$

The solutions (2.18) and (2.19) are very plausible from an economical perspective in so far as bonds whose coupons are equal to their discount rates are said to be "selling at par". Otherwise, if the coupons $\tilde{C}$ are below (above) the par yield $\tilde{Y}$, the bond is said to be "selling at a discount" ("selling at a premium").[11]

Remark 2.20. Observe that the par yields are independent of the tenor T of a bond in this context! Hence, we can treat the par yield as equal to the short rate plus $(1-R)$ times the

[10] This can be easily checked using

$$\int_s^{s+T} e^{-\int_s^u r_t+\gamma_t dt}(r_s+\gamma_s)ds = -\int_s^{s+T} d(e^{-\int_s^u r_t+\gamma_t dt}) = 1 - e^{\int_s^{s+T} r_t+\gamma_t dt}.$$

[11] Furthermore, for constant coupons, the par yield equals the so-called *yield to maturity*, as well, see CECCHETTI and DI CESARE (2012).

hazard rate at the corresponding infinitesimally small time instant. Remember that this is only an idealized assumption.[12]

Since we are claiming non-negative hazard rates in our considerations, $\tilde{Y}_t \geq Y_t$ always holds true for all $t \geq 0$. The difference $\tilde{Y} - Y$, which is then always non-negative, is called the par yield spread and has a specific interpretation:

Definition 2.21. (Par yield spread) For a defaultable and a default-free coupon bond with prices $\tilde{\Pi}^R_{\tilde{C}}$ and Π_C and corresponding coupon rates $\tilde{C}$ and C and par yields $\tilde{Y}$ and Y, respectively, the *par yield spread* $S^* := \{S^*_t\}_{t \geq 0}$, or simply *spread*, between these bonds is defined by

$$S^*_t = \tilde{Y}_t - Y_t, \ t \geq 0, \tag{2.20}$$

which is obviously equal to

$$S^*_t = (1 - R)\gamma_t, \ t \geq 0, \tag{2.21}$$

in the intensity-based framework.

Remark 2.22. (Credit spread) The default-free bond is usually called *reference* or *benchmark bond* in the context of Definition 2.21. The lower the credit quality of the defaultable bond compared to the reference bond with par yield $y = r$ and the lower his (expected) recovery rate R, the higher becomes his default intensity and therefore the additional yield premium $S^* = (1 - R)\gamma$ the investor is charging on top of r for taking on credit risk. For this intuitive reason, the par yield spread of a bond can be referred to as its *credit spread*.[13] Finally, claiming that $\tilde{Y}_t \geq Y_t$, $t \geq 0$, does indeed make sense (otherwise, an interpretation of a negative premium would correspond to a negative default intensity which is senseless). Notice at this point that, since the par yields do not depend on the maturity of the bond, neither does the respective spread!

Remark 2.23. Although in practice, there is profound empirical evidence and broad economical agreement that a bond's spread over a reference bond does not only consist of a credit component but is also influenced by liquidity, taxation and optionality components that are specific to the respective bond market or contract[14], we assume a frictionless, complete market, without transaction costs and other restrictions such as to short sellings and we neglect all determinants of the yield spread other than related to credit risk.

[12] For the sake of completeness and practical relevance, some formulas for constant coupon bonds are attached in Appendix A.1, which are more commonly used in practice.

[13] As we will see lateron in Section 2.2.2, this equals to the par premium chosen for a CDS contract on the same bond.

[14] see Longstaff, Mithal and Neis (2005) and references therein

Next, we consider derivatives on credit-risky bonds. As a very prominent representative the so-called *credit default swap* is discussed, which can be considered as an insurance against (or bet on) a default of a reference issuer. We will see that the premium of such an insurance contract coincides with to the credit spread S^* over a reference bond in our framework.

2.2.2. Credit default swaps

Like bonds, *credit default swaps* (*CDS*) are financial contracts between two parties for a specific time interval $[s, s+T]$ to exchange cashflows. In contrast, CDS are not securities: They are (credit) derivatives, since the cashflows during the life of the contract depend on a third party's obligation, called *reference obligation* or *reference entity* (normally a bond).

The *protection buyer* of a CDS usually seeks for a compensation by exactly the amount $L := 1 - R$ that is expected to be lost at the default of the reference obligation—which is in turn expected to repay only the fraction $R \in [0, 1)$ of its notional value. On the other hand, the *protection seller* is charging him a premium that is again, for the sake of generality, assumed to be payed continuously and denoted by an $\mathbb{F}$-predictable non-negative process $C^* := \{C^*_t\}_{t\geq 0}$.[15]

Hence, at the conclusion of a CDS contract at time $s \geq 0$ ("today"), the two parties agree upon the following subjects:

1. The reference entity (a concrete issuance by a particular third party)
2. The tenor T, such that the *date of expiry* (maturity date) is $s + T$
3. The compensation payment $L \in (0, 1]$ as a fraction of the reference obligation's (unit) notional amount that is payed when the reference issuer fails to meet his commitments.

Preliminarily, this is all we have to know about when it comes to determining the fair CDS premium $\{C^*_t\}_{t\geq 0}$. In the following, we want to illustrate how it can be obtained by the theory introduced in the sections before.

A first approach is based on Duffie (1999), who describes a no-arbitrage argument by replicating the cashflows of a CDS contract with selling a defaultable bond and buying a

[15] Albeit in practice, CDS premia are usually payed at scheduled points in time, e.g. quarterly or semi-annualy, similarly to the aforementioned coupons of bonds. For a treatment of the following formal steps in a constant premium context, see Appendix A.2.

default-free bond, which is in general void of any modelling assumptions with respect to the reference entity's default time τ. The second approach, actually based on the first, yields a more concrete formula for $\{C_t^*\}_{t\geq 0}$ assuming the intensity-based framework introduced in Section 2.1. In this thesis, we neglect "counterparty risk", i.e. the possibility that either the protection buyer or the protection seller themselves could fail to meet the agreed payments (premia and loss compensation, respectively), such that the premia only consist of credit risk components for the underlying reference entity.[16]

The fair CDS premium based on no-arbitrage arguments

CDS were initially developed for isolating credit risk that is immanent in a defaultable obligation (such as a bond) and transferring it to another party who is willing to be financially liable for a compensation of a potential loss. This is basically comparable to an insurance against default. The position that is actually left after "re-selling" the isolated default risk, should be the present value of an otherwise equivalent riskless investment, though with a smaller interest, reduced by the premium for the insurance.

This rather heuristic consideration is what we want to formalize in the following: Consider a defaultable bond with price $\tilde{\Pi}^R_{\tilde{C}}(s, s+T)$, recovery rate R and corresponding coupon process $\tilde{C}$ and a default-free bond trading at par, i.e. with price $\Pi_Y(s, s+T) = 1$ such that its coupon equals the par yield, $C = Y$. Recalling that $\Pi_Y = \tilde{\Pi}^1_Y$ (see Remark 2.17), consider a portfolio consisting of a short position in the defaultable bond and a long position in the default-free bond. This results in the formal calculation

$$\begin{aligned}
& - \quad \mathrm{dcf}^R_{\tilde{C}}(s, s+T) \\
& + \quad \mathrm{dcf}^1_Y(s, s+T) \\
= & - \left(\int_s^{s+T} D(s,u)\, \tilde{C}_u\, \mathbb{1}_{\{\tau>u\}} du \;+\; D(s,s+T)\, \mathbb{1}_{\{\tau>s+T\}} \;+\; D(s,\tau)\, R\, \mathbb{1}_{\{\tau\leq s+T\}} \right) \\
& + \left(\int_s^{s+T} D(s,u)\, Y_u\, \mathbb{1}_{\{\tau>u\}} du \;+\; D(s,s+T)\, \mathbb{1}_{\{\tau>s+T\}} \;+\; D(s,\tau)\, \mathbb{1}_{\{\tau\leq s+T\}} \right) \\
= & - \int_s^{s+T} D(s,u)\, (\tilde{C}_u - Y_u)\, \mathbb{1}_{\{\tau>u\}} du \;+\; D(s,\tau)\, \underbrace{(1-R)}_{L}\, \mathbb{1}_{\{\tau\leq s+T\}}
\end{aligned} \tag{2.22}$$

[16] For modelling approaches accounting for additional counterparty risk, see Hull and White (2001) for instance.

for the discounted cashflows of this portfolio.

We notice that the resulting cashflows replicate those of a CDS contract from the protection buyer's point of view, if we set the premium to $C^* = \tilde{C} - Y$.[17] The first term in line (2.22) describes the accumulated discounted cashflows of premia until maturity or default, whichever comes first. This is called the *premium leg* of the CDS contract (negative (positive) cashflow for the protection buyer (seller)). The second term is the discounted compensation at default—in case this event occurs before maturity—and is called the *protection leg* of the CDS (positive (negative) cashflow for the protection buyer (seller)).

The premium is called "fair" if the two opposed discounted cashflows are equal in terms of present values, that is,

$$\mathsf{E}\left[\int_s^{s+T} D(s,u)\, C_u^* \, \mathbb{1}_{\{\tau>u\}} du \mid \mathcal{G}_s\right] = L\, \mathsf{E}\left[D(s,\tau)\, \mathbb{1}_{\{\tau\leq s+T\}} \mid \mathcal{G}_s\right]. \qquad (2.23)$$

We refer equation (2.23) to as the *par CDS condition*. This implies that the short position in the defaultable bond and the long position in the default-free bond have to cancel out, as well. As a consequence,

$$\Pi_Y(s, s+T) = 1 = \tilde{\Pi}_{\tilde{C}}^R(s, s+T)$$

must hold true. Since this means that the defaultable bond is also traded at par, its coupon is also equal to the corresponding par yield and hence

$$C_t^* = \tilde{C}_t - Y_t = \tilde{Y}_t - Y_t = S_t^*, \quad t \geq 0, \qquad (2.24)$$

which is the par yield spread S^* as introduced in Definition 2.21 in the previous section.

This result suggests to use the par yield spread, which we assume to be a credit spread (i.e. it is not determined by other components) as the fair premium for a CDS contract.[18]

In reality one can observe a difference between the par yield spread of a bond and the premium of a CDS on the same bond. The difference $B^* := C^* - S^*$ is referred to as *CDS/bond basis* in the literature and has been subject to many empirical surveys that test the hypothesis of $B^* = 0$ (also called *zero CDS/bond basis hypothesis*), see, for example, studies of

[17] Otherwise, if $C^* \neq \tilde{C} - Y$, there would be arbitrage opportunities.

[18] Recall that we have already used the term "premium" for the par yield spread in our interpretation as an additional interest charged for taking on default risk in Remark 2.22. So eventually this turns out to be consistent with the word "premium" in the context of CDS.

Longstaff, Mithal and Neis (2005), Ericsson, Jacobs and Oviedo (2009), Blanco, Brennan and Marsh (2005) Nashikkar, Subrahmanyam and Mahanti (2011), Bai and Collin-Dufresne (2011), Levy (2009), Fontana (2011) and Zhu (2006). All of them can either confirm empirically an approximate parity *"CDS premium ≈ credit spread"* or find discrepancies due to economical reasons based on the properties of specific contracts or market characteristics. Another survey on the CDS/bond basis is provided by Hull, Predescu and White (2004). In contrast to the rest, they subtract CDS premia from corporate bond yields and draw conclusions about the benchmark interest rate. Comprehensive comments on this topic are also available in Mahanti, Nashikkar and Subrahmanyam (2007), who argue that

> *"...the advent of the CDS market makes it possible to isolate default risk in corporate bonds issued by a certain issuer without relying too heavily on a particular model of credit risk and a specific parameterization, since a direct reading of the market's pricing of credit risk is available. [...] since CDS contracts price default risk explicitly, they are a good benchmark for the pure credit risk of the firm [...] It must be noted that most corporate bonds issued by firms tend to be fixed-rate bonds, and thus, this equivalence does not hold exactly. More importantly, as shown by Longstaff et al. (2005), the pure corporate bond spread is a biased measure. [...] This is particularly true when there are frictions in the arbitrage mechanism between the CDS contract and the bond [...]"* (ib. Section 4.2, page 12)

Furthermore, see also Hull and White (2000), pages 14–18, for comments on the (idealized) approximation of the CDS premium by the credit spread of its reference entity.

As brought up before, we stick with these idealized assumptions as we are lateron only interested in modelling the time series behaviour of the spreads, anyway (as will be clear in Chapters 3 and 4).

Intensity-based valuation of CDS contracts

So far, we have not used any model assumption with regards to the default time τ of the reference entity. Treating both sides of the par CDS condition (2.23) separately, we first obtain the present values $\mathrm{PL}(C^*; s, s+T)$ and $\mathrm{DL}(L; s, s+T)$ of the premium leg and the default leg, namely

$$\begin{aligned}\mathrm{PL}(C^*; s, s+T) &:= \mathsf{E}\left[\int_s^{s+T} D(s,u)\, C_u^*\, \mathbb{1}_{\{\tau>u\}} du \mid \mathcal{G}_s\right] \\ &= \mathbb{1}_{\{\tau>s\}} \int_s^{s+T} \mathsf{E}\left[C_u^*\, e^{-\int_s^u r_t+\gamma_t dt} \mid \mathcal{F}_s\right] du\end{aligned}$$

and

$$\begin{aligned}\mathrm{DL}(L; s, s+t) &:= L\, \mathsf{E}\left[D(s,\tau)\, \mathbb{1}_{\{\tau\le s+T\}} \mid \mathcal{G}_s\right] \\ &= L\, \mathbb{1}_{\{\tau>s\}} \int_s^{s+T} \mathsf{E}\left[e^{-\int_s^u r_t+\gamma_t dt} \gamma_u \mid \mathcal{F}_s\right] du,\end{aligned}$$

respectively. Then, claiming the par CDS condition to hold, i.e. $\mathrm{PL}(C^*; s, s+T) = \mathrm{DL}(L; s, s+t)$, yields that

$$C_t^* = (1-R)\gamma_t, \quad t \ge 0, \tag{2.25}$$

and we are back in equation (2.24) with the intensity-based credit spread $C_t^* = S_t^*$ (2.21).

Concluding comments

Altogether, we have seen how prices of (non-)defaultable bonds and fair premia of CDS on such bonds are determined and how they are economically and formally interrelated.

As pointed out in Longstaff, Mithal and Neis (2005), pages 7–8, CDS contracts are characterized by

1. lower sensitivity to liquidity and supply/demand pressures,
2. theoretically unbounded supply (in terms of notional and availability of contracts and their "generic" nature) and
3. equal conditions for taking up positions of either a buyer or a seller, respectively,

in contrast to bonds.

Furthermore, CDS prices are proven to reflect new credit information on the reference issuer more rapidly than bond prices do.

In practice, CDS contracts usually agree on a premium $C^* = \{C_t^*\}_{t\ge0} \equiv const$ that is fixed instead of floating. In the basic literature about pricing CDS, such as Schönbucher

(2003), Martin, Reitz and Wehn (2006), Duffie and Singleton (2009), Bielecki and Rutkowski (2002), to name but a few, pricing formulas also follow this assumption. In Appendix A.2, we have collected modifications of the above formulae to the constant premium assumption that result in equations (A.1), (A.2) and (A.3) for the CDS spread $C^* = C^*(s, s+T)$ which do depend on the tenor T and can be found in the aforementioned literature. Of course, they reflect reality far better, however, these are much more useful for instance if we want to *calibrate* a term structure of hazard rates to a given term structure of CDS premia on a specific day $s \geq 0$. This means primarily replicating market equilibrium, i.e. to find the intensities implicitly assumed to get the given observed premia quoted by traders. For such purposes, deterministic intensity functions (most often piecewise constant or linear) are assumed. Then, the "local intensities" are stepwise identified, given all other quantities such as the discount factors (calculated from appropriate observable proxies for the short rate), the recovery rate (often assumed to be $R = 40\%$, hence $L = 60\%$) and the premia $C^*(s, s+T_1), ..., C^*(s, s+T_N)$ for the corresponding tenors $T_1 < ... < T_N$. This is usually referred to as *bootstrapping*, see, for example, Luo (2005), Chan-Lau (2006) and especially Martin, Reitz and Wehn (2006), Chapter 4, for extensive descriptions of this procedure. The idea of concatenating bootstrapped intensities on a day-by-day basis might be a starting point but is questionable from the point of view of a stochastic process/time series concept, though.

But this is not within the scope of this thesis. Instead, we try to fit a tractable yet sufficiently realistic time series model to describe the behaviour of several historical series of $C^{*i} = \{C_t^{*i}\}_{t \geq 0}$ on a daily basis for different reference entities $i \in \{1, ..., M\}$, $M \in \mathbb{N}$, observed at a large time horizon $[0, T^*]$, $T^* > 0$, all with a common fixed contract maturity $T > 0$, see Chapter 4.

Nevertheless, as a conclusion, it is worth mentioning that—be it a constant or a floating coupon (payed continuously or discretely)—one can convince oneself that all concepts coincide if one assumes constant intensities $\gamma_t \equiv \gamma \in [0, \infty)$. Then, for example, formulas (A.1) and (A.3) of Appendix A all lead to a constant CDS premium independent of the tenor T and short rate r (be it stochastic or not!), since

$$C^*(s, s+T) = (1-R)\, \frac{\int_s^{s+T} \mathsf{E}\left[e^{-\int_s^u r_t dt} \mid \mathcal{F}_s\right] e^{-\gamma(u-s)} \gamma du}{\int_s^{s+T} \mathsf{E}\left[e^{-\int_s^u r_t dt} \mid \mathcal{F}_s\right] e^{-\gamma(u-s)} du} = (1-R)\, \gamma,$$

(given $\tau > s$) and all other terms cancel out.

It has been argued with several justifications now that the realism is not lost and the time-varying approximation (2.25) is still valid though idealized within our framework. The following chapter is attended to provide the theoretical basis of a special class of stochastic processes that the premium process $C^* = \{C_t^*\}_{t \geq 0}$ of equation (2.25), and therefore implicitly the hazard rate, is tried to be modelled with. Afterwards, the idea will be applied in Chapter 4 and tested on real data in order to investigate how credit risk can appropriately be modelled by these concepts.

Chapter 3.

Continuous-time linear processes

In classical time series analysis, linear models for discrete-time processes such as *autoregressive moving-average* (ARMA) models have been applied to a wide range of practical problems in physics, engineering, finance and many more areas for several decades. In academic literature, recent works have focussed on so-called *CARMA* (*continuous-time ARMA*) *models*, trying to derive a continuation of the classical ARMA theory. With the seminal works of Brockwell (2001, 2004) as well as Brockwell and Lindner (2009), a foundation for further investigations and applications of CARMA processes was provided, emphasizing several advantages over their discrete-time counterparts, for instance, their ability to model high-frequent (nearly continuous) observations or irregularly spaced samples of a continuously modelled process, respectively.

In this chapter, we begin with a short motivation in Section 3.1, before formally more correct definitions and representations of (Lévy-driven) CARMA processes are following in Section 3.2. In order to get familiar with similarities and differences to the concepts of discrete-time ARMA modelling, necessary and sufficient conditions for *stationarity*, *causality* and *invertibility* are stated in Section 3.3. Especially the latter enables us to recover the background-driving process, in general a Lévy process, at least approximately, from discrete observations of CARMA processes—not quite so trivial as in the traditional ARMA case.

The intrinsic aim of the following sections is to provide a basis for practical applications in Chapter 4. There, we will make use of the manifold possibilities of the class of CARMA processes and apply the methods presented in this chapter to observations of the CDS premium time series introduced in Chapter 2. For readers interested in further topics, an overview on recent developments in the literature concludes this chapter (Section 3.4).

3.1. Motivation

The classical representation of an ARMA(p,q) process $X = \{X_t\}_{t\in\mathbb{Z}}$ with the (equidistant) time domain $\mathbb{Z}$ is stated in terms of their characteristic polynomials ϕ and ϑ by

$$\phi(B)X_t = \vartheta(B)\epsilon_t, \quad t \in \mathbb{Z}, \tag{3.1}$$

where B denotes the *backshift operator*, defined by $B^j X_t = X_{t-j}$, and $\epsilon = \{\epsilon_t\}_{t\in\mathbb{Z}}$ is a *white noise process*[1], also called the *innovation process* or simply *noise*. Usually, ϕ and ϑ are polynomials of degrees $p, q \in \mathbb{N} \cup \{0\}$, respectively, and parametrized with the coefficients $\phi_1, ..., \phi_p$ and $\vartheta_1, ..., \vartheta_q$ as

$$\begin{aligned} \phi(z) &:= 1 - \phi_1 z - ... - \phi_p z^p, \\ \vartheta(z) &:= 1 + \vartheta_1 z + ... + \vartheta_q z^q, \quad z \in \mathbb{C}. \end{aligned}$$

Therefore, equation (3.1) can be explicitly rewritten to

$$X_t = \sum_{j=1}^{p} \phi_j X_{t-j} + \sum_{k=1}^{q} \vartheta_k \epsilon_{t-k} + \epsilon_t, \quad t \in \mathbb{Z}.$$

In the following, a continuous-time counterpart of (3.1) is derived, i.e. on the time domain $\mathbb{R}$. To this end, the noise process $\epsilon = \{\epsilon_t\}_{t\in\mathbb{Z}}$ is replaced by a (two-sided[2]) Lévy process $L = \{L_t\}_{t\in\mathbb{R}}$, i.e. a process with $L_0 = 0$, independent stationary increments, càdlàg sample paths and being continuous in probability.

However, before we can present a formal definition, first of all, the continuation from the backshift operator B has to be reconsidered:

Definition 3.1. (Continuous-time backshift operator) Let $y = \{y_t\}_{t\in\mathbb{R}}$ denote an arbitrary continuous-time process. Then the *continuous-time backshift operator* D, notationally used as $y_t^{(j)} := D^j y_t$, $j \geq 0$, is defined by the following "symbolical" rules:

$$y_t^{(0)} = y_t, \; y_t^{(1)} dt = dy_t, \text{ and } y_t^{(j)} = D y_t^{(j-1)} \text{ for all } j \geq 2.$$

[1] This means that all ϵ_t, $t \in \mathbb{Z}$, are uncorrelated with zero mean and constant variance σ^2, usually denoted by $\epsilon \sim \mathsf{WN}(0, \sigma)$. If, for instance, the law is standard normal, then we shall write $\epsilon \sim \mathcal{N}(0, 1)$ and all ϵ_t, $t \in \mathbb{Z}$, are independent and identically distributed (iid); for a comprehensive introduction to traditional linear discrete-time modelling, we refer to Brockwell and Davis (1991, 2002).

[2] Usually a Lévy process is defined on the non-negative real half-line $[0, \infty)$ as time domain. The extension to the whole real line $\mathbb{R}$ is easily obtained by defining $L_t = \mathbb{1}_{[0,\infty)}(t)\, L_t^{(1)} + \mathbb{1}_{(-\infty,0)}(t)\, L_{-t-}^{(2)}$ for two usual independent Lévy processes $L^{(1)} = \{L_t^{(1)}\}_{t\geq 0}$ and $L^{(2)} = \{L_t^{(2)}\}_{t\geq 0}$.

One is easily inclined to treat D as the differential operator $D = \frac{d}{dt}$. Since not every stochastic continuous-time process is differentiable—in particular Brownian motions are well-known to be nowhere differentiable—this shall only be thought of for notational convenience.[3] Later, however, we will adopt an alternative notation in the context of the state-space representation of a CARMA(p,q) process, which treats D more carefully.

Example 3.2. (Continuous-time counterpart of AR(1)) As a first motivating example, think of an AR(1) process with Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma^2)$ given by

$$(1 - \phi B)X_t = \epsilon_t, \quad t \in \mathbb{Z}, \tag{3.2}$$

whose comparable continuous-time counterpart CAR(1) might take the following form (with a slightly different polynomial paramtrization):

$$(D + \alpha)y_t = DW_t \quad t \in \mathbb{R},$$

where $W = \{W_t\}_{t\in\mathbb{R}}$ denotes a standard Brownian motion. With Definition 3.1, this is equivalent to writing

$$dy_t + \alpha y_t dt = dW_t$$

which is the sde defining the famous *Ornstein-Uhlenbeck (OU) process*.[4] It is a well-known result that every solution y_t satisfies the Markovian relations

$$y_t = e^{-\alpha(t-s)}y_s + \int_s^t e^{-\alpha(t-u)}dW_u, \quad \forall s \leq t \in \mathbb{R}. \tag{3.3}$$

This process is known to be stationary if and only if $\alpha > 0$ and y_0 is independent of $\{W_t\}_{t\geq 0}$ having the same distribution as $\int_0^\infty e^{-\alpha u}dW_u$, which is $\mathcal{N}(0, \frac{1}{2\alpha})$. To convince oneself that CAR(1) and AR(1) are closely related not only with respect to their notational similarity, consider a discretely sampled version $\hat{y} = \{\hat{y}_t\}_{t\in\mathbb{Z}}$ of the OU process $y = \{y_t\}_{t\in\mathbb{R}}$ observed on the time grid $\mathbb{Z}$, which can be written as

$$\hat{y}_t = \phi\hat{y}_{t-1} + \epsilon_t, \quad t \in \mathbb{Z},$$

with $\phi = e^{-\alpha}$ and $\epsilon_t = \int_{t-1}^t e^{-\alpha(t-u)}dW_u$ which yields a Gaussian noise sequence $\epsilon \sim \mathcal{N}(0, \sigma^2)$ with appropriate σ^2, leading us back to equation (3.2). However, as $e^{-\alpha} > 0$, AR(1) processes

[3] For this reason, we have avoided the term "differential operator" for D and named it "continuous-time shift operator".

[4] Uhlenbeck and Ornstein (1930)

with $\phi \leq 0$ cannot be discretely sampled versions of CAR(1) processes.[5]

Next, we systematize this derivation to a general result that encompasses all CARMA(p,q) processes with orders $p, q \in \mathbb{N}$, $q < p$.

3.2. Definition and state-space representation

Consider two integers $p, q \in \mathbb{N}$ with $q < p$ and (real-valued) constants $\alpha_1, ..., \alpha_p$ and $\beta_0, ..., \beta_q$ which define the coefficients of the (complex-valued) polynomials

$$\alpha(z) := z^p + \alpha_1 z^{p-1} + ... + \alpha_p, \tag{3.4}$$

$$\beta(z) := \beta_0 + \beta_1 z + ... + \beta_q z^q, \quad z \in \mathbb{C}, \tag{3.5}$$

with $\beta_q = 1$ to avoid ambiguity.

Throughout this thesis, we consider only Lévy processes $L = \{L_t\}_{t\in\mathbb{R}}$ with $\mathsf{E}[L_1] = 0$ and $\mathsf{E}[L_1^2] = 1$.

Definition 3.3. (Lévy-driven CARMA(p,q) process) Let D denote the shift operator from Definition 3.1 and let $\sigma > 0$ be a constant. Then, a (zero-mean, complex-valued) process $y = \{y_t\}_{t\in\mathbb{R}}$ is called *(Lévy-driven) continuous-time autoregressive moving-average process of orders p and q* or simply *CARMA(p,q) process* if it is solving the formal p-th order sde,

$$\alpha(D)y_t = \sigma\beta(D)DL_t, \quad t \geq 0, \tag{3.6}$$

or, more explicitly,

$$y_t^{(p)} + \alpha_1 y_t^{(p-1)} + ... + \alpha_p y_t = \sigma\left(L_t^{(q+1)} + \beta_{q-1}L_t^{(q)} + ... + \beta_0 L_t^{(1)}\right), \quad t \geq 0. \tag{3.7}$$

The associated Lévy process L is called the *background-driving (noise) process.*

Due to the lack of differentiability, the "$D^j L_t$" do not exist in the usual sense. Therefore, an equivalent definition getting along without a shift operator D as on the right-hand side

[5] In that case, the AR(1) process is not stationary, anyway. This theoretical problem is commonly referred to as the *embedding problem*, which shall not be discussed in this thesis. For comprehensive treatments, we refer to Brockwell (1995), Brockwell and Brockwell (1999), Thornton and Chambers (2011) as well as Cochrane (2012), for instance.

of equations (3.6) or (3.7) is known as the *state-space representation* of a CARMA process, which is notationally less convenient but formally more precise and has several advantages for further considerations.

Definition 3.4. (State-space representation of a CARMA(p,q) process) A (zero-mean, complex-valued) CARMA(p,q) process $y = \{y_t\}_{t\in\mathbb{R}}$ can also be defined by the *observation equation*

$$y_t = \sigma \mathbf{b}^\top \mathbf{x}_t, \quad t \in \mathbb{R}, \tag{3.8}$$

where the *state vector* process $\mathbf{x} = \{\mathbf{x}_t\}_{t\in\mathbb{R}}$ with $\mathbf{x}_t = (x_t, x_t^{(1)}, ..., x_t^{(p-1)})^\top \in \mathbb{C}^p$, $t \in \mathbb{R}$, satisfies the *state equation*

$$d\mathbf{x}_t = \mathbf{A}\mathbf{x}_t dt + \mathbf{1}_p dL_t, \tag{3.9}$$

which is a first order multivariate sde with

$$\mathbf{A} := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_p & -\alpha_{p-1} & -\alpha_{p-2} & \dots & -\alpha_1 \end{bmatrix} \in \mathbb{R}^{p\times p},\ \mathbf{1}_p := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \text{ and } \mathbf{b} := \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-2} \\ \beta_{p-1} \end{pmatrix} \in \mathbb{R}^p,$$

such that $\beta_q = 1$ and $\beta_j = 0$ for every $q < j \leq p$. The pair of equations (3.8) and (3.9) forms the *state-space representation* of $y = \{y_t\}_{t\in\mathbb{R}}$ and is equivalent to the formal p-th order sde (3.6) introduced in Definition 3.3.

Remark 3.5. The matrix $\mathbf{A}$ in Definition 3.4 is called *companion matrix* (of the CARMA process y). Its characteristic and minimal polynomials coincide with the polynomial α in (3.4) (such that its eigenvalues are the roots of α).

Remark 3.6. Every solution $\mathbf{x}_t$ of the state equation (3.9) satisfies the relations

$$\mathbf{x}_t = e^{\mathbf{A}(t-s)}\mathbf{x}_s + \int_s^t e^{\mathbf{A}(t-u)}\mathbf{1}_p dL_u, \quad \forall s \leq t \in \mathbb{R}, \tag{3.10}$$

similarly to the univatiate Ornstein-Uhlenbeck result (3.3) in our motivating Example 3.2. The matrix exponential $e^{\mathbf{M}}$ is well-defined for every square matrix $\mathbf{M} \in \mathbb{C}^{p\times p}$ by traditional series expansion $e^{\mathbf{M}} := \sum_{n=0}^{\infty} \frac{\mathbf{M}^n}{n!}$, with the convention that $\mathbf{M}^0 = \mathbf{I}_p$, the $(p \times p)$-identity matrix and the stochastic integral in (3.10) is a special case of integration with respect to semimartingales.[6] Due to the fact that the increments of L are independent, we can furthermore infer that (3.10) represents a Markovian process $\mathbf{x}$.

[6] see for instance Protter (2004), Part II

Because $y = \{y_t\}_{t\in\mathbb{R}}$ is a linear combination of the marginal state processes $x^{(j)} = \{x_t^{(j)}\}_{t\in\mathbb{R}}$, $0 \leq j \leq p-1$, of the vector process $\mathbf{x} = (x, x^{(1)}, ..., x^{(p-1)})^\top$, stationarity of y is tantamount to stationarity of $\mathbf{x}$. In the following, we discuss several conditions that are necessary and sufficient for $\mathbf{x}_t$ to be a stationary solution of (3.9)/(3.10).

3.3. Stationarity

Before we establish a connection between traditional ARMA and the newer CARMA concepts regarding stationarity, let us briefly recall the definitions of some related terms.

Definition 3.7. (Stationarity) A stochastic process $y = \{y_t\}_{t\in\mathbb{R}}$ is said to be (*strictly*) *stationary* if its finite-dimensional distributions are invariant with respect to time shift, that is, for any $n \in \mathbb{N}$ and any finite time grid $\{t_1 < ... < t_n\} \subset \mathbb{R}$, the joint law of $(y_{t_1}, ..., y_{t_n})^\top$ is identical to that of $(y_{t_1+s}, ..., y_{t_n+s})^\top$ for any $s \geq 0$.

Thus, stationarity particularly implies that all moments of the finite-dimensional distribution are invariant under time shift, especially this yields $\mathsf{E}[y_t] = const$ for every $t \in \mathbb{R}$ and the invariance of autocovariance and autocorrelation (as far as the respective moments exist):

Definition 3.8. (Autocovariance and autocorrelation) Let $y = \{y_t\}_{t\in\mathbb{R}}$ denote a stationary stochastic process with finite second moments. Then, its

(i) *autocovariance function* (*ACVF*) is defined by

$$\gamma_y(s) := \mathsf{Cov}(y_t, y_{t+s}) \equiv \mathsf{Cov}(y_0, y_s), \quad s \in \mathbb{R},$$

independently of $t \in \mathbb{R}$,

(ii) *autocorrelation function* (*ACF*) is defined by

$$\rho_y(s) := \frac{\gamma_y(s)}{\gamma_y(0)}, \quad s \in \mathbb{R}. \tag{3.11}$$

If the finite-dimensional distribution of y exhibits a time-shift invariance only up to the second moments, then y is referred to as *weak* or *covariance stationary*.

As the autocorrelation function ρ_y is just a simple linear correlation coefficient, it quantifies the linear dependence of a time series on its past values. It is natural to expect that $|\rho_y(s)| \leq 1 = \rho_y(0)$ is decreasing as s is increasing.[7] We will make use of ACF plots ($\rho_y(s)$ against s) lateron in order to a priori detect (covariance) stationarity patterns in our data in Chapter 4.

For traditional ARMA processes, stationarity conditions are closely related to the roots of the corresponding characteristic polynomials ϕ and ϑ. As we will see in the following, it is the similar case with their continuous-time counterparts.

Denote by $\lambda_1, ..., \lambda_r$, $r \leq p$, and $\mu_1, ..., \mu_s$, $s \leq q$ the (distinct) roots of the polynomials α and β in (3.4) and (3.5), respectively, such that they can be factorized into

$$\alpha(z) = \prod_{i=1}^{r} (z - \lambda_i)^{m(\lambda_i)}, \quad z \in \mathbb{C},$$
$$\beta(z) = \prod_{i=1}^{s} (z - \mu_i)^{m(\mu_i)}, \quad z \in \mathbb{C},$$

with m denoting the multiplicity of each root. Moreover, let

$$\mathcal{R}^\alpha := \{\lambda \in \mathbb{C} : \alpha(\lambda) = 0\} = \{\lambda_1, ..., \lambda_r\} \text{ and } \mathcal{R}^\alpha_\pm := \{\lambda \in \mathcal{R}^\alpha : \Re(\lambda) \gtrless 0\} \subset \mathcal{R}^\alpha,$$
$$\mathcal{R}^\beta := \{\mu \in \mathbb{C} : \beta(\mu) = 0\} = \{\mu_1, ..., \mu_s\} \text{ and } \mathcal{R}^\beta_\pm := \{\mu \in \mathcal{R}^\beta : \Re(\mu) \gtrless 0\} \subset \mathcal{R}^\beta,$$

denote the sets of the roots of α and β (with positive and/or negative real parts, respectively). This will be of further importance in the next sections.

Now we are ready to concentrate our attention to stationarity of CARMA processes.

3.3.1. Necessary and sufficient conditions

Throughout this section, we will successively collect various assumptions with regards to the roots of the CARMA α and β, starting with the following which we will assume to hold for the rest of this thesis, without loss of generality and without mentioning it explicitly.

Condition 3.9. The characteristic CARMA polynomials α and β have no common zeroes,

[7] In particular, an iid white noise sequence is expected to have zero correlation for lags $s > 0$.

i.e. $\mathcal{R}^\alpha \cap \mathcal{R}^\beta = \emptyset$. In this case, the (meromorphic) function

$$z \mapsto \frac{\beta(z)}{\alpha(z)} = \frac{(z+\mu_1)^{m(\mu_1)} \cdots (z+\mu_s)^{m(\mu_s)}}{(z-\lambda_1)^{m(\lambda_1)} \cdots (z-\lambda_r)^{m(\lambda_r)}}, \quad z \in \mathbb{C},$$

has singularities exactly at the points of $\mathcal{R}^\alpha$.

According to Brockwell and Lindner (2009), Theorems 3.3 and 4.2, existence and uniqueness of a strictly stationary solution $\mathbf{x}$ of the state equation (3.9) is equivalent to $\mathsf{E}[(\log|L_1|)_+] < \infty$ and the following condition:

Condition 3.10. All zeroes $\lambda_1, ..., \lambda_r$ of α lie *outside* the imaginary axis, that is, the real parts are either strictly positive or strictly negative (i.e. $\mathcal{R}^\alpha = \mathcal{R}^\alpha_\pm$).

A given stationary solution of (3.9) is proven to be uniquely representable as

$$\mathbf{x}_t = \int_{-\infty}^{\infty} g_{\mathbf{x}}(t-u)dL_u, \quad t \in \mathbb{R},$$

where the $\mathbb{C}^p$-valued *kernel function* $g_{\mathbf{x}}$ is expressible by $g_{\mathbf{x}}(t) = e^{\mathbf{A}t}\mathbf{1}_p$, $t \in \mathbb{R}$.

We embed this result for a CARMA process in the following statement.

Proposition 3.11. (Stationarity of a CARMA process) *The CARMA(p,q) process y given by equations (3.8) and (3.9) is stationary if and only if Condition 3.10 is satisfied and* $\mathsf{E}[(\log|L_1|)_+] < \infty$. *In this case, y can be uniquely represented by*

$$y_t = \int_{-\infty}^{\infty} g(t-u)dL_u, \quad t \in \mathbb{R}, \tag{3.12}$$

with $\mathbb{C}$-valued kernel function

$$g(t) := \sigma\mathbf{b}^\top g_{\mathbf{x}}(t) = \sigma\mathbf{b}^\top e^{\mathbf{A}t}\mathbf{1}_p, \quad t \in \mathbb{R}.$$

This can be written in one of the two equivalent forms

$$g(t) = \frac{\sigma}{2\pi i}\int_\varsigma e^{tz}\frac{\beta(z)}{\alpha(z)}dz, \quad t \in \mathbb{R}, \tag{3.13}$$

$$= \sigma \sum_{\lambda \in \mathcal{R}^\alpha_\pm} \mathsf{Res}_{z=\lambda}\left(e^{zt}\frac{\beta(z)}{\alpha(z)}\right), \quad t \in \mathbb{R}. \tag{3.14}$$

In line (3.13), ζ is denoting a simple closed curve within the complex plane encircling all roots $\lambda \in \mathcal{R}_{\pm}^{\alpha}$, whereas in line (3.14), $\mathsf{Res}_{z=\lambda}(f(z))$ *denotes the residue of the function f at $z = \lambda$, that is*

$$\mathsf{Res}_{z=\lambda}\left(e^{zt}\frac{\beta(z)}{\alpha(z)}\right) := \frac{1}{(m(\lambda)-1)!}\left[\frac{d^{m(\lambda)-1}}{dz^{m(\lambda)-1}}\left((z-\lambda)^{m(\lambda)}e^{zt}\frac{\beta(z)}{\alpha(z)}\right)\right]_{z=\lambda} \tag{3.15}$$

for any root $\lambda \in \mathcal{R}_{\pm}^{\alpha}$.

Proof. See Brockwell and Lindner (2009), Theorem 3.3 (for the case under Condition 3.9 as we assume here w.l.o.g.) and Theorem 4.2 for the more general case of α and β possibly having common zeroes. □

One can furthermore show that the autocovariance functions of $\mathbf{x}$ and y can be given by

$$\gamma_{\mathbf{x}}(s) = \mathsf{Cov}(\mathbf{x}_{t+s}, \mathbf{x}_t) = e^{\mathbf{A}|s|}\Sigma,$$

and

$$\gamma_y(s) = \mathsf{Cov}(y_{t+s}, y_s) = \sigma^2 \mathbf{b}^\top e^{|s|\mathbf{A}}\Sigma\mathbf{b}, \tag{3.16}$$

respectively, for $s \in \mathbb{R}$ where

$$\Sigma := \mathsf{Var}(\mathbf{x}_t) = \int_0^\infty [g_{\mathbf{x}}(u)]^2\, du = \int_0^\infty e^{\mathbf{A}u}\mathbf{1}_p\mathbf{1}_p^\top e^{\mathbf{A}^\top u} du.$$

We will come back to further simplifications of equations (3.14) and (3.16) in the context of the following section.

3.3.2. Causality and invertibility

In a discrete time linear modelling setup, an ARMA(p,q) process $X = \{X_t\}_{t\in\mathbb{Z}}$ defined by (3.1) is called *causal* if its autoregressive polynomial ϕ has only roots outside the unit circle. In that case, X admits an MA(∞) representation with respect to the white noise sequence $\epsilon = \{\epsilon_t\}_{t\in\mathbb{Z}}$.[8] Moreover, X is said to be *invertible* if the roots of the moving-average polynomial ϑ lie outside the unit circle. In that case, the noise ϵ can be isolated with an AR(∞) representation with respect to X.

[8] A special case of the *Wold representation*, see for instance Brockwell and Davis (2002), Section 2.6 for more detailed basics.

Similar to the discrete-time case, one asks for CARMA counterparts of the terms causality and invertibility and respective characterizing conditions with respect to the roots of the AR polynomial α and the MA polynomial β.

Before stating formal definitions, observe that one can decompose the residual representation (3.14) of the kernel function g into

$$g(t) = \sigma \sum_{\lambda \in \mathcal{R}_-^\alpha} \mathsf{Res}_{z=\lambda} \left(e^{zt} \frac{\beta(z)}{\alpha(z)} \right) \mathbb{1}_{(0,\infty)}(t) - \sigma \sum_{\lambda \in \mathcal{R}_+^\alpha} \mathsf{Res}_{z=\lambda} \left(e^{zt} \frac{\beta(z)}{\alpha(z)} \right) \mathbb{1}_{(-\infty,0)}(t), \quad t \in \mathbb{R},$$

i.e. we take the sums separately on $\mathcal{R}_-^\alpha$ and $\mathcal{R}_+^\alpha$. This will be of further use in what follows.

Causality of a CARMA process

In general, a stationary process $y = \{y_t\}_{t\in\mathbb{R}}$ given by the representation (3.12) is said to be *causal* if the integral at time t is zero on future paths of L, i.e. (slightly more formally) y is independent of the σ-field generated by $\{L_u : u > t\}$. The following definition formalizes this for (the kernel function of) a stationary CARMA process.

Definition 3.12. (Causality) The Lévy-driven continuous-time stationary process $y = \{y_t\}_{t\in\mathbb{R}}$ with kernel function g is said to be *causal* if its integral representation (3.12) is equal to

$$y_t = \int_{-\infty}^{t} g(t-u) dL_u, \quad t \in \mathbb{R}, \tag{3.17}$$

which is equivalent to g being zero on $(-\infty, 0]$. For the particular CARMA cases (3.13) and (3.14), we obtain

$$\begin{aligned} g(t) &= \frac{\sigma}{2\pi i} \int_{\zeta} e^{\lambda t} \frac{\beta(\lambda)}{\alpha(\lambda)} dz \mathbb{1}_{(0,\infty)}(t) \\ &= \sigma \sum_{\lambda \in \mathcal{R}_-^\alpha} \mathsf{Res}_{z=\lambda} \left(e^{zt} \frac{\beta(z)}{\alpha(z)} \right) \mathbb{1}_{(0,\infty)}(t), \quad t \in \mathbb{R}, \end{aligned} \tag{3.18}$$

for the kernel function.

Finally, we can conclude from (3.18) that the following condition for causality of a stationary CARMA process is necessary and sufficient:

Condition 3.13. All roots of α have strictly negative real parts, i.e. $\mathcal{R}^\alpha = \mathcal{R}_-^\alpha$.

Of course, the causality condition $g \equiv 0$ on $(-\infty, 0]$ is valid in a more general setup beyond the CARMA context. Condition 3.13, however, is a special characterizing condition only for causality of CARMA processes. Notice further that causality ($\mathcal{R}^\alpha = \mathcal{R}^\alpha_-$) is stronger than (i.e. sufficient for) stationarity ($\mathcal{R}^\alpha = \mathcal{R}^\alpha_\pm$).

Invertibility of a CARMA process and recovery of the background driving noise

The notion of invertibility in the context of CARMA processes was first mentioned in Brockwell (2001), Remark 7, in a comparable manner to the term of causality and the corresponding notion well-known from discrete-time linear modelling. Very similar to Condition 3.13, though, the following assumption to the roots of the MA polynomial β is made there:

Condition 3.14. All roots of β have strictly negative real parts, i.e. $\mathcal{R}^\beta = \mathcal{R}^\beta_-$.

We formalize a definition of invertibility as, for instance, Ferrazzano and Fuchs (2013), Definition 3.1, do:

Definition 3.15. (Invertibility) The stationary CARMA process $y = \{y_t\}_{t\in\mathbb{R}}$ given by (3.8) and (3.9) is said to be *invertible* if Condition 3.14 is satisfied.

Invertibility of a CARMA process y is not as trivial as with ARMA processes. It goes back to the *minimum phase* spectral factorization, which we do not dicuss at this point; for further specifics, see Sayed and Kailath (2001). Condition 3.14 enables us to isolate the background driving Lévy process L in terms of past paths of y. Before explaining this in more detail, we want to summarize the characteristics of a CARMA process with respect to Conditions 3.10, 3.13 and 3.14 and state another particular assumption to the autoregressive roots which is of particular interest. Afterwards, the main steps of the recovery scheme of Brockwell, Davis and Yang (2007, 2011) will be presented and key concepts discussed so far will be illustrated by CAR(1) and CARMA(2,1).

The conditions for stationarity, causality and invertibility introduced in this section are summarized in the following box.

Stationarity is given if and only if Condition 3.10 is satisfied, i.e. $\mathcal{R}^\alpha = \mathcal{R}^\alpha_\pm$. In the discrete-time ARMA context, this corresponds to the condition that the roots of the AR polynomial ϕ are outside or inside the unit circle.

Causality is defined by Condition 3.13, i.e. $\mathcal{R}^\alpha = \mathcal{R}^\alpha_-$. In the discrete-time ARMA context, this corresponds to the condition that the roots of the AR polynomial ϕ lie outside the unit circle. The resulting (stationary) continuous-time or discrete-time ARMA process then only depends on past paths of the background driving/white noise process.

Invertibility is defined by Condition 3.14, i.e. $\mathcal{R}^\beta = \mathcal{R}^\beta_-$. In the discrete-time ARMA context, this corresponds to the condition that the roots of the MA polynomial ϑ lie outside the unit circle. The background driving/white noise process can then be isolated (recovered).

The next example illustrates the meaning of each of these conditions by the simplest representative of the CARMA class:

Example 3.16. (CAR(1) process) As motivated in Example (3.2), the CAR(1) process (or OU process) is given by the differential equation

$$(D + \alpha_1)y_t = \sigma D L_t,$$

whose equivalent state-space representation (3.8)–(3.9) is

$$\begin{aligned} y_t &= \sigma x_t, \\ \text{with } dx_t &= -\alpha_1 x_t dt + dL_t, \end{aligned}$$

with the $p(=1)$-dimensional state process $x = \{x_t\}_{t\in\mathbb{R}}$. Its solution admits the Markovian representation

$$y_t = e^{-\alpha_1(t-s)} \underbrace{\sigma x_s}_{y_s} + \sigma \int_s^t e^{-\alpha_1(t-u)} dL_u, \quad \forall s \le t \in \mathbb{R}. \tag{3.19}$$

Since the autoregressive polynomial of the process is of the form $\alpha(z) = z + \alpha_1$ with single root $\lambda = -\alpha_1$, Conditions 3.10 and 3.13 respectively correspond to the cases $\lambda \neq 0$ ($\alpha_1 \neq 0$) and $\lambda < 0$ ($\alpha_1 > 0$). Under Condition 3.13, the stationary OU process $y = \{y_t\}_{t\in\mathbb{R}}$ is causal and takes the integral form

$$y_t = \sigma \int_{-\infty}^{t} e^{-\alpha_1(t-u)} dL_u, \quad t \in \mathbb{R},$$

with kernel function $g(t) = \sigma e^{\lambda t} \mathbb{1}_{[0,\infty)}(t) = \sigma e^{-\alpha_1 t} \mathbb{1}_{[0,\infty)}(t)$.

Remark 3.17. (Inverting a CAR(1) process) In the CAR(1) context, invertibility in terms of the roots of the moving-average polynomial $\beta(z) \equiv 1$ (Condition 3.14) is not meaningful. Brockwell, Davis and Yang (2007), however, treat this special case of approximating the Lévy noise process out of a discretely observed OU process, which is described as follows:

1. They use Lemma 2.1 of Eberlein and Raible (1999) combined with an argument of Pham (1977) for Gaussian noise recovery in order to get the Lévy noise recovered by

$$\begin{aligned} L_t &= \frac{1}{\sigma}\left(y_t - y_0 + \alpha_1 \int_0^t y_u du\right) \\ &= L_s + \frac{1}{\sigma}\left(y_t - y_s + \alpha_1 \int_s^t y_u du\right) \end{aligned} \tag{3.20}$$

   for every $s, t \in [0, T]$ with $s \leq t$.

2. For the iid increments $\Delta L^{(h)} = \{\Delta L_t^{(h)}\}_{t\in\mathbb{R}}$ of the Lévy noise process of step size $h > 0$, this has the form

$$\Delta L_t^{(h)} := L_{th} - L_{(t-1)h} = \frac{1}{\sigma}\left(y_{th} - y_{(t-1)h} + \alpha_1 \int_{(t-1)h}^{th} y_u du\right), \quad t \in [0, T]. \tag{3.21}$$

3. The discrete version of this formula is then replaced by estimators for the OU parameters α_1 and σ and the observations of y (equidistantly spaced by h). Furthermore, the integral on the right-hand side is approximated by $(y_{(t-1)h} + y_{th})/2$ (trapezoidal rule). Conditional on the known parameters and observations, this yields a finite set of iid Lévy increments which may be subject to further analysis, e.g. estimation of an appropriate distribution.

Applications of this relatively easy procedure to real credit market data are, among other analyses, summarized in Chapter 4. We will later see that equations of the form (3.20) will be of greater interest in the recovery of higher order CARMA(p,q) noises.

From now on, we will restrict ourselves to the special case of roots $\lambda \in \mathcal{R}^\alpha_\pm$ only with multiplicity one which will turn out to have very interesting and useful consequences afterwards.

Condition 3.18. All (distinct) zeroes $\lambda_1, ..., \lambda_r$ of α have unit multiplicity, i.e. $r = p$ and $\alpha(z) = (z - \lambda_1) \cdots (z - \lambda_p)$.

For $p > 1$, the case of possibly multiple zeroes can also be covered by this condition when artificially considering such roots as distinct but close and converging to each other.

Remark 3.19. (Decomposition of a CARMA(p,q) process) Under Condition 3.18, the complex residue in (3.15) simplifies to $e^{\lambda t}\beta(\lambda)/\alpha'(\lambda)$ with the usual first derivative α' and, in the context of Proposition 3.11, the kernel g of the stationary CARMA process $y = \{y_t\}_{t\in\mathbb{R}}$ given by (3.8) and (3.9) can be written as

$$g(t) = \sigma \sum_{\lambda\in\mathcal{R}^\alpha_\pm} e^{\lambda t} \frac{\beta(\lambda)}{\alpha'(\lambda)}, \quad t \in \mathbb{R}. \tag{3.22}$$

such that the integral representation (3.12) takes the simple form

$$y_t = \sigma \sum_{\lambda\in\mathcal{R}^\alpha_\pm} \frac{\beta(\lambda)}{\alpha'(\lambda)} \int_{-\infty}^{\infty} e^{\lambda(t-u)} dL_u, \quad t \in \mathbb{R}.$$

This is a weighted sum (linear combination) of p different but *dependent* CAR(1) processes (see Example 3.16) all driven by the same Lévy process! Furthermore, the ACVF of y from equation (3.16) simplifies to

$$\gamma_y(s) = \sigma^2 \sum_{\lambda\in\mathcal{R}^\alpha_\pm} e^{\lambda|s|} \frac{b(\lambda)b(-\lambda)}{a'(\lambda)a(-\lambda)}, \quad s \in \mathbb{R}.$$

As a direct consequence of Remark 3.19, the problem of recovering the background driving Lévy process from a higher-order CARMA(p,q) process can be reduced under Condition 3.18 to the CAR(1) case in Remark 3.17, yielding p different but equivalent equations of the form (3.20), as we will see in the following.

We want to agree upon the validity of Condition 3.18 in addition to Conditions 3.10, 3.13 and 3.14 throughout the rest of this thesis, especially this is assumed for the practical implementations in Chapter 4. Using Remarks 3.17 and 3.19, one immediately obtains a useful concept of recovering (the increments of) the Lévy noise process from CARMA processes of arbitrary orders. As mentioned before, Brockwell (2001) briefly described such a recovery scheme, which is conceptualized more formally in the collaborative works of Brockwell, Davis and Yang (2007, 2011). Their arguments base on some useful observations, reflected in the following:

Remark 3.20. (Inverting a CARMA(p,q) process) Assume that we have observed a continuous path of a CARMA(p,q) process y on the interval $[0, T]$ and the coefficients $\alpha_1, ..., \alpha_p$ and $\beta_0, ..., \beta_q$ are already determined (e.g. by estimation) with appropriately chosen p and $q \geq 1$.

1. The first key idea is that the observation equation (3.8) can be "inverted" yielding a CAR(q) process denoted by $\mathbf{x}^q = \{\mathbf{x}_t^q\}_{t\in[0,T]}$ with

$$d\mathbf{x}_t^q = \left(\mathbf{B}\mathbf{x}_t^q + \frac{1}{\sigma}\mathbf{1}_q y_t\right) dt, \tag{3.23}$$

where $\mathbf{x}^q = \{(x_t^{(0)}, ..., x_t^{(q-1)})^\top\}_{t\in[0,T]}$ is the vector of the first q components of $\mathbf{x}$ and

$$\mathbf{B} := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\beta_0 & -\beta_1 & -\beta_2 & \dots & -\beta_{q-1} \end{bmatrix} \in \mathbb{R}^{q\times q} \text{ and } \mathbf{1}_q := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^q,$$

with $\mathbf{B} := -\beta_0$ and $\mathbf{1}_q := 1$ if $q = 1$. Note that—similarly to the state-space representation (3.8)–(3.9) of y in terms of $\mathbf{x}$—the reverse is now established for $\mathbf{x}$ in terms of y, with a background driving process involving the observable y instead of L. The first q components of $\mathbf{x}$ solving (3.23) satisfy the relations

$$\mathbf{x}_t^q = e^{\mathbf{B}t}\mathbf{x}_0^q + \frac{1}{\sigma}\int_0^t e^{\mathbf{B}(t-u)}\mathbf{1}_q y_u du, \quad t \in [0, T],$$

whereas the remaining parts are obtained recursively by $x^{(j)} = Dx^{(j-1)}$, $q \leq j \leq p-1$. By construction, the eigenvalues of $\mathbf{B}$ are the same as the roots of the polynomial β. Its role changes to that of an autoregressive polynomial in this inverted CAR(q) setting such that Condition 3.14 actually ensures causality (hence stationarity) of the inverted CAR(q) process $\mathbf{x}^q$.

2. The second step is to use the *canonical state vector* $\mathbf{y} := \{(y_{1,t}, ..., y_{p,t})^\top\}_{t\in[0,T]}$ given

by

$$\mathbf{y}_t = \sigma \int_{-\infty}^{t} \mathbf{g}(t-u) dL_u, \quad t \in [0,T],$$

with vector-valued kernel function

$$\mathbf{g}(t) := \left(\frac{\beta(\lambda_1)}{\alpha'(\lambda_1)} e^{\lambda_1 t}, ..., \frac{\beta(\lambda_p)}{\alpha'(\lambda_p)} e^{\lambda_p t} \right)^{\top}, \quad t \in [0,T].$$

The components of $\mathbf{y}$ are the CAR(1) processes already introduced in Remark 3.19 and hence sum up to the original CARMA(p,q) process y, i.e.

$$y_t = y_{1,t} + ... + y_{p,t}, \quad t \in [0,T].$$

This canonical state vector $\mathbf{y}$ (itself driven by L which is aimed to be recovered) can be re-expressed in terms of $\mathbf{x}$ determined by means of y (using the above representation (3.23)) by using the relation

$$\mathbf{y}_t = \beta(\mathbf{D})\mathbf{R}^{-1}\mathbf{x}_t, \quad t \in [0,T], \tag{3.24}$$

where $\mathbf{D} := \mathsf{diag}(\lambda_1, ..., \lambda_p)$ and $\mathbf{R} := (\lambda_j^{i-1})_{i,j=1,...,p}$ is the matrix of the right eigenvectors of $\mathbf{A}$, see Brockwell, Davis and Yang (2011), Remark 4.

3. Finally, in case of distinct zeroes $\lambda_1, ..., \lambda_p$ of α (Condition 3.18), combining Remarks 3.17 and 3.19 yields p *equivalent* representations of L in terms of *any* of the components of $\mathbf{y}$, namely

$$\begin{aligned} L_t &= \frac{\alpha'(\lambda_j)}{\sigma\beta(\lambda_j)} \left(y_{j,t} - y_{j,0} - \lambda_j \int_0^t y_{j,u} du \right), \quad 1 \le j \le p, \\ &= L_s + \frac{\alpha'(\lambda_j)}{\sigma\beta(\lambda_j)} \left(y_{j,t} - y_{j,s} - \lambda_j \int_s^t y_{j,u} du \right) \end{aligned} \tag{3.25}$$

for any $s, t \in [0,T]$ with $s \le t$.[9]

4. Finally, one can choose *any arbitrary* j in equation (3.25) and discretize it in order to approximate L. We end up at a discretized scheme for the iid Lévy increments, similar to Remark 3.17.[10]

[9] This is obtained similarly to equation (3.20) combining Lemma 2.1 of Eberlein and Raible (1999) and arguments by Pham (1977) as pointed out by Brockwell, Davis and Yang (2011).

[10] As the most computationally feasible choice Brockwell, Davis and Yang (2011) recommend an index j (if there is any) for which λ_j has no imaginary part.

Notice that, in the case of $p > 1$ and $q = 0$, we directly start with $x_t^{(0)} = \frac{1}{\sigma} y_t$, $t \in [0, T]$, instead of the more involved inverted relations (3.23) and proceed at step 2., equation (3.24), after having determined the remaining components of the vector $\mathbf{x}$ by $x^{(j)} = Dx^{(j-1)}$, $1 \leq j \leq p-1$. The simplest case $p = 1$ and $q = 0$ has already been studied in Remark 3.17.

We review the key steps in the following example:

Example 3.21. (CARMA(2,1) process) For the sake of simplicity, assume that $\sigma = 1$. In this case, y is given by the formal 2nd order sde

$$(D^2 + \alpha_1 D + \alpha_2) y_t = (\beta_0 + D) D L_t,$$

or the equivalent state-space representation

$$\begin{aligned} y_t &= \beta_0 x_t + x_t^{(1)}, \\ dx_t &= x_t^{(1)} dt, \\ dx_t^{(1)} &= (-\alpha_2 x_t - \alpha_1 x_t^{(1)}) dt + dL_t, \end{aligned}$$

which is solved by

$$y_t = (\beta_0, 1) e^{\mathbf{A}(t-s)} \begin{pmatrix} x_s \\ x_s^{(1)} \end{pmatrix} + \int_s^t (\beta_0, 1) e^{\mathbf{A}(t-u)} \begin{pmatrix} 0 \\ 1 \end{pmatrix} dL_u, \quad \forall s \leq t \in \mathbb{R}. \tag{3.26}$$

In our notation, the characteristic polynomials are of the form $\alpha(z) = z^2 + \alpha_1 z + \alpha_2$ and $\beta(z) = \beta_0 + z$. Then, the (possibly complex-valued) roots of α are generally satisfying $\lambda = -(\alpha_1 \pm \bar{\alpha})/2$ with $\bar{\alpha} := \sqrt{\alpha_1^2 - 4\alpha_2}$. Condition 3.10 (stationarity) corresponds to $\Re(\alpha_1 \pm \bar{\alpha}) \neq 0$ and Condition 3.13 (causality) is equivalent to $\Re(\alpha_1 \pm \bar{\alpha}) > 0$ which is given if

$$\text{either } \alpha_1^2 < 4\alpha_2 \quad \wedge \quad \alpha_1 > 0, \tag{3.27}$$

$$\text{or } \alpha_1^2 \geq 4\alpha_2 \quad \wedge \quad \alpha_1 > \bar{\alpha}. \tag{3.28}$$

Furthermore, Condition 3.14 (invertibility) is satisfied if the single (real-valued) root $\mu = -\beta_0$ is negative, i.e. $\beta_0 > 0$. In case of Condition 3.18 (non-multiple roots) which is fullfilled if and only if $\alpha_1^2 \neq 4\alpha_2$ in (3.28), we may divide by $\bar{\alpha}$ and the matrix exponential $e^{\mathbf{A}t}$ in (3.26) has the explicit expression

$$e^{\mathbf{A}t} = \frac{e^{-\frac{\alpha_1}{2}}}{2\bar{\alpha}} \begin{bmatrix} (\alpha_1 + \bar{\alpha}) e^{\frac{\bar{\alpha}t}{2}} - (\alpha_1 - \bar{\alpha}) e^{-\frac{\bar{\alpha}t}{2}} & 2\left(e^{\frac{\bar{\alpha}t}{2}} - e^{-\frac{\bar{\alpha}t}{2}}\right) \\ 2\alpha_2 \left(e^{-\frac{\bar{\alpha}t}{2}} - e^{\frac{\bar{\alpha}t}{2}}\right) & (\alpha_1 + \bar{\alpha}) e^{-\frac{\bar{\alpha}t}{2}} - (\alpha_1 - \bar{\alpha}) e^{\frac{\bar{\alpha}t}{2}} \end{bmatrix}, \quad t \in \mathbb{R}. \tag{3.29}$$

Hence, the kernel function of the stationary integral representation (3.26) can be written as

$$\begin{aligned}
g(t) &= (\beta_0, 1)^\top e^{\mathbf{A}t} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&= \frac{1}{\bar{\alpha}} \left[\left(\beta_0 - \frac{1}{2}(\alpha_1 - \bar{\alpha}) \right) e^{-\frac{1}{2}(\alpha_1 - \bar{\alpha})t} - \left(\beta_0 - \frac{1}{2}(\alpha_1 + \bar{\alpha}) \right) e^{-\frac{1}{2}(\alpha_1 + \bar{\alpha})t} \right].
\end{aligned}$$

In terms of the two distinct zeroes $\lambda_1 \neq \lambda_2$ (Condition 3.18), g can further be simplified using Remark 3.19: First calculating $\alpha'(z) = 2z + \alpha_1$, then observing that $\alpha'(\lambda) = \pm\bar{\alpha}$, yields

$$g(t) = (w_1 e^{\lambda_1 t} + w_2 e^{\lambda_2 t}) \mathbb{1}_{(0,\infty)}(t), \quad t \in \mathbb{R},$$

which turns out to be the weighted sum of two OU kernels $g_1(t) = e^{\lambda_1 t}\mathbb{1}_{(0,\infty)}(t)$ and $g_2(t) = e^{\lambda_2 t}\mathbb{1}_{(0,\infty)}(t)$. Controlling for the weights, one can easily check that

$$w_1 = \frac{\beta(\lambda_1)}{\alpha'(\lambda_1)} = \frac{\beta_0 + \lambda_1}{\bar{\alpha}} = \frac{1}{\bar{\alpha}} \left(\beta_0 - \frac{1}{2}(\alpha_1 - \bar{\alpha}) \right), \tag{3.30}$$

$$w_2 = \frac{\beta(\lambda_2)}{\alpha'(\lambda_2)} = -\frac{\beta_0 + \lambda_2}{\bar{\alpha}} = -\frac{1}{\bar{\alpha}} \left(\beta_0 - \frac{1}{2}(\alpha_1 + \bar{\alpha}) \right). \tag{3.31}$$

Remark 3.22. (Inverting a CARMA(2,1) process) Assume that we have observed a CARMA(2,1) process y continuously on the interval $[0, T]$. Applying Remark 3.20 for $p = 2$ and $q = 1$, we must determine $\mathbf{x} = (x^{(0)}, x^{(1)})^\top$ whose first component is given by the inverted CAR(1) relation (3.23),

$$x_t^{(0)} = e^{-\beta_0 t} x_0^{(0)} + \int_0^t e^{-\beta_0 (t-u)} y_u du, \quad t \in [0, T], \tag{3.32}$$

which is causal if and only if $\beta_0 > 0$ (Condition 3.18) and the second component is recursively obtained by

$$x_t^{(1)} = D x_t^{(0)} = -\beta_0 x_t^{(0)} + y_t, \quad t \in [0, T]. \tag{3.33}$$

Given (continuous) samples of $\{y_t\}_{t\in[0,T]}$ with known CARMA coefficients $\alpha_1, \alpha_2, \beta_0$ such that all Conditions 3.10, 3.13, 3.14 and 3.18 are satisfied and starting values $\mathbf{x}_0 = (x_0^{(0)}, x_0^{(1)})^\top$, one obtains the canonical state vector $\mathbf{y} = \{(y_{1,t}, y_{2,t})^\top\}_{t\in[0,T]}$ consisting of the two CAR(1) components from Remark 3.19 by equation (3.24), namely

$$
\begin{aligned}
\mathbf{y}_t &= \beta(\mathbf{D})\mathbf{R}^{-1}\mathbf{x}_t, \quad t \in [0,T], \\
&= \beta\left(\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\right)\begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix}^{-1}\begin{pmatrix} x_t^{(0)} \\ x_t^{(1)} \end{pmatrix} \\
&= \frac{1}{\lambda_1-\lambda_2}\begin{bmatrix} \beta_0+\lambda_1 & 0 \\ 0 & \beta_0+\lambda_2 \end{bmatrix}\begin{bmatrix} -\lambda_2 & 1 \\ \lambda_1 & -1 \end{bmatrix}\begin{pmatrix} x_t^{(0)} \\ x_t^{(1)} \end{pmatrix} \\
&= \frac{1}{\lambda_1-\lambda_2}\begin{bmatrix} -\lambda_2(\beta_0+\lambda_1) & \beta_0+\lambda_1 \\ \lambda_1(\beta_0+\lambda_2) & -(\beta_0+\lambda_2) \end{bmatrix}\begin{pmatrix} x_t^{(0)} \\ x_t^{(1)} \end{pmatrix} \\
&= \frac{1}{\bar{\alpha}}\begin{pmatrix} (\frac{1}{2}(\alpha_1+\bar{\alpha})\beta_0-\alpha_2)x_t^{(0)} + (\beta_0-\frac{1}{2}(\alpha_1-\bar{\alpha}))x_t^{(1)} \\ (-\frac{1}{2}(\alpha_1-\bar{\alpha})\beta_0+\alpha_2)x_t^{(0)} - (\beta_0-\frac{1}{2}(\alpha_1+\bar{\alpha}))x_t^{(1)} \end{pmatrix}.
\end{aligned}
$$

Inserting the two components (3.32) and (3.33) above yields two equivalent representations from (3.25) for the recovered Lévy process,

$$
\begin{aligned}
L_t &= \frac{1}{w_1}\left(y_{1,t}-y_{1,0}+\frac{1}{2}(\alpha_1-\bar{\alpha})\int_0^t y_{1,u}du\right), \quad t \in [0,T], \\
&= \frac{1}{w_2}\left(y_{2,t}-y_{2,0}+\frac{1}{2}(\alpha_1+\bar{\alpha})\int_0^t y_{2,u}du\right), \quad t \in [0,T],
\end{aligned}
$$

with weights w_1 and w_2 as expressed in equations (3.30) and (3.31).

Beyond the standard parametric approach of Brockwell, Davis and Yang (2007, 2011) for the retrieval of the background driving process presented above, several extending as well as alternative procedures were proposed recently as of the draft stadium of this thesis (early 2013). One generalization to multivariate CARMA (MCARMA) processes can be found in works by Schlemm (2011) and Brockwell and Schlemm (2013). The recovery idea is mainly based on the univariate procedure. Nevertheless, the crucial limitation behind these approaches is the dependence on *a priori* correctly determined degrees p and q such that the observed process is required to be $(p-q-1)$-times differentiable. For this reason, especially if p (q) is chosen too high (low), misspecification of the model can lead to analytical problems due to the lack of necessary derivatives. In order to compensate this limitation, Ferrazzano and Fuchs (2013) established a different strategy based on the discretely sampled version of a CARMA(p,q) process which is shown to be an ARMA(p,$p-1$) process. Their advantage lies in the fact that the orders p and q do not have to be specified a priori, however, at the cost of closed-form expressions for the approximating Lévy increments that are only available for $p \leq 2$ because of too involved calculations.

All these works draw upon the behaviour of discrete observations of a continuously modelled process on a finite time horizon. Therefore, one has to be aware of that the recovered Lévy process (or its iid increments) are only determined up to a certain accuracy which also depends on the step size $h > 0$. It has been shown though in Brockwell, Ferrazzano and Klüppelberg (2012a,b) that in the limiting case, as the step size of the discrete observation grid becomes infinitesimally small ($h \searrow 0$), the approximating Lévy increments converge to the real ones, for which reason CARMA models are of particular interest in estimation of high-frequency data.

3.4. Further topics and references

In this last technical chapter, the most important aspects on continuous-time linear modelling were presented in order to provide a theoretical basis for the recovery approaches to be applied to real credit derivatives prices in the next chapter.

Many applications of CARMA processes consider the special case of L being a subordinator, i.e. a non-decreasing Lévy processes (hence with non-negative increments). Together with a non-negative kernel function g, this results in non-negative CARMA processes and appropriate candidates for modelling variates that only live on the positive real half-line, for instance, stochastic volatility. One famous example is the subordinator-driven OU process first applied by Barndorff-Nielsen and Shephard (2001) to volatility modelling.

In the more general setting of a real-valued Lévy process, as is the case in this thesis, applications of CARMA models to high-frequential observations such as in electricity markets (Bernhardt, Klüppelberg and Meyer-Brandis (2008) and García, Klüppelberg and Müller (2011)) and wind turbulence (Brockwell, Ferrazzano and Klüppelberg (2012a,b)) were studied extensively in recent years. In that context, "high-frequential" is most often referred to as a higher observation rate than once per second.

In the following chapter, the concepts presented on the previous pages are brought together in order to find appropriate CARMA models for discretely (daily) observed premia of CDS on several North American and European reference obligors.

Chapter 4.

Appropriate models for CDS premia

Now we have collected all the basic concepts and tools such that we might search for and choose a suitable model to describe the time series behaviour of CDS premium rates making use of the CARMA class. Before turning to first appropriate proposals, we briefly turn our focus to an exploratory description of the data that will be subject to further technical analyses afterwards.

4.1. Exploratory analysis

4.1.1. The corporate CDS market

Within the credit derivatives markets, CDS have evolved into the most important instruments for buying and selling insurance against default of a reference issuer. At their trading launch in the early 1990s, only banks were predominantly the most active participants. As CDS contracts are exclusively traded over-the-counter, hence not on standardized exchange markets, it is still hard to measure exactly the size of the global markets in terms of, for example, liquidity or market capitalization. According to data assessed semi-annually by the Bank for International Settlement (BIS), CDS started to play a significant role around the turn of the millennium, when the size of notional amounts outstanding[1] grew rapidly from approximately \$300 billion in late 1998 to \$6 trillion in 2004 within only six years. During the next 3–4 years, the exponential extension of the market (fueled by speculators, among

[1] This is the sum of all notional values of each currently outstanding CDS contract.

other factors) reached its temporary high point at 2007/2008, the same time when the recent US subprime crisis peaked, with almost $60 trillion of notional outstanding. Since 2009, it has been decreasing again ending up at around $25 trillion reported in early 2012.[2] In contrast to that, the global bond markets, having existed far longer, showed an exponential growth in the long-term between 1989 ($15.4 trillion) and 2011 ($99.5 trillion).[3]

4.1.2. CDS names subject to our analysis

The fastest way to get a good overview of the evolution of the most actively traded CDS names is considering CDS indices and their respective members. There are two particular families of indices we take a closer look at in the following, namely the *iTraxx Europe* and the *CDX North America*, both provided by *Markit Group Ltd* and certain subsidiaries. Both iTraxx and CDX are themselves tradable yet completely standardized securities—In contrast to their ordinary single name members, they are traded on exchanges. For instance, they are used to hedge or transfer credit risk positions on a portfolio of defaultable obligors. The index levels are calculated on a daily basis as simple equally weighted averages of the CDS premia of their respective members (125 names in each index, the composition being updated twice every year[4]). Hence, they might also be used as benchmarks or leading indicators to reflect the performance in credit quality of an entire region (in this case, developed countries of Europe on the one hand and North America on the other hand). The standard versions of both indices only contain reference issuers that are rated *investment grade* (that is, AAA–BBB) though there are also related *high-yield indices* containing only *speculative grade* issuers (BB and lower), for instance. Figures 4.1.1–4.1.3 illustrate the absolute and relative composition of each of the two indices with respect to industry sector, country of domicile and current long-term issuer credit rating as of November 6, 2012.[5]

From these figures, one can easily read off that the European index has a much higher ratio of banks and other financial corporates (such as insurance companies) with together 36

[2] VAUSE (2010)

[3] For an extensive overview on the development of the CDS markets compared in terms of liquidity/market capitalization, consider for instance periodical statistical reports of the BIS, available online at `www.bis.org/publ`. For the evolution of the global bond market compared to that of the CDS market, see KUSHMA, CLASS and KURZWEIL (2012).

[4] For the current membership lists, see the *Credit Index Annexes* page on *Markit*'s website `www.markit.com`; at the time of download, November 6, 2012, the latest index rolls were series 18 (iTraxx) and 19 (CDX) respectively, see also Section 4.1.3 below.

[5] Credit ratings are categorical ranks, obtained from a one-to-one mapping from the set of disjoint sub-intervals of historical default probabilities in $[0, 1)$ in ascending order (the best grade AAA corresponds to the lowest sub-interval in $[0, 1)$).

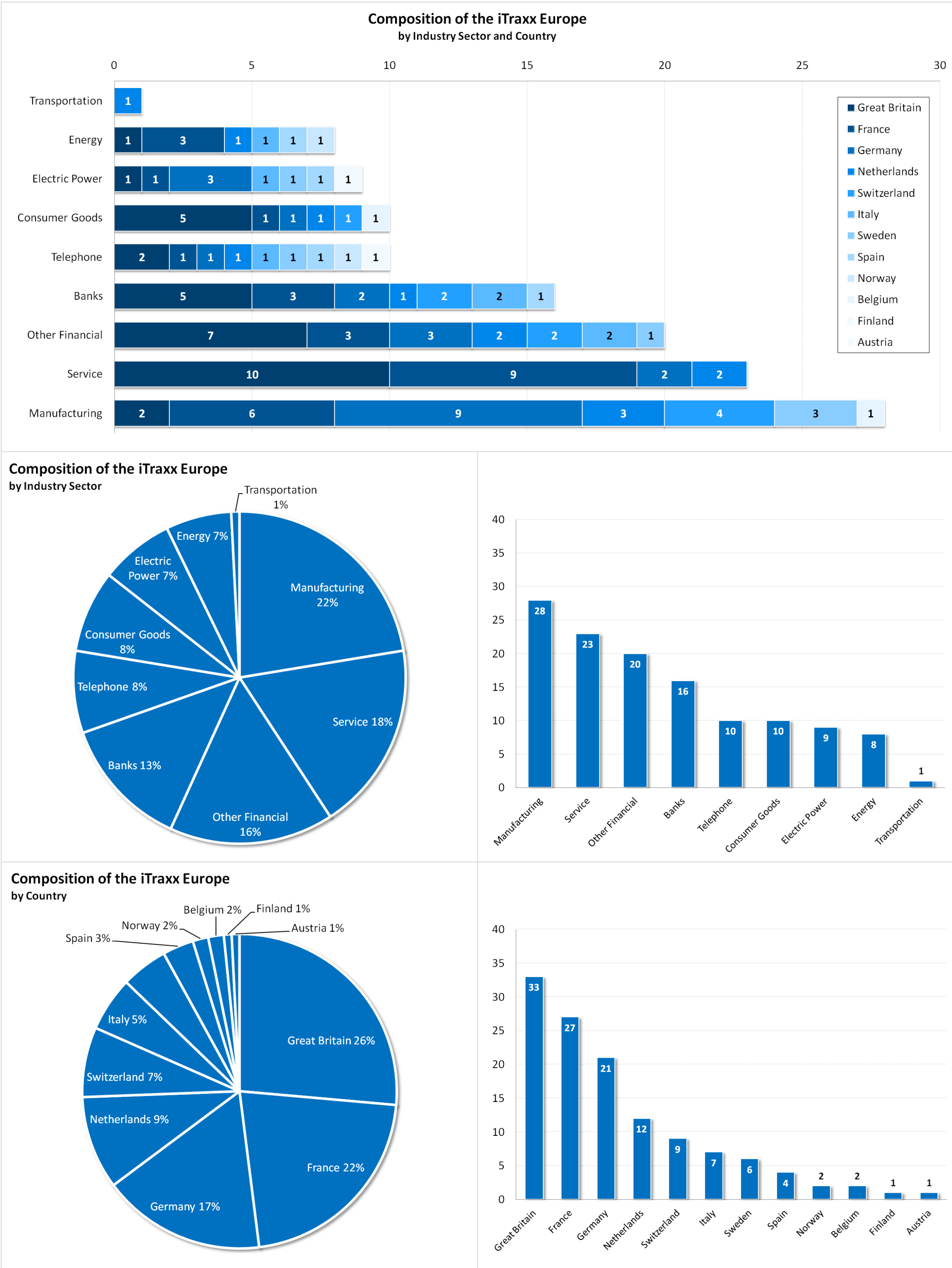

Figure 4.1.1.: Percental and absolute composition by industry sector and country of domicile of the iTraxx Europe Series 18 as of November 6, 2012. (Own graphics, resources: *Reuters/Markit*)

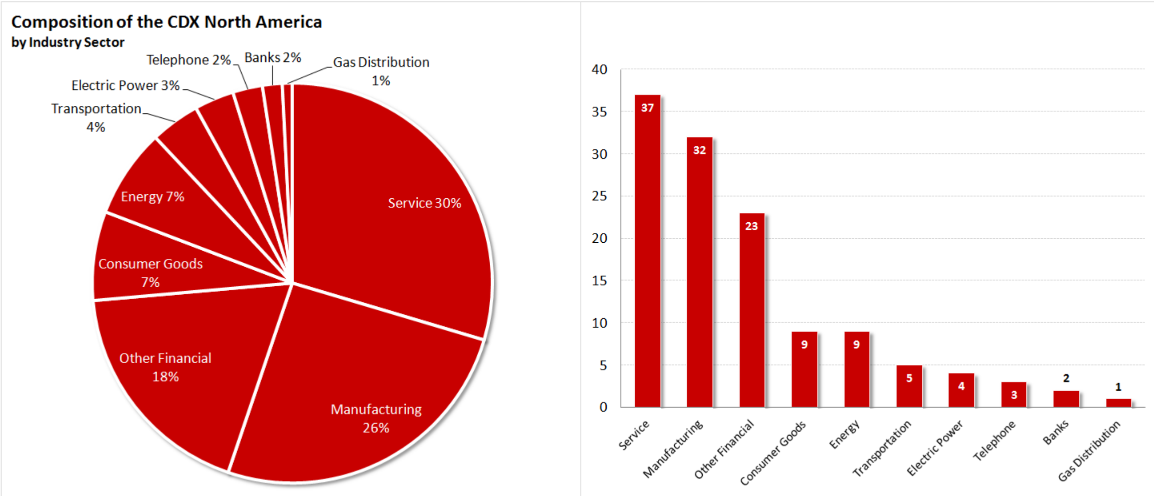

Figure 4.1.2.: Percental and absolute composition by industry sector of the CDX North America Series 19 as of November 6, 2012. The countries of domicile are United States (123 companies) and Canada (2 companies). (Own graphics, resources: *Reuters/Markit*)

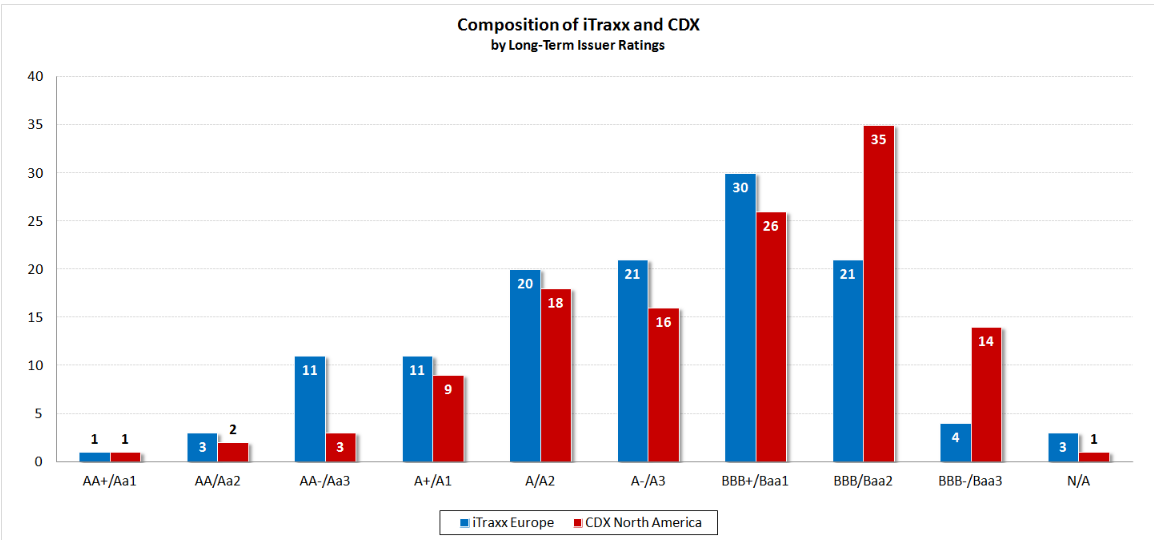

Figure 4.1.3.: Absolute composition of the iTraxx and CDX indices by long-term issuer credit rating as of November 6, 2012. (Own graphics, resources: *Reuters/Markit*)

reference entities (29%) than the North American index with only 25 financial institutions in total (20%), while the CDX consists of more service and manufacturing companies (69 in total or 56%) than the iTraxx (51 in total or 40%). In view of the ratings, the credit quality of the iTraxx companies is higher compared to CDX: The ratio between entities rated A-/A3 or above/below is 54:46 for Europe and 39:61 for North America. However, note that credit ratings only reflect a snapshot of the respective date, in contrast to sector and country belongings which are constant in time (within each fixed membership composition). Thus, the ratings composition might have changed by now.

Concerning the data availability, *Thomson Reuters*[6], one of the leading (financial) news agencies and data suppliers, was chosen as our primary data source for historical daily end-of-day prices from the 250 members of the iTraxx and CDX. As there are credit default swaps for different tenors, whereas our idealized framework assumes independence of the contracts' maturity[7], the most liquid (and therefore, probably best covered and most reliable) tenor $T = 5$ years was selected.

Next, an extensive description of the quality and quantity of the data is given, before appropriate models are tried to be estimated on them in Section 4.2.

4.1.3. Descriptive summary

In total, 250 data sets were downloaded on November 6, 2012, ranging from January 2, 2002 to November 5, 2012 with 2,830 days of possible observations (excluding weekends and holidays). Table 4.1.1 illustrates the coverage ratio of each index staggered by the k-th upper order statistic for several values of $k \in \{1, ..., 125\}$.[8]

At first glance, one can immediately infer from each line of Table 4.1.1 that the iTraxx entities have a higher coverage than those of the CDX index, as is also reflected by the overall average (last line of the table). The average coverage ratio among those companies with 50% or above (that is, $\geq$1,415 observations) amounts to 92.9% for the iTraxx and 84.7% for the CDX companies, respectively (second to last line of the table). Conversely, only 10 iTraxx and 23 CDX companies have a lower coverage. From those, only four companies in

[6] In particular, the software product *Thomson Reuters Eikon with CreditViews* and its interface for *Microsoft Excel* was used to retrieve the historical quotes.

[7] Recalling Chapter 2, we ended up at the model $C_t^* = (1 - R)\,\gamma_t$, $t \geq 0$, which was independent of the tenor T.

[8] The coverage ratio is calculated as the total number of observations divided by the maximum number of possible observations (2,830 in this case).

iTraxx	k	10	25	50	75	100	115	125
	$quantile\left(\frac{125-k}{125}\right)$	2,821	2,804	2,744	2,695	2,311	$1,415 = \frac{2{,}830}{2}$	0
	$avg.num.obs$	2,822	2,817	2,800	2,773	2,719	2,627	2,492
	$avg.cr$	99.7%	99.6%	99.0%	98.0%	96.1%	92.9%	88.1%
CDX	k	10	25	50	75	100	102	125
	$quantile\left(\frac{125-k}{125}\right)$	2,647	2,580	2,487	2,265	1,594	$1,415 = \frac{2{,}830}{2}$	0
	$avg.num.obs$	2,688	2,641	2,586	2,522	2,415	2,396	2,142
	$avg.cr$	95.0%	93.4%	91.4%	89.2%	85.4%	84.7%	75.7%

Table 4.1.1.: Average number of observations (avg.num.obs) and coverage ratios (avg.cr in percent) of the iTraxx Europe and CDX North America ranging from January 2, 2002 to November 5, 2012 (2,830 days of maximum possible observations). For each line, both are calculated conditional on "num.obs≥ k-th upper order statistic of the num.obs". The last line gives the overall (unconditional) avg.num.obs and avg.cr.

each index have less than 1,000 observations within the considered time range. To ensure a high data quality, these eight time series have been excluded, prior to any further analysis, resulting in 242 single-name CDS time series in total.

It might be worth mentioning that the yearly coverage ratios are increasing in the case of iTraxx, as is illustrated by Figure 4.1.4. This is, however, not the case with CDX which exhibits a first rapid growth phase in 2002–2005, followed by a decrease in 2006–2008 and—similarly to iTraxx—a quite constant almost perfect coverage in the latest period 2009–2012.

Now we turn to the quantitative aspects for the rest of this chapter.

4.2. Model selection

Denote by $\mathbb{T} := \{1, ..., T^* = 2830\}$ the discrete time grid of step size $h = 1$ day on which the corresponding CDS premia $\{C_t^*\}_{t\in\mathbb{T}}$ are observed. From the modelling point of view, in order to fit a model to observations, we require them to be sampled from a stationary process. First, consider the historical index levels[9] of iTraxx and CDX which are plotted

[9] Note that, since the series members are updated every six months, the actual iTraxx and CDX indices that are observable and tradable in reality can only be downloaded for these short periods. Thus, these are calculated as equally weighted averages of the CDS premia of the corresponding members at each day by hand for the entire period 2002–2012 ignoring the fact that, in the past, the index composition might have been completely different from what it is currently.

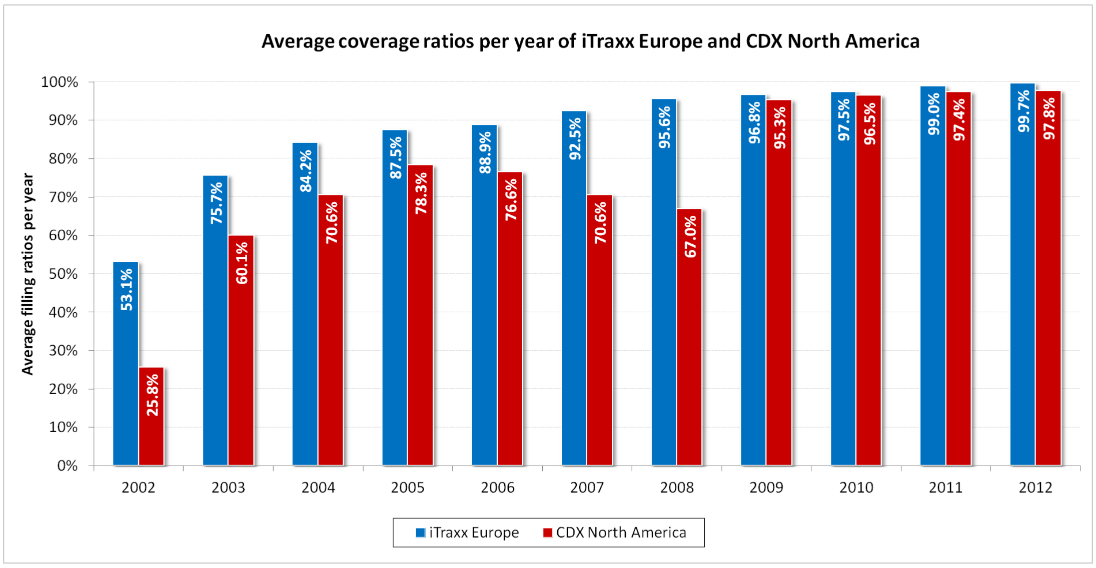

Figure 4.1.4.: Yearly coverage ratios for iTraxx Europe and CDX North America in the period 2002–2012

together in Figure 4.2.1. Note that the standard quotation unit of CDS is *basis points* [*bp*] with 1 bp = .01% = .0001.

4.2.1. Turning to log-differences

Taking first-order differences of the logarithms of a non-stationary process (*log-differences*, $\{\log(C_t^*) - \log(C_{t-1}^*)\}_{t\in\mathbb{T}\setminus\{1\}}$) has proved in many applications before to be an appropriate transformation for obtaining a possible sample of a stationary random process. That this transformation was necessary to be applied to our data, as well, was indeed confirmed by Dickey-Fuller stationarity tests which could not reject the hypothesis of non-stationarity for the index levels but for the log-differences it could. Figures 4.2.1–4.2.4 contain the time series and the ACF[10] plot of the iTraxx index levels and their respective log-differences, which clearify very well the effects of log-differencing. We transfer these insights from this index to every single-name CDS series, since the typical patterns of the time series behaviour are quite similar for all.

Thus, our studies presented in the following sections base solely on the log-differences of

[10] Recall Definition 3.8 for the ACF.

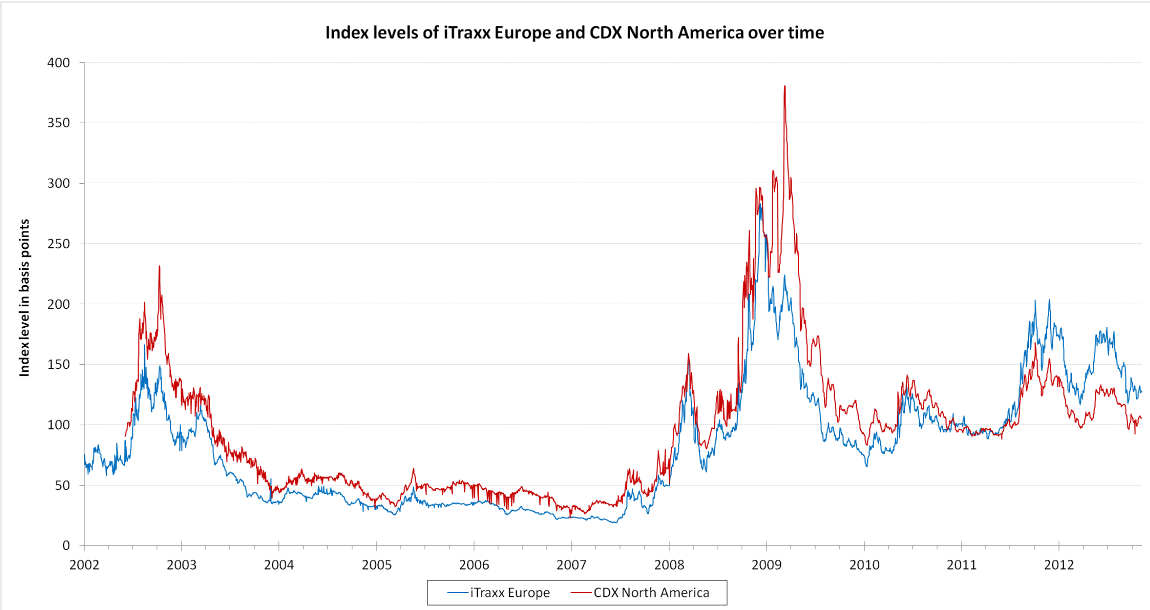

Figure 4.2.1.: Time series plot of the index levels of iTraxx Europe and CDX North America in the period 2002–2012

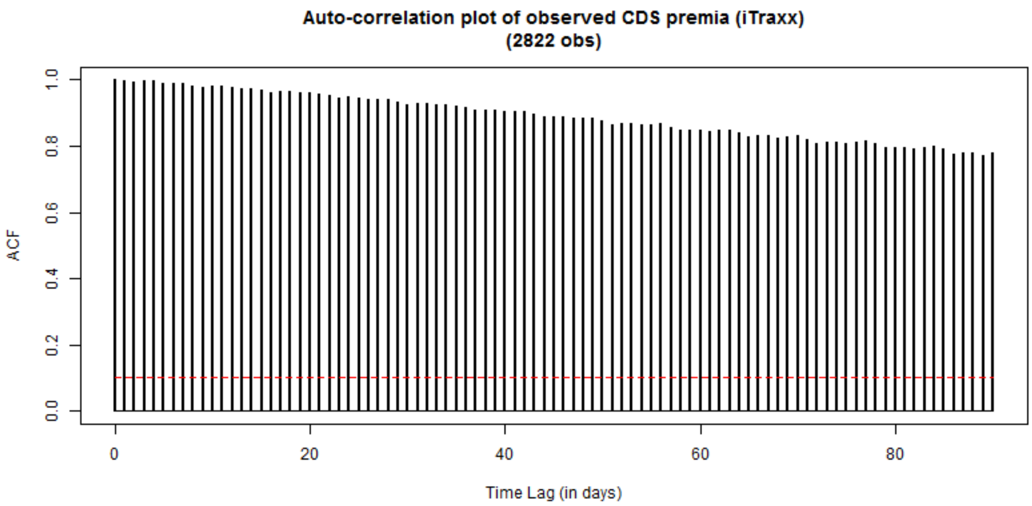

Figure 4.2.2.: Sample ACF plot of the index levels of iTraxx Europe

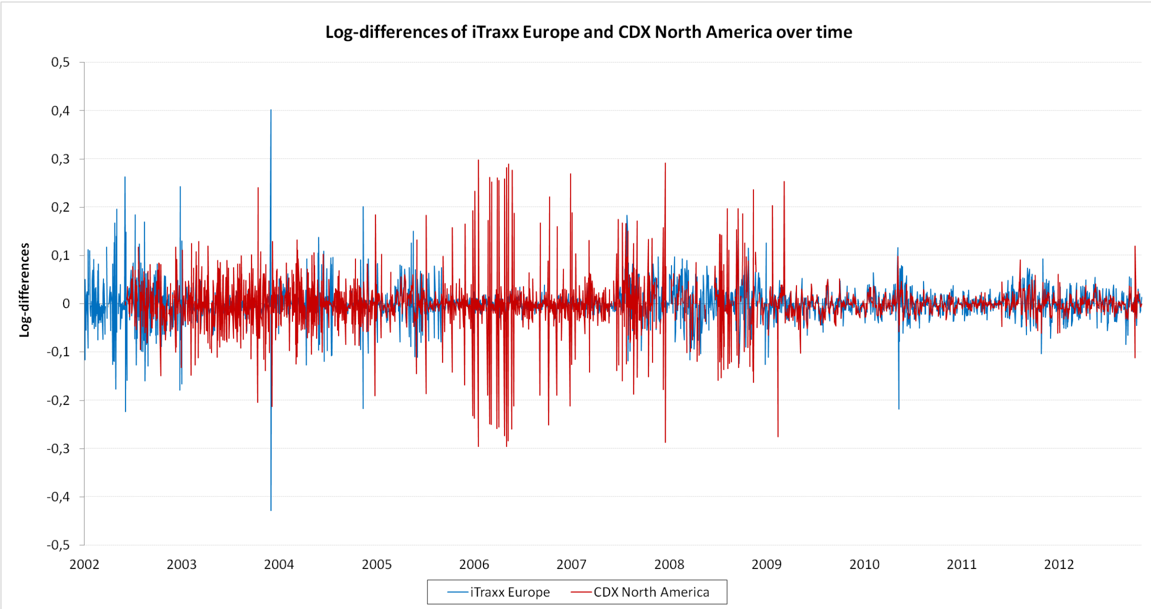

Figure 4.2.3.: Time series plot of the log-differenced index levels of iTraxx Europe and CDX North America in the period 2002–2012

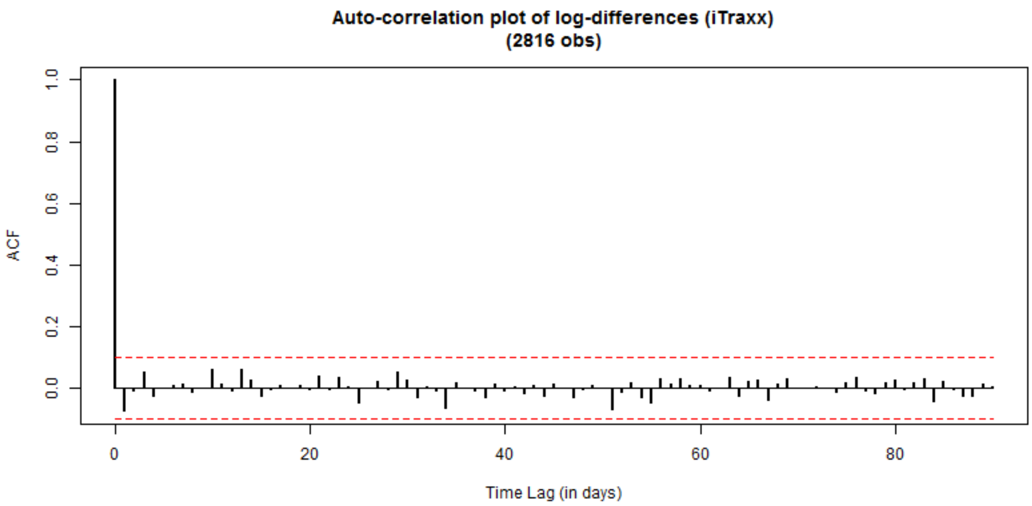

Figure 4.2.4.: Sample ACF plot of the log-differenced index levels of iTraxx Europe

each time series whose discretely sampled versions are henceforth denoted by

$$\hat{y}_t := \log(C_t^*) - \log(C_{t-1}^*), \quad t \in \mathbb{T},$$

with the slightly modified time grid $\mathbb{T} := \{1, ..., T^* - 1 = 2829\}$. Note that by calculating log-differences, each contiguous block of missing values in the CDS time series generates one additional missing value within the log-differenced time series. We accept losing information in favor of gaining possible samples from a stationary time series, though.

The aim of the following analyses is first to choose several continuous-time processes $y := \{y_t\}_{t\in\mathbb{R}}$ from the CARMA class introduced in Chapter 3 and to identify the most appropriate candidates in the sense that the log-differences $\hat{y} := \{\hat{y}_t\}_{t\in\mathbb{T}}$ are approximate samples of y on the discrete time grid $\mathbb{T}$.

As the primary aid for our quantitative analyses, the open-source programming environment `R` (R Development Core Team (2012)) is used together with several foreign libraries, all of them freely available at `www.R-project.org`. In particular, the estimation procedure of the CARMA parameters is done with the friendly assistance of the `R`-package `ctarma` by Tómasson (2012). It allows estimation approaches on either a frequency domain or a time domain based level. The latter is done by (exact) maximum likelihood (ML) whose goodness of fit can be measured by the *Bayesian information criterion* (BIC), defined by

$$BIC(\xi) = -2 \log \ell(\xi) + (p + q + 1) \log(T^* - 1)$$

where the log-likelihood function $\log \ell(\xi)$ is the objective function of the maximization problem

$$\max_{\xi\in\Xi} \log \ell(\xi) = \max_{\xi\in\Xi} \sum_{t=2}^{T^*-1} l(\hat{y}_t \mid \hat{y}_{t-1}, \xi)$$

with respect to the parameter vector ξ containing all coefficients of the AR and MA polynomials as well as $\sigma > 0$ and lying in the set of admissible values denoted by $\Xi \subset \mathbb{R}^{p+q+1} \times (0, \infty)$. The functions $l(\hat{y}_t \mid \hat{y}_{t-1}, \xi)$, $t \in \mathbb{T}\backslash\{1\}$, are the conditional log-likelihoods obtained by using the Kalman-filter (for further details consider Tómasson (2011) for instance).

The BIC is minimized if and only if the (log-)likelihood function is maximized—it will be used as our main criterion by which the appropriate choice of the degrees p and q will be decided in Section 4.2.2. Although the tools and techniques provided by this package are admittedly written for and based on Gaussian driving processes[11], we treat the resulting

[11] For theoretical and numerical aspects regarding the simulation as well as estimation of Gaussian CARMA

estimates as "pseudo-ML" estimates such that we allow for background driving processes different from Brownian motion. The recovery results are then discussed and further analyzed in Section 4.3.

Remark 4.1. As an aside, recall that as the parametrization of a CARMA(p,q) process, including σ, we chose the parameter vector $\xi := (\alpha_1, ..., \alpha_p, \beta_0, ..., \beta_q, \sigma)^\top$ with variable β_0 and fixed $\beta_q = 1$. In contrast to this, the routines of `ctarma` base on the parametrization $\xi' := (\alpha'_1, ..., \alpha'_p, \beta'_0, ..., \beta'_q, \sigma')^\top$ with fixed $\beta'_0 = 1$ and variable β'_q. We might switch from our original parametrization to that of Tómasson (2011) and back by dividing both sides of the CARMA sde (3.7) (or each component of ξ, ξ') by β_0 and β'_q, respectively. To avoid these calculations, we henceforth state the resulting output in terms of ξ'.

Remark 4.2. Furthermore, it is worth mentioning that the routines of `ctarma` are designed in such a way that the admissible parameters $\xi' := (\alpha'_1, ..., \alpha'_p, \beta'_0, ..., \beta'_q, \sigma')^\top$ satisfy the conditions of causality and invertibility, that is, the validity of Conditions 3.13 and 3.14 is ensured with respect to the zeroes of the polynomials α and β, respectively, see also the package manual of Tómasson (2012).

4.2.2. Estimation results and the choice of p and q

Among all possible CARMA(p,q) candidates for $p + q \leq 2$ and $q < p$, we choose that model exhibiting the least BIC value and most plausible parameter estimates (in terms of magnitude) and at the same time lowest possible standard errors. Table 4.2.1 gives a brief summary on how many times which model was optimal. For a complete tabular overview, we refer to the web address `mediatum.ub.tum.de/node?id=1140434` where several appendices are located.

Model	CAR(1)	CAR(2)	CARMA(2,1)
Number of optimal BICs	135	43	64

Table 4.2.1.: Number of optimal BICs for each CARMA(p,q) candidate with $p+q \leq 2$ and $q < p$. The numbers sum up to 242, the final number of analyzed time series.

According to that, more than half of the examined companies (135) fit optimally to a CAR(1) process, whereas the rest of the companies is sharing the other two models (43 CAR(2) and

processes, see Tómasson (2011), for Lévy-driven CARMA processes, see Todorov and Tauchen (2006). For alternative estimation methods such as least squares and their asymptotic properties see Brockwell, Davis and Yang (2011).

	BERKSHIRE		COMCAST	
LL		2531.4193		2592.6106
BIC		-5048.8600		-5171.2352
	Estimate	*(Std.error)*	*Estimate*	*(Std.error)*
$\hat{\alpha}_1$	0.6087	(0.0234)	0.8936	(0.0435)
$\hat{\sigma}$	0.0298	(0.0004)	0.0322	(0.0003)
	DAIMLER		**DEUTSCHE**	
LL		5336.6493		4961.2195
BIC		-10657.4174		-9906.5506
	Estimate	*(Std.error)*	*Estimate*	*(Std.error)*
$\hat{\alpha}_1$	1.3659	(0.0478)	1.5346	(0.0498)
$\hat{\sigma}$	0.0615	(0.0004)	0.0743	(0.0004)

Table 4.2.2.: Excerpt of the CAR(1) estimation results, including the selection criteria *LL* (log-likelihood) and *BIC*, the parameters estimates of $\hat{\alpha}_1$ and $\hat{\sigma}$ and their standard errors. The respective companies are: Berkshire Hathaway, Comcast, Daimler, Deutsche Bank.

64 CARMA(2,1), respectively). Comprehensive details on the estimation results can be found at the aforementioned online resource.

We proceed our investigations with the CAR(1) cases and try to find an answer of what law (the increments of) the background driving processes are following. To this end, they need to be approximated using the recovery methods presented in Chapter 3, in particular Remark 3.17. To conclude, Table 4.2.2 gives a brief overview on the estimated parameters for some picked companies. The full summary is available online as well.

4.3. The background driving process for CAR(1)

The aim of the next step is to illustrate the recovery scheme of Brockwell, Davis and Yang (2007, 2011) of the background driving (noise) process by the CAR(1) examples.

Recall the recovery formula (3.21) from Remark 3.17 for the increments of the background driving Lévy process. We replace the corresponding CAR(1) parameter α_1 as well as the volatility parameter σ by their corresponding (maximum likelihood) estimates $\hat{\alpha}_1$ and $\hat{\sigma}$ found in Section 4.2.2 and the CAR(1) process y by its discretely observed sample $\hat{y}$. Finally, setting $h = 1$ and approximating the Riemann integral by the trapezoidal rule, the recovery formula simplifies to

$$\widehat{\Delta L}_t := \frac{1}{\hat{\sigma}}\left(\hat{y}_t - \hat{y}_{t-1} + \hat{\alpha}_1 \int_{t-1}^{t} \hat{y}_u du\right)$$
$$\approx \frac{1}{\hat{\sigma}}\left(\hat{y}_t - \hat{y}_{t-1} + \frac{\hat{\alpha}_1}{2}(\hat{y}_{t-1} + \hat{y}_t)\right), \quad t \in \mathbb{T}, \tag{4.1}$$

yielding a set of approximate samples $\widehat{\Delta L} := \{\widehat{\Delta L}_t : t \in \mathbb{T}\}$ that is known from BROCKWELL, DAVIS and YANG (2007, 2011) to be iid samples of the Lévy increments ΔL. Checking the ACF plots in Figure 4.3.1 indicates that all remaining autocorrelation is filtered out to the greatest extent.

Thus, the remaining unknown is the distribution law of ΔL. Figure 4.3.2 contains exemplary histograms of the Lévy increments recovered from CDS log-differences by the above scheme (4.1). It can be observed that the distribution to be estimated is quite symmetric around zero and both the left as well as the right-hand tails are unbounded yet relatively heavy. We proceed with a set of candidates of possible parametric distributions that will be subject to further estimations and compared by goodness of fit afterwards.

We review briefly some well-known parametric families that are chosen to be candidates for the distribution of ΔL.

4.3.1. Possible candidates for the noise distribution

Hypothesis 4.3. The recovered iid Lévy increments $\{\widehat{\Delta L}_t\}_{t\in\mathbb{T}}$ are samples of one of the following five distribution laws, each with unbounded support $\mathbb{R}$:

1. *Normal distribution*, given by the density function

$$f_1(x; \mu, \sigma) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2}\right\}, \quad x \in \mathbb{R},$$

   where the parameters $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$ are location and scale parameters, respectively.

2. *Student's t-distribution*, given by the density function

$$f_2(x; \nu) := \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}\left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad x \in \mathbb{R},$$

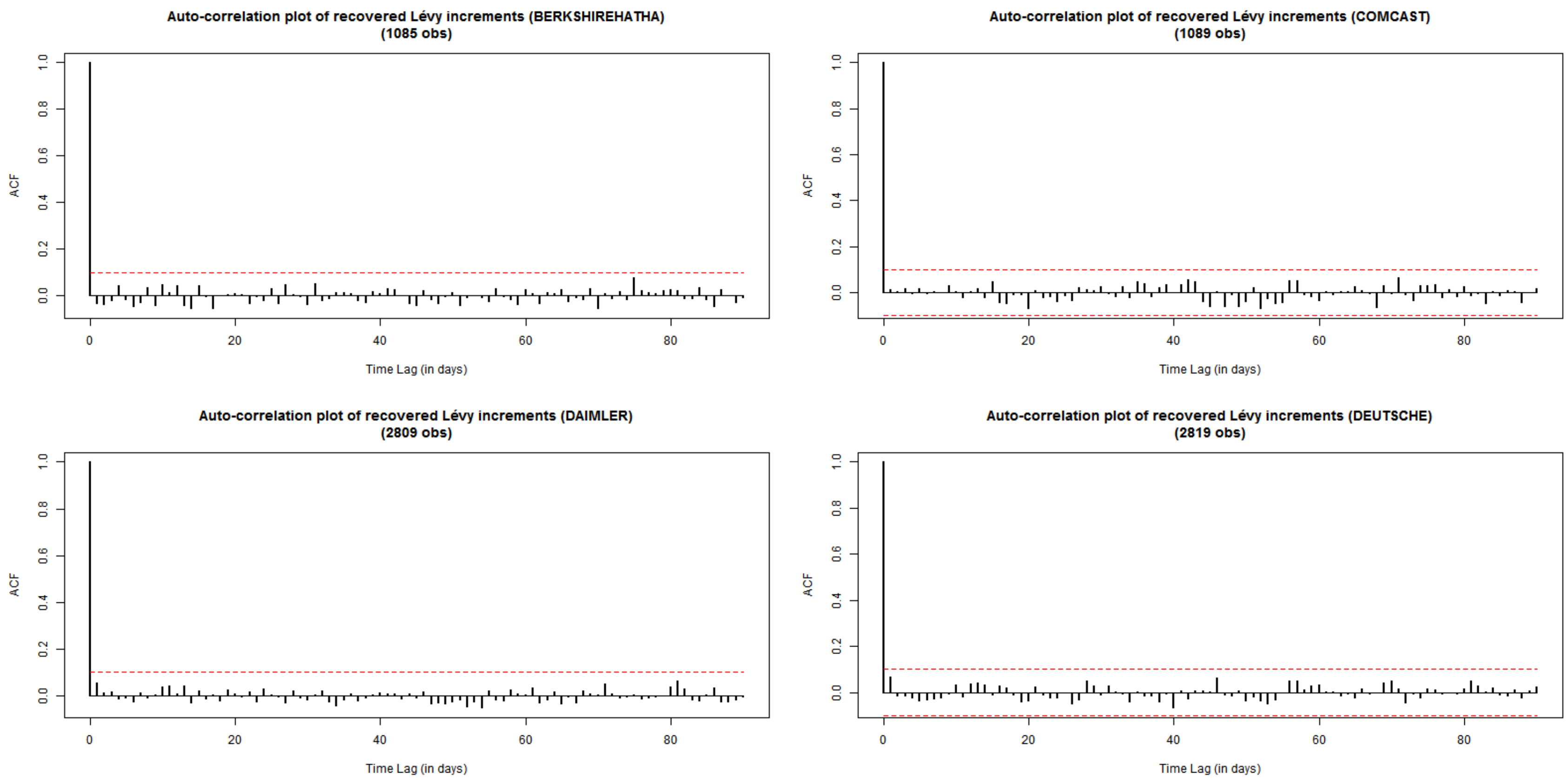

Figure 4.3.1.: Sample ACF plots of the Lévy increments recovered from the estimated CAR(1) models for the log-differenced CDS premium time series. The respective companies are: Berkshire Hathaway, Comcast, Daimler, Deutsche Bank.

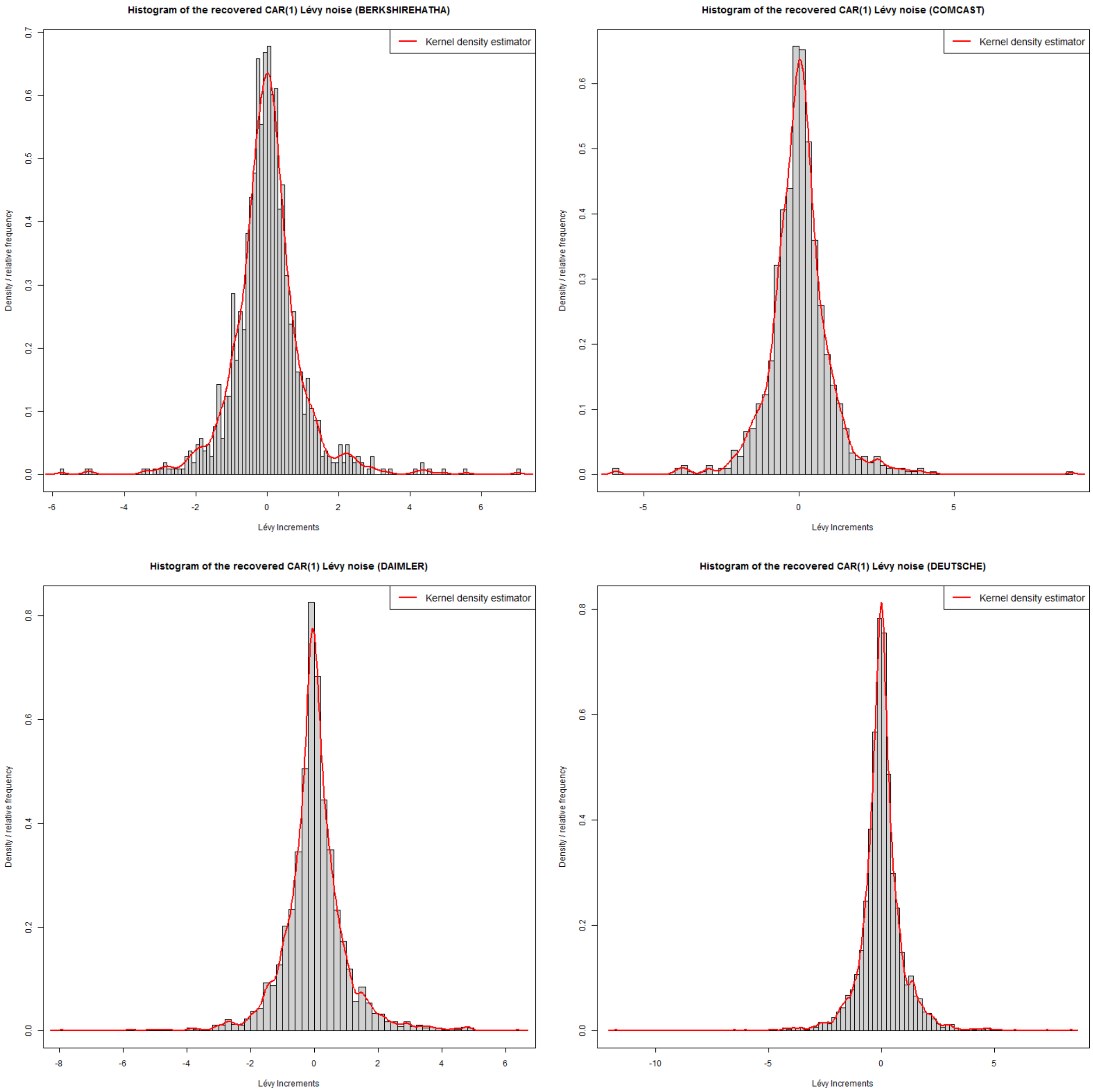

Figure 4.3.2.: Histogram plots of the Lévy increments recovered from the estimated CAR(1) models for the log-differenced CDS premium time series. The respective companies are: Berkshire Hathaway, Comcast, Daimler, Deutsche Bank. The histogram bars are shaded as a grey area and the kernel density estimate is plotted red solid lines.

where the parameter $\nu \in (0, \infty)$, called (*number of*) *degrees of freedom, is a shape parameter* and Γ is the gamma function.

3. *Laplace distribution*, given by the density function

$$f_3(x; \alpha, \beta, \mu) := \begin{cases} \frac{1}{\alpha+\beta} \exp\left\{-\frac{x-\mu}{\alpha}\right\}, & \text{if } x \in (-\infty, \mu], \\ \frac{1}{\alpha+\beta} \exp\left\{-\frac{x-\mu}{\beta}\right\}, & \text{if } x \in [\mu, \infty), \end{cases}$$

where the notation is adapted by Scott (2009) in the R-package `HyperbolicDist`. The parameters $\alpha, \beta \in (0, \infty)$ are shape parameters, $\mu \in \mathbb{R}$ is a location parameter. The Laplace distribution is actually a two-sided mixture of exponential distributions with *intensity parameters* α (for the lower, left-hand tail) and β (for the upper, right-hand tail), respectively.

4. *Normal inverse Gaussian (NIG) distribution*, given by the density function

$$f_4(x; \alpha, \beta, \delta, \mu) := \frac{\alpha\, \delta\, K_1\left(\alpha \sqrt{\delta^2 + (x-\mu)^2}\right)}{\pi \sqrt{\delta^2 + (x-\mu)^2}} \exp\left\{\delta\gamma + \beta(x-\mu)\right\}, \quad x \in \mathbb{R},$$

where the notation is based on Barndorff-Nielsen (1977) and implemented in the R-package `fBasics` by Wuertz and Rmetrics core team (2012). The parameters are

a) $\alpha \in (0, \infty)$: shape parameter for tail heaviness,

b) $\beta \in \mathbb{R}$ with $|\beta| \in (0, \alpha)$: skewness/asymmetry parameter,

c) $\delta \in [0, \infty)$: scale parameter,

d) $\mu \in \mathbb{R}$: location parameter.

and $\gamma := \sqrt{\alpha^2 - \beta^2}$. The function K_1 is the modified Bessel function of the second kind. This distribution is a Normal variance-mean mixture with inverse Gaussian mixing density, see Barndorff-Nielsen (1997) for details.

5. *α-stable distribution* which is not expressible in an analytical way (except for certain parameter values). It is parametrized by

a) $\alpha \in (0, 2]$: shape parameter for stability/tail heaviness,

b) $\beta \in [-1, 1]$: skewness/asymmetry parameter,

c) $\gamma \in (0, \infty)$: scale parameter,

d) $\delta \in \mathbb{R}$: location parameter.

In general, the support equals $\mathbb{R}$—however, in the special case $\alpha = 1$ and $\beta = \pm 1$, it is bounded by μ from below (if $\beta = 1$) and from above (if $\beta = -1$). The normal ($\alpha = 2$), the Cauchy ($\alpha = 1$, $\beta = 0$) and the Lévy ($\alpha = \frac{1}{2}$, $\beta = 1$) distribution all belong to the α-stable class. Since the density has no explicit expression, numerical approximations have to be used.[12]

While the distributions in 2.–4. are special cases of the class of *generalised hyperbolic (GH)* distributions that are characterized by semi-heavy tails, the class of α-stable distributions (5.) is a completely heavy tailed class, except for the normal distribution, which is the only light tailed distribution in this class.[13]

For each of the above five distribution laws, a maximum likelihood estimation is performed by maximizing the log-likelihood function $\log \ell_i(\theta_i)$ given explicitly by

$$\log \ell_i(\theta_i) := \log \left(\prod_{t=1}^{T^*-1} f_i(\widehat{\Delta L}_t \mid \theta_i) \right) = \sum_{t=1}^{T^*-1} \log \left(f_i(\widehat{\Delta L}_t \mid \theta_i) \right)$$

each in terms of the corresponding density $f_i(\cdot \mid \theta_i)$, $i \in \{1, ..., 5\}$, with parameters collected in the vector θ_i of length k_i lying in the admissible set denoted by $\Theta_i \subset \mathbb{R}^{k_i}$. Note that f_5 has no closed-form expression, hence its numerical approximation is used in `R`.

The most suitable model is then selected by the *AICC* (*Akaike information criterion, corrected* for finite sample sizes). This performance measure is calculated by means of the log-likelihood by

$$AICC_i(\theta_i) := 2k_i + \frac{2k_i(k_i+1)}{T^* - k_i - 2} - 2 \log \ell_i(\theta_i), \quad i \in \{1, ..., 5\}.$$

Like the *BIC* used for comparison of the goodness of fit of the CARMA parameters, the optimal model is attained at the minimum as well.

[12] For further details see Nolan (2009) who also provides a program called `STABLE`, available at http://academic2.american.edu/~jpnolan/stable/stable.html.

[13] In light of the properties that were detected already at first glance, the normal distribution was only added for comparison purposes.

4.3.2. Estimated parameters and goodness of fit

Table 4.3.1 confirms our conjectures that the tails being observed are too heavy in order to be appropriately described by the normal or even Student's t-distributions ($i \in \{1, 2\}$). Only the NIG distribution ($i = 4$) is nearly able to capture the behaviour of the extreme events. In all of the 17 cases, where either the Laplace or the α-stable distribution ($i \in \{3, 5\}$) is optimal, the estimation of the NIG estimation failed due to unresolved numerical reasons. In any other of the 118 cases, however, the NIG is clearly proven to be most appropriate across-the-board. For the previously picked exemplary companies, this is shown in (log-)histograms and quantile-quantile plots put together in Figures 4.3.3–4.3.6 for distributions $i \in \{3, 4, 5\}$. They contain the corresponding log-likelihood/AICC values as well as the estimated parameters of each distribution. More detailed tabular comparisons and plots are available at the aforementioned online resource.

Distribution $i =$	1	2	3	4	5
Number of optimal AICCs	0	0	5	118	12

Table 4.3.1.: Number of optimal AICCs for each distribution candidate from Hypothesis 4.3. Recall that the total number of analyzed CAR(1) time series was 135.

Finally, one might summarize an apparently appropriate model for the CDS premium rates $C^* = \{C^*_t\}_{t \geq 0}$ as follows:

$$C^*_t = (1 - R)\,\gamma_t = (1 - R)\,\gamma_0\, e^{\int_0^t y_s ds}, \quad t \geq 0,$$

where $y = \{y_t\}_{t \in \mathbb{R}}$ is an OU process driven by NIG distributed Lévy noise, i.e.

$$dy_t = -\alpha_1 y_t dt + \sigma dL_t,\ \Delta L \sim \mathrm{NIG}(\alpha, \beta, \delta, \mu).$$

The following chapter concludes this thesis with an outlook to questions that one can pick up on for further research and practical applications, followed by a summary.

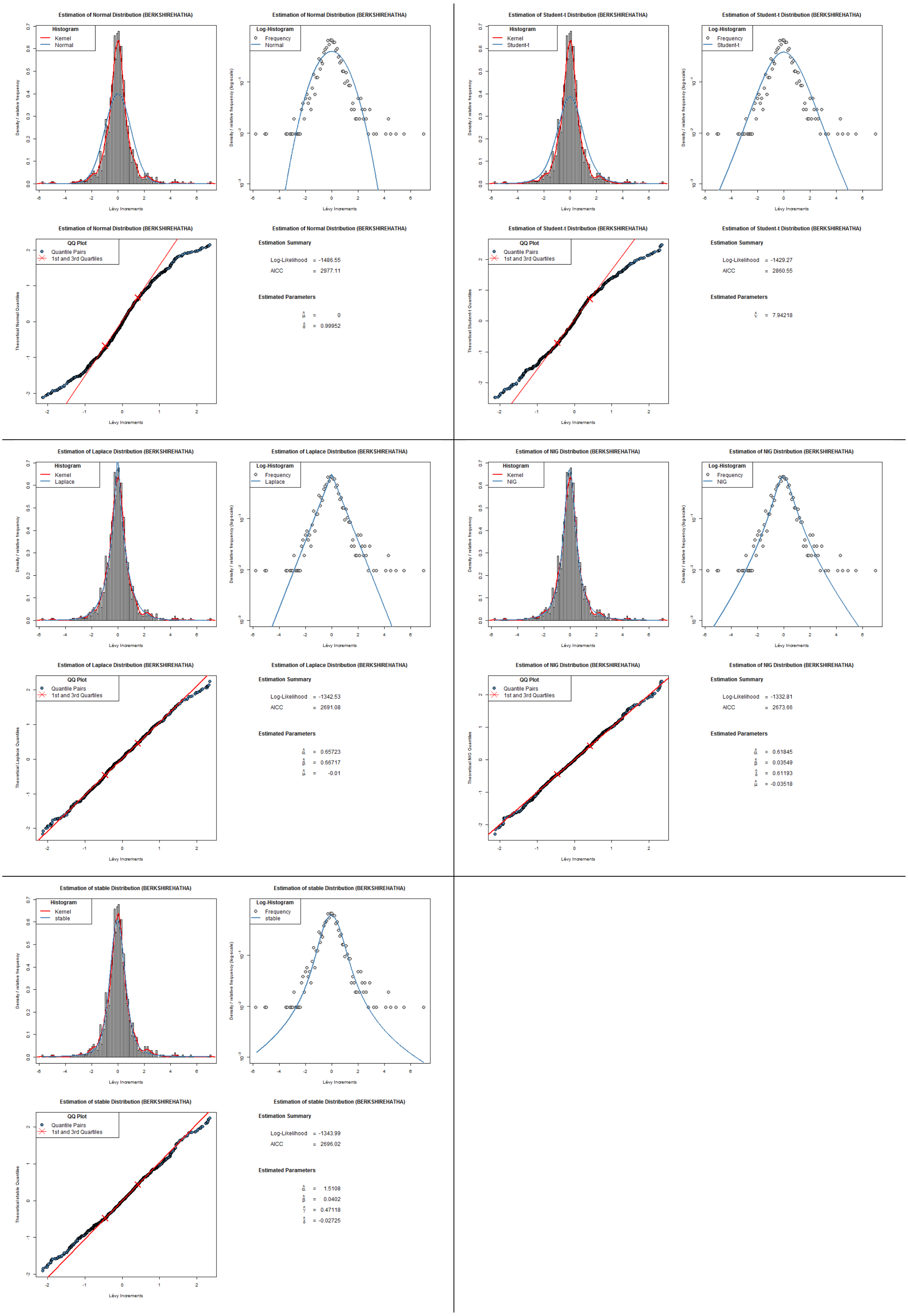

Figure 4.3.3.: (Log-)Histograms and quantile-quantile plots for fitted distributions of the recovered CAR(1) Lévy noise. The respective company is: Berkshire Hathaway.

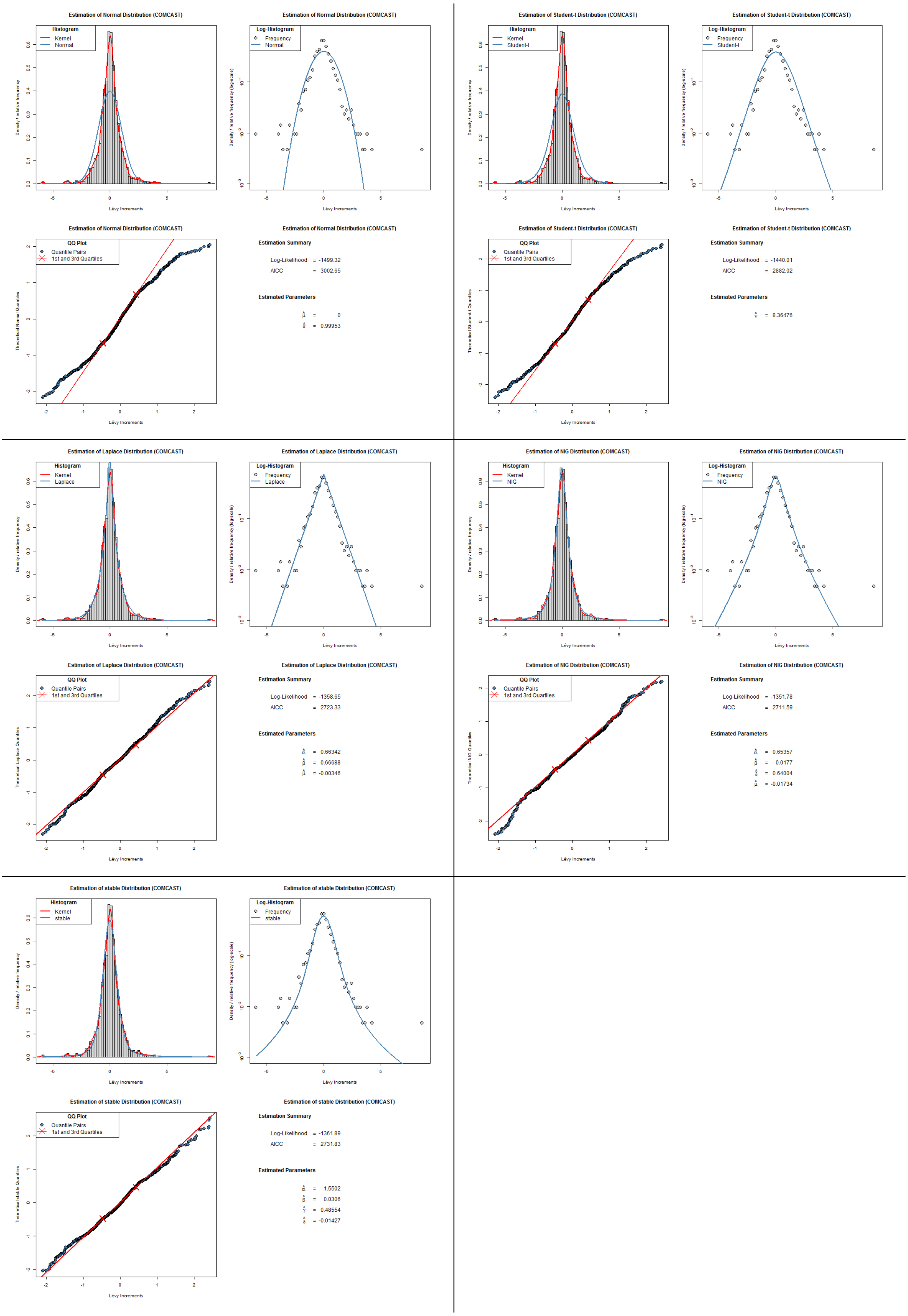

Figure 4.3.4.: (Log-)Histograms and quantile-quantile plots for fitted distributions of the recovered CAR(1) Lévy noise. The respective company is: Comcast.

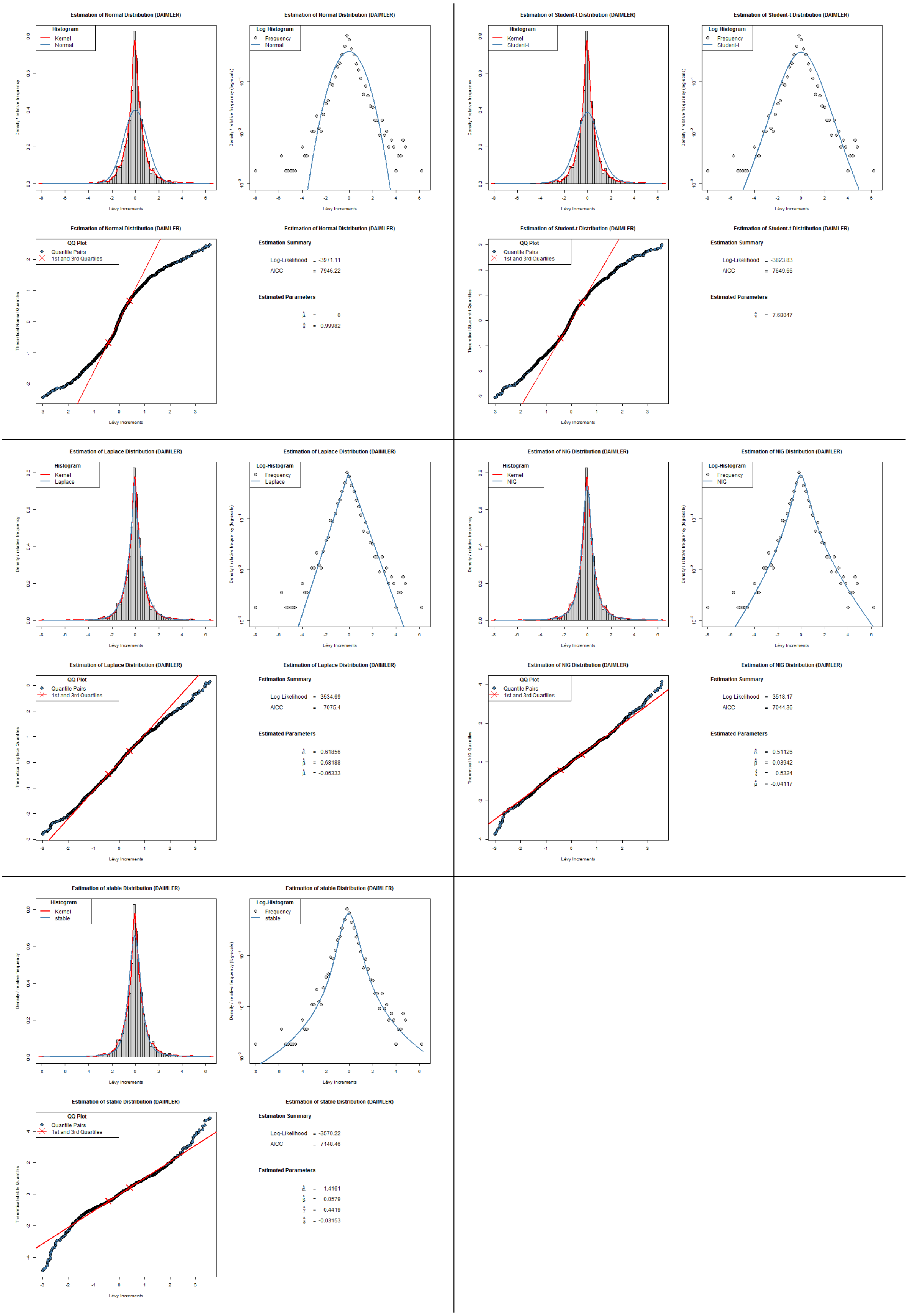

Figure 4.3.5.: (Log-)Histograms and quantile-quantile plots for fitted distributions of the recovered CAR(1) Lévy noise. The respective company is: Daimler.

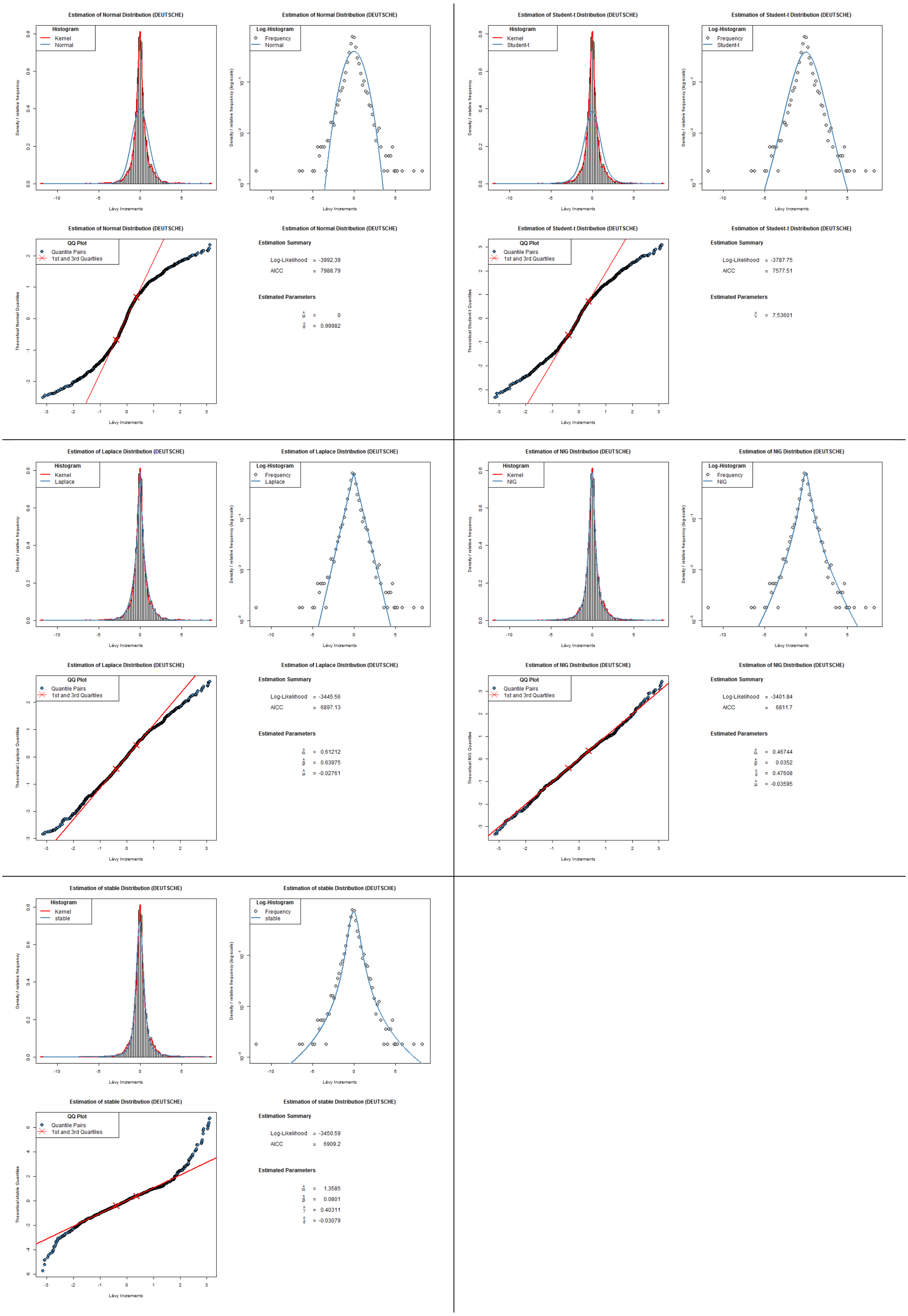

Figure 4.3.6.: (Log-)Histograms and quantile-quantile plots for fitted distributions of the recovered CAR(1) Lévy noise. The respective company is: Deutsche Bank.

Chapter 5.

Conclusions

5.1. Outlook

A brief outlook to possible extensions and further research topics, from different theoretical/modelling as well as practical/methodical aspects is following.

1. **Multivariate dependencies:**

    a) We have only considered each single name CDS on its own without taking dependence effects into account—which are obviously immanent, though: Just recall Figure 4.2.1 containing the very similar time series plots of iTraxx and CDX. To modell the dependence structure, one has different choices: For example, one can assume that all log-differences as a whole are observations from a vector-valued multivariate CARMA (MCARMA) process (see Brockwell and Schlemm (2013), for instance). Alternatively, one might fit each single process separately in a first step and model the joint distribution of the Lévy noises recovered from each individual CDS in a second step (for example using *copulas*[1]).

    b) Building different sub-portfolios—such as by industry sector, country of domicile or credit rating—instead of looking at all time series at once would lead to a different yet possibly more accurate view on the dependence structures.

[1] For a comprehensive introduction see Nelsen (2006)

2. **Volatility clustering:**

   Log-differencing has been proven to be a sufficient transformation in order to obtain a sample path of a presumably stationary stochastic process. Nevertheless, recalling Figure 4.2.3, one might consider groupings of several subsequent extremal log-difference values followed by groupings of values of relatively low magnitude as a sign of *volatility clustering*. To take this characteristic into additional account, a *stochastic volatility* modelling approach like (the continuous-time counterpart of) GARCH is one of the most often implemented applications for this.

   It is important to mention that the data basis predominantly described by CAR(1) processes in an apparently appropriate way is two-part, though: For the majority, the noise recovery procedure eliminates the correlation of the approximated Lévy increments both on the increments' level ($\widehat{\Delta L}$) as well as on the squared increments' level ($\widehat{\Delta L}^2$). This is commonly interpreted as an actually iid noise term. On the other hand, for some data fits, a linear process turns out to be non-sufficient. For example, the ACF plots of the $\widehat{\Delta L}$-levels exhibit zero correlation (see Figure 4.3.1), whereas the corresponding $\widehat{\Delta L}^2$-autocorrelations are still significantly high for several lags (see Figure 5.1.1). At this point, modelling the volatility stochastically is more appropriate, so our basic model probabibly still needs some little refinements.

3. **Estimation methods:**

   a) Within the scope of this thesis, only the noise recovery schemes of Brockwell, Davis and Yang (2007, 2011) have been illustrated for the CAR(1) case. One could either proceed correspondingly with the higher order cases as theoretically described in Remark 3.20 or try out the most recent methods introduced in Ferrazzano and Fuchs (2013) or Brockwell and Schlemm (2013), to get an extensive comparison between these estimation methods.

   b) Furthermore, alternative parameter estimation methods in addition to the maximum likelihood approach used here for CARMA and the noise distributions are conceivable.

4. **Risk management aspects:**

   a) Instead of the one-day log-differences, probably longer time horizons are of more interest (e.g., 30 or 90 trading days or even one year) for modelling stress scenarios.

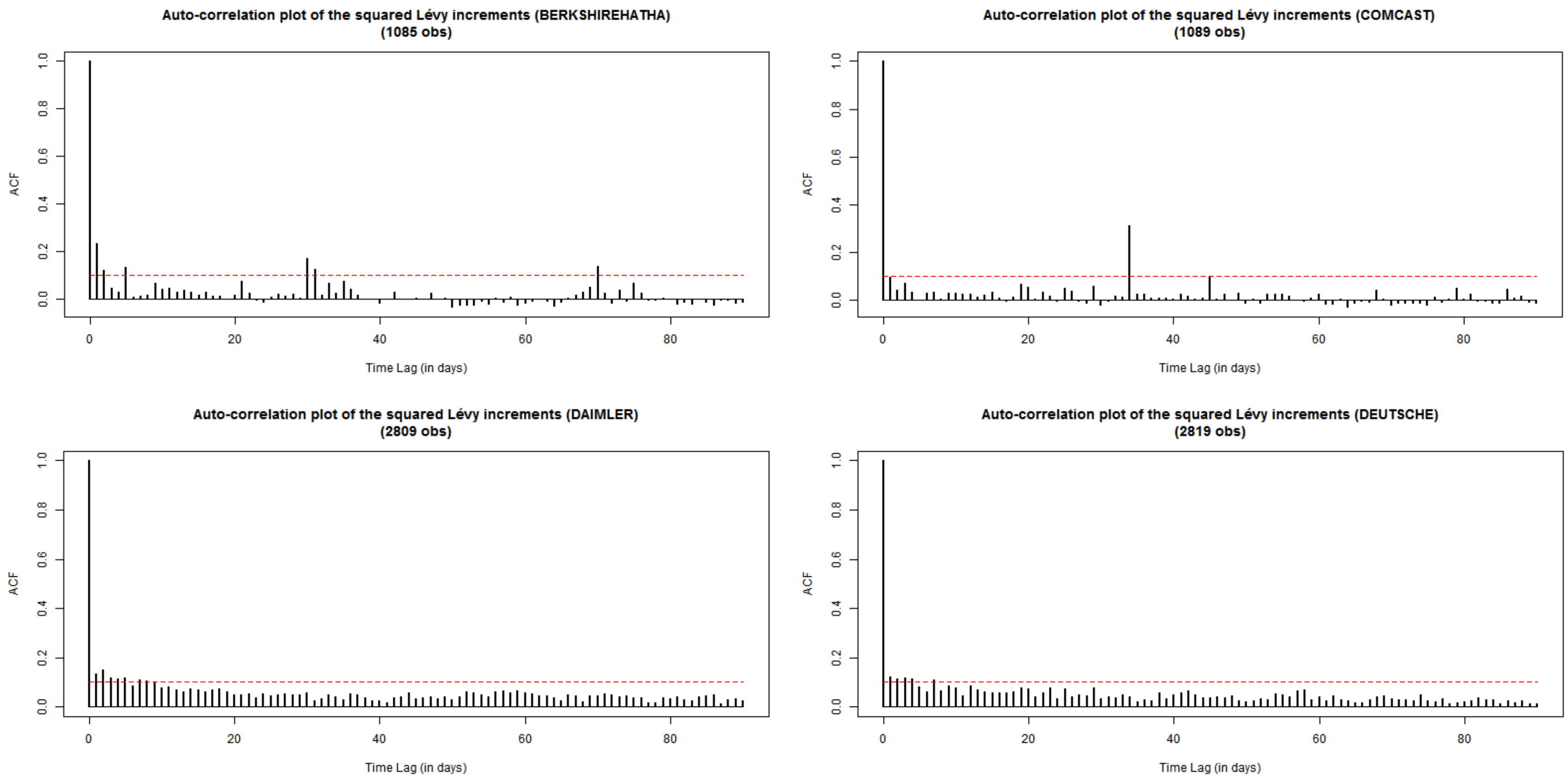

Figure 5.1.1.: Sample ACF plots of the squared Lévy increments recovered from the estimated CAR(1) models for the log-differenced CDS premium time series. The respective companies are: Berkshire Hathaway, Comcast, Daimler, Deutsche Bank.

To this end, one has to consider the h-day log-difference which is equal to the sum of h subsequent one-day log-differences, i.e.

$$\log C_t^* - \log C_{t-h}^* = \sum_{j=0}^{h-1} \left(\log C_{t-j}^* - \log C_{t-j-1}^*\right), \quad t \geq 0,$$

resulting in h-step iid noise increments and corresponding quantiles if one repeats the analyses on these variables.

b) Finally, from estimated log-differences of the CDS premia, one could draw conclusions on the behaviour of the *probabilities of default* (*PD*s) which we have already introduced theoretically in section 2.1 by equation (2.2), namely

$$F_t := \mathsf{E}[\mathbb{1}_{\{\tau \leq 0\}} \mid \mathcal{F}_t] = Q(\tau \leq t \mid \mathcal{F}_t), \quad t \geq 0,$$

which are in the intensity-based framework equal to

$$F_t = 1 - \exp\left\{-\int_0^t \gamma_s ds\right\}, \quad t \geq 0,$$

see Definition 2.10 (iii). Within our idealized assumptions resulting in a CDS premium C^* given by $C_t^* = (1-R)\gamma_t = S_t^*$, $t \geq 0$, its log-differences are trivially equal to those of the default rate γ_t. Then, we are not very far from being able to infer direct consequences on the process $\{F_t\}_{t\geq 0}$. Nevertheless, this would require successive integration/summation of the log-differences (modelled by CARMA process), then application of the exponential function in order to recover $\{\gamma_t\}_{t\geq 0}$ first and then integration and exponentiation again to end up finally at $\{F_t\}_{t\geq 0}$ which might be analytically anything but simple.

5.2. Summary

We have explored the theoretical properties and statistical behaviour of CDS premia over time. In the mostly theoretical chapters, the basic variable of interest, the continuously payed par premium $C^* = \{C_t^*\}_{t\geq 0}$ of a credit default swap contract, was mathematically derived to satisfy a simple yet idealized parity when neglecting other factors than credit risk determining the credit spread $S^* = \{S_t^*\}_{t\geq 0}$ of a bond.

To be able to apply statistical methods to (discrete observations of) this continuous-time premium process, we have introduced and made use of powerful techniques for modelling continuous-time linear time series by CARMA processes, extending the traditional discrete-time ARMA processes to infinitesimally small time steps. In this class of processes, the popular Ornstein-Uhlenbeck process is also contained as the simplest representative whose discrete counterpart is a classical autoregressive process of order one (AR(1) process). For the past two decades, these concepts have been enhanced and refined resulting in definitions and conditions for causality and invertibility (almost) analogously to the discrete-time theory. Nevertheless, we have learned that retrieving the background-driving (Lévy) process is less trivial. There are, however, several noise recovery approaches across the most recent literature, which base on the discrete samples of a CARMA process. One of the main goals of this thesis was to demonstrate a particular recovery scheme and to investigate the distributional properties of the resulting noise in order to derive several practical conclusions.

In the empirical part of this thesis, we focused on the discrete (daily) observations of CDS premia $C^* = \{C_t^*\}_{t\geq 0}$ on numerous European and North American reference issuers across different industry sectors and (investment grade) rating classes. However, on the premium level, they do not exhibit stationarity patterns required for a correct estimation. Since, in contrast, their log-differences do so—i.e. the non-stationarity hypothesis could be rejected—they have been chosen as the main subject of further considerations. In a second step, after the log-difference transformation, the data have been fitted via maximum likelihood to a set of possible CARMA(p,q) candidates. According to our primary selection criterion, CAR(1) seems predominantly to do best. The aim of the third step was to illustrate the noise recovery by means of a subset of time series that CAR(1) was optimally fitting to. The resulting iid noise samples have then further been explored: Among several contemplable parametric distribution families, the normal inverse Gaussian (NIG) has turned out to fit the entire empirical probability mass most suitably.

Appendix A.

Constant coupons and CDS premia

In the following sections, several formulae of the prices of constant-coupon bonds (as well as par yield spread) and premia of constant-premium CDS are put together, each of them in continuously paying and discreteley paying form. The latter is additionally broken down into immediately payed recovery/compensation (as assumed in this thesis) and payments postponed to the subsequent coupon date.

A.1. Constant coupon bonds

Each of the following lines starts with the model-free ansatz which is then further calculated within the intensity-based framework introduced in Chapter 2.

A.1.1. Continuously paying

Model-free framework:

$$\begin{aligned}\Pi^C(s, s+T) &= C\,\mathsf{E}\left[\int_s^{s+T} D(s,u)du \mid \mathcal{F}_s\right] \;+\; \mathsf{E}\left[D(s,s+T) \mid \mathcal{F}_s\right] \\ &= C\int_s^{s+T} \mathsf{E}\left[e^{-\int_s^u r_t dt} \mid \mathcal{F}_s\right] du \;+\; \mathsf{E}\left[e^{-\int_s^{s+T} r_t dt} \mid \mathcal{F}_s\right],\end{aligned}$$

Intensity-based framework:

$$
\begin{aligned}
\tilde{\Pi}^{\tilde{C},R}(s,s+T) &= \tilde{C}\,\mathsf{E}\left[\int_s^{s+T} D(s,u)\,\mathbb{1}_{\{\tau>u\}}du \mid \mathcal{G}_s\right] + \\
&\quad + \mathsf{E}\left[D(s,s+T)\,\mathbb{1}_{\{\tau>s+T\}} \mid \mathcal{G}_s\right] + \\
&\quad + R\,\mathsf{E}\left[D(s,\tau)\,\mathbb{1}_{\{\tau\leq s+T\}} \mid \mathcal{G}_s\right] \\
&= \tilde{C}\,\mathbb{1}_{\{\tau>s\}} \int_s^{s+T} \mathsf{E}\left[e^{-\int_s^u r_t+\gamma_t dt} \mid \mathcal{F}_s\right] du + \\
&\quad + \mathbb{1}_{\{\tau>s\}}\,\mathsf{E}\left[e^{-\int_s^{s+T} r_t+\gamma_t dt} \mid \mathcal{F}_s\right] + \\
&\quad + R\,\mathbb{1}_{\{\tau>s\}} \int_s^{s+T} E\left[e^{-\int_s^u r_t+\gamma_t dt}\gamma_u \mid \mathcal{F}_s\right] du.
\end{aligned}
$$

Using that $\Pi^C = \tilde{\Pi}^{C,1}$ for any defaultable issuer with random default time τ and intensity $\{\gamma_t\}_{t\geq 0}$ we obtain as coupon spread in general and par yield spread in particular:

$$
\Rightarrow\ \tilde{C} - C = \mathbb{1}_{\{\tau>s\}} \frac{\tilde{\Pi}^{\tilde{C},R} - \Pi^C + (1-R)\ \int_s^{s+T} E\left[e^{-\int_s^u r_t+\gamma_t dt}\gamma_u \mid \mathcal{F}_s\right] du}{\int_s^{s+T} \mathsf{E}\left[e^{-\int_s^u r_t+\gamma_t dt} \mid \mathcal{F}_s\right] du},
$$

$$
\Rightarrow\ S^* := \tilde{Y} - Y = \mathbb{1}_{\{\tau>s\}} \frac{(1-R)\ \int_s^{s+T} E\left[e^{-\int_s^u r_t+\gamma_t dt}\gamma_u \mid \mathcal{F}_s\right] du}{\int_s^{s+T} \mathsf{E}\left[e^{-\int_s^u r_t+\gamma_t dt} \mid \mathcal{F}_s\right] du}.
$$

A.1.2. Discretely paying $\{s = u_0 < \ldots < u_k = s+T\}$

Model-free framework:

$$
\begin{aligned}
\Pi^C(s,s+T) &= C\,\sum_{k=1}^{N} \Delta u_k\,\mathsf{E}\left[D(s,u_k) \mid \mathcal{F}_s\right] + \mathsf{E}\left[D(s,s+T) \mid \mathcal{F}_s\right] \\
&= C\,\sum_{k=1}^{N} \Delta u_k\,\mathsf{E}\left[e^{-\int_s^{u_k} r_t dt} \mid \mathcal{F}_s\right] + \mathsf{E}\left[e^{-\int_s^{s+T} r_t dt} \mid \mathcal{F}_s\right],
\end{aligned}
$$

Intensity-based framework:

$$\begin{aligned}
\tilde{\Pi}^{\tilde{C},R}(s,s+T) &= \tilde{C}\sum_{k=1}^{N}\Delta u_k\,\mathsf{E}\left[D(s,u_k)\,\mathbb{1}_{\{\tau>u_k\}}\mid\mathcal{G}_s\right]\;+\\
&+\;\mathsf{E}\left[D(s,s+T)\,\mathbb{1}_{\{\tau>s+T\}}\mid\mathcal{G}_s\right]\;+\\
&+\;R\,\mathsf{E}\left[D(s,\tau)\,\mathbb{1}_{\{\tau\le s+T\}}\mid\mathcal{G}_s\right]\\
&= \tilde{C}\,\mathbb{1}_{\{\tau>s\}}\sum_{k=1}^{N}\Delta u_k\,\mathsf{E}\left[e^{-\int_s^{u_k} r_t+\gamma_t dt}\mid\mathcal{F}_s\right]\;+\\
&+\;\mathbb{1}_{\{\tau>s\}}\,\mathsf{E}\left[e^{-\int_s^{s+T} r_t+\gamma_t dt}\mid\mathcal{F}_s\right]\;+\\
&+\;R\,\mathbb{1}_{\{\tau>s\}}\int_s^{s+T}\mathsf{E}\left[e^{-\int_s^{u} r_t+\gamma_t dt}\gamma_u\mid\mathcal{F}_s\right]du.
\end{aligned}$$

Using that $\Pi^C = \tilde{\Pi}^{C,1}$ for any defaultable issuer with random default time τ and intensity $\{\gamma_t\}_{t\ge0}$ we obtain as coupon spread in general and par yield spread in particular in the intensity-based framework:

$$\Rightarrow\;\tilde{C}-C=\mathbb{1}_{\{\tau>s\}}\frac{\tilde{\Pi}^{\tilde{C},R}-\Pi^C+(1-R)\int_s^{s+T}E\left[e^{-\int_s^{u} r_t+\gamma_t dt}\gamma_u\mid\mathcal{F}_s\right]du}{\sum_{k=1}^{N}\Delta u_k\,\mathsf{E}\left[e^{-\int_s^{u_k} r_t+\gamma_t dt}\mid\mathcal{F}_s\right]},$$

$$\Rightarrow\;S^*:=\tilde{Y}-Y=\mathbb{1}_{\{\tau>s\}}\frac{(1-R)\int_s^{s+T}E\left[e^{-\int_s^{u} r_t+\gamma_t dt}\gamma_u\mid\mathcal{F}_s\right]du}{\sum_{k=1}^{N}\Delta u_k\,\mathsf{E}\left[e^{-\int_s^{u_k} r_t+\gamma_t dt}\mid\mathcal{F}_s\right]}.$$

A.1.3. Recovery payment postponed to next coupon date

Equivalent to assuming that τ is not a continuous random variable but lives on the grid $\{u_1, ..., u_k\}$.

Intensity-based framework:

$$
\begin{aligned}
\tilde{\Pi}^{\tilde{C},R}(s,s+T) = \tilde{C}\,\mathbb{1}_{\{\tau>s\}} \sum_{k=1}^{N} \Delta u_k\, \mathsf{E}\left[e^{-\int_s^{u_k} r_t+\gamma_t dt} \mid \mathcal{F}_s\right] + \\
+\ \mathbb{1}_{\{\tau>s\}}\, \mathsf{E}\left[e^{-\int_s^{s+T} r_t+\gamma_t dt} \mid \mathcal{F}_s\right] + \\
+\ R\,\mathbb{1}_{\{\tau>s\}} \sum_{k=1}^{N} \mathsf{E}\left[e^{-\int_s^{u_k} r_t dt}\left(e^{-\int_s^{u_{k-1}} \gamma_t dt} - e^{-\int_s^{u_k} \gamma_t dt}\right) \mid \mathcal{F}_s\right].
\end{aligned}
$$

Using that $\Pi^C = \tilde{\Pi}^{C,1}$ for any defaultable issuer with random default time τ and intensity $\{\gamma_t\}_{t\geq 0}$ we obtain as coupon spread in general and par yield spread in particular in the intensity-based framework:

$$
\Rightarrow\ \tilde{C} - C = \mathbb{1}_{\{\tau>s\}} \frac{\tilde{\Pi}^{\tilde{C},R} - \Pi^C + (1-R)\sum_{k=1}^{N} \mathsf{E}\left[e^{-\int_s^{u_k} r_t dt}\left(e^{-\int_s^{u_{k-1}} \gamma_t dt} - e^{-\int_s^{u_k} \gamma_t dt}\right) \mid \mathcal{F}_s\right]}{\sum_{k=1}^{N} \Delta u_k\, \mathsf{E}\left[e^{-\int_s^{u_k} r_t+\gamma_t dt} \mid \mathcal{F}_s\right]},
$$

$$
\Rightarrow\ S^* := \tilde{Y} - Y = \mathbb{1}_{\{\tau>s\}} \frac{(1-R)\sum_{k=1}^{N} \mathsf{E}\left[e^{-\int_s^{u_k} r_t dt}\left(e^{-\int_s^{u_{k-1}} \gamma_t dt} - e^{-\int_s^{u_k} \gamma_t dt}\right) \mid \mathcal{F}_s\right]}{\sum_{k=1}^{N} \Delta u_k\, \mathsf{E}\left[e^{-\int_s^{u_k} r_t+\gamma_t dt} \mid \mathcal{F}_s\right]}.
$$

A.2. Constant premium CDS

Each of the following lines starts with the model-free ansatz for the par CDS condition, see equation (2.23), which is then further calculated within the intensity-based framework introduced in Chapter 2.

A.2.1. Continuously paying

Par CDS condition:

$$
C^*\,\mathsf{E}\left[\int_s^{s+T} D(s,u)\,\mathbb{1}_{\{\tau>u\}} du \mid \mathcal{G}_s\right] = (1-R)\,\mathsf{E}\left[D(s,\tau)\,\mathbb{1}_{\{\tau\leq s+T\}} \mid \mathcal{G}_s\right]
$$

Intensity-based framework:

$$\begin{aligned}\Leftrightarrow \quad C^* = C^*(s,s+T) &= (1-R)\,\frac{\mathsf{E}\left[D(s,\tau)\,\mathbb{1}_{\{\tau\le s+T\}} \mid \mathcal{G}_s\right]}{\mathsf{E}\left[\int_s^{s+T} D(s,u)\,\mathbb{1}_{\{\tau>u\}}du \mid \mathcal{G}_s\right]} \\ &= (1-R)\,\mathbb{1}_{\{\tau>s\}}\,\frac{\int_s^{s+T}\mathsf{E}\left[e^{-\int_s^u r_t+\gamma_t dt}\gamma_u \mid \mathcal{F}_s\right]du}{\int_s^{s+T}\mathsf{E}\left[e^{-\int_s^u r_t+\gamma_t dt} \mid \mathcal{F}_s\right]du} \\ &= S^*\end{aligned} \tag{A.1}$$

A.2.2. Discretely paying $\{s = u_0 < ... < u_k = s+T\}$

Par CDS condition:

$$C^* \sum_{k=1}^{N} \Delta u_k\,\mathsf{E}\left[D(s,u_k)\,\mathbb{1}_{\{\tau>u_k\}} \mid \mathcal{G}_s\right] = (1-R)\,\mathsf{E}\left[D(s,\tau)\,\mathbb{1}_{\{\tau\le s+T\}} \mid \mathcal{G}_s\right]$$

Intensity-based framework:

$$\begin{aligned}\Leftrightarrow \quad C^* = C^*(s,s+T) &= (1-R)\,\frac{\mathsf{E}\left[D(s,\tau)\,\mathbb{1}_{\{\tau\le s+T\}} \mid \mathcal{G}_s\right]}{\sum_{k=1}^{N}\Delta u_k\,\mathsf{E}\left[D(s,u_k)\,\mathbb{1}_{\{\tau>u_k\}} \mid \mathcal{G}_s\right]} \\ &= (1-R)\,\mathbb{1}_{\{\tau>s\}}\,\frac{\int_s^{s+T}\mathsf{E}\left[e^{-\int_s^u r_t+\gamma_t dt}\gamma_u \mid \mathcal{F}_s\right]du}{\sum_{k=1}^{N}\Delta u_k\,\mathsf{E}\left[e^{-\int_s^{u_k} r_t+\gamma_t dt} \mid \mathcal{F}_s\right]} \\ &= S^*\end{aligned} \tag{A.2}$$

A.2.3. Compensation payment postponed to next premium date

Par CDS condition:

$$C^*\,\Delta u_k \sum_{k=1}^{N} \mathsf{E}\left[D(s,u_k)\,\mathbb{1}_{\{\tau>u_k\}} \mid \mathcal{G}_s\right] = (1-R)\,\mathsf{E}\left[D(s,\tau)\,\mathbb{1}_{\{\tau\le s+T\}} \mid \mathcal{G}_s\right]$$

Intensity-based framework:

$$\Leftrightarrow \quad C^* = C^*(s, s+T) = (1-R)\,\frac{\mathsf{E}\left[D(s,\tau)\,\mathbb{1}_{\{\tau\le s+T\}} \mid \mathcal{G}_s\right]}{\sum_{k=1}^N \Delta u_k\,\mathsf{E}\left[D(s,u_k)\,\mathbb{1}_{\{\tau>u_k\}} \mid \mathcal{G}_s\right]}$$

$$= (1-R)\,\mathbb{1}_{\{\tau>s\}}\,\frac{\sum_{k=1}^N \mathsf{E}\left[e^{-\int_s^{u_k} r_t dt}\left(e^{-\int_s^{u_{k-1}}\gamma_t dt} - e^{-\int_s^{u_k}\gamma_t dt}\right) \mid \mathcal{F}_s\right]}{\sum_{k=1}^N \Delta u_k\,\mathsf{E}\left[e^{-\int_s^{u_k} r_t + \gamma_t dt} \mid \mathcal{F}_s\right]} \qquad \text{(A.3)}$$

$$= S^*$$

Bibliography

BAI, J., COLLIN-DUFRESNE, P. (2011). The Determinants of the CDS-Bond Basis during the Financial Crisis of 2007–2009 (Netspar Discussion Paper). http://ssrn.com/abstract=2180770.

BARNDORFF-NIELSEN, O. (1977). Exponentially Decreasing Distributions for the Logarithm of Particle Size. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **353**(1674):401–419.

BARNDORFF-NIELSEN, O. (1997). Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling. *Scandinavian Journal of statistics* **24**(1):1–13.

BARNDORFF-NIELSEN, O. E., SHEPHARD, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-based Models and some of their Uses in Financial Economics. *Journal of the Royal Statistical Society* **63**:167–241.

BERNHARDT, C., KLÜPPELBERG, C., MEYER-BRANDIS, T. (2008). Estimating High Quantiles for Electricity Prices by Stable Linear Models. *Journal of Energy Markets* **1**(1):3–19.

BIELECKI, T. R., RUTKOWSKI, M. (2002). *Credit Risk: Modeling, Valuation and Hedging.* Springer, Springer Finance.

BLACK, F., COX, J. C. (1976). Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. *Journal of Finance* **31**(2):351–367.

BLANCO, R., BRENNAN, S., MARSH, I. W. (2005). An Empirical Analysis of the Dynamic Relation between Investment-Grade Bonds and Credit Default Swaps. *Journal of Finance* 60(5):2255–2281.

BRENNAN, M. J., SCHWARTZ, E. S. (1980). Analyzing Convertible Bonds. *Journal of Financial and Quantitative Analysis* **15**(4):907–929.

Brockwell, A. E., Brockwell, P. J. (1999). A Class of Non-Embeddable ARMA Processes. *Journal of Time Series Analysis* **20**(5):483–486.

Brockwell, P. J. (1995). A Note on the Embedding of Discrete-Time ARMA Processes. *Journal of Time Series Analysis* **16**(5):451–460.

Brockwell, P. J. (2001). Lévy-driven CARMA processes. *Annals of Applied Probability* **53**(1):113–124.

Brockwell, P. J. (2004). Representations of Continuous-Time ARMA Processes. *Journal of Applied Probability* **41A**:375–382.

Brockwell, P. J., Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer, Springer Series in Statistics 2nd edition edition.

Brockwell, P. J., Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer, Springer Texts in Statistics 2nd edition edition.

Brockwell, P. J., Davis, R. A., Yang, Y. (2007). Estimation for Non-Negative Lévy-driven Ornstein-Uhlenbeck Processes. *Journal of Applied Probability* **44**:987–989.

Brockwell, P. J., Davis, R. A., Yang, Y. (2011). Estimation for Non-Negative Lévy-driven CARMA Processes. *Journal of Business & Economic Statistics* **29**(2):250–259.

Brockwell, P. J., Ferrazzano, V., Klüppelberg, C. (2012a). High-Frequency Sampling and Kernel Estimation for Continuous-Time Moving Average Processes (Preprint, to appear in the *Journal of Time Series Analysis*). http://mediatum.ub.tum.de/node?id=1108438.

Brockwell, P. J., Ferrazzano, V., Klüppelberg, C. (2012b). High-Frequency Sampling of a Continuous-Time ARMA Process. *Journal of Time Series Analysis* **33**(1):152–160.

Brockwell, P. J., Lindner, A. (2009). Existence and Uniqueness of Stationary Lévy-driven CARMA Processes. *Stochastic Processes and their Applications* **119**:2660–2681.

Brockwell, P. J., Schlemm, E. (2013). Parametric Estimation of the Driving Lévy Process of Multivariate CARMA Processes from Discrete Observations. *Journal of Multivariate Analysis* **115**:217–251.

CECCHETTI, S., DI CESARE, A. (2012). The Mathematics of the Relationship between the Default Risk and Yield-to-Maturity of Coupon Bonds (Preprint). http://arxiv.org/pdf/1203.6723.pdf.

CHAN-LAU, J. A. (2006). *Market-based Estimation of Default Probabilities and its Application to Financial Market Surveillance*. International Monetary Fund, IMF Working Papers.

COCHRANE, J. H. (2012). Continuous-Time Linear Models (Working Paper). http://faculty.chicagobooth.edu/john.cochrane/research/papers/continuous_time_linear_models.pdf.

DUFFEE, G. R. (1998). The Relation between Treasury Yields and Corporate Bond Yield Spreads. *Journal of Finance* **53**(6):2225–2241.

DUFFIE, D. (1999). Credit Swap Valuation. *Financial Analysts Journal* **55**:73–87.

DUFFIE, D., SINGLETON, K. J. (1999). Modeling Term Structures of Defaultable Bonds. *Review of Financial Studies* **12**(4):687–720.

DUFFIE, D., SINGLETON, K. J. (2009). *Credit Risk: Pricing, Measurement, and Management*. Princeton University Press, Princeton Series in Finance.

EBERLEIN, E., RAIBLE, S. (1999). Term Structure Models Driven by General Lévy Processes. *Mathematical Finance* **9**(1):31–53.

ERICSSON, J., JACOBS, K., OVIEDO, R. (2009). The Determinants of Credit Default Swap Premia. *Journal of Financial and Quantitative Analysis* **44**(1):109–132.

FERRAZZANO, V., FUCHS, F. (2013). Noise Recovery for Lévy-driven CARMA Processes and High-Frequency Behaviour of Approximating Riemann Sums (Preprint). http://arxiv.org/abs/1210.1389.

FÖLLMER, H., SCHIED, A. (2011). *Stochastic Finance: An Introduction in Discrete Time*. de Gruyter graduate. de Gruyter.

FONTANA, A. (2011). The Negative CDS-Bond Basis and Convergence Trading during the 2007/09 Financial Crisis (Working Paper). http://ssrn.com/abstract=1688330.

GARCÍA, I., KLÜPPELBERG, C., MÜLLER, G. (2011). Estimation of Stable CARMA Models with an Application to Electricity Spot Prices. *Statistical Modelling* **11**(5):447–470.

HULL, J., PREDESCU, M., WHITE, A. (2004). The Relationship between Credit Default Swap Spreads, Bond Yields, and Credit Rating Announcements. *Journal of Banking & Finance* **28**(11):2789–2811.

HULL, J., WHITE, A. (2000). Valuing Credit Default Swaps I: No Counterparty Default Risk. *Journal of Derivatives* **8**(1):29–40.

HULL, J., WHITE, A. (2001). Valuing Credit Default Swaps II: Modeling Default Correlations. *Journal of Derivatives* **8**(3):12–22.

JARROW, R., TURNBULL, S. (1995). Pricing Options on Financial Securities Subject to Default Risk. *Journal of Finance* **50**:53–86.

JEANBLANC, M., RUTKOWSKI, M. (2000a). Modelling of Default Risk: Mathematical Tools (Workshop *Fixed Income and Credit Risk Modeling and Management*). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.139.678&rep=rep1&type=pdf.

JEANBLANC, M., RUTKOWSKI, M. (2000b). Modelling of Default Risk: An Overview. *Mathematical Finance: Theory and Practice* pages 171–269.

JEANBLANC, M., RUTKOWSKI, M. (2000c). Modelling of Default Risk: An Overview. *Mathematical Finance: Theory and Practice* pages 281–312.

KUSHMA, M. B., CLASS, R., KURZWEIL, P. (2012). The Evolution of the Global Bond Market. *Investment Management Journal, Morgan Stanley* **2**(2):3–18.

LEVY, A. (2009). The CDS Bond Spread in Emerging Markets: Liquidity and Counterparty Risk Effects (Draft). http://econ.haifa.ac.il/~todd/seminars/papers08-09/CDSPaper(draft).pdf.

LONGSTAFF, F., MITHAL, S., NEIS, E. (2005). Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit-Default Swap Market. *Journal of Finance* **60**:2213–2253.

LUO, L. S. J. (2005). Bootstrapping Default Probability Curves. *Journal of Credit Risk* **1**(4):169–179.

MAHANTI, S., NASHIKKAR, A., SUBRAHMANYAM, M. (2007). Latent Liquidity and Corporate Bond Yield Spreads (Working Paper). http://pages.stern.nyu.edu/~msubrahm/papers/CDSPaperold.pdf.

Martin, M. R. W., Reitz, S., Wehn, C. (2006). *Kreditderivate und Kreditrisikomodelle: Eine mathematische Einführung*. Friedr. Vieweg & Sohn Verlag GWV Fachverlage GmbH (GWV).

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance* **29**(2):449–470.

Nashikkar, A., Subrahmanyam, M. G., Mahanti, S. (2011). Liquidity and Arbitrage in the Market for Credit Risk. *Journal of Financial and Quantitative Analysis* 46(03):627–656.

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer Series in Statistics. Springer.

Nolan, J. P. (2009). Stable Distributions - Models for Heavy Tailed Data (In progress). http://academic2.american.edu/~jpnolan/stable/chap1.pdf.

Pan, J., Singleton, K. J. (2008). Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads. *Journal of Finance* **63**(5):2345–2384.

Pham, D. T. (1977). Estimation of Parameters of a Continuous Time Gaussian Stationary Process with Rational Spectral Density. *Biometrika* **64**(2):385–399.

Protter, P. (2004). *Stochastic Integration and Differential Equations*. Applications of Mathematics Series. Springer.

R Development Core Team (2012). R: A Language and Environment for Statistical Computing (Manual). http://www.R-project.org/.

Sayed, A., Kailath, T. (2001). A Survey of Spectral Factorization Methods. *Numerical Linear Algebra with Applications* **8**(6-7):467–496.

Schläfer, T. S., Uhrig-Homburg, M. (2009). Estimating Market-Implied Recovery Rates from Credit Default Swap Premia (Presented at the *13th Conference of the Swiss Society for Financial Market Research*). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.183.8641&rep=rep1&type=pdf.

Schlemm, E. (2011). Estimation of Continuous-Time ARMA Models and Random Matrices with Dependent Entries. (PhD thesis at Technische Universität München). http://mediatum.ub.tum.de/node?id=1075059.

Schönbucher, P. J. (2003). *Credit Derivative Pricing Models: Models, Pricing and Implementation.* Wiley, Wiley Finance Series.

Scott, D. (2009). `HyperbolicDist`: The Hyperbolic Distribution (Manual). http://CRAN.R-project.org/package=HyperbolicDist.

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance* **19**(3):425–442.

Thornton, M. A., Chambers, M. J. (2011). On the Equivalence of Discrete Time Representations of Continuous Time ARMA Processes (Unpublished). http://www-users.york.ac.uk/~mt795/Papers/carma_supp.pdf.

Tómasson, H. (2011). Some Computational Aspects of Gaussian CARMA Modelling (Preprint). http://www.ihs.ac.at/publications/eco/es-274.pdf.

Tómasson, H. (2012). `ctarma`: Estimation and Simulation of CARMA(p,q) (Manual). http://CRAN.R-project.org/package=ctarma.

Todorov, V., Tauchen, G. (2006). Simulation Methods for Lévy-driven Continuous-Time Autoregressive Moving Average (CARMA) Stochastic Volatility Models. *Journal of Business & Economic Statistics* **24**(4):455–469.

Uhlenbeck, G. E., Ornstein, L. S. (1930). On the Theory of the Brownian Motion. *Physical Review* **36**(5):823–841.

Uhrig-Homburg, M. (2002). Valuation of Defaultable Claims – A Survey. *Schmalenbach Business Review* **54**:24–57.

Vause, N. (2010). Counterparty Risk and Contract Volumes in the Credit Default Swap Market. *BIS Quarterly Review* December 2010:59–69.

Wuertz, D., Rmetrics core team (2012). `fBasics`: Rmetrics – Markets and Basic Statistics (Manual). http://CRAN.R-project.org/package=fBasics.

Zhu, H. (2006). An Empirical Comparison of Credit Spreads between the Bond Market and the Credit Default Swap Market. *Journal of Financial Services Research* **29**(3):211–235.