Modelling of Multivariate Asset Processes using Squared Bessel Processes

Inaugural-Dissertation

zur Erlangung des akademischen Grades
"Doktor der Naturwissenschaften"

eingereicht beim

Fachbereich Mathematik und Informatik, Physik,

Geographie
der Justus-Liebig-Universität Gießen

von
Rolf Klaas

Giessen, 2009

D-26

Dekan:	Prof. Dr. Bernd Baumann (Justus-Liebig-Universität Gießen)
Gutachter:	Prof. Dr. Ludger Overbeck (Justus-Liebig-Universität Gießen)
	Prof. Dr. Winfried Stute (Justus-Liebig-Universität Gießen)

Zusammenfassung

Das Risiko, dass ein Kreditnehmer oder allgemeiner ein Geschäftspartner seinen Forderungen, bedingt durch beispielsweise Insolvenz, nicht nachkommen kann, wird als Kreditrisiko oder Kreditausfallrisiko bezeichnet. Betrachtet man einen einzelnen Kreditnehmer, sind dabei für eine Bank drei Größen von besonderer Bedeutung:

- die Höhe der ausstehenden Forderungen zum Zeitpunkt des Ausfalls
- die Verwertungsrate, d.h. der Anteil, der aufgrund von Sicherheiten noch zurückgezahlt werden kann
- die Ausfallwahrscheinlichkeit

Die Ausfallwahrscheinlichkeit bezieht sich dabei immer auf einen festgelegten Zeithorizont. Die Ein-Jahres-Ausfallwahrscheinlichkeit beschreibt beispielsweise die Wahrscheinlichkeit eines Ausfalls innerhalb des nächsten Jahres. Betrachtet man ein Portfolio mit mehreren Kreditnehmern, müssen auch noch Abhängigkeiten zwischen den Kreditnehmern berücksichtigt werden. Für die Modellierung von Kreditrisiken haben sich im Wesentlichen zwei Modellansätze durchgesetzt:

- Intensitätsbasierte Modelle
- Firmenwertbasierte Modelle

Der Fokus der vorliegenden Arbeit liegt dabei auf firmenwertbasierten Modellen. In der Einleitung wird die wegweisende, mit dem Nobelpreis ausgezeichnete Pionierarbeit von Robert C. Merton, (Merton, 1974), sowie einige darauf basierende Erweiterungen vorgestellt. Im Merton-Modell wird der Firmenwert als geometrische Brownsche Bewegung modelliert. Ein Ausfall findet statt, falls der Firmenwert zu einem bestimmten Zeitpunkt unterhalb einer festgelegten Schranke liegt. Dieses Modell wurde in (Black and Cox, 1976) dahingehend erweitert, dass ein Ausfall zu jedem Zeitpunkt stattfinden konnte. Als Ausfallzeitpunkt wurde der Zeitpunkt definiert, zu dem der Firmenwert zum ersten Mal eine bestimmte Schranke erreicht.

Basierend auf diesem Ansatz wurden zahlreiche so genannte strukturelle Modelle entwickelt, deren Hauptmerkmal ist, dass der Ausfallzeitpunkt als Erstaustrittszeit eines stochastischen Prozesses, nicht notwendigerweise des tatsächlichen Firmenwertprozesses, modelliert wird. Als Beispiel wird im zweiten Kapitel das in (Albanese et al., 2003) vorgestellte Modell näher betrachtet. Als Ausfallschranke dient dort die Null. Der Prozess, der die Kreditqualität repräsentiert, wird als driftfreier stochastischer Prozess mit lokaler Volatilität modelliert. Dieser Prozess resultiert aus einer stochastischen Transformation, angewendet auf einen Squared Bessel Prozess. Das Modell kann jedoch lediglich Kreditrisiken einzelner Kreditnehmer modellieren. Um jedoch auch Abhängigkeiten zwischen verschiedenen Kreditnehmern berücksichtigen zu können, muss ein Modell auch gemeinsame Ausfallwahrscheinlichkeiten zweier Kreditnehmer produzieren können.

Basierend auf dem Merton-Ansatz sind so genannte Faktormodelle wie beispielsweise Creditmetrics, (JPMorgan, 1997), entwickelt worden. Da dort der Ausfall nur an einem bestimmten Zeitpunkt stattfinden kann, genügt es, die Kreditwürdigkeit durch eine Zufallsvariable zu modellieren. Die Ausfallschranke ist dann durch das Quantil der Ausfallwahrscheinlichkeit gegeben. Um Abhängigkeiten zu modellieren, wird diese Zufallsvariable als Summe zweier unabhängiger Zufallsvariablen modelliert. Dabei repräsentiert die erste Zufallsvariable das individuelle Risiko und die zweite Zufallsvariable das gemeinsame Risiko. Im dritten Kapitel wird dieses Konzept auf Erstaustrittszeitmodelle übertragen. Dazu muss der Kreditwürdigkeitsprozess als Summe zweier unabhängiger Prozesse modelliert werden. Bezug nehmend auf das vorangegangene Kapitel werden aufgrund ihrer Additivitätseigenschaft dazu Squared Bessel Prozesse gewählt. Als Ausfallschranke dient wiederum die Null. Um die gemeinsame Wahrscheinlichkeit, d.h. die Wahrscheinlichkeit, dass die Prozesse beider Kreditnehmer bis zu einem bestimmten Zeitpunkt die Null erreichen, zu berechnen, muss auf den gemeinsamen Prozess bis zum festgelegten Zeithorizont bedingt werden. Da dies jedoch nicht realisierbar ist, wird, um dieses Problem zu beheben, der gemeinsame Prozess durch einen gestoppten Prozess ersetzt. Falls dieser also die Null erreicht, bleibt er dort. Der Kreditwürdigkeitsprozess besteht danach nur noch aus dem individuellen Prozess. Dadurch erhält das Modell die Charakteristika eines so genannten Common Shock Modells. Der Ausfall kann nur durch ein bestimmtes, alle Kreditnehmer beeinflussendes Ereignis ermöglicht werden, in diesem Fall dem Erreichen der Null des gemeinsamen Prozesses. Der Ausfallzeitpunkt und damit der Erstaustrittszeitpunkt des Kreditwürdigkeitsprozesses setzt sich zusammen als Summe des Erstaustrittszeitpunktes des gemeinsamen Prozesses und der Zeit bis zum ersten Eintreffen in der Null des individuellen Prozesses nach diesem Zeitpunkt. Damit lassen sich dann für jeden gegebenen Zeithorizont Ausfallwahrscheinlichkeiten sowie gemeinsame Ausfallwahrscheinlichkeiten bestimmen.

Im vierten Kapitel werden zwei weitere Ansätze kurz vorgestellt. Im vorangegangenen Modell kann ein Ausfall nur stattfinden, falls ein alle Kreditnehmer beeinflussendes Ereignis vorher eingetreten ist. Damit ein Ausfall auch vor einem solchen Ereignis stattfinden kann, wirkt sich in diesem Modell der Schock lediglich auf die Parameter der Firmenwertprozesse aus. Werden diese also durch geometrische Brownsche Bewegungen modelliert, könnte der Schock beispielsweise eine Erhöhung der Volatilität hervorrufen. In dem zweiten vorgestellten Ansatz wird der Driftparameter selbst als Zufallsvariable modelliert.

Das letzte Kapitel gibt einen bereits veröffentlichten Artikel über die Verwendung von Copulas in dem Kreditrisikomodell CreditRisk+, (Ebmeyer et al., 2007), wieder.

Contents

1 Introduction **1**

1.1 Merton Approach 3

1.2 First Passage Approach 4

1.2.1 Black and Cox (1976) 4

1.2.2 Time-dependent Default Barrier 5

1.3 Distance to Default Models 6

1.3.1 Hull and White (2001) 6

1.3.2 Avellaneda and Zhu (2001) 9

2 Credit Barrier model by Albanese et al. **12**

2.1 Introduction 12

2.1.1 The Credit Quality Process 13

2.2 Stochastic Transformation 14

2.2.1 Change of Measure 15

2.2.2 Doob's h-transform 27

2.2.3 Removal of Drift 31

2.3 Default Probability 35

2.4 Stochastic Time Change 37

3 A Structural Squared Bessel Model **39**

3.1 Squared Bessel Processes 41

3.1.1 Transition Densities 46

3.2 Simulation of Squared Bessel Processes 47

3.2.1 Euler-Maruyama Scheme 47

3.2.2 An implicit Scheme . . . 48
3.2.3 Exact Simulation . . . 50
3.2.4 Some Paths of Squared Bessel Processes . . . 51
3.3 First Hitting Time of Squared Bessel Processes . . . 56
3.4 The Default Time . . . 67
3.5 Joint Default Probability . . . 72
3.6 Examples . . . 74
3.6.1 Conclusions and Remarks . . . 81
3.7 Large Homogeneous Portfolio Approximation . . . 81
3.7.1 Vasicek Approach . . . 82
3.7.2 Squared Bessel Approach . . . 84
3.7.3 Comparison . . . 86

4 Alternative models 92
4.1 A Structural Markov Switching Model - Single Event Case . . . 92
4.1.1 The Default Time . . . 92
4.1.2 Joint Default Probability . . . 98
4.1.3 A Hybrid Model . . . 99
4.2 Random drift . . . 100
4.2.1 Model Description . . . 100
4.2.2 Examples . . . 102
4.2.3 Joint Default Probability . . . 105

5 The Role of Copulas in the CreditRisk+™Framework 108
5.1 Introduction . . . 109
5.2 The CreditRisk+ Default Copula . . . 112
5.3 The Compound Gamma Model . . . 116
5.4 The Characteristic Function . . . 119
5.5 The Limit Distribution . . . 121
5.5.1 The Case of Bernoulli - Default Variables . . . 126

5.6 Measuring Diversification 130

Bibliography **138**

Chapter 1

Introduction

Credit risk can be defined as the risk that a counterparty or a bank borrower will not be able to settle its obligations according to agreed terms. Since credit risk is still the leading source of problems for a bank, a main task of its risk management is to analyze the credit risk of the entire portfolio as well as the risk of individual credits or transactions. Considering a single counterparty, for a bank there are three issues to pay attention to:

- credit exposure, i.e., the amount of the outstanding obligations in case of default.
- recovery rate, i.e., the fraction of the credit exposure, that can be recovered in case of default.
- default probability, i.e., the likelihood that the counterparty will default on its obligations over some specified time horizon.

At least as important as the single counterparty risk is the consideration of credit risk at the portfolio level. One objective of assessing portfolio credit risk is the increase of diversification, that reduces risk as well as to decrease credit concentration effects. There are several reasons for the high attention that is paid to portfolio credit risk modelling. On the one hand there is an interest in getting the Basel II accord permission to use internal models for calculating credit risk capital charges and to improve the capital allocation for portfolio credit risk. On the other hand there is the necessity to price credit derivatives that are linked to baskets of defaultable obligations and the increasing importance of collateralized debt obligations. Hence, an important task in risk management is the study of the dependence structure of different obligors.

The main objective is to develop appropriate models that do not just produce reasonable single default probabilities but also provide joint default probabilities for at least two obligors. The credit risk literature often distinguishes between two different types of credit risk models:

- Reduced form models
- Firm value based models

Reduced form models were introduced for the first time in (Jarrow and Turnbull, 1992) and studied in (Jarrow and Turnbull, 1995), (Duffie and Singleton, 1999) and others.

In contrast to firm value models, the time of default in reduced form models is not linked to the value of the firm. It is modelled as the first jump of an exogenously given jump process. The intensity rate of this process, called default intensity represents the instantaneous default probability. The simplest case is a constant default intensity λ. Then the jump process is a Poisson process with intensity λ and the default time τ is exponentially distributed with parameter λ. In more advanced models the jump process is a Cox process. Then the default intensity itself is a stochastic process. In order to model dependence between defaults of different firms, it is possible to introduce a dependence structure for the different default intensities.

In firm value based models, also called structural models, the credit quality of the firm itself is modelled as a stochastic process. Thus, a default takes place if the process lies below a certain barrier. This process, also called ability-to-pay process and the default barrier can be directly linked to the assets and the liabilities of the firm. The structural models are based on the Merton model, introduced in (Merton, 1974), where the option pricing theory developed by (Black and Scholes, 1973) is applied. In this chapter the Merton model as well as some extensions are introduced.

1.1 Merton Approach

As already mentioned, the Merton model introduced in (Merton, 1974) uses the option pricing framework of (Black and Scholes, 1973) to create a link between the default risk and the asset and liability structure of the firm. In this simplified approach it is assumed that the firm has issued only equity and debt, where the equity receives no dividends. The debt is a single bond with promised payment K at time T. Since the debt is senior to the equity, the value of the equity is 0 at time T if the asset value V_T lies below K and is equal to $V_T - K$ otherwise, i.e.,

$$E_T = \max(V_T - K, 0).$$

Thus, the equity can be seen as a call option on the asset value with strike K. Applying the option price model of (Black and Scholes, 1973) leads to

$$E_t = V_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

with

$$\begin{aligned} d_1 &= \frac{\log\left(\frac{V_t}{K}\right) + r(T-t) + \frac{1}{2}\sigma\sqrt{T-t}}{\sigma\sqrt{T-t}} \\ d_2 &= d_1 - \sigma\sqrt{T-t} \end{aligned}$$

where r is the risk free interest rate and σ is the volatility of the asset value. Now we are interested in the probability of default under the real world measure. Since the payment is due only at maturity T, a default can only happen at the fixed date T. In the case of default, the firm's asset value at time T does not suffice to pay the amount of K. If we define τ to be the random variable representing the default time, τ can only take the values T and ∞:

$$\tau = \begin{cases} T & \text{if } V_T \leq K \\ \infty & \text{otherwise.} \end{cases}$$

Since the Black-Scholes environment is assumed to hold, the asset value $(V_t)_{0\leq t\leq T}$ follows a geometric Brownian motion, i.e., its dynamics are represented by

$$dV_t \;=\; \mu V_t dt \;+\; \sigma V_t dW_t \quad, V_0 > 0$$

where $\mu \in \mathbb{R}$ is called the drift parameter and $\sigma > 0$ the volatility parameter. W_t denotes a standard Brownian motion. Applying Itô's Lemma, we get

$$V_t = V_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right).$$

Then the default probability is

$$\begin{aligned}
\mathbb{P}(\tau = T) &= \mathbb{P}(V_T \leq K) \\
&= \mathbb{P}\left(V_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma W_T\right) \leq K\right) \\
&= \mathbb{P}\left(\sigma W_T \leq \log\left(\frac{K}{V_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)T\right) \\
&= \Phi\left(\frac{\log\left(\frac{K}{V_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)
\end{aligned}$$

The obvious shortcoming of this approach is that a default can only happen at maturity of the zero coupon bond.

1.2 First Passage Approach

1.2.1 Black and Cox (1976)

The first approach to overcome this problem was made by (Black and Cox, 1976). They extended the Merton model to the case when the firm can default at any time. So the default time τ becomes a continuous random variable on $(0, \infty]$. In their approach the default time τ describes the first time the firm value falls below a certain deterministic time-dependent threshold $D(t) = De^{\lambda t}$. So τ is defined as

$$\tau = \inf\{t : V_t < De^{\lambda t}\}.$$

Nevertheless we first describe the case of a constant threshold D. Still modelling the firm value V as a geometric Brownian motion and the default barrier D as fixed and deterministic we get for the default probability, i.e., the probability that the firm defaults before time t as

$$\begin{aligned}\mathbb{P}(\tau \leq t) &= \mathbb{P}(\min_{0\leq s\leq t} V_s < D)\\ &= \mathbb{P}\left(\min_{0\leq s\leq t} V_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)s + \sigma W_s\right) < D\right)\\ &= \mathbb{P}\left(\min_{0\leq s\leq t}\left(\left(\mu - \frac{1}{2}\sigma^2\right)s + \sigma W_s\right) < \log\left(\frac{D}{V_0}\right)\right).\end{aligned}$$

This is the probability, that the running minimum of the arithmetic Brownian motion $X_t = \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t$ lies below $\log\left(\frac{D}{V_0}\right)$ and is given as

$$\Phi\left(\frac{\log\left(\frac{D}{V_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)t}{\sigma\sqrt{t}}\right) + \left(\frac{D}{V_0}\right)^{\frac{2\mu}{\sigma^2}-1}\Phi\left(\frac{\log\left(\frac{D}{V_0}\right) + \left(\mu - \frac{1}{2}\sigma^2\right)t}{\sigma\sqrt{t}}\right).$$

1.2.2 Time-dependent Default Barrier

The constant default barrier D is now replaced by the deterministic default barrier function $D(t) = De^{\lambda t}$. Thus, τ is the first time the process V_t lies below the time-dependent barrier $D(t)$. Since

$$\{V_t < D(t)\} = \{V_t < De^{\lambda t}\} = \{V_t e^{-\lambda t} < D\}\,\forall t$$

we get the following default probability:

$$\begin{aligned}
\mathbb{P}(\tau \leq t) &= \mathbb{P}\left(\min_{0\leq s\leq t} V_s e^{-\lambda s} < D\right) \\
&= \mathbb{P}\left(\min_{0\leq s\leq t} V_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2 - \lambda\right) s + \sigma W_s\right) < D\right) \\
&= \mathbb{P}\left(\min_{0\leq s\leq t} \left(\left(\mu - \frac{1}{2}\sigma^2 - \lambda\right) s + \sigma W_s\right) < \log\left(\frac{D}{V_0}\right)\right) \\
&= \Phi\left(\frac{\log\left(\frac{D}{V_0}\right) - \left(\mu - \frac{1}{2}\sigma^2 - \lambda\right) t}{\sigma\sqrt{t}}\right) \\
&+ \left(\frac{D}{V_0}\right)^{\frac{2(\mu-\lambda)}{\sigma^2}-1} \Phi\left(\frac{\log\left(\frac{D}{V_0}\right) + \left(\mu - \frac{1}{2}\sigma^2 - \lambda\right) t}{\sigma\sqrt{t}}\right).
\end{aligned}$$

The approach by (Zhou, 1997) proposes to set the parameter λ of the default barrier to $\lambda = \mu - \frac{1}{2}\sigma^2$. So the default probability reduces to

$$\mathbb{P}(\tau \leq t) = 2\Phi\left(\frac{\log\left(\frac{D}{V_0}\right)}{\sigma\sqrt{t}}\right).$$

1.3 Distance to Default Models

In this section we present two structural models, where although the ability-to-pay processes indicate the creditworthiness of the firm, the processes are not directly identified with the firm's asset value. The first model considers a discrete time grid such that a default can only happen at one of these pre-defined time points. The second model extends the first model to the continuous time case.

1.3.1 Hull and White (2001)

In (Hull and White, 2001) the authors extend their own proposal to the valuation of Credit Default Swaps, see (Hull and White, 2000), by allowing the payoff to be contingent on defaults by multiple reference entities. Thus they

have to incorporate default correlation of those entities.
Here we restrict ourselves on the calculation of the default correlation. We will not address the issue of valuation of Credit Default Swaps. Although following the ideas of (Merton, 1974), (Black and Cox, 1976) and (Zhou, 1997), this approach is based on risk-neutral default probabilities, since in order to price contingent claims, risk-neutral probabilities shall be used. Considering a set of companies $\mathscr{J} = \{1, \dots, N\}$, for each company $j \in \mathscr{J}$, the credit index $(X_j(t))_{t \geq 0}$ (a kind of ability to pay process) is defined as a standard Brownian motion. The instantaneous correlation between two firms $j, k \in \mathscr{J}$ is denoted by ρ_{jk}. The default correlation $\beta_{jk}(T)$ is defined as

$$\beta_{jk}(T) = \frac{P_{jk}(T) - Q_j(T)Q_k(T)}{\sqrt{Q_j(T) - Q_j(T)^2}\sqrt{Q_k(T) - Q_k(T)^2}}$$

where $P_{jk}(T)$ denotes the joint default probability of j and k and $Q_j(T)$ the default probability of company j by time T. To obtain $\beta_{jk}(T)$ from the credit index correlation ρ_{jk}, one has to simulate the credit indices, since $P_{jk}(T)$ cannot be specified directly. The default probability $Q_j(T)$ is defined as

$$Q_j(T) = \int_0^T q_j(t)dt$$

where $q_j(t)$ denotes the default probability density. These risk-neutral default probability densities are assumed to be estimated from bond prices or CDS spreads. Now the time dependent default barrier $K_j(t)$ has to be determined such that $q_j(t)$ coincides with the first passage time density, i.e., $K_j(t)$ has to be chosen so that with the first passage time

$$\tau_j := \inf\{s \geq 0 : X_j(s) \leq K_j(s)\}$$

we get

$$\mathbb{P}(\tau_j \leq t) = Q_j(t).$$

For this the default probability density is discretized so that defaults can only happen at times t_i, $1 \leq i \leq n$. Now q_{ij} is defined as the risk-neutral probability of default by company j at time t_i, i.e.,

$$q_{ij} = \mathbb{P}(\tau_j = t_i).$$

A default occurs at time t_i if the credit index $X_j(t_i)$ lies below the default barrier $K_{ij} = K_j(t_i)$ and no default has happened before t_i. Furthermore we define $f_{ij}(x)$ to be the density of $X_j(t_i)$ under the condition that no default happened prior to t_i. These definitions imply that the cumulative probability of company j defaulting by time t_i satisfies

$$\mathbb{P}(\tau_j \leq t_i) = \sum_{k=1}^{i} q_{kj} = 1 - \int_{K_{ij}}^{\infty} f_{ij}(x)dx.$$

Then K_{ij} and $f_{ij}(x)$ have to be determined inductively from the risk-neutral default probabilities. Since X_j follows a standard Brownian motion, we get

$$f_{1j}(x) = \frac{1}{\sqrt{2\pi t_1}} \exp\left(-\frac{x^2}{2t_1}\right)$$

and

$$q_{1j} = \mathbb{P}(X_j(t_1) \leq K_{1j}) = \mathbb{P}(N(0,t_1) \leq K_{1j}) = \Phi\left(\frac{K_{1j}}{\sqrt{t_1}}\right).$$

So the default barrier for t_1 is

$$K_{1j} = \sqrt{t_1}\Phi^{-1}(q_{1j}).$$

Then for $2 \leq i \leq n$, the probability of a default at time t_i, conditioned on the event $\{X_j(t_{i-1}) = u, u > K_{i-1,j}\}$ equals

$$\begin{aligned} q_{ij}(u) &:= \mathbb{P}(X_j(t_i) \leq K_{ij} | X_j(t_{i-1}) = u) \\ &= \mathbb{P}(N(u, t_i - t_{i-1}) \leq K_{ij}) = \Phi\left(\frac{K_{ij} - u}{\sqrt{t_i - t_{i-1}}}\right). \end{aligned}$$

Thus

$$q_{ij} = \int_{K_{i-1,j}}^{\infty} q_{ij}(u) f_{i-1,j}(u) du = \int_{K_{i-1,j}}^{\infty} f_{i-1,j}(u) \Phi\left(\frac{K_{ij} - u}{\sqrt{t_i - t_{i-1}}}\right) du. \qquad (1.1)$$

Equivalently $f_{ij}(x)$ is determined for $x > K_{i-1,j}$:

$$f_{ij}(x) = \int\limits_{K_{i-1,j}}^{\infty} f_{i-1,j}(u) \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left(-\frac{(x-u)^2}{2(t_i - t_{i-1})} \right) du. \tag{1.2}$$

Now equations 1.1 and 1.2 have to be solved numerically.

1.3.2 Avellaneda and Zhu (2001)

In (Avellaneda and Zhu, 2001) the authors describe an extension to the approach of (Hull and White, 2001), introduced in the last subsection. Here the continuous time case is considered and the credit index is modelled as a general diffusion process

$$X(t) = X(0) + \int\limits_0^t a(X(s), s)ds + \int\limits_0^t \sigma(X(s), s)dW(s)$$

where $W(t)$ is a standard Brownian motion. The default time is the first time the credit index hits the barrier function $b(t)$, i.e.,

$$\tau = \inf\{t \geq 0 | X(t) \leq b(t)\}.$$

The survival probability density function of $X(t)$ is given as

$$f(x,t)dx = \mathbb{P}(x < X(t) < x + dx, \tau \geq t).$$

Then $f(x,t)$ satisfies the forward Fokker-Planck equation

$$f_t = \frac{1}{2}(\sigma^2(x,t)f)_{xx} - (a(x,t)f)_x, \ t > 0, \ x > b(t) \tag{1.3}$$

with the following conditions

$$f(x,t)|_{t=0} = \delta(x - X_0) \tag{1.4}$$

$$f(x,t)|_{x=b(t)} = 0. \tag{1.5}$$

Thus the default probability is given by

$$P(t) = 1 - \int\limits_{b(t)}^{\infty} f(x,t)dx.$$

Now $b(t)$ has to be determined so that $P(t)$ coincides with the market-given default probabilities. Therefore we first determine the derivative of $P(t)$:

$$\begin{aligned}
P'(t) &= -\int_{b(t)}^{\infty} \frac{\partial f}{\partial t} dx + f(b(t),t)b'(t) \\
&\underset{1.5}{=} -\int_{b(t)}^{\infty} \frac{\partial f}{\partial t} dx \\
&\underset{1.3}{=} -\frac{1}{2}\int_{b(t)}^{\infty} (\sigma^2(x,t)f)_{xx} dx + \int_{b(t)}^{\infty} (a(x,t)f)_x dx \\
&\underset{1.5}{=} \frac{1}{2}\frac{\partial}{\partial x}(\sigma^2 f)\bigg|_{x=b(t)} .
\end{aligned}$$

Thus $b(t)$ has to satisfy

$$f(b(t),t) = 0 \text{ and } P'(t) = \frac{1}{2}\frac{\partial}{\partial x}(\sigma^2 f)\bigg|_{x=b(t)}$$

A risk-neutral distance-to-default (RNDD) process, i.e., calibrated to the market implied default probability $P(t)$, is defined as

$$Y(t) = Y(0) + \int_0^t \tilde{a}(Y,t)dt + \int_0^t \tilde{\sigma}(Y,t)dW(t), \; Y(0) > 0$$

so that

$$P(t) = \mathbb{P}\left(\inf_{0\le s\le t} Y(s) \le 0\right).$$

Then

$$Y(t) = X(t) - b(t)$$

is a RNDD process with

$$\tilde{a}(Y,t) = a(X,t) - b'(t).$$

Its survival density equals

$$u(y,t) = f(y + b(t), t).$$

Thus, the Fokker-Planck equation is given by

$$\begin{aligned} u_t &= \frac{1}{2}(\sigma^2 u)_{yy} - (\tilde{a}u)_y \\ &= \frac{1}{2}(\sigma^2 u)_{yy} - a u_y + b'(t) u_y. \end{aligned}$$

Furthermore we have

$$P'(t) = \frac{1}{2} \left(\frac{\partial}{\partial y} (\sigma^2 u) \right) \Big|_{y=0}.$$

Hence, $b'(t)$ has to be determined so that this condition is satisfied for all $t > 0$. That means, the free boundary problem for the credit index in (Hull and White, 2001) is transformed into a control problem for the RNDD.

Chapter 2

Credit Barrier model by Albanese et al.

2.1 Introduction

In (Albanese et al., 2003) the authors extend the distance to default models presented in the preceding subsections. In this article, the credit quality of an obligor is represented by a credit quality process, whose current state can be identified with the obligor's credit rating. Since the calibration to default probabilities as well as migration rates failed in the diffusion case, they introduced jumps by using a stochastic time change. The presented model is a single obligor model and cannot be extended to the multidimensional case in order to determine joint default probabilities. Nevertheless in this chapter we will review the development of the credit quality process in the diffusion case. This process results from a two-stage transformation, called stochastic transformation, applied to a Squared Bessel process. In the next chapter the Squared Bessel process is used to develop a multi issuer model which disregards rating transition probabilities and instead focusses on default probabilities and joint default probabilities. Nevertheless the stochastic transformation used to produce the credit quality process will be presented here in detail. First we summarize the requirements for the credit quality process.

2.1.1 The Credit Quality Process

The credit quality process $(Y_t)_{t\geq 0}$ is to be identified with the credit rating of the obligor. Next we list the properties we require from this process.

- Y_t is nonnegative
- the lower boundary 0 is absorbing
- the upper boundary $b_{\max}$ is unattainable
- the volatility is state dependent
- the process is driftless
- the process has stationary increments.

Since Y_t can be seen as a distance-to-default it is chosen to be nonnegative. The state 0 represents the default state. Thus this state is absorbing. The interval $(0, b_{\max})$ is divided into K subintervals, where $(0, b_1)$ represents the worst rating and $[b_{K-1}, b_{\max})$ represents the best rating. The subintervals do not have to be equidistant. Since higher quality ratings are less volatile than low quality ratings the volatility is chosen to be state dependent. Finally the process has stationary increments since migration probabilities remain consistent over time and through economic cycles.

The migration probabilities are determined as follows. If the obligor has rating i at time 0, i.e., the initial state y_0^i satisfies $y_0^i \in [b_{i-1}, b_i)$, the probability to have rating j at time $t > 0$ is given as

$$p_j(y_0^i, t) = \int_{b_{j-1}}^{b_j} p_Y(t, y_0^i, y)dy$$

where $p_Y(t, y_0^i, y)$ is the transition probability density of the credit quality process $(Y_t)_{t\geq 0}$. There it is not clear, how to determine the initial state y_0^i.

The default probability is given as

$$p^D(y_0^i, t) = 1 - \int_0^{b_{\max}} p_Y(t, y_0^i, y)dy.$$

In the next section the state dependent volatility $\sigma(Y_t)$ is determined in order to satisfy the requirements listed above. Then the process Y_t satisfies the following integral equation:

$$Y_t = Y_0 + \int_0^t \sigma(Y_s)dW_s$$

where W_t is a standard Brownian motion. Thus, Y_t is a martingale with local volatility function.
In order to achieve analytical tractability of the probability kernel $p_Y(t, y_0, y)$, the process Y_t is chosen to be the result of a stochastic transformation, as described in (Albanese and Campolieti, 2001), consisting of a change of measure and a nonlinear transformation of a process X_t with a closed form transition probability density. The stochastic transformation preserves the analytical tractability by allowing the probability kernel p_Y to be expressed in terms of the probability kernel p_X of the process X_t. Moreover the stochastic transformation has to provide sufficient parameters for the purpose of calibration.

2.2 Stochastic Transformation

The underlying process X_t is chosen to be a Squared Bessel process, i.e., X_t satisfies

$$X_t = x_0 + \delta t + 2\int_0^t \sqrt{X_s}dW_s \tag{2.1}$$

with the well known probability kernel

$$p_X(t, x_0, x) = \frac{1}{2t}\left(\frac{x}{x_0}\right)^{\frac{\theta}{2}} \exp\left(-\frac{x + x_0}{2t}\right) I_\theta\left(\frac{\sqrt{x x_0}}{t}\right)$$

where $\theta = \frac{\delta}{2} - 1$ and I_θ is the modified Bessel function of the first kind of order θ. For a deeper survey of Squared Bessel processes see section 3.1.

As already mentioned the stochastic transformation consists of two steps. The first step is a change of measure and the second step is a removal of the drift. These two steps will be described in the next subsections.

2.2.1 Change of Measure

The Squared Bessel process X_t is a Markov process on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_{x_0})$ according to the following definition (see ,e.g., (Borodin and Salminen, 1996)).

Definition 2.2.1 *An adapted stochastic process $(X_t)_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and taking values in a general state space $(E, \mathscr{E})$ is called a time-homogeneous Markov process if for every t there exists a kernel $P_t : E \times \mathscr{E} \mapsto [0, 1]$, called a transition function, such that*

1. *for all $A \in \mathscr{E}$ and $t, s \geq 0$,*

$$\mathbb{P}(X_{t+s} \in A|\mathcal{F}_t) = P_s(X_t, A) \quad a.s.$$

2. *for all $A \in \mathscr{E}$ and t $P_t(\cdot, A)$ is measurable*

3. *for all $x \in E$ and t $P_t(x, \cdot)$ is a probability measure on $(E, \mathscr{E})$*

4. *for all $A \in \mathscr{E}$ and $x \in E$,*

$$P_0(x, A) = \delta_x(A)$$

where $\delta_x(\cdot)$ is the Dirac measure, i.e.,

$$\delta_x(A) = \begin{cases} 1 & \textit{if } x \in A \\ 0 & \textit{if } x \notin A \end{cases}$$

5. *The Chapman-Kolmogorov equation holds:*

$$P_{t+s}(x, A) = \int_E P_t(x, dy) P_s(y, A) \quad \forall x \in E, A \in \mathscr{E}, s, t \geq 0$$

The process starts almost surely in x_0, i.e., $\mathbb{P}_{x_0}(X_0 = x_0) = 1$. As from now we can unambiguously omit the subscript x_0 from the probability measure. The new probability measure $\mathbb{Q}$ has to be locally absolutely continuous with respect to $\mathbb{P}$ according to the following definition:

Definition 2.2.2 *A probability measure $\mathbb{Q}$ defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$ is said to be locally absolutely continuous with respect to $\mathbb{P}$, defined on the same filtered space, if, for all $t \geq 0$, $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}_t$, i.e., for all $F \in \bigcup_{t\in[0,\infty)} \mathcal{F}_t$,*

$$\mathbb{P}(F) = 0 \Rightarrow \mathbb{Q}(F) = 0.$$

Then for every $t \geq 0$ the Radon-Nikodým derivative of $\mathbb{Q}$ relative to $\mathbb{P}$ with respect to $\mathcal{F}_t$ is given by

$$L_t = \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t}.$$

The process L_t is a positive $\mathbb{P}$-martingale. Now the following result holds:

Lemma 2.2.3 *Let $\mathbb{Q}$ be locally absolutely continuous with respect to $\mathbb{P}$. The random variable ζ is defined by*

$$\zeta = \inf\{t > 0 | L_t = 0\}$$

and $\zeta = \infty$ if $\{L_t = 0\} = \emptyset\, \forall t$. Then

1. *$\zeta = \infty \quad \mathbb{Q} - a.s.$*

2. *Let V_t be an $\mathcal{F}_t$-adapted process, then for $s \leq t$*

$$\mathbb{E}_{\mathbb{Q}}(V_t|\mathcal{F}_s) = L_s^{-1}\mathbb{E}_{\mathbb{P}}(L_t V_t|\mathcal{F}_s)\mathbb{1}_{\{L_s\neq 0\}} \tag{2.2}$$

3. *Let $\tilde{M}$ be an adapted process. Then*

$$\tilde{M} \in \mathscr{M}^{loc}(\mathbb{Q}) \Leftrightarrow \tilde{M}L \in \mathscr{M}^{loc,\zeta}(\mathbb{P})$$

where $\mathscr{M}^{loc}(\mathbb{Q})$ is the set of local $\mathbb{Q}$-martingales and $\mathscr{M}^{loc,\zeta}(\mathbb{P})$ are the local $\mathbb{P}$-martingales up to ζ.

Proof:

1. Since $\{\zeta < t\} \in \mathcal{F}_t$ we have

$$\mathbb{Q}(\zeta < t) = \mathbb{E}_{\mathbb{P}}(L_t \mathbb{1}_{\{\zeta<t\}}) = \mathbb{E}_{\mathbb{P}}(L_\zeta \mathbb{1}_{\{\zeta<t\}}) = 0$$

2. Both sides of 2.2 are $\mathcal{F}_s$-measurable. According to the definition of the conditional expectation $\mathbb{E}_{\mathbb{Q}}(V_t|\mathcal{F}_s)$ is unique with

$$\mathbb{E}_{\mathbb{Q}}(\mathbb{1}_A V_t) = \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_A \mathbb{E}_{\mathbb{Q}}(V_t|\mathcal{F}_s))$$

for any set $A \in \mathcal{F}_s$. Thus we have to show that

$$\mathbb{E}_{\mathbb{Q}}(\mathbb{1}_A L_s^{-1} \mathbb{1}_{\{L_s \neq 0\}} \mathbb{E}_{\mathbb{P}}(L_t V_t|\mathcal{F}_s)) = \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_A V_t).$$

For this we make use of the following:
For $t > 0$ we have

$$\mathbb{E}_{\mathbb{Q}}(Y) = \mathbb{E}_{\mathbb{P}}(L_t Y) \tag{2.3}$$

for any $\mathcal{F}_t$-measurable random variable Y.
Then

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_A V_t) \underset{(2.3)}{=} \mathbb{E}_{\mathbb{P}}(\mathbb{1}_A L_t V_t) \\ = \; & \mathbb{E}_{\mathbb{P}}(\mathbb{1}_A \mathbb{1}_{\{L_s\neq 0\}} L_t V_t) = \mathbb{E}_{\mathbb{P}}(\mathbb{1}_A \mathbb{1}_{\{L_s\neq 0\}} \mathbb{E}_{\mathbb{P}}(L_t V_t|\mathcal{F}_s)) \\ = \; & \mathbb{E}_{\mathbb{P}}(\mathbb{1}_A L_s L_s^{-1} \mathbb{E}_{\mathbb{P}}(L_t V_t|\mathcal{F}_s)) \\ \underset{(2.3)}{=} \; & \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_A L_s^{-1} \mathbb{1}_{\{L_s\neq 0\}} \mathbb{E}_{\mathbb{P}}(L_t V_T|\mathcal{F}_s)). \end{aligned}$$

3. $'' \Leftarrow''$: Let $\tilde{M}L$ be a $\mathbb{P}$-martingale. Then

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}(\tilde{M}_t|\mathcal{F}_s) &= L_s^{-1} \mathbb{E}_{\mathbb{P}}(\tilde{M}L_t|\mathcal{F}_s) \mathbb{1}_{\{L_s\neq 0\}} \\ &= L_s^{-1} \tilde{M}_s L_s \mathbb{1}_{\{L_s\neq 0\}} = \tilde{M}_s \quad \mathbb{Q}-a.s. \end{aligned}$$

$'' \Rightarrow''$: Let $\tilde{M}$ be a $\mathbb{Q}$-martingale. Then

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}(\tilde{M}L_t|\mathcal{F}_s) \mathbb{1}_{\{L_s\neq 0\}} &= L_s \mathbb{1}_{\{L_s\neq 0\}} \mathbb{E}_{\mathbb{Q}}(\tilde{M}_t|\mathcal{F}_s) \\ &= \tilde{M}_s L_s \mathbb{1}_{\{L_s\neq 0\}}. \end{aligned}$$

$\square$

Theorem 2.2.4 *(Girsanov's Theorem) Let $\mathbb{Q}$ be locally absolutely continuous with respect to $\mathbb{P}$ and set*

$$L_t = \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t}.$$

If $(M_t)_{t\geq 0}$ is a local $\mathbb{P}$-martingale, then

$$\tilde{M}_t := M_t - \int_0^t L_s^{-1} d\langle M, L\rangle_s$$

is a local martingale under $\mathbb{Q}$.

Proof: To show that $\tilde{M} \in \mathscr{M}^{loc}(\mathbb{Q})$, we can also show that $\tilde{M}L \in \mathscr{M}^{loc,\zeta}(\mathbb{P})$. Let

$$A_t := \int_0^t L_s^{-1} d\langle M, L\rangle_s.$$

For

$$(\tilde{M}L)_t = L_t(M_t - A_t)$$

with Itô's product formula it follows that

$$\begin{aligned}
\tilde{M}_t L_t &= M_0 L_0 + \int_0^t (M_s - A_s) dL_s + \int_0^t L_s dM_s - \int_0^t L_s dA_s + \langle M, L\rangle_t \\
&= M_0 L_0 + \int_0^t (M_s - A_s) dL_s + \int_0^t L_s dM_s - \langle M, L\rangle_t + \langle M, L\rangle_t \\
&= M_0 L_0 + \int_0^t \tilde{M} dL_s + \int_0^t L_s dM_s.
\end{aligned}$$

$\square$

Applied to the special case of a Brownian motion we get

Corollary 2.2.5 *Let W be a Brownian motion under $\mathbb{P}$ and let $\mathbb{Q}$ be locally absolutely continuous with respect to $\mathbb{P}$. Then*

$$\tilde{W}_t = W_t - \int_0^t L_s^{-1} d\langle W, L\rangle_s$$

is a Brownian motion under $\mathbb{Q}$.

Proof: $\tilde{W}$ is a $\mathbb{Q}$-martingale with $\langle \tilde{W}\rangle_t = t$. □

If it is required that the process X_t remains a Markov process under the new measure $\mathbb{Q}$, the Radon-Nikodým derivative L_t must be a function of X_t and t, i.e.,

$$L_t = g(X_t, t)$$

with $g(X_0, 0) = 1$. Then for the $\mathbb{Q}$-Brownian motion $\tilde{W}$ it follows

$$\tilde{W}_t = W_t - \int_0^t \frac{1}{g(X_s, s)} d\langle W, g(X_., \cdot)\rangle_s.$$

According to Itô's formula we have

$$g(X_t, t) = g(X_0, 0) + \int_0^t g_x(X_s, s) dX_s + \int_0^t g_{xx}(X_s, s) d\langle X\rangle_s + \int_0^t g_s(X_s, s) ds$$

where $g_x(X_s, s)$ is the first derivative and $g_{xx}(X_s, s)$ is the second derivative with respect to the first component. $g_s(X_s, s)$ is the first derivative with respect to the second component. Thus we have

$$\begin{aligned}
\langle W, g(X_., \cdot)\rangle_t &= \int_0^t g_x(X_s, s) d\langle W, X\rangle_s \\
&\quad + \int_0^t g_{xx}(X_s, s) d\langle W, \langle X\rangle\rangle_s + \int_0^t g_s(X_s, s) d\langle W, s\rangle_s \\
&= \int_0^t g_x(X_s, s) d\langle W, X\rangle_s.
\end{aligned}$$

With

$$\langle W, X\rangle_t = \int_0^t 2\sqrt{X_s}d\langle W\rangle_s = \int_0^t 2\sqrt{X_s}ds$$

we have

$$\begin{aligned}\langle W, g(X_\cdot, \cdot)\rangle_t &= \int_0^t g_x(X_s, s)d\langle W, X\rangle_s \\ &= \int_0^t g_x(X_s, s)2\sqrt{X_s}ds\end{aligned}$$

Altogether we get

$$\tilde{W}_t = W_t - \int_0^t \frac{g_x(X_s, s)}{g(X_s, s)}2\sqrt{X_s}ds.$$

Under the new measure $\mathbb{Q}$ the process X_t has the following form

$$\begin{aligned}X_t^{\mathbb{Q}} &= x_0 + \int_0^t \delta ds + \int_0^t 2\sqrt{X_s}d\left(\tilde{W} + \int_0^t \frac{g_x(X_s, s)}{g(X_s, s)}2\sqrt{X_s}ds\right) \\ &= x_0 + \int_0^t \left(\delta + 4X_s\frac{g_x(X_s, s)}{g(X_s, s)}\right)ds + \int_0^t 2\sqrt{X_s}d\tilde{W}_s.\end{aligned}$$

The process is also required to be stationary. This is only possible if $\frac{g_x(X_t,t)}{g(X_t,t)}$ is independent of t. This implies that

$$g(X_t, t) = g(t)h(X_t).$$

Applying Itô's formula again leads to

$$\begin{aligned}dg(X_t, t) &= (g_x(X_t, t)\delta + g_{xx}(X_t, t)2X_t + g_t(X_t, t))dt + 2\sqrt{X_t}g_x(X_t, t)dW_t \\ &= (g(t)h'(X_t)\delta + g(t)h''(X_t)2X_t + g'(t)h(X_t))dt \\ &\quad +2\sqrt{X_t}g(t)h'(X_t)dW_t \\ &= (g(t)(h'(X_t)\delta + h''(X_t)2X_t) + g'(t)h(X_t))dt \\ &\quad +2\sqrt{X_t}g(t)h'(X_t)dW_t.\end{aligned}$$

Since $g(X_t, t)$ is a martingale under $\mathbb{P}$ it has to be driftless, i.e., g has to satisfy

$$g(t)(h'(x)\delta + h''(x)2x) + g'(t)h(x) = 0$$

Since $g(x,t)$ is strictly positive we can rearrange the above equation to obtain

$$\frac{g'(t)}{g(t)} = -\frac{h'(x)\delta + h''(x)2x}{h(x)}.$$

The right hand side does not depend on t. Thus it must be constant, i.e.,

$$\frac{h'(x)\delta + h''(x)2x}{h(x)} = \rho.$$

Then

$$\frac{g'(t)}{g(t)} = -\rho$$

leads to

$$g(t) = e^{-\rho t}.$$

Thus

$$g(x,t) = e^{-\rho t} h(x).$$

The drift condition becomes

$$\begin{aligned} & e^{-\rho t}(h'(x)\delta + h''(x)2x) - \rho e^{-\rho t} h(x) = 0 \\ \Leftrightarrow \quad & h'(x)\delta + h''(x)2x = \rho h(x). \end{aligned}$$

This corresponds to the equation

$$\mathscr{G}_X h(x) = \rho h(x)$$

where $\mathscr{G}_X$ is the infinitesimal generator of the process X. To verify this we first need some definitions and remarks, see (Revuz and Yor, 1991).

Definition 2.2.6 *(Feller semi-group) Let E be a metric space and let $C_0(E)$ be the space of continuous functions on E which vanish at infinity. A Feller semi-group on $C_0(E)$ is a family $T_t, t \geq 0$, of positive linear operators on $C_0(E)$ such that*

1. $T_0 = Id$ *and* $||T_t|| \leq 1$ *for every* t

2. $T_{t+s} = T_t \circ T_s$ *for any pair* $s, t \geq 0$

3. $\lim_{t\downarrow 0} ||T_t f - f|| = 0$ *for every* $f \in C_0(E)$

The transition function of a Markov process (cf. 2.2.1) forms a semi-group and even a Feller semi-group if the following proposition holds.

Proposition 2.2.7 *A transition function* $(P_t)_{t\geq 0}$ *is a Feller semi-group if and only if*

1. $P_t C_0 \subset C_0$ *for each* t

2. $\forall f \in C_0, \forall x \in E,\ \lim_{t\downarrow 0} P_t f(x) = f(x)$

Proof: See (Revuz and Yor, 1991) □

If the transition function of a Markov process is a Feller semi-group, the process is a Feller process. Now we can define the infinitesimal generator of a Feller process.

Definition 2.2.8 *Let* $(V_t)_{t\geq 0}$ *be a Feller process. A function* f *in* C_0 *is said to belong to the domain* $\mathscr{D}$ *of the infinitesimal generator of* V *if the limit*

$$\mathscr{G}_V f = \lim_{t\downarrow 0} \frac{P_t f - f}{t}$$

exists in C_0. *The operator* $\mathscr{G}_V : \mathscr{D} \to C_0$ *is called the infinitesimal generator of* V *or of the semi-group* P_t.

If the process V_t is a diffusion process and satisfies the following stochastic integral equation

$$V_t = v_0 + \int_0^t a(V_s)ds + \int_0^t b(V_s)dW_s$$

with Borel-measurable functions $a(x)$ and $b(x)$, then the infinitesimal generator is given by

$$\mathscr{G}_V f(x) = a(x)f'(x) + \frac{1}{2}b^2(x)f''(x).$$

In the case of the Squared Bessel process we have

$$\mathscr{G}_X f(x) = \delta f'(x) + 2x f''(x).$$

As already mentioned, the function $h(x)$ has to satisfy the following equation

$$\mathscr{G}_V h(x) = \rho h(x).$$

This problem is solved in section 4.6 of (Itô and McKean, 1974). There two solutions are presented that are linearly independent and span all solutions of the above equation.

If we choose H_q to be the first time the process V reaches the value $q \in \mathbb{R}$, i.e.,

$$H_q = \inf\{t \geq 0 | V_t = q\}$$

these solutions are given by

$$\begin{aligned} \psi_\rho^V(x) &= \mathbb{E}_x(e^{-\rho H_q}) \quad \text{if } x \leq q \\ \phi_\rho^V(x) &= \mathbb{E}_x(e^{-\rho H_q}) \quad \text{if } x \geq q. \end{aligned}$$

According to (Rogers and Williams, 1987, p. 292), changing the reference point q will only change $\psi_\rho^V(x)$ and $\phi_\rho^V(x)$ by a multiplicative constant. Furthermore, due to the fact that

$$\mathbb{E}_x(e^{-\rho H_q}) = \mathbb{E}_x(e^{-\rho H_z}) \cdot \mathbb{E}_z(e^{-\rho H_q}) \quad \text{for } x \leq z \leq q$$

and

$$\mathbb{E}_x(e^{-\rho H_q}) = \mathbb{E}_x(e^{-\rho H_z}) \cdot \mathbb{E}_z(e^{-\rho H_q}) \quad \text{for } q \leq z \leq z$$

we get

$$\mathbb{E}_x(e^{-\rho H_z}) = \begin{cases} \frac{\psi_\rho^V(x)}{\psi_\rho^V(z)} & \text{if } x \leq z \\ \frac{\varphi_\rho(x)^V}{\varphi_\rho^V(z)} & \text{otherwise.} \end{cases} \tag{2.4}$$

Thus the function $\psi_\rho^V(x)$ is increasing and the function $\phi_\rho(x)^V$ is decreasing. For a Squared Bessel process of order θ we get, see (Borodin and Salminen, 1996, chap. 4, 2.0.1), that

$$\mathbb{E}_x(e^{-\rho H_z}) = \begin{cases} \frac{x^{-\frac{\theta}{2}} I_\theta(\sqrt{2\rho x})}{z^{-\frac{\theta}{2}} I_\theta(\sqrt{2\rho z})} & x \leq z \\ \frac{x^{-\frac{\theta}{2}} K_\theta(\sqrt{2\rho x})}{z^{-\frac{\theta}{2}} K_\theta(\sqrt{2\rho z})} & x \geq z. \end{cases}$$

Thus, we can set

$$\psi_\rho^X(x) = \frac{I_\theta(\sqrt{2\rho x})}{x^{\frac{\theta}{2}}}$$

and

$$\varphi_\rho^X(x) = \frac{K_\theta(\sqrt{2\rho x})}{x^{\frac{\theta}{2}}}$$

where $K_\theta(x)$ is the modified Bessel function of the second kind of order θ. Thus, in order to obtain two further parameters q_1 and q_2, which in addition ensure that $h(x_0) = 1$, the function $g(x,t)$ defining the Radon-Nikodým derivative L_t, i.e.,

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = g(X_t, t)$$

is given by

$$\begin{aligned} g(x,t) &= e^{-\rho t}h(x) \\ &= e^{-\rho t}(q_1\varphi_\rho^X(x) + q_2\psi_\rho^X(x)) \\ &= e^{-\rho t}x^{-\frac{\theta}{2}}\left(q_1 I_\theta\left(\sqrt{2\rho x}\right) + q_2 K_\theta\left(\sqrt{2\rho x}\right)\right). \end{aligned} \tag{2.5}$$

The next theorem shows that $g(X_t, t)$ is a martingale.

Theorem 2.2.9 *The stochastic process $G_t = g(X_t, t)$, where $g(x,t)$ is given by (2.5), satisfies the following stochastic differential equation*

$$dG_t = e^{-\rho t}2\sqrt{X_t}\frac{\partial h}{\partial x}dW_t.$$

Proof: The function h satisfies

$$\begin{aligned} & \mathscr{G}h = \rho h \\ \Leftrightarrow\quad & \delta\frac{\partial h}{\partial x} + 2x\frac{\partial^2 h}{\partial x^2} = \rho\nu \\ \Leftrightarrow\quad & -\rho h + \delta\frac{\partial h}{\partial x} + 2x\frac{\partial^2 h}{\partial x^2} = 0. \end{aligned}$$

By multiplying both sides with $e^{-\rho t}$ we get

$$-\rho e^{-\rho t}h + \delta e^{-\rho t}\frac{\partial h}{\partial x} + 2xe^{-\rho t}\frac{\partial^2 h}{\partial x^2} = 0.$$

Since

$$\frac{\partial g}{\partial t} = e^{-\rho t} h \qquad \frac{\partial g}{\partial x} = e^{-\rho t}\frac{\partial h}{\partial x} \qquad \frac{\partial^2 g}{\partial x^2} = e^{-\rho t}\frac{\partial^2 h}{\partial x^2}$$

we have

$$\frac{\partial g}{\partial t} + \delta\frac{\partial g}{\partial x} + 2x\frac{\partial^2 g}{\partial x^2} = 0.$$

Thus, by applying Itô's Lemma, $g(X_t, t)$ is driftless, since

$$\begin{aligned} dg_t &= \underbrace{\left(\frac{\partial g}{\partial t} + \delta\frac{\partial g}{\partial x} + 2x\frac{\partial^2 g}{\partial x^2}\right)}_{=0} dt + 2\sqrt{X_t}\frac{\partial g}{\partial x} dW_t \\ &= e^{-\rho t} 2\sqrt{X_t}\frac{\partial h}{\partial x} dW_t. \end{aligned}$$

□

Under the new measure $\mathbb{Q}$, the process has the following form

$$\begin{aligned} X_t^{\mathbb{Q}} &= x_0 + \int_0^t \left(\delta + 4X_s\frac{g_x(X_s, s)}{g(X_s, s)}\right) ds + \int_0^t 2\sqrt{X_s} d\tilde{W}_s \\ &= x_0 + \int_0^t \left(\delta + 4X_s\frac{h'(X_s)}{h(X_s)}\right) ds + \int_0^t 2\sqrt{X_s} d\tilde{W}_s. \end{aligned}$$

Since $g(X_t, t)$ is a positive $\mathbb{P}$-martingale, it satisfies

$$h(x_0) = g(x_0, 0) = \mathbb{E}_{x_0}(g(X_t, t)) = \mathbb{E}_{x_0}(e^{-\rho t} h(X_t)) = e^{-\rho t}\mathbb{E}_{x_0}(h(X_t)).$$

Let I be an Interval with left endpoint $l \geq -\infty$ and right endpoint $r \leq \infty$, such that the diffusion process V takes values in I. Then the function $h(x)$ is ρ-invariant according to the following definition.

Definition 2.2.10 *A nonnegative measurable function $h : I \mapsto \mathbb{R} \cup \{\infty\}$ is called ρ-excessive, $\rho \geq 0$, for V if it has the following two properties:*

- $e^{-\rho t}\mathbb{E}_x(h(V_t)) \leq h(x)\ \forall x$
- $e^{-\rho t}\mathbb{E}_x(h(V_t)) \to h(x), \quad t \to 0\ \forall x.$

A ρ-excessive function h is called ρ-invariant if for all $x \in I$ and $t \geq 0$

$$e^{-\rho t}\mathbb{E}_x(h(V_t)) = h(x).$$

Furthermore, according to (Borodin and Salminen, 1996) we can use the following lemma.

Lemma 2.2.11 *A nonnegative function h is ρ-excessive if and only if h is continuous and satisfies*

$$\mathbb{E}_x(e^{-\rho T}h(V_T)) \leq h(x)$$

where $T := \inf\{t : V_t \in \Gamma\}$ is the first passage time of an arbitrary compact set Γ, with $h(V_T) = 0$ if $T = \infty$.

If we choose H_z to be the first time the process V reaches the value $z \in I$, i.e.,

$$H_z = \inf\{t : V_t = z\},$$

then $V_{H_z} = z$ a.s. when $H_z < \infty$. Since V has a.s. continuous paths and thus

$$\begin{aligned} & \mathbb{E}_x(e^{-\rho H_z}h(z)) \leq h(x) \\ \Leftrightarrow \quad & \mathbb{E}_x(e^{-\rho H_z}) \leq \frac{h(x)}{h(z)}, \end{aligned}$$

with (2.4) we get

$$\frac{h(x)}{h(z)} \geq \begin{cases} \frac{\psi_\rho^V(x)}{\psi_\rho^V(z)} & \text{if } x \leq z \\ \frac{\varphi_\rho^V(x)}{\varphi_\rho^V(z)} & \text{otherwise.} \end{cases}$$

The following lemma will show that $h(x) = q_1\varphi_\rho^V(x) + q_2\psi_\rho^V(x)$ is a ρ-excessive function.

Lemma 2.2.12 *The linear combination*

$$h(x) = q_1\psi_\rho^V(x) + q_2\varphi_\rho^V(x)$$

where $\psi_\rho^V(x)$ and $\varphi_\rho^V(x)$ are the fundamental solutions of $\mathscr{G}f = \rho f$, is a ρ-excessive function.

Proof: For $x \leq z$, we have to show that

$$\frac{q_1\psi_\rho^V(x) + q_2\varphi_\rho^V(x)}{q_1\psi_\rho^V(z) + q_2\varphi_\rho^V(z)} \geq \frac{\psi_\rho^V(x)}{\psi_\rho^V(z)}.$$

Multiplying both sides with the left denominator leads to

$$\begin{aligned} & q_1\psi_\rho^V(x) + q_2\varphi_\rho^V(x) \geq q_1\psi_\rho^V(x) + q_2\frac{\psi_\rho^V(x)\varphi_\rho^V(z)}{\psi_\rho^V(z)} \\ \Leftrightarrow \quad & q_2\varphi_\rho^V(x) \geq q_2\frac{\psi_\rho^V(x)\varphi_\rho^V(z)}{\psi_\rho^V(z)}. \end{aligned}$$

If $q_2 = 0$ this inequality is satisfied, otherwise we get

$$\frac{\varphi_\rho^V(x)}{\varphi_\rho^V(z)} \geq \frac{\psi_\rho^V(x)}{\psi_\rho^V(z)}.$$

Since φ_ρ^V is decreasing and ψ_ρ^V is increasing the inequality is satisfied. The case $x \geq z$ is analog. $\square$

As a conclusion the process $(X_t^{\mathbb{Q}})_{t\geq 0}$ can be interpreted as a so-called Doob's h-transform of X. We will illustrate the concept of Doob's h-transform, sometimes also called ρ-excessive transform, in the next subsection.

2.2.2 Doob's h-transform

First we have to introduce some basic characteristics of diffusion processes, namely the scale function and the speed measure (see e.g. (Rogers and Williams, 1987)). Let V_t be again a diffusion process with state space $I = (l, r)$, where $-\infty \leq l < r \leq \infty$, satisfying

$$dV_t = a(V_t)dt + b(V_t)dW_t. \tag{2.6}$$

Then the scale function $s_V(x)$ is a solution of the following equation

$$\mathscr{G}_V s(x) = 0,$$

i.e.,

$$s'(x)a(x) + \frac{1}{2}s''(x)b^2(x) = 0. \tag{2.7}$$

Furthermore the function $s(x)$ is assumed to have a strictly positive derivative. Though in the literature it is often unambiguously characterized as **the** scale function it is not unique, since $\alpha s(x)+\beta$ with $\alpha > 0$ and $\beta \in \mathbb{R}$ satisfies the same criteria.

Using Itô's formula, we get for the process $Y_t = s_V(V_t)$

$$\begin{aligned} dY_t &= (s'(V_t)a(V_t) + \frac{1}{2}s''(V_t)b^2(V_t))dt + s'(V_t)b(V_t)dW_t \\ &= s'(V_t)b(V_t)dW_t. \end{aligned} \tag{2.8}$$

Thus, Y_t is a local martingale, since it is driftless.

This fact will also be important for the second step of the stochastic transformation.

Assuming that

$$b^2(x) > 0 \quad \forall x \in \mathbb{R}$$

and

$$\forall x \in \mathbb{R}\ \exists \varepsilon > 0 \text{ such that } \int\limits_{x-\varepsilon}^{x+\varepsilon} \frac{|a(y)|}{b^2(y)} dy < \infty,$$

equation (2.7) is an ordinary linear differential equation of order one for $s'(x)$ and is solved by

$$s'(x) = \exp\left(-\int\limits_c^x \frac{2a(z)}{b^2(z)} dz\right)$$

for a fixed number $c \in \mathbb{R}$. Thus the scale function is

$$s(x) = \int\limits_c^x \exp\left(-\int\limits_c^y \frac{2a(z)}{b^2(z)} dz\right) dy.$$

The speed measure M is a measure on $\mathscr{B}(I)$ with $0 < M((a,b)) < \infty$ for $l < a < b < r$. For $t > 0$ and $x \in I$, the probability measure $P_t(x,\cdot)$ on $\mathscr{B}(I)$ is absolutely continuous with respect to M, i.e.,

$$P_t(x, A) = \int_A p_M(t, x, y) M(dy)$$

where $p_M(t,x,y)$ is the transition density of X with respect to the speed measure. The speed density, i.e., $m(x)$ with

$$M(dx) = m(x)dx$$

is given by

$$m(x) = \frac{1}{b^2(x)s(x)}.$$

Furthermore we introduce the Green function.

Definition 2.2.13 *The Green function of X is given by*

$$G_\rho(x,y) = \int_0^\infty e^{-\rho t} p_M(t,x,y)dt,$$

where $p_M(t,x,y)$ is the transition density of X with respect to the speed measure.

The Green function can also be written as

$$G_\rho = \begin{cases} W_\rho^{-1}\psi_\rho(x)\varphi_\rho(y) & x \le y \\ W_\rho^{-1}\psi_\rho(y)\varphi_\rho(x) & x \ge y \end{cases}$$

(see (Itô and McKean, 1974, p. 150)), where W_ρ is the Wronskian of $\psi_\rho(x)$ and $\varphi_\rho(x)$, defined as

$$W_\rho(x) = \psi_\rho(x)\phi'_\rho(x) - \psi'_\rho(x)\varphi_\rho(x).$$

In this case the Wronskian is a positive constant. Then with Lemma 2.2.11 and (2.4) we get

Lemma 2.2.14 *$x \mapsto G_\rho(x,y)$ is a ρ-excessive function.*

Proof: For $x \le y \le z$ we have

$$\frac{G_\rho(x,y)}{G_\rho(z,y)} = \frac{W_\rho^{-1}\psi_\rho(x)\varphi_\rho(y)}{W_\rho^{-1}\psi_\rho(z)\varphi_\rho(y)} = \frac{\psi_\rho(x)}{\psi_\rho(z)},$$

while for $x \leq z \leq y$ we have

$$\begin{aligned}\frac{G_\rho(x,y)}{G_\rho(z,y)} &= \frac{W_\rho^{-1}\psi_\rho(x)\varphi_\rho(y)}{W_\rho^{-1}\psi_\rho(y)\varphi_\rho(z)} = \frac{\psi_\rho(x)\varphi_\rho(y)}{\psi_\rho(y)\varphi_\rho(z)} \\ &\geq \frac{\psi_\rho(x)}{\psi_\rho(y)} \geq \frac{\psi_\rho(x)}{\psi_\rho(z)}.\end{aligned}$$

The other cases follow analogously. □

According to the following lemma the functions $\psi_\rho(x)$, $\varphi_\rho(x)$ and $G_\rho(x,y)$ are minimal in the sense that an arbitrary non-trivial ρ-excessive function h can be represented as a linear combination of them, see (Borodin and Salminen, 1996, II.5.30).

Lemma 2.2.15 *Let $h(x)$ be a positive function on $I = [l,r]$ such that $h(x_0) = 1$. Then h is ρ-excessive if and only if there exists a probability measure μ on $[l,r]$ such that for all $x \in I$*

$$h(x) = \int\limits_{(l,r)} \frac{G_\rho(x,y)}{G_\rho(x_0,y)}\mu(dy) + \frac{\varphi_\rho(x)}{\varphi_\rho(x_0)}\mu(\{l\}) + \frac{\psi_\rho(x)}{\psi_\rho(x_0)}\mu(\{r\}).$$

The measure μ is called the representing measure of h.

The representing measure for the function

$$h(x) = q_1\psi_\rho(x) + q_2\varphi_\rho(x)$$

is given by

$$\mu((l,r)) = 0 \quad \mu(\{l\}) = q_1\psi_\rho(x_0) \quad \mu(\{r\}) = q_2\varphi_\rho(x_0).$$

This leads to the following condition for the parameters q_1 and q_2:

$$\frac{q_1}{q_2} = \frac{\varphi_\rho(x_0)}{\psi_\rho(x_0)}.$$

Now we can construct a new measure $\mathbb{P}^h$. For this let $h(x)$ be a ρ-excessive function for X. Then

$$\mathbb{P}^h_{x_0}(A) := e^{-\rho t}\mathbb{E}_{x_0}\left(\frac{h(X_t)}{h(x_0)}\mathbb{1}_A\right)$$

for $A \in \mathcal{F}_t$. The process X under the new measure $\mathbb{P}^h$ or shortly X^h is again a diffusion process and is called Doob's h-transform. The next lemma describes the above mentioned characteristics of an h-transform.

Lemma 2.2.16 *Let X be a diffusion process satisfying (2.6). Then the h-transform X^h is a diffusion process satisfying*

$$dX_t^h = a^h(X_t)dt + b^h(X_t)dW_t^h$$

where W_t^h is a Brownian motion under $\mathbb{P}^h$ with

$$a^h(x) = a(x) + b^2(x)\frac{h'(x)}{h(x)}$$

and

$$b^h(x) = b(x).$$

For the scale function we get

$$(s^h)'(x) = \frac{s'(x)}{h^2(x)}.$$

The density of the speed measure is given by

$$m^h(x) = h^2(x)m(x)$$

The transition density with respect to the speed measure is

$$p_M^h(t,x,y) = \frac{e^{-\rho t}}{h(x)h(y)} p_M(t,x,y).$$

Proof: see (Albanese and Kuznetsov, 2007) □

2.2.3 Removal of Drift

The second step of the stochastic transformation removes the drift from the process X_t^h. As we have already seen in the last subsection the scale function applied to the process leads to a local martingale, i.e., it has no drift component. According to Lemma 2.2.16 we have

$$(s^h)'(x) = \frac{s'(x)}{h^2(x)}.$$

Thus we first determine the derivative of $s(x)$. With $a(x) = \delta$ and $b^2(x) = 4x$ we get

$$\begin{aligned} s'(x) &= \exp\left(-2\int_c^x \frac{\delta}{4y}dy\right) = \exp\left(-\int_c^x \frac{\delta}{2y}dy\right) \\ &= \exp\left(-\frac{\delta}{2}\int_c^x \frac{1}{y}dy\right) = \exp\left(-\frac{\delta}{2}(\log y)_c^x\right) \\ &= \exp\left(-\frac{\delta}{2}(\log x - \log c)\right) \\ &= \exp\left(-\frac{\delta}{2}\log\left(\frac{x}{c}\right)\right) \\ &= \exp\left(\log\left(\left(\frac{x}{c}\right)^{-\frac{\delta}{2}}\right)\right) \\ &= \left(\frac{x}{c}\right)^{-\frac{\delta}{2}} = c_1 x^{-\frac{\delta}{2}} = c_1 x^{-\theta-1} \end{aligned}$$

where $\theta = \frac{\delta}{2} - 1$ is the index of the Squared Bessel process and $c_1 = c^{\frac{\delta}{2}} = c^{\theta+1}$. Since

$$h(x) = x^{-\frac{\theta}{2}}\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)$$

we have

$$\begin{aligned} (s^h)'(x) &= \frac{s'(x)}{h^2(x)} \\ &= \frac{c_1 x^{-\theta-1}}{x^{-\theta}\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2} \\ &= \frac{c_1}{x\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2}. \end{aligned}$$

Now we have to determine $s^h(x)$ as the primitive of $(s^h)'(x)$.

Lemma 2.2.17 *The primitive of*

$$(s^h)'(x) = \frac{c_1}{x(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x}))^2}$$

where $c_1 = c^{\theta+1}$, is given by

$$s^h(x) = \frac{2c_1 I_\theta(\sqrt{2\rho x})}{q_2(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x}))}.$$

Proof:

$$\begin{aligned}
(s^h)'(x) &= \frac{2c_1 \frac{\rho}{\sqrt{2\rho x}} I'_\theta(\sqrt{2\rho x})(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x}))}{q_2\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2} \\
&- \frac{2c_1 \frac{\rho}{\sqrt{2\rho x}} I_\theta(\sqrt{2\rho x})(q_1 I'_\theta(\sqrt{2\rho x}) + q_2 K'_\theta(\sqrt{2\rho x}))}{q_2\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2} \\
&= \frac{2c_1 \frac{\rho}{\sqrt{2\rho x}} \left(q_1 I_\theta(\sqrt{2\rho x}) I'_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x}) I'_\theta(\sqrt{2\rho x})\right)}{q_2\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2} \\
&+ \frac{2c_1 \frac{\rho}{\sqrt{2\rho x}} \left(-q_1 I_\theta(\sqrt{2\rho x}) I'_\theta(\sqrt{2\rho x}) - q_2 K'_\theta(\sqrt{2\rho x}) I_\theta(\sqrt{2\rho x})\right)}{q_2\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2} \\
&= \frac{2c_1 \frac{\rho}{\sqrt{2\rho x}} \left(K_\theta(\sqrt{2\rho x}) I'_\theta(\sqrt{2\rho x}) - K'_\theta(\sqrt{2\rho x}) I_\theta(\sqrt{2\rho x})\right)}{\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2}.
\end{aligned}$$

Here we can use the Wronskian relation, see (Abramowitz and Stegun, 1972, 9.6.14)

$$K_\theta(x) I'_\theta(x) - I_\theta(x) K'_\theta(x) = \frac{1}{x}$$

for any index θ.

Thus, we get

$$\begin{aligned}
(s^h)'(x) &= \frac{2c_1 \frac{\rho}{\sqrt{2\rho x}} \left(\frac{1}{\sqrt{2\rho x}}\right)}{\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2} \\
&= \frac{c_1}{x\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2}.
\end{aligned}$$

□

Then according to (2.8) the process

$$Y_t = s^h(X_t^h)$$

solves the stochastic differential equation

$$dY_t = \tilde{\sigma}(Y_t) d\tilde{W}_t$$

under $\mathbb{P}^h$ with

$$\begin{aligned}
\tilde{\sigma}(y) &= (s^h)'((s^h)^{-1}(y))b^h((s^h)^{-1}(y)) \\
&= (s^h)'(x)b^h(x) \\
&= 2\sqrt{x}\frac{c_1}{x\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2} \\
&= \frac{2c_1}{\sqrt{x}\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)^2}
\end{aligned}$$

where $s^h(x) = y$.

Since the process X_t^h is nonnegative and the scale function $s^h(x)$ is non-decreasing, the process Y_t is also nonnegative. Thus, the lower boundary 0 is absorbing. To determine the upper boundary $b_{\max}$, we can make use of the following approximations for large $x \gg \theta$, following (Abramowitz and Stegun, 1972, 9.7.1 and 9.7.2):

$$\begin{aligned}
I_\theta(x) &\sim \frac{1}{\sqrt{2\pi x}}e^x \\
K_\theta(x) &\sim \sqrt{\frac{\pi}{2x}}e^{-x}.
\end{aligned}$$

So for large x we have

$$\begin{aligned}
s^h(x) &= \frac{2c_1 I_\theta(\sqrt{2\rho x})}{q_2\left(q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})\right)} \\
&\approx \frac{2c_1\frac{e^{\sqrt{2\rho x}}}{\sqrt{2\pi\sqrt{2\rho x}}}}{q_2(q_1\frac{e^{\sqrt{2\rho x}}}{\sqrt{2\pi\sqrt{2\rho x}}} + q_2\sqrt{\frac{\pi}{2\sqrt{2\rho x}}}e^{-\sqrt{2\rho x}})} \\
&= \frac{2c_1 e^{\sqrt{2\rho x}}}{q_2\left(q_1 e^{\sqrt{2\rho x}} + q_2\pi e^{-\sqrt{2\rho x}}\right)} \\
&\underset{x\to\infty}{\longrightarrow} \frac{2c_1}{q_1 q_2}.
\end{aligned}$$

Thus, we get

$$b_{\max} = \frac{2c_1}{q_1 q_2} - s^h(c).$$

2.3 Default Probability

One of the main reasons for this two-stage transformation is the preservation of the analytical tractability. Especially the transition probability has to be expressed in terms of the well known transition probability of the Squared Bessel process. Thus, the next proposition will clarify the relation between the transition probability of the process X_t and the transition probability of the transformed process Y_t.

Proposition 2.3.1 *Let $p_X(t, x_0, x_1)$ be the transition probability density of the process X_t and let $p_Y(t, y_0, y_1)$ be the transition probability density of the process $Y_t = s^h(X_t^h)$. Then*

$$p_Y(t, y_0, y_1) = \frac{g(x_1, t)}{g(x_0, 0)} \cdot \frac{1}{(s^h)'(x_1)} p_X(t, x_0, x_1),$$

where $y_0 = s^h(x_0)$ and $y_1 = s^h(x_1)$.

Proof: Since $s^h(x)$ is a diffeomorphism we have

$$\begin{aligned} p_Y(t, y_0, y_1) dy_1 &= p_{X^h}(t, x_0, x_1) \frac{1}{(s^h)'(x_1)} dx_1 \\ &= \frac{1}{(s^h)'(x_1)} \frac{g(x_1, t)}{g(x_0, 0)} p_X(t, x_0, x_1) dx_1. \end{aligned}$$

□

With the following abbreviation

$$w(x) := q_1 I_\theta(\sqrt{2\rho x}) + q_2 K_\theta(\sqrt{2\rho x})$$

we get

$$\frac{g(x_1, t)}{g(x_0, s)} = e^{-\rho t} \left(\frac{x_0}{x_1} \right)^{\frac{\theta}{2}} \frac{w(x_1)}{w(x_0)}$$

and

$$(s^h)'(x) = \frac{c_1}{x w(x)^2}.$$

Thus, the transition probability density becomes

$$p_Y(t, y_0, y_1) = e^{-\rho t} \left(\frac{x_0}{x_1} \right)^{\frac{\theta}{2}} \frac{x_1 w(x_1)^3}{c_1 w(x_0)} p_X(t, x_0, x_1).$$

As we have already seen, the transition probability density for the Squared Bessel process is given as

$$p_X(t,x_0,x_1) = \frac{1}{2t}\left(\frac{x_1}{x_0}\right)^{\frac{\theta}{2}} \exp\left(-\frac{x_1+x_0}{2t}\right) I_\theta\left(\frac{\sqrt{xx_0}}{t}\right).$$

So, we finally get

$$p_Y(t,y_0,y_1) = \exp\left(-\rho t - \frac{x_1+x_0}{2t}\right)\frac{x_1}{2c_1 t}\frac{w(x_1)^3}{w(x_0)} I_\theta\left(\frac{\sqrt{xx_0}}{t}\right).$$

Since the absorbing state 0 has been proposed to be the default state, the probability of default by time t is given as

$$p^D(y_0^i,t) = 1 - \int_0^{b_{\max}} p_Y(t,y_0^i,y)dy$$

where y_0^i is the initial state. The index i indicates the rating class of the obligor at time 0.
If we differentiate p^D with respect to t, we get

$$\frac{\partial p^D(y_0^i,t)}{\partial t} = -\int_0^{b_{\max}} \frac{\partial p_Y(t,y_0^i,y)}{\partial t}dy. \tag{2.9}$$

As we have already seen, the process Y_t is driftless. Thus, the forward Kolmogorov equation for the transition probability density p_Y reduces the integrand to

$$\frac{\partial p_Y}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(\tilde{\sigma}^2(y)p_Y\right).$$

If we plug this into equation (2.9), we get

$$\begin{aligned}\frac{\partial p^D(y_0^i,t)}{\partial t} &= -\int_0^{b_{\max}} \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(\tilde{\sigma}^2(y)p_Y\right)dy \\ &= -\left(\frac{1}{2}\frac{\partial}{\partial y}\left(\tilde{\sigma}^2(y)p_Y\right)\Big|_{y=b_{\max}} - \frac{1}{2}\frac{\partial}{\partial y}\left(\tilde{\sigma}^2(y)p_Y\right)\Big|_{y=0}\right).\end{aligned}$$

The upper boundary $b_{\max}$ is unattainable, so the following boundary condition

$$p_Y(t,y_0,b_{\max}) = 0$$

leads to

$$\frac{1}{2}\frac{\partial}{\partial y}\left(\tilde{\sigma}^2(y)p_Y\right)\Big|_{y=b_{\max}} = 0.$$

Thus we get

$$\frac{dp_i^D(t)}{dt} = \frac{1}{2}\frac{\partial}{\partial y}\left(\tilde{\sigma}^2(y)p_Y\right)\Big|_{y=0}.$$

Using an integral representation of the modified Bessel function K_θ and a series expansion for I_θ, the derivative of the default probability is equal to

$$\frac{dp_i^D(t)}{dt} = \frac{q_2 e^{-\rho t - \frac{x_0}{2t}}}{2w(x_0)}\left(\frac{x_0}{2\rho}\right)^{\frac{\theta}{2}}.$$

Integrating this expression leads to

$$p_i^D(t) = \frac{q_2}{2w(x_0)}\left(\frac{x_0}{2\rho}\right)^{\frac{\theta}{2}} \int_0^t \frac{e^{-\rho s - \frac{x_0}{2s}}}{s^{1+\theta}} ds.$$

2.4 Stochastic Time Change

Since the calibration to historical migration rates was unsuccessful in this diffusion model, especially for rating migrations of several rating classes, Albanese et al. include jumps in their model. These jumps are introduced by a stochastic time change. Thus, the process is obtained by evaluating the process Y at a random time given by a gamma process, similar to the Variance Gamma model, as introduced, e.g., in (Madan et al., 1998). The gamma process $\gamma(t, 1, \beta)$ with mean 1 and variance rate β is a process with independent gamma distributed increments. Thus, the transition probability density $\tilde{\Gamma}(s, t)$ is given by a gamma distribution with shape parameter $\frac{t}{\beta}$ and scale parameter β:

$$\tilde{\Gamma}(s, t) = \frac{s^{\frac{t}{\beta}-1} e^{-\frac{s}{\beta}}}{\Gamma\left(\frac{t}{\beta}\right)\beta^{\frac{t}{\beta}}}$$

where Γ is the gamma function.

Thus the rating transition probability under the stochastic time change is given by

$$\begin{aligned}\tilde{p}_{ij}(t) &= \int\limits_{b_{j-1}}^{b_j} p_{\tilde{Y}}(t, y_0, y)dy \\ &= \int\limits_{b_{j-1}}^{b_j} \int\limits_{0}^{\infty} p_Y(s, y_0, y)\tilde{\Gamma}(s, t)dsdy.\end{aligned}$$

Equivalently, we get for the default probability

$$\begin{aligned}\tilde{p}_i^D(t) &= 1 - \int\limits_{0}^{b_{\max}} p_{\tilde{Y}}(t, y_0, y)dy \\ &= \frac{q_2}{2w(x_0)} \left(\frac{x_0}{2\rho}\right)^{\frac{\theta}{2}} \int\limits_{0}^{\infty} \tilde{\Gamma}(z, t) \int\limits_{0}^{z} \frac{e^{-\rho s - \frac{x_0}{2s}}}{s^{1+\theta}} dsdz.\end{aligned}$$

Chapter 3

A Structural Squared Bessel Model

The article (Zhou, 1997) was the first to present an analytical formula for the joint default probability of two firms in a first-passage time model. The two firm values follow geometric Brownian motions where the dependence is put in the underlying Brownian motions. Another possibility to calculate default correlations based on a first-passage time model is to adopt the concept of factor models such as Creditmetrics, see (JPMorgan, 1997). Similarly, in this chapter, we introduce and discuss a time continuous structural model where the ability to pay process of every firm is composed of one individual process part and one common process part. These two processes are assumed to be independent of each other. Since this process is not necessarily directly connected to the asset value of the firm, but rather represents an abstract indicator of the ability of the firm to fulfill its obligations, we denote this process as ability to pay process. Thus the ability to pay process of firm $i \in I = \{1, 2\}$ turns out to be

$$X_t^i = Y_t + Y_t^i \quad \forall t \in [0, \infty)$$

where $(Y_t)_{t \geq 0}$ and $(Y_t^i)_{t \geq 0}$ are independent of each other. In order to study the properties of $(X_t^i)_{t \geq 0}$ we have to be able to determine the process resulting from the addition of two independent processes. Thus we require the processes to have the so-called additivity property. A similar concept, though not designed as ability to pay processes, has already been applied in reduced form models, see (Duffie and Singleton, 2003). There, the default probability is given as

$$\mathbb{P}(\tau \leq t) = 1 - \mathbb{E}\left(\exp\left(-\int_0^t \lambda(u)du\right)\right).$$

The intensity process $\lambda(t)$ is modelled as a stochastic process with state space $\mathbb{R}^+$ satisfying

$$d\lambda(t) = \kappa(\theta - \lambda(t))dt + \sigma\sqrt{\lambda(t)}dW(t) + \Delta J(t).$$

A process $\lambda(t)$ satisfying this stochastic differential equation is called a basic affine process with parameters $(\kappa, \theta, \sigma, \mu, l)$, where μ is the mean of the exponentially distributed jump sizes and l is the mean jump arrival rate of the pure jump process J. In order to introduce dependencies between different obligors, Duffie proposed to describe the intensity process λ_i of obligor i as the sum of two basic affine processes, i.e.,

$$\lambda_i = X_c + X_i.$$

Here the parameters of X_c are $(\kappa, \theta_c, \sigma, \mu, l_c)$ and X_i has the parameters $(\kappa, \theta_i, \sigma, \mu, l_i)$. Then λ_i is a basic affine process with parameters $(\kappa, \theta, \sigma, \mu, l)$ where $\theta = \theta_c + \theta_i$ and $l = l_c + l_i$. In our first passage time model we have to consider the special case $l = 0$, i.e., with no jumps at all. This leads to the Cox-Ingersoll-Ross process

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t$$

or in a more general form

$$dX_t = (\delta + bX_t)dt + c\sqrt{X_t}dB_t$$

where $\delta \geq 0,\ b \in \mathbb{R},\ c > 0$ and $(B_t)_{t\geq 0}$ a standard Brownian motion. In this chapter we will concentrate on the case $b = 0,\ c = 2$, which leads to the Squared Bessel process. Some of the following definitions and properties can be found in (Göing-Jaeschke and Yor, 2003).

3.1 Squared Bessel Processes

In this section we will highlight some of the main properties of the Squared Bessel Process (see (Revuz and Yor, 1991)).

Definition 3.1.1 *A process $(X_t)_{t\geq 0}$ satisfying the equation*

$$X_t = x + \delta t + 2\int_0^t \sqrt{X_t}dB_t$$

with $\delta \geq 0$ is called a δ-dimensional Squared Bessel process with start in $x \geq 0$ and is denoted by $BESQ^\delta(x)$.

The next Theorem makes clear, why the parameter δ is called the dimension of the Squared Bessel process.

Theorem 3.1.2 *Let $(B_t)_{t\geq 0}$ be a δ-dimensional Brownian motion, where $\delta \in \mathbb{N}$ with start in $B_0 \in \mathbb{R}^\delta$ and*

$$X_t = ||B_t||^2$$

where $||\cdot||$ denotes the Euclidian norm in $\mathbb{R}^\delta$.
Then the stochastic process $(X_t)_{t\geq 0}$ with start in $x = ||B_0||^2$ satisfies the following equation

$$X_t = x + \delta t + 2\int_0^t \sqrt{X_t}d\beta_t$$

where $(\beta_t)_{t\geq 0}$ is a standard Brownian motion. According to this the δ-dimensional Squared Bessel process equals the square of the norm of a δ-dimensional Brownian motion if $\delta \in \mathbb{N}$.

Proof: We have

$$||B_t||^2 = (B_t^1)^2 + \ldots + (B_t^\delta)^2.$$

Applying the δ-dimensional Itô-formula to $f(x) = ||x||^2 = \sum_{i=1}^\delta x_i^2$ leads to

$$f(B_t) = f(B_0) + \sum_{i=1}^{\delta}\int_0^t \left.\frac{\partial f}{\partial x_i}\right|_{x=B_s} dB_s^i + \frac{1}{2}\sum_{i,j=1}^{\delta}\int_0^t \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{x=B_s} d\langle B_i, B_j\rangle_s.$$

Since

$$\langle B_i, B_j \rangle_t = \delta_{ij} t$$

where δ_{ij} denotes the Kronecker-Delta, we get

$$f(B_t) = f(B_0) + \sum_{i=1}^{\delta} \int_0^t \left.\frac{\partial f}{\partial x_i}\right|_{x=B_s} dB_s^i + \frac{1}{2} \sum_{i=1}^{\delta} \int_0^t \left.\frac{\partial^2 f}{\partial x_i^2}\right|_{x=B_s} ds.$$

The partial derivatives are

$$\begin{aligned} \frac{\partial f}{\partial x_i} &= 2x_i \\ \frac{\partial^2 f}{\partial x_i^2} &= 2 \end{aligned}$$

for $i = 1, ..., \delta$. Hence

$$\begin{aligned} f(B_t) &= f(B_0) + 2 \sum_{i=1}^{\delta} \int_0^t B_s^i dB_s^i + \frac{1}{2} \sum_{i=1}^{\delta} \int_0^t 2ds \\ &= f(B_0) + 2 \sum_{i=1}^{\delta} \int_0^t B_s^i dB_s^i + \delta \cdot t. \end{aligned}$$

Furthermore

$$\sum_{i=1}^{\delta} \int_0^t B_s^i dB_s^i = \int_0^t \sum_{i=1}^{\delta} B_s^i dB_s^i$$

$$= \int_0^t \rho_s \sum_{i=1}^{\delta} \frac{B_s^i}{\rho_s} dB_s^i = \int_0^t \rho_s d\beta_s$$

where $\rho_t = ||B_t|| = \sqrt{f(B_t)}$ and

$$\beta_t = \sum_{i=1}^{\delta} \int_0^t \frac{B_s^i}{\rho_s} dB_s^i.$$

Now we show that $\langle\beta\rangle_t = t$ so that β_t is a Brownian motion. In fact,

$$\begin{aligned}
\langle\beta\rangle_t &= \langle\sum_{i=1}^{\delta}\int_0^{\cdot}\frac{B_s^i}{\rho_s}dB_s^i\rangle_t \\
&= \sum_{i,j=1}^{\delta}\langle\int_0^{\cdot}\frac{B_s^i}{\rho_s}dB_s^i,\int_0^{\cdot}\frac{B_s^j}{\rho_s}dB_s^j\rangle_t = \sum_{i=1}^{\delta}\langle\int_0^{\cdot}\frac{B_s^i}{\rho_s}dB_s^i\rangle_t \\
&= \sum_{i=1}^{\delta}\int_0^{\cdot}\left(\frac{B_s^i}{\rho_s}\right)^2 d\langle B\rangle_s = \sum_{i=1}^{\delta}\int_0^{t}\left(\frac{B_s^i}{\rho_s}\right)^2 ds \\
&= \int_0^t \frac{\sum_{i=1}^{\delta}(B_s^i)^2}{\rho_s^2}ds = \int_0^t 1ds = t.
\end{aligned}$$

With $X_t = f(B_t)$ we get

$$X_t = x + \delta t + 2\int_0^t \sqrt{X_t}d\beta_t.$$

□

The next Theorem will prove the additivity property for the Squared Bessel process, i.e., the sum of two independent squared Bessel processes Y_t^1 and Y_t^2 with start in y_1 and y_2 and dimensions δ_1 and δ_2, respectively, is a Squared Bessel process itself with start in $y_1 + y_2$ and dimension $\delta_1 + \delta_2$. The law of a $BESQ^\delta(x)$ on $C(\mathbb{R}^+, \mathbb{R})$ is denoted by Q_x^δ.

Definition 3.1.3 *Let P and Q be two probability measures, then $P \star Q$ is defined to be the convolution of P and Q.*

Then the following result holds:

Theorem 3.1.4 *For every $\delta_1, \delta_2 \geq 0$ and $y_1, y_2 \geq 0$,*

$$Q_{y_1}^{\delta_1} \star Q_{y_2}^{\delta_2} = Q_{y_1+y_2}^{\delta_1+\delta_2}.$$

Proof: We set $X_t = Y_t^1 + Y_t^2$, where Y^1 is a $BESQ^{\delta_1}(y_1)$ and Y^2 is a $BESQ^{\delta_2}(y_2)$ independent of Y^1. Now we have to show, that X is a

$BESQ^{\delta_1+\delta_2}(y_1+y_2)$.

$$\begin{aligned} X_t &= Y_t^1 + Y_t^2 \\ &= y_1 + y_2 + 2\int_0^t (\sqrt{Y_s^1}dB_s^1 + \sqrt{Y_s^2}dB_s^2) + (\delta_1+\delta_2)t. \end{aligned}$$

Now we can write

$$\begin{aligned} & \sqrt{Y_s^1}dB_s^1 + \sqrt{Y_s^2}dB_s^2 \\ = & \sqrt{X_s}\mathbb{1}_{\{X_s>0\}}\frac{\sqrt{Y_s^1}dB_s^1 + \sqrt{Y_s^2}dB_s^2}{\sqrt{X_s}} \\ = & \sqrt{X_s}d\gamma_s \end{aligned}$$

where

$$\gamma_t = \int_0^t \mathbb{1}_{\{X_s>0\}}\frac{\sqrt{Y_s^1}dB_s^1 + \sqrt{Y_s^2}dB_s^2}{\sqrt{X_s}}.$$

We can conclude

$$X_t = y_1 + y_2 + 2\int_0^t \sqrt{X_s}d\gamma_s + (\delta_1+\delta_2)t.$$

Now we have to show that γ_t is a Brownian motion. For this we again show, that its quadratic variation equals t:

$$\begin{aligned} \langle\gamma\rangle_t &= \langle\int_0^{\cdot} \mathbb{1}_{\{X_s>0\}}\frac{\sqrt{Y_s^1}dB_s^1 + \sqrt{Y_s^2}dB_s^2}{\sqrt{X_s}}\rangle_t \\ &= \langle\int_0^{\cdot} \mathbb{1}_{\{X_s>0\}}\frac{\sqrt{Y_s^1}}{\sqrt{X_s}}dB_s^1\rangle_t + \langle\int_0^{\cdot} \mathbb{1}_{\{X_s>0\}}\frac{\sqrt{Y_s^2}}{\sqrt{X_s}}dB_s^2\rangle_t \\ &= \int_0^t \mathbb{1}_{\{X_s>0\}}\frac{Y_s^1}{X_s}ds + \int_0^t \mathbb{1}_{\{X_s>0\}}\frac{Y_s^2}{X_s}ds \\ &= \int_0^t \mathbb{1}_{\{X_s>0\}}\frac{Y_s^1+Y_s^2}{X_s}ds = \int_0^t \mathbb{1}_{\{X_s>0\}}ds \\ &= t \end{aligned}$$

and γ is a Brownian motion. □

Since we are going to identify the credit barrier with zero, it is important to study the behavior of the paths of Squared Bessel process dependent on its dimension δ:

(i) for $\delta = 0$ the point 0 is absorbing (i.e., after it reaches 0, the process will stay there forever);

(ii) for $0 < \delta < 2$ the point 0 is instantaneously reflecting (i.e., after reaching 0, the process will immediately move away from 0);

(iii) for $0 < \delta < 2$ the point 0 is reached a.s.

(iv) for $\delta \geq 2$ the point 0 is unattainable.

The term Squared Bessel process leads to the following definition:

Definition 3.1.5 *The square root of a $BESQ^\delta(x)$ is called δ-dimensional Bessel process starting at $\sqrt{x}$ and is denoted by $BES^\delta(\sqrt{x})$.*

In the following we will set $a := \sqrt{x}$ and we write $BES^\delta(a)$ instead of $BES^\delta(\sqrt{x})$. Now we can determine the SDE satisfied by a $BES^\delta(a)$:

Lemma 3.1.6 *The process $(R_t)_{t\geq 0}$ is a $BES^\delta(a)$ if it is a solution of the SDE*

$$R_t = a + B_t + \frac{\delta - 1}{2} \int_0^t \frac{1}{R_s} ds.$$

Proof: We have to apply the Itô-formula

$$f(X_t) = f(X_0) + \int_0^t \frac{\partial f}{\partial x} dX_s + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2} d\langle X \rangle_s$$

to the $BESQ^\delta(x)$ X with $x = a^2$ and $f(x) = \sqrt{x}$. We have

$$\frac{\partial f}{\partial x} = \frac{1}{2\sqrt{x}}\ , \ \frac{\partial^2 f}{\partial x^2} = -\frac{1}{4x\sqrt{x}}$$

and

$$dX_t = \delta dt + 2\sqrt{X_t} dB_t \ , \ d\langle X \rangle_t = 4X_t dt.$$

This leads to

$$\begin{aligned} R_t = \sqrt{X_t} &= a + \int_0^t \frac{1}{2\sqrt{X_s}} (\delta dt + 2\sqrt{X_s} dB_s) - \frac{1}{2} \int_0^t \frac{1}{4X_s\sqrt{X_s}} 4X_s ds \\ &= a + \int_0^t 1 dB_s + \int_0^t \frac{\delta - 1}{2\sqrt{X_s}} ds \\ &= a + B_t + \frac{\delta - 1}{2} \int_0^t \frac{1}{R_s} ds. \end{aligned}$$

□

3.1.1 Transition Densities

Since the Squared Bessel process is a Markov process, its transition densities are known explicitly. For $\delta > 0$, the transition density is given as

$$p(t, x, y) = \frac{1}{2t} \left(\frac{y}{x} \right)^{\frac{\nu}{2}} \exp\left(-\frac{x + y}{2t} \right) I_\nu \left(\frac{\sqrt{xy}}{t} \right). \tag{3.1}$$

There $t, x > 0$ and $\nu = \frac{\delta}{2} - 1$ is the index of the Squared Bessel process. Moreover, I_ν is the modified Bessel function of the first kind with index ν and is defined as

$$I_\nu(z) = \sum_{n=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{\nu+2n}}{n!\Gamma(n + \nu + 1)}, \quad z \in \mathbb{C}.$$

For $x = 0$ and $y > 0$ the transition density becomes

$$p(t, 0, y) = \frac{y^{\frac{\delta}{2}-1}}{(2t)^{\frac{\delta}{2}} \Gamma\left(\frac{\delta}{2}\right)} e^{-\frac{y}{2t}}.$$

From these transition densities, one can obtain the transition densities for the BES^δ by a change of variables, i.e., for $a = \sqrt{x}$ and $b = \sqrt{y}$ we get

$$\begin{aligned} &\frac{1}{2t}\left(\frac{b^2}{a^2}\right)^{\frac{\nu}{2}} \exp\left(-\frac{a^2+b^2}{2t}\right) I_\nu\left(\frac{ab}{t}\right) db^2 \\ = \; &\frac{b}{t}\left(\frac{b}{a}\right)^{\nu} \exp\left(-\frac{a^2+b^2}{2t}\right) I_\nu\left(\frac{ab}{t}\right) db \end{aligned}$$

and thus

$$p(t,a,b) = \frac{b}{t}\left(\frac{b}{a}\right)^{\nu} \exp\left(-\frac{a^2+b^2}{2t}\right) I_\nu\left(\frac{ab}{t}\right).$$

Similarly, we get for $a = 0$

$$p(t,0,b) = \frac{y^{2\hat{\nu}+1}}{2^{\hat{\nu}} t^{\hat{\nu}+1} \Gamma(\hat{\nu}+1)} \exp\left(-\frac{y^2}{2t}\right). \tag{3.2}$$

3.2 Simulation of Squared Bessel Processes

In this section we want to present some methods to simulate a Squared Bessel process on a given interval $[0, T]$, $T > 0$. For this we divide the interval into $n + 1$ equidistant grid points

$$\{t_i = \frac{iT}{n} | i = 0, ..., n\}$$

and generate the vector $(\hat{X}_{t_0}, ..., \hat{X}_{t_n})$, where $\hat{X}_{t_0} = x$ is the starting point of the $BESQ^\delta(x)$. Although the Squared Bessel process can be simulated exactly, as seen at the end of this section, sometimes it is easier or at least faster to use an approximation scheme. The next two discretization schemes can be found for Cox-Ingersoll-Ross processes in (Alfonsi, 2005).

3.2.1 Euler-Maruyama Scheme

The easiest way to simulate a process is the explicit Euler-Maruyama scheme. This scheme extends the Euler-scheme in not just discretizing the time, but also the Brownian motion. This leads to

$$\hat{X}_{t_{i+1}} = \hat{X}_{t_i} + \delta\frac{T}{n} + 2\sqrt{\hat{X}_{t_i}}(W_{t_{i+1}} - W_{t_i})$$

and thus

$$\hat{X}_{t_{i+1}} = \hat{X}_{t_i} + \delta\frac{T}{n} + 2\sqrt{\hat{X}_{t_i}}\xi_i$$

where $\{\xi_i\}_{i=0,\dots,n}$ are i.i.d. $N(0,\frac{T}{n})$-distributed and $\hat{X}_{t_0} = x$. Since the Gaussian increments are not bounded from below, this scheme can lead to negative values. In order to be able to compute $\hat{X}_{t_{i+1}}$ if $\hat{X}_{t_i}$ is negative one can instead use

$$\hat{X}_{t_{i+1}} = \hat{X}_{t_i} + \delta\frac{T}{n} + 2\sqrt{\hat{X}_{t_i}\mathbb{1}_{\{\hat{X}_{t_i}>0\}}}(W_{t_{i+1}} - W_{t_i}).$$

Negative values can be totally avoided by using

$$\hat{X}_{t_{i+1}} = \left|\hat{X}_{t_i} + \delta\frac{T}{n} + 2\sqrt{\hat{X}_{t_i}}(W_{t_{i+1}} - W_{t_i})\right|.$$

3.2.2 An implicit Scheme

The positivity can also be obtained by using an implicit scheme. To derive this implicit scheme we first have to rewrite the Squared Bessel process.
Since

$$\int_0^t \delta ds = \lim_{n\to\infty} \sum_{i=1,\ t_i<t}^{n} \delta\frac{T}{n}$$

and

$$\int_0^t \sqrt{X_s}dW_s = \lim_{n\to\infty} \sum_{i=1,\ t_i<t}^{n} \sqrt{X_{t_i}}(W_{t_{i+1}} - W_{t_i})$$

we have

$$\begin{aligned}
X_t &= x + \int_0^t \delta ds + 2\int_0^t \sqrt{X_s}dW_s \\
&= x + \lim_{n\to\infty}\left\{\sum_{i=1,\ t_i<t}^{n} \delta\frac{T}{n} + 2\sum_{i=1,\ t_i<t}^{n} \sqrt{X_{t_i}}(W_{t_{i+1}} - W_{t_i})\right\} \\
&= x + \lim_{n\to\infty}\left\{\sum_{i:t_i<t} \delta\frac{T}{n} + 2\sum_{i:t_i<t} \sqrt{X_{t_{i+1}}}(W_{t_{i+1}} - W_{t_i})\right. \\
&\quad \left. -2\sum_{i:t_i<t}(\sqrt{X_{t_{i+1}}} - \sqrt{X_{t_i}})(W_{t_{i+1}} - W_{t_i})\right\}.
\end{aligned}$$

Now

$$\begin{aligned}\lim_{n\to\infty}\sum_{i:t_i<t}(\sqrt{X_{t_{i+1}}}-\sqrt{X_{t_i}})(W_{t_{i+1}}-W_{t_i})\\ =\langle\sqrt{X},W\rangle_t\end{aligned}$$

is the quadratic covariation of the Bessel process $(\sqrt{X_t})_{t\geq 0}$ and the underlying Brownian motion $(W_t)_{t\geq 0}$. Since the Bessel process is a semi-martingale (at least for $\delta > 1$) with the Brownian motion as the martingale-part, the quadratic covariation is

$$\langle\sqrt{X},W\rangle_t=\langle W\rangle_t=t.$$

Thus we get

$$\begin{aligned}X_t &= x+\delta t-2t+2\lim_{n\to\infty}\left\{\sum_{i:t_i<t}\sqrt{X_{t_{i+1}}}(W_{t_{i+1}}-W_{t_i})\right\}\\ &= x+(\delta-2)t+2\lim_{n\to\infty}\left\{\sum_{i:t_i<t}\sqrt{X_{t_{i+1}}}(W_{t_{i+1}}-W_{t_i})\right\}.\end{aligned}$$

Then for sufficiently large n we can consider the following implicit scheme:

$$\hat{X}_{t_{i+1}}=\hat{X}_{t_i}+(\delta-2)\frac{T}{n}+2\sqrt{\hat{X}_{t_{i+1}}}(W_{t_{i+1}}-W_{t_i}).$$

Since $\hat{X}_{t_{i+1}}$ appears on both sides of the equation, we set $\hat{x}=\sqrt{\hat{X}_{t_{i+1}}}$ to get a quadratic equation in $\hat{x}$:

$$\hat{x}^2-2(W_{t_{i+1}}-W_{t_i})\hat{x}-\hat{X}_{t_i}-(\delta-2)\frac{T}{n}=0.$$

This equation has two solutions:

$$\hat{x}_{1,2}=(W_{t_{i+1}}-W_{t_i})\pm\sqrt{(W_{t_{i+1}}-W_{t_i})^2+\hat{X}_{t_i}+(\delta-2)\frac{T}{n}}.$$

But only for $\delta > 2$ one can guarantee that there is only one positive solution:

$$\hat{x}_1=(W_{t_{i+1}}-W_{t_i})+\sqrt{(W_{t_{i+1}}-W_{t_i})^2+\hat{X}_{t_i}+(\delta-2)\frac{T}{n}}.$$

Then we have

$$\hat{X}_{t_{i+1}}=\left((W_{t_{i+1}}-W_{t_i})+\sqrt{(W_{t_{i+1}}-W_{t_i})^2+\hat{X}_{t_i}+(\delta-2)\frac{T}{n}}\right)^2.$$

3.2.3 Exact Simulation

The last method we introduce and we will actually use to simulate the Squared Bessel process is to determine the increments of the process. Therefore we generate the vector

$$(\hat{X}_{t_0}, ..., \hat{X}_{t_n})$$

for a given initial value $\hat{X}_{t_0} = \hat{X}_0 = x$ by means of the increments

$$\hat{X}_{t_{i+1}} = \hat{X}_{t_i} + (\hat{X}_{t_{i+1}} - \hat{X}_{t_i}).$$

As already mentioned the Squared Bessel process is a Markov process, so we have for $A \in \sigma(X_u, u \leq t)$

$$P(X_t \in A | \sigma(X_u, u \leq s))$$

$$= P(X_t \in A | X_s), \quad s \leq t$$

and thus

$$P(X_{t_{i+1}} \in A | X_{t_i} = x) = \int_A p(\frac{T}{n}, x, y) dy$$

where $p(\frac{T}{n}, x, y)$ is the transition density of the Squared Bessel process, introduced in 3.1. If we substitute now $x = bt$ and $y = at$ in the transition density $p(t, x, y)$, we get

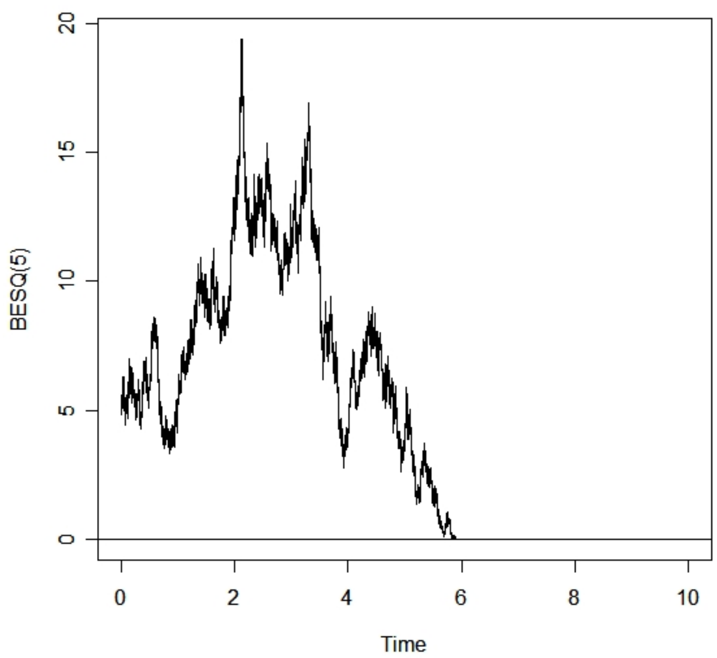
$$p(t, x, y) = p_b(a) = \frac{1}{2} \left(\frac{a}{b}\right)^{\frac{\nu}{2}} \exp\left(-\frac{a+b}{2}\right) I_\nu\left(\sqrt{ab}\right).$$

But this is exactly the probability density of a noncentral χ^2-distribution with δ degrees of freedom and noncentrality-parameter b. Thus we can draw $\hat{X}_{t_{i+1}}$ as a random number from a conditional noncentral χ^2-distribution with noncentrality-parameter $\hat{X}_{t_i}$, i.e.,

$$\hat{X}_{t_{i+1}} \sim \chi^2_{\delta, \frac{\hat{X}_{t_i} n}{T}}.$$

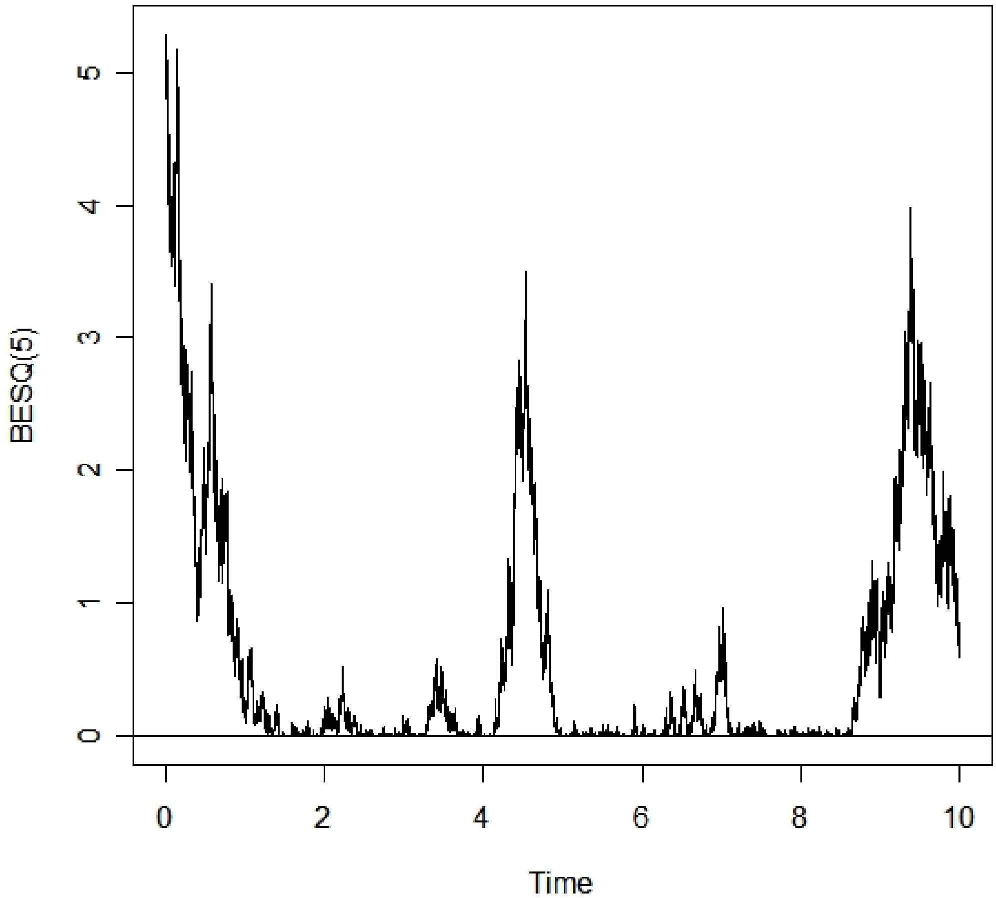
3.2.4 Some Paths of Squared Bessel Processes

In this subsection we will present some figures with paths of Squared Bessel processes. For this we use the exact simulation introduced in the last subsection. The figures will show some sample paths to illustrate the behavior of the Squared Bessel process in dependence of the dimension δ as itemized in section 3.1. As initial value we choose $x = 5$, while the dimensions are $\delta_1 = 0, \delta_2 = 0.5, \delta_3 = 1, \delta_4 = 1.5, \delta_5 = 2$ and $\delta_6 = 3$.



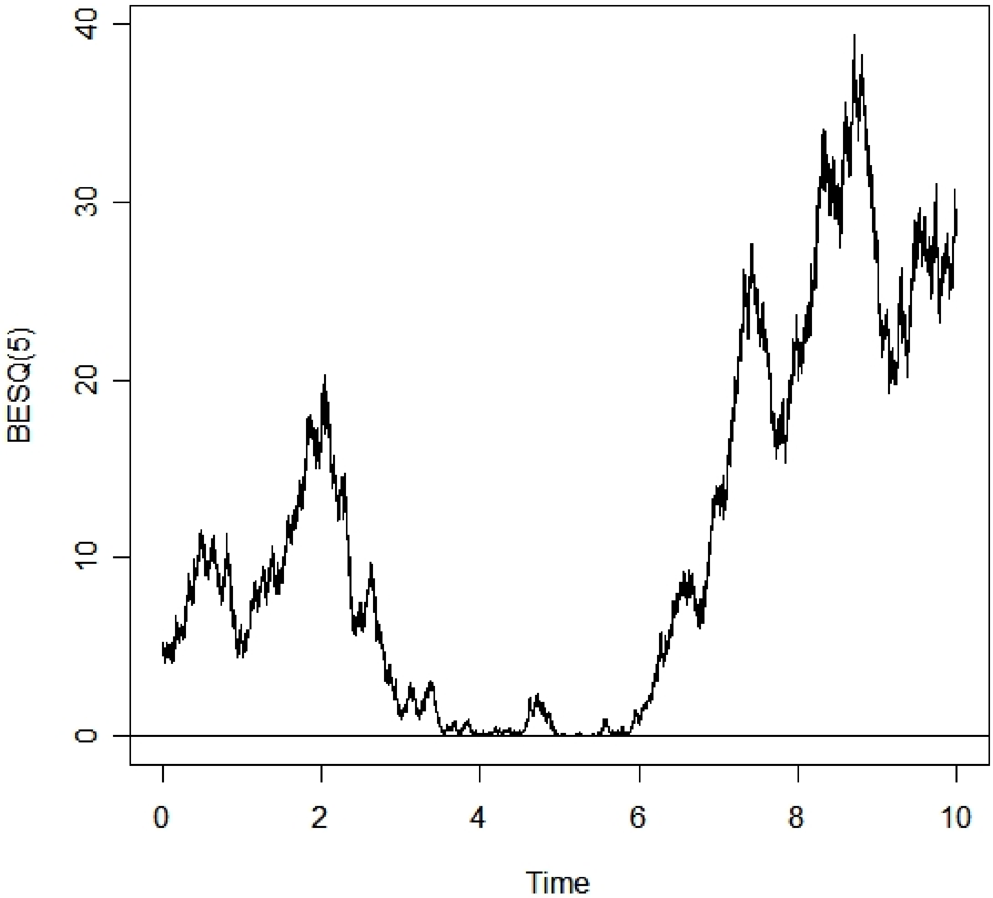
(a) Squared Bessel process with $\delta = 0$

Figure 3.1: This figure illustrates, that a $BESQ^0(x)$ dies as soon as it reaches zero, i.e., it stays there without return.



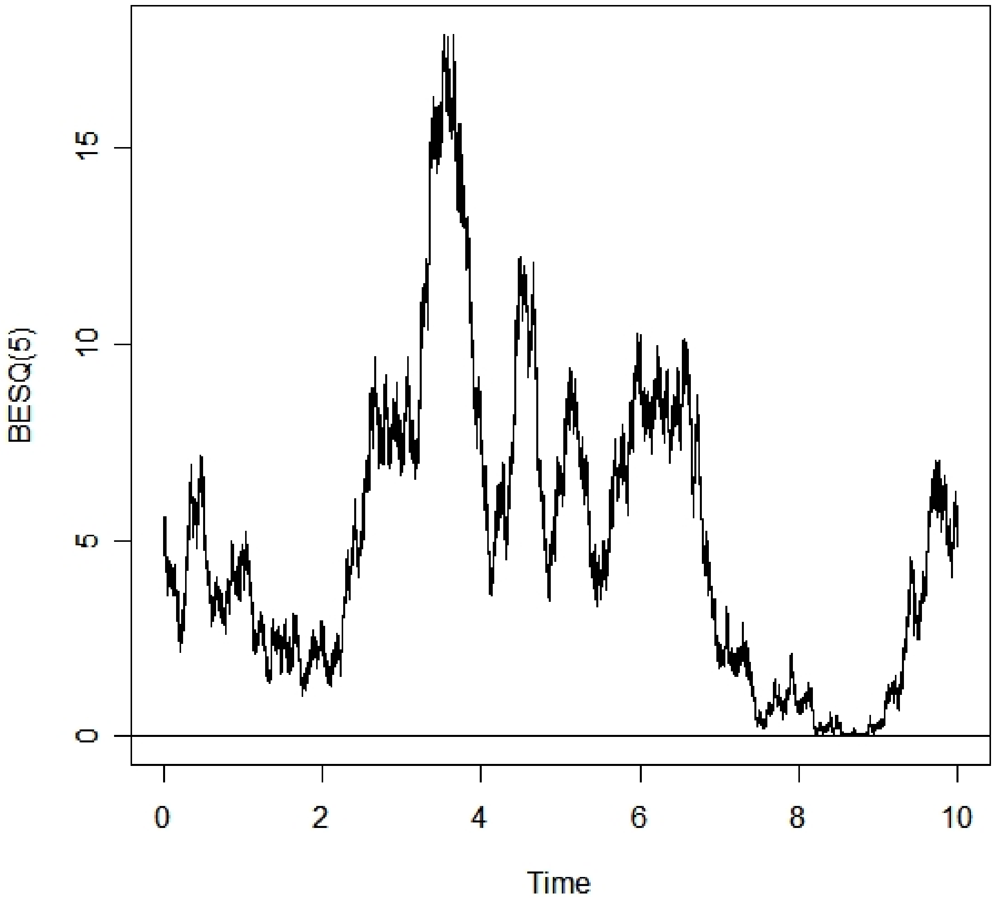
(a) Squared Bessel process with $\delta = 0.5$

Figure 3.2: This figure illustrates, that a $BESQ^{0.5}(x)$ hits 0 very often and is reflected immediately.



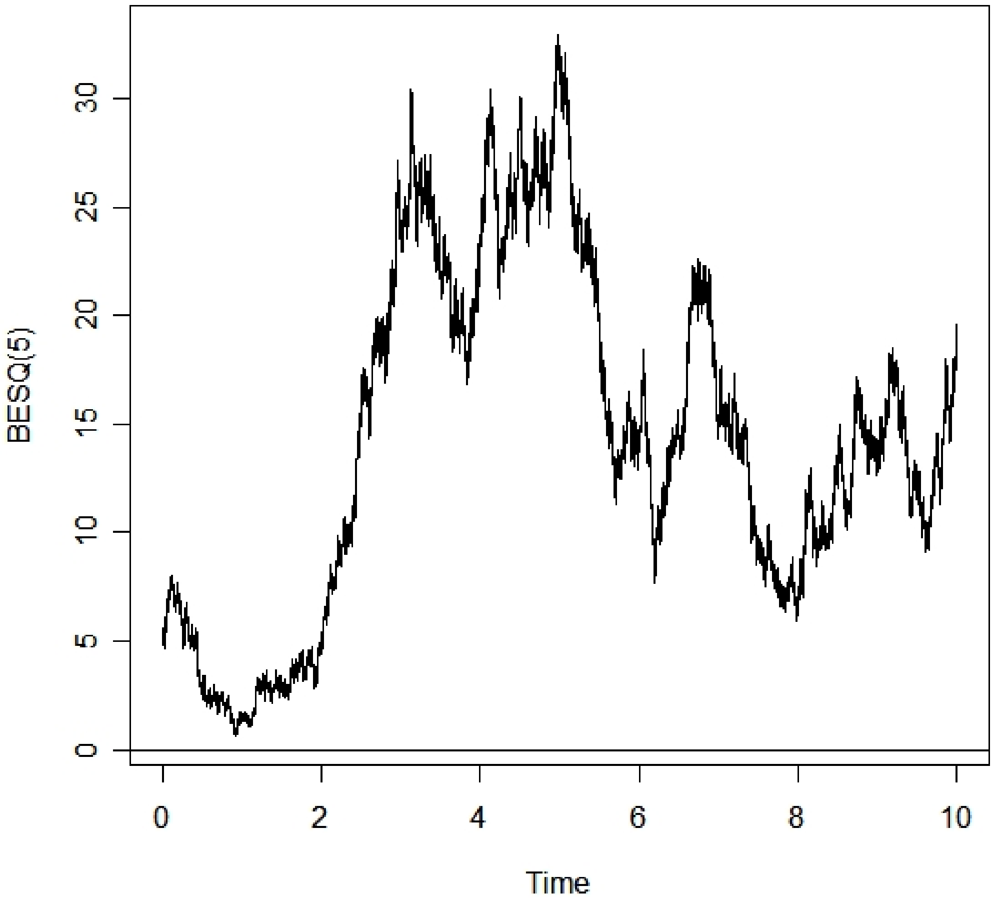
(a) Squared Bessel process with $\delta = 1$

Figure 3.3: This process equals the square of a one-dimensional Brownian motion with start in $\sqrt{5}$.



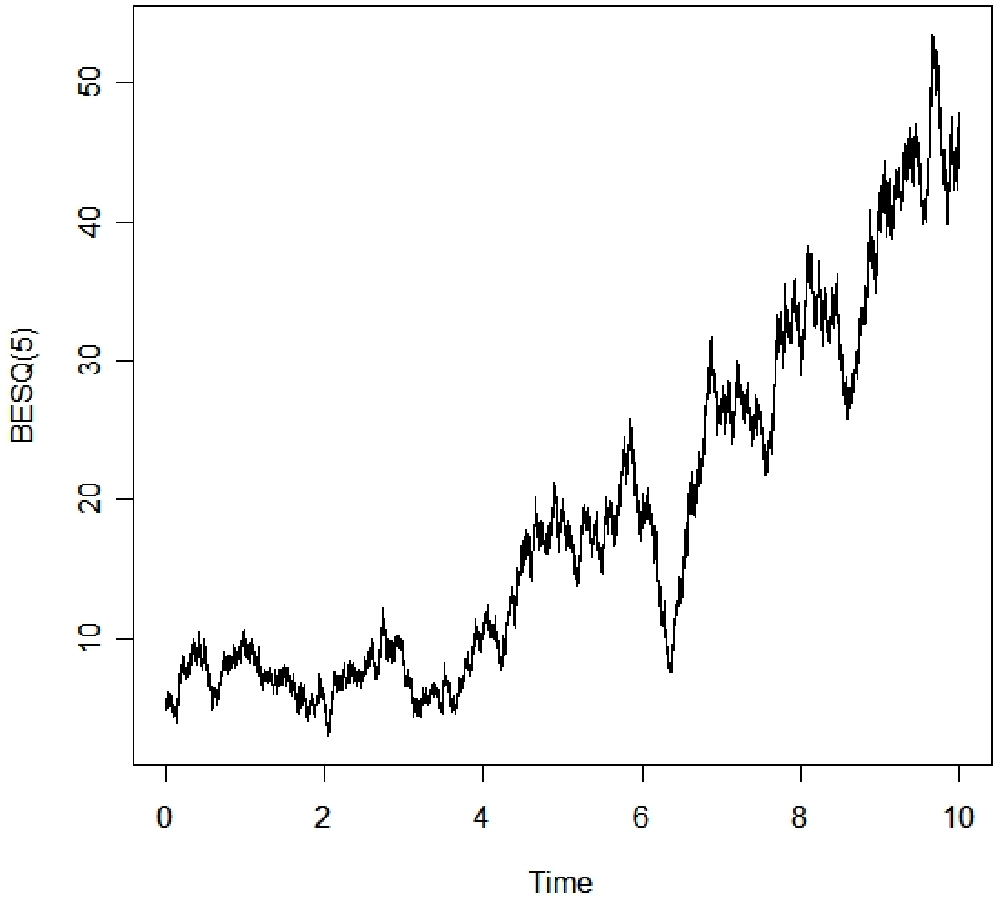
(a) Squared Bessel process with $\delta = 1.5$

Figure 3.4: This figure illustrates, that a $BESQ^{\delta}(x)$ with $1 < \delta < 2$ still reaches zero with positive probability.



(a) Squared Bessel process with $\delta = 2$

Figure 3.5: This figure shall illustrate, that a $BESQ^2(x)$ never reaches zero.



(a) Squared Bessel process with $\delta = 3$

Figure 3.6: This figure shall illustrate, that a $BESQ^3(x)$ is a transient process.

3.3 First Hitting Time of Squared Bessel Processes

In this section we will determine the probability density of the first time a Squared Bessel process with dimension $\delta < 2$ hits zero. Therefore we denote the first hitting time of a $BESQ^\delta(x)$ process $(X_t)_{t\geq 0}$ with T_0, i.e.,

$$T_0 = \inf\{t \geq 0 | X_t = 0\}.$$

Let R_t be the associated Bessel process. Then it is obvious, that the first hitting time of the Squared Bessel process X_t and the Bessel process R_t coincide:

$$T_0 = \inf\{t \geq 0 | X_t = 0\} = \inf\{t \geq 0 | R_t^2 = 0\} = \inf\{t \geq 0 | R_t = 0\}.$$

Thus we can as well determine the first hitting time of a Bessel process.

Theorem 3.3.1 *Let R_t be a δ-dimensional Bessel process with start in $r = \sqrt{x}$ and let $T_0 = \inf\{t \geq 0 | R_t = 0\}$ be its first hitting time in zero. Then we have*

$$\mathbb{P}(T_0 \in dt) = \frac{1}{t\Gamma(1-\frac{\delta}{2})} \left(\frac{x}{2t}\right)^{1-\frac{\delta}{2}} e^{-\frac{x}{2t}} dt.$$

To prove this result we will make use of the fact that the first hitting time in zero is equal in distribution to the last exit time of the time reversed process. The time reversed process of R_t is denoted by $\hat{R}_t$ according to the following definition:

Definition 3.3.2 *If we just consider the process R_t on the stochastic interval $[0, T_0]$ we can define the time reversed process $\hat{R}_t$ as follows:*

$$\hat{R}_t = R_{T_0 - t} \quad \textit{if } 0 \leq t \leq T_0.$$

Now we have to define the last exit time of a transient diffusion.

Definition 3.3.3 *Let X be a transient diffusion, living on $\mathbb{R}_+$ and starting at x_0, the last exit time of $a \geq 0$ is defined as follows*

$$L_a = \sup\{u | X_u = a\}.$$

Then the following result, see (Göing-Jaeschke and Yor, 2003), can be applied:

Theorem 3.3.4 *Let X be a transient diffusion, living on $\mathbb{R}_+$ and starting at 0. Let $\tilde{X}$ be the time reversed process, starting at a. Then for the last exit time of X, $L_a = \sup\{u|X_u = a\}$ and the first hitting time of $\tilde{X}$, $T_0 = \inf\{u \geq 0|\tilde{X}_u = 0\}$ we have*

$$\{X_u, u \leq L_a\} \stackrel{\mathcal{L}}{=} \{\tilde{X}_{T_0 - u}, u \leq T_0\}.$$

This leads to the following proposition, see (Getoor and Sharpe, 1979):

Proposition 3.3.5 *If we denote $\mathbb{P}_r^\delta$ the probability measure of the $BES^\delta(r)$ and analogously $\mathbb{P}_0^{\hat{\delta}}$ the probability measure of $BES^{\hat{\delta}}(0)$, we have*

$$\mathbb{P}_r^\delta(T_0 \in dt) = \mathbb{P}_0^{\hat{\delta}}(L_r \in dt).$$

Thus to prove Theorem 3.3.1, we can equivalently determine the last exit time of r of the time reversed Bessel process $\hat{R}$. Therefore we need to determine the dimension of the process $\hat{R}$.

The time reversed process can be obtained as a Doob's h-transform, (for Doob's h-transform see section 2.2.2), where the excessive function is given by the scale function of the process.

Lemma 3.3.6 *The scale function of the Bessel process with dimension δ is given by*

$$s(x) = \begin{cases} \frac{x^{-(\delta-2)}}{\delta-2} & \text{if } 0 < \delta < 2 \\ \log(x) & \text{if } \delta = 2 \\ -\frac{x^{-(\delta-2)}}{\delta-2} & \text{if } \delta > 2. \end{cases}$$

Proof:

$$\begin{aligned}
s(x) &= \int_{c_1}^{x} \exp\left(-\int_{c_2}^{y} \frac{2(\delta-1)}{2z}dz\right)dy \\
&\underset{c_2=1}{=} \int_{c_1}^{x} \exp\left(-\int_{1}^{y} \frac{\delta-1}{z}dz\right)dy \\
&= \int_{c_1}^{x} \exp\left(-\log(z)(\delta-1)\right)dy \\
&\underset{c_1=0}{=} \int_{0}^{x} z^{-(\delta-1)}dy.
\end{aligned}$$

□

Thus according to Lemma 2.2.16 the drift parameter of the time reversed process $\hat{R}_t$ is given by

$$\begin{aligned}
a^h(x) &= a(x) + b^2(x)\frac{s'(x)}{s(x)} \\
&= \frac{\delta-1}{2x} - (\delta-2)\frac{x^{-(\delta-1)}}{x^{-(\delta-2)}} \\
&= \frac{\delta-1}{2x} - \frac{\delta-2}{x} = \frac{3-\delta}{2x} = \frac{\hat{\delta}-1}{2x}
\end{aligned}$$

with $\hat{\delta} = 4-\delta$. Thus the time reversed process is again a Bessel process with dimension $\hat{\delta} = 4-\delta$ and is thus transient.

With these preliminaries we can prove now Theorem 3.3.1 by determining the last exit time of a Bessel process with dimension $\hat{\delta} > 2$. This problem has been solved in (Getoor, 1979) for integer dimensions only. Thus in this section we will extend this result to the more general non-integer case. Another way to determine the last exit time is given in (Borodin and Salminen, 1996, IV.44). Accordingly the probability density of the last exit time is given by

$$\mathbb{P}_0^{\hat{\delta}}(L_r \in dt) = \frac{p_m(t,0,r)}{-s(r)}dt$$

where $p_m(t,0,r)$ is the transition probability density with respect to the speed measure $m(x)$, i.e., such that

$$P(t,0,A) = \int_A p_m(t,0,r)m(dr).$$

For the Bessel process with dimension δ we have

$$p_m(t,0,r) = \frac{1}{(2t)^{\frac{\delta}{2}}\Gamma(\frac{\delta}{2})} \exp\left(-\frac{r^2}{2t}\right).$$

As seen in Lemma 3.3.6 the scale function $s(x)$ is given by

$$s(x) = -\frac{x^{-(\delta-2)}}{\delta-2}.$$

Thus we get

$$\begin{aligned} \mathbb{P}_0^{\hat{\delta}}(L_r \in dt) &= \frac{p_m(t,0,r)}{-s(r)}dt \\ &= \frac{2(\frac{\hat{\delta}}{2}-1)r^{\frac{\hat{\delta}}{2}}}{(2t)^{\frac{\hat{\delta}}{2}}\Gamma(\frac{\hat{\delta}}{2})} \exp\left(-\frac{r^2}{2t}\right). \end{aligned}$$

The following property of the Gamma function

$$\Gamma(x+1) = \Gamma(x)\cdot x \quad \forall x \in \mathbb{R}^+ \tag{3.3}$$

and $\hat{\delta} = 4-\delta$ as well as $x = r^2$ leads to

$$\mathbb{P}_x^{\delta}(T_0 \in dt) = \frac{1}{t\Gamma(1-\frac{\delta}{2})}\left(\frac{x}{2t}\right)^{1-\frac{\delta}{2}} \exp\left(-\frac{x}{2t}\right).$$

As already mentioned, in (Getoor, 1979) a proof for the last exit time of a Bessel process with integer dimension $d \geq 3$ is presented. There for $r > 0$ he defines

$$L_r = \sup\{t : |X_t| \leq r\},$$

called the Brownian escape process, where X is a d-dimensional Brownian motion with start in zero and $d \geq 3$. In the following we will show that this proof can be extended to the non-integer case as well. Therefore we have to define another process:

$$F_t = \inf\{\hat{R}_s | s > t\}, \quad t \geq 0.$$

Obviously $t \to F_t$ is increasing and continuous. Furthermore

$$\lim_{t\to\infty} F_t = \infty \qquad \mathbb{P}_0^{\hat{\delta}} - f.s.$$

If the last exit time of r is less than t, the process $\hat{R}$ stays above that level and thus the minimum of the process after t is greater than r, i.e.,

$$\mathbb{P}_0^{\hat{\delta}}(L_r < t) = \mathbb{P}_0^{\hat{\delta}}(F_t > r).$$

Now we introduce the $\sigma(\hat{R}_s, s \geq t)$-measurable shift-operator θ_t. If ω is an arbitrary path of the Bessel process, such that $\hat{R}_t(\omega) = \omega(t)$, then the shift operator θ_t cuts off the path up to time t, i.e.,

$$\theta_t(\omega)(u) = \omega(u + t).$$

Thus

$$\begin{aligned} F_t(\omega) &= \inf\{\hat{R}_u(\omega) | u > t\} \\ &= \inf\{\omega(u) | u > t\} = \inf\{\omega(s + t) | s > 0\} \\ &= \inf\{\hat{R}_s(\omega(\cdot + t)) | s > 0\} = F_0(\omega(\cdot + t)) \\ &= F_0(\theta_t(\omega)) = F_0 \circ \theta_t(\omega). \end{aligned}$$

This leads to

$$\begin{aligned} \mathbb{P}_0^{\hat{\delta}}(F_t > r) &= \mathbb{P}_0^{\hat{\delta}}(F_0 \circ \theta_t > r) \\ &= \mathbb{E}_0^{\hat{\delta}}(\mathbb{P}_{\hat{R}_t}^{\hat{\delta}}(F_0 > r)). \end{aligned}$$

As the next step we determine the probability distribution, respectively the survival function of F_0 for

$$F_0 = \inf\{\bar{R}_s | s > 0\},$$

where $\bar{R}$ is a $BES^{\hat{\delta}}(u)$ for an arbitrary $u > 0$. Since

$$\mathbb{P}_u^{\hat{\delta}}(F_0 > a) = \mathbb{P}_u^{\hat{\delta}}(\inf\{\bar{R}_s | s > 0\} > a),$$

the process evaluated at the exit time of an interval (a, b) for an arbitrary $b > a$ $\mathbb{P}_u^{\hat{\delta}}$-almost surely will be b on the event $\inf\{\bar{R}_s | s > 0\} > a$. Letting $b \to \infty$ means thus

$$\{\bar{R}_{T_{ab}} = b\} \to \{\inf\{\bar{R}_s | s > 0\} > a\} \quad \text{for } b \to \infty$$

where

$$T_{ab} = \inf\{t | \bar{R}_t \notin (a, b)\}.$$

As can be found, e.g., in (Borodin and Salminen, 1996), we have

$$\mathbb{P}_u^{\hat{\delta}}(\bar{R}_{T_{ab}} = b) = \frac{a^{-(\hat{\delta}-2)} - u^{-(\hat{\delta}-2)}}{a^{-(\hat{\delta}-2)} - b^{-(\hat{\delta}-2)}}.$$

Then we can conclude

$$\begin{aligned} \lim_{b \to \infty} \mathbb{P}_u^{\hat{\delta}}(\bar{R}_{T_{ab}} = b) &= \frac{a^{-(\hat{\delta}-2)} - u^{-(\hat{\delta}-2)}}{a^{-(\hat{\delta}-2)}} \\ &= 1 - \left(\frac{a}{u}\right)^{\hat{\delta}-2}. \end{aligned}$$

Hence it follows that

$$\mathbb{P}_u^{\hat{\delta}}(F_0 > a) = 1 - \left(\frac{a}{u}\right)^{\hat{\delta}-2}.$$

Bringing together the facts from above, we get

$$\begin{aligned} \mathbb{P}_0^{\hat{\delta}}(L_r < t) &= \mathbb{P}_0^{\hat{\delta}}(F_t > r) \\ &= \mathbb{P}_0^{\hat{\delta}}(F_0 \circ \theta_t > r) = \mathbb{E}_0^{\hat{\delta}}(\mathbb{P}_{\hat{R}_t}^{\hat{\delta}}(F_0 > r)). \end{aligned}$$

Since we can consider $(\mathbb{P}_{\hat{R}_t}^{\hat{\delta}}(F_0 > r)$ as a function in $\hat{R}_t$, we have

$$\mathbb{E}_0^{\hat{\delta}}(\mathbb{P}_{\hat{R}_t}^{\hat{\delta}}(F_0 > r)) = \int_0^\infty \mathbb{P}_y^{\hat{\delta}}(F_0 > r) p_{\hat{\delta}}(t, 0, y) dy.$$

There $p_{\hat{\delta}}(t,0,y)$ denotes the transition density of the $BES^{\hat{\delta}}(0)$. As seen in 3.2 this equals

$$\begin{aligned} p_{\hat{\delta}}(t,0,y) &= \frac{y^{2\hat{\nu}+1}}{2^{\hat{\nu}}t^{\hat{\nu}+1}\Gamma(\hat{\nu}+1)} \exp\left(-\frac{y^2}{2t}\right) \\ &= \frac{y^{\hat{\delta}-1}}{2^{\frac{\hat{\delta}-2}{2}}t^{\frac{\hat{\delta}}{2}}\Gamma(\frac{\hat{\delta}}{2})} \exp\left(-\frac{y^2}{2t}\right) dy. \end{aligned}$$

Obviously we have to assume $0 < r < y$, since otherwise

$$\mathbb{P}_y^{\hat{\delta}}(F_0 > r) = 0.$$

With

$$\mathbb{P}_y^{\hat{\delta}}(F_0 > r) = 1 - \left(\frac{r}{y}\right)^{\hat{\delta}-2}$$

it follows that

$$\begin{aligned} &\mathbb{E}_0^{\hat{\delta}}(\mathbb{P}_{\hat{R}_t}^{\hat{\delta}}(F_0 > r)) \\ = &\int_r^\infty \left[1 - \left(\frac{r}{y}\right)^{\hat{\delta}-2}\right] \cdot \frac{y^{\hat{\delta}-1}}{2^{\frac{\hat{\delta}-2}{2}}t^{\frac{\hat{\delta}}{2}}\Gamma(\frac{\hat{\delta}}{2})} \exp\left(-\frac{y^2}{2t}\right) dy \\ = &\left[2^{\frac{\hat{\delta}-2}{2}}t^{\frac{\hat{\delta}}{2}}\Gamma(\frac{\hat{\delta}}{2})\right]^{-1} \int_r^\infty \left[1 - \left(\frac{r}{y}\right)^{\hat{\delta}-2}\right] y^{\hat{\delta}-1} \exp\left(-\frac{y^2}{2t}\right) dy \\ = &\left[2^{\frac{\hat{\delta}-2}{2}}t^{\frac{\hat{\delta}}{2}}\Gamma(\frac{\hat{\delta}}{2})\right]^{-1} \left[\int_r^\infty y^{\hat{\delta}-1} \exp\left(-\frac{y^2}{2t}\right) dy - \int_r^\infty r^{\hat{\delta}-2} y \exp\left(-\frac{y^2}{2t}\right) dy\right]. \end{aligned}$$

This term has to be rearranged suitably to evaluate the integrals. Next we get

$$\left[\Gamma(\frac{\hat{\delta}}{2})\right]^{-1} \left[\int_r^\infty \frac{y^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}} \frac{y}{t} \exp\left(-\frac{y^2}{2t}\right) dy - \frac{r^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}} \int_r^\infty \frac{y}{t} \exp\left(-\frac{y^2}{2t}\right) dy\right].$$

Now we consider both integrals separately. First:

$$\int_r^\infty \frac{y^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}} \frac{y}{t} \exp\left(-\frac{y^2}{2t}\right) dy.$$

Here we can substitute v for $\frac{y^2}{2t}$. Hence

$$dv = d\left(\frac{y^2}{2t}\right) = \frac{y}{t}dy.$$

Then we can write

$$\frac{y^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}} = \left(\frac{y^2}{2t}\right)^{\frac{\hat{\delta}-2}{2}} = v^{\frac{\hat{\delta}-2}{2}}.$$

This leads to

$$\int_r^\infty \frac{y^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}} \frac{y}{t} \exp\left(-\frac{y^2}{2t}\right) dy = \int_{\frac{r^2}{2t}}^\infty v^{\frac{\hat{\delta}-2}{2}} \exp(-v)dv.$$

Now we consider the second integral without the coefficient $\frac{r^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}}$:

$$\int_r^\infty \frac{y}{t} \exp(-\frac{y^2}{2t})dy.$$

Here no substitution is necessary, since we get directly

$$\begin{aligned} \int_r^\infty \frac{y}{t} \exp(-\frac{y^2}{2t})dy &= -\exp\left(-\frac{y^2}{2t}\right)\Big|_r^\infty \\ &= \exp\left(-\frac{r^2}{2t}\right). \end{aligned}$$

Summarizing, we get

$$\mathbb{P}_r^{\hat{\delta}}(T_0 < t) = \left[\Gamma\left(\frac{\hat{\delta}}{2}\right)\right]^{-1} \left[\int_{\frac{r^2}{2t}}^\infty v^{\frac{\hat{\delta}-2}{2}} \exp(-v)dv - \frac{r^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}} \exp(-\frac{r^2}{2t})\right].$$

In order to get the density of the first hitting time we have to differentiate this expression with respect to t. For the first part we have

$$
\begin{aligned}
&\frac{\partial}{\partial t}\left(\int_{\frac{r^2}{2t}}^{\infty} v^{\frac{\hat{\delta}-2}{2}} \exp(-v)dv\right)\\
= \;& \frac{r^2}{2t^2}\frac{r^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}}\exp(-\frac{r^2}{2t})\\
= \;& \frac{1}{t}\frac{r^{\hat{\delta}}}{(2t)^{\frac{\hat{\delta}}{2}}}\exp(-\frac{r^2}{2t}) = \frac{1}{t}\left(\frac{r^2}{2t}\right)^{\frac{\hat{\delta}}{2}}\exp(-\frac{r^2}{2t}).
\end{aligned}
$$

Now we determine the derivative of the second part:

$$
\begin{aligned}
&\frac{\partial}{\partial t}\left(\frac{r^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}}\exp(-\frac{r^2}{2t})\right)\\
= \;& \frac{-\frac{\hat{\delta}-2}{2} r^{\hat{\delta}-2}}{2^{\frac{\hat{\delta}-2}{2}} t^{\frac{\hat{\delta}}{2}}}\exp(-\frac{r^2}{2t}) + \frac{r^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}}\frac{r^2}{2t^2}\exp(-\frac{r^2}{2t})\\
= \;& -\frac{\hat{\delta}-2}{2}\left(\frac{r^2}{2t}\right)^{\frac{\hat{\delta}-2}{2}}\frac{1}{t}\exp(-\frac{r^2}{2t}) + \frac{1}{t}\left(\frac{r^2}{2t}\right)^{\frac{\hat{\delta}}{2}}\exp(-\frac{r^2}{2t}).
\end{aligned}
$$

Hence it follows that

$$
\begin{aligned}
&\mathbb{P}(T_0 \in dt)\\
= \;& \frac{\partial}{\partial t}\left[\Gamma\left(\frac{\hat{\delta}}{2}\right)\right]^{-1}\left[\int_{\frac{r^2}{2t}}^{\infty} v^{\frac{\hat{\delta}-2}{2}}\exp(-v)dv - \frac{r^{\hat{\delta}-2}}{(2t)^{\frac{\hat{\delta}-2}{2}}}\exp(-\frac{r^2}{2t})\right]\\
= \;& \left[\Gamma\left(\frac{\hat{\delta}}{2}\right)\right]^{-1}\left[\frac{1}{t}\left(\frac{r^2}{2t}\right)^{\frac{\hat{\delta}}{2}}\exp(-\frac{r^2}{2t})\right.\\
+ \;& \left.\frac{\hat{\delta}-2}{2}\left(\frac{r^2}{2t}\right)^{\frac{\hat{\delta}-2}{2}}\frac{1}{t}\exp(-\frac{r^2}{2t}) - \frac{1}{t}\left(\frac{r^2}{2t}\right)^{\frac{\hat{\delta}}{2}}\exp(-\frac{r^2}{2t})\right]\\
= \;& \left[\Gamma\left(\frac{\hat{\delta}}{2}\right)\right]^{-1}\frac{\hat{\delta}-2}{2}\left(\frac{r^2}{2t}\right)^{\frac{\hat{\delta}-2}{2}}\frac{1}{t}\exp(-\frac{r^2}{2t}).
\end{aligned}
$$

Applying (3.3) we get

$$\mathbb{P}(T_0 \in dt) = \left[\Gamma\left(\frac{\hat{\delta}-2}{2}\right)\right]^{-1} \left(\frac{r^2}{2t}\right)^{\frac{\hat{\delta}-2}{2}} \frac{1}{t} \exp\left(-\frac{r^2}{2t}\right).$$

Since $\hat{\delta} = 4 - \delta$, the above expression becomes

$$\mathbb{P}(T_0 \in dt) = \left[\Gamma\left(1-\frac{\delta}{2}\right)\right]^{-1} \left(\frac{r^2}{2t}\right)^{1-\frac{\delta}{2}} \frac{1}{t} \exp\left(-\frac{r^2}{2t}\right).$$

Hence with $x = r^2$ the assertion from Theorem 3.3.1 follows.

As the next lemma shows, this is exactly the density of the Inverse Gamma distribution.

Lemma 3.3.7 *The first hitting time T_0 has an Inverse Gamma distribution with shape parameter $\alpha = 1 - \frac{\delta}{2}$ and scale parameter $\beta = \frac{2}{x}$.*

Proof: To show that T_0 has an Inverse Gamma distribution, we can use that $\frac{1}{T_0}$ has a Gamma distribution.

$$\begin{aligned} \mathbb{P}(T_0 \le t) &= \mathbb{P}\left(\frac{1}{T_0} \ge \frac{1}{t}\right) \\ &= \int_{\frac{1}{t}}^{\infty} \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1} e^{-\frac{y}{\beta}} dy. \end{aligned}$$

The substitution $s = \frac{1}{y}$ leads to

$$\begin{aligned} \mathbb{P}(T_0 \le t) &= -\int_0^t \frac{1}{\Gamma(\alpha)\beta^\alpha} \left(\frac{1}{s}\right)^{\alpha-1} e^{-\frac{1}{\beta s}} \left(-\frac{1}{s^2}\right) ds \\ &= \int_0^t \frac{1}{\Gamma\left(1-\frac{\delta}{2}\right)} \left(\frac{x}{2s}\right)^\alpha \frac{1}{s} e^{-\frac{x}{2s}} ds. \end{aligned}$$

This integrand is exactly the density from Theorem 3.3.1. □

3.4 The Default Time

In order to determine default probabilities as well as joint default probabilities, we consider a finite set $\mathscr{I}$ of obligors, whereas for our purposes it suffices to consider $\mathscr{I} = \{1, 2\}$. We assign to each obligor $i \in \mathscr{I}$ an ability to pay process X^i with

$$X_t^i = Y_t + Y_t^i \text{ for } t \in [0, \infty)$$

where $(Y_t)_{t \geq 0}$ and $\{(Y_t^i)_{t \geq 0} | i \in \mathscr{I}\}$ are independent Squared Bessel processes:

$$\begin{aligned} Y_t^i &= y_0^i + \delta_i t + 2 \int_0^t \sqrt{Y_s^i} dW_s^i \\ Y_t &= y_0 + \delta t + 2 \int_0^t \sqrt{Y_s} dW_s \end{aligned}$$

with dimensions $\delta, \delta_i \in (0, 2)$. As already mentioned, we identify the default with the first time the ability to pay process reaches a certain level, called default barrier. In our setting this default barrier will be zero for all obligors. Thus, the default time of obligor i, T_i is identified with the first hitting time of the process X^i, i.e.,

$$T_i = \inf\{t | X_t^i = 0\}.$$

Since X^i is composed of two other processes, the default time is the first time **both** processes hit the barrier at the same time. This circumstance would not pose a problem for the single default probability, because due to the additivity property we can directly compute the first hitting time of X^i regardless of Y and Y^i if $\delta + \delta_i < 2$ is satisfied. However, we are also interested in the joint default probability of two firms, who are driven by the same common process $(Y_t)_{t \geq 0}$. In analogy to the single period factor models one wants to exploit the conditional independence of the asset values, respectively ability to pay processes. Therefore we would have to condition on the complete path of Y up to time t and the knowledge of the first hitting time of a Squared Bessel process would no longer be of any use for us, since the first hitting time of X^i almost surely does not coincide with the first hitting time of the idiosyncratic part Y^i. This makes it impossible to compute the joint default probability. The only way out is to force the common process to stay in zero after the first hitting time, i.e., to consider the stopped process. Then a default can only happen after the "default" of the common process. That means the model can be seen as a kind of common shock model where one certain event triggers the possibility of default for all obligors. Thus, instead of Y_t henceforth we consider the stopped process

$$Y_t^{T_0} := Y_{T_0 \wedge t},$$

where

$$T_0 = \inf\{t | Y_t = 0\}.$$

Then, for $t \geq T_0$ we have $Y_t^{T_0} = 0$, so that

$$\begin{aligned} X_t^i &= Y_t^i + Y_t^{T_0} \\ &= \left(Y_t^i + Y_t^{T_0}\right) \mathbb{1}_{\{T_0 > t\}} + \left(Y_t^i + Y_t^{T_0}\right) \mathbb{1}_{\{T_0 \leq t\}} \\ &= \left(Y_t^i + Y_t^{T_0}\right) \mathbb{1}_{\{T_0 > t\}} + Y_t^i \mathbb{1}_{\{T_0 \leq t\}} \\ &= \left(y_0^i + y_0 + (\delta_i + \delta) \cdot t + 2 \int_0^t \sqrt{X_s^i} dB_s^i \right) \cdot \mathbb{1}_{\{T_0 > t\}} \\ &+ \left(y_0^i + \delta_i \cdot t + 2 \int_0^t \sqrt{Y_s^i} dW_s^i \right) \cdot \mathbb{1}_{\{T_0 \leq t\}} \end{aligned}$$

where

$$B_t^i = \int_0^t \frac{\sqrt{Y_s^i} dW_s^i + \sqrt{Y_s} dW_s}{\sqrt{X_s^i}}.$$

Thus we can also decompose the default time T_i in the first hitting time T_0 and the first hitting time of the remaining process after T_0. We denote this stopping time by $\tilde{T}_i$:

$$\tilde{T}_i = \inf\{t | Y_{T_0+t}^i = 0\},$$

i.e., we have

$$T_i = T_0 + \tilde{T}_i.$$

Since the ability to pay process is no longer a pure Squared Bessel process, we also have to compute the single default probability by conditioning on the first hitting time T_0:

$$\begin{aligned}
\mathbb{P}(T_i \leq t) &= \mathbb{P}(T_0 + \tilde{T}_i \leq t) \\
&= \mathbb{E}(\mathbb{P}(\tilde{T}_i \leq t - T_0 | T_0)) \\
&= \int_0^\infty \mathbb{P}(\tilde{T}_i \leq t - s)\mathbb{P}(T_0 \in ds) \\
&= \int_0^\infty \mathbb{P}(\tilde{T}_i \leq t - s)\mathbb{1}_{\{s \leq t\}}\mathbb{P}(T_0 \in ds) \\
&+ \int_0^\infty \mathbb{P}(\tilde{T}_i \leq t - s)\mathbb{1}_{\{s > t\}}\mathbb{P}(T_0 \in ds).
\end{aligned}$$

For the second integral we get

$$\begin{aligned}
&\int_0^\infty \mathbb{P}(\tilde{T}_i \leq t - s)\mathbb{1}_{\{s > t\}}\mathbb{P}(T_0 \in ds) \\
= &\int_t^\infty \mathbb{P}(\tilde{T}_i \leq t - s)\mathbb{P}(T_0 \in ds) \\
= \ &0.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\mathbb{P}(T_i \leq t) &= \int_0^\infty \mathbb{P}(\tilde{T}_i \leq t - s)\mathbb{1}_{\{s \leq t\}}\mathbb{P}(T_0 \in ds) \\
&= \int_0^t \mathbb{P}(\tilde{T}_i \leq t - s)\mathbb{P}(T_0 \in ds).
\end{aligned}$$

Using the definition of $\tilde{T}_i$ leads to

$$\mathbb{P}(T_i \leq t) = \int_0^t \mathbb{P}(\inf\{u | Y^i_{s+u} = 0\} \leq t - s)\mathbb{P}(T_0 \in ds).$$

To be able to proceed, we have to use the Strong Markov property of the Squared Bessel process. The Strong Markov property is stated in the following Theorem:

Theorem 3.4.1 *If Z is a $\mathcal{F}_\infty$-measurable and positive (or bounded) random variable and T is a stopping time, for any initial measure ν,*

$$\mathbb{E}_\nu(Z \circ \theta_T | \mathcal{F}_T) = E_{X_T}(Z)$$

$\mathbb{P}_\nu - a.s.$ on the set $\{T < \infty\}$.

Proof: see (Revuz and Yor, 1991) $\square$

Due to the Strong Markov property, conditional on a realization of Y^i, evaluated at the random time T_0, i.e., on the event $\{Y^i_{T_0} = y\}$, the conditional probability distribution of $\tilde{T}_i$ is equal to the distribution of the first hitting time of a Squared Bessel process, starting at y. Now we can apply Theorem 3.4.1 to compute $\mathbb{P}(\inf\{u|Y^i_{T_0+u} = 0\} \leq t-s)$. Therefore we write this probability as the expectation of the indicator function:

$$\begin{aligned} & \mathbb{P}(\inf\{u|Y^i_{T_0+u} = 0\} \leq t-s) \\ = \ & \mathbb{E}(\mathbb{1}_{[0,t-s]}(\inf\{u|Y^i_{T_0+u} = 0\})). \end{aligned}$$

Applying the tower property leads to

$$\begin{aligned} & \mathbb{E}(\mathbb{1}_{[0,t-s]}(\inf\{u|Y^i_{T_0+u} = 0\})) \\ = \ & \mathbb{E}(\mathbb{E}^{\delta_i}_{y^i_0}(\mathbb{1}_{[0,t-s]}(\inf\{u|Y^i_{T_0+u} = 0\})|\mathcal{F}_{T_0})). \end{aligned}$$

Now we can apply Theorem 3.4.1 for

$$Z = \mathbb{1}_{[0,t-s]}(\inf\{u|Y^i_{T_0+u} = 0\})$$

and get

$$\begin{aligned} & \mathbb{E}(\mathbb{E}^{\delta_i}_{y^i_0}(\mathbb{1}_{[0,t-s]}(\inf\{u|Y^i_{T_0+u} = 0\})|\mathcal{F}_{T_0})) \\ = \ & \mathbb{E}(\mathbb{E}^{\delta_i}_{y^i_0}(\mathbb{1}_{[0,t-s]}(\inf\{u|Y^i_u = 0\}) \circ \theta_{T_0}|\mathcal{F}_{T_0})) \\ = \ & \mathbb{E}(\mathbb{E}^{\delta_i}_{Y^i_{T_0}}(\mathbb{1}_{[0,t-s]}(\inf\{u|Y^i_u = 0\}))) \\ = \ & \mathbb{E}(\mathbb{P}^{\delta_i}_{Y^i_{T_0}}(\inf\{u|Y^i_u = 0\} \leq t-s)). \end{aligned}$$

As already seen, the transition density of a Squared Bessel process X is given by 3.1. We denote the associated semi-group by $(P_t)_{t\geq 0}$. Then the following property applies to the transition kernel $P_t(\cdot,\cdot)$:

$$\begin{aligned} P_t(x,A) &= \int_A P_t(x,dy) \\ &= \int_A p(t,x,y)dy. \end{aligned}$$

Furthermore, we have

$$\mathbb{E}_x(f(X_t)) = \int_0^\infty f(y)p(t,x,y)dy = P_t f(x).$$

As an application of the Strong Markov property, this extends to

$$\mathbb{E}_x(f(X_T)) = P_T f(x)$$

for any stopping time T. Thus with

$$\mathbb{P}^{\delta_i}_{Y^i_{T_0}}(\inf\{u|Y^i_u = 0\} \leq t-s) =: f(Y^i_{T_0})$$

we have

$$\begin{aligned} &\mathbb{E}(\mathbb{P}^{\delta_i}_{Y^i_{T_0}}(\inf\{u|Y^i_u = 0\} \leq t-s)) \\ =\ & P_{T_0}\mathbb{P}^{\delta_i}_{y^i_0}(\inf\{u|Y^i_u = 0\} \leq t-s) = \int_0^\infty \mathbb{P}^{\delta_i}_y(\inf\{u|Y^i_u = 0\} \leq t-s)p(T_0, y^i_0, y)dy. \end{aligned}$$

Then, on the event $\{T_0 = s\}$, we get

$$\int_0^\infty \mathbb{P}^{\delta_i}_y(\inf\{u|Y^i_u = 0\} \leq t-s)p_s(y^i_0, y)dy.$$

Finally, the probability distribution of the default time T_i evolves into

$$\begin{aligned}
&\mathbb{P}(T_i \leq t) \\
=& \int_0^t \int_0^\infty \mathbb{P}_y^{\delta_i}(\inf\{u|Y_u^i = 0\} \leq t-s) p(s, y_0^i, y) dy \mathbb{P}(T_0 \in ds) \\
=& \int_0^t \int_0^\infty \left[\int_0^{t-s} \Gamma\left(1-\frac{\delta_i}{2}\right)^{-1} \frac{1}{u} \left(\frac{y_0^i}{2u}\right)^{1-\frac{\delta_i}{2}} e^{-\frac{y_0^i}{2u}} du \right] \\
& \frac{1}{2s}\left(\frac{y}{y_0^i}\right)^{\frac{\delta_i - 1}{4}} \exp\left(-\frac{y_0^i + y}{2s}\right) I_{\frac{\delta_i-1}{2}}\left(\frac{\sqrt{y_0^i y}}{s}\right) dy \\
& \Gamma\left(1-\frac{\delta}{2}\right)^{-1} \frac{1}{s}\left(\frac{y_0}{2s}\right)^{1-\frac{\delta}{2}} e^{-\frac{y_0}{2s}} ds.
\end{aligned}$$

Now we denote the distribution function of the Inverse Gamma distribution by $F_{IG}(x; \alpha, \beta)$ and its probability density by $f_{IG}(x; \alpha, \beta)$, where α is the shape parameter and β is the scale parameter.

Accordingly the probability density of the noncentral-χ^2-distribution is denoted by $f_{NC\chi^2}(x; k, \lambda)$ with k degrees of freedom and noncentrality-parameter λ.

With these notations, the distribution function of the default time T_i becomes

$$\begin{aligned}
\mathbb{P}(T_i \leq t) = \int_0^t \int_0^\infty \quad & F_{IG}\left(t-s; 1-\frac{\delta_i}{2}, \frac{2}{y}\right) \\
& f_{NC\chi^2}\left(\frac{y}{s}; 1-\frac{\delta_i}{2}, \frac{y_0^i}{s}\right) f_{IG}\left(s; 1-\frac{\delta}{2}, \frac{2}{y_0}\right) \frac{1}{s} dy ds.
\end{aligned}$$

3.5 Joint Default Probability

In order to compute the default correlation of two firms which is given by

$$\rho_{ij}(t) = \frac{\mathbb{P}(T_i \leq t, T_j \leq t) - \mathbb{P}(T_i \leq t)\mathbb{P}(T_j \leq t)}{\sqrt{\mathbb{P}(T_i \leq t)(1-\mathbb{P}(T_i \leq t))\mathbb{P}(T_j \leq t)(1-\mathbb{P}(T_j \leq t))}} \tag{3.4}$$

we need to determine the joint default probability, i.e., the probability that both firms have defaulted by time t. The joint default probability of two firms that are related by the common process Y_t can now be easily be determined. We have

$$\begin{aligned}
\mathbb{P}(T_1 \leq t, T_2 \leq t) &= \mathbb{P}(T_0 + \tilde{T}_1 \leq t, T_0 + \tilde{T}_2 \leq t) \\
&= \mathbb{E}(\mathbb{P}(\tilde{T}_1 \leq t - T_0, \tilde{T}_2 \leq t - T_0 | T_0)) \\
&= \int_0^\infty \mathbb{P}(\tilde{T}_1 \leq t - s, \tilde{T}_2 \leq t - s)\mathbb{P}(T_0 \in ds) \\
&= \int_0^\infty \mathbb{P}(\tilde{T}_1 \leq t - s)\mathbb{P}(\tilde{T}_2 \leq t - s)\mathbb{P}(T_0 \in ds).
\end{aligned}$$

Here we have used the independence of $\tilde{T}_1$ and $\tilde{T}_2$.
In the last section we have determined the distribution function of $\tilde{T}_i$:

$$\mathbb{P}(\tilde{T}_i \leq t - s) = \int_0^\infty \mathbb{P}_y^{\delta_i}(\inf\{u | Y_u = 0\} \leq t - s) p_s(y_0^i, y) dy \mathbb{P}(T_0 \in ds).$$

Thus, for the joint default probability we get

$$\begin{aligned}
&\mathbb{P}(T_1 \leq t, T_2 \leq t) \\
= &\int_0^t \mathbb{P}(\tilde{T}_1 \leq t - s)\mathbb{P}(\tilde{T}_2 \leq t - s)\mathbb{P}(T_0 \in ds) \\
= &\int_0^t \int_0^\infty \mathbb{P}_{y_1}^{\delta_1}(\inf\{u | Y_u = 0\} \leq t - s) p_s(y_0^1, y_1) dy_1 \\
&\int_0^\infty \mathbb{P}_{y_2}^{\delta_2}(\inf\{u | Y_u = 0\} \leq t - s) p_s(y_0^2, y_2) dy_2 \mathbb{P}(T_0 \in ds) \\
= &\int_0^t \int_0^\infty F_{IG}\left(t - s; 1 - \frac{\delta_1}{2}, \frac{2}{y_1}\right) f_{NC\chi^2}\left(\frac{y_1}{s}; 1 - \frac{\delta_1}{2}, \frac{y_0^1}{s}\right) \frac{1}{s} dy_1 \\
&\int_0^\infty F_{IG}\left(t - s; 1 - \frac{\delta_2}{2}, \frac{2}{y_2}\right) f_{NC\chi^2}\left(\frac{y_2}{s}; 1 - \frac{\delta_2}{2}, \frac{y_0^2}{s}\right) \frac{1}{s} dy_2 \\
&f_{IG}\left(s; 1 - \frac{\delta}{s}, \frac{2}{y_0}\right) ds.
\end{aligned}$$

3.6 Examples

The next table shows the average S&P default rates.

years	1	2	3	4	5	6	7	8	9	10
rating										
AAA	0.00	0.00	0.03	0.06	0.10	0.17	0.25	0.38	0.43	0.48
AA	0.01	0.03	0.08	0.16	0.27	0.39	0.53	0.65	0.75	0.85
A	0.05	0.15	0.28	0.44	0.62	0.81	1.03	1.25	1.52	1.82
BBB	0.37	0.94	1.52	2.34	3.20	4.02	4.74	5.40	5.99	6.68
BB	1.38	4.07	7.16	9.96	12.34	14.65	16.46	18.02	19.60	20.82
B	6.20	13.27	19.07	23.45	26.59	29.08	31.41	33.27	34.58	35.87
CCC	27.87	36.02	41.79	46.26	50.46	52.17	53.60	54.36	56.16	57.21

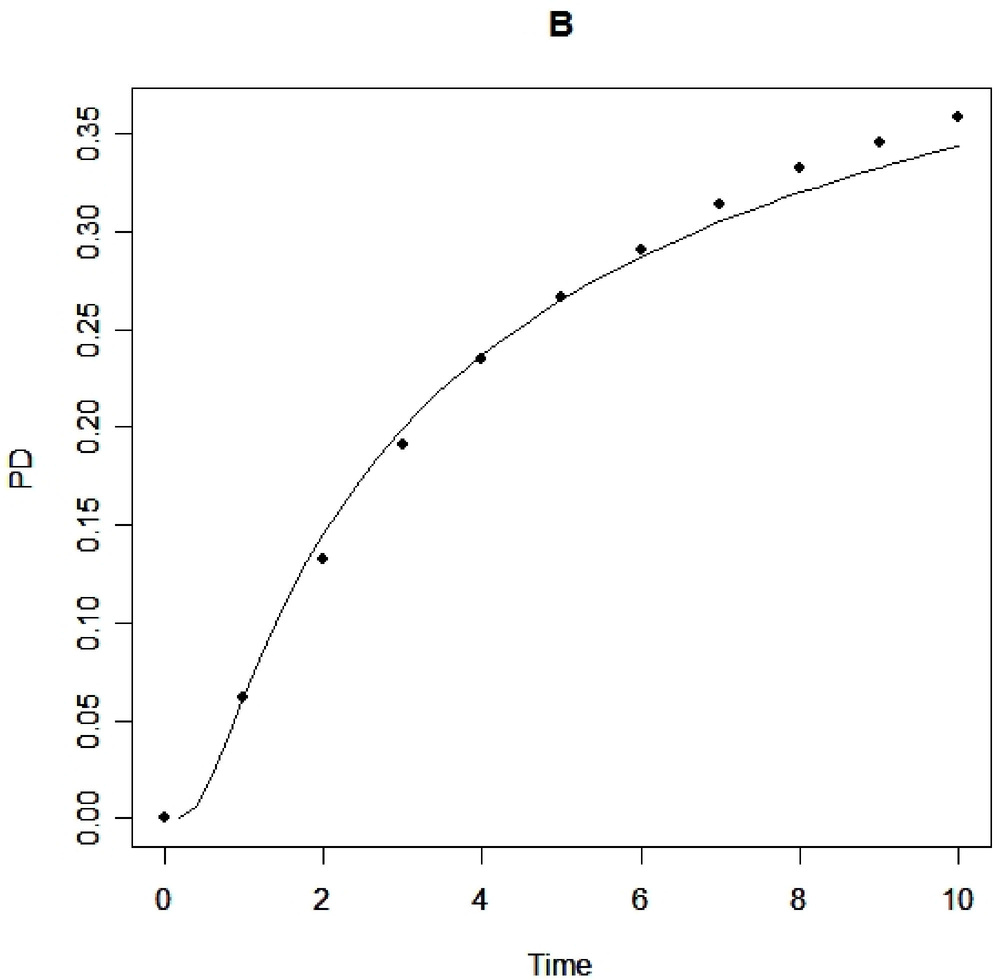
Table 3.1: Average S&P default rates [in %]

In order to calibrate the model to these values, there are four parameters available $(\delta, y_0, \delta_1, y_0^1)$. As examples we have chosen the ratings B and CCC. Then the following parameters result from the calibration.

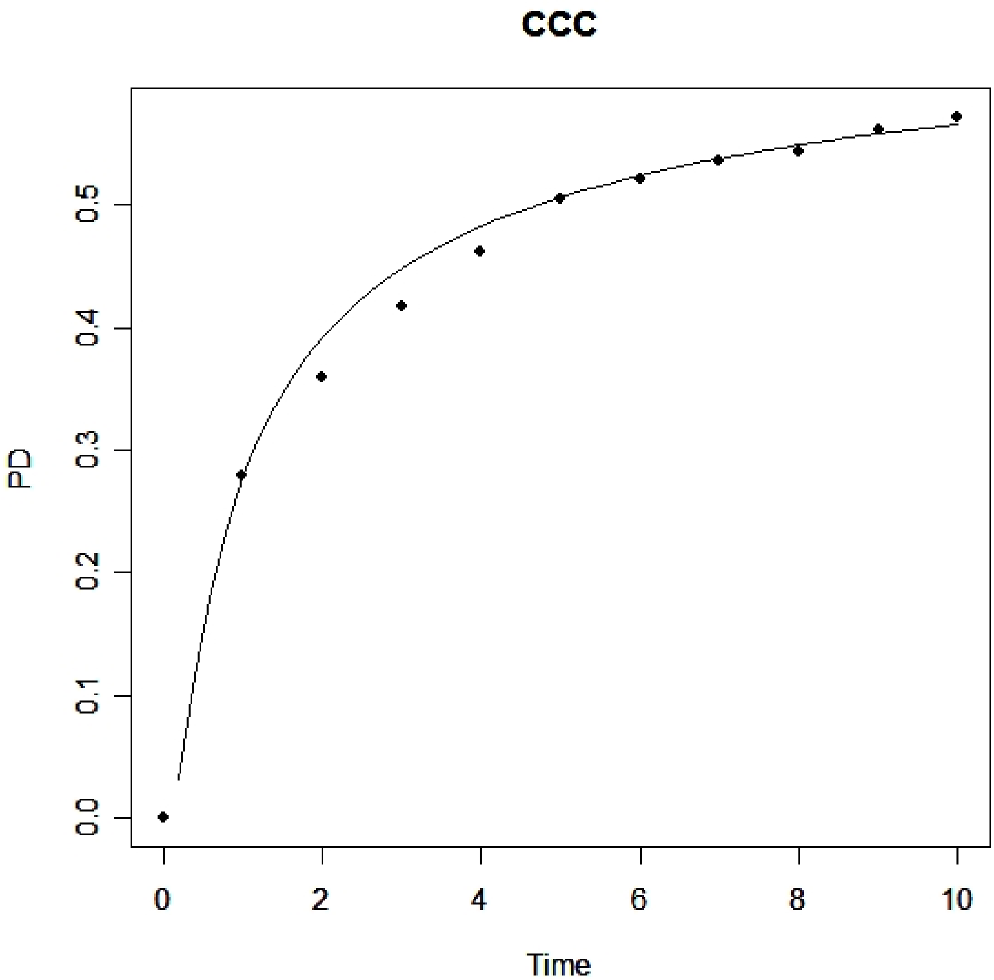
Parameters	δ	y_0	δ_1	y_0^1
Rating				
B	1.45	1.02	0.38	0.8
CCC	0.941	0.76	0.051	0.025

Table 3.2: calibration parameters

This leads to the following graphs where the long term default probabilities are displayed.

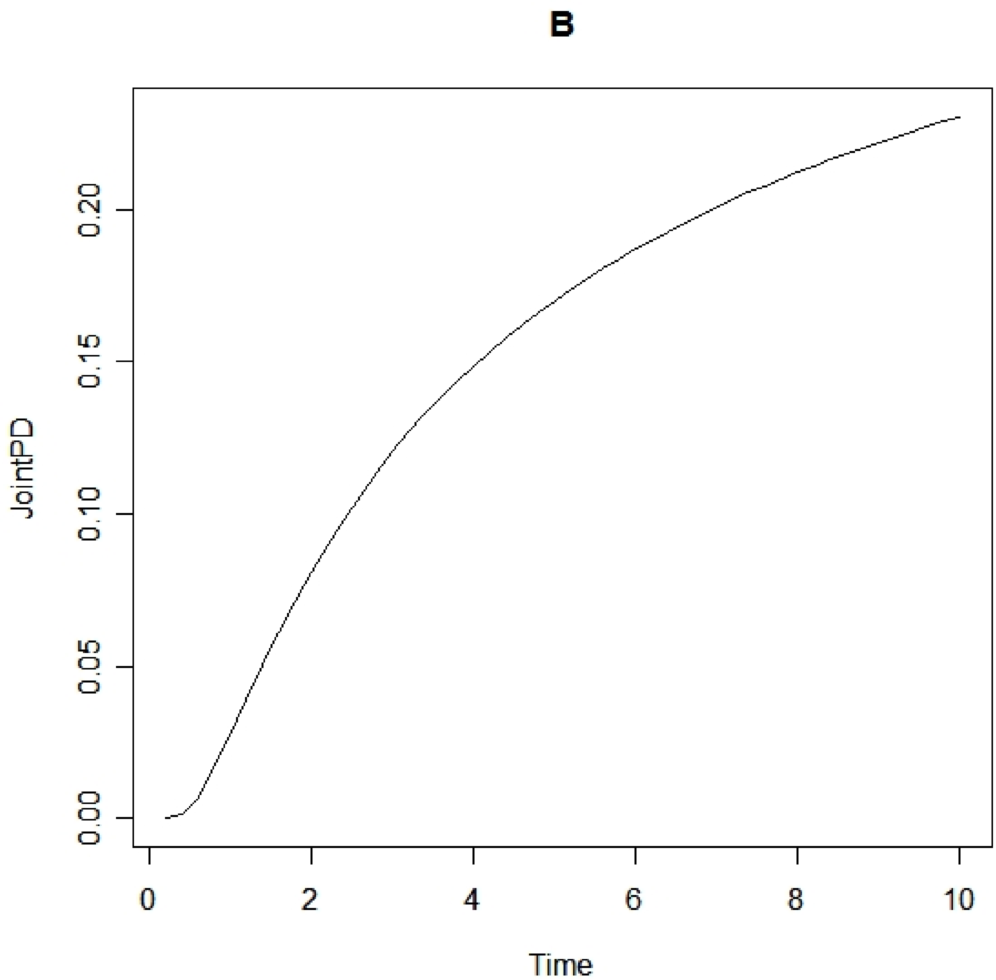


(a) Model and S & P default probabilities for B

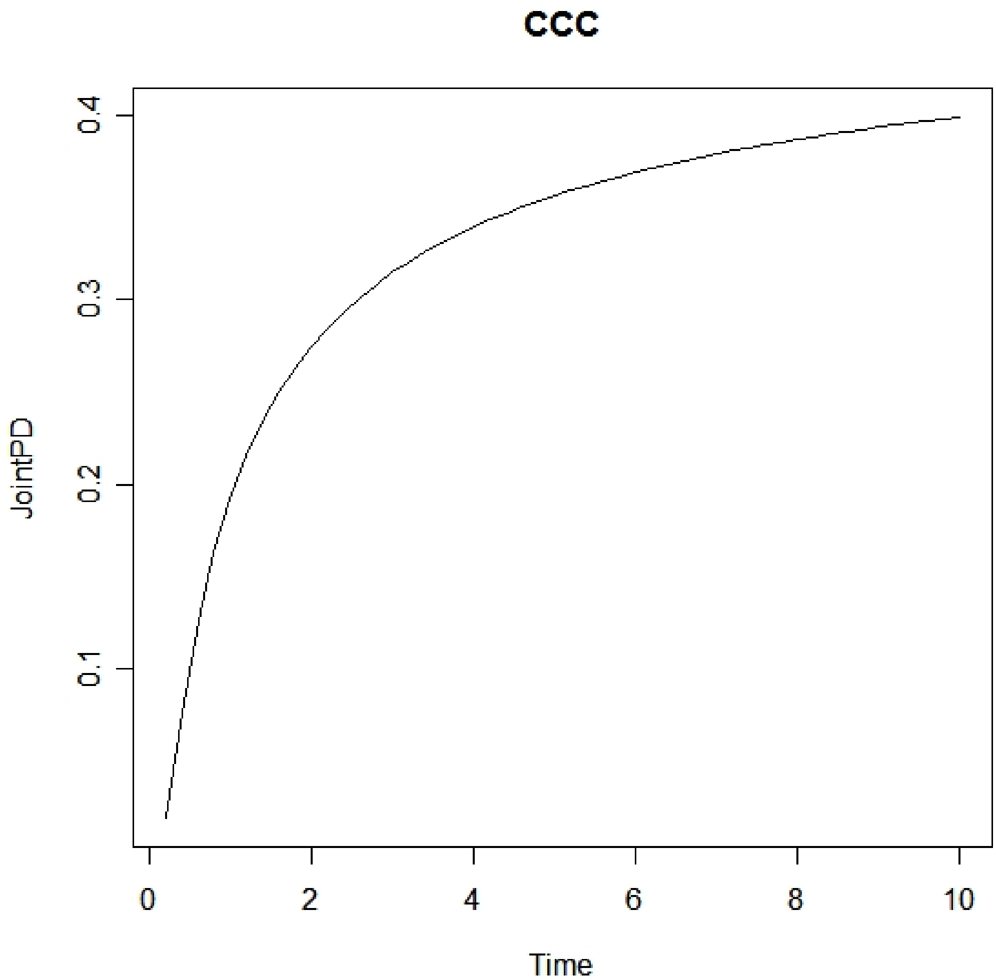


(b) Model and S & P default probabilities for CCC

Now we can compute the joint default probabilities for two B-rated firms as well as for two CCC-rated firms. We get

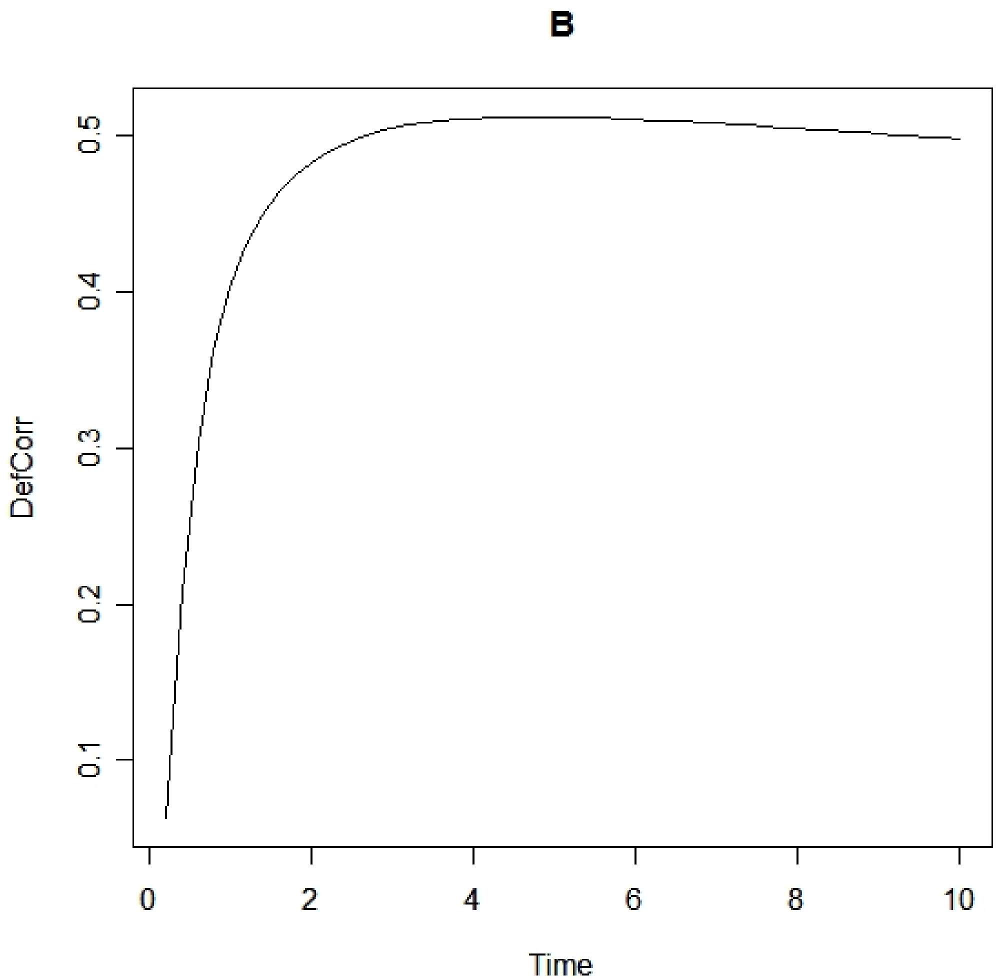


(c) Joint default probabilities for B-rated firms

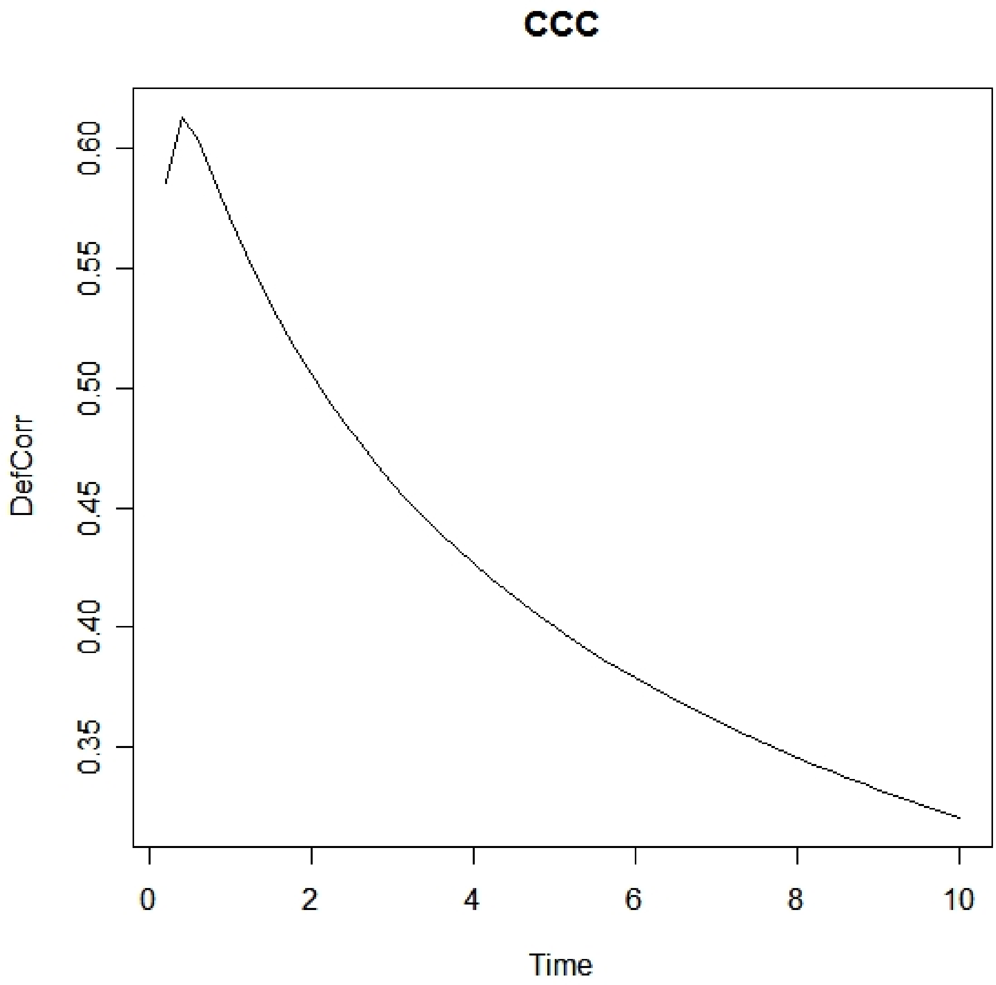


(d) Joint default probabilities for CCC-rated firms

With these results we can determine the default correlations, see 3.4.



(e) Default correlation for B-rated firms



(f) Default correlation for CCC-rated firms

Thus for the time horizon of one year we have a default correlation of 0.4027 for the B-rated firms and 0.5716 for the CCC-rated firms. For these default correlations we can determine their respective implied asset correlations, i.e., the correlation we have to plug into the normal copula such that the resulting joint default probability equals the one we computed before, i.e., we are looking for a ρ such that

$$\mathbb{P}(T_1 \leq 1, T_2 \leq 1) = \Phi_2(\Phi^{-1}(\mathbb{P}(T_1 \leq 1)), \Phi^{-1}(\mathbb{P}(T_2 \leq 1)); \rho)$$

where Φ_2 denotes the two-dimensional normal distribution.
In the case of the two B-rated firms we have

$$0.0273 = \Phi_2(\Phi^{-1}(0.062), \Phi^{-1}(0.062); \rho).$$

Thus, in this case we get an implied asset correlation of

$$\rho = 0.73$$

3.6.1 Conclusions and Remarks

As we have seen within one rating class it is possible to fit approximately the whole default probability term-structure. Since the Squared Bessel model is a one-factor model the parameters δ and y_0, i.e., the dimension and the starting value of the common process have to remain unchanged for two obligors that are assumed to depend on each other. Thus for two obligors within the same rating class the model is able to produce high joint default probabilities. If we fix one time horizon and fix the parameters of the common process it is possible to fit default probabilities across different rating classes. But the model is unfortunately not flexible enough to fit the whole term-structure or even at least two different time horizons of two different rating classes with fixed parameters δ and y_0. The reason is that the range of possible term structures is too narrow if we only vary the parameters of the individual process.

3.7 Large Homogeneous Portfolio Approximation

Since in the structural Squared Bessel model, the ability to pay process is composed of an individual and a common part, the default risk can be divided into a systematic risk and an idiosyncratic risk. Thus, we can consider a large homogenous portfolio of infinitely many issuers, where loosely speaking, the idiosyncratic risk is diversified away. Based on the framework of the Merton model, introduced in (Merton, 1974), see section 1.1, Vasicek has derived a simple closed-form solution for the loss distribution of an asymptotically large, homogeneous portfolio in (Vasicek, 1987) and (Vasicek, 2002). Though the limiting distribution in the Squared Bessel setting is far from being available in closed form, nevertheless we can compare the shape of the resulting distribution with the Vasicek distribution. In the first subsection, the Vasicek approach will be introduced.

3.7.1 Vasicek Approach

As already mentioned, this approach is based on the Merton framework. Thus, the asset value of every firm follows a geometric Brownian motion, i.e., for firm i the dynamics of the asset value are given by

$$dV_t^i = \mu_i V_t^i dt + \sigma_i V_t^i dW_t^i$$

where $\mu_i \in \mathbb{R}$ and $\sigma_i > 0$ are constants and W^i is a standard Brownian motion. We assume that every issued bond matures at date T. That means, a default of firm i happens if its firm value V_T^i lies below the threshold K_i. The firm value V_T^i at time T is given by

$$V_T^i = V_0^i \exp\left(\left(\mu_i - \frac{1}{2}\sigma_i^2\right)T + \sigma_i W_T^i\right).$$

Thus, if τ_i denotes the default time of firm i, the default probability is

$$p_i = \mathbb{P}(\tau_i = T) = \Phi(c_i)$$

where

$$c_i = \frac{\log\left(\frac{K}{V_0^i}\right) - \left(\mu_i - \frac{1}{2}\sigma_i^2\right)T}{\sigma_i\sqrt{T}}.$$

That means, the probability of default is equal to the probability that a standard normally distributed random variable X_i falls below a threshold c_i. In order to introduce a dependence structure, we assume that the standard

Brownian motions all have pairwise instantaneous correlation ρ. Thus, for two firms i and j, we have

$$Cor(X_i, X_j) = \rho.$$

This is fulfilled, if we set

$$X_i = Y\sqrt{\rho} + Z_i\sqrt{1-\rho}$$

where $Y, Z_1, Z_2, ...$ are i.i.d. standard normally distributed random variables. Then Y can be interpreted as a common factor and Z_i represents the idiosyncratic risk.

Now we consider a portfolio consisting of n loans. To simplify matters every loan has an equal amount. Furthermore, every default probability is assumed to be equal to p. The random variable L^n denotes the portfolio loss as a percentage of the total portfolio:

$$L^n = \frac{1}{n}\sum_{i=1}^{n} L_i$$

where

$$L_i = \mathbb{1}_{\{X_i \le \Phi^{-1}(p)\}}.$$

If we condition on a realization of the common factor, $Y = y$, the expectation of the individual losses, i.e., the default probability conditional on the realization is given by

$$\begin{aligned}
p(y) = \mathbb{E}(L_i|Y = y) &= \mathbb{E}(\mathbb{1}_{\{X_i \le \Phi^{-1}(p)\}}|Y = y) \\
&= \mathbb{P}(X_i \le \Phi^{-1}(p)|Y = y) \\
&= \mathbb{P}(Y\sqrt{\rho} + Z_i\sqrt{1-\rho} \le \Phi^{-1}(p)|Y = y) \\
&= \mathbb{P}(y\sqrt{\rho} + Z_i\sqrt{1-\rho} \le \Phi^{-1}(p)) \\
&= \mathbb{P}\left(Z_i \le \frac{\Phi^{-1}(p) - y\sqrt{\rho}}{\sqrt{1-\rho}}\right) \\
&= \Phi\left(\frac{\Phi^{-1}(p) - y\sqrt{\rho}}{\sqrt{1-\rho}}\right).
\end{aligned}$$

Conditional on a realization y of Y the L_i are independent. That means, the law of large numbers gives

$$\mathbb{P}(\lim_{n\to\infty} L^n = p(y)|Y = y) = 1 \tag{3.5}$$

Thus, for $L := \lim_{n\to\infty} L^n$ we get

$$\begin{aligned}
\mathbb{P}(L \le x) &= \mathbb{E}(\mathbb{P}(L \le x|Y)) \\
&= \int_{-\infty}^{\infty} \mathbb{P}(L \le x|Y = y)\phi(y)dy \\
&\underset{3.5}{=} \int_{-\infty}^{\infty} \mathbb{1}_{\{p(y)\le x\}}\phi(y)dy
\end{aligned}$$

Since $p(y)$ is a decreasing function, we get

$$\begin{aligned}
\mathbb{P}(L \le x) &= \int_{-\infty}^{\infty} \mathbb{1}_{\{y\ge p^{-1}(x)\}}\phi(y)dy \\
&= \int_{p^{-1}(x)}^{\infty} \phi(y)dy \\
&= 1 - \Phi(p^{-1}(x)) = \Phi(-p^{-1}(x)) \\
&= \Phi\left(\frac{\sqrt{1-\rho}\Phi^{-1}(x) - \Phi^{-1}(p)}{\sqrt{\rho}}\right).
\end{aligned}$$

3.7.2 Squared Bessel Approach

The main difference to the last approach is, that the Squared Bessel model is a first passage model, i.e., if we also consider the interval $[0, T]$, a default can happen at every time. As already shown in the last section, the default time is given as

$$T_i = T_0 + \tilde{T}_i.$$

In order to be in accordance with the Vasicek approach, it is assumed that every obligor has the same default probability over the considered time horizon, i.e.,

$$p_T = \mathbb{P}(T_i \le T) \quad \forall i = 1, ..., n.$$

If we consider the same portfolio again, the portfolio loss is given by

$$L_T^n = \frac{1}{n}\sum_{i=1}^{n} L_i$$

where

$$L_i = \mathbb{1}_{\{T_i \le T\}}.$$

If we condition on a realization of T_0, i.e., $T_0 = s$, the remaining random variable $\tilde{T}_i$ is the time until the first hitting of zero after time s for the idiosyncratic process Y^i. The expectation of L_i conditional on $T_0 = s$ is given by

$$\begin{aligned}
L_T(s) &= \mathbb{E}(L_i|T_0 = s) = \mathbb{E}(\mathbb{1}_{\{T_i \le T\}}|T_0 = s) \\
&= \mathbb{P}(T_i \le T|T_0 = s) = \mathbb{P}(T_0 + \tilde{T}_i \le T|T_0 = s) \\
&= \mathbb{P}(s + \tilde{T}_i \le T) = \mathbb{P}(\tilde{T}_i \le T - s) \\
&= \mathbb{P}_{y_0}^{\delta_i}(\inf\{u|Y_{s+u}^i = 0\} \le T - s) \\
&= \mathbb{E}_{Y_s^i}(\mathbb{P}_{Y_s^i}^{\delta_i}(\inf\{u|Y_u^i = 0\} \le T - s)) \\
&= \int_0^\infty \frac{\Gamma\left(1 - \frac{\delta_i}{2}, \frac{2}{y(T-s)}\right)}{\Gamma\left(1 - \frac{\delta_i}{2}\right)} p_s(y_0^i, y)dy.
\end{aligned}$$

Conditional on a realization s of T_0 the L_i are independent. That means, the law of large numbers gives

$$\mathbb{P}(\lim_{n\to\infty} L_T^n = L_T(s)|T_0 = s) = 1 \tag{3.6}$$

Thus, for $L_T = \lim_{n\to\infty} L_T^n$ we get

$$\begin{aligned}
\mathbb{P}(L_T \le x) &= \mathbb{E}(\mathbb{P}(L_T \le x|T_0) \\
&= \int_0^\infty \mathbb{P}(L_T \le x|T_0 = s)\mathbb{P}(T_0 \in ds) \\
&\underset{3.6}{=} \int_0^\infty \mathbb{1}_{\{L_T(s) \le x\}}\mathbb{P}(T_0 \in ds)
\end{aligned}$$

Since $L_T(s)$ is a decreasing function, we get

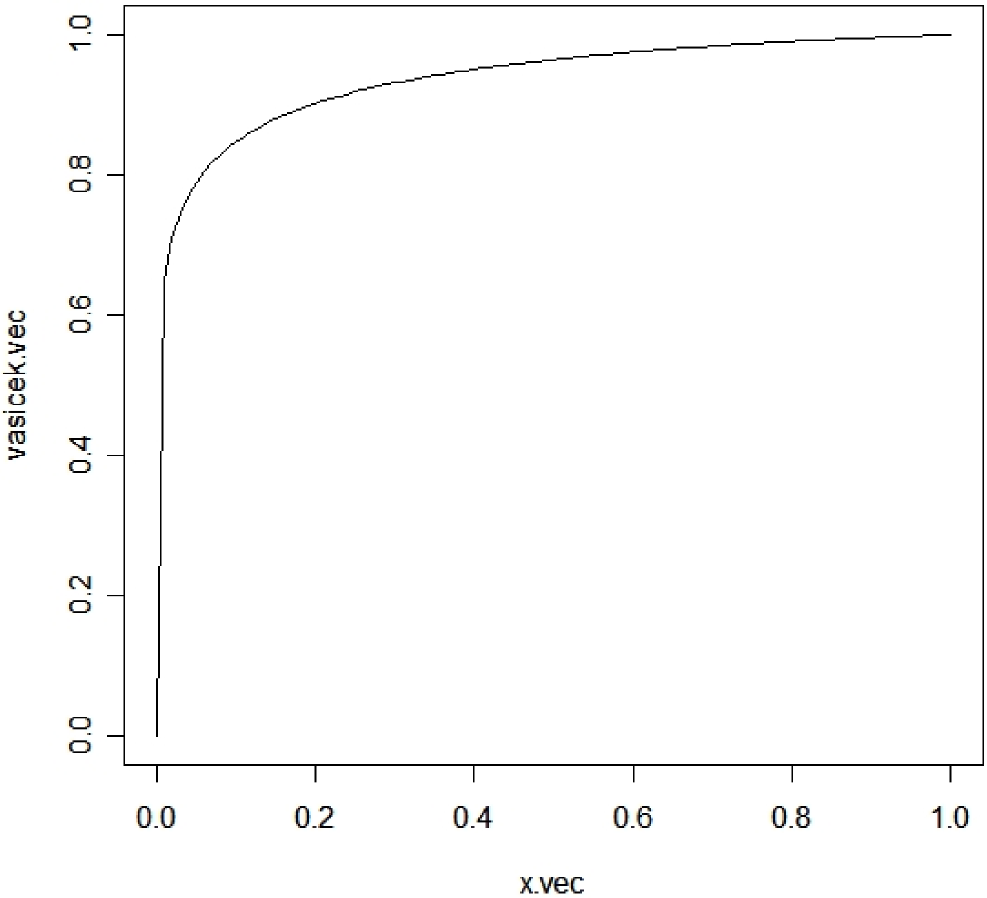
$$\begin{aligned}
\mathbb{P}(L_T \leq x) &= \int_0^\infty \mathbb{1}_{\{s \geq L_T^{-1}(x)\}} \mathbb{P}(T_0 \in ds) \\
&= \int_{L_T^{-1}(x)}^\infty \mathbb{P}(T_0 \in ds) = \mathbb{P}(T_0 \geq L_T^{-1}(x)) \\
&= 1 - \mathbb{P}(T_0 \leq L^{-1}(x)) \\
&= 1 - \frac{\Gamma\left(1 - \frac{\delta}{2}, \frac{2}{y_0 L^{-1}(x)}\right)}{\Gamma\left(1 - \frac{\delta}{2}\right)}.
\end{aligned}$$

3.7.3 Comparison

In order to be able to compare these two approaches we use the parameters from the last section for a B-rated obligor. In the Squared Bessel approach we use the following parameters:

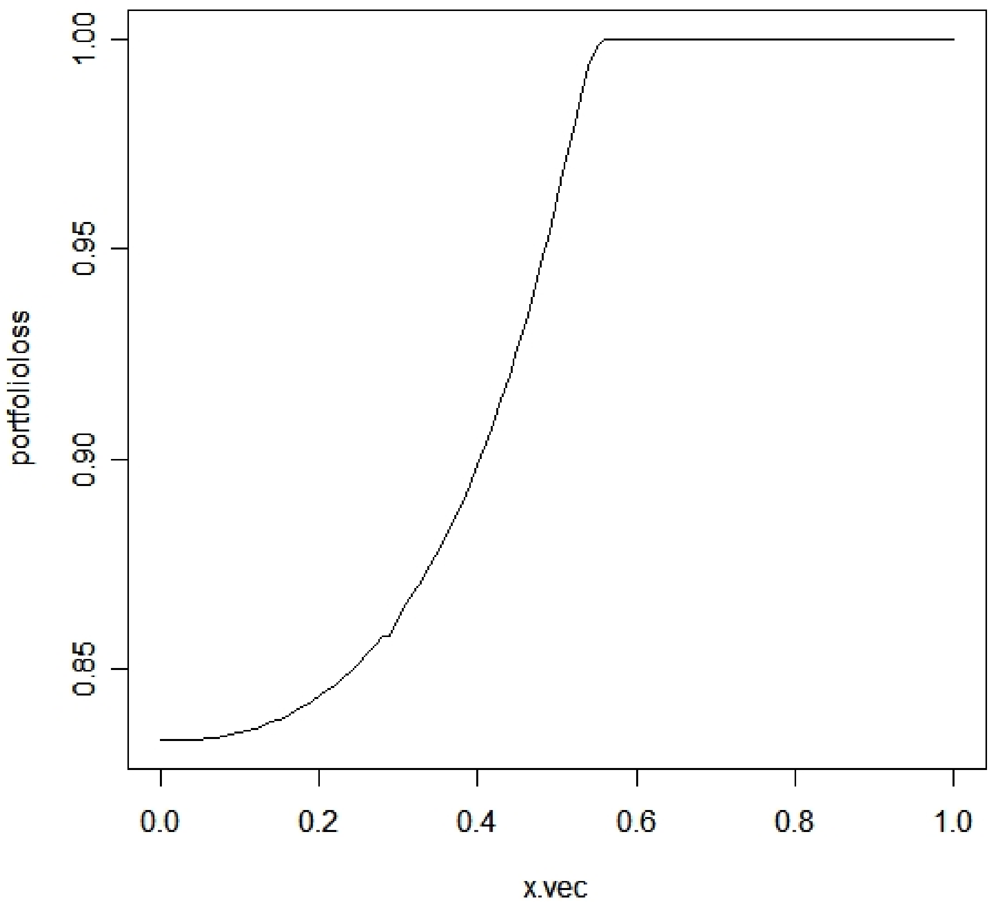
$$\delta = 1.45 \quad y_0 = 1.02 \quad \delta_1 = 0.38 \quad y_0^1 = 0.8$$

Then the resulting one-year default probability was given by $p = 0.062$ and the implied asset correlation was $\rho = 0.73$. These two parameters will be used for the Vasicek approach. Then the portfolio loss distribution function in the Vasicek approach is given by



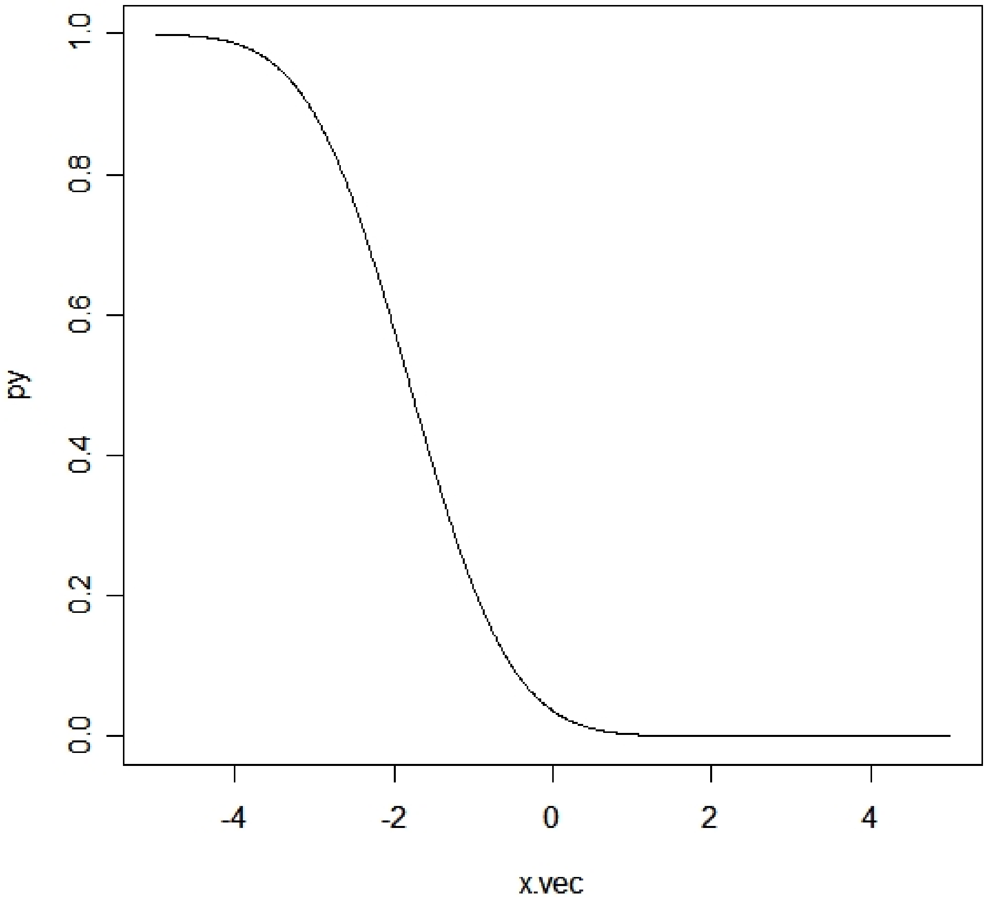
(g) Vasicek loss distribution

In the Squared Bessel approach the portfolio loss distribution function looks as follows



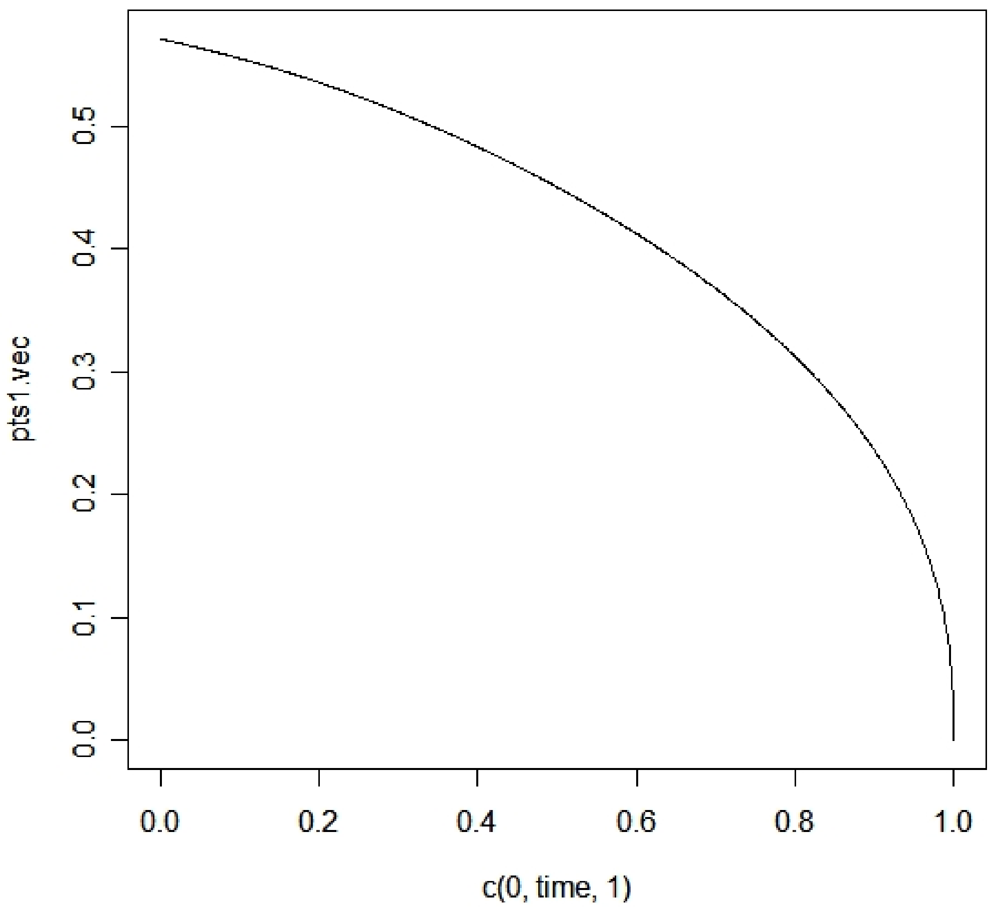
(h) Squared Bessel loss distribution

To explain this result we have to take a look at the decreasing functions $p : \mathbb{R} \to [0, 1]$ in the Vasicek approach and $L : \mathbb{R}_+ \to [0, 1]$ in the Squared Bessel approach. We have



(i) The decreasing function p in the Vasicek approach

and



(j) The decreasing function L in the Squared Bessel approach

Thus at the lower bound we have in the Vasicek approach

$$p(s) \to 1 \text{ for } s \to -\infty$$

and then

$$\mathbb{P}(p(Y) \leq 1) = \mathbb{P}(Y \geq -\infty)) = 1.$$

In the Squared Bessel approach we have

$$L(0) = F_{IG}(1, 1 - \frac{\delta_1}{2}, \frac{2}{y_0^1}) = 0.57$$

and thus we have

$$\mathbb{P}(L(T_0) \leq 0.57) = \mathbb{P}(T_0 \geq L^{-1}(0.57)) = \mathbb{P}(T_0 \geq 0) = 1.$$

That means the portfolio loss does not exceed 0.57 with probability 1.
Due to the common shock-like behavior of the model the probability that the common process survives the fixed time horizon is bigger than zero. Thus the probability that no default occurs at all is even higher. Since this is an unfavorable property for a large pool approximation the squared Bessel model is not suited to serve as an alternative to the Vasicek approach.

Chapter 4

Alternative models

In this chapter we present two ideas for alternative approaches which may overcome the drawbacks of the Squared Bessel model. In order to assess the quality of the models and most notably the possibility to be calibrated, further research is necessary. In this rather introductory chapter in both models we take geometric Brownian motions as a basis for the ability to pay process, since this is sufficient to explain the functionality of these approaches.

4.1 A Structural Markov Switching Model - Single Event Case

4.1.1 The Default Time

In the Squared Bessel model of the last chapter, the ability to pay processes were defined as

$$X_t^i = Y_t^i + Y_t^{T_0} \quad i \in \mathcal{I} = \{1, ..., n\}$$

with

$$\begin{aligned} Y_t^i &= Y_0^i + \int_0^t \delta_i ds + \int_0^t 2\sqrt{Y_s^i} dW_s^i \\ Y_t &= Y_0 + \int_0^t \delta ds + \int_0^t 2\sqrt{Y_s} dW_s \end{aligned}$$

defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\mathcal{F}_t = \sigma\left(((W_s^i)_{i\in\mathcal{I}}, W_s), 0 \leq s \leq t\right)$.
The default time was defined to be the first time the process X_t^i hits zero, i.e.,

$$T_i = \inf\{t \geq 0 | X_t^i = 0\}.$$

Since this stopping time can only happen after the common process Y_t has reached zero, T_i can be written as

$$T_i = T_0 + \tilde{T}_i$$

where

$$T_0 = \inf\{t \geq 0 | Y_t = 0\}$$

and

$$\tilde{T}_i = \inf\{t \geq 0 | Y_{T_0+t}^i = 0\}.$$

As already mentioned, the default event is triggered by the first hitting time of the common process. This can also be described by a single event process ξ_t with

$$\xi_t = \mathbb{1}_{\{T_0 \leq t\}}.$$

Thus a default is only possible on the event $\{\xi_t = 1\}$ and we get

$$T_i = \inf\{t \geq 0 | (Y_t^i = 0) \wedge (\xi_t = 1)\}.$$

The single event process, determined by the common process describes two possible states of the economy:

- $\{\xi_t = 0\}$ means no default is possible at all

- $\{\xi_t = 1\}$ means default is possible.

For the next approach, the underlying single event process ξ_t is not necessarily dependent on the first hitting time of a stochastic process. For sake of simplicity, the jump time of the process ξ_t will be assumed to be exponentially distributed, i.e.,

$$\xi_t = \mathbb{1}_{\{T \leq t\}}$$

with

$$\mathbb{P}(T \leq t) = 1 - \exp(-\lambda t).$$

Again the values of ξ_t represent the states of the economy, i.e., at the jump time the economy deteriorates. But in this approach the jump event does not trigger the default of an obligor. It just increases the probability of default, i.e., the main difference to the Squared Bessel model is that a default can also happen before the jump time T. The increase of the default probability at the jump time can be implemented through a change of the parameters of the asset value process. That means, the asset value processes are assumed to be geometric Brownian motions, where the drift and the volatility parameter take different values depending on the values of ξ_t, i.e.,

$$dX_t^i = \mu_i(\xi_t) X_t^i dt + \sigma_i(\xi_t) X_t^i dW_t^i$$

with

$$\begin{aligned} \mu_i(0) &= \mu_0^i \quad \mu_i(1) = \mu_1^i \\ \sigma_i(0) &= \sigma_0^i \quad \sigma_i(1) = \sigma_1^i. \end{aligned}$$

To formalize these assumptions, the Brownian motions $(W_t^i)_{i \in \mathscr{I}}$ are defined on a probability space $(\Omega_W, \mathcal{F}^W, \mathbb{P}_W)$ with the filtration $\mathcal{F}_t^W = \sigma\left((W_s^i)_{i \in \mathscr{I}}, 0 \leq s \leq t\right)$. Obviously X_t^i is not adapted to $\mathcal{F}_t^W$, thus we define another probability space $(\Omega_\xi, \mathcal{F}^\xi, \mathbb{P}_\xi)$ with the filtration $\mathcal{F}_t^\xi$, such that ξ_t is $\mathcal{F}_t^\xi$ -adapted. From these probability spaces we can define a new probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the filtration $\mathcal{F}_t$, where

- $\Omega := \Omega_W \times \Omega_\xi$

- $\mathcal{F}_t := \mathcal{F}_t^W \bigvee \mathcal{F}_t^\xi$

- $\mathbb{P} := \mathbb{P}_W \otimes \mathbb{P}_\xi$.

Since under these assumptions, X_t^i is a $\mathcal{F}_t$-adapted process, any first passage time of X_t^i will be a $\mathcal{F}_t$ stopping time, we can define the default time as the first passage time of X_t^i of a constant barrier K_i, i.e.,

$$T_i = \inf\{t \geq 0 | X_t^i \leq K_i\}.$$

In order to determine the default probability, i.e., $\mathbb{P}(T_i \leq t)$, we have to consider the cases $\{\xi_t = 0\}$ and $\{\xi_t = 1\}$, i.e., if the process ξ_t has already jumped before the time t or not. Thus, we get for the default probability

$$\begin{aligned} \mathbb{P}(T_i \leq t) &= \mathbb{E}(\mathbb{P}(T_i \leq t | \xi_t)) \\ &= \mathbb{E}\left(\mathbb{P}(T_i \leq t | \xi_t = 0)\mathbb{1}_{\{\xi_t = 0\}} + \mathbb{P}(T_i \leq t | \xi_t = 1)\mathbb{1}_{\{\xi_t = 1\}}\right) \\ &= \mathbb{E}\left(\mathbb{P}(T_i \leq t | \xi_t = 0)\mathbb{1}_{\{\xi_t = 0\}}\right) + \mathbb{E}\left(\mathbb{P}(T_i \leq t | \xi_t = 1)\mathbb{1}_{\{\xi_t = 1\}}\right). \end{aligned}$$

First, we consider the case $\xi_t = 0$:
The probability $\mathbb{P}(T_i \leq t | \xi_t = 0)$ is equal to the probability distribution function of the first passage time T_i^0 with

$$T_i^0 = \inf\{t \geq 0 | X_t^{i,0} \leq K_i\}$$

where $X_t^{i,0}$ solves the following stochastic differential equation

$$dX_t^{i,0} = \mu_i^0 X_t^{i,0} dt + \sigma_i^0 X_t^{i,0} dW_t^i$$

and

$$X_0^{i,0} = X_0^i.$$

Thus, we have

$$\begin{aligned}\mathbb{E}\left(\mathbb{P}(T_i \le t|\xi_t = 0)\mathbb{1}_{\{\xi_t=0\}}\right) &= \mathbb{E}\left(\mathbb{P}(T_i^0 \le t)\mathbb{1}_{\{\xi_t=0\}}\right)\\ &= \mathbb{P}(T_i^0 \le t)\mathbb{E}(\mathbb{1}_{\{\xi_t=0\}}) = \mathbb{P}(T_i^0 \le t)\mathbb{P}(\xi_t = 0)\\ &= \mathbb{P}(T_i^0 \le t)\mathbb{P}(T > t).\end{aligned}$$

Now we consider the case $\xi_t = 1$:
This means, the jump of ξ_t has already occurred, i.e., $T \le t$. Thus, we have to differentiate, if the default happened before the jump or vice versa. We get

$$\begin{aligned}&\mathbb{E}\left(\mathbb{P}(T_i \le t|\xi_t = 1)\mathbb{1}_{\{\xi_t=1\}}\right) = \mathbb{E}\left(\mathbb{P}(T_i \le t|T \le t)\mathbb{1}_{\{T\le t\}}\right)\\ = \;&\mathbb{E}\left(\left(\mathbb{P}(T_i \le T|T \le t) + \mathbb{P}(T < T_i \le t|T \le t)\right)\mathbb{1}_{\{T\le t\}}\right).\end{aligned}$$

On the set $\{T \le t\}$, the probability $\mathbb{P}(T_i \le T)$ is equal to the probability distribution function of T_i^0 evaluated at the random time T, whereas due to the strong Markov property of the two-dimensional process (X_t, ξ_t), the probability $\mathbb{P}(T < T_i \le t)$ equals the product of the probability $\mathbb{P}(T_i^0 > s)$ and the probability distribution function of the first passage time T_i^1 evaluated at $t - T$ with

$$T_i^1 = \inf\{t \ge 0|X_t^{i,1} \le K_i\}$$

where $X_t^{i,1}$ solves the following stochastic differential equation

$$dX_t^{i,1} = \mu_i^1 X_t^{i,1} dt + \sigma_i^1 X_t^{i,1} dW_t^i$$

and

$$X_0^{i,1} = X_T^i.$$

Thus, we have

$$
\begin{aligned}
& \mathbb{E}\left(\mathbb{P}(T_i \le t|\xi_t = 1)\mathbb{1}_{\{\xi_t=1\}}\right) = \mathbb{E}\left(\mathbb{P}(T_i \le t|T \le t)\mathbb{1}_{\{T\le t\}}\right) \\
= \ & \mathbb{E}\left(\left(\mathbb{P}(T_i^0 \le t) + \mathbb{P}(T_i^0 > s)\mathbb{P}_{X_T^i}(T_i^1 \le t - T)\right)\mathbb{1}_{\{T\le t\}}\right) \\
= \ & \mathbb{E}\left(\left(\mathbb{P}(T_i^0 \le T) + \mathbb{P}(T_i^0 > s)\int_0^\infty \mathbb{P}_y(T_i^1 \le t - T)p^i(T, X_0^i, y)dy\right)\mathbb{1}_{\{T\le t\}}\right) \\
= \ & \int_0^t \left(\mathbb{P}(T_i^0 \le s) + \mathbb{P}(T_i^0 > s)\int_0^\infty \mathbb{P}_y(T_i^1 \le t - s)p^i(s, X_0^i, y)dy\right) f_T(s)ds
\end{aligned}
$$

where $p^i(s,\cdot,\cdot)$ is the transition density function of the process X_t^i and f_T is the probability density function of the jump time T.

Altogether, we have

$$
\begin{aligned}
& \mathbb{P}(T_i \le t) \\
= \ & \mathbb{P}(T_i^0 \le t)\mathbb{P}(T > t) \\
+ \ & \int_0^t \left(\mathbb{P}(T_i^0 \le s) + \mathbb{P}(T_i^0 > s)\int_0^\infty \mathbb{P}_y(T_i^1 \le t - s)p^i(s, X_0^i, y)dy\right) f_T(s)ds.
\end{aligned}
$$

At full length, the default probability is given as

$$
\begin{aligned}
&\mathbb{P}(T_i \leq t) \\
= \; & \Bigg[\Phi\left(\frac{\log\left(\frac{K_i}{X_0^i}\right) - \left(\mu_i^0 - \frac{1}{2}(\sigma_i^0)^2\right) t}{\sigma_i^0 \sqrt{t}} \right) \\
& \quad + \left(\frac{K_i}{X_0^i}\right)^{\frac{2\mu_i^0}{(\sigma_i^0)^2} - 1} \Phi\left(\frac{\log\left(\frac{K_i}{X_0^i}\right) + \left(\mu_i^0 - \frac{1}{2}(\sigma_i^0)^2\right) t}{\sigma_i^0 \sqrt{t}} \right) \Bigg] \cdot \exp(-\lambda t) \\
+ \; & \int_0^t \Bigg(\Bigg[\Phi\left(\frac{\log\left(\frac{K_i}{X_0^i}\right) - \left(\mu_i^0 - \frac{1}{2}(\sigma_i^0)^2\right) s}{\sigma_i^0 \sqrt{s}} \right) \\
& \qquad + \left(\frac{K_i}{X_0^i}\right)^{\frac{2\mu_i^0}{(\sigma_i^0)^2} - 1} \Phi\left(\frac{\log\left(\frac{K_i}{X_0^i}\right) + \left(\mu_i^0 - \frac{1}{2}(\sigma_i^0)^2\right) s}{\sigma_i^0 \sqrt{s}} \right) \Bigg] \\
& \quad + \Bigg[1 - \Phi\left(\frac{\log\left(\frac{K_i}{X_0^i}\right) - \left(\mu_i^0 - \frac{1}{2}(\sigma_i^0)^2\right) s}{\sigma_i^0 \sqrt{s}} \right) \\
& \qquad + \left(\frac{K_i}{X_0^i}\right)^{\frac{2\mu_i^0}{(\sigma_i^0)^2} - 1} \Phi\left(\frac{\log\left(\frac{K_i}{X_0^i}\right) - \left(\mu_i^0 - \frac{1}{2}(\sigma_i^0)^2\right) s}{\sigma_i^0 \sqrt{s}} \right) \Bigg] \\
& \quad \times \int_0^\infty \Bigg[\Phi\left(\frac{\log\left(\frac{K_i}{y}\right) - \left(\mu_i^1 - \frac{1}{2}(\sigma_i^1)^2\right)(t-s)}{\sigma_i^1 \sqrt{t-s}} \right) \\
& \qquad + \left(\frac{K_i}{y}\right)^{\frac{2\mu_i^1}{(\sigma_i^1)^2} - 1} \Phi\left(\frac{\log\left(\frac{K_i}{y}\right) + \left(\mu_i^1 - \frac{1}{2}(\sigma_i^1)^2\right)(t-s)}{\sigma_i^1 \sqrt{t-s}} \right) \Bigg] \\
& \quad \frac{1}{y\sigma_i^0\sqrt{2\pi s}} \exp\left(\frac{-\left(\log\left(\frac{y}{X_0^i}\right) - \left(\mu_i^1 - \frac{1}{2}(\sigma_i^0)^2\right) s\right)^2}{2(\sigma_i^0)^2 s} \right) dy \\
& \quad \lambda \exp(-\lambda s) ds.
\end{aligned}
$$

4.1.2 Joint Default Probability

In this subsection we will determine the joint default probability of two firms. As already mentioned, the single event process ξ_t influences both firm values.

Conditioned on the process ξ_t we can consider the case that the two firms are independent of each other and the case that the underlying Brownian motions are correlated.
Analog to the one-dimensional case we get

$$\begin{aligned}
& \mathbb{P}(T_1 \leq t, T_2 \leq t) \\
= \; & \mathbb{P}(T_1^0 \leq t, T_2^0 \leq t)\mathbb{P}(T > t) \\
+ \; & \int_0^t \Bigg(\mathbb{P}(T_1^0 \leq s, T_2^0 \leq s) \\
+ \; & \int_0^\infty \int_0^\infty \mathbb{P}_{y_1,y_2}(T_1^1 \leq t-s, T_2^1 \leq t-s) \\
& \qquad \times p_s^{1,2}((X_0^1, X_0^2), (y_1, y_2)) dy_1 dy_2 \Bigg) f_T(s) ds.
\end{aligned}$$

If the two firm value processes are independent conditional on ξ_t, we have

$$\begin{aligned}
\mathbb{P}(T_1^0 \leq t, T_2^0 \leq t) &= \mathbb{P}(T_1^0 \leq t)\mathbb{P}(T_2^0 \leq t) \\
\mathbb{P}_{y_1,y_2}(T_1^1 \leq t-s, T_2^1 \leq t-s) &= \mathbb{P}_{y_1}(T_1^1 \leq t-s)\mathbb{P}_{y_2}(T_2^1 \leq t-s).
\end{aligned}$$

Due to the independence the transition probability density is given by

$$p_s^{1,2}((X_0^1, X_0^2), (y_1, y_2)) = p_s^1(X_0^1, y_1) p_s^2(X_0^2, y_2).$$

If the firm values are correlated, we can use the results from (Zhou, 1997).

4.1.3 A Hybrid Model

So far, the single event process only changed the parameters of the firm value processes at its jump time, which was interpreted as a sudden deterioration of economy. Now this concept is extended such that this event can also cause a sudden default of a firm independent of the actual value of its firm value process. That means, we have to introduce Bernoulli variables, that indicate the sudden default for every obligor and are revealed at the jump time of the single event process:

$$Y_i = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{with probability } 1 - p_i. \end{cases}$$

where every Y_i is $\mathcal{F}_T$-measurable. With the same notations as in the last subsections, we have to distinguish now between three cases for the default time:

$$T_i = \begin{cases} T_i^0 & \text{if } T_i^0 \leq T \\ T & \text{if } Y_i = 1 \\ T_i^1 + T & \text{if } Y_i = 0. \end{cases}$$

The default probability is then given by

$$\begin{aligned} & \mathbb{P}(T_i \leq t) \\ = \; & \mathbb{P}(T_i^0 \leq t)\mathbb{P}(T > t) \\ + \; & \int_0^t \Bigg(\mathbb{P}(T_i^0 \leq s) \\ & \quad + \mathbb{P}(T_i^0 > s) \left[p_i + (1 - p_i) \int_0^\infty \mathbb{P}_y(T_i^1 \leq t - s) p_s^i(X_0^i, y) dy \right] \Bigg) f_T(s) ds. \end{aligned}$$

4.2 Random drift

4.2.1 Model Description

In this chapter, the firm value process follows a geometric Brownian motion, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with

- $\Omega := \Omega_\mu \times \Omega_W$
- $\mathcal{F} := \mathcal{F}^\mu \otimes \mathcal{F}^W$
- $\mathbb{P} := \mathbb{P}_\mu \otimes \mathbb{P}_W.$

The filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is given by

$$\mathcal{F}_t = \mathcal{F}_\mu \otimes \mathcal{F}_t^W$$

where $\mathcal{F}_t^W$ is the natural filtration of the Brownian motion W_t. Furthermore, the Brownian motion W_t is independent of $\mathcal{F}_\mu$, i.e.,

$$W_t(\omega_1, \omega_2) = W_t(\omega_2) \text{ for } (\omega_1, \omega_2) \in \Omega.$$

The drift μ is a random variable on the probability space $(\Omega, \mathcal{F}^\mu, \mathbb{P}_\mu)$ and is revealed at time 0. Thus, a path of V satisfies

$$V_t(\omega_1, \omega_2) = V_0 + \int_0^t \mu(\omega_1) V_s ds + \int_0^t \sigma V_s dW_s(\omega_2).$$

The default time is again defined as the first passage time of the constant barrier K, i.e.,

$$T = \inf\{t \geq 0 | X_t \leq K\}.$$

Conditional on a realization μ_0 of the random variable μ, we get the following default probability

$$\begin{aligned} & \mathbb{P}(T \leq t | \mu = \mu_0) \\ = \quad & \Phi\left(\frac{\log\left(\frac{K}{V_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)t}{\sigma\sqrt{t}}\right) + \left(\frac{K}{V_0}\right)^{\frac{2\mu}{\sigma^2}-1} \Phi\left(\frac{\log\left(\frac{K}{V_0}\right) + \left(\mu - \frac{1}{2}\sigma^2\right)t}{\sigma\sqrt{t}}\right). \end{aligned}$$

Thus, in order to determine the default probability, we have to take the expectation with respect to the drift μ, i.e.,

$$\begin{aligned} \mathbb{P}(T \leq t) &= \mathbb{E}(\mathbb{P}(T \leq t|\mu)) \\ &= \int_{\Omega_\mu} \mathbb{P}(T \leq t|\mu(\omega_1))\mathbb{P}_\mu(d\omega_1) \\ &= \int_{-\infty}^{\infty} \mathbb{P}(T \leq t|\mu = \mu_0) f_\mu(\mu_0) d\mu_0 \end{aligned}$$

where f_μ is the probability density of the drift μ.

4.2.2 Examples

In the following examples, we will assume, that the drift μ is Gamma-distributed. The default probabilities will be calibrated to the following cumulative average default rates

years	1	2	3	4	5	6	7	8	9	10
rating										
AAA	0.00	0.00	0.03	0.06	0.10	0.17	0.25	0.38	0.43	0.48
AA	0.01	0.03	0.08	0.16	0.27	0.39	0.53	0.65	0.75	0.85
A	0.05	0.15	0.28	0.44	0.62	0.81	1.03	1.25	1.52	1.82
BBB	0.37	0.94	1.52	2.34	3.20	4.02	4.74	5.40	5.99	6.68
BB	1.38	4.07	7.16	9.96	12.34	14.65	16.46	18.02	19.60	20.82
B	6.20	13.27	19.07	23.45	26.59	29.08	31.41	33.27	34.58	35.87
CCC	27.87	36.02	41.79	46.26	50.46	52.17	53.60	54.36	56.16	57.21

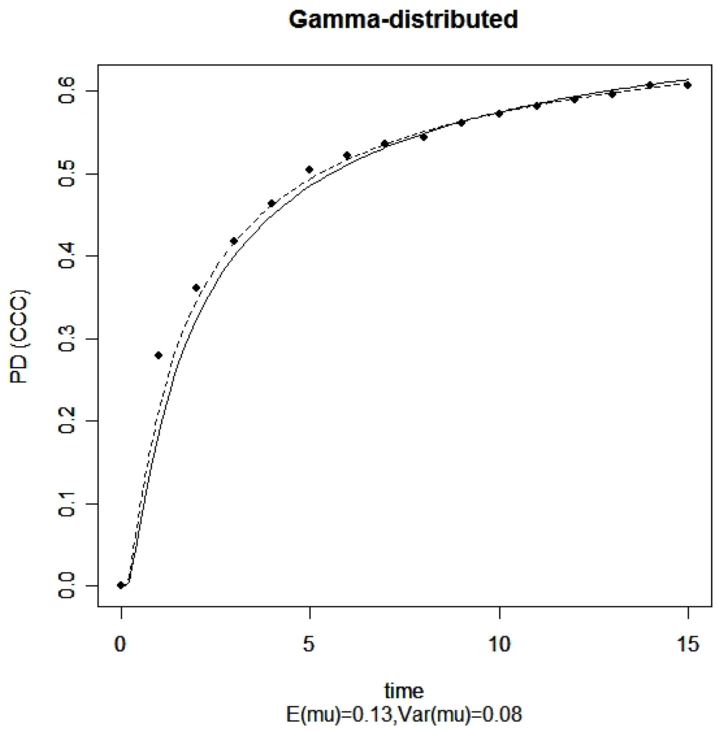
Table 4.1: Average S&P default rates [in %]

The first graph shows the average default rates for the rating CCC, the calibrated default probability for a deterministic drift with the following parameters:

$$\begin{aligned}\mu &= 0.0441\\ \sigma &= 0.0188\\ \frac{K}{V_0} &= 0.795\end{aligned}$$

The dashed line shows the default probability with random drift with the following parameters:

$$\begin{aligned}\mathbb{E}(\mu) &= 0.13\\ Var(\mu) &= 0.08\\ \sigma &= 0.22\\ \frac{K}{V_0} &= 0.78\end{aligned}$$



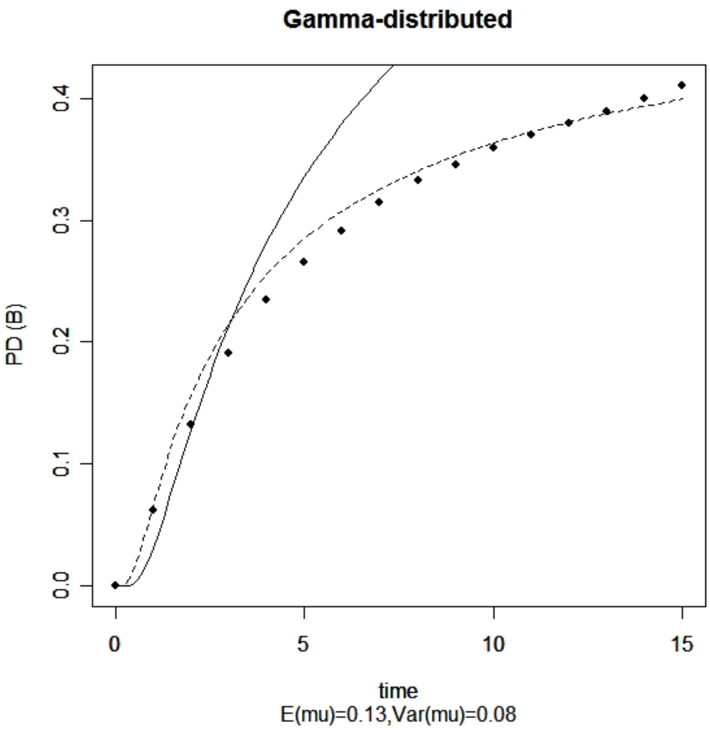
(k) PDs for CCC

The second graph shows the average default rates for the rating B, the calibrated default probability for a deterministic drift with the following parameters:

$$\begin{aligned} \mu &= 0.02 \\ \sigma &= 0.2 \\ \frac{K}{V_0} &= 0.65 \end{aligned}$$

The dashed line shows the default probability with random drift with the following parameters:

$$\begin{aligned} \mathbb{E}(\mu) &= 0.185 \\ Var(\mu) &= 0.08 \\ \sigma &= 0.22 \\ \frac{K}{V_0} &= 0.7 \end{aligned}$$



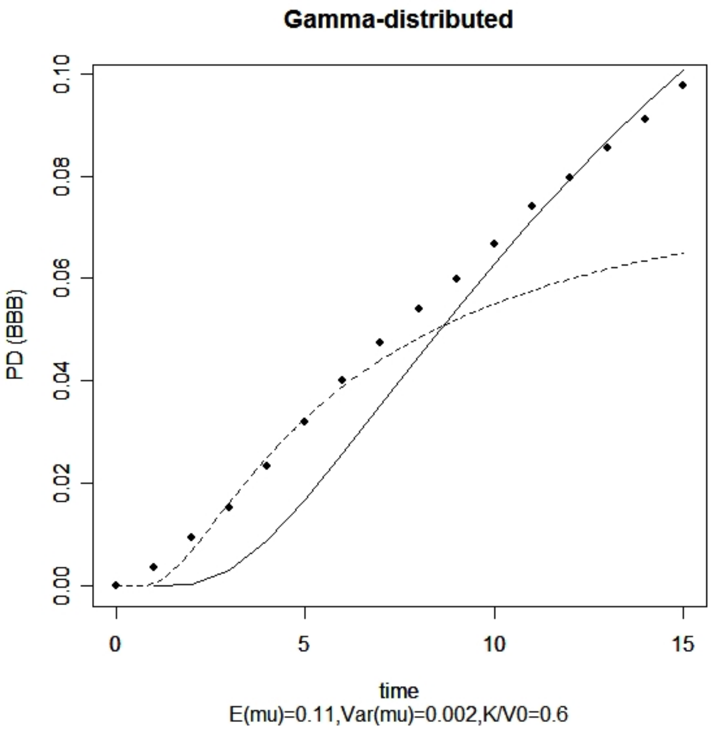
(l) PDs for B

The third graph shows the average default rates for the rating BBB, the calibrated default probability for a deterministic drift with the following parameters:

$$\begin{aligned} \mu &= 0.04 \\ \sigma &= 0.17 \\ \frac{K}{V_0} &= 0.45 \end{aligned}$$

The dashed line shows the default probability with random drift with the following parameters:

$$\begin{aligned} \mathbb{E}(\mu) &= 0.11 \\ Var(\mu) &= 0.002 \\ \sigma &= 0.17 \\ \frac{K}{V_0} &= 0.6 \end{aligned}$$



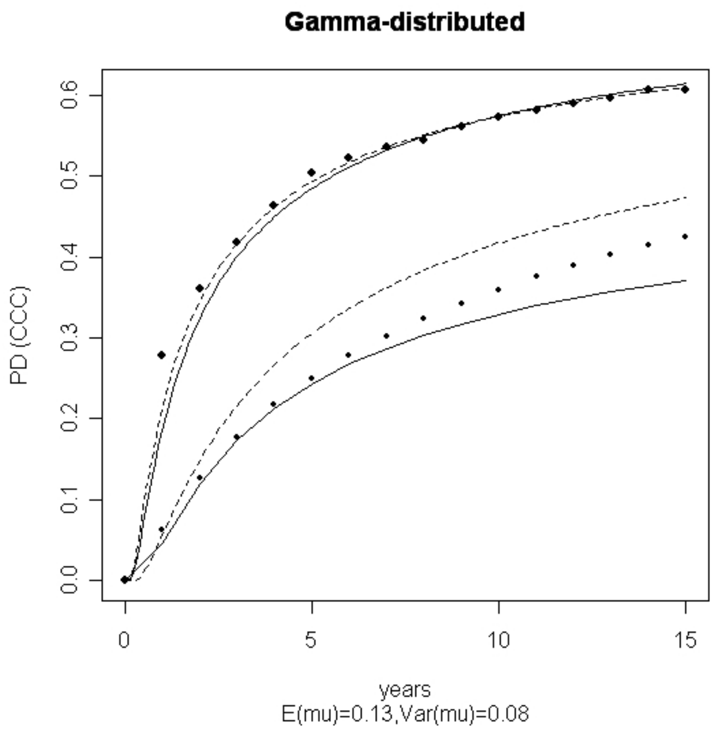
(m) PDs for BBB

4.2.3 Joint Default Probability

The random drift μ can also be used to increase the dependence between two obligors. The strongest dependence can be achieved, if the two firms are assumed to have exactly the same random drift μ. If the two firm value processes are independent conditional on the drift, we get

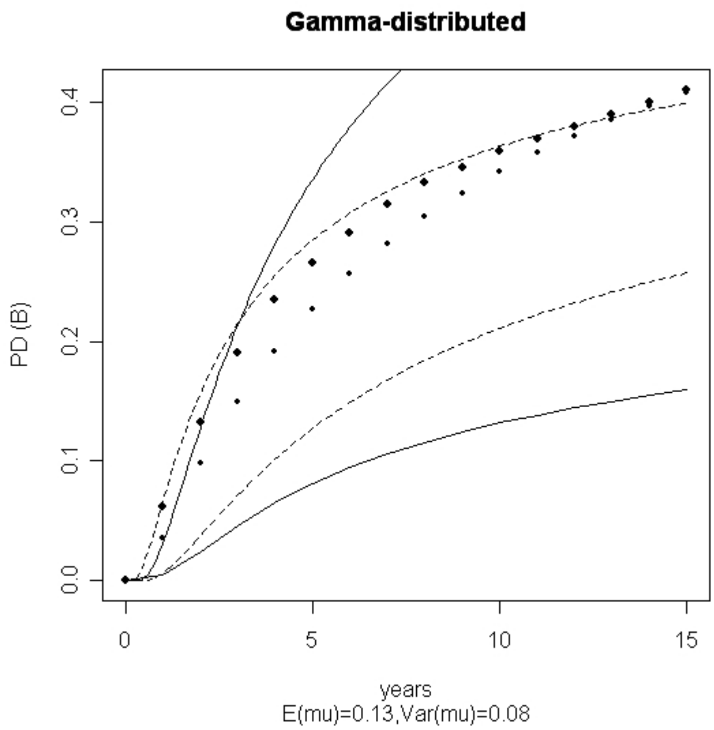
$$\begin{aligned} & \mathbb{P}(T_1 \leq t, T_2 \leq t) \\ = \; & \mathbb{E}(\mathbb{P}(T_1 \leq t, T_2 \leq t|\mu)) \\ = \; & \mathbb{E}(\mathbb{P}(T_1 \leq t|\mu)\mathbb{P}(T_1 \leq t|\mu)). \end{aligned}$$

The next two graphs show the joint probability and the default correlation for two CCC-rated firms with identical parameters and two B-rated firms as in the above examples.



(n) Joint PDs for CCC

In this figure the upper dots show the average default rates for CCC. The dashed line shows the calibrated default probabilities for a random drift and the continuous line shows the calibrated default probabilities for a deterministic drift. The lower dashed line shows the joint default probabilities for two firms with the same random drift. The dots show the associated default correlation. The continuous line shows the joint default probability in the case of independence.



(o) Joint PDs for B

In this figure the upper dots show the average default rates for B. The dashed line shows the calibrated default probabilities for a random drift and the continuous line shows the calibrated default probabilities for a deterministic drift. The lower dashed line shows the joint default probabilities for two firms with the same random drift. The dots show the associated default correlation. The continuous line shows the joint default probability in the case of independence.

Chapter 5

The Role of Copulas in the CreditRisk+™Framework

In chapter 3 we used Squared Bessel processes for our first passage time model. These processes belong to the family of square root processes, also known as Cox-Ingersoll-Ross processes. As we already mentioned, these processes are often used as intensity processes in reduced form models, see (Duffie and Singleton, 2003). In order to consider a static model, i.e., with a fixed time horizon, the Poisson process can be substituted by a Poisson variable. The intensity process can be substituted by its invariant distribution. In this case, the invariant distribution of a Cox-Ingersoll-Ross process is the Gamma distribution. This leads directly to the well-known CreditRisk+™model. The following chapter depicts and to some degree explains in greater detail the major part of an already published article (see (Ebmeyer et al., 2007)). The article was a contribution to a textbook about copulas and their application in finance, edited by Jörn Rank from d-fine, see (Rank, 2007). In our article we tried to emphasize the importance of copulas in the CreditRisk+™model and presented a kind of large portfolio approximation in order to get a measure for diversification.

5.1 Introduction

In this section we shortly present the CreditRisk+™model developed by (CSFP, 1997) in its original setting. Therefore, we consider a finite set of obligors, denoted by $\mathscr{I} = \{1, ..., m\}$, each of whom is characterized by his exposure $\tilde{\nu}_l \in (0, \infty)$ and his *expected* probability of default $\bar{p}_l \in (0, 1)$ for any $l \in \mathscr{I}$. Then the loss is expressed as the integer multiple of a basic loss unit ν_0, i.e. $\nu_l = \lfloor \frac{\tilde{\nu}_l}{\nu_0} \rfloor$. Furthermore D_l denotes obligor l's default variable, defined as

$$D_l = \begin{cases} 1 & \text{if obligor } l \text{ defaults} \\ 0 & \text{otherwise.} \end{cases}$$

So the portfolio loss variable $\tilde{L}$, expressed in multiples of the basic loss unit is given as

$$\tilde{L} := \sum_{l=1}^{m} D_l \, \nu_l, \tag{5.1}$$

For a given default probability p_l, we get for the probability generating function of D_l for $z \in [-1, 1]$

$$\mathbb{E}(z^{D_l}) = p_l z + (1 - p_l) = 1 + p_l(z - 1).$$

For small default probabilities the following approximation can be used:

$$1 + p_l(z - 1) \approx \exp(p_l(z - 1)).$$

Since this is the probability generating function of a Poisson distributed random variable, in CreditRisk+ every Bernoulli default indicator D_l is now replaced by a Poisson random variable N_l with stochastic intensity $p_l(\gamma)$. Thus, the portfolio loss variable is now being defined as

$$L := \sum_{l=1}^{m} N_l \nu_l. \tag{5.2}$$

The reason for performing this approximation is that the resulting probability generating function for the portfolio loss variable has a much simpler form than for the original loss variable. The default probability $p_l(\gamma)$ of each obligor is a random variable, whose realization depends linearly on a set of $d \geq 1$ systematic risk factors, denoted by $\gamma = (\gamma_1, ..., \gamma_d)^\top$. Each γ_k is an almost surely strictly positive random variable satisfying $\mathbb{E}(\gamma_k) = 1$ and $Var(\gamma_k) = \sigma_k^2$. We get

$$p_l(\gamma) \; = \; \bar{p}_l \; \cdot \; \sum_{k=1}^{d} w_k^l \, \gamma_k \qquad \text{for any } 1 \leq l \leq m. \tag{5.3}$$

In the standard CreditRisk+ approach the risk factors $\gamma_1, ..., \gamma_d$ have a Gamma distribution and are independent of each other.

Definition 5.1.1 *A random variable γ is called Gamma-distributed with shape parameter α and scale parameter β, i.e. $\gamma \sim \Gamma(\alpha, \beta)$ if its probability distribution function is given as follows*

$$P(\gamma \leq x) \; = \; \int_0^x \frac{1}{\beta^\alpha \, \bar{\Gamma}(\alpha)} \, t^{\alpha-1} \, e^{-\frac{t}{\beta}} \, dt, \qquad x \geq 0,$$

where $\bar{\Gamma}(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1} dx$ denotes the Gamma Function. Its expectation is given as $\mathbb{E}(\gamma) = \alpha\beta$ and its variance as $Var(\gamma) = \alpha\beta^2$

In this case, the parameters of the k-th systematic risk factor are

$$\alpha_k = \frac{1}{\sigma_k^2} \text{ and } \beta_k = \sigma_k^2.$$

As already mentioned, the concept of the probability generating function plays a central role in the CreditRisk+ model. Since the portfolio loss variable is expressed in integer multiples of the basic loss unit, its probability generating function can be represented as

$$G(z) = \sum_{n=0}^{\infty} \mathbb{P}(L = n) z^n.$$

Due to the Poisson approximation, the probability generating function can be represented in a closed analytical form, such that the coefficients can be determined with numerical algorithms. For the probability generating function of the portfolio loss we obtain

$$\begin{aligned}
\mathbb{E}(z^L) &= \mathbb{E}(\mathbb{E}(z^L|\gamma)) = \mathbb{E}\left(\mathbb{E}(z^{\sum_{l=1}^m N_l\nu_l}|\gamma)\right) \\
&= \mathbb{E}\left(\prod_{l=1}^m \mathbb{E}((z^{\nu_l})^{N_l}|\gamma)\right) = \mathbb{E}\left(\prod_{l=1}^m e^{p_l(\gamma)(z^{\nu_l}-1)}\right) \\
&= \mathbb{E}\left(e^{\sum_{l=1}^m \sum_{k=1}^d \bar{p}_l w_l^k \gamma_k (z^{\nu_l}-1)}\right) \\
&= \mathbb{E}\left(e^{\sum_{k=1}^d \mathscr{P}_k(z)\gamma_k}\right)
\end{aligned}$$

with

$$\mathscr{P}_k(z) = \sum_{l=1}^m w_k^l \bar{p}_l (z^{\nu_l} - 1) \ \ \forall\ k = 1, ..., d.$$

Then due to the independence of the systematic risk factors, we get

$$\begin{aligned}
\mathbb{E}(z^L) &= \mathbb{E}\left(e^{\sum_{k=1}^d \mathscr{P}_k(z)\gamma_k}\right) = \prod_{k=1}^d \mathbb{E}\left(e^{\mathscr{P}_k(z)\gamma_k}\right) \\
&= \prod_{k=1}^d (1 - \sigma_k^2 \mathscr{P}_k(z))^{-\frac{1}{\sigma_k^2}} \\
&= \exp\left(\log\left(\prod_{k=1}^d (1 - \sigma_k^2 \mathscr{P}_k(z))^{-\frac{1}{\sigma_k^2}}\right)\right) \\
&= \exp\left(\sum_{k=1}^d \log\left(1 - \sigma_k^2 \mathscr{P}_k(z)\right)^{-\frac{1}{\sigma_k^2}}\right) \\
&= \exp\left(-\sum_{k=1}^d \frac{1}{\sigma_k^2} \log\left(1 - \sigma_k^2 \sum_{l=1}^m w_k^l \bar{p}_l (z^{\nu_l} - 1)\right)\right). \qquad (5.4)
\end{aligned}$$

In the original document, the Panjer recursion scheme is used to compute the probabilities $\mathbb{P}(L = n)$. Since this algorithm is numerically unstable, (Haaf et al., 2004) propose a numerically stable method, where the probability generating function has to result from a series of convoluted operations on the polynomials $\mathscr{P}_k(z)$. These operations are:

- Multiplication with a scalar
- Addition of a constant

- Addition of polynomials
- log and exp

For log and exp we consider two polynomials $P(z)$ and $Q(z)$

$$P(z) = \sum_{n=0}^{\infty} p_n z^n \text{ and } Q(z) = \sum_{n=0}^{\infty} q_n z^n$$

so that $P(z) = \exp(Q(z))$. With Taylor expansion we get for the coefficients

$$p_n = \sum_{k=1}^{n} \frac{k}{n} p_{n-k} q_k \text{ and } p_0 = \exp(q_0).$$

For $Q(z) = \log(P(z))$ follows

$$q_n = \frac{1}{p_0} \left(p_n - \sum_{k=1}^{n-1} \frac{k}{n} p_{n-k} q_k \right)$$

Then the probability generating function in the form of 5.4 satisfies these conditions and thus the coefficients can be computed using the algorithm proposed in (Haaf et al., 2004).

5.2 The CreditRisk+ Default Copula

As already mentioned the default variable of obligor l in the CreditRisk+™ model is represented by a Poisson random variable N_l. Thus the default of obligor l coincides with the event $\{N_l \geq 1\}$. That means, we can identify the default indicator D_l with the indicator function $\mathbb{1}_{\{N_l \geq 1\}}$, i.e.,

$$D_l = \begin{cases} 1 & \text{if } N_l \geq 1 \\ 0 & \text{if } N_l = 0 \end{cases}$$

In this section we will determine the joint distribution of the random vector of default indicators $D = (D_1, ..., D_m)^\top$ and figure out the copula that describes its dependence structure. For this we first have to clarify the relationship of multivariate distribution functions and copulas. First of all we have to define copulas:

Definition 5.2.1 *A d-dimensional copula* $C : [0,1]^d \to [0,1]$ *is a joint distribution function such that every marginal distribution is uniform on the interval* $[0,1]$.

The relationship of multivariate distribution functions and copulas is established by the following theorem:

Theorem 5.2.2 (Sklar) *Consider a d-dimensional distribution function* F *with marginals* $F_1, ..., F_d$. *There exists a copula* C, *such that*

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)) \tag{5.5}$$

for all x_i *in* $[-\infty, \infty], i = 1, ..., d$. *If* F_i *is continuous for all* $i = 1, ..., d$ *then* C *is unique, otherwise* C *is uniquely determined only on* $RanF_1 \times ... \times RanF_d$, *where* $RanF_i$ *denotes the range of the distribution function* F_i.

On the other hand, consider a copula C *and univariate distribution functions* $F_1, ..., F_d$. *Then* F *as defined in equation 5.5 is a multivariate distribution function with marginals* $F_1, ..., F_d$.

An important family of copulas are the Archimedean copulas. A n-dimensional Archimedean copula can be written as follows:

$$C(u_1, ..., u_n) = \phi^{-1}\left(\sum_{i=1}^{n} \phi(u_i)\right)$$

for all $0 \leq u_1, ..., u_n \leq 1$ and the function ϕ satisfies

- $\phi(1) = 0$
- $\forall u \in (0,1)\ \phi'(u) < 0$,i.e., ϕ is decreasing
- $\forall u \in (0,1)\ \phi''(u) \geq 0$,i.e., ϕ is convex

The function ϕ is called generator.
A bivariate example of an Archimedean copula is the Clayton copula. The generator of the Clayton copula is given by

$$\phi(x) = \frac{1}{\Theta}\left(x^{-\Theta} - 1\right) \text{ with } \Theta > 0$$

Then we have

$$\phi^{-1}(x) = (\Theta x + 1)^{-\frac{1}{\Theta}}$$

and the resulting bivariate Clayton copula is

$$C(u_1, u_2) = \left(u_1^{-\Theta} + u_2^{-\Theta} - 1\right)^{-\frac{1}{\Theta}}$$

A short introduction to copulas is given in (Schmidt, 2007). For a deeper survey we refer to (Nelsen, 2006).

Now we return to the default indicator D_l. Conditioned on the set of systematic factors γ we get

$$\mathbb{P}(D_l = 0|\gamma) = \mathbb{P}(N_l = 0|\gamma) = e^{-p_l(\gamma)}$$

and thus

$$\begin{aligned}\mathbb{P}(D_l = 1|\gamma) &= \mathbb{P}(N_l \geq 1|\gamma) = 1 - \mathbb{P}(N_l = 0|\gamma) \\ &= 1 - e^{-p_l(\gamma)}\end{aligned}$$

Here we consider just one systematic risk factor, i.e., $d = 1$, so we have

$$P(D_l = 1|\gamma) = 1 - e^{-\bar{p}_l \gamma}.$$

It follows

$$\begin{aligned}\mathbb{P}(D_l = 1) &= \mathbb{E}(\mathbb{P}(D_l = 1|\gamma)) = 1 - \mathbb{E}(\mathbb{P}(D_l = 0|\gamma)) \\ &= 1 - \mathbb{E}(e^{-\bar{p}_l\gamma}) = 1 \ - \int_0^\infty e^{-\bar{p}_l \ x}\, g_{\alpha,\beta}(x)\, dx \\ &= 1 \ - \ \mathscr{L}_\gamma(\bar{p}_l) \qquad (5.6)\end{aligned}$$

where $\mathscr{L}_\gamma(\bar{p}_l)$ denotes the Laplace Transform of the random variable γ evaluated at $\bar{p}_l$. Now let F_l denote the distribution function of D_l, i.e., $F_l(x) = \mathbb{P}(D_l \leq x)$. For $x \geq 1$ it is obvious that

$$F_l(x) = \mathbb{P}(D_l \leq x) = 1.$$

For $0 \leq x < 1$ we get

$$F_l(x) = \mathbb{P}(D_l \leq x) = \mathbb{P}(D_l = 0) = \mathscr{L}_\gamma(\bar{p}_l)$$

Thus

$$F_l(x) \;=\; \begin{cases} \mathscr{L}_\gamma(\bar{p}_l) & \text{for } 0 \le x < 1 \\ 1 & \text{for } x \ge 1. \end{cases}$$

Then the joint distribution function $F : [0,\infty)^n \to [0,1]$ of the indicator random vector $D = (D_1, ..., D_m)^\top$ is given as

$$\begin{aligned} F(x_1, ..., x_m) &= \mathbb{P}(D_1 \le x_1, ..., D_l \le x_m) \\ &= \mathbb{E}(\mathbb{P}(D_1 \le x_1, ..., D_m \le x_m|\gamma)) \\ &= \mathbb{E}(\mathbb{P}(D_1 \le x_1|\gamma) \cdot ... \cdot \mathbb{P}(D_m \le x_m|\gamma)) \\ &= \mathbb{E}\left(e^{-\bar{p}_1 \mathbb{1}_{[0,1[}(x_1)\gamma} \cdot ... \cdot e^{-\bar{p}_m \mathbb{1}_{[0,1[}(x_m)\gamma}\right) \\ &= \mathbb{E}\left(e^{(-\sum_{l=1}^m \bar{p}_l \mathbb{1}_{[0,1[}(x_l))\,\gamma}\right) \\ &= \mathscr{L}_\gamma\left(\sum_{l=1}^m \bar{p}_l \mathbb{1}_{[0,1[}(x_l)\right). \end{aligned}$$

Now for $0 \le x_l < 1$ we have

$$\bar{p}_l = \mathscr{L}_\gamma^{-1}\left(\mathscr{L}_\gamma(\bar{p}_l)\right) = \mathscr{L}_\gamma^{-1}\left(F_l(x_l)\right)$$

and since $\mathscr{L}_\gamma(0) = 1$ for $x_l \ge 1$ we get

$$0 = \mathscr{L}_\gamma^{-1}(1) = \mathscr{L}_\gamma^{-1}\left(F_l(x_l)\right)$$

and thus

$$F(x_1, ..., x_m) = \mathscr{L}_\gamma\left(\sum_{l=1}^m \mathscr{L}_\gamma^{-1}(F_l(x_l))\right)$$

Since $\mathscr{L}_\gamma(\cdot)$ is a convex, decreasing function satisfying $\mathscr{L}_\gamma^{-1}(1) = 0$ we have

$$\begin{aligned} F(x_1, ..., x_m) &= \phi^{-1}\left(\sum_{l=1}^m \phi(F_l(x_l))\right) \\ &= C(F_1(x_1), ..., F_m(x_m)) \end{aligned}$$

with an Archimedean copula C with generator $\phi(\cdot) \equiv \mathscr{L}_\gamma^{-1}(\cdot)$. Since

$$\mathscr{L}_\gamma(x) \;=\; (1 \,+\, \beta\, x)^{-\frac{1}{\beta}},$$

we see that $\mathscr{L}_\gamma(x)$ is the inverse of the generator of the Clayton copula. Thus we may write the joint default copula of the CreditRisk+™ model as

$$C(u_1, ..., u_m) \;=\; \left(1 \;-\; m \;+\; \sum_{i=1}^{m} u_i^{-\beta}\right)^{-1/\beta}$$

which is the m - dimensional extension of the Clayton copula, or the *Cook–Johnson copula*.

5.3 The Compound Gamma Model

In the original setup the systematic risk factors are assumed to be independent so that default correlations between different obligors may only implicitly be defined by assigning positive weights to one or more common risk factors. In this section, an approach proposed by (Giese, 2004) is presented, where this shortcoming is overcome. Here, the systematic risk factors are assumed to be Compound Gamma distributed. That means, conditional on some realization $\{\gamma^0 = s\}$ of an additional Gamma-distributed random variable γ^0, the factors $\gamma_1, ..., \gamma_d$ are independent and $\Gamma(\alpha_k s, \beta_k)$-distributed for $k = 1, ..., d$. Thus, a Compound Gamma distributed random vector is defined as follows

Definition 5.3.1
A nonnegative d - dimensional random vector $\gamma = (\gamma_1, ..., \gamma_d)^\top$ is called **Compound Gamma distributed***, if its distribution function is absolutely continuous with density*

$$f_\gamma(t_1, ..., t_d) \;=\; \int_0^\infty \prod_{k=1}^{d} g_{\alpha_k \cdot s, \beta_k}(t_k)\; g_{\alpha^0, \beta^0}(s)\; ds, \quad (t_1, ..., t_d) \in \mathbb{R}_+^d,$$

where each $g_{\alpha_k, \beta_k}(\cdot)$ denotes the density of a Gamma variate with parameters $(\alpha_k, \beta_k) \in (0, \infty)^2$ satisfying $\alpha_k = 1/\beta_k$ for $k = 0, ..., d$.

Again we set

$$\mathbb{E}(\gamma_k) = 1 \text{ for } k = 1, ..., d$$

such that $\alpha_k = \frac{1}{\beta_k}$. Then the following lemma delivers the covariance matrix of a Compound Gamma distributed random vector.

Lemma 5.3.2 *The covariance matrix $(C^\gamma)_{k,l=1,...,d}$ for a Compound Gamma distributed random vector $(\gamma_1, ..., \gamma_d)$ is given as follows*

$$C^\gamma_{k,l} = Cov(\gamma_k, \gamma_l) = \delta_{k,l}\beta_k + \beta^0 \quad \forall k, l = 1, ..., d$$

where $\delta_{k,j}$ denotes Kronecker's delta function

$$\delta_{k,j} = \begin{cases} 1 & \text{for } k = j \\ 0 & \text{for } k \neq j. \end{cases}$$

Proof: We have

$$Cov(\gamma_k, \gamma_l) = \mathbb{E}(Cov(\gamma_k, \gamma_l|\gamma^0)) + Cov(\mathbb{E}(\gamma_k|\gamma^0), \mathbb{E}(\gamma_l|\gamma^0))$$

Since the risk factors are conditionally independent, we get

$$\mathbb{E}(Cov(\gamma_k, \gamma_l|\gamma^0)) = 0$$

and thus for $k \neq l$

$$\begin{aligned} Cov(\gamma_k, \gamma_l) &= Cov(\mathbb{E}(\gamma_k|\gamma^0), \mathbb{E}(\gamma_l|\gamma^0)) \\ &= Cov(\alpha_k\gamma^0\beta_k, \alpha_l\gamma^0\beta_l) \\ &= \underbrace{\alpha_k\beta_k}_{=1}\underbrace{\alpha_l\beta_l}_{=1} Cov(\gamma^0, \gamma^0) \\ &= Var(\gamma^0) = \beta^0. \end{aligned}$$

For $k = l$ we get

$$\begin{aligned} Var(\gamma_k) &= \mathbb{E}(Var(\gamma_k|\gamma_0)) + Var(\mathbb{E}(\gamma_k|\gamma_0)) \\ &= \mathbb{E}(\alpha_k\gamma_0\beta_k^2) + Var(\alpha_k\gamma_0\beta_k) \\ &= \alpha_k\beta_k^2\mathbb{E}(\gamma_0) + \alpha_k^2\beta_k^2 Var(\gamma_0) \\ &= \beta_k + \beta_0. \end{aligned}$$

□

According to definition 5.3.1 the marginal distribution function $F_{\alpha_k,\beta_k}(\cdot)$ of some single systematic risk factor γ_k is given as

$$F_{\alpha_k,\beta_k}(x) \;:=\; \int_0^x \int_0^\infty g_{\alpha_k\, s,\beta_k}(t)\; g_{\alpha_0,\beta_0}(s)\, ds\; dt \quad \text{for any } x \in [0,\infty). \tag{5.7}$$

Denoting the joint distribution of $(\gamma_1, ..., \gamma_d)$ by F_γ, i.e.

$$F_\gamma(x_1, ..., x_d) \;=\; \int_0^{x_1} \cdots \int_0^{x_d} f_\gamma(t_1, ..., t_d)\, dt_1 \;\cdots\; dt_d,$$

then the copula Function describing the joint distribution is given as

$$C(u_1, ..., u_d) \;=\; F_\gamma(F^{-1}_{\alpha_1,\beta_1}(u_1), ..., F^{-1}_{\alpha_d,\beta_d}(u_d)), \quad (u_1, ..., u_d) \in [0,1]_d.$$

We shall henceforth impose the following assumption:

Assumption 5.3.3
The default variables N_l, $l \in \{1, ..., m\}$ are pairwise independent conditional on $\sigma(\gamma)$, i.e., the σ–algebra generated by $\gamma = (\gamma_1, ..., \gamma_d)^\top$.

The portfolio loss distribution can again be computed with the recursion scheme introduced in (Haaf et al., 2004). Therefore, the probability generating function must have a similar form as in 5.4. In this case, we get

$$\begin{aligned}
\mathbb{E}(z^L) &= \mathbb{E}\left(e^{\sum_{k=1}^d \mathscr{P}_k(z)\gamma_k}\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(e^{\sum_{k=1}^d \mathscr{P}_k(z)\gamma_k}|\gamma_0\right)\right) \\
&= \mathbb{E}\left(\prod_{k=1}^d \mathbb{E}\left(e^{\mathscr{P}_k(z)\gamma_k}|\gamma_0\right)\right) \\
&= \mathbb{E}\left(\prod_{k=1}^d (1-\beta_k\mathscr{P}_k(z))^{-\frac{\gamma_0}{\beta_k}}\right) \\
&= \mathbb{E}\left(\exp\left(\log\left(\prod_{k=1}^d (1-\beta_k\mathscr{P}_k(z))^{-\frac{\gamma_0}{\beta_k}}\right)\right)\right) \\
&= \mathbb{E}\left(\exp\left(\gamma_0 \log\left(\prod_{k=1}^d (1-\beta_k\mathscr{P}_k(z))^{-\frac{1}{\beta_k}}\right)\right)\right).
\end{aligned}$$

This is the moment generating function of the gamma distributed random variable γ^0 evaluated at

$$\log\left(\prod_{k=1}^{d}(1-\beta_k\mathscr{P}_k(z))^{-\frac{1}{\beta_k}}\right).$$

Thus

$$\begin{aligned}
&\mathbb{E}\left(\exp\left(\gamma^0\log\left(\prod_{k=1}^{d}(1-\beta_k\mathscr{P}_k(z))^{-\frac{1}{\beta_k}}\right)\right)\right)\\
=&\left(1-\beta_0\log\left(\prod_{k=1}^{d}(1-\beta_k\mathscr{P}_k(z))^{-\frac{1}{\beta_k}}\right)\right)^{-\frac{1}{\beta_0}}\\
=&\exp\left(-\frac{1}{\beta_0}\log\left(1+\beta_0\sum_{k=1}^{d}\frac{1}{\beta_k}\log\left(1-\beta_k\mathscr{P}_k(z)\right)\right)\right).
\end{aligned}$$

5.4 The Characteristic Function

The purpose of this section is to derive the Characteristic Function $\phi(\cdot)$ of the loss variable defined in (5.2). Note that by its very definition, we have

$$\phi(t) := \mathbb{E}\left(e^{it\,L}\right) \qquad \text{for any } t\in\mathbb{R}.$$

To evaluate the above integral, it is convenient to first consider the *conditional* Characteristic Function $\phi_\gamma(\cdot)$ of L, defined as

$$\phi_\gamma(t) := \mathbb{E}\left(e^{it\,L}\,\middle|\,\gamma\right) \qquad \text{for any } t\in\mathbb{R}.$$

To this end, fix some $t\in\mathbb{R}$ and set

$$z := e^{it}.$$

By successively exploiting the conditional independence of the Poisson default variables, we may rewrite $\phi_\gamma(t)$ as

$$\begin{aligned}
\phi_\gamma(t) \;=\; \mathbb{E}\left(z^L \,|\gamma\right) &= \mathbb{E}\left(z^{\sum_{l=1}^m N_l \nu_l} \,\middle|\, \gamma \right) \\
&= \mathbb{E}\left(\prod_{l=1}^m (z^{\nu_l})^{N_l} \,\middle|\, \gamma \right) \\
&= \prod_{l=1}^m \mathbb{E}\left((z^{\nu_l})^{N_l} \,\middle|\, \gamma \right) \\
&= \prod_{l=1}^m \; e^{p_l(\gamma)(z^{\nu_l}-1)} \\
&= e^{\sum_{l=1}^m (\sum_{k=1}^d \bar{p}_l w_k^l \gamma_k)(z^{\nu_l}-1)} \\
&= e^{\sum_{k=1}^d \mathscr{P}^k(z)\,\gamma^k}, \qquad (5.8)
\end{aligned}$$

where

$$\mathscr{P}^k(z) \;:=\; \sum_{l=1}^m w_k^l\, \bar{p}_l \,(z^{\nu_l} - 1) \qquad \text{for any } k \in \{1, ..., d\}.$$

The above defined polynomial $\mathscr{P}^k(\cdot)$ will henceforth be called the *sector polynomial* of sector k. It is merely an elegant and condensed way of summarizing the possibly highly complex exposure distribution within a given sector. To obtain the unconditional Characteristic Function, we merely integrate over all realizations of γ, or formally,

$$\begin{aligned}
\phi(t) &= \int_{\mathbb{R}_+^d} \phi_{s^1,...,s^d}(t)\; f_\gamma(s^1, ..., s^d)\; ds^1 \cdots ds^d \\
&= \int_{\mathbb{R}_+^d} \left(e^{\sum_{k=1}^d \mathscr{P}^k(z)\, s^k} \right)\; f_\gamma(s^1, ..., s^d)\, ds^1 \cdots ds^d.
\end{aligned}$$

We shall summarize the above sequence of equations with the following Lemma:

Lemma 5.4.1 *The Characteristic Function $\phi(\cdot)$ of the portfolio loss variable L is given as*

$$\begin{aligned}\phi(t) &= \mathbb{E}\left(e^{it\,L}\right)\\ &= \mathbb{E}\left(e^{\sum_{k=1}^{d}\mathscr{P}^k(e^{it})\,\gamma^k}\right)\\ &= \int\limits_{\mathbb{R}_+^d}\left(e^{\sum_{k=1}^{d}\mathscr{P}^k(e^{it})\,s^k}\right)\ f_\gamma(s^1,...,s^d)\,ds^1\cdots ds^d.\end{aligned}$$

5.5 The Limit Distribution

The purpose of this section is to examine the limiting loss distribution which results after a specific replication process. In the n - th stage of replication, we consider the same types of obligors, but instead of a single obligor of a certain type, there are n obligors, each of whom has an exposure equal to $1/n$ - times his original exposure. Each clone of a particular obligor is considered as being a conditionally independent copy of the original obligor and is associated with the same expected probability of default and the same sector allocation. The resulting sequence of portfolios is often associated with diversification, since concentration of exposures in single obligors is reduced step by step and the idiosyncratic part of the portfolio risk is slowly diversified away.

Formally, we are interested in the limiting distribution of the sequence of random variables $(S_n)_{n\in\mathbb{N}}$ defined as

$$S_n \;:=\; \frac{1}{n}\sum_{j=1}^{n} L_j \qquad \text{for any } n\geq 1, \tag{5.9}$$

where each L_j is a precise copy of the original portfolio L. In order to study the limiting behavior of the distribution corresponding to each S_n, we shall make use of the following result, known as Lévy's Continuity Theorem:

Lemma 5.5.1 (Lévy's Continuity Theorem)
Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of distribution functions, whose associated characteristic functions are given respectively as $(\varphi_n(\cdot))_{n\in\mathbb{N}}$. If $(\varphi_n(\cdot))_{n\in\mathbb{N}}$ converges

pointwise to some function $\varphi^*(\cdot)$ and if $\varphi^*(\cdot)$ is continuous at $t = 0$, then there exists a distribution function F, such that $F_n \longrightarrow F$ weakly as $n \to \infty$ and the characteristic function associated with F is $\varphi^*(\cdot)$.

For a complete proof see for example (Feller, 1971). As mentioned before, after exactly n portfolio replications, the corresponding portfolio S_n consists exactly of n clones of the original set of obligors, each of whom has a credit exposure equal to $1/n$-times his original exposure. More precisely, instead of m obligors with exposures $\{\nu_1, ..., \nu_m\}$ as considered in the original portfolio L, the portfolio S_n consists of $n \cdot m$ obligor with exposures

$$\Big\{\underbrace{\frac{\nu_1}{n}, ..., \frac{\nu_1}{n}}_{n\ times}, \dots, \underbrace{\frac{\nu_m}{n}, ..., \frac{\nu_m}{n}}_{n\ times}\Big\}.$$

Each of the first n clones of some obligor l in the original portfolio L has the same probability of default $p_l(\gamma)$, with

$$p_l(\gamma) \;=\; \bar{p}_l \;\cdot\; \sum_{k=1}^{d} w_k^l\, \gamma_k,$$

that is, each clone underlies the same systematic risk factors as the corresponding obligor in the original portfolio. If we wish to set up the characteristic function of S_n, then we first have to determine the corresponding sector polynomials. For sector k, it is obviously given as $\mathscr{P}_n^k(z)$, where

$$\begin{aligned}
\mathscr{P}_n^k(z) \;&=\; \sum_{j=1}^{n} \sum_{l=1}^{m} w_l^k\, \bar{p}_l\, (z^{\frac{1}{n}\nu_l} \,-\, 1) \\
&=\; \sum_{l=1}^{m} n\, w_l^k\, \bar{p}_l\, (z^{\frac{1}{n}\nu_l} \,-\, 1) \qquad \text{for any } z \in \mathbb{C}.
\end{aligned} \tag{5.10}$$

Just as in the case of the original portfolio, we may set up the characteristic function $\phi_n(\cdot)$ of the portfolio S_n. For any given $n \in \mathbb{N}$, it is given as

$$\phi_n(t) \;=\; \mathbb{E}\left(e^{it\, S_n}\right) \tag{5.11}$$

$$=\; \int_{\mathbb{R}_+^d} \left(e^{\left(\sum_{k=1}^{d} \mathscr{P}_n^k(e^{it})\; s_k\right)}\; f_\gamma(s_1, ..., s_d)\right)\; ds_1 \,\cdots\, ds_d$$

for any given $t \in \mathbb{R}$.
In order to study the pointwise limit of $(\phi_n(t))_{n\in\mathbb{N}}$, note first that the expansion of the complex exponential function reads

$$e^{it} = \sum_{j=0}^{\infty} \frac{(it)^j}{j!} \quad \text{for any } t \in \mathbb{R},$$

so that in particular

$$n\,(e^{it\frac{\nu_l}{n}} - 1) = n\left((1 + it\frac{\nu_l}{n} - 1) + \mathscr{O}(\frac{1}{n^2})\right) \quad \text{for any } t \in \mathbb{R}.$$

It follows that

$$\lim_{n\to\infty} n\,(e^{it\frac{\nu_l}{n}} - 1) = i\,t\,\nu_l \quad \text{for any given } t \in \mathbb{R}. \tag{5.12}$$

It is now straightforward to write down the pointwise limit of the sequence of sector polynomials:

Lemma 5.5.2
For any $t \in \mathbb{R}$ and $k \in \{1, ..., d\}$, the sequence $(\mathscr{P}_n^k(e^{it}))_{n\in\mathbb{N}}$ converges in $\mathbb{C}$ to $EL_k \cdot it$ where EL_k is defined as follows

$$\begin{aligned} EL_k &:= \mathbb{E}(\sum_{l=1}^{m} w_l^k\, N_l\, \nu_l) = \sum_{l=1}^{m} w_l^k\, \nu_l\, \mathbb{E}[\,\mathbb{E}(N_l \,|\, \sigma(\gamma))\,] \\ &= \sum_{l=1}^{m} w_l^k\, \nu_l\, \mathbb{E}(\,p_l(\gamma)\,) = \sum_{l=1}^{m} w_l^k\, \bar{p}_l\, \nu_l. \end{aligned} \tag{5.13}$$

Proof: Fix $k \in \{1, ..., d\}$ and $t \in \mathbb{R}$ and note that

$$\begin{aligned} \lim_{n\to\infty} \mathscr{P}_n^k(e^{it}) &= \lim_{n\to\infty} \sum_{l=1}^{m} w_k^l\, \bar{p}_l\, n(e^{it\,\frac{\nu_l}{n}} - 1) \qquad (5.14) \\ &= \sum_{l=1}^{m} w_k^l\, \bar{p}_l \lim_{n\to\infty} n\,(e^{it\,\frac{\nu_l}{n}} - 1) \\ &= \sum_{l=1}^{m} w_k^l\, \bar{p}_l\, \nu_l\, it \\ &= EL_k \cdot it. \end{aligned}$$

□

The above Lemma together with the Dominated Convergence Theorem now quite naturally lead to the limiting distribution of the sequence $(S_n)_{n\in\mathbb{N}}$:

Proposition 5.5.3 (Limit Distribution)
The sequence $(S_n)_{n\in\mathbb{N}}$ defined in (5.9) converges in distribution to the random variable

$$S^* \; := \; \sum_{k=1}^{d} \; EL_k \, \cdot \, \gamma_k. \tag{5.15}$$

That is, for any $x \in (0,\infty)$, we have

$$\lim_{n\to\infty} \; P(S_n \, \leq \, x) \; = \; F(x) \; := \; \int\limits_{\mathscr{D}(x)\cap\mathbb{R}^d_+} \; f_\gamma(s_1, ..., s_d) \; ds_1 \cdots ds_d,$$

where the set $\mathscr{D}(x)$ denotes the hyperplane in $\mathbb{R}^d$ defined as

$$\mathscr{D}(x) \; := \; \left\{ \; s = (s_1, ..., s_d)^\top \, \in \, \mathbb{R}^d \; \middle| \; \sum_{k=1}^{d} \, s_k \, \cdot \, EL_k \, \leq \, x \; \right\}.$$

Proof: Applying equation (5.8) to the portfolio S_n entails that

$$\mathbb{E}\left(e^{it\,S_n} \,\middle|\, \gamma \right) \; = \; e^{\sum_{k=1}^{d} \mathscr{P}_n(e^{it})\,\gamma_k}.$$

for any $t \in \mathbb{R}$. Moreover, according to Lemma 5.5.2, we have

$$\lim_{n\to\infty} \mathscr{P}^k_n(e^{it}) \; = \; EL_k \, it$$

for any k, which, by the continuity of the complex exponential function, implies that

$$\lim_{n\to\infty} \mathbb{E}\left(e^{it\,S_n} \,\middle|\, \gamma \right) \; = \; \lim_{n\to\infty} \; e^{\sum_{k=1}^{d} \mathscr{P}^k_n(e^{it})\,\gamma_k} \; = \; e^{it\,\sum_{k=1}^{d} EL_k\,\gamma_k}.$$

Next, note that since $|e^{it\,S_n}| \leq 1$, we also have

$$\left| \mathbb{E}\left(e^{it\,S_n} \,\middle|\, \gamma \right) \right| \; \leq \; 1,$$

so that the Dominated Convergence Theorem implies that

$$\begin{aligned}\lim_{n\to\infty} \phi_n(t) &= \lim_{n\to\infty} \mathbb{E}\left(\mathbb{E}\left(e^{it\,S_n} \,\middle|\, \gamma \right)\right) \\ &= \mathbb{E}\left(\lim_{n\to\infty} \mathbb{E}\left(e^{it\,S_n} \,\middle|\, \gamma \right)\right) \\ &= \mathbb{E}\left(e^{it\,\sum_{k=1}^{d} EL_k\,\gamma_k} \right)\end{aligned}$$

which is the characteristic function of the random variable S^* defined in (5.15). □

Thus, the entire remaining risk in the limiting portfolio is of systematic nature and is solely driven by the joint distribution of the sector risk factors $\gamma_1, ..., \gamma_d$. The case $d = 1$ is particularly intuitive:

$$\lim_{n\to\infty} P\left(\frac{S_n}{EL} \le x \right) = F_{\alpha_1,\beta_1}(x).$$

Put differently, as the degree of diversification increases step by step, the distribution of the random variable S_n/EL approaches the distribution of the underlying systematic risk factor. This may alternatively be formulated as follows:

Corollary 5.5.4
Referring to the case $d = 1$, for any given confidence level $\alpha \in (0,1)$ we have

$$\lim_{n\to\infty} P\left(S_n \le \mathbb{E}(S_n \,|\, \gamma_1 = q_\alpha(\gamma_1))\right) = \alpha, \tag{5.16}$$

where

$$q_\alpha(\gamma_1) := \inf\{x \ge 0 \mid P(\gamma_1 \le x) \ge \alpha\}$$

denotes the α - quantile of γ_1.

That is, in a large and well-diversified portfolio, the expected loss of the portfolio times the quantile of the underlying factor is already a good estimate for the quantile of the corresponding portfolio, or formally

$$EL \cdot q_\alpha(\gamma_1) \approx q_\alpha(L).$$

Equivalently, the fraction $q_\alpha(L)/EL$ is equal to the quantile of the univariate risk factor distribution. The next Corollary refers to the remaining risk in the portfolio as measured by its variance:

Corollary 5.5.5
The variance $V(S^*)$ of the limiting portfolio S^* is given as

$$V(S^*) \;=\; \sum_{k,j=1}^{d} EL_k \; EL_j \; Cov(\gamma_k, \gamma_j)$$

The proof follows immediately from the representation (5.15).

5.5.1 The Case of Bernoulli - Default Variables

The purpose of this section is to deliver the arguments for the foregoing convergence result to hold if we replace the assumption of a Poisson distributed default variable by the assumption of a Bernoulli default variable. Recall that the above convergence result was formulated for the portfolio loss variable

$$L \;=\; \sum_{l=1}^{m} N_l \; \nu_l$$

where each N_l is Poisson distributed with stochastic intensity $p_l(\gamma)$. As previously remarked this approximation is made for computational reasons, since the probability generating function has a particularly simple form. The above convergence result does not depend however on this choice. Since the detailed arguments are quite similar to the Poisson case, we will only sketch the results. So consider the random variable

$$\tilde{L} \;=\; \sum_{l=1}^{m} D_l \; \nu_l$$

where D_l denotes the Bernoulli indicator variable

$$D_l \;=\; \begin{cases} 1 & \text{if obligor } l \text{ defaults} \\ 0 & \text{otherwise,} \end{cases}$$

with $\mathbb{E}(D_l|\gamma) = p_l(\gamma)$ for any $l \in \mathscr{I}$. Henceforth, we shall make the following assumption which is supposed to parallel Assumption 5.3.3:

Assumption 5.5.6
The default variables D_l, $l \in \{1, ..., m\}$ *are pairwise independent conditional on* $\sigma(\gamma)$.

For a single obligor l, the conditional characteristic function $\tilde{\phi}^l_\gamma(\cdot)$ reads

$$\begin{aligned} \tilde{\phi}^l_\gamma(t) \;&:=\; \mathbb{E}\left(e^{it\, D_l\, \nu_l} \,\middle|\, \gamma \right) \\ &=\; p_l(\gamma)\, e^{it\,\nu_l} \;+\; 1 \;-\; p_l(\gamma). \end{aligned} \tag{5.17}$$

for any $t \in \mathbb{R}$. By Assumption 5.5.6, the single default variables D_l are conditionally independent, so that the conditional characteristic function of the loss variable $\tilde{L}$ is given as

$$\tilde{\phi}_\gamma(t) \;:=\; \mathbb{E}\left(e^{it\,\tilde{L}} \,\middle|\, \gamma \right) \;=\; \prod_{l=1}^{m} \tilde{\phi}^l_\gamma(t) \quad \forall\, t \in \mathbb{R}.$$

Thus, the unconditional characteristic function $\tilde{\phi}(\cdot)$ is the integral of the above expression and given as

$$\begin{aligned} \tilde{\phi}(t) \;&:= \mathbb{E}\left(e^{it\,\tilde{L}} \right) \\ = \int\limits_{\mathbb{R}^d_+} \prod_{l=1}^{m} &\left(1 + (e^{it\nu_l} - 1)\bar{p}_l \sum_{k=1}^{d} w^l_k s_k \right) f_\gamma(s_1, ..., s_d) ds_1 \cdots ds_d. \end{aligned}$$

As in the previous section, we are interested in the limiting distribution of the sequence $(\tilde{S})_{n\in\mathbb{N}}$, defined as

$$\tilde{S}_n \;:=\; \frac{1}{n} \sum_{j=1}^{n} \tilde{L}_j \quad \text{for any } n \in \mathbb{N}, \tag{5.18}$$

where each $\tilde{L}_j$ is a copy of $\tilde{L}$ with the same conventions as in the previous section. Analogously, we wish to examine the pointwise limit of the sequence of characteristic functions $(\tilde{\phi}^n(t))_{n\in\mathbb{N}}$, where

$$\begin{aligned} \tilde{\phi}^n(t) &:= \mathbb{E}\left(e^{it\,\tilde{S}_n} \right) \qquad (5.19) \\ = \int\limits_{\mathbb{R}^d_+} \prod_{l=1}^{m} &\left(1 + (e^{it\frac{\nu_l}{n}} - 1)\bar{p}_l \sum_{k=1}^{d} w^l_k s_k \right)^n f_\gamma(s_1, ..., s_d) ds_1 \cdots ds_d. \end{aligned}$$

To this end, we shall make use of the following elementary Lemma:

Lemma 5.5.7
For any sequence $(z_n)_{n\in\mathbb{N}}$ of complex numbers, we have

$$\lim_{n\to\infty} z_n \;=\; z \quad\Longrightarrow\quad \lim_{n\to\infty}(1+z_n/n)^n \;=\; e^z.$$

Proof: For any complex number ζ satisfying $|\zeta|<1$, the principal value of $\log(1+\zeta)$ has the power series expansion

$$\log(1+\zeta) \;=\; \sum_{n=1}^{\infty}(-1)^{n-1}\frac{\zeta^n}{n}$$

so that for $|\zeta|\le 1/2$ we have

$$\begin{aligned}|\log(1+\zeta)-\zeta| \;&\le\; \frac{|\zeta|^2}{2} \;+\; \frac{|\zeta|^3}{3} \;+\; ...\\ &\le\; \frac{|\zeta|^2}{2}\,(1+|\zeta|+|\zeta|^2+...) \;=\; \frac{|\zeta|^2}{2}\,\frac{1}{1-|\zeta|} \;\le\; |\zeta|^2.\end{aligned}$$

Now suppose that $z_n \to z \in \mathbb{C}$. Since there exists some n_0 with $\left|\frac{z_n}{n}\right| \le \frac{1}{2}\ \forall n \ge n_0$, the above implies that

$$n\;(\log(1+z_n/n)) \;=\; n\left((z_n/n) \;+\; \mathscr{O}(1/n^2)\right) \;=\; z_n \;+\; \mathscr{O}(1/n) \;\to\; z$$

so that

$$(1+z_n/n)^n \;=\; \exp(n\log(1+z_n/n)) \;\to\; \exp(z).$$

□

The next assertion is the analogue of Proposition 5.5.3 adapted to the present modified setup:

Proposition 5.5.8
The sequence of random variables (5.18) converges in distribution to the random variable

$$S^* \;=\; \sum_{k=1}^{d} EL_k\,\gamma_k.$$

Proof: Consider first the sequence of *conditional* characteristic functions

$$\tilde{\phi}^n_{s_1,\ldots,s_d}(t) \;:=\; \mathbb{E}\left(e^{it\,\tilde{S}_n} \;\middle|\; \gamma = (s_1, ..., s_d) \right) \qquad n = 1, 2, ...$$

for some fixed $t \in \mathbb{R}$. From the definition of $\tilde{S}_n$ it is clear that

$$\tilde{\phi}^n_{s_1,\ldots,s_d}(t) \;=\; \prod_{l=1}^{m} \left(1 \;+\; (e^{it\,\nu_l/n} \;-\; 1)\; \bar{p}_l \;\sum_{k=1}^{d} w^l_k\, s_k \right)^n$$

for any $n \in \mathbb{N}$. We wish to derive the limit of the sequence $(\tilde{\phi}^n_{s_1,\ldots,s_d}(t))_{n\in\mathbb{N}}$ as $n \to \infty$. To this end, note that by an argument similar to (5.12), we have

$$\lim_{n\to\infty} n\; (e^{it\,\nu_l/n} \;-\; 1)\; \bar{p}_l \sum_{k=1}^{d} w^l_k\, s_k \;=\; it\,\nu_l\,\bar{p}_l \sum_{k=1}^{d} w^l_k\, s_k$$

so that by Lemma 5.5.7, we have

$$\begin{aligned}
\lim_{n\to\infty} \tilde{\phi}^n_{s_1,\ldots,s_d}(t) \;&=\; \prod_{l=1}^{m} \lim_{n\to\infty} \left(1 \;+\; (e^{it\,\nu_l/n} \;-\; 1)\; \bar{p}_l \sum_{k=1}^{d} w^l_k\, s_k \right)^n \\
&=\; \prod_{l=1}^{m} e^{it\,\nu_l\,\bar{p}_l\,\sum_{k=1}^{d} w^l_k\, s_k} \\
&=\; e^{it\;\sum_{l=1}^{m} \nu_l\,\bar{p}_l\,(\sum_{k=1}^{d} w^k_l\, s_k)} \\
&=\; e^{it\,\sum_{k=1}^{d} EL_k\, s_k}. \qquad (5.20)
\end{aligned}$$

Note further that since

$$\left| e^{it\,\tilde{S}_n} \right| \;\leq\; 1 \;\;\text{for any } n \in \mathbb{N},$$

we also have

$$|\, \tilde{\phi}^n_{s_1,\ldots,s_d}(t) \,| \;=\; \left| \mathbb{E}\left(e^{it\,\tilde{S}_n} \;\middle|\; \gamma = (s_1, ..., s_d) \right) \right| \;\leq\; 1$$

for any $n \in \mathbb{N}$. Thus the Dominated Convergence Theorem implies that

$$\begin{aligned}
\lim_{n\to\infty} \tilde{\phi}^n(t) &= \lim_{n\to\infty} \int\limits_{\mathbb{R}_+^d} \tilde{\phi}^n_{s_1,\dots,s_d}(t)\ f_\gamma(s_1,\dots,s_d)\, ds_1 \cdots ds_d \\
&= \int\limits_{\mathbb{R}_+^d} \lim_{n\to\infty} \tilde{\phi}^n_{s_1,\dots,s_d}(t)\ f_\gamma(s_1,\dots,s_d)\, ds_1 \cdots ds_d, \\
&= \int\limits_{\mathbb{R}_+^d} e^{it\sum_{k=1}^d EL_k\, s_k}\ f_\gamma(s_1,\dots,s_d)\, ds_1 \cdots ds_d,
\end{aligned}$$

which proves the assertion. □

Thus, the result does not depend on the particular form of the default variables and holds for the Bernoulli case as well.

5.6 Measuring Diversification

In this section we wish to relate the foregoing results to a particular measure of diversification. In order to measure the distance of a given portfolio from its maximum degree of diversification, it is tempting to compare an appropriate risk measure applied to the original portfolio with the corresponding risk measure applied to the limiting portfolio S^*. For a given confidence level, one may for example consider the quantile or the economic capital of some given credit portfolio and relate it to the quantile or economic capital of the corresponding limit portfolio to get an information about how much of the portfolio risk is made up of idiosyncratic risk and how much is due to pure systematic risk and thus not diversifiable. Such a comparison yields the theoretic potential of a given portfolio to reduce its risk by dividing large exposures into smaller ones with different obligors.

Formally, we shall consider a **risk measure** as a mapping $\rho : \mathscr{L}^0(P) \to [0,\infty)$, where $\mathscr{L}^0(P)$ denotes the set of measurable and real-valued functions on some fixed probability space $(\Omega, \mathscr{F}, P)$. A risk measure ρ is called

(i) positively homogeneous, if $\rho(c\,X) = c\rho(X)$ for any $c > 0$ and $X \in \mathscr{L}^0(P)$;

(ii) subadditive, if $\rho(X+Y) \leq \rho(X) + \rho(Y)$ for all $X, Y \in \mathscr{L}^0(P)$;

(iii) translation invariant, if $\rho(X+a) = \rho(X)+a$ for any $X \in \mathscr{L}^0(P)$ and $a \in \mathbb{R}_+$;

(iv) monotonic, if $X \geq Y$ in $\mathscr{L}^0(P)$ implies $\rho(X) \geq \rho(Y)$.

A risk measure which satisfies all of the above–mentioned items is commonly called a *coherent* risk measure (cf. (Artzner et al., 1999)). Recall that the quantile of a given loss distribution is no coherent risk measure as it lacks the property of subadditivity. Instead of considering the economic capital of some portfolio, we shall consider the expected shortfall as the relevant risk measure and examine the sequence of shortfall values corresponding to the sequence $(\tilde{S}_n)_{n\in\mathbb{N}}$ defined in (5.18).

Given a confidence level $\alpha \in (0,1)$, the expected shortfall of the original portfolio $\tilde{L}$ is defined as

$$ES_\alpha(\tilde{L}) := \frac{1}{1-\alpha}\left(\mathbb{E}\left[\tilde{L}\mathbf{1}_{\{\tilde{\mathbf{L}}>\mathbf{q}_\alpha(\tilde{\mathbf{L}})\}}\right] + q_\alpha(\tilde{L})\{P(\tilde{L} \leq q_\alpha(\tilde{L})) - \alpha\}\right) \tag{5.21}$$

where

$$q_\alpha(\tilde{L}) \;:=\; \inf\{\, x \,\geq\, 0 \,|\, P(\tilde{L} \,\leq\, x\,) \,\geq\, \alpha \,\}$$

denotes the α - quantile of $\tilde{L}$. Meanwhile, it is well-known that Expected Shortfall defines a coherent risk measure (cf. (Acerbi and Tasche, 2002)) . For the sequel, it is useful to rewrite the representation (5.21) in terms of an integral as

$$ES_\alpha(\tilde{L}) \;=\; \frac{1}{1-\alpha}\int_\alpha^1 q_\theta(\tilde{L})\,d\theta. \tag{5.22}$$

For a derivation of the above formula we refer again to (Acerbi and Tasche, 2002). As previously mentioned, in order to gain some insight in how much of the portfolio risk is of idiosyncratic nature, it may be useful to relate the risk of the original portfolio to the risk of the limit portfolio in the sense of Proposition 5.5.8. So define

$$\kappa(\tilde{L}) := \frac{ES_\alpha(S^*)}{ES_\alpha(\tilde{L})}. \tag{5.23}$$

Moreover, let $F_n(\cdot)$ denote the distribution function of $\tilde{S}_n$, i.e.,

$$F_n(x) := P(\tilde{S}_n \leq x) \quad \text{for any } x \geq 0.$$

The following assertion states that the expected shortfall of the limit portfolio S^* is the limit of the sequence of shortfall values along the sequence $(\tilde{S}_n)_{n\in\mathbb{N}}$.

Proposition 5.6.1
Given a confidence level $\alpha \in (0,1)$, the sequence $(\tilde{S}_n)_{n\in\mathbb{N}}$ of portfolios defined in (5.18) satisfies

$$\lim_{n\to\infty} ES_\alpha(\tilde{S}_n) = ES_\alpha(S^*). \tag{5.24}$$

Moreover, we have

$$0 \leq \frac{ES_\alpha(S^*)}{ES_\alpha(\tilde{L})} \leq 1. \tag{5.25}$$

Proof: We wish to show that (5.24) holds. Fix an arbitrary confidence level $\theta \in (0,1)$ and let

$$q_\theta(\tilde{S}_n) := \inf\{x \geq 0 \mid P(\tilde{S}_n \leq x) \geq \theta\}$$

denote the θ – quantile of the portfolio $\tilde{S}_n$. We claim that

$$\lim_{n\to\infty} q_\theta(\tilde{S}_n) = q_\theta(S^*). \tag{5.26}$$

To see this, choose an arbitrary $r > 0$ and pick $x > 0$, such that

$$q_\theta(S^*) - r < x < q_\theta(S^*).$$

The above inequalities imply that $F(x) < \theta$ and since $F_n(x) \to F(x)$, there exists $n_0 \in \mathbb{N}$, such that

$$F_n(x) < \theta \quad \text{for all } n \geq n_0.$$

Thus, it follows that

$$q_\theta(S^*) - r < x \le q_\theta(\tilde{S}_n) \quad \text{for all } n \ge n_0$$

and hence,

$$q_\theta(S^*) \le \liminf_{n\to\infty} q_\theta(\tilde{S}_n) \tag{5.27}$$

since $r > 0$ was arbitrary. Next, choose $u \in (\theta, 1)$ and let $r > 0$. Pick some $x \ge 0$, such that

$$q_u(S^*) < x < q_u(S^*) + r.$$

It follows that

$$\theta < u < F(x) \text{ and } \theta < F_n(x) \quad \text{for } n \text{ sufficiently large.}$$

We may thus conclude that for n sufficiently large, we have

$$q_\theta(\tilde{S}_n) \le x < q_u(S^*) + r$$

and thus,

$$\limsup_{n\to\infty} q_\theta(\tilde{S}_n) \le q_u(S^*).$$

Letting $u \to \theta$, we may conclude that

$$\limsup_{n\to\infty} q_\theta(\tilde{S}_n) \le q_\theta(S^*). \tag{5.28}$$

Collecting (5.27) and (5.28), it follows that

$$\lim_{n\to\infty} q_\theta(\tilde{S}_n) = q_\theta(S^*).$$

From the definition of the portfolio $\tilde{S}_n$, it follows that

$$\tilde{S}_n = \frac{1}{n}\sum_{i=1}^{n} \tilde{L}_i = \frac{1}{n}\sum_{i=1}^{n}\sum_{l=1}^{m} D_l\,\nu_l \le \sum_{l=1}^{m} \nu_l < \infty$$

for any $n \in \mathbb{N}$. Thus, the Dominated Convergence Theorem implies that

$$\begin{aligned}
\lim_{n\to\infty} ES_\alpha(\tilde{S}_n) &= \frac{1}{1-\alpha} \lim_{n\to\infty} \int_\alpha^1 q_\theta(\tilde{S}_n)\, d\theta \\
&= \frac{1}{1-\alpha} \int_\alpha^1 \lim_{n\to\infty} q_\theta(\tilde{S}_n)\, d\theta \qquad (5.29) \\
&= \frac{1}{1-\alpha} \int_\alpha^1 q_\theta(S^*)\, d\theta \;=\; ES_\alpha(S^*)
\end{aligned}$$

which proves the first assertion. Finally, since ES_α is positively homogeneous and subadditive, we have

$$ES_\alpha(\tilde{S}_n) \;=\; ES_\alpha\left(\frac{1}{n}\sum_{i=1}^n \tilde{L}_i\right) \;\le\; \frac{1}{n}\sum_{i=1}^n ES_\alpha(\tilde{L}_i) \;=\; ES_\alpha(\tilde{L})$$

for any $n \in \mathbb{N}$, so that $ES_\alpha(S^*) = \lim_{n\to\infty} ES_\alpha(\tilde{S}_n) \le ES_\alpha(\tilde{L})$ as well and the assertion is proved.

Bibliography

Abramowitz, M. and I. Stegun (1972). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.* John Wiley & Sons.

Acerbi, C. and D. Tasche (2002). Expected Shortfall: A natural coherent Alternative to Value at Risk. *Economic Notes 32*(2), 379–388.

Albanese, C. and G. Campolieti (2001). Integrability by Quadratures of Pricing Equations. `www3.imperial.ac.uk/portal/pls/portallive/docs/1/11166.PDF`.

Albanese, C., G. Campolieti, O. Chen, and A. Zavidanov (2003). Credit Barrier Models. *RISK 16*(6), 125–129.

Albanese, C. and A. Kuznetsov (2007). Transformations of Markov Processes and Classification Scheme for Solvable Driftless Diffusions. `http://arxiv.org/abs/0710.1596v1`.

Alfonsi, A. (2005). On the Discretization Schemes for the CIR (and Bessel Squared) Processes. *Monte Carlo Methods and Applications 11*(4), 355–384.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent Measures of Risk. *Mathematical Finance 9*(3), 203–228.

Avellaneda, M. and Z. Zhu (2001). Distance to Default. *RISK 14*(12), 125–129.

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

Black, F. and M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy 81*, 637–654.

Borodin, A. and P. Salminen (1996). *Handbook of Brownian Motion - Facts and Formulae*. Birkhäuser-Verlag Basel Boston Berlin.

CSFP (1997). CreditRisk+. *Technical Document*.

Duffie, D. and K. Singleton (1999). Modelling Term Structures of Defaultable Bonds. *Review of Financial Studies 12*(4), 197–226.

Duffie, D. and K. Singleton (2003). *Credit Risk: Pricing, Measurement and Management*. Princeton University Press.

Ebmeyer, D., R. Klaas, and P. Quell (2007). The Role of Copulas in the CreditRisk+™Framework. *appeared in: Copulas: From Theory to Application in Finance*.

Feller, W. (1971). *An Introduction to Probability Theory and Applications, 2nd ed., Volume II*. John Wiley & Sons.

Getoor, R. (1979). The Brownian Escape Process. *The Annals of Probability 7*(5), 864–867.

Getoor, R. and M. Sharpe (1979). Excursions of Brownian motion and Bessel Processes. *Z. Wahrsch. Verw. Gebiete 47*, 83–106.

Giese, G. (2004). Enhanced CreditRisk+. *appeared in: CreditRisk+ in the Banking Industry*, 77–88.

Göing-Jaeschke, A. and M. Yor (2003). A Survey and some Generalizations of Bessel Processes. *Monte Carlo Methods and Applications 9*(2), 313–349.

Haaf, H., O. Reiß, and J. Schoenmakers (2004). Numerically stable Computation of CreditRisk+. *Journal of Risk 6*(4).

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

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

Itô, K. and H. McKean (1974). *Diffusion Processes and their Sample Paths.* Springer-Verlag Berlin Heidelberg.

Jarrow, R. and S. Turnbull (1992). Credit Risk: Drawing the Analogy. *RISK 5*(9).

Jarrow, R. and S. Turnbull (1995). Pricing Derivatives on Financial Securities subject to Credit Risk. *Journal of Finance 50*(1), 53–85.

JPMorgan (1997). Credit Metrics. `www.creditmetrics.com`.

Madan, D., P. Carr, and E. Chang (1998). The Variance Gamma Process and Option Pricing. *European Finance Review 2*, 79–105.

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

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

Rank, J. (2007). *Copulas: From Theory to Application in Finance.* RISK Book.

Revuz, D. and M. Yor (1991). *Continuous Martingales and Brownian Motion.* Springer-Verlag Berlin Heidelberg.

Rogers, L. and D. Williams (1987). *Diffusions, Markov Processes and Martingales, Vol.2: Itô Calculus.* John Wiley & Sons.

Schmidt, T. (2007). Coping with Copulas. *appeared in: Copulas: From Theory to Application in Finance*.

Vasicek, O. (1987). Probability of Loss on Loan Portfolio. KMV Corporation.

Vasicek, O. (2002). Loan Portfolio Value. *RISK 15*(12), 160–162.

Zhou, C. (1997). Default Correlation: An Analytic Result. www.defaultrisk.com/pp_corr_02.htm.