Analytical approximations for the value of embedded options in unit-linked insurance and profit sharing

by

Almer Moggré

Supervisors SNS Reaal
Dr. F. Boshuizen,
Drs. R. Smit.

Supervisor Utrecht University
Dr. A.V. Gnedin.

13-08-2010

Abstract

Many insurance companies sell products that involve embedded options. The value of such an option represents the expected future liability and therefore it is important that insurers can value the options they sold. Since most of these options are very complex, they are valued using Monte Carlo simulations. This requires considerable computation resources and therefore methods have been developed to approximate the option values analytically. In this thesis we study two of these options and give analytical approximations for their values. The first option is a guarantee in unit-linked insurance for which upper and lower bounds are derived using the concept of comonotonicity as developed by Dhaene, Denuit, Goovaerts, Kaas and Vyncke (2002a, 2002b). This is done in the Black-Scholes model as well as in the Hull-White-Black-Scholes model, where the latter has the additional feature of stochastic interest rates. The lower bound is the same as derived in Schrager and Pelsser (2004), but the derivation by explicitly applying the concept of comonotonicity was not given before. The second option is a call option on an average of swap rates as used in profit sharing. The value is approximated by using approximate swap rate dynamics as developed by Schrager and Pelsser (2006). Finally, the quality of the approximations is determined by comparing them to Monte Carlo simulations. It turns out that the lower bound for the guarantee in unit-linked insurance is a very accurate approximation.

Contents

1 Introduction

Many insurance companies sell products that contain embedded options. Similar to stock options, the seller of the option is at risk, because the buyer has the right to exercise the option, but not the obligation. This means that the buyer will only exercise the option if it has a positive terminal value and in that case, the seller has to pay. An example of such an embedded option is a guarantee in unit-linked insurance. This is a contract where a customer pays a yearly premium that is invested in an investment fund. At maturity of the product, the customer gets the total amount of money accumulated over the years. This amount will depend on the performance of the investment fund, but for many customers this would be to risky. Imagine what happens when you are saving for your retirement and the fund performs really bad. Therefore, insurers guarantee to pay a minimum amount at maturity. This is basically a put option that the insurance company sells to its customers. If the fund performs bad, the customer exercises the option and gets the guaranteed amount. If the fund performs well, the accumulated value will be higher than the guaranteed amount and the customer will not exercise the option.

The problem with embedded options is that the insurer is not definite about how much he has to pay in the future. However, with the recent financial crises in mind, it is not hard to imagine that the amount an insurer has to pay can be significant. To prepare for this, insurance companies need to ensure that they have enough capital to cover these liabilities.

So the question that arises is how much capital should this be? The answer is the value of the option, because the value of the option reflects the expected future liability. This looks very simple, but what is the value of the option? That is not so easy, because embedded options are not traded. So we cannot simply look at the market to see what the value of the option is.

To answer this question, we need other techniques and mathematics can help us out. Since the 1970's a whole framework of pricing financial products was developed and the basic idea is that the value of any financial product is given by the expected value under the risk neutral measure of the discounted payoff (see Harrison and Kreps (1979), Harrison and Pliska (1981)). In most cases it is not possible to compute this value analytically and simulations are used to get approximations. These approximations are almost exact if enough simulations are used.

This technique of determining the value of financial products (an embedded option is also a financial product) is now widespread. However, many financial products are so complex that simulations take huge amounts of time. This is very inconvenient if an insurance company wants to compute values in different scenario's. Therefore, it is important to find good analytical approximations for the value of embedded options that can replace the simulations.

In the literature we can find different analytical approximations for different kinds of embedded options. In this thesis we will study two of these options. We will give analytical approximations for the value of these options and we will test them. As we will see, some of these approximations are very accurate.

The first embedded option we will study is the guarantee in unit-linked insurance that we introduced above. As will be discussed in chapter three, we have to compute the following expectation under the risk neutral measure to obtain the value of the guarantee:

$$\mathbb{E}\left[\left(G - \sum_{i=0}^{T-1} P_i \frac{S_T}{S_i}\right)_+\right],$$

where G is the guaranteed payoff, T the maturity of the product, P_i the premium paid by the customer at time i for $i = 0, \ldots, T-1$ and S_i is the value of the investment fund at time i. Also, the notation z_+ is adopted for the positive part $z_+ = max(z, 0)$. The fractions $\frac{S_T}{S_i}$ are unknown at time 0, since we don't know what the value of the investment fund will be in the future. So this means that the guarantee in unit-linked insurance is a put option on a sum of the random variables $P_i \frac{S_T}{S_i}$. These random variables are even dependent which makes it impossible to compute the exact value.

In this thesis we will therefore derive an accurate approximation of the option value. For this we will use bounds on the value of the option as developed by Dhaene, Denuit, Goovaerts, Kaas, Vyncke (2002a,2002b). These bounds are based on the theory of comonotonicity and the parts from the theory that we will heavily use, are given in chapter 2. In the appendix, a complete treatment of the theory is given, since it can also be used in similar valuation problems. Here I would like to stress that this is copied from Dhaene, Denuit, Goovaerts, Kaas, Vyncke (2002a,2002b) except that I extended some derivations to make them easier to understand for students in the Master's. Furthermore, I would like to thank the authors for providing me with the tex version of their work, saving me a lot of writing.

In order to be able to derive concrete bounds on the value of the option, we need a model for the value of the investment fund S_t at any time $0 \leq t \leq T$. In chapter three we will use a Black-Scholes model. This model has the advantages that it is easy to use and well-known in practice. A big disadvantage is that the interest rate is fixed. Therefore, we will use the Hull-White-Black-Scholes model in chapter four, where the interest rate is stochastic. The modelling in this chapter comes from Schrager and Pelsser (2004), but they use a Jamshidian decomposition (see also Jamshidian (1989), Nielsen and Sandmann(2002)) instead of comonotonicity to find an approximation for the value of the option. Both approaches give exactly the same result, but the derivation of this result by explicitly applying the concept of comonotonicity is not given before.

In chapter five we will study the second embedded option. This option is called profit sharing and basically means that a customer always gets a guaranteed rate of return $TR(t)$ at time t, but if a certain reference rate $R(t)$ is higher than the guaranteed rate of return, he gets this higher rate. Often swap rates are used as the reference rate and the chapter will start with an explanation of swap rate dynamics. Then, to determine the value of the option, we have to compute an expectation like

$$\mathbb{E}\left[(R(t) - TR(t))_+\right],$$

where the reference rate $R(t)$ (now the swap rate) at time t is unknown at time 0. This swap rate will therefore be modelled, but it will turn out that the swap rate has a stochastic volatility. This makes it again impossible to compute the above expectation and therefore we will approximate the stochastic volatility by a deterministic one. This is an interesting idea developed by Schrager and Pelsser (2006) and applied to profit sharing by Plat and Pelsser (2009). Chapter five comes from Plat and Pelsser (2009), although we extend a lot of derivations to make them easier to understand. Also, we have tested the approximation in our own test environment.

2 The theory of comonotonicity

As mentioned in the introduction we will approximate the value of the guarantee in unit-linked insurance using bounds. These bounds are based on the theory of comonotonicity. In this chapter we will give an overview of the theory and the bounds it produces. The results are limited to the ones we need in the following chapters and stated without proof. For a complete overview of the theory including proofs we refer to the appendix.

2.1 Convex order

Recall from the introduction that the value of the guarantee in unit-linked insurance depends on:

$$\mathbb{E}\left[\left(G-\sum_{i=0}^{T-1} P_i \frac{S_T}{S_i}\right)_+\right].$$

To get bounds on this value, we need random variables X and Y such that

$$\mathbb{E}\left[(G-X)_+\right] \leq \mathbb{E}\left[\left(G-\sum_{i=0}^{T-1} P_i \frac{S_T}{S_i}\right)_+\right] \leq \mathbb{E}\left[(G-Y)_+\right],$$

and such that we can compute $\mathbb{E}\left[(G-X)_+\right]$ and $\mathbb{E}\left[(G-Y)_+\right]$ explicitly. The approach we take will be based on convex order:

Definition 2.1 *Consider two random variables X and Y. Then X is said to precede Y in the convex order sense, notation $X \leq_{cx} Y$, if and only if*

$$\begin{aligned} \mathbb{E}\left[X\right] &= \mathbb{E}\left[Y\right], \\ \mathbb{E}\left[(d-X)_+\right] &\leq \mathbb{E}\left[(d-Y)_+\right], \qquad -\infty < d < \infty. \end{aligned} \tag{1}$$

Here it is important to note that we will always consider random variables with finite mean. This makes sure that the expectations as above exist. Furthermore, we can see from the definition that we are looking for random variables X and Y such that they have the same mean as $\sum_{i=0}^{T-1} P_i \frac{S_T}{S_i}$ and such that $X \leq_{cx} \sum_{i=0}^{T-1} P_i \frac{S_T}{S_i}$ and $\sum_{i=0}^{T-1} P_i \frac{S_T}{S_i} \leq_{cx} Y$.

It can also be proved that $X \leq_{cx} Y$ if and only if $\mathbb{E}\left[v(X)\right] \leq \mathbb{E}\left[v(Y)\right]$ for all convex functions v, provided the expectations exist. This explains the name convex order. Also, we have $X \leq_{cx} Y$ if and only if $\mathbb{E}\left[X\right] = \mathbb{E}\left[Y\right]$ and $\mathbb{E}\left[u(-X)\right] \geq \mathbb{E}\left[u(-Y)\right]$ for all non-decreasing concave functions u, provided the expectations exist. Hence, in a utility context, convex order represents the common preferences of all risk-averse decision makers between random variables with equal mean. More explicit, suppose you can choose between two random payments X and Y that you have to make in one year and that your utility function u is a non-decreasing concave function. If $X \leq_{cx} Y$, then X and Y have the same mean. Furthermore it holds that $\mathbb{E}\left[u(-X)\right] \geq \mathbb{E}\left[u(-Y)\right]$, i.e. the expected utility of making payment X is larger than the expected utility of making payment Y. Since rational individuals want to maximize their utility, they will choose X.

If we look at the inequality $\mathbb{E}\left[v(X)\right] \leq \mathbb{E}\left[v(Y)\right]$ for the specific convex function $v(x) = x^2$, we immediately see that $X \leq_{cx} Y$ implies $Var(X) \leq Var(Y)$. Furthermore, we have the following relation for the variance:

$$\frac{1}{2} Var(X) = \int_{-\infty}^{\infty} \mathbb{E}\left[(X-t)_+\right] - (\mathbb{E}\left[X\right] - t)_+ dt, \tag{2}$$

which is proved in the appendix (equation (101)). This will play an important role when we derive one of the estimates for the value of the embedded option in unit-linked insurance.

2.2 Comonotonicity

When valuing the option in unit-linked insurance, we are dealing with a sum of dependent random variables. That is, we are dealing with random variables of the type $S = \sum_{i=1}^{n} X_i$ where the X_i are not mutually independent. Since it is not possible to compute the value of an option on such a sum, we are looking for bounds on this value. Therefore, given the marginal distributions of the terms in a random variable $S = \sum_{i=1}^{n} X_i$, we will look for the joint distribution with the largest sum in the convex order sense. This will give an upper bound and as we will see in section 2.4, the convex-largest sum of the components of a random vector with given marginals will be obtained in the case that the random vector $(X_1, X_2, \ldots, X_n)$ has the *comonotonic* distribution, which means that each two possible outcomes $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ of $(X_1, X_2, \ldots, X_n)$ are ordered component wise.
We start by defining comonotonicity of a set of n-vectors in $\mathbb{R}^n$. A n-vector $(x_1, \ldots, x_n)$ will be denoted by $\vec{x}$. For two n-vectors $\vec{x}$ and $\vec{y}$, the notation $\vec{x} \leq \vec{y}$ will be used for the component wise order which is defined by $x_i \leq y_i$ for all $i = 1, \ldots, n$.

Definition 2.2 *The set $A \subseteq \mathbb{R}^n$ is said to be comonotonic if for all $\vec{x}$ and $\vec{y}$ in A, either $\vec{x} \leq \vec{y}$ or $\vec{y} \leq \vec{x}$ holds.*

So, a set $A \subseteq \mathbb{R}^n$ is comonotonic if for any $\vec{x}$ and $\vec{y}$ in A we have that if $x_i < y_i$ for some i, then $\vec{x} \leq \vec{y}$ must hold. Hence, a comonotonic set is simultaneously non-decreasing in each component. Notice that a comonotonic set is a thin set: it cannot contain any subset of dimension larger than 1. Any subset of a comonotonic set is also comonotonic.

Next, we will define the notion of support of an n-dimensional random vector $\vec{X} = (X_1, \ldots, X_n)$. Any closed subset $A \subseteq \mathbb{R}^n$ will be called a support of $\vec{X}$ if $\mathbb{P}\left[\vec{X} \in A\right] = 1$ holds true. In general we will be interested in supports which are " as small as possible ". Informally, the smallest support of a random vector $\vec{X}$ is the subset of $\mathbb{R}^n$ that is obtained by subtracting from $\mathbb{R}^n$ all points which have a zero probability neighbourhood with respect to $\vec{X}$. This support can be interpreted as the set of all possible outcomes of $\vec{X}$. We will now define the important concept of comonotonicity of random vectors:

Definition 2.3 *A random vector $\vec{X} = (X_1, \ldots, X_n)$ is said to be comonotonic if it has a comonotonic support.*

From the definition we can conclude that comonotonicity is a very strong positive dependency structure. Indeed, if $\vec{x}$ and $\vec{y}$ are elements of the comonotonic support of $\vec{X}$, i.e. $\vec{x}$ and $\vec{y}$ are possible outcomes of $\vec{X}$, then they must be ordered component wise. This is also the intuition behind comonotonicity. A random vector $\vec{X}$ with components X_i is comonotonic if the set of all possible outcomes is ordered component wise.

In the following theorem, some useful equivalent characterizations will be given for comonotonicity of a random vector, but we first need to introduce the cumulative distribution function (cdf) of a random variable X and some related properties. The cdf $F_X(x) = \mathbb{P}[X \leq x]$ of a random variable X is a right-continuous (r.c.) non-decreasing function with

$$F_X(-\infty) = \lim_{x \to -\infty} F_X(x) = 0, \qquad F_X(\infty) = \lim_{x \to \infty} F_X(x) = 1,$$

where right-continuous means that for any x and any sequence x_n decreasing to x we have

$$\lim_{n \to \infty} F_X(x_n) = F_X(x).$$

The usual definition of the inverse of a distribution function is the non-decreasing and left-continuous (l.c.) function $F_X^{-1}(p)$ defined by

$$F_X^{-1}(p) = \inf\{x \in \mathbb{R} | F_X(x) \geq p\}, \qquad p \in [0,1] \tag{3}$$

with $\inf \emptyset = \infty$ by convention. Left-continuous means that for any x and any sequence x_n increasing to x we have

$$\lim_{n \to \infty} F_X(x_n) = F_X(x).$$

From these definitions it follows that for all $x \in \mathbb{R}$ and $p \in [0,1]$, we have

$$F_X^{-1}(p) \leq x \iff p \leq F_X(x). \tag{4}$$

Furthermore, it can be proved from (4) that (see (109) in the appendix)

$$X =_d F_X^{-1}(U), \tag{5}$$

where $=_d$ denotes equality in distribution and U is a $uniform(0,1)$ random variable, i.e. $F_U(p) = p$ and $F_U^{-1}(p) = p$ for all $0 < p < 1$.

The theorem is now as follows (for the proof see appendix theorem 8.7).

Theorem 2.4 *A random vector $\vec{X} = (X_1, \ldots, X_n)$ is comonotonic if and only if one of the following equivalent conditions holds:*

(1) *$\vec{X}$ has a comonotonic suppport.*

(2) *For all $\vec{x} = (x_1, \ldots, x_n)$, we have*

$$F_{\vec{X}}(\vec{x}) = \min\left(F_{X_1}(x_1), \ldots, F_{X_n}(x_n)\right), \tag{6}$$

where $F_{\vec{X}}(\vec{x}) = \mathbb{P}\left[\vec{X} \leq \vec{x}\right]$ is the multivariate cdf of $\vec{X}$.

(3) *There exists a random variable $U =_d$ Uniform(0,1), such that*

$$\vec{X} =_d \left(F_{X_1}^{-1}(U), \ldots, F_{X_n}^{-1}(U)\right). \tag{7}$$

(4) *There exists a random variable Z and non-decreasing functions f_i, $i = 1, \ldots, n$, such that*

$$\vec{X} =_d \left(f_1(Z), \ldots, f_n(Z)\right). \tag{8}$$

From (6) we see that, in order to find the probability of all the outcomes of n comonotonic risks X_i being less than x_i, $(i = 1, \ldots, n)$, one simply takes the probability of the least likely of these events. It is obvious that for any random vector $(X_1, \ldots, X_n)$, not necessarily comonotonic, the following inequality holds:

$$\mathbb{P}\left[X_1 \leq x_1, \ldots, X_n \leq x_n\right] \leq \min\{F_{X_1}(x_1), \ldots, F_{X_n}(x_n)\}, \tag{9}$$

and from theorem 2.4 we have that the function $\min\{F_{X_1}(x_1), \ldots, F_{X_n}(x_n)\}$ is the multivariate cdf of the random vector $\left(F_{X_1}^{-1}(U), \ldots, F_{X_n}^{-1}(U)\right)$, which has the same marginals as $(X_1, \ldots, X_n)$ (this follows from (5)). The inequality (9) states that in the class of all random vectors $(X_1, \ldots, X_n)$ with the same marginals, the probability that all X_i simultaneously realize 'small' values is maximized if the vector is comonotonic, suggesting that comonotonicity is indeed a very strong positive dependency structure.
In the sequel, for any random vector $(X_1, \ldots, X_n)$, the notation $(X_1^c, \ldots, X_n^c)$ will be used to indicate a comonotonic random vector with the same marginals as $(X_1, \ldots, X_n)$.

Example As an example of comonotonicity, consider the random vector $(S(T), (S(T) - K)_+)$ where K is some constant. We can think of $S(T)$ as the price of a certain stock at time T. Then this vector is a stock and a call option on the stock. This vector is comonotonic: if $S(T)$ increases both components will increase or stay equal and if $S(T)$ decreases both components will decrease or stay equal.

Similarly, $(S(T), (K - S(T))_+)$ is not comonotonic. If $S(T)$ increases the first component increases, but the second component might decrease.

2.3 Sums of comonotonic random variables

In section 2.1 we started with the observation that we need random variables X and Y such that

$$\mathbb{E}\left[(G-X)_+\right] \leq \mathbb{E}\left[\left(G-\sum_{i=0}^{T-1} P_i \frac{S_T}{S_i}\right)_+\right] \leq \mathbb{E}\left[(G-Y)_+\right],$$

and such that we can compute $\mathbb{E}\left[(G-X)_+\right]$ and $\mathbb{E}\left[(G-Y)_+\right]$ explicitly.
It will turn out that we can choose X and Y such that they are comonotonic sums. In this section we will show how to compute an expectation like $\mathbb{E}\left[(G-X)_+\right]$, when X is a comonotonic sum.
For this, we first define $F_X^{-1+}(p)$ for the random variable X by

$$F_X^{-1+}(p) = \sup\{x \in \mathbb{R} | F_X(x) \leq p\}, \qquad p \in [0,1]. \tag{10}$$

Then $F_X^{-1+}(p)$ is a non-decreasing and right-continuous function and this is an alternative definition for the inverse distribution function of the random variable X. Note that $F_X^{-1}(0) = -\infty$, $F_X^{-1+}(1) = \infty$ and that all probability mass of X is contained in the interval $[F_X^{-1+}(0), F_X^{-1}(1)]$. Also note that $F_X^{-1}(p)$ and $F_X^{-1+}(p)$ are finite for all $p \in (0,1)$ and that $F_X^{-1}(p)$ and $F_X^{-1+}(p)$ differ only for values of p where F_X is flat.

The main theorem is now as follows (theorem 8.10 in the appendix).

Theorem 2.5 *Suppose that the random vector $(X_1, \ldots, X_n)$ has strictly increasing and continuous marginals and let S^c be the sum of the components of the comonotonic random vector $(X_1^c, \ldots, X_n^c)$. Then we have*

$$\mathbb{E}\left[(S^c - d)_+\right] = \sum_{i=1}^{n} \mathbb{E}\left[(X_i - d_i)_+\right], \qquad F_{S^c}^{-1(+)}(0) < d < F_{S^c}^{-1}(1), \tag{11}$$

with the d_i given by

$$d_i = F_{X_i}^{-1}(F_{S^c}(d)), \qquad i = 1, \ldots, n \tag{12}$$

From this theorem it follows that we also need an expression for the distribution function F_{S^c} of S^c in the point d. This is obtained from the inverse distribution function $F_{S^c}^{-1}$ which follows from the following theorem (theorem 8.9 in the appendix).

Theorem 2.6 *The inverse distribution function $F_{S^c}^{-1}$ of a sum S^c of comonotonic random variables $(X_1^c, \ldots, X_n^c)$ is given by*

$$F_{S^c}^{-1}(p) = \sum_{i=1}^{n} F_{X_i}^{-1}(p), \qquad 0 < p < 1. \tag{13}$$

So we get $F_{S^c}(d)$ by solving the following equation

$$d = F_{S^c}^{-1}(F_{S^c}(d)) = \sum_{i=1}^{n} F_{X_i}^{-1}(F_{S^c}(d)), \tag{14}$$

from which we immediately see that it is enough to know the marginal distributions of the X_i. This will be very important in chapters three and four. Also note that by definition of the d_i we have $\sum_{i=1}^n d_i = \sum_{i=1}^n F_{X_i}^{-1}(F_{S^c}(d)) = F_{S^c}^{-1}(F_{S^c}(d)) = d$.

Now, the only missing piece is that we are not interested in $\mathbb{E}\left[(S^c - d)_+\right]$ but in $\mathbb{E}\left[(d - S^c)_+\right]$. Application of the relation $E[(X-d)_+] = E\left[(d-X)_+\right] + E\left[X\right] - d$ for S^c and the X_i in (11) leads to

$$E\left[(d - S^c)_+\right] = \sum_{i=1}^{n} E\left[(d_i - X_i)_+\right], \qquad F_{S^c}^{-1+}(0) < d < \ F_{S^c}^{-1}(1), \tag{15}$$

with the d_i as defined in (12).

So the above derivation gives the tools to compute an expectation like $\mathbb{E}\left[(G-Y)_{+}\right]$, when Y is a comonotonic sum of which we know the marginal distributions. This derivation is only done for strictly increasing and continuous marginals, but in the appendix an extensive treatment for general marginal distributions is given.

Before we go to an example, we state the following useful theorem (theorem 8.3 in the appendix)

Theorem 2.7 *Let X and $g(X)$ be real-valued random variables, and let $0<p<1$.*

(a) *If g is non-decreasing and continuous, then*

$$F_{g(X)}^{-1}(p)=g(F_X^{-1}(p)).$$

(b) *If g is non-increasing and continuous, then*

$$F_{g(X)}^{-1}(p)=g(F_X^{-1+}(1-p)).$$

Example Consider a random vector $(\alpha_1X_1,\alpha_2X_2,\ldots,\alpha_nX_n)$ of which the α_i are non-zero real numbers and the X_i are lognormally distributed: $\ln(X_i)\sim N\left(\mu_i,\sigma_i^2\right)$. We have that

$$\begin{aligned}\mathbb{E}\left[X_i\right] &= e^{\mu_i+\frac{1}{2}\sigma_i^2}, &(16)\\ Var\left[X_i\right] &= e^{2\mu_i+\sigma_i^2}\left(e^{\sigma_i^2}-1\right). &(17)\end{aligned}$$

As an example we can think of the situation where the α_i are deterministic payments at times i, and the X_i are the corresponding lognormally distributed discount factors. Then $(\alpha_1X_1,\alpha_2X_2,\ldots,\alpha_nX_n)$ is the vector of the stochastically discounted deterministic payments. From $\Phi^{-1}(1-p)=-\Phi^{-1}(p)$, and Theorem 2.7 ((a) for $\alpha_i>0$ and (b) for $\alpha_i<0$) we find that

$$F_{\alpha_iX_i}^{-1}(p)=\alpha_i\ e^{\mu_i+\mathrm{sgn}(\alpha_\mathrm{i})\ \sigma_\mathrm{i}\Phi^{-1}(\mathrm{p})},\qquad 0<p<1, \tag{18}$$

where $\mathrm{sgn}\left(\alpha_\mathrm{i}\right)$ equals 1 if $\alpha_i>0$ and -1 if $\alpha_i<0$. In particular, we find that the product of n comonotonic lognormal random variables is again lognormal(this is not always the case, for example, if the individual normal distributions do not constitute a multivariate normal distribution):

$$\Pi_{i=1}^nF_{X_i}^{-1}(U)=_d e^{\sum_{i=1}^n\mu_i+\sum_{i=1}^n\sigma_i\Phi^{-1}(U)}. \tag{19}$$

Also, for X_i lognormal we have

$$\mathbb{E}[(X_i-d_i)_+]=e^{\mu_i+\frac{\sigma_i^2}{2}}\ \Phi(d_{i,1})-d_i\ \Phi(d_{i,2}),\qquad d_i>0. \tag{20}$$

where $d_{i,1}$ and $d_{i,2}$ are determined by

$$d_{i,1}=\frac{\mu_i+\sigma_i^2-\ln\left(d_i\right)}{\sigma_i},\quad d_{i,2}=d_{i,1}-\sigma_i. \tag{21}$$

This result can be proved as follows. Differentiating $\mathbb{E}[(X_i-d_i)_+]$ with respect to d_i using $\mathbb{E}\left[(X-d)_+\right]=\int_d^\infty(1-F_X(x))dx$ we find that the derivative is given by $F_{X_i}(d_i)-1$. Similarly, differentiation of $e^{\mu_i+\frac{\sigma_i^2}{2}}\ \Phi(d_{i,1})-d_i\ \Phi(d_{i,2})$ with respect to d_i also gives $F_{X_i}(d_i)-1$ (this requires quite some computation). Also, for $d_i\to\infty$, both sides in (20) go to zero. So both sides end up with the same value and have the same derivative everywhere. But then they should be equal everywhere.
For the lower tails we find

$$\mathbb{E}[(d_i-X_i)_+]=-e^{\mu_i+\frac{\sigma_i^2}{2}}\ \Phi(-d_{i,1})+d_i\ \Phi(-d_{i,2}),\qquad d_i>0. \tag{22}$$

As $\mathbb{E}[(\alpha_i(X_i-d_i))_+]=-\alpha_i\ \mathbb{E}[(d_i-X_i)_+]$ if α_i is negative, we find from (20) and (22)

$$\mathbb{E}[(\alpha_i(X_i-d_i))_+]=\alpha_i\ e^{\mu_i+\frac{\sigma_i^2}{2}}\ \Phi(\mathrm{sgn}(\alpha_\mathrm{i})\ \mathrm{d_{i,1}})-\alpha_\mathrm{i}\ \mathrm{d_i}\ \Phi(\mathrm{sgn}(\alpha_\mathrm{i})\ \mathrm{d_{i,2}}),\qquad \mathrm{d_i}>0, \tag{23}$$

with $d_{i,1}$ and $d_{i,2}$ as defined in (21).

Now, let $S = \alpha_1 X_1 + \ldots + \alpha_n X_n$, and S^c its comonotonic counterpart: $S^c = F^{-1}_{\alpha_1 X_1}(U) + \ldots + F^{-1}_{\alpha_n X_n}(U)$. As the marginal distribution functions are strictly increasing and continuous, we find that the distribution function $F_{S^c}(x)$ is implicitly defined by $F^{-1}_{S^c}(F_{S^c}(x)) = x$, which is by (13) and (18) equivalent to,

$$\sum_{i=1}^{n} \alpha_i \; e^{\mu_i + \mathrm{sgn}(\alpha_i) \;\; \sigma_i \Phi^{-1}(F_{S^c}(x))} = x, \qquad F^{-1+}_{S^c}(0) < x < F^{-1}_{S^c}(1). \tag{24}$$

For $F^{-1+}_{S^c}(0) < d < F^{-1}_{S^c}(1)$, we have from theorem 2.5 and (18):

$$\begin{aligned}
\mathbb{E}[(S^c - d)_+] &= \sum_{i=1}^{n} \mathbb{E}[(\alpha_i X_i - F^{-1}_{\alpha_i X_i}(F_{S^c}(d)))_+] \\
&= \sum_{i=1}^{n} \mathbb{E}[\Big(\alpha_i \Big(X_i - e^{\mu_i + \mathrm{sgn}(\alpha_i) \;\; \sigma_i \Phi^{-1}(F_{S^c}(d))}\Big)\Big)_+].
\end{aligned}$$

Now, from (21) we find for $d_i = e^{\mu_i + \mathrm{sgn}(\alpha_i) \;\; \sigma_i \Phi^{-1}(F_{S^c}(d))}$ that

$$\begin{aligned}
d_{i,1} &= \frac{\mu_i + \sigma_i^2 - \ln(d_i)}{\sigma_i} \\
&= \frac{\mu_i + \sigma_i^2 - (\mu_i + \mathrm{sgn}(\alpha_i) \;\; \sigma_i \Phi^{-1}(F_{S^c}(d)))}{\sigma_i} \\
&= \frac{\sigma_i^2 - \mathrm{sgn}(\alpha_i) \;\; \sigma_i \Phi^{-1}(F_{S^c}(d)))}{\sigma_i} = \sigma_i - \mathrm{sgn}(\alpha_i) \;\; \Phi^{-1}(F_{S^c}(d))),
\end{aligned}$$

and $d_{i,2} = d_{i,1} - \sigma_i = -\mathrm{sgn}(\alpha_i) \;\; \Phi^{-1}(F_{S^c}(d)))$. So for this d_i we have

$$\begin{aligned}
\sum_{i=1}^{n} \alpha_i \; d_i \; \Phi(\mathrm{sgn}(\alpha_i) \; d_{i,2}) &= \sum_{i=1}^{n} \alpha_i \; d_i \; \Phi(-\mathrm{sgn}(\alpha_i) \; \mathrm{sgn}(\alpha_i) \;\; \Phi^{-1}(F_{S^c}(d)))) \\
&= \sum_{i=1}^{n} \alpha_i \; d_i \; \Phi(- \; \Phi^{-1}(F_{S^c}(d)))) \\
&= \sum_{i=1}^{n} \alpha_i \; d_i \; \Phi(\; \Phi^{-1}(1 - F_{S^c}(d)))) \\
&= \sum_{i=1}^{n} \alpha_i \; e^{\mu_i + \mathrm{sgn}(\alpha_i) \;\; \sigma_i \Phi^{-1}(F_{S^c}(d))} \; (1 - F_{S^c}(d)) \\
&= d(1 - F_{S^c}(d)),
\end{aligned}$$

where we used (24) for the last equality. Plugging this in (23) we find for $F^{-1+}_{S^c}(0) < d < F^{-1}_{S^c}(1)$

$$\mathbb{E}[(S^c - d)_+] = \sum_{i=1}^{n} \alpha_i \; e^{\mu_i + \frac{\sigma_i^2}{2}} \Phi\left(\mathrm{sgn}(\alpha_i) \;\; \sigma_i - \Phi^{-1}(F_{S^c}(d))\right) - d \;\; (1 - F_{S^c}(d)) \,. \tag{25}$$

Similarly, the lower tails are given by

$$\mathbb{E}[(d - S^c)_+] = -\sum_{i=1}^{n} \alpha_i \; e^{\mu_i + \frac{\sigma_i^2}{2}} \Phi\left(-\mathrm{sgn}(\alpha_i) \;\; \sigma_i + \Phi^{-1}(F_{S^c}(d))\right) + d \; F_{S^c}(d), \tag{26}$$

for $F^{-1+}_{S^c}(0) < d < F^{-1}_{S^c}(1)$.

In the sequel, equation (26) will play a very important role, because in computing the value of the option in unit-linked insurance we will work with lognormally distributed random variables.

2.4 Convex bounds for sums of random variables

In the previous section we already mentioned that the bounds on the value of the guarantee in unit-linked insurance can be obtained from comonotonic random vectors. Also, we have seen that if a random vector $\vec{X}$ is comonotonic, we can compute the value of a put option using only the marginal distributions (equation (15)). In this section we will derive explicitly the comonotonic random vectors that will produce the bounds.

An upper bound is obtained from the following theorem (appendix theorem 8.11)

Theorem 2.8 *For any random vector $(X_1, X_2, \ldots, X_n)$ we have*

$$S = X_1 + X_2 + \ldots + X_n \leq_{cx} X_1^c + X_2^c + \ldots + X_n^c = S^c. \tag{27}$$

This theorem states that an upper bound can be obtained by assuming that the components of a sum of random variables have the comonotonic dependency.

We can also derive an improved upper bound. For this, we assume that we have some additional information available concerning the stochastic nature of $(X_1, \ldots, X_n)$. More precisely, we assume that there exists some random variable Λ with a given distribution function, such that we know the conditional cdf's, given $\Lambda = \lambda$, of the random variables X_i, for all possible values of λ. We will show that in this case we can derive improved upper bounds in terms of convex order for S which are smaller in convex order than the upper bound S^c. Essentially, the idea is to determine comonotonic upper bounds for the sum S, conditionally given $\Lambda = \lambda$. Next, we mix the resulting distributions with weights $dF_\Lambda(\lambda)$. By this procedure, convex order is maintained. The upper bound obtained in this way turns out to be sharper than the comonotonic upper bound S^c because it still has the right marginal cdf's for its terms.

In the following theorem (appendix theorem 8.12), we introduce the notation $F^{-1}_{X_i|\Lambda}(U)$ for the random variable $f_i(U, \Lambda)$, where the function f_i is defined by $f_i(u, \lambda) = F^{-1}_{X_i|\Lambda=\lambda}(u)$.

Theorem 2.9 *Let U be uniform(0,1), and independent of the random variable Λ. Then we have*

$$X_1 + X_2 + \ldots + X_n \leq_{cx} F^{-1}_{X_1|\Lambda}(U) + F^{-1}_{X_2|\Lambda}(U) + \ldots + F^{-1}_{X_n|\Lambda}(U). \tag{28}$$

Note that the random vector $\left(F^{-1}_{X_1|\Lambda}(U), F^{-1}_{X_2|\Lambda}(U), \ldots, F^{-1}_{X_n|\Lambda}(U)\right)$ has marginals $F_{X_1}, F_{X_2}, \ldots, F_{X_n}$, because

$$\begin{aligned}
F_{X_i}(x) &= \int_{-\infty}^{\infty} \mathbb{P}\left[X_i \leq x \mid \Lambda = \lambda\right] dF_\Lambda(\lambda) \\
&= \int_{-\infty}^{\infty} \mathbb{P}\left[F^{-1}_{X_i|\Lambda=\lambda}(U) \leq x\right] dF_\Lambda(\lambda) \\
&= \int_{-\infty}^{\infty} \mathbb{P}\left[f_i(U, \lambda) \leq x\right] dF_\Lambda(\lambda) \\
&= \mathbb{P}\left[f_i(U, \Lambda) \leq x\right].
\end{aligned}$$

From theorem 2.8 it follows that for a random vector with given marginal distributions the comonotonic sum is the largest possible sum in the convex order sense. This implies

$$F^{-1}_{X_1|\Lambda}(U) + \ldots + F^{-1}_{X_n|\Lambda}(U) \leq_{cx} F^{-1}_{X_1}(U) + \ldots + F^{-1}_{X_n}(U), \tag{29}$$

which means that the upper bound derived is indeed an improved upper bound.
If Λ is independent of all $X_1, X_2, \ldots, X_n$, then we actually do not have any extra information at all and the improved upper bound reduces to the comonotonic upper bound derived in Theorem 2.8.

Finally, we would like to have a lower bound. Again we assume that there exists some random variable Λ with a given distribution function, such that we know the conditional cdf's, given $\Lambda = \lambda$, of the random variables X_i , for all possible values of λ. The idea is then to observe that the expectation of a random variable is always smaller than or equal in convex order than the random variable itself, and also that convex order is maintained under mixing (appendix theorem 8.13):

Theorem 2.10 *For any random vector X and any random variable Λ, we have*

$$\mathbb{E}\left[X_1 \mid \Lambda\right] + \mathbb{E}\left[X_2 \mid \Lambda\right] + \ldots + \mathbb{E}\left[X_n \mid \Lambda\right] \leq_{cx} X_1 + X_2 + \ldots + X_n. \tag{30}$$

Now let $S = X_1 + X_2 + \ldots + X_n$ and let S^l be defined by

$$S^l = \mathbb{E}\left[S \mid \Lambda\right] \tag{31}$$

Note that if Λ and S are mutually independent, we find the trivial result

$$\mathbb{E}\left[S\right] \leq_{cx} S. \tag{32}$$

On the other hand, if Λ and S have a one-to-one relation (i.e. Λ completely determines S), the lower bound coincides with S. Note further that $\mathbb{E}\left[\mathbb{E}\left[X_i \mid \Lambda\right]\right] = \mathbb{E}\left[X_i\right]$ always holds, but $Var\left[\mathbb{E}\left[X_i \mid \Lambda\right]\right] < Var\left[X_i\right]$ unless $\mathbb{E}\left[Var\left[X_i \mid \Lambda\right]\right] = 0$. This follows from the identity

$$Var(X_i) = \mathbb{E}\left[Var\left[X_i \mid \Lambda\right]\right] + Var\left[\mathbb{E}\left[X_i \mid \Lambda\right]\right]$$

and this means that X_i, given $\Lambda = \lambda$, is degenerate for each λ. This implies that the random vector $\left(\mathbb{E}\left[X_1 \mid \Lambda\right], \mathbb{E}\left[X_2 \mid \Lambda\right], \ldots, \mathbb{E}\left[X_n \mid \Lambda\right]\right)$ will in general not have the same marginal distribution functions as X. But if we can find a conditioning random variable Λ with the property that all random variables $\mathbb{E}\left[X_i \mid \Lambda\right]$ are non-increasing functions of Λ (or all are non-decreasing functions of Λ), the lower bound S^l is a sum of n comonotonic random variables by theorem 2.4(4).
To judge the quality of the stochastic lower bound $\mathbb{E}[S \mid \Lambda]$, we might look at its variance. To maximize it, i.e. to make it as close as possible to $Var\left[S\right]$, the average value of $Var[S \mid \Lambda = \lambda]$ should be minimized. In other words, to get the best lower bound, Λ and S should be as alike as possible.

To compute the lower bound, we need to compute conditional expectations like $\mathbb{E}\left[Y \mid \Lambda\right]$. An important case for us will be that we have a random vector $(Y_1, Y_2, \ldots, Y_n)$ with a multivariate normal distribution and where Y and Λ are linear combinations of the variates: $Y = \sum_{i=1}^n \alpha_i Y_i$ and $\Lambda = \sum_{i=1}^n \beta_i Y_i$. Then also $(Y,\ \Lambda)$ has a bivariate normal distribution. Further, if $(Y,\ \Lambda)$ has a bivariate normal distribution, then, conditionally given $\Lambda = \lambda$, Y has a univariate normal distribution with mean and variance given by

$$E\left[Y \mid \Lambda = \lambda\right] = E\left[Y\right] + r\left(Y, \Lambda\right) \frac{\sigma_Y}{\sigma_\Lambda} \left(\lambda - E\left[\Lambda\right]\right) \tag{33}$$

and

$$Var\left[Y \mid \Lambda = \lambda\right] = \sigma_Y^2 \ \left(1 - r\left(Y, \Lambda\right)^2\right), \tag{34}$$

where $r\left(Y, \Lambda\right) = \frac{Cov(Y,\Lambda)}{\sigma_\Lambda \sigma_Y}$ is the correlation coefficient for the couple (Y, Λ).

3 Pricing guarantees in unit-linked insurance

In the previous chapter we studied the concept of comonotonicity and we used this to derive convex bounds on sums of random variables. By definition of convex order the convex bounds also give bounds on a put option on the particular sum of random variables.
In this chapter we will apply this theory to the guarantee in unit-linked insurance. We will start with an explanation of this contract and our main goal will then be to get bounds on the value of this contract using the theory from chapter 2. These bounds will, of course, depend on the modelling assumptions. In this chapter we will start with a simple model, in which the value of the investment fund is modelled as a geometric Brownian motion (Black-Scholes model). The advantage of this model is that computations are relatively easy. Furthermore, the model is well-known in the financial industry. One of the disadvantages of the model is that it assumes a fixed interest rate and it is obvious that in practice that is not the case. Therefore we will extend the model in chapter 4 by assuming the interest rate is not fixed but stochastic.

3.1 Introduction

Recall from section one that unit-linked insurance is an insurance contract where a customer pays a premium, typically on an annual basis, that will be invested in an investment fund (for example, this investment fund could be a portfolio of stocks, a portfolio of bonds or a portfolio of stocks and bonds). At maturity, the customer then gets the total capital accumulated over the years. As an example we can imagine someone who has just started working and who wants to save for retirement. The problem for this person is, that there is a significant amount of risk. If the investments go really bad, he/she will end up with a little money. Therefore, insurers often guarantee to pay at least a guaranteed amount at maturity. This makes the product attractive for customers: if the investments go well, the customer gets the full profit, but if the investment go bad, he/she always gets at least the guaranteed amount. Note that this means that the customer has a put option on the investment fund.
This put option comes from the insurance company and the insurance company is now at risk. If the investments go bad, they will have to pay the difference between the guaranteed amount and the accumulated amount. In the past, insurance companies did not pay much attention to this, because they were used to a situation where investment returns were such that the accumulated value was bigger than the guaranteed value. But in the beginning of 2000, bearish stock markets and new accounting principles made insurers realize that the options could have significant value and that they had to have an idea about this value. This is important, because the value reflects the amount of money you expect to pay for the guarantees in the future.

Mathematically we can formalize the contract as follows. Suppose the customer pays a premium P_i in year i where $i = 0, \dots, T-1$. This premium will be invested in an investment fund and at maturity T the customer gets back the total amount of money accumulated over the years. If we denote S_t the value of the investment fund at time t, then the accumulated value at time T (denoted U_T) will be

$$U_T = \sum_{i=0}^{T-1} P_i \frac{S_T}{S_i}.$$

Note that U_T is a sum of the random variables $P_i \frac{S_T}{S_i}$, which have a certain dependency structure. In the sequel we will denote U_T^c the analogous sum where we assume the random variables in the sum to have the comonotonic dependency structure, i.e. $U_T^c = \sum_{i=0}^{T-1} F^{-1}_{P_i \frac{S_T}{S_i}}(U)$, where $U =_d$Uniform(0,1). Similarly we will denote $U_T^l = \sum_{i=0}^{T-1} \mathbb{E}\left[P_i \frac{S_T}{S_i} \mid \Lambda\right]$ and $U_T^u = \sum_{i=0}^{T-1} F^{-1}_{P_i \frac{S_T}{S_i} \mid \Lambda}(U)$, where Λ is a conditioning random variable.

In case the customer gets a guaranteed amount G, the payoff to the customer after T years will be $\max(U_T, G)$. Since U_T is totally financed by the customer and the performance of the investment fund, the cost to the insurer of the guarantee at time T equals

$$C_T = (G - U_T)_+$$

where again $z_+ = \max(z, 0)$.
We are now interested in pricing this guarantee at time $t = 0$. To compute this, we will suppose a Black-Scholes model for S_t, i.e. under the risk neutral measure we have:

$$S_t = S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W_t},$$

where r is the (constant) risk free interest rate, σ the volatility of the fund and W_t a standard Brownian motion. As we mentioned, the investment fund could be a portfolio of stocks, a portfolio of bonds, a portfolio of stocks and bonds or a portfolio of stocks, bonds and other securities. In the Black-Scholes model, the composition of the fund is captured by the volatility parameter. For example, a volatility of 20 % would be typical for stocks.
We will denote $V(T, G)$ the present value of this guarantee option with maturity T and guarantee level G. From the basic principles of risk-neutral valuation (see Harrison and Kreps (1979), Harrison and Pliska (1981)) we have for the present value of the option

$$V(T, G) = \mathbb{E}\left[e^{-rT}(G - U_T)_+\right],$$

where the expectation is taken with respect to the risk neutral measure. Plugging in U_T we get the following risk neutral valuation formula for $V(T, G)$:

$$V(T, G) = \mathbb{E}\left[e^{-rT}(G - U_T)_+\right] = \mathbb{E}\left[e^{-rT}(G - \sum_{i=0}^{T-1} P_i \frac{S_T}{S_i})_+\right] = e^{-rT}\ \mathbb{E}\left[(G - \sum_{i=0}^{T-1} P_i \frac{S_T}{S_i})_+\right]$$

Using standard properties of Brownian motion we also get

$$P_i\ \frac{S_T}{S_i} = P_i\ \frac{S_0 e^{(r-\frac{\sigma^2}{2})T+\sigma W_T}}{S_0 e^{(r-\frac{\sigma^2}{2})i+\sigma W_i}} = P_i\ e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma(W_T - W_i)} =_d P_i\ e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma W_{T-i}} \tag{35}$$

where $=_d$ means equality in distribution.
Now, $V(T, G)$ cannot be computed explicitly since it involves a sum of dependent random variables. Therefore, we will apply the theory from chapter 2 to derive upper and lower bounds on the value of the guarantee.

3.2 An upper bound on the value of the guarantee

We have the following upper bound on the value of the guarantee:

Theorem 3.1 *Let $V(T,G)$ be the value of the guarantee with maturity T and guarantee level G. Then*

$$V(T,G) \leq -\sum_{i=0}^{T-1} P_i e^{-ri}\Phi\left(-\sigma\sqrt{T-i}+\Phi^{-1}(F_{U_T^c}(G))\right)+e^{-rT}\ G\ F_{U_T^c}(G), \tag{36}$$

where $F_{U_T^c}(G)$ follows from solving

$$\sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma\sqrt{T-i}\Phi^{-1}(F_{U_T^c}(G))}=G. \tag{37}$$

Proof This result is derived as follows. From the introduction we have that

$$V(T,G)=\mathbb{E}\left[e^{-rT}(G-U_T)_+\right].$$

From theorem 2.8 we have that $U_T \leq_{cx} U_T^c$. By definition of convex order this means

$$V(T,G)=\mathbb{E}\left[e^{-rT}(G-U_T)_+\right] \leq \mathbb{E}\left[e^{-rT}(G-U_T^c)_+\right].$$

The marginal distributions of the terms in the sum U_T are given by (35). Equivalently, we can write $P_i\ \frac{S_T}{S_i} =_d P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma\sqrt{T-i}\Phi^{-1}(U)}$, where Φ^{-1} denotes the cumulative distribution function of the standard normal distribution and U is again a uniformly distributed random variable on the unit interval. From this we see that the marginals are lognormal with parameters $\mu_i=(r-\frac{\sigma^2}{2})(T-i)$ and $\sigma_i^2=\sigma^2(T-i)$ multiplied by the constant P_i. From (26) we therefore have for all $F_{U_T^c}^{-1+}(0)<G<F_{U_T^c}^{-1}(1)$

$$\mathbb{E}\left[(G-U_T^c)_+\right]=-\sum_{i=0}^{T-1} P_i e^{\mu_i+\frac{\sigma_i^2}{2}}\Phi\left(-\sigma_i+\Phi^{-1}(F_{U_T^c}(G))\right)+GF_{U_T^c}(G). \tag{38}$$

Inserting the parameters $\mu_i=(r-\frac{\sigma^2}{2})(T-i)$ and $\sigma_i^2=\sigma^2(T-i)$ we have that for all $0<G<\infty$

$$\mathbb{E}\left[(G-U_T^c)_+\right]=-\sum_{i=0}^{T-1} P_i e^{r(T-i)}\Phi\left(-\sigma\sqrt{T-i}+\Phi^{-1}(F_{U_T^c}(G))\right)+GF_{U_T^c}(G).$$

So we conclude that

$$\begin{aligned}
V(T,G) &= e^{-rT}\ \mathbb{E}\left[(G-U_T\right] \ \leq\ e^{-rT}\ \mathbb{E}[(G-U_T^c)_+] \\
&= e^{-rT}\ \left[-\sum_{i=0}^{T-1} P_i e^{r(T-i)}\Phi\left(-\sigma\sqrt{T-i}+\Phi^{-1}(F_{U_T^c}(G))\right)+GF_{U_T^c}(G)\right] \\
&= -\sum_{i=0}^{T-1} P_i e^{-ri}\Phi\left(-\sigma\sqrt{T-i}+\Phi^{-1}(F_{U_T^c}(G))\right)+e^{-rT}\ G\ F_{U_T^c}(G),
\end{aligned}$$

which gives (36). Also, from (24) with $\alpha_i=P_i$ and μ_i and σ_i^2 as above we immediately find

$$\sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma\sqrt{T-i}\Phi^{-1}(F_{U_T^c}(G))}=G,$$

which is equivalent to (37). ∎

In terms of computations this means that we first need to solve (37) and then the result can be plugged in (36). Note that (37) cannot be solved analytically. However, $F_{U_T^c}(G)$ will be between 0 and 1 and therefore the equation can easily be solved numerically.

3.3 A lower bound for the value of the guarantee

A lower bound on the value of the guarantee is given by:

Theorem 3.2 *Let $V(T,G)$ be the value of the guarantee with maturity T and guarantee level G. Then*

$$V(T,G) \geq -\sum_{i=0}^{T-1} P_i e^{-ri}\Phi\left(-\sigma R_{T-i}\sqrt{T-i} + \Phi^{-1}(F_{U_T^l}(G))\right) + e^{-rT}\ GF_{U_T^l}(G), \tag{39}$$

where $F_{U_T^l}(G)$ follows from

$$\sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2}R_{T-i}^2)(T-i)+\sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(F_{U_T^l}(G))} = G. \tag{40}$$

Furthermore, R_{T-i} is given by

$$R_{T-i} = \frac{\sum_{j=0}^{T-1} P_j e^{(r-\frac{\sigma^2}{2})(T-j)} \min(T-i, T-j)}{\sigma_\Lambda\sqrt{T-i}}, \tag{41}$$

where

$$\sigma_\Lambda^2 = \sum_{i=0}^{T-1}\sum_{j=0}^{T-1} P_i P_j e^{(r-\frac{\sigma^2}{2})(2T-i-j)} \min(T-i, T-j). \tag{42}$$

Proof This lower bound is derived as follows. From theorem 2.10 we have that $U_T^l \leq_{cx} U_T$ for any conditioning random variable Λ. As noted in section 2.4 we need to take Λ as alike to U_T to get the best lower bound. Therefore we consider the conditioning random variable

$$\Lambda = \sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)} W_{T-i}$$

where W is a standard Brownian motion. This Λ is a linear transformation of a first order approximation of U_T as is shown by the following computation:

$$\begin{aligned} U_T &= \sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma W_{T-i}} \\ &\approx \sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)}(1+\sigma W_{T-i}) \\ &= \sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)} + \sigma \sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)} W_{T-i}. \end{aligned}$$

The variance of Λ (denoted σ_Λ^2) is given by

$$\begin{aligned} Var(\Lambda) = Cov(\Lambda,\Lambda) &= Cov\left(\sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)} W_{T-i}, \sum_{j=0}^{T-1} P_j e^{(r-\frac{\sigma^2}{2})(T-j)} W_{T-j}\right) \\ &= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1} P_i P_j e^{(r-\frac{\sigma^2}{2})(2T-i-j)} Cov(W_{T-i}, W_{T-j}) \\ &= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1} P_i P_j e^{(r-\frac{\sigma^2}{2})(2T-i-j)} \min(T-i, T-j), \end{aligned}$$

which gives (42).
Next we define the correlation coefficient R_{T-i} for the couple (W_{T-i},Λ):

$$R_{T-i}=\frac{Cov\left(W_{T-i},\Lambda\right)}{\sigma_\Lambda\sqrt{T-i}} = \frac{Cov\left(W_{T-i},\sum_{j=0}^{T-1}P_je^{(r-\frac{\sigma^2}{2})(T-j)}W_{T-j}\right)}{\sigma_\Lambda\sqrt{T-i}} = \frac{\sum_{j=0}^{T-1}P_je^{(r-\frac{\sigma^2}{2})(T-j)}\min(T-i,T-j)}{\sigma_\Lambda\sqrt{T-i}}.$$

So indeed, R_{T-i} is given by (41). Now, we have that $(W_{T-(T-1)},W_{T-(T-2)},\ldots,W_T)=(W_1,\ldots,W_T)$ has a multivariate normal distribution. So given $\Lambda=\lambda$, it follows from equations (33) and (34) that the random variable W_{T-i} is normally distributed with mean

$$\mathbb{E}\left[W_{T-i}\ |\Lambda=\lambda\right]=\mathbb{E}\left[W_{T-i}\right]+R_{T-i}\frac{\sigma_{W_{T-i}}}{\sigma_\Lambda}(\lambda-\mathbb{E}\left[\Lambda\right])=R_{T-i}\frac{\sqrt{T-i}}{\sigma_\Lambda}\lambda$$

and variance

$$Var(W_{T-i}\ |\Lambda=\lambda)=\sigma^2_{W_{T-i}}(1-R^2_{T-i})=(T-i)(1-R^2_{T-i}).$$

By definition we have $U_T^l=\sum_{i=0}^{T-1}\mathbb{E}\left[P_i\frac{S_T}{S_i}\ |\Lambda\right]$. Plugging in $P_i\frac{S_T}{S_i}$ from (35) gives

$$U_T^l=\sum_{i=0}^{T-1}\mathbb{E}\left[P_ie^{(r-\frac{\sigma^2}{2})(T-i)+\sigma W_{T-i}}|\Lambda\right]=\sum_{i=0}^{T-1}P_ie^{(r-\frac{\sigma^2}{2})(T-i)}\mathbb{E}\left[e^{\sigma W_{T-i}}|\Lambda\right]$$

Now, let B be normally distributed with parameters a and b^2. If $c>0$, then cB is normally distributed with parameters ca and c^2b^2. So, since $\sigma>0$, we have that $e^{\sigma W_{T-i}}|\Lambda=\lambda$ is lognormal with parameters $\mu_i=\sigma R_{T-i}\frac{\sqrt{T-i}}{\sigma_\lambda}\lambda$ and $\sigma_i^2=\sigma^2(T-i)(1-R^2_{T-i})$. From (16) we have that the expectation of a lognormally distributed random variable is $e^{\mu_i+\frac{\sigma_i^2}{2}}$, so, given $\Lambda=\lambda$, we get

$$U_T^l=\sum_{i=0}^{T-1}P_ie^{(r-\frac{\sigma^2}{2})(T-i)}e^{\sigma R_{T-i}\frac{\sqrt{T-i}}{\sigma_\lambda}\lambda+\frac{\sigma^2}{2}(T-i)(1-R^2_{T-i})}=\sum_{i=0}^{T-1}P_ie^{(r-\frac{\sigma^2}{2}R^2_{T-i})(T-i)+\sigma R_{T-i}\frac{\sqrt{T-i}}{\sigma_\lambda}\lambda}.$$

Since Λ is normally distributed with mean zero we have that $\frac{\Lambda}{\sigma_\lambda}$ is standard normally distributed. Hence we have

$$U_T^l=_d\sum_{i=0}^{T-1}P_ie^{(r-\frac{\sigma^2}{2}R^2_{T-i})(T-i)+\sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(U)} \tag{43}$$

where U is as before a random variable with uniform distribution on the unit interval. From this expression we see that U_T^l is a comonotonic sum of lognormal random variables $X_{T-i}=e^{(r-\frac{\sigma^2}{2}R^2_{T-i})(T-i)+\sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(U)}$ multiplied by P_i. So from (26) it follows that for all $0<G<\infty$

$$\mathbb{E}\left[(G-U_T^l)_+\right]=-\sum_{i=0}^{T-1}P_ie^{r(T-i)}\Phi\left(-\sigma R_{T-i}\sqrt{T-i}+\Phi^{-1}(F_{U_T^l}(G))\right)+GF_{U_T^l}(G)$$

where $F_{U_T^l}(G)$ is computable as in (24) from

$$\sum_{i=0}^{T-1}P_ie^{(r-\frac{\sigma^2}{2}R^2_{T-i})(T-i)+\sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(F_{U_T^l}(G))}=G, \tag{44}$$

which gives (40).

From $U_T^l\leq_{cx}U_T$ it follows by definition of convex order that

$$V(T,G)=e^{-rT}\ \mathbb{E}\left[(G-U_T)_+\right]\geq e^{-rT}\ \mathbb{E}\left[(G-U_T^l)_+\right]. \tag{45}$$

So we conclude

$$
\begin{aligned}
V(T,G) &\geq e^{-rT}\ \mathbb{E}\left[(G-U_T^l)_+\right] \\
&= e^{-rT}\ \left(-\sum_{i=0}^{T-1} P_i e^{r(T-i)}\Phi\left(-\sigma R_{T-i}\sqrt{T-i}+\Phi^{-1}(F_{U_T^l}(G))\right)+GF_{U_T^l}(G)\right) \\
&= -\sum_{i=0}^{T-1} P_i e^{-ri}\Phi\left(-\sigma R_{T-i}\sqrt{T-i}+\Phi^{-1}(F_{U_T^l}(G))\right)+e^{-rT}\ GF_{U_T^l}(G),
\end{aligned}
$$

which gives (39). ∎

3.4 An improved upper bound

From chapter 2 we know that it is also possible to derive an improved upper bound by conditioning on a random variable Λ. This improved upper bound is given by:

Theorem 3.3 *Let $V(T,G)$ be the value of the guarantee with maturity T and guarantee level G. Then*

$$V(T,G) \leq$$

$$e^{-rT}\int_0^1 -\sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2}R_{T-i}^2)(T-i)+\sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(v)}\Phi\left(-\sigma\sqrt{T-i}\sqrt{1-R_{T-i}^2}+\Phi^{-1}(F_{U_T^u|V=v}(G))\right) +GF_{U_T^u|V=v}(G)dv,$$

where $F_{U_T^u|V=v}(G)$ follows from

$$\sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(v)+\sigma\sqrt{T-i}\sqrt{1-R_{T-i}^2}\Phi^{-1}(F_{U_T^u|V=v}(G))} = G. \tag{46}$$

Proof To derive this result we start with $\Lambda = W_T$ as the conditioning random variable. The reason to take this Λ different from the one in the previous section is that this choice will keep the computations tractable. From equations (33) and (34) we then have

$$\mathbb{E}\left[\sigma W_{T-i}\ |\Lambda=\lambda\right] = \sigma\mathbb{E}\left[W_{T-i}\right] + \sigma R_{T-i}\frac{\sigma_{W_{T-i}}}{\sigma_\Lambda}(\lambda - \mathbb{E}\left[\Lambda\right]) = \sigma R_{T-i}\frac{\sqrt{T-i}}{\sigma_\Lambda}\lambda$$

and

$$Var(\sigma W_{T-i}\ |\Lambda=\lambda) = \sigma^2\sigma_{W_{T-i}}^2(1-R_{T-i}^2) = \sigma^2(T-i)(1-R_{T-i}^2),$$

where

$$\begin{aligned} R_{T-i} = \frac{Cov\left(W_{T-i},\Lambda\right)}{\sigma_\Lambda\sqrt{T-i}} &= \frac{Cov\left(W_{T-i},W_T\right)}{\sigma_\Lambda\sqrt{T-i}} \\ &= \frac{T-i}{\sqrt{T-i}\sqrt{T}} = \frac{\sqrt{T-i}}{\sqrt{T}}. \end{aligned}$$

From this it follows that $(r-\frac{\sigma^2}{2})(T-i)+\sigma(W_{T-i}\mid\Lambda=\lambda)$ is normally distributed with parameters $\mu_i = (r-\frac{\sigma^2}{2})(T-i)+\sigma R_{T-i}\frac{\sqrt{T-i}}{\sigma_\Lambda}\lambda$ and $\sigma_i^2 = \sigma^2(T-i)(1-R_{T-i}^2)$. So $\left(P_i\frac{S_T}{S_i}|\Lambda=\lambda\right) = P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma W_{T-i}}|\Lambda=\lambda$ is P_i times a lognormally distributed random variable with parameters μ_i and σ_i^2. So we find from equation (18)

$$F^{-1}_{P_i\frac{S_T}{S_i}|\Lambda=\lambda}(p) = P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma R_{T-i}\sqrt{T-i}\frac{\lambda}{\sigma_\Lambda}+\sigma\sqrt{T-i}\sqrt{1-R_{T-i}^2}\Phi^{-1}(p)}.$$

So it follows that

$$(U_T^u|\Lambda=\lambda) =_d \sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma R_{T-i}\sqrt{T-i}\frac{\lambda}{\sigma_\Lambda}+\sigma\sqrt{T-i}\sqrt{1-R_{T-i}^2}\Phi^{-1}(U)}.$$

Since $\frac{\lambda}{\sigma_\Lambda}$ is standard normally distributed this is the same as

$$(U_T^u|V=v) =_d \sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(v)+\sigma\sqrt{T-i}\sqrt{1-R_{T-i}^2}\Phi^{-1}(U)},$$

where V is uniformly distributed on the unit interval and independent from U.

From $\sigma\sqrt{T-i}\sqrt{1-R_{T-i}^2} \geq 0$ for all $i = 0, \ldots, T-1$ we can conclude that, given $V = v$, U_T^u is a comonotonic sum with lognormal marginals with parameters $\mu_i = (r - \frac{\sigma^2}{2})(T-i) + \sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(v)$ and $\sigma_i^2 = \sigma^2(T-i)(1-R_{T-i}^2)$ multiplied by P_i. So from (26) it now follows that

$$\mathbb{E}\left[(G - U_T^u)_+ | V = v\right] =$$

$$-\sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2}R_{T-i}^2)(T-i)+\sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(v)}\Phi\left(-\sigma\sqrt{T-i}\sqrt{1-R_{T-i}^2} + \Phi^{-1}(F_{U_T^u|V=v}(G))\right) + GF_{U_T^u|V=v}(G).$$

From theorem 2.9 it follows that $U_T \leq_{cx} U_T^u$ and equation (29) tells us that U_T^u is indeed an improved upper bound. By definition of convex order we therefore find the improved upper bound as $V(T,G) = \mathbb{E}\left[(G - U_T)_+\right] \leq \mathbb{E}\left[(G - U_T^u)_+\right] =$

$$\int_0^1 -\sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2}R_{T-i}^2)(T-i)+\sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(v)}\Phi\left(-\sigma\sqrt{T-i}\sqrt{1-R_{T-i}^2} + \Phi^{-1}(F_{U_T^u|V=v}(G))\right)$$
$$+GF_{U_T^u|V=v}(G)dv.$$

Furthermore, equation (24) gives that $F_{U_T^u|V=v}(G)$ follows from

$$\sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma R_{T-i}\sqrt{T-i}\Phi^{-1}(v)+\sigma\sqrt{T-i}\sqrt{1-R_{T-i}^2}\Phi^{-1}(F_{U_T^u|V=v}(G))} = G.$$

Together the previous two equations give the improved upper bound as stated. ∎

In theory, the improved upper bound is more interesting than the original upper bound if we want to use the bounds as estimates for the value of the guarantee. However, as we can see above, the expression for the improved upper bound is much more complicated. In particular, there is an integral to compute, which makes it harder to implement. For this reason, we will not test the performance of the improved upper bound.

3.5 Another estimate for $V(T,G)$

The upper and lower bounds for $V(T,G)$ derived in the previous sections can serve as estimates for $V(T,G)$. These estimates were derived by using the relation $U_T^l \leq_{cx} U_T \leq_{cx} U_T^c$, which is equivalent to

$$\mathbb{E}\left[(t-U_T^l)_+\right] \leq \mathbb{E}\left[(t-U_T)_+\right] \leq \mathbb{E}\left[(t-U_T^c)_+\right], \quad \forall t \in \mathbb{R}.$$

It is also possible to improve these estimates by considering the random variable U_T^m defined by

$$\mathbb{E}\left[(t-U_T^m)_+\right] = z\mathbb{E}\left[(t-U_T^l)_+\right] + (1-z)\mathbb{E}\left[(t-U_T^c)_+\right], \quad 0 \leq z \leq 1, \tag{47}$$

which lies between the lower and upper bound. In this case the estimate for $V(T,G)$ is given by

$$\hat{V}(T,G) = e^{-rT}\mathbb{E}\left[(G-U_T^m)_+\right] = e^{-rT}\left(z\mathbb{E}\left[(G-U_T^l)_+\right] + (1-z)\mathbb{E}\left[(G-U_T^c)_+\right]\right), \tag{48}$$

where $\mathbb{E}\left[(G-U_T^l)_+\right]$ and $\mathbb{E}\left[(G-U_T^c)_+\right]$ can be computed as in the previous sections. Of course, the result depends on z and we will show that a reasonable choice for z is given by

$$z = \frac{Var(U_T^c) - Var(U_T)}{Var(U_T^c) - Var(U_T^l)}. \tag{49}$$

For this choice of z we will see that U_T^m has the same mean and variance as U_T.

We start with the identities

$$\lim_{t\to\infty}\left(t - \mathbb{E}\left[(t-X)_+\right]\right) = \mathbb{E}\left[X\right], \tag{50}$$

and

$$\mathbb{E}\left[(t-X)_+\right] = \int_{-\infty}^{t} F_X(x)dx, \tag{51}$$

which are proved in the appendix (equation (100) and the derivation thereafter). From this we find

$$\begin{aligned}
\mathbb{E}\left[U_T^m\right] = \lim_{t\to\infty}\left(t - \mathbb{E}\left[(t-U_T^m)_+\right]\right) &= z\lim_{t\to\infty}\left(t - \mathbb{E}\left[(t-U_T^l)_+\right]\right) + (1-z)\lim_{t\to\infty}\left(t - \mathbb{E}\left[(t-U_T^c)_+\right]\right) \\
&= z\mathbb{E}\left[U_T^l\right] + (1-z)\mathbb{E}\left[U_T^c\right]
\end{aligned}$$

Since U_T, U_T^l and U_T^c all have the same expectation we conclude $\mathbb{E}\left[U_T^m\right] = \mathbb{E}\left[U_T\right]$.

Of course we want U_T^m as close as possible to S. A natural choice of z is such that

$$\int_{-\infty}^{\infty} \mathbb{E}\left[(t-U_T^m)_+\right] - \mathbb{E}\left[(t-U_T)_+\right] dt = 0. \tag{52}$$

How such z could be determined? To answer this we recall equation (2), which tells us that $Var(X) = 2\int_{-\infty}^{\infty} \mathbb{E}\left[(X-t)_+\right] - (\mathbb{E}\left[X\right]-t)_+ dt$. It also holds that $\mathbb{E}\left[(X-t)_+\right] = \mathbb{E}\left[X\right] - t + \mathbb{E}\left[(t-X)_+\right]$. So we get

$$Var(X) = 2\int_{-\infty}^{\infty} \mathbb{E}\left[X\right] - t + \mathbb{E}\left[(t-X)_+\right] - (\mathbb{E}\left[X\right]-t)_+ dt.$$

Using

$$\mathbb{E}\left[X\right] - t - (\mathbb{E}\left[X\right]-t)_+ = \begin{cases} 0 & \text{if } t \leq \mathbb{E}\left[X\right] \\ \mathbb{E}\left[X\right] - t & \text{if } t > \mathbb{E}\left[X\right] \end{cases} = -(t-\mathbb{E}\left[X\right])_+ ,$$

we find

$$Var(X) = 2\int_{-\infty}^{\infty} \mathbb{E}\left[(t-X)_+\right] - (t-\mathbb{E}\left[X\right])_+ dt. \tag{53}$$

By definition of U_T^m we now see that (52) is equivalent to

$$\int_{-\infty}^{\infty} z\mathbb{E}\left[(t-U_T^l)_+\right] + (1-z)\mathbb{E}\left[(t-U_T^c)_+\right] - \mathbb{E}\left[(t-U_T)_+\right] dt = 0.$$

Adding and subtracting $(t-\mathbb{E}[U_T])_+$ in the integral and using $\mathbb{E}\left[U_T^l\right]=\mathbb{E}[U_T^m]=\mathbb{E}[U_T^c]=\mathbb{E}[U_T]$ we obtain that this is equivalent to

$$\begin{aligned} & z\int_{-\infty}^{\infty}\mathbb{E}\left[(t-U_T^l)_+\right]-(t-\mathbb{E}\left[U_T^l\right])_+dt \quad + \quad (1-z)\int_{-\infty}^{\infty}\mathbb{E}\left[(t-U_T^c)_+\right]-(t-\mathbb{E}\left[U_T^c\right])_+dt \\ = & \int_{-\infty}^{\infty}\mathbb{E}\left[(t-U_T)_+\right]-(t-\mathbb{E}\left[U_T\right])_+dt, \end{aligned}$$

which is by (53) equivalent to

$$z\frac{1}{2}Var(S^l)+(1-z)\frac{1}{2}Var(S^c)=\frac{1}{2}Var(S).$$

From this we conclude that (52) holds for

$$z=\frac{Var(U_T^c)-Var(U_T)}{Var(U_T^c)-Var(U_T^l)},$$

which is exactly (49).

For this z we have that (52) holds. Assuming this z, adding and subtracting $(t-\mathbb{E}[U_T])_+$ in (52) and using $\mathbb{E}[U_T^m]=\mathbb{E}[U_T]$ we find that in this situation (52) is equivalent to

$$\int_{-\infty}^{\infty}\mathbb{E}\left[(t-S^m)_+\right]-(t-\mathbb{E}\left[S^m\right])_+dt=\int_{-\infty}^{\infty}\mathbb{E}\left[(t-S)_+\right]-(t-\mathbb{E}\left[S\right])_+dt,$$

which is by (53) equivalent to $\frac{1}{2}Var(U_T^m)=\frac{1}{2}Var(U_T)$, so $Var(U_T^m)=Var(U_T)$. So indeed, for z given by (49) we have that U_T^m can be seen as a moment estimator.

Let us now apply the preceding derivations to the model of the previous sections. For this we recall that for X lognormally distributed with parameters μ_i and σ_i^2 we have $\mathbb{E}[X]=e^{\mu_i+\frac{1}{2}\sigma_i^2}$ and $Var(X)=e^{2\mu_i+\sigma_i^2}(e^{\sigma_i^2}-1)$. (see also equations (16) and (17))
From the definition of U_T^c we have

$$U_T^c=\sum_{i=0}^{T-1}P_ie^{(r-\frac{\sigma^2}{2})(T-i)+\sigma\sqrt{T-i}\Phi^{-1}(U)}.$$

Defining $\mu_i=(r-\frac{\sigma^2}{2})(T-i)$ and $\sigma_i^2=\sigma^2(T-i)$ we find $Var(U_T^c)$:

$$\begin{aligned} Var(U_T^c) &= Cov(U_T^c,U_T^c)=Cov\left(\sum_{i=0}^{T-1}P_ie^{\mu_i+\sigma_i\Phi^{-1}(U)},\sum_{j=0}^{T-1}P_je^{\mu_j+\sigma_j\Phi^{-1}(U)}\right) \\ &= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1}P_iP_je^{\mu_i+\mu_j}Cov(e^{\sigma_i\Phi^{-1}(U)},e^{\sigma_j\Phi^{-1}(U)}) \\ &= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1}P_iP_je^{\mu_i+\mu_j}\left(\mathbb{E}\left[e^{\sigma_i\Phi^{-1}(U)}e^{\sigma_j\Phi^{-1}(U)}\right]-\mathbb{E}\left[e^{\sigma_i\Phi^{-1}(U)}\right]\mathbb{E}\left[e^{\sigma_j\Phi^{-1}(U)}\right]\right) \\ &= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1}P_iP_je^{\mu_i+\mu_j}\left(\mathbb{E}\left[e^{(\sigma_i+\sigma_j)\Phi^{-1}(U)}\right]-e^{\frac{1}{2}\sigma_i^2}e^{\frac{1}{2}\sigma_j^2}\right) \\ &= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1}P_iP_je^{(r-\frac{\sigma^2}{2})(2T-i-j)}\left(e^{\frac{1}{2}(\sigma_i+\sigma_j)^2}-e^{\frac{1}{2}\sigma_i^2}e^{\frac{1}{2}\sigma_j^2}\right) \\ &= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1}P_iP_je^{r(2T-i-j)}\left(e^{\sigma_i\sigma_j}-1\right) \end{aligned}$$

$$= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1} P_i P_j e^{r(2T-i-j)} \left(e^{\sigma^2 \sqrt{(T-i)(T-j)}} - 1 \right).$$

From equation (43) and a similar computation we also find

$$Var(U_T^l) = \sum_{i=0}^{T-1}\sum_{j=0}^{T-1} P_i P_j e^{r(2T-i-j)} \left(e^{\sigma^2 R_{T-i} R_{T-j} \sqrt{(T-i)(T-j)}} - 1 \right).$$

Finally, $Var(U_T)$ is given by

$$\begin{aligned}
Var(U_T) &= Cov\left(\sum_{i=0}^{T-1} P_i e^{(r-\frac{\sigma^2}{2})(T-i)+\sigma W_{T-i}}, \sum_{j=0}^{T-1} P_j e^{(r-\frac{\sigma^2}{2})(T-j)+\sigma W_{T-j}} \right) \\
&= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1} P_i P_j e^{(r-\frac{\sigma^2}{2})(2T-i-j)} \left(\mathbb{E}\left[e^{\sigma W_{T-i}} e^{\sigma\sigma W_{T-j}} \right] - \mathbb{E}\left[e^{\sigma W_{T-i}} \right] \mathbb{E}\left[e^{\sigma W_{T-j}} \right] \right) \\
&= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1} P_i P_j e^{(r-\frac{\sigma^2}{2})(2T-i-j)} \left(\mathbb{E}\left[e^{\sigma(W_{T-i}+W_{T-j})} \right] - e^{\frac{1}{2}\sigma^2(T-i)} e^{\frac{1}{2}\sigma^2(T-j)} \right) \\
&= \sum_{i=0}^{T-1}\sum_{j=0}^{T-1} P_i P_j e^{r(2T-i-j)} \left(e^{\sigma^2 \min((T-i),(T-j))} - 1 \right).
\end{aligned}$$

This gives all the ingredients to compute z as given by (49) explicitly. The estimate for $V(T,G)$ is then given by (48), where $\mathbb{E}\left[(G-U_T^l)_+\right]$ and $\mathbb{E}\left[(G-U_T^c)_+\right]$ have to be computed as in the previous sections. In summary, the estimator for the value of the guarantee is given by:

Theorem 3.4 *Let $V(T,G)$ be the value of the guarantee with maturity T and guarantee level G. Then an estimator of $V(T,G)$ is given by*

$$\hat{V}(T,G) = e^{-rT}\left(z\mathbb{E}\left[(G-U_T^l)_+\right] + (1-z)\mathbb{E}\left[(G-U_T^c)_+\right] \right),$$

where z is given by

$$z = \frac{Var(U_T^c) - Var(U_T)}{Var(U_T^c) - Var(U_T^l)}.$$

3.6 Numerical results

This section provides numerical results for the bounds derived in this chapter. In order to judge the quality of the bounds, we compare them to the value of the guarantee obtained by Monte Carlo simulation. The Monte Carlo simulations are performed in ALS (copyright by Ortec), which is the programme that SNS Reaal uses for valuing the guarantees. We have tested the bounds for 5 different model points which are also used by SNS for valuing their portfolio of guarantees. A model point is a package of individual contracts and they are used so that not each contract needs to be valued separately. A model point is characterized by gross premium, investment percentage, maturity, current fund value and guaranteed value which together determine the option value. The gross premium $GP(i)$ and the investment percentage give the net premium that is invested in the investment fund. If we denote the investment percentage at time i by $I(i)$, the net premium $NP(i)$ invested in the fund at time i is given by $NP(i) = GP(i) \times I(i)$. Here $(1-I(i)) \times GP(i)$ are (fixed) costs that are subtracted from the gross premium and that SNS Reaal uses for example to finance guarantees for which they have to pay. The investment fund in which the premiums are invested is the SNS Garantie Mixfonds. This is a fund consisting of 70 % fixed income securities and 30 % stocks. The management costs of this fund are 0.82 % (denoted c = 0.0082), which means that 0.82 % of the fund value at the end of each year (or the beginning of each year, since the fund value does not change between December 31 and January 1) is subtracted. This is something we did not consider so far, but we can deal with this as follows. Define

$$\widetilde{P}_i = NP_i - c\sum_{j=0}^{i-1} \widetilde{P}_j \frac{S_i}{S_j}.$$

Then $\sum_{j=0}^{i-1} \widetilde{P}_j \frac{S_i}{S_j}$ is the fund value at the end of year $i-1$ (=beginning of year i) and $c\sum_{j=0}^{i-1} \widetilde{P}_j \frac{S_i}{S_j}$ are the management costs that are subtracted at the end of year $i-1$ (=beginning of year i). So $\widetilde{P}_i$ is the net premium at time i minus the management costs that are subtracted at the same time the net premium is invested in the fund. This means that $\widetilde{P}_i$ is the net value added to the fund. According to our assumption, the customer does not pay a premium at maturity T ($NP(T) = 0$). However, at maturity, management costs are still deducted for the last year. So we get $\widetilde{P}_T = -c\sum_{j=0}^{T-1} \widetilde{P}_j \frac{S_T}{S_j}$ and the payoff of the guarantee will be

$$\left(G - \sum_{i=0}^{T}\left(NP(i) - c\sum_{j=0}^{i-1} \widetilde{P}_j \frac{S_i}{S_j}\right)\frac{S_T}{S_i}\right)_+ .$$

In this form the option payoff becomes very complicated, but Schrager and Pelsser (2004) show that this is equivalent to

$$\left(G - \sum_{i=0}^{T} P_i \frac{S_T}{S_i}\right)_+ = \left(G - \sum_{i=0}^{T-1} P_i \frac{S_T}{S_i}\right)_+$$

where $P_i = NP(i) \times (1-c)^{T-i}$ and where the equality follows since $NP(T) = 0$. So what we did in the numerical comparison is that we adapted the net premiums as above and we used the P_i as defined in the computation of our bounds. This we did because this is the way it is done in practice and also the Monte Carlo simulation can work with the management costs deduction every year.

Another subtlety we had to deal with, is the fact that the model points we considered did not contain contracts that started today, but somewhere earlier. So there already is a certain amount of accumulated premiums and returns at time 0. To make this fit in our model, we simply have to add this current fund value to the premium in year 0.

Since we consider it inappropriate to disclose all the information on the model points we used, we will only give the time to maturity and the guaranteed value. We do this, because these parameters are the main characteristics of a model point. They are as follows:

Model point	Time to maturity	Guaranteed value
1	26	1,217,993.65
2	22	10,019,706.79
3	17	5,412,609.16
4	15	20,097.40
5	6	132,834.02

Table 1: Time to maturity in years and guaranteed value in Euro's of the model points.

For the risk free rate we used 3.922 % (the continuous equivalent of an annual rate of 4 %) and the volatility (σ) of the investment fund is approximately 6%. Since the volatility varies a lot for different funds, we also studied a scenario with a volatility of 20 %. For the Monte Carlo simulation we used 10,000 simulations and the results are as follows.

σ	Model point	UB	LB	EST	MC (s.e.)
0.06	1	14,274	10,902	10,912	10,560 (241)
0.06	2	114,375	89,083	89,147	85,804 (2,135)
0.06	3	71,688	57,402	57,433	55,276 (1,328)
0.06	4	374	304	303	298 (6)
0.06	5	2,280	2,114	2,114	2,081 (43)
0.2	1	109,192	94,618	95,548	94,079 (570)
0.2	2	992,752	874,392	880,354	871,283 (5,734)
0.2	3	593,941	528,565	531,099	525,861 (3,722)
0.2	4	2,290	2,021	2,024	2,017 (13)
0.2	5	15,469	14,787	14,791	14,771 (125)

Table 2: Upper bound (UB), lower bound (LB) and the estimate of section 3.5 (Est) for the value of the guarantee in the different model points, compared to Monte Carlo estimates (MC) and their standard error (s.e.). All numbers are in Euro's.

From these results we immediately see that the lower bound is a very accurate estimate for the value of the guarantee for all different maturities and guaranteed values. In all cases, the value of the lower bound is even higher than the value of the Monte Carlo, which indicates that the realized Monte Carlo value is below the true value of the guarantee. So the scenario's in the Monte Carlo simulation have given an above average investment return.

Also, in all cases, the lower bound is less than two standard deviations away from the Monte Carlo simulation so it lies in the 95 % confidence interval. The estimate of section 3.6 is always very close to the lower bound. Even though it might be a bit more accurate we would advise to work with the lower bound, since this avoids the computation of the moments of the theoretical bounds.

The upper bound is a bad estimate for the option value. This is not a surprise, since the assumption that the $P_i \frac{S_T}{S_i}$ for $i = 0, \ldots, T-1$ are comonotonic is crude.

4 A model with stochastic interest rates

In chapter 3 we derived bounds on the value of the guarantee under the assumption that the value of the investment fund follows a geometric Brownian motion. This is a standard assumption in the financial industry, but one of the disadvantages is that the interest rate is fixed. In practice, interest rates are stochastic. Working with a stochastic interest rate instead of a fixed interest rate has the biggest impact if we study products which have a long maturity. This is exactly the case with unit-linked insurance, so it makes sense to use a model with stochastic interest rates.
In this chapter we will first give some general theory about stochastic interest rate models. Then we will make some modelling assumptions and apply the theory to derive the distribution of $\frac{S_T}{S_i}$ under this modelling assumptions. This distribution will appear to be lognormal again and therefore we will be able to derive bounds as in the previous chapter. As a concrete model, we will consider the Hull-White-Black-Scholes model.

4.1 General theory

Throughout this thesis we used risk neutral valuation to derive bounds on the value of the guarantee. This is based on the fact that in a complete and arbitrage free market the unique value of any financial claim equals the expectation of the payoff normalized by the money market account under some equivalent measure. Under this measure (from now on denoted by $\mathbb{Q}$) the expected return on all assets equals the risk free rate. Furthermore, the price processes of the normalized assets are martingales under $\mathbb{Q}$. The normalizing asset (in this setting the money market account) is also called the numeraire. Geman et al. (1995) show how not only the money market account, but every strictly positive self-financing portfolio of traded assets, can be used as numeraire. Their change of numeraire theorem shows how an expectation under a probability measure $\mathbb{Q}^N$ associated with numeraire N is related to an expectation under an equivalent probability measure $\mathbb{Q}^M$ associated with numerarie M. More specifically, their theorem states that in an arbitrage-free and complete market, for any numeraires N and M with associated measures $\mathbb{Q}^N$ and $\mathbb{Q}^M$, the following holds for the price of an asset H at time $t \leq T$:

$$H(t) = N(t)\mathbb{E}_t^N\left[\frac{H(T)}{N(T)}\right] = M(t)\mathbb{E}_t^M\left[\frac{H(T)}{M(T)}\right],$$

where $\mathbb{E}_t^N$ and $\mathbb{E}_t^M$ denote conditional expectations on the information available at time t under $\mathbb{Q}^N$ and $\mathbb{Q}^M$ respectively. The Radon-Nikodym derivative associated with a change of measure from $\mathbb{Q}^N$ to $\mathbb{Q}^M$ is given by

$$\frac{d\mathbb{Q}^M}{d\mathbb{Q}^N} = \frac{\frac{M(T)}{M(t)}}{\frac{N(T)}{N(t)}}.$$

So if the price of a payoff $H(T)$, known at time T, can be calculated by taking a risk-neutral expectation it can also be calculated by changing numeraires.

So far we have modelled the value of the investment fund S_t as

$$dS_t = rS_t dt + \sigma S_t dW_t,$$

where r is the (constant) risk free interest rate, σ the volatility and W_t a standard Brownian motion under the risk neutral measure. In this model, with a constant risk free interest rate, the value of the money market account at time t is given by

$$B_t = e^{rt}.$$

With the money market account the numeraire associated with the risk neutral measure $\mathbb{Q}$, we have that the process $\frac{S_t}{B_t}$ is a martingale under $\mathbb{Q}$. Furthermore, the price of a zero coupon bond with maturity T (paying 1 unit of currency at time T) at time t (denoted $D(t,T)$) is given by

$$D(t,T) = e^{-r(T-t)}.$$

If we now assume the risk free rate r_t to be stochastic, the value of the money market account at time t is given by

$$B_t = e^{\int_0^t r_s ds},$$

and the price of a zero coupon bond is given by

$$D(t,T) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s ds}\right].$$

Under suitable assumptions it holds that for each maturity T the process $\frac{D(t,T)}{B_t}$ is a martingale under $\mathbb{Q}$. Now, $D(t,T)$ is strictly positive and therefore we can also use $D(t,T)$ as a numeraire. The associated measure is called the T-forward measure, denoted $\mathbb{Q}^T$. From the change of numeraire theorem we now have that the forward stock process $F_t^T = \frac{S_t}{D_{t,T}}$ and the forward bond process $D^T(t,U) = \frac{D(t,U)}{D(t,T)}$ for each $t < U \leq T$ are martingales under $\mathbb{Q}^T$.
Our choice, which is a common choice, is to model them as lognormal martingales:

$$dF_t^T = \sigma_F(t) F_t^T dW_t^T, \tag{54}$$

and

$$dD^T(t,U) = \sigma_U(t) D^T(t,U) dW_t^{UT}, \tag{55}$$

where W_t^T and W_t^{UT} are brownian motions under the T-forward measure and $\sigma_F(t)$ and $\sigma_U(t)$ are deterministic functions of time. Correlations between those Brownian motions are given by $dW_t^T dW_t^{UT} = \rho_{F,U}(t)dt$ and $dW_t^{UT} dW_t^{VT} = \rho_{UV}(t)dt$. When using time points t_i and t_j we will write W_t^{iT}, $\sigma_i(t)$, $\rho_{F,i}(t)$ and $\rho_{ij}(t)$ for $W_t^{t_iT}$, $\sigma_{t_i}(t)$, $\rho_{F,t_i}(t)$ and $\rho_{t_it_j}(t)$.

We are now able to derive the distribution of $\frac{S_T}{S_{t_i}}$ at time t under the dynamics (54) and (55). At this moment, it is good to stress that this is really what we are after. The dynamics (54) and (55) represent a model with stochastic interest and once we know the distribution of $\frac{S_T}{S_{t_i}}$ for each i we can go back to valuing our guarantee. It will turn out that $\frac{S_T}{S_{t_i}}$ is lognormal, so we can basically repeat chapter 3 once we know the parameters of the lognormal distribution.

First suppose that $t \leq t_i$. From (54) it follows that

$$F_{t_i}^T = F_t^T e^{-\frac{1}{2}\int_t^{t_i} \sigma_F^2(s)ds + \int_t^{t_i} \sigma_F(s) dW_s^T}$$

Plugging in the definition of F_t^T we find

$$\frac{S_{t_i}}{D(t_i,T)} = \frac{S_t}{D(t,T)} e^{-\frac{1}{2}\int_t^{t_i} \sigma_F^2(s)ds + \int_t^{t_i} \sigma_F(s) dW_s^T},$$

and similarly

$$\frac{S_T}{D(T,T)} = S_T = \frac{S_t}{D(t,T)} e^{-\frac{1}{2}\int_t^{T} \sigma_F^2(s)ds + \int_t^{T} \sigma_F(s) dW_s^T}.$$

Dividing the previous two equations we find

$$\frac{S_T}{S_{t_i}} = \frac{1}{D(t_i,T)} e^{-\frac{1}{2}\int_{t_i}^{T} \sigma_F^2(s)ds + \int_{t_i}^{T} \sigma_F(s) dW_s^T}. \tag{56}$$

From (55) we obtain for $t' \geq t_i \geq t$

$$\frac{D(t_i,t')}{D(t_i,T)} = \frac{D(t,t')}{D(t,T)} e^{-\frac{1}{2}\int_t^{t_i} \sigma_{t'}^2(s)ds + \int_t^{t_i} \sigma_{t'}(s) dW_s^{t'T}}.$$

Taking $t' = t_i$ we get

$$\frac{1}{D(t_i,T)} = \frac{D(t,t_i)}{D(t,T)} e^{-\frac{1}{2}\int_t^{t_i} \sigma_i^2(s)ds + \int_t^{t_i} \sigma_i(s) dW_s^{iT}}.$$

If we insert this result in (56) we obtain

$$\frac{S_T}{S_{t_i}} = \frac{D(t,t_i)}{D(t,T)} e^{-\frac{1}{2}\int_t^{t_i} \sigma_i^2(s)ds - \frac{1}{2}\int_{t_i}^T \sigma_F^2(s)ds + \int_t^{t_i} \sigma_i(s)dW_s^{iT} + \int_{t_i}^T \sigma_F(s)dW_s^T}. \tag{57}$$

For $t_i \le t$ we can do a similar derivation and find

$$\frac{S_T}{S_{t_i}} = \frac{1}{S_{t_i}} \frac{S_t}{D(t,T)} e^{-\frac{1}{2}\int_t^T \sigma_F^2(s)ds + \int_t^T \sigma_F(s)dW_s^T}. \tag{58}$$

In order to keep things clear, we simplify the notation. We define

$$\int_t^T \bar{\sigma_i}(s)d\bar{W}_s^i = \int_t^{max(t,t_i)} \sigma_i(s)dW_s^{iT} + \int_{max(t,t_i)}^T \sigma_F(s)dW_s^T,$$

$$\bar{\mu_i}(t) = \frac{D(t,t_i)}{D(t,T)} I_{[0,t_i]}(t) + \frac{1}{S_{t_i}} \frac{S_t}{D(t,T)} I_{[t_i,T]}(t).$$

From standard properties of Brownian motion it follows that $\int_t^{max(t,t_i)} \sigma_i(s)dW_s^{iT}$ and $\int_{max(t,t_i)}^T \sigma_F(s)dW_s^T$ are independent normally distributed. So it follows that

$$\begin{aligned}
\int_t^T \bar{\sigma_i}^2(s)ds = Var(\int_t^T \bar{\sigma_i}(s)d\bar{W}_s^i) &= Var(\int_t^{max(t,t_i)} \sigma_i(s)dW_s^{iT}) + Var(\int_{max(t,t_i)}^T \sigma_F(s)dW_s^T) \\
&= \int_t^{max(t,t_i)} \sigma_i^2(s)ds + \int_{max(t,t_i)}^T \sigma_F^2(s)ds.
\end{aligned}$$

Using the developed notation, we can now write

$$\frac{S_T}{S_{t_i}} = \bar{\mu_i}(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma_i}^2(s)ds + \int_t^T \bar{\sigma_i}(s)d\bar{W}_s^i}. \tag{59}$$

From this expression we see that $\frac{S_T}{S_{t_i}}$ is indeed lognormally distributed. Therefore we should be able to use the theory of comonotonicity to derive bounds on the value of the guarantee at time t

$$V_t(T,G) = e^{\int_0^t r_s ds} \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_0^T r_s ds} (G - \sum_{i=0}^{T-1} P_i \frac{S_T}{S_i})_+ \right] = D(t,T) \mathbb{E}_t^{\mathbb{Q}^T} \left[(G - \sum_{i=0}^{T-1} P_i \frac{S_T}{S_i})_+ \right],$$

where the last equality follows from the change of numeraire theorem.

4.2 Bounds under stochastic interest rates

In the previous section we have found the distribution of $\frac{S_T}{S_{t_i}}$ in a model with stochastic interest given by (54) and (55). As noted, the distribution is lognormal, so applying the theory of comonotonicity will enable us to derive bounds on the value of the guarantee. Here we will first derive an upper bound and then proceed with the lower bound. This we will do for an arbitrary time $0 \leq t \leq T$, i.e. we assume that the contract started at time 0 and that we are now at time t. So we already know the investment fund values between time 0 and t. Similar to chapter three we will denote $U_T = \sum_{i=0}^{T-1} P_i \frac{S_T}{S_{t_i}}$, $U_T^c = \sum_{i=0}^{T-1} F^{-1}_{P_i \frac{S_T}{S_{t_i}}}(U)$ and $U_T^l = \sum_{i=0}^{T-1} \mathbb{E}_t \left[P_i \frac{S_T}{S_i} \mid \Lambda \right]$, given the information at time t.
An upper bound on the value of the guarantee is given by

Theorem 4.1 *Let $V_t(T,G)$ be the value of the guarantee with maturity T and guarantee level G at time t. Then*

$$V_t(T,G) \leq D(t,T) \left[-\sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) \Phi \left(-\sqrt{\int_t^T \bar{\sigma_i}^2(s)ds} + \Phi^{-1}(F_{U_T^c}(G)) \right) + G F_{U_T^c}(G) \right], \tag{60}$$

where $F_{U_T^c}(G)$ follows from solving

$$\sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma_i}^2(s)ds + \sqrt{\int_t^T \bar{\sigma_i}^2(s)ds}\ \Phi^{-1}(F_{U_T^c}(G))} = G. \tag{61}$$

To derive this we start with the following observation

$$\int_t^T \bar{\sigma_i}(s) d\bar{W}_s^i =_d \sqrt{\int_t^T \bar{\sigma_i}^2(s)ds}\ \Phi^{-1}(U),$$

where U is uniformly distributed on the unit interval under $\mathbb{Q}^T$.
If we combine this with (59), we find that the comonotonic counter part of $U_T = \sum_{i=0}^{T-1} P_i \frac{S_T}{S_{t_i}}$, given the information at time t, is

$$U_T^c \quad = \quad \sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma_i}^2(s)ds + \sqrt{\int_t^T \bar{\sigma_i}^2(s)ds}\ \Phi^{-1}(U)}.$$

Using (26) this gives

$$\begin{aligned} V_t(T,G) &= D(t,T)\mathbb{E}_t^{\mathbb{Q}^T}\left[(G-U_T)_+\right] \\ &\leq D(t,T)\mathbb{E}_t^{\mathbb{Q}^T}\left[(G-U_T^c)_+\right] \\ &= D(t,T) \left[-\sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) \Phi \left(-\sqrt{\int_t^T \bar{\sigma_i}^2(s)ds} + \Phi^{-1}(F_{U_T^c}(G)) \right) + G F_{U_T^c}(G) \right], \end{aligned}$$

which gives (60). Also from equation (14) we have that $F_{U_T^c}(G)$ follows from $F^{-1}_{U_T^c}(F_{U_T^c}(G)) = G$, or equivalently

$$\sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma_i}^2(s)ds + \sqrt{\int_t^T \bar{\sigma_i}^2(s)ds}\ \Phi^{-1}(F_{U_T^c}(G))} = G,$$

which gives (61).
A lower bound on the value of the guarantee is given by

Theorem 4.2 *Let $V_t(T,G)$ be the value of the guarantee with maturity T and guarantee level G at time t. Then*

$$V_t(T,G) \geq D(t,T) \left[-\sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) \Phi \left(-\mu_{i|Z}(t) + \Phi^{-1}(F_{U_T^l}(G)) \right) + G F_{U_T^l}(G) \right], \tag{62}$$

where $F_{U_T^l}(G)$ follows from solving

$$\sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) e^{\mu_{i|Z}(t)\Phi^{-1}(F_{U_t^l}(G)) - \frac{1}{2}\mu_{i|Z}^2(t)} = G. \tag{63}$$

Furthermore, $\mu_{i|Z}(t)$ is given by $\mu_{i|Z}(t) = \mathbb{E}_t^{\mathbb{Q}^T}\left[Z \int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i\right]$, where Z is a standard normally distributed random variable.

Explicit expressions for Z and $\mu_{i|Z}(t)$ are given below.

To derive this result we proceed as in chapter 3.3. The only difference is that we want the conditioning random variable Λ to be standard normally distributed and we will call it Z. So we have to compute $\mathbb{E}_t^{\mathbb{Q}^T}\left[P_i \frac{S_T}{S_{t_i}} | \ Z\right]$, for a suitable conditioning random variable Z. If Z is standard normally distributed we get by equations (33) and (34) (where we plug in the definition of $r(Y, \Lambda)$)

$$\mathbb{E}_t^{\mathbb{Q}^T}\left[\int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i | \ Z\right] = Cov\left(\int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i, Z\right) Z = \mathbb{E}_t^{\mathbb{Q}^T}\left[Z \int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i\right] Z,$$

and

$$\begin{aligned} Var(\int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i | \ Z) &= \int_t^T \bar{\sigma}_i{}^2(s) ds - Cov\left(\int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i, Z\right)^2 \\ &= \int_t^T \bar{\sigma}_i{}^2(s) ds - \mathbb{E}_t^{\mathbb{Q}^T}\left[Z \int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i\right]^2. \end{aligned}$$

We now define $\mu_{i|Z}(t) = \mathbb{E}_t^{\mathbb{Q}^T}\left[Z \int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i\right]$. Applying (16) then gives

$$\begin{aligned} U_T^l = \mathbb{E}_t^{\mathbb{Q}^T}\left[P_i \frac{S_T}{S_{t_i}} | \ Z\right] &= P_i \bar{\mu}_i(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma}_i{}^2(s) ds} \mathbb{E}_t^{\mathbb{Q}^T}\left[e^{\int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i} | Z\right] \\ &= P_i \bar{\mu}_i(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma}_i{}^2(s) ds} e^{\mu_{i|Z}(t) Z + \frac{1}{2}\int_t^T \bar{\sigma}_i{}^2(s) ds - \frac{1}{2}\mu_{i|Z}^2(t)} \\ &= P_i \bar{\mu}_i(t) e^{\mu_{i|Z}(t) Z - \frac{1}{2}\mu_{i|Z}^2(t)}. \end{aligned}$$

If $\mu_{i|Z}(t)$ is positive for all i, then $U_T^l = \sum_{i=0}^{T-1} \mathbb{E}_t^{\mathbb{Q}^T}\left[P_i \frac{S_T}{S_{t_i}} | \ Z\right]$ will be comonotone (with lognormal marginals) and this will enable us to compute the lower bound. So once we know $\mu_{i|Z}(t) = \mathbb{E}_t^{\mathbb{Q}^T}\left[Z \int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i\right]$, we can compute the lower bound.

As in chapter 3 we will now choose Z to be a first order approximation of U_T. The first order approximation is

$$\begin{aligned} U_T &= \sum_{i=0}^{T-1} P_i \frac{S_T}{S_{t_i}} = \sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma}_i{}^2(s) ds + \int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i} \\ &\approx \sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma}_i{}^2(s) ds} (1 + \int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i) \\ &= C + \sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma}_i{}^2(s) ds} \int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i, \end{aligned}$$

where C is the appropriate constant. So as a conditioning random variable we can take

$$Z = \sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma}_i{}^2(s) ds} \int_t^T \bar{\sigma}_i(s) d\bar{W}_s^i.$$

This Z is normally distributed but, we want Z to be standard normally distributed. Therefore, we have to divide by the standard deviation. This standard deviation will be denoted α_t and the conditioning standard normal random variable becomes

$$Z = \frac{1}{\alpha_t} \sum_{i=0}^{T-1} P_i \bar{\mu_i}(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma_i}^2(s)ds} \int_t^T \bar{\sigma_i}(s) d\bar{W}_s^i.$$

Performing the calculation we find

$$\alpha_t = \sum_{i=0}^{T-1}\sum_{j=0}^{T-1} P_i \bar{\mu_i}(t) P_j \bar{\mu_j}(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma_i}^2(s)ds - \frac{1}{2}\int_t^T \bar{\sigma_j}^2(s)ds} Cov\left(\int_t^T \bar{\sigma_i}(s) d\bar{W}_s^i, \int_t^T \bar{\sigma_j}(s) d\bar{W}_s^j\right),$$

where for $i \leq j$ (i.e. $t_i \leq t_j$) we have

$$Cov\left(\int_t^T \bar{\sigma_i}(s) d\bar{W}_s^i, \int_t^T \bar{\sigma_j}(s) d\bar{W}_s^j\right) =$$

$$Cov\left(\int_t^{max(t,t_i)} \sigma_i(s) dW_s^{iT} + \int_{max(t,t_i)}^T \sigma_F(s) dW_s^T, \int_t^{max(t,t_j)} \sigma_j(s) dW_s^{jT} + \int_{max(t,t_j)}^T \sigma_F(s) dW_s^T\right) =$$

$$\int_t^{\max(t,t_i)} \rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds + \int_{\max(t,t_i)}^{\max(t,t_j)} \rho_{F,j}(s)\sigma_F(s)\sigma_j(s)ds + \int_{\max(t,t_j)}^T \sigma_F^2(s)ds.$$

As mentioned before we also need an explicit expression for $\mu_{i|Z}(t) = \mathbb{E}_t^{\mathbb{Q}^T}\left[Z\int_t^T \bar{\sigma_i}(s)d\bar{W}_s^i\right]$. Plugging in all the definitions and doing the calculation gives

$$\begin{aligned}
\mu_{i|Z}(t) &= \mathbb{E}_t^{\mathbb{Q}^T}\left[Z\int_t^T \bar{\sigma_i}(s)d\bar{W}_s^i\right] \\
&= \mathbb{E}_t^{\mathbb{Q}^T}\left[\frac{1}{\alpha_t}\left(\sum_{j=0}^{T-1} P_j\bar{\mu_j}(t)e^{-\frac{1}{2}\int_t^T \bar{\sigma_j}^2(s)ds}\int_t^T \bar{\sigma_j}(s)d\bar{W}_s^j\right)\int_t^T \bar{\sigma_i}(s)d\bar{W}_s^i\right] \\
&= \frac{1}{\alpha_t}\sum_{j=0}^{i} P_j\bar{\mu_j}(t)e^{-\frac{1}{2}\int_t^T \bar{\sigma_j}^2(s)ds}\left[\int_t^{\max(t,t_j)} \rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds + \int_{\max(t,t_j)}^{\max(t,t_i)} \rho_{F,i}(s)\sigma_F(s)\sigma_i(s)ds\right. \\
&+ \left.\int_{\max(t,t_i)}^T \sigma_F^2(s)ds\right] + \\
&\quad \frac{1}{\alpha_t}\sum_{j=i+1}^{T-1} P_j\bar{\mu_j}(t)e^{-\frac{1}{2}\int_t^T \bar{\sigma_j}^2(s)ds}\left[\int_t^{\max(t,t_i)} \rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds + \int_{\max(t,t_i)}^{\max(t,t_j)} \rho_{F,j}(s)\sigma_F(s)\sigma_j(s)ds\right. \\
&+ \left.\int_{\max(t,t_j)}^T \sigma_F^2(s)ds\right].
\end{aligned}$$

As noted in Schrager and Pelsser (2006), for reasonable values of the correlations (mainly the correlation between the forward stock and forward bond processes) the $\mu_{i|Z}(t)$ are positive for all i. Assuming this we can compute the lower bound. We have

$$V_t(T,G) = D(t,T)\mathbb{E}_t^{\mathbb{Q}^T}\left[(G-U_T)_+\right] \geq D(t,T)\mathbb{E}_t^{\mathbb{Q}^T}\left[(G-U_T^l)_+\right],$$

where $U_T^l =_d \sum_{i=0}^{T-1} P_i\bar{\mu_i}(t)e^{\mu_{i|Z}(t)\Phi^{-1}(U)-\frac{1}{2}\mu_{i|Z}^2(t)}$ and $\mu_{i|Z}(t)$ as above. By (26) we conclude

$$V_t(T,G) \geq D(t,T)\left[-\sum_{i=0}^{T-1} P_i\bar{\mu_i}(t)\Phi\left(-\mu_{i|Z}(t) + \Phi^{-1}(F_{U_T^l}(G))\right) + GF_{U_T^l}(G)\right],$$

which is (62). Also, by (14), we have that $F_{U_T^l}(G)$ follows from

$$\sum_{i=0}^{T-1} P_i \bar{\mu}_i(t) e^{\mu_{i|Z}(t)\Phi^{-1}(F_{U_T^l}(G)) - \frac{1}{2}\mu_{i|Z}^2(t)} = G,$$

which gives (63).

4.3 The Hull-White-Black-Scholes model

In the previous section we derived bounds on the value of the guarantee for a general model with stochastic interest given by (54) and (55). In this section we will focus on a specific model, namely the Hull-White-Black-Scholes model (HWBS model). We will start with the definition of the model and then we will show that the model satisfies (54) and (55) for certain values of the parameters. Using these parameters we are able to derive all relevant expressions that we need to compute the bounds. So this means that we can compute a lower bound and an upper bound on the value of the guarantee in this particular model. This is very interesting, since in practice this model is used to compute the value of the guarantee. However, in practice it is done by simulation and we do it with bounds that can be used as approximations.

The Hull-White-Black-Scholes model is a combination of the Black-Scholes model for the development of the value of the investment fund S_t and the Hull-White model for the development of the short rate r_t. More specifically, the dynamics under the risk neutral measure are given by the following stochastic differential equations

$$dS_t = r_t S_t dt + \sqrt{1-\rho^2}\sigma_S S_t dW_{1,t} + \rho\sigma_S S_t dW_{2,t}, \tag{64}$$

$$dr_t = (\theta_t - a r_t)dt + \sigma_r dW_{2,t}, \tag{65}$$

where ρ is the correlation between the stock and short rate, σ_S is the volatility of the stock, a is the mean reversion parameter, σ_r is the volatility of the short rate, $W_{1,t}$ and $W_{2,t}$ are Brownian motions under the risk neutral measure and θ_t is a function of time determined by the initial forward rates observed in the market.

To determine the dynamics of the T-forward stock price $F_t^T = \frac{S_t}{D_{t,T}}$ and the T-forward bond price with maturity t_i, $D^T(t,t_i) = \frac{D(t,t_i)}{D(t,T)}$ we need to know the bond prices. It is well known that in the Hull-White model for the short rate these are given by

$$D(t,T) = e^{A(t,T)-B(t,T)r_t},$$

where

$$\begin{aligned}
B(t,T) &= \frac{1}{a}(1-e^{-a(T-t)}) \\
A(t,T) &= \frac{\sigma_r^2}{2}\int_t^T B(s,T)^2 ds - \int_t^T \theta_s B(s,T)ds.
\end{aligned}$$

Furthermore, for $\theta(t) = \frac{\partial f(0,t)}{\partial t} + a\ f(0,t) + \frac{\sigma_r^2}{2a}(1-e^{-2at})$ the model fits the initial term structure, where $f(0,t)$ is the forward rate curve as observed in market.

Applying Ito's lemma it follows that

$$\begin{aligned}
dD(t,T) &= D(t,T)\left(-\frac{\sigma_r^2}{2}B(t,T)^2 + \theta_t B(t,T) + r_t e^{-a(T-t)}\right)dt - D(t,T)B(t,T)dr_t \\
&+ \frac{1}{2}B(t,T)^2 D(t,T)d < r_t, r_t > \\
&= D(t,T)\left(-\frac{\sigma_r^2}{2}B(t,T)^2 + \theta_t B(t,T) + r_t e^{-a(T-t)}\right)dt \\
&- D(t,T)B(t,T)\left((\theta_t - a r_t)dt + \sigma_r dW_{2,t}\right) + \frac{1}{2}B(t,T)^2 D(t,T)\sigma_r^2 dt,
\end{aligned}$$

where $< X, Y >$ denotes the quadratic covariation process of X and Y. Rearranging the terms we conclude

$$\frac{dD(t,T)}{D(t,T)} = \left(r_t e^{-a(T-t)} + aB(t,T)r_t\right)dt - \sigma_r B(t,T)dW_{2,t} = r_t dt - \sigma_r B(t,T)dW_{2,t}. \tag{66}$$

To derive the dynamics of the T-forward bond price, we now look at $d\left(\frac{D(t,t_i)}{D(t,T)}\right)$. Applying Ito's lemma we get

$$
\begin{aligned}
d\left(\frac{D(t,t_i)}{D(t,T)}\right) &= \frac{1}{D(t,T)}dD(t,t_i) - \frac{D(t,t_i)}{D(t,T)^2}dD(t,T) - \frac{1}{D(t,T)^2}d < D(t,t_i), D(t,T) > \\
&+ \frac{1}{2}\frac{2D(t,t_i)}{D(t,T)^3}d < D(t,T), D(t,T) > .
\end{aligned}
$$

Plugging in (66) it follows that

$$
\begin{aligned}
d\left(\frac{D(t,t_i)}{D(t,T)}\right) &= \frac{1}{D(t,T)}\left(r_t D(t,t_i)dt - \sigma_r D(t,t_i)B(t,t_i)dW_{2,t}\right) \\
&- \frac{D(t,t_i)}{D(t,T)^2}\left(r_t D(t,T)dt - \sigma_r D(t,T)B(t,T)dW_{2,t}\right) \\
&- \frac{1}{D(t,T)^2}\sigma_r^2 D(t,t_i)D(t,T)B(t,t_i)B(t,T)dt + \frac{1}{2}\frac{2D(t,t_i)}{D(t,T)^3}\sigma_r^2 D(t,T)^2 B(t,T)^2 dt.
\end{aligned}
$$

Rearranging the terms we conclude

$$
\frac{d\left(\frac{D(t,t_i)}{D(t,T)}\right)}{\frac{D(t,t_i)}{D(t,T)}} = \sigma_r^2\left(B(t,T)^2 - B(t,T)B(t,t_i)\right)dt + \sigma_r\left(B(t,T) - B(t,t_i)\right)dW_{2,t}. \tag{67}
$$

Similarly, applying Ito's lemma to $\frac{S_t}{D(t,T)}$ we get

$$
\frac{S_t}{D(t,T)} = \frac{1}{D(t,T)}dS_t - \frac{S_t}{D(t,T)^2}dD(t,T) - \frac{1}{D(t,T)^2}d < S_t, D(t,T) > + \frac{1}{2}\frac{2S_t}{D(t,T)^3}d < D(t,T), D(t,T) > .
$$

Plugging in (64) and (66) we obtain

$$
\begin{aligned}
\frac{d\left(\frac{S_t}{D(t,T)}\right)}{\frac{S_t}{D(t,T)}} &= r_t dt + \sqrt{1-\rho^2}\sigma_S dW_{1,t} + \rho\sigma_S dW_{2,t} - r_t dt + \sigma_r B(t,T)dW_{2,t} \\
&+ \rho\sigma_S\sigma_r B(t,T)dt + \sigma_r^2 B(t,T)^2 dt,
\end{aligned}
$$

So rearranging gives

$$
\frac{d\left(\frac{S_t}{D(t,T)}\right)}{\frac{S_t}{D(t,T)}} = B(t,T)\left(\rho\sigma_S\sigma_r + \sigma_r^2 B(t,T)\right)dt + \sqrt{1-\rho^2}\sigma_S dW_{1,t} + \rho\sigma_S dW_{2,t} + \sigma_r B(t,T)dW_{2,t}. \tag{68}
$$

Now we change measure from the risk neutral measure to the T-forward measure. From Girsanov's theorem (see for example Karatzas and Shreve (2000)) we know that this can be done in such a way that we get $dW_{1,t} = dW_{1,t}^T$ and $dW_{2,t} = dW_{2,t}^T - \sigma_r B(t,T)dt$, with $W_{1,t}^T$ and $W_{2,t}^T$ independent Brownian motions under the new measure. Defining as before $F_t^T = \frac{S_t}{D(t,T)}$ and $D^T(t,t_i) = \frac{D(t,t_i)}{D(t,T)}$, (67) becomes

$$
\frac{dD^T(t,t_i)}{D^T(t,t_i)} = \sigma_r\left(B(t,T) - B(t,t_i)\right)dW_{2,t}^T, \tag{69}
$$

and for (68) we get

$$
\begin{aligned}
\frac{dF_t^T}{F_t^T} = B(t,T)\left(\rho\sigma_S\sigma_r + \sigma_r^2 B(t,T)\right)dt + \sqrt{1-\rho^2}\sigma_S dW_{1,t}^T + \rho\sigma_S dW_{2,t}^T - \rho\sigma_S\sigma_r B(t,T)dt \\
+\sigma_r B(t,T)dW_{2,t}^T - \sigma_r^2 B(t,T)^2 dt,
\end{aligned}
$$

so

$$\frac{dF_t^T}{F_t^T} = \sqrt{1-\rho^2}\sigma_S dW_{1,t}^T + \rho\sigma_S dW_{2,t}^T + \sigma_r B(t,T) dW_{2,t}^T. \tag{70}$$

For (70) we can write equivalently (in weak SDE solution terms)

$$dF_t^T = \sqrt{\sigma_S^2 + 2\rho\sigma_S\sigma_r B(t,T) + \sigma_r^2 B(t,T)^2} F_t^T dZ_t^T, \tag{71}$$

where Z_t^T is a Brownian motion under the T-forward measure. This can be seen as follows. First note that it follows from (70) that F_t^T is a martingale. Also,

$$dF_t^T = \begin{pmatrix} \sqrt{1-\rho^2}\sigma_S F_t^T, & \rho\sigma_S F_t^T + \sigma_r B(t,T) F_t^T \end{pmatrix} \begin{pmatrix} W_{1,t}^T \\ W_{2,t}^T \end{pmatrix}.$$

The covariance matrix of $(W_{1,t}^T, W_{2,t}^T)$ is given by

$$\begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}.$$

So the variance of F_t^T is

$$\begin{pmatrix} \sqrt{1-\rho^2}\sigma_S F_t^T, & \rho\sigma_S F_t^T + \sigma_r B(t,T) F_t^T \end{pmatrix} \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} \begin{pmatrix} \sqrt{1-\rho^2}\sigma_S F_t^T \\ \rho\sigma_S F_t^T + \sigma_r B(t,T) F_t^T \end{pmatrix} =$$

$$\left(\sigma_S^2 + 2\rho\sigma_S\sigma_r B(t,T) + \sigma_r^2 B(t,T)^2\right) \left(F_t^T\right)^2 t,$$

and we conclude that a weak solution is indeed given by (71).

Comparing (69) and (71) with (54) and (55) we conclude that the HWBS model is indeed of our general form with

$$\sigma_F(t) = \sqrt{\sigma_S^2 + 2\rho\sigma_S\sigma_r B(t,T) + \sigma_r^2 B(t,T)^2},$$

and

$$\sigma_i(t) = \sigma_r \left(B(t,T) - B(t,t_i)\right).$$

Now, to get the lower bound explicitly we need to compute integrals like $\int_t^{\max(t,t_i)} \rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds$, $\int_{\max(t,t_i)}^{\max(t,t_j)} \rho_{F,j}(s)\sigma_F(s)\sigma_j(s)ds$ and $\int_{\max(t,t_j)}^{T} \sigma_F^2(s)ds$ (these are used in the expression for α_t and $\mu_{i|Z}(t)$ in 4.2). Therefore it is necessary to find the expressions for $\rho_{ij}(s)\sigma_i(s)\sigma_j(s)$, $\rho_{F,j}(s)\sigma_F(s)\sigma_j(s)$ and $\sigma_F^2(s)$. For this we first note that the solutions of (54) and (55) are given by

$$F_t^T = F_0^T e^{-\frac{1}{2}\int_0^t \sigma_F^2(s)ds + \int_0^t \sigma_F(s)dW_s^T}$$

and

$$D^T(t,t_j) = D^T(0,t_j) e^{-\frac{1}{2}\int_0^t \sigma_j^2(s)ds + \int_0^t \sigma_j(s)dW_s^{jT}}.$$

From this we see that

$$\mathbb{E}^{\mathbb{Q}^T}\left[F_t^T D^T(t,t_j)\right] = \mathbb{E}^{\mathbb{Q}^T}\left[F_0^T D^T(0,t_j) e^{-\frac{1}{2}\int_0^t \sigma_F^2(s)ds - \frac{1}{2}\int_0^t \sigma_j^2(s)ds + \int_0^t \sigma_F(s)dW_s^T + \int_0^t \sigma_j(s)dW_s^{jT}}\right].$$

Using properties of the quadratic covariation we find that the covariance of $\int_0^t \sigma_F(s)dW_s^T$ and $\int_0^t \sigma_j(s)dW_s^{iT}$ is given by $\int_0^t \sigma_F(s)\sigma_j(s)\rho_{F,j}(s)ds$. Therefore, $\int_0^t \sigma_F(s)dW_s^T + \int_0^t \sigma_j(s)dW_s^{jT}$ is normally distributed with mean 0 and variance $\int_0^t \sigma_F^2(s)ds + \int_0^t \sigma_j^2(s)ds + 2\int_0^t \sigma_F(s)\sigma_j(s)\rho_{F,j}(s)ds$. Hence, using the standard properties of the expectation of lognormal random variables (16) we find

$$\mathbb{E}^{\mathbb{Q}^T}\left[F_t^T D^T(t,t_j)\right] = F_0^T D^T(0,t_j) e^{\int_0^t \sigma_F(s)\sigma_j(s)\rho_{F,j}(s)ds}.$$

The expectation of the above product is almost equal to the covariation between F_t^T and $D^T(t,t_j)$. The difference is that for the covariation we still have to subtract $\mathbb{E}^{\mathbb{Q}^T}\left[F_t^T\right]\mathbb{E}^{\mathbb{Q}^T}\left[D^T(t,t_j)\right]$, but since both processes are martingales, that equals $F_0^T D^T(0,t_j)$. So, the instantaneous covariance at time t is given by

$$\begin{aligned} e^{\int_0^{t+dt}\sigma_F(s)\sigma_j(s)\rho_{F,j}(s)ds} &- e^{\int_0^{t}\sigma_F(s)\sigma_j(s)\rho_{F,j}(s)ds} \\ \approx 1+\int_0^{t+dt}\sigma_F(s)\sigma_j(s)\rho_{F,j}(s)ds &- \left(1+\int_0^{t}\sigma_F(s)\sigma_j(s)\rho_{F,j}(s)ds\right) \\ &= \sigma_F(t)\sigma_j(t)\rho_{F,j}(t)dt. \end{aligned}$$

Also, from (69) and (70) we find that the instantaneous covariance between F_t^T and $D^T(t,t_j)$ is given by

$$\begin{aligned} &< \sqrt{1-\rho^2}\sigma_S dW_{1,t}^T+\rho\sigma_S dW_{2,t}^T+\sigma_r B(t,T)dW_{2,t}^T\ ,\ \sigma_r\left(B(t,T)-B(t,t_j)\right)dW_{2,t}^T > \\ &\qquad = (\rho\sigma_S+\sigma_r B(t,T))(\sigma_r\left(B(t,T)-B(t,t_j)\right))dt \\ &\qquad = \rho\sigma_S\sigma_r\left(B(t,T)-B(t,t_j)\right)+\sigma_r^2 B(t,T)^2-\sigma_r^2 B(t,T)B(t,t_j)dt. \end{aligned}$$

So combining both expressions for the instantaneous covariance we find the first expression we will need to compute our bounds on the value of the guarantee:

$$\sigma_F(t)\sigma_j(t)\rho_{F,j}(t)=\rho\sigma_S\sigma_r\left(B(t,T)-B(t,t_j)\right)+\sigma_r^2 B(t,T)^2-\sigma_r^2 B(t,T)B(t,t_j). \tag{72}$$

From (71) it is easy to see that

$$\sigma_F^2(t)=\sigma_S^2+2\rho\sigma_S\sigma_r B(t,T)+\sigma_r^2 B(t,T)^2. \tag{73}$$

Finally, the correlation between bonds with maturity t_i and t_j normalized by the bond with maturity T is one, since they are both driven by the same Brownian motion $W_{2,t}^T$. From (69) we therefore find

$$\rho_{ij}(t)\sigma_i(t)\sigma_j(t)=\sigma_r^2\left(B(t,T)-B(t,t_i)\right)\left(B(t,T)-B(t,t_j)\right). \tag{74}$$

The equations (72), (73) and (74) are the first ingredients for computing α_t and $\mu_{i|Z}(t)$. But as can be seen in the expressions for α_t and $\mu_{i|Z}(t)$ we finally need these terms integrated. For this we will assume $t<t_i<t_j$. We will first concentrate on $\int_t^{\max(t,t_i)}\rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds$. According to (74) this is now given by $\int_t^{t_i}\sigma_r^2\left(B(s,T)-B(s,t_i)\right)\left(B(s,T)-B(s,t_j)\right)ds$. Recall that $B(s,T)=\frac{1}{a}(1-e^{-a(T-s)})$. From this it follows that $B(s,T)-B(s,t_i)=\frac{1}{a}(e^{-a(t_i-s)}-e^{-a(T-s)})$. So

$$\begin{aligned} \left(B(s,T)-B(s,t_i)\right)\left(B(s,T)-B(s,t_j)\right) &= \frac{1}{a}(e^{-a(t_i-s)}-e^{-a(T-s)})\frac{1}{a}(e^{-a(t_j-s)}-e^{-a(T-s)})= \\ \frac{1}{a^2}\left(e^{-a(t_i+t_j-2s)}-e^{-a(t_i+T-2s)}-e^{-a(T+t_j-2s)}+e^{-a(2T-2s)}\right). \end{aligned}$$

Inserting this, we get the following

$$\begin{aligned} \int_t^{\max(t,t_i)}\rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds &= \int_t^{t_i}\frac{\sigma_r^2}{a^2}\left(e^{-a(t_i+t_j-2s)}-e^{-a(t_i+T-2s)}-e^{-a(T+t_j-2s)}+e^{-a(2T-2s)}\right)ds \\ &= \frac{\sigma_r^2}{a^2}\left[\frac{1}{2a}e^{-a(t_i+t_j-2s)}\right]_t^{t_i}-\frac{\sigma_r^2}{a^2}\left[\frac{1}{2a}e^{-a(t_i+T-2s)}\right]_t^{t_i} \\ &- \frac{\sigma_r^2}{a^2}\left[\frac{1}{2a}e^{-a(T+t_j-2s)}\right]_t^{t_i}+\frac{\sigma_r^2}{a^2}\left[\frac{1}{2a}e^{-a(2T-2s)}\right]_t^{t_i} \end{aligned}$$

$$\begin{aligned} &= \frac{\sigma_r^2}{2a^3}\left[e^{-a(t_j-t_i)}-e^{-a(t_i+t_j-2t)}\right]- \frac{\sigma_r^2}{2a^3}\left[e^{-a(T-t_i)}-e^{-a(t_i+T-2t)}\right] \\ &-\frac{\sigma_r^2}{2a^3}\left[e^{-a(T+t_j-2t_i)}-e^{-a(T+t_j-2t)}\right]+ \frac{\sigma_r^2}{2a^3}\left[e^{-a(2T-2t_i)}-e^{-a(2T-2t)}\right]. \end{aligned}$$

At first sight, this is a terrible expression, but using the definition of $B(t,T)$ we can rewrite this to get the following

$$\int_t^{\max(t,t_i)} \rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds = \frac{\sigma_r^2}{2a^2}\left[B(2t,t_i+t_j)\right.$$
$$\left.-B(t_i,t_j) - B(2t,t_i+T) + B(t_i,T) - B(2t,T+t_j) + B(2t_i,T+t_j) + B(2t,2T) - B(2t_i,2T)\right].$$

To compute $\int_{t_i}^{t_j} \rho_{F,j}(s)\sigma_F(s)\sigma_j(s)ds$ we proceed in a similar way. From (72) this integral is given by $\int_{t_i}^{t_j} \rho\sigma_S\sigma_r\left(B(s,T) - B(s,t_j)\right) + \sigma_r^2 B(s,T)\left(B(s,T) - B(s,t_j)\right)ds$. Using the expression for $(B(s,T) - B(s,t_j))$ that we already derived, we find

$$\begin{aligned} B(s,T)\left(B(s,T) - B(s,t_j)\right) &= \frac{1}{a}(1 - e^{-a(T-s)})\frac{1}{a}(e^{-a(t_j-s)} - e^{-a(T-s)}) \\ &= \frac{1}{a^2}(e^{-a(t_j-s)} - e^{-a(T-s)} - e^{-a(T+t_j-2s)} + e^{-a(2T-2s)}). \end{aligned}$$

So

$$\begin{aligned} \int_{t_i}^{t_j} \rho_{F,j}(s)\sigma_F(s)\sigma_j(s)ds &= -\rho\sigma_S\sigma_r \int_{t_i}^{t_j} \frac{1}{a}(e^{-a(T-s)} - e^{-a(t_j-s)})ds \\ &+ \sigma_r^2 \int_{t_i}^{t_j} \frac{1}{a^2}(e^{-a(t_j-s)} - e^{-a(T-s)} - e^{-a(T+t_j-2s)} + e^{-a(2T-2s)})ds. \end{aligned}$$

If we compute this expression explicitly, we get

$$\int_{t_i}^{t_j} \rho_{F,j}(s)\sigma_F(s)\sigma_j(s)ds = -\frac{\rho\sigma_S\sigma_r}{a}\left[\frac{1}{a}e^{-a(T-s)}\right]_{t_i}^{t_j} + \frac{\rho\sigma_S\sigma_r}{a}\left[\frac{1}{a}e^{-a(t_j-s)}\right]_{t_i}^{t_j}$$
$$+\frac{\sigma_r^2}{a^2}\left[\frac{1}{a}e^{-a(t_j-s)}\right]_{t_i}^{t_j} - \frac{\sigma_r^2}{a^2}\left[\frac{1}{a}e^{-a(T-s)}\right]_{t_i}^{t_j} - \frac{\sigma_r^2}{a^2}\left[\frac{1}{2a}e^{-a(T+t_j-2s)}\right]_{t_i}^{t_j} + \frac{\sigma_r^2}{a^2}\left[\frac{1}{2a}e^{-a(2T-2s)}\right]_{t_i}^{t_j}$$

$$= \frac{\rho\sigma_S\sigma_r}{a^2}\left[e^{-a(T-t_i)} - e^{-a(T-t_j)} - e^{-a(t_j-t_i)} + 1\right]$$
$$+\frac{\sigma_r^2}{a^3}\left[1 - e^{-a(t_j-t_i)} - e^{-a(T-t_j)} + e^{-a(T-t_i)} - \frac{1}{2}e^{-a(T-t_j)} + \frac{1}{2}e^{-a(T+t_j-2t_i)}\right.$$
$$\left.+\frac{1}{2}e^{-a(2T-2t_j)} - \frac{1}{2}e^{-a(2T-2t_i)}\right].$$

Using that $B(t_j,T)B(t_i,t_j) = \frac{1}{a^2}\left(1 + e^{-a(T-t_i)} - e^{-a(T-t_j)} - e^{-a(t_j-t_i)}\right)$, we can rewrite this to obtain

$$\int_{t_i}^{t_j} \rho_{F,j}(s)\sigma_F(s)\sigma_j(s)ds = \rho\sigma_S\sigma_r B(t_j,T)B(t_i,t_j) + \frac{\sigma_r^2}{a}B(t_j,T)B(t_i,t_j)$$
$$+\frac{\sigma_r^2}{2a^2}\left[B(t_j,T) - B(2t_i,T+t_j)\right] + \frac{\sigma_r^2}{2a^2}\left[B(2t_i,2T) - B(2t_j,2T)\right].$$

Now we still have to compute (use (73)) $\int_{t_j}^T \sigma_F^2(s)ds = \int_{t_j}^T \sigma_S^2 + 2\rho\sigma_S\sigma_r B(s,T) + \sigma_r^2 B(s,T)^2 ds$. We find

$$\int_{t_j}^T \sigma_F^2(s)ds = \sigma_S^2(T-t_j) + 2\rho\sigma_S\sigma_r\int_{t_j}^T \frac{1}{a}(1 - e^{-a(T-s)})ds$$
$$+\sigma_r^2\int_{t_j}^T \frac{1}{a^2}(1 - 2e^{-a(T-s)} + e^{-a(2T-2s)})ds$$
$$= \sigma_S^2(T-t_j) + \frac{2\rho\sigma_S\sigma_r}{a}(T-t_j) - \frac{2\rho\sigma_S\sigma_r}{a^2}\left[1 - e^{-a(T-t_j)}\right]$$

$$
\begin{aligned}
&+\frac{\sigma_r^2}{a^2}(T-t_j) - \frac{2\sigma_r^2}{a^3}\left[1-e^{-a(T-t_j)}\right] + \frac{\sigma_r^2}{2a^3}\left[1-e^{-a(2T-2t_j)}\right] \\
&= \left(\sigma_S^2 + \frac{2\rho\sigma_S\sigma_r}{a} + \frac{\sigma_r^2}{a^2}\right)(T-t_j) - B(t_j,T)\left(\frac{2\rho\sigma_S\sigma_r}{a} + \frac{2\sigma_r^2}{a^2}\right) \\
&\qquad + \frac{\sigma_r^2}{2a^2}B(2t_j, 2T).
\end{aligned}
$$

We have now derived all quantities that we need to compute α_t and $\mu_{i|Z}(t)$. So all we have to do to find the lower bound, is to plug these quantities in formulas (62) and (63). The only thing that remains, is that we need an explicit expression for $D(t,T)$. As noted, $D(t,T) = e^{A(t,T)-B(t,T)r_t}$, with $A(t,T)$ and $B(t,T)$ as given before. So what we still have to do, is find an explicit expression for $A(t,T) = \frac{\sigma_r^2}{2}\int_t^T B(s,T)^2 ds - \int_t^T \theta_s B(s,T)ds$. For this we shall assume that the initial term structure is flat, i.e. $f(0,t) = f$ for some constant rate f. Then $\theta(t) = af + \frac{\sigma_r^2}{2a}(1-e^{-2at})$. From the computation of $\int_{t_j}^T \sigma_F^2(s)ds$ we easily derive that

$$
\frac{\sigma_r^2}{2}\int_t^T B(s,T)^2 ds = \frac{\sigma_r^2}{2a^2}(T-t) - \frac{\sigma_r^2}{a^3}\left[1-e^{-a(T-t)}\right] + \frac{\sigma_r^2}{4a^3}\left[1-e^{-a(2T-2t)}\right],
$$

so it remains to compute $\int_t^T \theta_s B(s,T)ds$. Plugging in the definitions we find

$$
\begin{aligned}
\int_t^T \theta_s B(s,T)ds &= \int_t^T \left(af + \frac{\sigma_r^2}{2a}(1-e^{-2as})\right)\left(\frac{1}{a}(1-e^{-a(T-s)})\right)ds \\
= \int_t^T f(1-e^{-a(T-s)})ds &+ \int_t^T \frac{\sigma_r^2}{2a^2}(1-e^{-2as})(1-e^{-a(T-s)})ds \\
= f\left((T-t) - \frac{1}{a} + \frac{1}{a}e^{-a(T-t)}\right) &+ \frac{\sigma_r^2}{2a^2}\int_t^T 1 - e^{-a(T-s)} - e^{-2as} + e^{-a(T+s)}ds \\
= f\left((T-t) - B(t,T)\right) &+ \frac{\sigma_r^2}{2a^2}(T-t) - \frac{\sigma_r^2}{2a^3}(1-e^{-a(T-t)}) + \frac{\sigma_r^2}{4a^3}(e^{-2aT} - e^{-2at}) \\
&- \frac{\sigma_r^2}{2a^3}(e^{-2aT} - e^{-a(T+t)}).
\end{aligned}
$$

Putting things together, we get

$$
\begin{aligned}
A(t,T) &= \frac{\sigma_r^2}{2a^2}(T-t) - \frac{\sigma_r^2}{a^3}\left[1-e^{-a(T-t)}\right] + \frac{\sigma_r^2}{4a^3}\left[1-e^{-a(2T-2t)}\right] \\
&- f\left((T-t)-B(t,T)\right) - \frac{\sigma_r^2}{2a^2}(T-t) + \frac{\sigma_r^2}{2a^3}(1-e^{-a(T-t)}) - \frac{\sigma_r^2}{4a^3}(e^{-2aT} - e^{-2at}) \\
&+ \frac{\sigma_r^2}{2a^3}(e^{-2aT} - e^{-a(T+t)}) \\
&= -f\left((T-t)-B(t,T)\right) - \frac{\sigma_r^2}{4a^3}\Big(2 - 2e^{-a(T-t)} - 1 + e^{-a(2T-2t)} + e^{-2aT} \\
&- e^{-2at} - 2e^{-2aT} + 2e^{-a(T+t)}\Big) \\
&= -f\left((T-t)-B(t,T)\right) - \frac{\sigma_r^2}{4a^3}\left(e^{-aT} - e^{-at}\right)^2\left(e^{2at} - 1\right).
\end{aligned}
$$

This concludes the computation of the terms used in the bounds of section 4.2 for the HWBS model. The expressions are complicated, but they are all explicit. In summary, the results in the HWBS model with flat initial term structure are as follows:

Theorem 4.3 *An upper bound on the value of the guarantee is given by*

$$
V_t(T,G) \le D(t,T)\left[-\sum_{i=0}^{T-1} P_i \bar{\mu}_i(t)\Phi\left(-\sqrt{\int_t^T \bar{\sigma_i}^2(s)ds} + \Phi^{-1}(F_{U_T^c}(G))\right) + GF_{U_T^c}(G)\right], \tag{75}
$$

where $F_{U_T^c}(G)$ follows from solving

$$\sum_{i=0}^{T-1} P_i \bar{\mu_i}(t) e^{-\frac{1}{2}\int_t^T \bar{\sigma_i}^2(s)ds + \sqrt{\int_t^T \bar{\sigma_i}^2(s)ds}\ \Phi^{-1}(F_{U_T^c}(G))} = G. \tag{76}$$

A lower bound on the value of the guarantee is given by

$$V_t(T,G) \geq D(t,T)\left[-\sum_{i=0}^{T-1} P_i\bar{\mu_i}(t)\Phi\left(-\mu_{i|Z}(t) + \Phi^{-1}(F_{U_T^l}(G))\right) + GF_{U_T^l}(G)\right], \tag{77}$$

where $F_{U_T^l}(G)$ follows from solving

$$\sum_{i=0}^{T-1} P_i\bar{\mu_i}(t) e^{\mu_{i|Z}(t)\Phi^{-1}(F_{U_t^l}(G)) - \frac{1}{2}\mu_{i|Z}^2(t)} = G. \tag{78}$$

Furthermore, we have the following expressions for the terms in the bounds:

$$D(t,T) = e^{A(t,T)-B(t,T)r_t},$$

$$\begin{aligned}
\mu_{i|Z}(t) \quad &= \quad \frac{1}{\alpha_t}\sum_{j=0}^{i} P_j\bar{\mu_j}(t)e^{-\frac{1}{2}\int_t^T \bar{\sigma_j}^2(s)ds}\left[\int_t^{\max(t,t_j)} \rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds + \int_{\max(t,t_j)}^{\max(t,t_i)} \rho_{F,i}(s)\sigma_F(s)\sigma_i(s)ds\right. \\
&+ \quad \left.\int_{\max(t,t_i)}^T \sigma_F^2(s)ds\right] + \\
& \quad\ \frac{1}{\alpha_t}\sum_{j=i+1}^{T-1} P_j\bar{\mu_j}(t)e^{-\frac{1}{2}\int_t^T \bar{\sigma_j}^2(s)ds}\left[\int_t^{\max(t,t_i)} \rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds + \int_{\max(t,t_i)}^{\max(t,t_j)} \rho_{F,j}(s)\sigma_F(s)\sigma_j(s)ds\right. \\
&+ \quad \left.\int_{\max(t,t_j)}^T \sigma_F^2(s)ds\right],
\end{aligned}$$

$$\alpha_t \quad = \quad \sum_{i=0}^{T-1}\sum_{j=0}^{T-1} P_i\bar{\mu_i}(t)P_j\bar{\mu_j}(t)e^{-\frac{1}{2}\int_t^T \bar{\sigma_i}^2(s)ds - \frac{1}{2}\int_t^T \bar{\sigma_j}^2(s)ds} Cov\left(\int_t^T \bar{\sigma_i}(s)d\bar{W}_s^i, \int_t^T \bar{\sigma_j}(s)d\bar{W}_s^j\right),$$

where for $i \leq j$ (i.e. $t_i \leq t_j$) we have

$$Cov\left(\int_t^T \bar{\sigma_i}(s)d\bar{W}_s^i, \int_t^T \bar{\sigma_j}(s)d\bar{W}_s^j\right) =$$

$$\int_t^{\max(t,t_i)} \rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds + \int_{\max(t,t_i)}^{\max(t,t_j)} \rho_{F,j}(s)\sigma_F(s)\sigma_j(s)ds + \int_{\max(t,t_j)}^T \sigma_F^2(s)ds.$$

Finally, $A(t,T)$, $B(t,T)$ and the integrals $\int_t^{\max(t,t_i)} \rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds$, $\int_{\max(t,t_i)}^{\max(t,t_j)} \rho_{F,j}(s)\sigma_F(s)\sigma_j(s)ds$ and $\int_{\max(t,t_j)}^T \sigma_F^2(s)ds$ are as computed in this section.

4.4 Numerical Results

This section provides numerical results for the lower bound in the HWBS model. The reason we only study the lower bound, is that the upper bound is not accurate as we have already seen in chapter 3. Again, the lower bound is compared to the value obtained by 10,000 Monte Carlo simulations.
We have computed the lower bound in Excel, but because it was hard to implement the formulas in excel, we only studied one contract which has a long maturity. The longer the maturity, the worse the approximation (Schrager and Pelsser (2004)) and therefore we believe that the results obtained are representative. Furthermore, the lower bound and Monte Carlo simulation are additive. So if the lower bound for our representative contract is accurate, this will also hold for a model point (package of contracts) or the whole portfolio (all model points together).
The maturity of the contract we used is 29 years and the guaranteed value is 24,764.61. We have tested the lower bound for various values of the parameters that are typical in practice. For the mean reversion a we used 0.01 and 0.03, for σ_r we used 0.005, 0.01 and 0.015, for σ_S we used 0.06 and 0.20, for ρ we used 0 and -0.02 and in all situations we assumed the initial term structure to be flat with rate 3.922 %. The results are as follows.

For $\sigma_S = 0.06$ and $\rho = 0$

σ_r	a	LB	MC (s.e.)
0.005	0.01	618	622 (15)
	0.03	486	482 (11)
0.01	0.01	1,457	1,415 (46)
	0.03	1,122	1,156 (32)
0.015	0.01	2,341	2,478 (105)
	0.03	1,839	1,928 (68)

Table 3: Lower bound (LB) compared to Monte Carlo estimate (MC) and their standard error (s.e.) for $\sigma_S = 0.06$ and $\rho = 0$

For $\sigma_S = 0.20$ and $\rho = 0$

σ_r	a	LB	MC (s.e.)
0.005	0.01	1,926	1,905 (20)
	0.03	1,864	1,838 (17)
0.01	0.01	2,396	2,438 (50)
	0.03	2,193	2,204 (36)
0.015	0.01	3,003	3,137 (108)
	0.03	2,646	2,725 (72)

Table 4: Lower bound (LB) compared to Monte Carlo estimate (MC) and their standard error (s.e.) for $\sigma_S = 0.20$ and $\rho = 0$

For $\sigma_S = 0.06$ and $\rho = -0.02$

σ_r	a	LB	MC (s.e.)
0.005	0.01	610	614 (15)
	0.03	479	475 (11)
0.01	0.01	1,478	1,505 (46)
	0.03	1,113	1,147 (31)
0.015	0.01	2,333	2,469 (105)
	0.03	1,830	1,918 (68)

Table 5: Lower bound (LB) compared to Monte Carlo estimate (MC) and their standard error (s.e.) for $\sigma_S = 0.06$ and $\rho = -0.02$

From these results we immediately conclude that the lower bound is again a very accurate estimate for the value of the guarantee. The lower bound is always within two standard deviations of the Monte Carlo, which means that it lies in the 95 % confidence interval.
We can also see that the higher the values of σ_S and σ_r, the higher the option value. This is explained by the fact that the option value is higher if the total volatility of the investment returns is higher and this is determined by σ_S and σ_r. Also, the standard deviation of the Monte Carlo is higher when σ_S and σ_r are higher, which has the same explanation.
Another observation is that the higher the mean reversion a, the lower the option value. This is because a higher mean reversion dampens the volatility of the short rate so that total volatility goes down.

5 Profit Sharing

In chapters 3 and 4 we studied bounds on the value of the guarantee in unit-linked insurance. For this we heavily relied on the theory developed in chapter 2. From now on we will study the second product mentioned in the introduction: profit sharing.

As opposed to unit-linked insurance, profit sharing is a form of non-linked insurance. There are many variations but the product as we will study looks as follows. A customer pays a premium (for example, this could be yearly) until maturity. The insurance company will then give a yearly guaranteed return on the total value of premiums and accumulated returns built up until that moment. This guaranteed return is called the technical interest rate $TR(t)$. This rate is typically low and therefore the product is combined with profit sharing. This refers to a situation where some reference return $R(t)$ is paid out to the policy holder if this reference return exceeds the technical interest rate. So the interest rate return above the technical interest rate is given by

$$(R(t) - TR(t))_+$$

and therefore the embedded option in this product is a call option on the rate $R(t)$.

There are two basic ways in which the profit sharing can be paid out. One possibility is that the customer gets the profit sharing bonus the moment it is determined. This is called direct payment. The other possibility is that the profit sharing bonus is not paid out until maturity, but added to the total value already accumulated (compounding profit sharing). The advantage of this is that a customer can get profit sharing over profit sharing and this method is mostly used in practice. Also, there are different types of reference rates that are used and in most cases these rates are very complex. For example (Plat and Pelsser (2008)), in the Netherlands the most common form of profit sharing is based on a moving average of the so-called u-rate. The u-rate is the 3-months average of u-rate-parts, where the subsequent u-rate-parts are weighted averages of an effective return on a basket of government bonds. This leads to a complicated expression. Therefore, it is common practice to approximate the u-rate by a moving average of swap rates.

A moving average of swap rates can also be used to approximate other profit sharing rates. For this reason we will assume in this chapter that the profit sharing is based on an average of swap rates. The rest of the chapter will then look as follows. First we introduce the swap rate and derive a model for the dynamics of this rate. Using these dynamics, we will give approximations for the value of the option in profit sharing based on an average of swap rates in the case of direct payment and compounding profit sharing. Finally, we will test the quality of the approximation in the case of compounding profit sharing, since this is the most important case in practice. The concrete model we use for that will be the 1-factor Hull-White model.

5.1 Swap rate dynamics

To compute the value of the call option in insurance with profit sharing, we need to have a model for the interest rate on which the profit sharing is based. This will be an average of swap rates and therefore we will start with the derivation of an approximation of the swap rate dynamics which will turn out to work well. The swap rate is dependent on the bond prices and therefore we will first need a model for the short rates which determines the bond prices. We will assume an m-factor gaussian short rate model, which means that the short rate is described as follows (Schrager and Pelsser 2006):

$$r(t) = \mathbf{1}'Y(t) + \alpha(t), \tag{79}$$

$$dY(t) = -AY(t)dt + \Sigma dW^{\mathbb{Q}}(t), \tag{80}$$

where $W^{\mathbb{Q}}(t)$ is a m-dimensional brownian motion under the risk-neutral measure $\mathbb{Q}$, $\mathbf{1}$ is a column vector of ones, $'$ denotes transposition and A and Σ are $m \times m$ matrices with A a diagonal matrix. The function $\alpha(t)$ is chosen in such a way that the fit with the initial term structure is perfect. The covariance matrix of the Y-variables is $\hat{\Sigma} = \Sigma\Sigma'$. Note that for $m = 1$, $\Sigma = \sigma$, $A = a$, $Y(t) = r(t)$ and $\frac{\partial \alpha(t)}{\partial t} = \theta(t)$ we have the one-factor Hull-White model as in (65).

It is well known that this model is an affine term structure model. That is, the bond price at time t of a bond paying one unit of currency at maturity T, is given by

$$D(t,T) = e^{C(t,T) - \sum_{i=1}^{m} B^i(t,T)Y^i(t)}, \tag{81}$$

where $B^i(t,T) = \frac{1}{A_{ii}}\left(1 - e^{-A_{ii}(T-t)}\right)$. The expression for $C(t,T)$ will not be important for us, and therefore we omit it here.

Now, the swap rate is defined using the zero coupon bond prices as we will now explain. Let us first give the definition of the forward LIBOR rate. The forward LIBOR rate $L_{TS}(t)$ is the interest rate one can contract at time t to put money in a money-market account for the time period $[T, S]$. Precisely:

$$L_{TS}(t) = \frac{1}{\Delta^L_{TS}} \frac{D(t,T) - D(t,S)}{D(t,S)},$$

where Δ^L_{TS} is the LIBOR market convention for the calculation of the daycount fraction for the period $[T, S]$. In the market, the tenor of the LIBOR rate, $S - T$, is usually fixed at 3 or 6 months. In financial markets there can be different definitions for how many days there are in a year and this makes the daycount fraction important. As an example, we can have that there are 360 days in a year and that $S - T$ is 180 days. Then the daycount fraction will be $\frac{1}{2}$.

An interest rate swap is a contract in which two parties agree to exchange a set of fixed cash flows, consisting of a fixed rate K on the swap principal A, for a set of floating rate payments, consisting of the LIBOR rate on the principal A. In a payer swap you pay the fixed side and receive floating, in a receiver swap you receive the fixed side and pay the floating side. Given a set of dates $T_i, i = n+1, \ldots, N$, at which swap payments are to be made, the value at time t of a (payer) swap contract starting at T_n (paying out for the first time at T_{n+1}) and lasting until T_N with a principal of 1 and fixed payments at rate K is given by

$$V^{pay}_{n,N}(t) = V^{flo}_{n,N}(t) - V^{fix}_{n,N}(t) = \{D(t,T_n) - D(t,T_N)\} - K \sum_{i=n+1}^{N} \Delta^Y_{i-1} D(t,T_i), \tag{82}$$

where Δ^Y_{i-1} is the market convention for the calculation of the daycount fraction for the swap payment at T_i. This formula follows, since the value of the fixed leg at time T_n is given by discounting the payments of size $K\Delta^Y_{i-1}$ at times T_i to time T_n. This gives the value at time T_n of $K \sum_{i=n+1}^{N} \Delta^Y_{i-1} D(T_n, T_i)$. Discounting this to time t to get the value at time t, we get $V^{fix}_{n,N}(t) = K \sum_{i=n+1}^{N} \Delta^Y_{i-1} D(t, T_i)$. For the floating side we have that the value at time T_n is given by $1 - D(T_n, T_N)$. This is a bit harder to see, but the explanation can be found in for example (Neftci: Principles of financial engineering, 2004). Discounting this to time t

we find $V^{flo}_{n,N}(t) = D(t,T_n) - D(t,T_N)$.

Again, given a set of payment dates T_i, a forward par swap rate $y_{n,N}(t)$ is defined by the fixed rate for which the value of the (forward starting) swap equals zero. Solving (82) gives

$$y_{n,N}(t) = \frac{D(t,T_n) - D(t,T_N)}{\sum_{i=n+1}^{N} \Delta^Y_{i-1} D(t,T_i)} = \frac{D(t,T_n) - D(t,T_N)}{P_{n+1,N}(t)}. \tag{83}$$

The rate $y_{n,N}(t)$ is the arbitrage free rate at which at time t a person would like to enter into a swap contract starting at time T_n(paying out for the first time at time T_{n+1}) and lasting until T_N.

Now, under the basic modelling assumptions that there exists a money market account and a risk neutral measure $\mathbb{Q}$ such that for all maturities T the discounted bond prices are martingales under $\mathbb{Q}$, we can also use $P_{n+1,N}(t)$ as a numeraire since it is strictly positive. Using $P_{n+1,N}(t)$ as numeraire, we have that all asset prices normalized by $P_{n+1,N}(t)$ must be martingales under the measure $\mathbb{Q}^{n+1,N}$, associated with this numeraire. This follows from the change of numeraire theorem. But then we find from (83) that $y_{n,N}$ is a martingale under this so called swap measure.

For our purpose of pricing the embedded call option that is dependent on the swap rate, we will need the dynamics of the swap rate under the swap measure. As we will see, the dynamics of the swap rate under this measure in our model will have a stochastic volatility. This makes it hard to work with and therefore we will approximate the stochastic volatility by a deterministic volatility. It will turn out that this approximation works well, since the volatility of the stochastic volatility is low and therefore it is already almost deterministic.

We will now proceed with the derivation of the swap rate dynamics under the swap measure. For this, we will use Ito's lemma under the risk neutral measure and then change measure to the swap measure using Girsanov's theorem. Since we already know that the swap rate is a martingale under the swap measure, we can forget about the dt terms appearing when we apply Ito's lemma under the risk neutral measure, since when we then change measure all these dt terms will necessarily cancel. This also implies that we can find the swap rate dynamics under the swap measure by finding all the $dW^{\mathbb{Q}}(t)$ terms that appear when we apply Ito's lemma and in the end replace the $dW^{\mathbb{Q}}(t)$ by $dW^{n+1,N}(t)$, which is a m-dimensional Brownian motion under the swap measure. This is true because from Girsanov's theorem it follows that $dW^{\mathbb{Q}}(t) = dW^{n+1,N}(t) + \gamma(t)dt$ for some function $\gamma(t)$. The $\gamma(t)$ will cancel all dt terms and in the $dW^{\mathbb{Q}}(t)$ terms, $dW^{\mathbb{Q}}(t)$ is replaced by $dW^{n+1,N}(t)$. Additionally, we will use that in our model we are only dealing with stochastics driven by Brownian motions. This means that quadratic covariations will be of the order dt and hence can be forgotten. With these remarks in mind, we start by deriving the dynamics under the risk neutral measure of $D(t,T)$. From (81) we find by Ito's lemma

$$dD(t,T) = \ldots dt - D(t,T)\sum_{i=1}^{m} B^i(t,T)dY^i(t).$$

Plugging in (80) we get

$$dD(t,T) = \ldots dt - D(t,T)\sum_{i=1}^{m} B^i(t,T)\sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t),$$

where Σ^{ij} denotes the (i,j)-th element of Σ and $dW^{\mathbb{Q},j}(t)$ is the j-th element of $dW^{\mathbb{Q}}(t)$. Using the definition of $P_{n+1,N}(t)$ and the previous equation we get by Ito's lemma

$$\begin{aligned} d\left(\frac{D(t,T_n)}{P_{n+1,N}(t)}\right) &= \ldots dt + \frac{1}{P_{n+1,N}(t)} dD(t,T_n) - \frac{D(t,T_n)}{P^2_{n+1,N}(t)} dP_{n+1,N}(t) \\ &= \ldots dt + \frac{1}{P_{n+1,N}(t)}\left(-D(t,T_n)\sum_{i=1}^{m} B^i(t,T_n)\sum_{j=1}^{m}\Sigma^{ij} dW^{\mathbb{Q},j}(t)\right) \end{aligned}$$

$$
\begin{aligned}
&- \frac{D(t,T_n)}{P^2_{n+1,N}(t)} \left(- \sum_{k=n+1}^{N} \Delta^Y_{k-1} D(t,T_k) \sum_{i=1}^{m} B^i(t,T_k) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t) \right) \\
&= \ldots dt + \frac{-D(t,T_n)}{P_{n+1,N}(t)} \sum_{i=1}^{m} B^i(t,T_n) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t) \\
&+ \frac{D(t,T_n)}{P^2_{n+1,N}(t)} \sum_{k=n+1}^{N} \Delta^Y_{k-1} D(t,T_k) \sum_{i=1}^{m} B^i(t,T_k) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t).
\end{aligned}
$$

So

$$
\begin{aligned}
dy_{n,N}(t) &= d\left(\frac{D(t,T_n)}{P_{n+1,N}(t)} - \frac{D(t,T_N)}{P_{n+1,N}(t)} \right) = d\left(\frac{D(t,T_n)}{P_{n+1,N}(t)} \right) - d\left(\frac{D(t,T_N)}{P_{n+1,N}(t)} \right) \\
&= \ldots dt + \frac{1}{P_{n+1,N}(t)} \left(-D(t,T_n) \sum_{i=1}^{m} B^i(t,T_n) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t) \right. \\
&+ \left. D(t,T_N) \sum_{i=1}^{m} B^i(t,T_n) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t) \right) \\
&+ \frac{1}{P^2_{n+1,N}(t)} \left([D(t,T_n) - D(t,T_N)] \sum_{k=n+1}^{N} \Delta^Y_{k-1} D(t,T_k) \sum_{i=1}^{m} B^i(t,T_k) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t) \right).
\end{aligned}
$$

We now define $D^p(t,T_n) = \frac{D(t,T_n)}{P_{n+1,N}(t)}$ (the bond price normalized by the numeraire). Then we get from the previous equations

$$
\begin{aligned}
dy_{n,N}(t) &= \ldots dt + \sum_{i=1}^{m} -B^i(t,T_n) D^p(t,T_n) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t) \\
&+ \sum_{i=1}^{m} B^i(t,T_N) D^p(t,T_N) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t) \\
&+ \frac{D(t,T_n) - D(t,T_N)}{P_{n+1,N}(t)} \sum_{i=1}^{m} \sum_{k=n+1}^{N} \Delta^Y_{k-1} D^p(t,T_k) B^i(t,T_k) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t) \\
&= \ldots dt + \left(\sum_{i=1}^{m} \left(-B^i(t,T_n) D^p(t,T_n) + B^i(t,T_N) D^p(t,T_N) \right) \right. \\
&+ \left. y_{n,N}(t) \sum_{k=n+1}^{N} \Delta^Y_{k-1} D^p(t,T_k) B^i(t,T_k) \right) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t).
\end{aligned}
$$

If we now change to the swap measure, the dt term will cancel out and $dW^{\mathbb{Q}}(t)$ is replaced by $dW^{n+1,N}(t)$. Furthermore, if we switch to vector notation, we find the following dynamics of $y_{n,N}(t)$ under the swap measure

$$
dy_{n,N}(t) = \frac{\partial y_{n,N}(t)}{\partial Y(t)} \Sigma dW^{n+1,N}(t), \tag{84}
$$

with the i-th element of the m-dimensional vector $\frac{\partial y_{n,N}(t)}{\partial Y(t)}$ (denoted $\frac{\partial y_{n,N}(t)}{\partial Y^i(t)}$) given by

$$
-B^i(t,T_n) D^p(t,T_n) + B^i(t,T_N) D^p(t,T_N) + y_{n,N}(t) \sum_{k=n+1}^{N} \Delta^Y_{k-1} D^p(t,T_k) B^i(t,T_k).
$$

Note that $\frac{\partial y_{n,N}(t)}{\partial Y(t)}$ suggests some sort of derivative, but it is just notation for the vector we specified above. As already mentioned, the volatility of the swap rate under the swap measure $\frac{\partial y_{n,N}(t)}{\partial Y(t)} \Sigma$ is stochastic, since

$\frac{\partial y_{n,N}(t)}{\partial Y(t)}$ depends on the stochastic $D^p(t,T_i)$. However, for all i we have that $D^p(t,T_i)$ is a martingale under the swap measure, since it is the asset $D(t,T_i)$ normalized by the numeraire $P_{n+1,N}(t)$. Also, these martingales have a low variance and Schrager and Pelsser (2006) show that we can approximate these low-variance martingales by their time zero values. They show that prices of options computed under this approximated swap rate dynamics are very close to the prices of options computed under the exact model. Therefore, we will also work with this approximation, in which the volatility of the swap rate becomes a deterministic function of time. More precisely, the approximated dynamics is given by

$$dy_{n,N}(t) = \widetilde{\frac{\partial y_{n,N}(t)}{\partial Y(t)}} \Sigma dW^{n+1,N}(t), \tag{85}$$

with the i-th element of the m-dimensional vector $\widetilde{\frac{\partial y_{n,N}(t)}{\partial Y(t)}}$ given by

$$-B^i(t,T_n)D^p(0,T_n) + B^i(t,T_N)D^p(0,T_N) + y_{n,N}(0)\sum_{k=n+1}^{N}\Delta_{k-1}^{Y}D^p(0,T_k)B^i(t,T_k).$$

Recall that $B^i(t,T) = \frac{1}{A_{ii}}\left(1 - e^{-A_{ii}(T-t)}\right)$. If we insert this in the previous expression, we see that this expression can be split in a constant and a time dependent part. Also note that terms coming from $\frac{1}{A_{ii}}$ cancel. This can be seen by simply writing everything down and inserting the definitions of $y_{n,N}(0)$ and $P_{n+1,N}(0)$. Finally we get

$$\begin{aligned}\widetilde{\frac{\partial y_{n,N}(t)}{\partial Y^i(t)}} &= \frac{1}{A_{ii}}e^{A_{ii}t}\Big[e^{-A_{ii}T_n}D^p(0,T_n) - e^{-A_{ii}T_N}D^p(0,T_N) \\ &- y_{n,N}(0)\sum_{k=n+1}^{N}\Delta_{k-1}^{Y}e^{-A_{ii}T_k}D^p(0,T_k)\Big] = e^{A_{ii}t}C_{n,N}^i.\end{aligned}$$

All this results in the following expression for the swap rate at time T_n (by integration of (84))

$$\begin{aligned}y_{n,N}(T_n) &= y_{n,N}(0) + \int_0^{T_n}\frac{\partial y_{n,N}(t)}{\partial Y(t)}\Sigma dW^{n+1,N}(t) \\ &\approx y_{n,N}(0) + \int_0^{T_n}\widetilde{\frac{\partial y_{n,N}(t)}{\partial Y(t)}}\Sigma dW^{n+1,N}(t) \\ &= y_{n,N}(0) + \int_0^{T_n}e^{At'}diag(C_{n,N})\Sigma dW^{n+1,N}(t),\end{aligned}$$

where e^{At} is the vector with elements $e^{A_{ii}t}$ and $diag(C_{n,N})$ is a diagonal matrix with element (i,i) equal to $C_{n,N}^i$ and all other elements equal to zero. Furthermore, the integrated variance of $y_{n,N}$ over the interval $[0,T_n]$ is

$$\begin{aligned}\sigma_{n,N}^2 &\approx \int_0^{T_n}e^{As'}diag(C_{n,N})\Sigma\Sigma' diag(C_{n,N})e^{As}ds \\ &= \sum_{i=1}^{m}\sum_{j=1}^{m}\hat{\Sigma_{ij}}C_{n,N}^iC_{n,N}^j\frac{e^{(A_{ii}+A_{jj})T_n}-1}{A_{ii}+A_{jj}}.\end{aligned}$$

So $y_{n,N}(T_n)$ is approximately normally distributed with mean $y_{n,N}(0)$ and variance $\sigma_{n,N}^2$.

5.2 Direct payment

With the approximate swap rate dynamics of the previous section at hand, we are now able to value the call option in profit sharing insurance products. In general, the profit sharing payoff $PS(t)$ in year t is

$$PS(t) = L(t)Max\{c(R(t) - K(t)), 0\},$$

where $L(t)$ is the profit sharing basis, c is the percentage that is distributed to the policyholder and $K(t)$ is the strike of the option. The strike equals the sum of the technical interest $TR(t)$ and a margin. In most cases, either margin or c is used for the benefits of the insurer. In this case, $R(t)$ is the profit sharing rate and will be taken as a weighted average swap rates. Again, it is clear from the above expression that profit sharing is a call option on the rate $R(t)$ and therefore we will use option pricing techniques to find its value. In the case of direct payment, the client will directly receive the profit sharing the moment it is determined. This is done on certain dates between initiation and maturity of the contract. So if the contract runs from time 0 to time T_n and the profit sharing is determined at times $T_i, i = 1, \ldots, n$, then the client receives an amount $L(T_i)Max\{c(R(T_i) - K(T_i)), 0\}$ at time T_i. This means that in case of direct payment, the embedded option is in fact a strip of options that mature at time T_i and lead to a direct payment depending on $R(T_i)$ on these dates.

We will assume that $R(T_i)$ is a weighted average of τ-year maturity swap rates with weights w_k and the averaging period is from time T_{i-s} to time T_i:

$$R(T_i) = \sum_{k=T_{i-s}}^{T_i} w_k y_{k,k+\tau}(k), \tag{86}$$

where $\sum w_k = 1$.

As we have seen in the previous section, the $y_{k,k+\tau}(k)$ are approximately normally distributed and therefore the $R(T_i)$ are also approximately normally distributed. So to value the option the expectation and the variance of $R(T_i)$ have to be approximated under the T_i-forward measure and feed into a gaussian option valuation formula for each time T_i. For determining the variance of $R(T_i)$ the covariances of the $y_{k,k+\tau}(k)$ with the $y_{l,l+\tau}(l)$ have to be specified.

5.2.1 Mean of $R(T_i)$

The above means that each individual option has to be priced in the T_i-forward measure. To come to the expectation of $R(T_i)$ under the right measure the following steps have to be done:

(a) For each (forward) swap rate $y_{k,k+\tau}$ a change of measure has to be done from the swap measure $\mathbb{Q}^{k+1,k+\tau}$ to the T_k-forward measure $\mathbb{Q}^{T_k}$

(b) If the payoff of the option on the average of the swap rates is at time T_i, for each of the individual swap rates in the average ($y_{k,k+\tau}$ for $T_{i-s} \leq k \leq T_i$), a change of measure has to be done from the T_k-forward measure to the T_i-forward measure.

The corrections mentioned above can be interpreted as a convexity correction (a), which is a change of measure where the numeraire changes to a different asset and a timing correction (b), which is a change of measure where the numeraire is the same kind of asset, but with a different maturity. For a mathematical foundation of convexity correction see (Pelsser (2003)).

Now, let us first change measure from the swap measure $\mathbb{Q}^{k+1,k+\tau}$ to the T_k-forward measure $\mathbb{Q}^{T_k}$. From the change of numeraire theorem we have that the density $\rho(t)$ is given by

$$\frac{d\mathbb{Q}^{T_k}}{d\mathbb{Q}^{k+1,k+\tau}} = \rho(t) = \frac{D(t, T_k)/D(0, T_k)}{P_{k+1,k+\tau}(t)/P_{k+1,k+\tau}(0)}.$$

As in the previous section, we have by Ito's lemma:

$$\begin{aligned} d\left(\frac{D(t,T_k)}{P_{k+1,k+\tau}(t)}\right) &= \dots dt + \frac{-D(t,T_k)}{P_{k+1,k+\tau}(t)} \sum_{i=1}^{m} B^i(t,T_k) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t) \\ &+ \frac{D(t,T_k)}{P^2_{k+1,k+\tau}(t)} \sum_{l=k+1}^{k+\tau} \Delta^Y_{l-1} D(t,T_l) \sum_{i=1}^{m} B^i(t,T_l) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q},j}(t). \end{aligned}$$

From the change of numeraire theorem, we already know that $\rho(t)$ is a martingale under the T_k-forward measure. So when we change measure in the previous equation to the T_k-forward measure, the dt terms will cancel and the $dW^{\mathbb{Q},j}(t)$ will be replaced by $dW^{\mathbb{Q}^{T_k}}(t)$. Furthermore, taking into account the constant $\frac{P_{k+1,k+\tau}(0)}{D(0,T_k)}$, we get

$$d\rho(t) = \rho(t) \sum_{i=1}^{m} \left(-B^i(t,T_k) + \sum_{l=k+1}^{k+\tau} \Delta^Y_{l-1} D^p(t,T_l) B^i(t,T_l) \right) \sum_{j=1}^{m} \Sigma^{ij} dW^{\mathbb{Q}^{T_k},j}(t),$$

where again $D^p(t,T_l) = \frac{D(t,T_l)}{P_{k+1,k+\tau}(t)}$. So in vector notation we have

$$d\rho(t) = \rho(t)\kappa(t)\Sigma dW^{\mathbb{Q}^{T_k}}(t), \tag{87}$$

where $\kappa(t)$ is a $1 \times m$ vector with i-th element

$$\kappa^i(t) = -B^i(t,T_k) + \sum_{l=k+1}^{k+\tau} \Delta^Y_{l-1} D^p(t,T_l) B^i(t,T_l).$$

Again, $\kappa(t)$ is stochastic which makes it hard to work with. As before we will therefore replace the stochastic terms $D^p(t,T_i)$ by their time zero values $D^p(0,T_i)$. Using this and plugging in $B^i(t,T) = \frac{1}{A_{ii}}\left(1 - e^{-A_{ii}(T-t)}\right)$ we find

$$\kappa^i(t) = \frac{1}{A_{ii}} e^{A_{ii}t} \left[e^{-A_{ii}T_k} - \sum_{l=k+1}^{k+\tau} \Delta^Y_{l-1} e^{-A_{ii}T_l} D^p(0,T_l) \right] = e^{A_{ii}t} G^i_{k,k+\tau}.$$

From (87) we now see that $\rho(t)$ is the stochastic exponential of $(\kappa(t)\Sigma)'$. So, by Girsanov's theorem, we have that $dW^{T_k}(t) = dW^{k+1,k+\tau}(t) - \kappa(t)\Sigma dt$. Plugging this into (85) we find the dynamics of $y_{k,k+\tau}$ under the new measure:

$$\begin{aligned} dy_{k,k+\tau}(t) &= \frac{\partial \widetilde{y_{k,k+\tau}(t)}}{\partial Y(t)} \Sigma dW^{k+1,k+\tau}(t) \\ &= \frac{\partial \widetilde{y_{k,k+\tau}(t)}}{\partial Y(t)} \Sigma dW^{T_k}(t) + \frac{\partial \widetilde{y_{k,k+\tau}(t)}}{\partial Y(t)} \Sigma \kappa(t) \Sigma dt. \end{aligned}$$

Using the expression $\frac{\partial \widetilde{y_{k,k+\tau}(t)}}{\partial Y^i(t)} = e^{A_{ii}t} C^i_{k,k+\tau}$, the expression for $\kappa(t)$, $\hat{\Sigma} = \Sigma\Sigma'$ and the notation developed before, we get

$$dy_{k,k+\tau}(t) = e^{At'} diag(C_{k,k+\tau}) \Sigma dW^{T_k}(t) + e^{At'} diag(C_{k,k+\tau}) \hat{\Sigma} diag(G_{k,k+\tau}) e^{At} dt. \tag{88}$$

Integrating this and taking the expectation under $\mathbb{Q}^{T_k}$, we get

$$\mathbb{E}^{\mathbb{Q}^{T_k}}\left[y_{k,k+\tau}(T_k)\right] = y_{k,k+\tau}(0) + \int_0^{T_k} e^{At'} diag(C_{k,k+\tau}) \hat{\Sigma} diag(G_{k,k+\tau}) e^{At} dt.$$

Now note that under $\mathbb{Q}^{k+1,k+\tau}$ we have $\mathbb{E}^{\mathbb{Q}^{k+1,k+\tau}}\left[y_{k,k+\tau}(T_k)\right] = y_{k,k+\tau}(0)$ and therefore the term $\int_0^{T_k} e^{At'} diag(C_{k,k+\tau}) \hat{\Sigma} diag(G_{k,k+\tau}) e^{At} dt$ is called the convexity correction for not taking the expectation under $\mathbb{Q}^{k+1,k+\tau}$ but under $\mathbb{Q}^{T_k}$. This convexity correction term will be denoted $CC_{k,k+\tau}(T_k)$. Performing the integration leads to

$$CC_{k,k+\tau}(T_k) = \sum_{i=1}^{m}\sum_{j=1}^{m} \hat{\Sigma}_{ij} C^i_{k,k+\tau} G^j_{k,k+\tau} \frac{e^{(A_{ii}+A_{jj})T_k} - 1}{A_{ii} + A_{jj}}. \tag{89}$$

To compute the value of the option in the direct payment case, we need to know the mean and variance of $R(T_i)$ under the T_i-forward measure. The above analysis gives us the mean of each $y_{k,k+\tau}(k)$ in $\sum_{k=T_{i-s}}^{T_i} w_k y_{k,k+\tau}(k)$ under the T_k-forward measure, so we still have to derive the mean under the T_i-forward measure. This means that we still have to do step (b). From the change of numeraire theorem we have that the density $\rho(t)$ for a change of measure from the T_k-forward measure to the T_i-forward measure is given by (where $T_k \leq T_i$)

$$\begin{aligned}
\frac{d\mathbb{Q}^{T_i}}{d\mathbb{Q}^{T_k}} &= \rho(t) = \frac{D(t,T_i)/D(0,T_i)}{D(t,T_k)/D(0,T_k)} \\
&= \frac{D(0,T_k)}{D(0,T_i)} \exp\left[C(t,T_i) - C(t,T_k) - \left(\sum_{q=1}^{m} B^q(t,T_i) - \sum_{q=1}^{m} B^q(t,T_k)\right) Y^q(t)\right],
\end{aligned}$$

where we used (81). If we apply Ito's lemma using (80), we find

$$d\rho(t) = \ldots dt + \rho(t) \sum_{q=1}^{m} \left(B^q(t,T_k) - B^q(t,T_i)\right) \sum_{j=1}^{m} \Sigma^{qj} dW^{\mathbb{Q},j}(t).$$

Again. from the change of numeraire theorem, we already know that $\rho(t)$ is a martingale under the T_i-forward measure. So when we change measure in the previous equation to the T_i-forward measure, the dt terms will cancel and the $dW^{\mathbb{Q},j}(t)$ will be replaced by $dW^{\mathbb{Q}^{T_i},j}(t)$. So, switching to vector/matrix notation we have

$$d\rho(t) = \rho(t)\kappa(t)\Sigma dW^{\mathbb{Q}^{T_i}}(t), \tag{90}$$

where $\kappa(t)$ is a $1 \times m$ vector with q-th element

$$\kappa^q(t) = B^q(t,T_k) - B^q(t,T_i).$$

Plugging in $B^q(t,T) = \frac{1}{A_{qq}}\left(1 - e^{-A_{qq}(T-t)}\right)$ we find

$$\kappa^q(t) = \frac{1}{A_{qq}} e^{A_{qq}t}\left[e^{-A_{qq}T_i} - e^{-A_{qq}T_k}\right] = e^{A_{qq}t} H^q_{k,i}.$$

From (90) we get that $\rho(t)$ is the stochastic exponential of $(\kappa(t)\Sigma)'$. So, by Girsanov's theorem, we have that $dW^{T_i}(t) = dW^{T_k}(t) - \kappa(t)\Sigma dt$. Plugging this into (88) we find the dynamics of $y_{k,k+\tau}$ under the new measure:

$$\begin{aligned}
dy_{k,k+\tau}(t) &= e^{At'} diag(C_{k,k+\tau})\Sigma dW^{T_i}(t) + e^{At'} diag(C_{k,k+\tau})\hat{\Sigma} diag(H_{k,i}) e^{At} dt \\
&+ e^{At'} diag(C_{k,k+\tau})\hat{\Sigma} diag(G_{k,k+\tau}) e^{At} dt.
\end{aligned}$$

Integrating this and taking the expectation under $\mathbb{Q}^{T_i}$, we get

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{T_i}}\left[y_{k,k+\tau}(T_k)\right] &= y_{k,k+\tau}(0) + \int_0^{T_k} e^{At'} diag(C_{k,k+\tau})\hat{\Sigma} diag(H_{k,i}) e^{At} dt \\
&+ \int_0^{T_k} e^{At'} diag(C_{k,k+\tau})\hat{\Sigma} diag(G_{k,k+\tau}) e^{At} dt. \\
&= y_{k,k+\tau}(0) + CC_{k,k+\tau}(T_k) + \int_0^{T_k} e^{At'} diag(C_{k,k+\tau})\hat{\Sigma} diag(H_{k,i}) e^{At} dt,
\end{aligned}$$

with $CC_{k,k+\tau}(T_k)$ the convexity correction term as in (89). Now note that changing from the measure $\mathbb{Q}^{T_k}$ to the measure $\mathbb{Q}^{T_i}$ gives the extra term $\int_0^{T_k} e^{At'} diag(C_{k,k+\tau})\hat{\Sigma} diag(H_{k,i}) e^{At} dt$ in the expectation. This term is called the timing correction and will be denoted by $TC_{k,k+\tau}(T_k, T_i)$. If we work out the integration we get

$$TC_{k,k+\tau}(T_k, T_i) = \sum_{q=1}^{m} \sum_{j=1}^{m} \hat{\Sigma}_{qj} C^q_{k,k+\tau} H^j_{k,i} \frac{e^{(A_{qq}+A_{jj})T_k} - 1}{A_{qq} + A_{jj}}. \tag{91}$$

So, we are now able to compute the expectation of $R(T_i)$ under the T_i-forward measure. Indeed, for every $y_{k,k+\tau}(k)$ in $\sum_{k=T_{i-s}}^{T_i} w_k y_{k,k+\tau}(k)$, we have that the expectation under the T_i-forward measure is $y_{k,k+\tau}(0) + CC_{k,k+\tau}(k) + TC_{k,k+\tau}(k, T_i)$. We conclude that

$$\mu(R(T_i)) = \mathbb{E}^{\mathbb{Q}^{T_i}} \left[\sum_{k=T_{i-s}}^{T_i} w_k y_{k,k+\tau}(k) \right] = \sum_{k=T_{i-s}}^{T_i} w_k \left(y_{k,k+\tau}(0) + CC_{k,k+\tau}(k) + TC_{k,k+\tau}(k, T_i) \right). \tag{92}$$

There is still one remark to make. If $T_k \leq 0$, then $y_{k,k+\tau}(k) = y_{k,k+\tau}(T_k)$ is known. So in that case $CC_{k,k+\tau}(k)$ and $TC_{k,k+\tau}(k, T_i)$ are 0.

5.2.2 variance of $R(T_i)$

In order to compute the option value in profit sharing, we still have to find an expression for the variance of $R(T_i)$ under the T_i-forward measure. In the derivation for the mean, we found that under the T_i-forward measure we have

$$\begin{aligned} dy_{k,k+\tau}(t) &= e^{At'} diag(C_{k,k+\tau}) \Sigma dW^{T_i}(t) + e^{At'} diag(C_{k,k+\tau}) \hat{\Sigma} diag(H_{k,i}) e^{At} dt \\ &+ e^{At'} diag(C_{k,k+\tau}) \hat{\Sigma} diag(G_{k,k+\tau}) e^{At} dt. \end{aligned}$$

The drift terms are deterministic and therefore we get the following the following expression for the variance:

$$\begin{aligned} \sigma^2_{R(T_i)} &= \sum_{k=T_{i-s}}^{T_i} \sum_{l=T_{i-s}}^{T_i} w_k w_l Cov(y_{k,k+\tau}(k), y_{l,l+\tau}(l)) \\ &= \sum_{k=T_{i-s}}^{T_i} \sum_{l=T_{i-s}}^{T_i} w_k w_l Cov\left(\int_0^{T_k} e^{At'} diag(C_{k,k+\tau}) \Sigma dW^{T_i}(t), \int_0^{T_l} e^{At'} diag(C_{l,l+\tau}) \Sigma dW^{T_i}(t) \right), \end{aligned}$$

where

$$\begin{aligned} Cov&\left(\int_0^{T_k} e^{At'} diag(C_{k,k+\tau}) \Sigma dW^{T_i}(t), \int_0^{T_l} e^{At'} diag(C_{l,l+\tau}) \Sigma dW^{T_i}(t) \right) \\ &= \int_0^{\min(T_l, T_k)} e^{At'} diag(C_{k,k+\tau}) \Sigma \Sigma' diag(C_{l,l+\tau}) e^{At} dt \\ &= \sum_{i=1}^{m} \sum_{j=1}^{m} \hat{\Sigma}_{ij} C^i_{k,k+\tau} C^j_{l,l+\tau} \frac{e^{(A_{ii}+A_{jj})\min(T_l,T_k)} - 1}{A_{ii} + A_{jj}}. \end{aligned}$$

Here we can make a similar remark as we made for the mean. Namely, if $\min(T_l, T_k) \leq 0$, then $Cov(y_{k,k+\tau}(k), y_{l,l+\tau}(l)) = 0$.

5.2.3 Option price

The total value of the profit sharing option in the direct payment case is now the sum of call options on the rate $R(T_i)$ for times $T_i, i = 1, \ldots, n$. Each option has a profit sharing payoff $PS(T_i) = L(T_i)Max\{c(R(T_i) - K(T_i)), 0\}$ at time T_i. We can compute the value of this option at time 0 using that $R(T_i)$ is normally distributed with mean $\mu(R(T_i))$ and variance $\sigma^2_{R(T_i)}$ under the T_i-forward measure. For this, let $\phi_{\mu,\sigma}(.)$ be the density of a normal random variable with mean μ and standard deviation σ, $\Phi_{\mu,\sigma}$ the corresponding distribution function and $\Phi = \Phi_{0,1}$. Then the value at time 0 of the profit sharing payoff $PS(T_i)$ is:

$$\begin{aligned}
V[PS(T_i)] &= D(0,T_i)L(T_i)c\mathbb{E}^{T_i}\left[Max\{(R(T_i) - K(T_i)), 0\}\right] \\
&= D(0,T_i)L(T_i)c\int_{K(T_i)}^{\infty}(x - K(T_i))\phi_{\mu(R(T_i)),\sigma_{R(T_i)}}(x)dx \\
&= D(0,T_i)L(T_i)c\left[\int_{K(T_i)}^{\infty}x\phi_{\mu(R(T_i)),\sigma_{R(T_i)}}(x)dx - K(T_i)\int_{K(T_i)}^{\infty}\phi_{\mu(R(T_i)),\sigma_{R(T_i)}}(x)dx\right].
\end{aligned}$$

Using $\Phi_{\mu,\sigma}(z) = \Phi(\frac{z-\mu}{\sigma})$, $1 - \Phi(z) = \Phi(-z)$ and making the changes of variables $x = y + \mu$ and $\frac{y}{\sigma_{R(T_i)}} = z$, we get

$$\begin{aligned}
V[PS(T_i)] &= D(0,T_i)L(T_i)c\left[\int_{K(T_i)-\mu(R(T_i)}^{\infty}(y + \mu(R(T_i))\phi_{0,\sigma_{R(T_i)}}(y)dy\right. \\
&- \left.K(T_i)(1 - \Phi_{\mu(R(T_i)),\sigma_{R(T_i)}}(K(T_i))\right] \\
&= D(0,T_i)L(T_i)c\left[\int_{\frac{K(T_i)-\mu(R(T_i)}{\sigma_{R(T_i)}}}^{\infty}\sigma_{R(T_i)}z\phi_{0,1}(z)dz\right. \\
&+ \left.\mu(R(T_i)\int_{K(T_i)-\mu(R(T_i)}^{\infty}\phi_{0,\sigma_{R(T_i)}}(x)dx - K(T_i)\Phi\left(\frac{\mu(R(T_i) - K(T_i)}{\sigma_{R(T_i)}}\right)\right] \\
&= D(0,T_i)L(T_i)c\left[\left(\frac{-\sigma_{R(T_i)}}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}\right)_{\frac{K(T_i)-\mu(R(T_i)}{\sigma_{R(T_i)}}}^{\infty} + (\mu(R(T_i) - K(T_i))\Phi\left(\frac{\mu(R(T_i) - K(T_i)}{\sigma_{R(T_i)}}\right)\right] \\
&= D(0,T_i)L(T_i)c\left[\sigma_{R(T_i)}\phi\left(\frac{K(T_i) - \mu(R(T_i)}{\sigma_{R(T_i)}}\right) + (\mu(R(T_i) - K(T_i))\Phi\left(\frac{\mu(R(T_i) - K(T_i)}{\sigma_{R(T_i)}}\right)\right].
\end{aligned}$$

So, the total value of the profit sharing at time 0 is

$$V[PS] = \sum_{i=1}^{n}V[PS(T_i)].$$

When the profit sharing payoff at a time > 0 is dependent on swap rates observed at times < 0 (so they are known), a slight adjustment has to be done. In that case the expectation to be valued is:

$$\begin{aligned}
V[PS(T_i)] &= D(0,T_i)L(T_i)c\mathbb{E}^{T_i}\left[Max\{(R(T_i) - K(T_i)), 0\}\right] \\
&= D(0,T_i)L(T_i)c\mathbb{E}^{T_i}\left[Max\{(R(T_i)_{t>0} + R(T_i)_{t\leq 0} - K(T_i)), 0\}\right] \\
&= D(0,T_i)L(T_i)c\mathbb{E}^{T_i}\left[Max\{(R(T_i)_{t>0} - K^*(T_i)), 0\}\right],
\end{aligned}$$

where $R(T_i)_{t>0} = \sum_{k=T_{i-s}}^{T_j=0} w_k y_{k,k+\tau}(k)$, $R(T_i)_{t\leq 0} = \sum_{k=T_j}^{T_i=0} w_k y_{k,k+\tau}(k)$ and $K^*(T_i) = K(T_i) - R(T_i)_{t\leq 0}$.

The computation is similar as before and therefore we conclude that the option in the profit sharing with direct payment is relatively easy to compute and implement with a gaussian option formula.

The quality of this approximation is tested in Plat and Pelsser (2009). They report that the approximation is very accurate. We will not test the approximation ourselves, but we will proceed with compounding profit sharing, which is commonly used in practice. Furthermore, we will need the formulas derived in this section to get the approximation of the option value in the compounding profit sharing case.

5.3 Compounding profit sharing

In the case of direct payment, a client gets the profit sharing payoff each time it is determined. However, it is often the case that the profit sharing is not paid immediately, but is compounded and paid out at the end of the contract. If the maturity of the contract is T_n, then there will only be a payoff at time T_n and the payoff $L(T_n)$ is of the form (in case the customers only pays the amount $P(0)$ at the beginning of the contract)

$$L(T_n) = P(0) \prod_{i=0}^{n} \left[1 + TR(T_i) + Max\{c(R(T_i) - K(T_i)), 0\}\right], \tag{93}$$

where the definition of the variables is as before. The distribution of the right term is very complicated and there is no analytical expression for this payoff. However, if we assume that the $R(T_i)$ are independent, the expectation of $L(T_n)$ under the T_n-forward measure is

$$\begin{aligned}
\mathbb{E}^{T_n}\left[L(T_n)\right] &= \mathbb{E}^{T_n}\left[P(0) \prod_{i=0}^{n} \left[1 + TR(T_i) + Max\{c(R(T_i) - K(T_i)), 0\}\right]\right] \\
&= P(0) \prod_{i=0}^{n} \left[1 + TR(T_i) + \mathbb{E}^{T_n}\left[Max\{c(R(T_i) - K(T_i)), 0\}\right]\right].
\end{aligned}$$

To compute this expectation, we need to know the distribution of $R(T_i)$ under the T_n-forward measure. This can be done in a similar way as in the previous section. There we first made a convexity correction and then a timing correction to get the distribution of the $y_{k,k+\tau}$ in $R(T_i) = \sum_{k=T_{i-s}}^{T_i} w_k y_{k,k+\tau}(k)$ under the T_i-forward measure. The procedure is now exactly the same, except that the timing correction has to be made to T_n instead of T_i by using $TC_{k,k+\tau}(k, T_n)$ instead of $TC_{k,k+\tau}(k, T_i)$. Then we can use option pricing from section 5.2.3.

The value of the compounding profit sharing option is now the payoff $L(T_n)$ the client receives at T_n minus the payoff the client would have received without the option (which equals $P(0) \prod_{i=0}^{n} \left[1 + TR(T_i)\right]$) discounted to time 0. So the option value is

$$V[PS] = D(0, T_n) \left[\mathbb{E}^{T_n}\left[L(T_n)\right] - P(0) \prod_{i=0}^{n} \left[1 + TR(T_i)\right]\right]. \tag{94}$$

In case the customer pays a premium $P(T_j)$ at time T_j ($j = 0, \ldots, n-1$) it is easy to see that the payoff $L(T_n)$ at time T_n is given by

$$L(T_n) = \sum_{j=0}^{n-1} P(T_j) \prod_{i=j}^{n} \left[1 + TR(T_i) + Max\{c(R(T_i) - K(T_i)), 0\}\right], \tag{95}$$

Similar reasoning as in the case where the customer only pays at the beginning of the contract, gives that the expected payoff under the T_n-forward measure is

$$\begin{aligned}
\mathbb{E}^{T_n}\left[L(T_n)\right] &= \mathbb{E}^{T_n} \sum_{j=0}^{n-1} \left[P(T_j) \prod_{i=j}^{n} \left[1 + TR(T_i) + Max\{c(R(T_i) - K(T_i)), 0\}\right]\right] \\
&= \sum_{j=0}^{n-1} P(T_j) \prod_{i=j}^{n} \left[1 + TR(T_i) + \mathbb{E}^{T_n}\left[Max\{c(R(T_i) - K(T_i)), 0\}\right]\right].
\end{aligned}$$

This expectation can again be computed using the convexity and timing correction and the option pricing formulas from section 5.2.3.
The value of the compounding profit sharing option is now the payoff $L(T_n)$ the client receives at T_n minus the payoff the client would have received without the option (which equals $\sum_{j=0}^{n-1} P(T_j) \prod_{i=j}^{n} \left[1 + TR(T_i)\right]$)

discounted to time 0. So the option value is

$$V[PS] = D(0,T_n)\left[\mathbb{E}^{T_n}\left[L(T_n)\right] - \sum_{j=0}^{n-1} P(T_j)\prod_{i=j}^{n}\left[1+TR(T_i)\right]\right]. \tag{96}$$

Now, in both cases, the assumption that the $R(T_i)$ are independent is very crude, because they are obviously not independent. However the analytical approximation could still work well: When the expected $R(T_i)$ are low, the impact of the compounding effect is relatively low, resulting in a relatively good approximation of the time value of the option. When the expected $R(T_i)$ are high, the impact of the compounding effect is relatively high and the quality of the approximation will be less (in terms of time value). However, in this case the total value of the option will also be high and the impact of approximation errors in the time value on the total value will be less. This reasoning is tested by Plat and Pelsser (2009) and they show that the approximation is indeed reasonable. In the next section we will compare this to our test results.

5.4 Numerical results

We will now present test results for the approximation of the option value in the case of compounding profit sharing, where we use a 1-factor Hull-White model for the development of the short rate (i.e. in (79) and (80) we have $m = 1$, $\Sigma = \sigma_r$, $A = a$, $Y(t) = r(t)$ and $\frac{\partial \alpha(t)}{\partial t} = \theta(t)$). The product we study is one where a customer pays a premium of 100 every year starting at time 0 and continuing until time T_{n-1}. The customer receives a guaranteed rate $TR(t)$ of 4 % for every t and $K(t) = TR(t)$. Furthermore, we assume $c = 1$. The results will be given for different maturities T_n and different parameters in the Hull-White model (mean reversion a and σ_r). In calibrating the Hull-White model we assumed the initial term structure to be flat at rate 3.848 %. Also, we use 10-year swap rates and an averaging period of 10-years. The swap rates for the past are all assumed to be 3.848 %.

For $\sigma_r = 0.0075$ we have the following results.

Mean reversion(a)	Maturity (years)	Value by simulation (s.e.)	Analytical approximation
0.1	10	5.50 (0.19)	5.28
	20	41.49 (1.10)	39.10
	30	101.27 (2.15)	91.71
0.03	10	10.03 (0.31)	9.64
	20	82.47 (2.02)	76.02
	30	219.74 (4.23)	188.05
0.01	10	11.98 (0.37)	11.51
	20	102.50 (2.43)	93.36
	30	284.91 (5.15)	235.24

Table 6: Comparison of the option value as obtained by Monte Carlo simulation and by analytical approximation in compounding profit sharing. Here $\sigma_r = 0.0075$ and (s.e.) denotes the standard error.

For $\sigma_r = 0.01$ we have

Mean reversion(a)	Maturity (years)	Value by simulation (s.e.)	Analytical approximation
0.1	10	7.97 (0.25)	7.63
	20	57.38 (1.41)	53.46
	30	138.96 (2.67)	123.99
0.03	10	14.16 (0.42)	13.55
	20	112.17 (2.49)	101.26
	30	295.10 (4.90)	243.05
0.01	10	16.83 (0.48)	16.10
	20	138.82 (2.95)	123.43
	30	379.74 (5.76)	297.95

Table 7: Comparison of the option value as obtained by Monte Carlo simulation and by analytical approximation in compounding profit sharing. Here $\sigma_r = 0.01$ and (s.e.) denotes the standard error.

From these results we conclude the following. For short maturities (10 years) the approximation works well, since the approximation lies in all cases within two standard deviations of the Monte Carlo. So the approximation is in the 95 % confidence interval of the true value.

For longer maturities the approximation deteriorates and is smaller than the value obtained from the Monte Carlo simulation. This can be explained by the fact that we assume the $R(T_i)$ to be independent although they have a positive dependence. The total volatility for independent $R(T_i)$ is lower than the total volatility for positively dependent $R(T_i)$ and therefore the option value is lower. Intuitively, this can be seen as follows. If $R(T_{i-1})$ is higher, then $R(T_i)$ tends to be higher. Hence we get a path that tends to go up. Similarly, if $R(T_{i-1})$ is lower, then $R(T_i)$ tends to be lower. So we get a path that tends be lower. This gives on average a large volatility. Now, if we assume independence, then $R(T_{i-1})$ does not say anything about $R(T_i)$, so this can be either higher or lower with the same probability. The result is that each path will be less volatile and the option value will be lower.

The reasoning from Plat and Pelsser (2009) that a worse approximation is compensated by a larger option value does not hold in our situation. It is true that the higher the option value, the worse the approximation. However, the higher option value cannot compensate enough to make to approximation still relatively close to the option value. For example, with a maturity of 20 years, $\sigma_r = 0.0075$ and $a = 0.01$, the approximation is 8.9 % lower than the Monte Carlo. This is way to far off to use the analytical approximation for pricing.

6 Conclusion

In this thesis we studied two embedded options that are used in insurance and we derived analytical approximations for their values. The first option we studied was a guarantee in unit-linked insurance. We derived bounds on the value of this option using the concept of comonotonicity in the Black-Scholes model and in the Hull-White-Black-Scholes model. The derivation of the bounds making explicit use of the concept of comonotonicity has not been given before and was an effective way to derive the bounds. Then it turned out that the lower bound is always very close to the true option value and can therefore be used for pricing. This is very interesting, because it avoids the use of time-consuming Monte Carlo simulations. The upper bound turned out to be a very bad estimate for the option value. Finally, we combined the upper and lower bound which gave us the estimator of section 3.5 for the option value. This estimator is always very close to the lower bound and therefore it is my advice to use the lower bound as the approximation which avoids the computation of the moments of the upper and lower bound.

The second option we studied was profit sharing. This we modelled as a call option on an average of swap rates. To get an approximation for the option value we approximated the stochastic volatility of the swap rate by a deterministic one. Plat and Pelsser (2009) have shown that this is very accurate in case of direct payment and that it also works well in the case of compounding profit sharing. We have tested this in the case of compounding profit sharing and it turned out that for our product the approximation only works well for short maturities (10 years). For the longer maturities the approximation is to low and it is more difficult to price the contract.

7 Discussion

Looking back, there are some issues I like to discuss. The first thing is that in chapter three and four of this thesis, we assumed the initial term structure to be flat. This made the computations easier because it gives explicit formulas for all parameters in the Hull-White-Black-Scholes model. However, in practice the initial term structure will not be flat. The good thing is that using a realistic term structure does not change the underlying ideas, but only some formulas. Furthermore, Schrager and Pelsser (2004) used a realistic term structure and they concluded that the lower bound is still a very accurate estimate.
Secondly, if we look carefully to the results in chapter four, we see that the option value and standard deviation of the Monte Carlo increases if σ_r and σ_S increase. Also, in absolute terms, the difference between the lower bound and the Monte Carlo increases in this situation. This indicates that for even larger values of σ_r and σ_S the lower bound might not be a good approximation any more. Therefore, if we want to use this model with more extreme volatilities we first need to do more research to the performance of the lower bound in case of extreme volatilities. A more subtle thing is that we assumed all premiums to be invested in an investment fund which value we modelled as a geometric Brownian motion. In many cases, the investment fund will be a mix of stocks and fixed income and it is therefore also possible to model the value of stocks and fixed income separately. The stocks would again be modelled as a geometric Brownian motion, but the fixed income would depend on the short rate which was modelled by a 1-factor Hull-White model. This implies that the fixed income process has a large autocorrelation. In some sense, such a model would be more realistic, but also much more complex. Interesting research would be to see if it is still possible to get accurate analytical approximations in such a model.
Finally, the approximation for the option value in profit sharing was only accurate for short maturities. This is due to the crude assumption of independence of the underlying process. So there is still work to do to see if different techniques can lead to better approximations.

Acknowledgements

I would like to express my special thanks to Jan Dhaene for providing me with the tex versions of the articles Dhaene, Denuit, Goovaerts, Kaas and Vyncke (2002a, 2002b), Twan Possen and David van Bragt for being so helpful with all my questions concerning ALS, Willem-Jan de Goeij for helping me with the programming to get the numerical results of chapter five, my supervisors within SNS, Frans Boshuizen and Rob Smit, for the feedback they gave me every week so that I was able to move on again and my supervisor within Utrecht University, Alexander Gnedin, for his feedback when I was finishing the thesis. Finally, I would like to thank my other colleagues within SNS, Jan-Paul van der Waal, Maarten Hoedjes, Arno Boon, Aram Brouwer and Maarten Heyse for answering all my questions about the financial world and the field of risk management and for giving me a good time at SNS in general. Thanks to them I also learned a lot of other things than treated in this thesis.

8 Appendix: The theory of comonotonicity

In this appendix an extensive treatment of the theory of comonotonicity and the bounds it produces on options on sums of random variables is given. The most important parts are already given in chapter 2, but here the theory is presented in full generality and with proofs. We do this because this theory forms the basis of our research and because it is interesting on its own. For example, it could also be used to derive bounds on other financial products than are treated in this thesis. Again, I would like to stress that this theory is copied from Dhaene, Denuit, Goovaerts, Kaas, Vyncke (2002a,2002b), except that I extended a few derivations to make them easier to understand for students in the Master's degree.

8.1 Ordering Random Variables

In the sequel we will always consider random variables with finite mean. This implies that for any Random Variable X we have that $\lim_{x\to\infty} x(1-F_X(x)) = \lim_{x\to-\infty} xF_X(x) = 0$, where $F_X(x)$ denotes the cumulative distribution function (cdf) of X. Under this assumption we can derive, by partial integration, the following expression for the expectation of the random variable X:

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x dF_X(x) = \int_{-\infty}^{0} x dF_X(x) - \int_{0}^{\infty} x d(1-F_X(x)) \\
&= xF_X(x)|_{-\infty}^{0} - \int_{-\infty}^{0} F_X(x)dx - \left(x(1-F_X(x))|_0^{\infty} - \int_0^{\infty}(1-F_X(x))dx \right) \\
&= -\int_{-\infty}^{0} F_X(x)dx + \int_0^{\infty}(1-F_X(x))dx.
\end{aligned}$$

In order to derive bounds on the value of embedded options, we will replace random variables by more attractive or less attractive ones, which have a simpler structure. This will make it easier to compute the distribution functions of the random variables. Of course, we have to clarify what we mean by a more or less attractive random variable and for this purpose we first introduce the notation of "stop-loss premium". The stop-loss premium at level d of random variable X is defined by $\mathbb{E}[(X-d)_+]$, with the notation $(X-d)_+ = \max(X-d, 0)$. Using again an integration by parts, we find for $-\infty < d < \infty$

$$\begin{aligned}
\mathbb{E}[(X-d)_+] &= -\int_d^{\infty}(x-d)d(1-F_X(x)) \\
&= -\left((x-d)_+(1-F_X(x))|_d^{\infty} - \int_d^{\infty}(1-F_X(x))dx \right).
\end{aligned}$$

Using the assumption that X has finite mean we conclude that

$$\mathbb{E}[(X-d)_+] = \int_d^{\infty}(1-F_X(x))dx \qquad -\infty < d < \infty, \tag{97}$$

from which we see that the stop-loss at level d can be considered as the weight of an upper tail of (the distribution function) of X: it is the surface between the cdf F_X of X and the constant function 1, from d on. Also useful is the observation that $\mathbb{E}[(X-d)_+]$ is a decreasing continuous function of d, with derivative $F_X(d) - 1$ at d, which vanishes at $+\infty$.
Now we are able to define the stop-loss order between random variables.

Definition 8.1 *Consider two random variables X and Y. Then X is said to precede Y in the stop-loss order sense, notation $X \leq_{sl} Y$, if and only if X has lower stop-loss premiums than Y:*

$$\mathbb{E}[(X-d)_+] \leq \mathbb{E}[(Y-d)_+], \qquad -\infty < d < \infty. \tag{98}$$

Hence, $X \leq_{sl} Y$ means that X has uniformly smaller upper tails than Y, which in turn means that making a payment Y is less attractive than a payment X. Indeed, for each value of d, the probability that Y is larger than d is bigger than the probability that X is larger than d and therefore everybody would prefer to

pay the uncertain amount X.
Stop-loss order also has a natural economic interpretation in terms of expected utility. It can be shown that $X \leq_{sl} Y$ if and only if $\mathbb{E}[u(-X)] \geq \mathbb{E}[u(-Y)]$ holds for all non-decreasing concave real functions u for which the expectation exists. This means that any risk-averse decision maker would prefer to pay X instead of Y (the utility of $-X$ is larger than the utility of $-Y$).
The characterization of stop-loss order in terms of utility functions is equivalent to $\mathbb{E}[v(X)] \leq \mathbb{E}[v(Y)]$ holding for all non-decreasing convex functions v for which the expectations exist. Therefore, stop-loss order is often called increasing convex order and denoted by $\leq_{icx}$. For more details and properties of stop-loss order in a general context, see Shaked & Shanthikumar(1994), or Kaas, Van Heerwaarden & Goovaerts (1994).

Stop-loss order between random variables X and Y implies a corresponding ordering of their means. To prove this, assume that $d < 0$. From (97) we obtain

$$\begin{aligned} d + \mathbb{E}[(X-d)_+] &= -\int_d^0 1dx + \int_d^0 (1 - F_X(x))dx + \int_0^\infty (1 - F_X(x))dx \\ &= -\int_d^0 F_X(x)dx + \int_0^\infty (1 - F_X(x))dx, \end{aligned}$$

so combining this with the expression for the expectation of a random variable we get

$$\lim_{d \to -\infty} (d + \mathbb{E}[(X-d)_+]) = \mathbb{E}[X].$$

Hence, adding d to both sides of the inequality (98) in definition 8.1, and taking the limit for $d \to -\infty$, we find $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

Recall that to derive approximations for the value of embedded options, we will replace random variables X by more attractive or less attractive Y, which have a simpler structure. If $X \leq_{sl} Y$, then also $\mathbb{E}[X] \leq \mathbb{E}[Y]$ and it is intuitively clear that we get the best approximations in the case where $\mathbb{E}[X] = \mathbb{E}[Y]$. This leads to the so-called convex order.

Definition 8.2 *Consider two random variables X and Y. Then X is said to precede Y in the convex order sense, notation $X \leq_{cx} Y$, if and only if*

$$\begin{gathered} \mathbb{E}[X] = \mathbb{E}[Y], \\ \mathbb{E}[(X-d)_+] \leq \mathbb{E}[(Y-d)_+], \qquad -\infty < d < \infty. \end{gathered} \tag{99}$$

From $\mathbb{E}[(X-d)_+] - \mathbb{E}[(d-X)_+] = \mathbb{E}[X] - d$, we find

$$X \leq_{cx} Y \iff \begin{cases} \mathbb{E}[X] = \mathbb{E}[Y], \\ \mathbb{E}[(d-X)_+] \leq \mathbb{E}[(d-Y)_+], \quad -\infty < d < \infty. \end{cases}$$

Partial integration leads to

$$\begin{aligned} \mathbb{E}[(d-X)_+] &= \int_{-\infty}^d (d-x)dF_X(x) \\ &= \left((d-x)_+ F_X(x)|_{-\infty}^d + \int_{-\infty}^d F_X(x)dx \right). \end{aligned}$$

Using the assumption that X has finite mean we conclude that

$$\mathbb{E}[(d-X)_+] = \int_{-\infty}^d F_X(x)dx, \tag{100}$$

which means that $\mathbb{E}[(d-X)_+]$ can be interpreted as the weight of a lower tail of X: it is the surface between the constant function 0 and the cdf of X, from $-\infty$ to d. We have seen that stop-loss order entails uniformly heavier upper tails. The additional condition of equal means implies that convex order also leads to uniformly

heavier lower tails.
If we let $d > 0$ we find from (100)

$$d - \mathbb{E}\left[(d-X)_+\right] = \int_0^d 1dx - \int_{-\infty}^d F_X(x)dx = -\int_{-\infty}^0 F_X(x)dx + \int_0^d (1-F_X(x))dx.$$

So from the expression for the expectation of X it follows that

$$\lim_{d\to\infty} (d - \mathbb{E}\left[(d-X)_+\right]) = \mathbb{E}\left[X\right].$$

This implies that convex order can also be characterized as follows:

$$X \leq_{cx} Y \iff \begin{cases} \mathbb{E}\left[(X-d)_+\right] \leq \mathbb{E}\left[(Y-d)_+\right], & -\infty < d < \infty, \\ \mathbb{E}\left[(d-X)_+\right] \leq \mathbb{E}\left[(d-Y)_+\right], & -\infty < d < \infty. \end{cases}$$

Indeed, the $\Leftarrow$ implication follows from observing that the upper tail inequalities imply $\mathbb{E}[X] \leq \mathbb{E}[Y]$, while the lower tail inequalities imply $\mathbb{E}[X] \geq \mathbb{E}[Y]$, hence $\mathbb{E}[X] = \mathbb{E}[Y]$ must hold.
It can also be proved that $X \leq_{cx} Y$ if and only if $\mathbb{E}[v(X)] \leq \mathbb{E}[v(Y)]$ for all convex functions v, provided the expectations exist. This explains the name convex order. Note that when characterizing stop-loss order, the convex functions v are additionally required to be non-decreasing. Hence, stop-loss order is weaker: more pairs of random variables are ordered.
Finally, we also have $X \leq_{cx} Y$ if and only if $\mathbb{E}[X] = \mathbb{E}[Y]$ and $\mathbb{E}[u(-X)] \geq \mathbb{E}[u(-Y)]$ for all non-decreasing concave functions u, provided the expectations exist. Hence, in a utility context, convex order represents the common preferences of all risk-averse decision makers between random variables with equal mean.
If we look at the inequality $\mathbb{E}[v(X)] \leq \mathbb{E}[v(Y)]$ for the specific convex function $v(x) = x^2$, we immediately see that $X \leq_{cx} Y$ implies $Var(X) \leq Var(Y)$. Furthermore, the following relation between variances and stop-loss premiums holds:

$$\frac{1}{2}Var(X) = \int_{-\infty}^{\infty} \mathbb{E}\left[(X-t)_+\right] - (\mathbb{E}\left[X\right] - t)_+ dt. \tag{101}$$

To prove this relation we note that $\mathbb{E}[(X-t)_+] - (\mathbb{E}[X] - t) = \mathbb{E}[(t-X)_+]$. So we get

$$\begin{aligned} \int_{-\infty}^{\infty} \mathbb{E}\left[(X-t)_+\right] - (\mathbb{E}\left[X\right]-t)_+ dt &= \int_{-\infty}^{\mathbb{E}[X]} \mathbb{E}\left[(X-t)_+\right] - (\mathbb{E}\left[X\right]-t)dt + \int_{\mathbb{E}[X]}^{\infty} \mathbb{E}\left[(X-t)_+\right] dt \\ &= \int_{-\infty}^{\mathbb{E}[X]} \mathbb{E}\left[(t-X)_+\right] dt + \int_{\mathbb{E}[X]}^{\infty} \mathbb{E}\left[(X-t)_+\right] dt. \end{aligned}$$

Plugging in (100), interchanging the order of integration and using partial integration for the first term gives

$$\begin{aligned} \int_{-\infty}^{\mathbb{E}[X]} \mathbb{E}\left[(t-X)_+\right] dt &= \int_{-\infty}^{\mathbb{E}[X]} \int_{-\infty}^{t} F_X(x)dxdt = \int_{-\infty}^{\mathbb{E}[X]} \int_{x}^{\mathbb{E}[X]} F_X(x)dtdx \\ = \int_{-\infty}^{\mathbb{E}[X]} F_X(x)(\mathbb{E}\left[X\right]-x)dx &= -F_X(x)\frac{1}{2}(\mathbb{E}\left[X\right]-x)^2\Big|_{-\infty}^{\mathbb{E}[X]} + \frac{1}{2}\int_{-\infty}^{\mathbb{E}[X]} (x-\mathbb{E}\left[X\right])^2 dF_X(x) \\ &= \frac{1}{2}\int_{-\infty}^{\mathbb{E}[X]} (x-\mathbb{E}\left[X\right])^2 dF_X(x). \end{aligned}$$

Similarly, we find for the second term

$$\int_{\mathbb{E}[X]}^{\infty} \mathbb{E}\left[(X-t)_+\right] dt = \frac{1}{2}\int_{\mathbb{E}[X]}^{\infty} (x-\mathbb{E}\left[X\right])^2 dF_X(x).$$

This proves (101), since by definition we have $Var(X) = \int_{-\infty}^{\mathbb{E}[X]} (x-\mathbb{E}[X])^2 dF_X(x) + \int_{\mathbb{E}[X]}^{\infty} (x-\mathbb{E}[X])^2 dF_X(x)$.
As we have seen before $X \leq_{cx} Y$ implies $\mathbb{E}[X] = \mathbb{E}[Y]$. Using this, we deduce from (101) that for $X \leq_{cx} Y$ we have

$$\int_{-\infty}^{\infty} \mathbb{E}\left[(Y-t)_+\right] - \mathbb{E}\left[(X-t)_+\right] dt = \frac{1}{2}(Var(Y) - Var(X)). \tag{102}$$

Now, $X \leq_{cx} Y$ implies $\mathbb{E}\left[(Y-t)_+\right] - \mathbb{E}\left[(X-t)_+\right] \geq 0$. Thus, if $X \leq_{cx} Y$, their stop-loss distance, i.e. the integrated absolute difference of their respective stop-loss premiums, equals half the variance difference between these two random variables. Also, the integrand above is non-negative, so if in addition $Var(Y) = Var(X)$, then X and Y must necessary have equal stop-loss premiums, which implies that they are equal in distribution. Furthermore, if $X \leq_{cx} Y$, and X and Y are not equal in distribution, then $Var(X) < Var(Y)$ must hold. Note that (101) and (102) have been derived under the additional conditions that both $lim_{x\to\infty} x^2(1-F_X(x))$ and $lim_{x\to-\infty} x^2 F_X(x)$ are equal to 0(and similar for Y). A sufficient condition for this is that X and Y have finite second moments.

8.2 Inverse distribution functions

The cdf $F_X(x) = \mathbb{P}[X \leq x]$ of a random variable X is a right-continuous (r.c.) non-decreasing function with

$$F_X(-\infty) = \lim_{x\to-\infty} F_X(x) = 0, \qquad F_X(\infty) = \lim_{x\to\infty} F_X(x) = 1,$$

where right-continuous means that for any x and any sequence x_n decreasing to x we have

$$\lim_{n\to\infty} F_X(x_n) = F_X(x).$$

The usual definition of the inverse of a distribution is the non-decreasing and left-continuous (l.c.) function $F_X^{-1}(p)$ defined by

$$F_X^{-1}(p) = \inf\{x \in \mathbb{R} | F_X(x) \geq p\}, \qquad p \in [0,1] \tag{103}$$

with $\inf \emptyset = \infty$ by convention. Left-continuous means that for any x and any sequence x_n increasing to x we have

$$\lim_{n\to\infty} F_X(x_n) = F_X(x).$$

From these definitions it follows that for all $x \in \mathbb{R}$ and $p \in [0,1]$, we have

$$F_X^{-1}(p) \leq x \iff p \leq F_X(x). \tag{104}$$

Later on we will introduce the concept of comonotonicity and to treat this subject well, we will use a more sophisticated definition for inverses of distribution functions. For any real $p \in [0,1]$, a possible choice for the inverse of F_X in p is any point in the closed interval

$$[\inf\{x \in \mathbb{R} | F_X(x) \geq p\}, \sup\{x \in \mathbb{R} | F_X(x) \leq p\}],$$

where $\sup \emptyset = -\infty$. Taking the left hand border of this interval to be the value of the inverse cdf at p, we get $F_X^{-1}(p)$. Similarly, we define $F_X^{-1+}(p)$ as the right hand border of the interval:

$$F_X^{-1+}(p) = \sup\{x \in \mathbb{R} | F_X(x) \leq p\}, \qquad p \in [0,1] \tag{105}$$

which is a non-decreasing and right-continuous function. Note that $F_X^{-1}(0) = -\infty$, $F_X^{-1+}(1) = \infty$ and that all probability mass of X is contained in the interval $[F_X^{-1+}(0), F_X^{-1}(1)]$. Also note that $F_X^{-1}(p)$ and $F_X^{-1+}(p)$ are finite for all $p \in (0,1)$. In the sequel, we will always use p as a variable ranging over the open interval $(0,1)$, unless stated otherwise.

For any $\alpha \in [0,1]$, we define the α-mixed inverse function of F_X as follows:

$$F_X^{-1(\alpha)}(p) = \alpha F_X^{-1}(p) + (1-\alpha) F_X^{-1+}(p), \qquad p \in (0,1), \tag{106}$$

which is again a non-decreasing function. In particular, we find $F_X^{-1(0)}(p) = F_X^{-1+}(p)$ and $F_X^{-1(1)}(p) = F_X^{-1}(p)$. Also, for all $\alpha \in [0,1]$ we have,

$$F_X^{-1}(p) \leq F_X^{-1(\alpha)}(p) \leq F_X^{-1+}(p), \qquad p \in (0,1). \tag{107}$$

Note that the definition of $F_X^{-1(\alpha)}(p)$ is such that only for values of p corresponding to a horizontal segment of F_X lead to different values of $F_X^{-1(\alpha)}(p)$, $F_X^{-1+}(p)$ and $F_X^{-1}(p)$.

Now, let d be such that $0 < F_X(d) < 1$. Then $F_X^{-1}(F_X(d))$ and $F_X^{-1+}(F_X(d))$ are finite, and $F_X^{-1}(F_X(d)) \leq d \leq F_X^{-1+}(F_X(d))$. So for some value $\alpha_d \in [0,1]$, d can be expressed as $d = \alpha_d F_X^{-1}(F_X(d)) + (1-\alpha_d) F_X^{-1+}(F_X(d)) = F_X^{-1(\alpha_d)}(F_X(d))$. This implies that for any random variable X and any d with $0 < F_X(d) < 1$, there exists an $\alpha_d \in [0,1]$ such that $F_X^{-1(\alpha_d)} = d$.

Later on we will not only use the sophisticated inverse distribution function $F_X^{-1(\alpha)}(p)$, but we will also need the relation between inverse distribution functions of the random variables X and $g(X)$ for a monotone function g

Theorem 8.3 *Let X and $g(X)$ be real-valued random variables, and let $0 < p < 1$.*

(a) *If g is non-decreasing and l.c., then*

$$F_{g(X)}^{-1}(p) = g(F_X^{-1}(p)).$$

(b) *If g is non-decreasing and r.c., then*

$$F_{g(X)}^{-1+}(p) = g(F_X^{-1+}(p)).$$

(c) *If g is non-increasing and l.c., then*

$$F_{g(X)}^{-1+}(p) = g(F_X^{-1}(1-p)).$$

(d) *If g is non-increasing and r.c., then*

$$F_{g(X)}^{-1}(p) = g(F_X^{-1+}(1-p)).$$

Proof We will only prove (a), since the other results can be proved similarly. Let $0 < p < 1$ and consider a non-decreasing and l.c. function g. For any real x we find from (104) that

$$F_{g(X)}^{-1}(p) \leq x \iff p \leq F_{g(X)}(x).$$

As g is non-decreasing and l.c., we have that

$$g(z) \leq x \iff z \leq \sup\{y|g(y) \leq x\} \tag{108}$$

holds for all real z and x. From this it follows that $F_{g(X)}(x) = \mathbb{P}\left[g(X) \leq x\right] = \mathbb{P}\left[X \leq \sup\{y|g(y) \leq x\}\right]$ and therefore we have

$$p \leq F_{g(X)}(x) \iff p \leq F_X\left[\sup\{y|g(y) \leq x\}\right].$$

If $\sup\{y|g(y) \leq x\}$ is finite, then we find from (104) and the equivalence above

$$p \leq F_X\left[\sup\{y|g(y) \leq x\}\right] \iff F_X^{-1}(p) \leq \sup\{y|g(y) \leq x\}.$$

If $\sup\{y|g(y) \leq x\}$ is ∞ or $-\infty$ we cannot use (104), but then the equivalence still holds. Indeed, if the supremum equals $-\infty$, then the equivalence becomes $p \leq 0 \iff F_X^{-1}(p) \leq -\infty$, which is true. If the supremum equals ∞, then the equivalence becomes $p \leq 1 \iff F_X^{-1}(p) \leq \infty$, which is also true.
From the previous equivalence and (108) we now find

$$F_X^{-1}(p) \leq \sup\{y|g(y) \leq x\} \iff g(F_X^{-1}(p)) \leq x.$$

Finally, combining the equivalences we get that

$$F_{g(X)}^{-1}(p) \leq x \iff g(F_X^{-1}(p)) \leq x$$

holds for all values of x, which means that (a) must hold ∎

For the special case that g and F_X are continuous and strictly increasing on $[F_X^{-1+}(0), F_X^{-1}(1)]$ the proof is simpler. In this case we have

$$F_{g(X)}(x) = \mathbb{P}\left[g(X) \leq x\right] = \mathbb{P}\left[X \leq g^{-1}(x)\right] = (F_X \circ g^{-1})(x),$$

which is a continuous and strictly increasing function of x. The results (a) and (b) then follow by inversion of this relation.
Similarly, if g is continuous and strictly decreasing and F_X is continuous and strictly increasing we have

$$F_{g(X)}(x) = \mathbb{P}\left[g(X) \leq x\right] = \mathbb{P}\left[X \geq g^{-1}(x)\right] = 1 - (F_X \circ g^{-1})(x).$$

From this it follows that $(F_X \circ g^{-1})(x)$ is strictly decreasing (since $F_{g(X)}(x)$ is strictly increasing) and inversion gives

$$\begin{aligned} F_{g(X)}^{-1}(p) &= \inf\{x \in \mathbb{R} | F_{g(X)}(x) \geq p\} = \inf\{x \in \mathbb{R} | 1 - (F_X \circ g^{-1})(x) \geq p\} \\ &= \inf\{x \in \mathbb{R} | (F_X \circ g^{-1})(x) \leq 1 - p\} = (F_X \circ g^{-1})^{-1}(1-p) = g(F_X^{-1}(1-p)). \end{aligned}$$

So indeed, (c) and (d) hold.

Hereafter, we will reserve the notation U for a $uniform(0,1)$ random variable, i.e. $F_U(p) = p$ and $F_U^{-1}(p) = p$ for all $0 < p < 1$. Then one can prove that for all $\alpha \in [0,1]$,

$$X =_d F_X^{-1}(U) =_d F_X^{-1+}(U) =_d F_X^{-1(\alpha)}(U), \tag{109}$$

where $=_d$ denotes equality in distribution. The first equality follows immediately from (104), since (104) implies

$$F_{F_X^{-1}(U)}(x) = \mathbb{P}\left[F_X^{-1}(U) \leq x\right] = \mathbb{P}\left[U \leq F_X(x)\right] = F_X(x).$$

Note that F_X has at most a countable number of horizontal segments, implying that the last three random variables in (109) only differ in a null-set of values of U. This implies that these random variables are equal with probability one.

8.3 Comonotonicity

In valuing the option in unit-linked insurance, we have to deal with a sum of dependent random variables. That is, we have to deal with random variables of the type $S = \sum_{i=1}^{n} X_i$ where the X_i are not mutually independent. In some situations it can also be that the multivariate distribution function of the random vector $X = (X_1, X_2, \ldots, X_n)$ is not completely specified because we only know the marginal distribution functions of the random variables X_i or that the dependence structure is to cumbersome to work with. In case the random variable S represents a payment the insurance company has to do, it may be helpful to find the dependence structure for the random vector $(X_1, X_2, \ldots, X_n)$ producing the least favourable payment S with the given marginals X_i. This is a prudent strategy, since the insurance company will then make decisions based on the least favourable outcome.

Therefore, given the marginal distributions of the terms in a random variable $S = \sum_{i=1}^{n} X_i$, we will look for the joint distribution with the largest sum in the convex order sense. As we will prove in section 6.5, the convex-largest sum of the components of a random vector with given marginals will be obtained in the case that the random vector $(X_1, X_2, \ldots, X_n)$ has the *comonotonic* distribution, which means that each two possible outcomes $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ of $(X_1, X_2, \ldots, X_n)$ are ordered component wise.

We start by defining comonotonicity of a set of n-vectors in $\mathbb{R}^n$. A n-vector $(x_1, \ldots, x_n)$ will be denoted by $\vec{x}$. For two n-vectors $\vec{x}$ and $\vec{y}$, the notation $\vec{x} \leq \vec{y}$ will be used for the component wise order which is defined by $x_i \leq y_i$ for all $i = 1, \ldots, n$.

Definition 8.4 *The set $A \subseteq \mathbb{R}^n$ is said to be comonotonic if for any $\vec{x}$ and $\vec{y}$ in A, either $\vec{x} \leq \vec{y}$ or $\vec{y} \leq \vec{x}$ holds.*

So, a set $A \subseteq \mathbb{R}^n$ is comonotonic if for any $\vec{x}$ and $\vec{y}$ in A we have that if $x_i < y_i$ for some i, then $\vec{x} \leq \vec{y}$ must hold. Hence, a comonotonic set is simultaneously non-decreasing in each component. Notice that a comonotonic set is a thin set: it cannot contain any subset of dimension larger than 1. Any subset of a comonotonic set is also comonotonic.

We will denote the (i, j) projection of a set A in $\mathbb{R}^n$ by $A_{i,j}$ It is defined by

$$A_{i,j} = \{(x_i, x_j) | \vec{x} \in A\}$$

Lemma 8.5 *$A \subseteq \mathbb{R}^n$ is comonotonic if and only if $A_{i,j}$ is comonotonic for all $i \neq j$ in $1, 2, \ldots, n$.*

Proof It is clear that if $A \subseteq \mathbb{R}^n$ is comonotonic then $A_{i,j}$ is comonotonic for all $i \neq j$ in $1, 2, \ldots, n$.

Now suppose that $A_{i,j}$ is comonotonic for all $i \neq j$ in $1, 2, \ldots, n$ and suppose that $\vec{x}$ and $\vec{y}$ in A such that neither $\vec{x} \leq \vec{y}$ nor $\vec{y} \leq \vec{x}$ holds. Then we can find $i \in 1, 2, \ldots, n$ and $j \in 1, 2, \ldots, n$ $(i \neq j)$ such that either $x_i > y_i$ and $x_j \leq y_j$ or $x_i < y_i$ and $x_j \geq y_j$. But (x_i, x_j) and (y_i, y_j) are in $A_{i,j}$, which is comonotonic. So this cannot happen and therefore we have for $\vec{x}$ and $\vec{y}$ in A that either $\vec{x} \leq \vec{y}$ or $\vec{y} \leq \vec{x}$ must hold. Hence, A is comonotonic. ∎

Next, we will define the notion of support of an n-dimensional random vector $\vec{X} = (X_1, \ldots, X_n)$. Any subset $A \subseteq \mathbb{R}^n$ will be called a support of $\vec{X}$ if $\mathbb{P}\left[\vec{X} \in A\right] = 1$ holds true. In general we will be interested in supports which are " as small as possible ". Informally, the smallest support of a random vector $\vec{X}$ is the subset of $\mathbb{R}^n$ that is obtained by subtracting from $\mathbb{R}^n$ all points which have a zero probability neighbourhood with respect to $\vec{X}$. This support can be interpreted as the set of all possible outcomes of $\vec{X}$.

We will now define the important concept of comonotonicity of random vectors:

Definition 8.6 *A random vector $\vec{X} = (X_1, \ldots, X_n)$ is said to be comonotonic if it has a comonotonic support.*

From the definition we can conclude that comonotonicity is a very strong positive dependency structure. Indeed, if $\vec{x}$ and $\vec{y}$ are elements of the comonotonic support of X, i.e. $\vec{x}$ and $\vec{y}$ are possible outcomes of $\vec{X}$, then they must be ordered component wise. This is also the intuition behind comonotonicity. A random vector $\vec{X}$ with components X_i is comonotonic if the set of all possible outcomes is ordered component wise. In the following theorem, some useful equivalent characterizations are given for comonotonicity of a random vector.

Theorem 8.7 *A random vector $\vec{X} = (X_1, \ldots, X_n)$ is comonotonic if and only if one of the following equivalent conditions holds:*

(1) *$\vec{X}$ has a comonotonic suppport.*

(2) *For all $\vec{x} = (x_1, \ldots, x_n)$, we have*

$$F_{\vec{X}}(\vec{x}) = \min\left(F_{X_1}(x_1), \ldots, F_{X_n}(x_n)\right), \tag{110}$$

where $F_{\vec{X}}(\vec{x}) = \mathbb{P}\left[\vec{X} \leq \vec{x}\right]$ is the multivariate cdf of $\vec{X}$.

(3) *For $U =_d Uniform(0,1)$, we have*

$$\vec{X} =_d \left(F_{X_1}^{-1}(U), \ldots, F_{X_n}^{-1}(U)\right). \tag{111}$$

(4) *There exists a random variable Z and non-decreasing functions f_i, $i = 1, \ldots, n$, such that*

$$\vec{X} =_d \left(f_1(Z), \ldots, f_n(Z)\right). \tag{112}$$

Proof $(\mathbf{1}) \Rightarrow (\mathbf{2})$: Assume that $\vec{X}$ has comonotonic support B. Let $\vec{x} \in \mathbb{R}^n$ and let A_j be defined by

$$A_j = \{\vec{y} \in B | y_j \leq x_j\}, \qquad j = 1, \ldots, n.$$

Because of the comonotonicity of B, there exists an i such that $A_i = \cap_{j=1}^n A_j$. Hence, we find

$$\begin{aligned} F_{\vec{X}}(\vec{x}) &= \mathbb{P}\left[\vec{X} \in \cap_{j=1}^n A_j\right] = \mathbb{P}\left[\vec{X} \in A_i\right] = F_{X_i}(x_i) \\ &= \min(F_{X_1}(x_1), \ldots, F_{X_n}(x_n)). \end{aligned}$$

The last equality follows from $A_i \subseteq A_j$, so that $F_{X_i}(x_i) \leq F_{X_j}(x_j)$ holds for all values of j.
$(\mathbf{2}) \Rightarrow (\mathbf{3})$: Now assume that $F_{\vec{X}}(\vec{x}) = \min\left(F_{X_1}(x_1), \ldots, F_{X_n}(x_n)\right)$ for all $\vec{x} = (x_1, \ldots, x_n)$. Then we find (using (104) for the first equality):

$$\begin{aligned} \mathbb{P}\left[F_{X_1}^{-1}(U) \leq x_1, \ldots, F_{X_n}^{-1}(U) \leq x_n\right] &= \mathbb{P}\left[U \leq F_{X_1}(x_1), \ldots, U \leq F_{X_n}(x_n)\right] \\ &= \mathbb{P}\left[U \leq \min_{j=1,\ldots,n}\left(F_{X_j}(x_j)\right)\right] \\ &= \min_{j=1,\ldots,n}\left(F_{X_j}(x_j)\right) = F_{\vec{X}}(\vec{x}), \end{aligned}$$

which means that (111) holds.
$(\mathbf{3}) \Rightarrow (\mathbf{4})$: this is clear.
$(\mathbf{4}) \Rightarrow (\mathbf{1})$: Assume (4) and let B be the support of Z. The set of possible of outcomes of X is then $\{(f_1(z), \ldots, f_n(z)) | z \in B\}$. Since the functions f_i are non-decreasing this is comonotonic, which implies that $\vec{X}$ is indeed comonotonic. ∎

From (110) we see that, in order to find the probability of all the outcomes of n comonotonic risks X_i being less than x_i, $(i = 1, \ldots, n)$, one simply takes the probability of the least likely of these events. It is obvious that for any random vector $(X_1, \ldots, X_n)$, not necessarily comonotonic, the following inequality holds:

$$\mathbb{P}\left[X_1 \leq x_1, \ldots, X_n \leq x_n\right] \leq \min\{F_{X_1}(x_1), \ldots, F_{X_n}(x_n)\}, \tag{113}$$

and we have seen in the proof of theorem 8.7 that the function $\min\{F_{X_1}(x_1), \ldots, F_{X_n}(x_n)\}$ is the multivariate cdf of the random vector $\left(F_{X_1}^{-1}(U), \ldots, F_{X_n}^{-1}(U)\right)$, which has the same marginals as $(X_1, \ldots, X_n)$. The inequality (113) states that in the class of all random vectors $(X_1, \ldots, X_n)$ with the same marginals, the probability that all X_i simultaneously realize 'small' values is maximized if the vector is comonotonic, suggesting that comonotonicity is indeed a very strong positive dependency structure.

In the sequel, for any random vector $(X_1, \ldots, X_n)$, the notation $(X_1^c, \ldots, X_n^c)$ will be used to indicate a comonotonic random vector with the same marginals as $(X_1, \ldots, X_n)$. From (111), we find that for any random vector $\vec{X}$ the outcome of its comonotonic counterpart $\vec{X}^c = (X_1^c, \ldots, X_n^c)$ is with probability 1 in the set

$$\{F_{X_1}^{-1}(p), \ldots, F_{X_n}^{-1}(p) | 0 < p < 1\}. \tag{114}$$

This support is not necessarily a connected curve. Indeed, all horizontal segments of the cdf of X_i lead to missing pieces in this curve. By connecting the endpoints of consecutive pieces in this curve by straight lines, we obtain a comonotonic connected curve in $\mathbb{R}^n$. Hence, it may be traversed in a direction which is upwards for all components simultaneously. We will call this set the connected support of $\vec{X}^c$. To give a possible parametrization, we note that the random vectors $\left(F_{X_1}^{-1}(U), \ldots, F_{X_n}^{-1}(U)\right)$ and $\left(F_{X_1}^{-1(\alpha_1)}(U), \ldots, F_{X_n}^{-1(\alpha_n)}(U)\right)$ as defined previously are equal with probability 1. So comonotonicity of $\vec{X}$ can also be characterized by

$$\vec{X} =_d \left(F_{X_1}^{-1(\alpha_1)}(U), \ldots, F_{X_n}^{-1(\alpha_n)}(U)\right), \tag{115}$$

for $U =_d Uniform(0, 1)$ and real numbers $\alpha_i \in [0, 1]$. From this it follows that a possible parametrization of the connected support of $\vec{X}^c$ is given by

$$\{\left(F_{X_1}^{-1(\alpha)}(p), \ldots, F_{X_n}^{-1(\alpha)}(p)\right) | 0 < p < 1, 0 \leq \alpha \leq 1\}. \tag{116}$$

Before we go to an example, we state the following useful theorem.

Theorem 8.8 *A random vector $\vec{X}$ is comonotonic if and only if (X_i, X_j) is comonotonic for all $i \neq j$ in $\{1, \ldots, n\}$.*

Proof The $\Rightarrow$ implication is clear.
For the proof of the $\Leftarrow$ implication, consider the set $A \in \mathbb{R}^n$ defined by

$$A = \{(F_{X_1}^{-1}(p), \ldots, F_{X_n}^{-1}(p)) | 0 < p < 1\}.$$

Its (i, j)-projections are given by

$$A_{i,j} = \{(F_{X_i}^{-1}(p), F_{X_j}^{-1}(p)) | 0 < p < 1\}.$$

Since (X_i, X_j) is comonotonic for all $i \neq j$, we have for all $i \neq j$ that $(X_i, X_j) \in A_{i,j}$ with probability 1. But the event $(X_i, X_j) \in A_{i,j}$ for all $i \neq j$ is equivalent to the event $\vec{X} \in A$. So $\mathbb{P}\left[\vec{X} \in A\right] = 1$ and therefore the comonotonic set A is a support of $\vec{X}$. Hence, $\vec{X}$ is a comonotonic random vector. ∎

This theorem states that comonotonicity of a random vector is equivalent with pairwise comonotonicity.

Example As an example of comonotonicity, consider the distributions $\vec{X} =_d Uniform\{0, 1, 2, 3\}$ and $Y =_d Binomial(3, \frac{1}{2})$. The random vector (X, Y) is not comonotonic, because two possible outcomes are for example $(0, 1)$ and $(1, 0)$. And for these outcomes it does not hold that $(0, 1) \leq (1, 0)$ nor does it hold that$(0, 1) \geq (1, 0)$. Now, we can compute that for $0 < p < 1$ we have

$$\begin{aligned}
(F_X^{-1}(p), F_Y^{-1}(p)) &= (0,0) \ for \ 0 < p \leq \frac{1}{8}, \\
&= (0,1) \ for \ \frac{1}{8} < p \leq \frac{2}{8}, \\
&= (1,1) \ for \ \frac{2}{8} < p \leq \frac{4}{8}, \\
&= (2,2) \ for \ \frac{4}{8} < p \leq \frac{6}{8}, \\
&= (3,2) \ for \ \frac{6}{8} < p \leq \frac{7}{8}, \\
&= (0,1) \ for \ \frac{7}{8} < p < 1.
\end{aligned}$$

The support of (X^c, Y^c) is just these six points and the connected support arises by simply connecting them consecutively with straight lines.

8.4 Sums of comonotonic random variables

Recall again that if we want to value the option in unit-linked insurance, we have to deal with a sum of dependent random variables. That is, we will have to deal with random variables of the type $S = \sum_{i=1}^n X_i$ where the X_i are not mutually independent. In the sequel, we will use the notation S^c for the comonotonic counterpart of S. That is, $S^c = \sum_{i=1}^n X_i^c$, where $(X_1^c, \ldots, X_n^c)$ is the comonotonic counterpart of the random vector $(X_1, \ldots, X_n)$.
Further on, we will prove that approximating the distribution function of S by the distribution function of the comonotonic sum S^c is a prudent strategy for an insurer in the sense that $S \leq_{cx} S^c$. This means that computing the value of the option in unit-linked insurance using S^c instead of S will lead to a higher option value and since the option is a liability for the insurer, the insurer will base its decisions on a worst case scenario.
Of course, this approximation will only be meaningful if we can easily determine the distribution function and the stop-loss premiums of S^c (since we will need those to compute the option value). In this section we will show that these quantities can indeed easily be determined from the marginal distribution functions of the terms in the sum. In the next theorem we will first prove that the inverse distribution function of a sum of comonotonic random variables is simply the sum of the inverse distribution functions of the marginal distributions.

Theorem 8.9 *The α-inverse distribution function $F_{S^c}^{-1(\alpha)}$ of a sum S^c of comonotonic random variables $(X_1^c, \ldots, X_n^c)$ is given by*

$$F_{S^c}^{-1(\alpha)}(p) = \sum_{i=1}^n F_{X_i}^{-1(\alpha)}(p), \qquad 0 < p < 1, \qquad 0 \leq \alpha \leq 1. \tag{117}$$

Proof Consider $(X_1, \ldots, X_n)$ and its comonotonic counterpart $(X_1^c, \ldots, X_n^c)$. Then we have from theorem 8.7 (3) that $(X_1^c, \ldots, X_n^c) =_d \left(F_{X_1}^{-1}(U), \ldots, F_{X_n}^{-1}(U)\right)$, with U uniformly distributed on $(0,1)$. So $S^c = \sum_{i=1}^n X_i^c =_d g(U)$ with the function g defined by

$$g(u) = \sum_{i=1}^n F_{X_i}^{-1}(u), \qquad 0 < u < 1.$$

By definition of the inverse distribution function $F_X^{-1}(p)$ of a random variable X, we have that this function is left-continuous and non-decreasing. So g is also left-continuous and non-decreasing. Applying theorem 8.3(a) we therefore get

$$F_{S^c}^{-1}(p) = F_{g(U)}^{-1}(p) = g(F_U^{-1}(p)) = g(p), \qquad 0 < p < 1,$$

so the inverse distribution function of S^c can be computed from

$$F_{S^c}^{-1}(p) = \sum_{i=1}^n F_{X_i}^{-1}(p), \qquad 0 < p < 1.$$

Similarly, we find from (109) that we also have $S^c = \sum_{i=1}^n X_i^c =_d h(U)$ with the function h defined by

$$h(u) = \sum_{i=1}^n F_{X_i}^{-1(+)}(u), \qquad 0 < u < 1.$$

Applying theorem 8.3(b) then gives

$$F_{S^c}^{-1(+)}(p) = h(p) = \sum_{i=1}^n F_{X_i}^{-1(+)}(p), \qquad 0 < p < 1.$$

Multiplying the last equality by $1-\alpha$, the equivalent equality for $F_{S^c}^{-1}(p)$ by α and adding up, we get the result ∎

Now, from section 6.2 we know that $F_{X_i}^{-1(+)}$ is a right-continuous function on $[0,1]$ and $F_{X_i}^{-1}$ is a left-continuous function on $[0,1]$. So, taking limits (respectively from the right and left) in (117) we find

$$F_{S^c}^{-1(+)}(0) = \sum_{i=1}^{n} F_{X_i}^{-1(+)}(0), \tag{118}$$

and

$$F_{S^c}^{-1}(1) = \sum_{i=1}^{n} F_{X_i}^{-1}(1). \tag{119}$$

Hence, the minimal value of the comonotonic sum equals the sum of the minimal values of each term. Similarly, the maximal value of the comonotonic sum equals the sum of the maximal values of each term. The number $\sum_{i=1}^{n} F_{X_i}^{-1(+)}(0)$, which is either finite or $-\infty$, is the minimum possible value of S^c, and $\sum_{i=1}^{n} F_{X_i}^{-1}(1)$ is the maximum.

For any $(X_1,\ldots,X_n)$, we have that $S = X_1 + \ldots + X_n \geq \sum_{i=1}^{n} F_{X_i}^{-1(+)}(0)$ must hold with probability 1. This holds because we have seen in section 6.2 that $X_i \geq F_{X_i}^{-1(+)}(0)$ with probability 1 for every i. This implies

$$\sum_{i=1}^{n} F_{X_i}^{-1(+)}(0) \leq F_S^{-1(+)}(0), \tag{120}$$

because the minimal value of the sum S, given by $F_S^{-1(+)}(0)$, must be bigger then the sum of minimal values. Similarly, we find

$$F_S^{-1}(1) \leq \sum_{i=1}^{n} F_{X_i}^{-1}(1). \tag{121}$$

This means that the sum S of the components of any random vector $(X_1,\ldots,X_n)$ has a support that is contained in the interval $\left[\sum_{i=1}^{n} F_{X_i}^{-1(+)}(0), \sum_{i=1}^{n} F_{X_i}^{-1}(1)\right]$.

So far, we only derived an expression for $F_{S^c}^{-1(\alpha)}(p)$ and we derived some properties. However, we are interested in the distribution function F_{S^c} of S^c. For us, an important case will be that the marginal distribution functions $F_{X_i}, i = 1,\ldots,n$ of the comonotonic random vector $(X_1^c,\ldots,X_n^c)$ are strictly increasing and continuous. Where, from now on, the expression F_X is strictly increasing should be interpreted as F_X is strictly increasing on $\left(F_X^{-1(+)}(0), F_X^{-1}(1)\right)$. In this case we can find a useful relation that enables us to compute the distribution function F_{S^c} of S^c once we know $F_{S^c}^{-1(\alpha)}(p)$. For this we first observe that for any random variable X, the following equivalences hold:

$$F_X \textit{ is strictly increasing} \iff F_X^{-1} \textit{ is continuous on } (0,1), \tag{122}$$

and also

$$F_X \textit{ is continuous} \iff F_X^{-1} \textit{ is strictly increasing on } (0,1). \tag{123}$$

So if the marginal distribution functions $F_{X_i}, i = 1,\ldots,n$ of the comonotonic random vector $(X_1^c,\ldots,X_n^c)$ are strictly increasing and continuous, then each inverse distribution function $F_{X_i}^{-1}$ is continuous on $(0,1)$ by (122). This implies that $F_{S^c}^{-1}$ is continuous on $(0,1)$ because $F_{S^c}^{-1}(p) = \sum_{i=1}^{n} F_{X_i}^{-1}(p)$ holds for $0 < p < 1$. This in turn implies that F_{S^c} is strictly increasing on $(F_{S^c}^{-1(+)}(0), F_{S^c}^{-1}(1))$ by using (122) the other way around. By a similar reasoning using (123) we find that F_{S^c} is continuous.
Hence, in case of strictly increasing and continuous marginals, for any $F_{S^c}^{-1(+)}(0) < x < F_{S^c}^{-1}(1)$, the probability $F_{S^c}(x)$ is uniquely determined by $F_{S^c}^{-1}(F_{S^c}(x)) = x$, or equivalently from theorem 8.9

$$\sum_{i=1}^{n} F_{X_i}^{-1}(F_{S^c}(x)) = x, \qquad F_{S^c}^{-1(+)}(0) < x < F_{S^c}^{-1}(1). \tag{124}$$

It suffices to solve the latter equation to get $F_{S^c}(x)$.

As mentioned before, we will also be interested in the stop-loss premium of a sum of comonotonic random variables. In the following theorem we prove that this stop-loss premium can also be obtained form the stop-loss premiums of the terms.

Theorem 8.10 *The stop-loss premiums of the sum S^c of the components of the comonotonic random vector $(X_1^c, \ldots, X_n^c)$ are given by*

$$\mathbb{E}\left[(S^c - d)_+\right] = \sum_{i=1}^{n} \mathbb{E}\left[(X_i - d_i)_+\right], \qquad F_{S^c}^{-1(+)}(0) < d < F_{S^c}^{-1}(1), \tag{125}$$

with the d_i given by

$$d_i = F_{X_i}^{-1(\alpha_d)}(F_{S^c}(d)), \qquad i = 1, \ldots, n \tag{126}$$

and $\alpha_d \in [0,1]$ determined by

$$F_{S^c}^{-1(\alpha_d)}(F_{S^c}(d)) = d. \tag{127}$$

Proof Let $d \in (F_{S^c}^{-1(+)}(0), F_{S^c}^{-1}(1))$. Then $0 < F_{S^c}(d) < 1$. As the connected support of X^c as defined in (116) is comonotonic, it can have at most one point of intersection with the hyperplane $\{\vec{x} | x_1 + \ldots + x_n = d\}$. This is because the hyperplane clearly cannot contain different points $\vec{x}$ and $\vec{y}$ such that $\vec{x} \leq \vec{y}$ or $\vec{x} \geq \vec{y}$ holds.
We will now proof that the vector $\vec{d} = (d_1, \ldots, d_n)$ as defined above is the unique point of this intersection. As $0 < F_{S^c}(d) < 1$ must hold, we know from section 6.2 that there exists an $\alpha_d \in [0,1]$ that fulfils condition (127). Also note that by theorem 8.9, we have that $\sum_{i=1}^n d_i = \sum_{i=1}^n F_{X_i}^{-1(\alpha_d)}(F_{S^c}(d)) = F_{S^c}^{-1(\alpha_d)}(F_{S^c}(d)) = d$. Hence, the vector $\vec{d}$ is an element of both the connected support of X^c and the hyperplane $\{\vec{x} | x_1 + \ldots + x_n = d\}$. We can conclude that $\vec{d}$ is the unique point of intersection of the connected support and the hyperplane. Now, let $\vec{x}$ be an element of the connected support of X^c. Then the following equality holds:

$$(x_1 + \ldots + x_n - d)_+ = (x_1 - d_1)_+ + \ldots + (x_n - d_n)_+.$$

This is because $\vec{x}$ and $\vec{d}$ are both elements of the connected support of X^c, so if there exists any j such that $x_j > d_j$ holds, then we also have $x_k \geq d_k$ for all k (the connected support is comonotonic), and the left hand side equals the right hand side because $\sum_{i=1}^n d_i = d$. On the other hand, when all $x_j \leq d_j$, both sides are 0. Now replacing constants by the corresponding random variables in the equality above and taking expectations, we find (125) ∎

Note that we also have

$$\mathbb{E}\left[(S^c - d)_+\right] = \sum_{i=1}^{n} \mathbb{E}\left[X_i\right] - d, \qquad if\ d \leq F_{S^c}^{-1+}(0) \tag{128}$$

and

$$\mathbb{E}\left[(S^c - d)_+\right] = 0, \qquad if\ d \geq F_{S^c}^{-1}(1). \tag{129}$$

This follows since for $d \leq F_{S^c}^{-1(+)}(0)$ we have $S^c \geq d$ with probability 1, $X_i \geq d_i = F_{X_i}^{-1(\alpha_d)}(F_{S^c}(d))$ with probability 1 and $\sum_{i=1}^n d_i = d$. Also, for $d \geq F_{S^c}^{-1}(1)$ we have $S^c \leq d$ with probability 1.

So from (118), (119), (128), (129) and Theorem 8.10 we can conclude that for any real d, there exist d_i with $\sum_{i=1}^n d_i = d$, such that $\mathbb{E}\left[(S^c - d)_+\right] = \sum_{i=1}^n \mathbb{E}\left[(X_i - d_i)_+\right]$ holds.

The expression for the stop-loss premiums of a comonotonic sum S^c can also be written in terms of the usual inverse distribution functions. Indeed, for any $d \in \left(F_{S^c}^{-1(+)}(0), F_{S^c}^{-1}(1)\right)$, we have

$$\begin{aligned} & \mathbb{E}\left[(X_i - F_{X_i}^{-1(\alpha_d)}\left(F_{S^c}(d)\right))_+\right] \\ = \quad & \mathbb{E}\left[(X_i - F_{X_i}^{-1}\left(F_{S^c}(d)\right))_+\right] - \left(F_{X_i}^{-1(\alpha_d)}\left(F_{S^c}(d)\right) - F_{X_i}^{-1}\left(F_{S^c}(d)\right)\right)(1 - F_{S^c}(d)) \end{aligned}$$

Summing over i, and taking into account the definition of α_d,we find for any $d \in \left(F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1)\right)$:

$$\begin{aligned}\mathbb{E}\left[(S^c - d)_+\right] &= \sum_{i=1}^{n} \mathbb{E}\left[(X_i - F_{X_i}^{-1}\left(F_{S^c}(d)\right))_+\right] \\ &\quad - \left(d - F_{S^c}^{-1}\left(F_{S^c}(d)\right)\right)\left(1 - F_{S^c}(d)\right). \end{aligned} \tag{130}$$

In the important case where the marginal cdf's F_{X_i} are strictly increasing, (130) reduces to

$$E[(S^c - d)_+] = \sum_{i=1}^{n} \mathbb{E}\left[(X_i - F_{X_i}^{-1}\left(F_{S^c}(d)\right)_+\right], \qquad d \in \left(F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1)\right). \tag{131}$$

From Theorem 8.10, we can conclude that any stop-loss premium of a sum of comonotonic random variables can be written as the sum of stop-loss premiums for the individual random variables involved. The theorem provides an algorithm for directly computing stop-loss premiums of sums of comonotonic random variables, without having to compute the entire distribution function of the sum itself. Indeed, in order to compute the stop-loss premium with level d, we only need to know $F_{S^c}(d)$, which, in the case of strictly increasing and continuous marginals, can be computed from (124).

Application of the relation $E[(X-d)_+] = E\left[(d-X)_+\right] + E\left[X\right] - d$ for S^c and the X_i in relation (125) leads to the following expression for the lower tails of a sum of comonotonic random variables:

$$E\left[(d - S^c)_+\right] = \sum_{i=1}^{n} E\left[(d_i - X_i)_+\right], \qquad F_{S^c}^{-1+}(0) < d < F_{S^c}^{-1}(1), \tag{132}$$

with the d_i as defined in (126) and (127).

Example Consider a random vector $\vec{X}$ with exponential marginals: $X_i \sim \mathrm{Exp}(1/\beta_i)$. Then

$$F_{X_i}(x) = 1 - e^{-\frac{x}{\beta_i}}, \qquad \beta_i > 0,\ x \geq 0. \tag{133}$$

We find the following expression for the inverse distribution function:

$$F_{X_i}^{-1}(p) = -\beta_i\ \ln\left(1-p\right), \qquad 0 < p < 1. \tag{134}$$

One can easily verify that the stop-loss premium with level d is given by

$$\int_d^{\infty}(1 - F_{X_i})dx = E[(X_i - d)_+] = \beta_i\ e^{-\frac{d}{\beta_i}}, \qquad 0 < d < \infty \tag{135}$$

The inverse distribution function of the comonotonic sum S^c is then given by

$$F_{S^c}^{-1}(p) = -\left(\sum_{i=1}^{n} \beta_i\right)\ \ln\left(1-p\right), \qquad 0 < p < 1. \tag{136}$$

This means that the comonotonic sum of exponentially distributed random variables is again exponentially distributed with parameter $\beta = \sum_{i=1}^{n} \beta_i$. The stop-loss premiums of S^c are given by

$$E\left[(S^c - d)_+\right] = \beta\ e^{-\frac{d}{\beta}}, \qquad 0 < d < \infty. \tag{137}$$

8.5 Convex bounds for sums of random variables

Our ultimate goal is to find approximations for a put option on a sum of dependent random variables and in the previous section we have seen that if a random vector $\vec{X}$ is comonotonic, we can compute the value of a put option using only the marginal distributions (equation (132)). In this section we will show that it is possible to find an upper bound and a lower bound on a put option on a sum of dependent random variables by using the concept of comonotonicity.

8.5.1 The comonotonic upper bound for $\sum_{i=1}^{n} X_i$

To derive an upper bound on the value of an option on a sum of dependent random variables, we first need an upper bound on the sum of random variables itself. This upper bound will be an upper bound in the convex order sense and therefore we will call it a convex bound. The upper bound that we will derive here is attainable in the class of all random vectors with given marginals. It is reached by the comonotonic random vector in this class. So, the upper bound is a supremum in the sense of convex order.

Theorem 8.11 *For any random vector $(X_1, X_2, \ldots, X_n)$ we have*

$$X_1 + X_2 + \ldots + X_n \leq_{cx} X_1^c + X_2^c + \ldots + X_n^c. \tag{138}$$

Proof It is obvious that the means of these two sums are equal, so from definition 8.2 we have that it is sufficient to prove stop-loss order. Hence, we have to prove that

$$\mathbb{E}[(X_1 + X_2 + \ldots + X_n - d)_+] \leq E[(X_1^c + X_2^c + \ldots + X_n^c - d)_+]$$

holds for all d with $d \in \left(F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1)\right)$, since the stop-loss premiums are clearly equal for other values of d.
The following holds for all $(x_1, x_2, \ldots, x_n)$ when $d_1 + d_2 + \ldots + d_n = d$:

$$\begin{aligned}
& (x_1 + x_2 + \ldots + x_n - d)_+ \\
= \;& ((x_1 - d_1) + (x_2 - d_2) + \ldots + (x_n - d_n))_+ \\
\leq \;& ((x_1 - d_1)_+ + (x_2 - d_2)_+ + \ldots + (x_n - d_n)_+)_+ \\
= \;& (x_1 - d_1)_+ + (x_2 - d_2)_+ + \ldots + (x_n - d_n)_+.
\end{aligned}$$

Now replacing constants by the corresponding random variables in the inequality above and taking expectations, we get that

$$\mathbb{E}[(X_1 + X_2 + \ldots + X_n - d)_+] \leq \mathbb{E}[(X_1 - d_1)_+] + \mathbb{E}[(X_2 - d_2)_+] + \ldots + \mathbb{E}[(X_n - d_n)_+] \tag{139}$$

holds for all d and d_i such that $\sum_{i=1}^{n} d_i = d$.
By choosing $d \in \left(F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1)\right)$ and the d_i as in Theorem 8.10, we have by (125)

$$E[(X_1 - d_1)_+] + E[(X_2 - d_2)_+] + \ldots + E[(X_n - d_n)_+] = \mathbb{E}\left[(S^c - d)_+\right].$$

Now $S^c = X_1^c + X_2^c + \ldots + X_n^c$, so combining this with (124) we proved what we had to prove. ∎

Theorem 8.11 states that if we see the random vector $(X_1, \ldots, X_n)$ as a vector of payments, the least attractive one with given marginals F_i, in the sense that the sum of their components is largest in convex order, has the *comonotonic* joint distribution, which means that it has the joint distribution of $\left(F_1^{-1}(U), F_2^{-1}(U), \ldots, F_n^{-1}(U)\right)$. The components of this random vector are maximally dependent, all components being non-decreasing functions of the same random variable.

If we write $S = X_1 + X_2 + \ldots + X_n$ and $S^c = X_1^c + \ldots + X_n^c$, then we have from definition 8.2 and theorem 8.10 for all $d \in \left(F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1)\right)$ that

$$\begin{aligned}
\mathbb{E}\left[(S - d)_+\right] &\leq \mathbb{E}\left[(S^c - d)_+\right] \\
&= \sum_{i=1}^{n} \mathbb{E}[(X_i - F_{X_i}^{-1(\alpha_d)}\left(F_{S^c}(d)\right))_+].
\end{aligned}$$

Also, from (139), we have for all $d \in \left(F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1)\right)$ and for all d_i such that $\sum_{i=1}^n d_i = d$

$$\mathbb{E}\left[(S^c - d)_+\right] \leq \sum_{i=1}^n \mathbb{E}[(X_i^c - d_i)_+].$$

Combining these two inequalities gives that

$$\begin{aligned} \mathbb{E}[(X_1 + X_2 + \ldots + X_n - d)_+] &\leq \sum_{i=1}^n \mathbb{E}[(X_i - F_{X_i}^{-1(\alpha_d)}\left(F_{S^c}\left(d\right)\right))_+] \\ &\leq \sum_{i=1}^n \mathbb{E}[(X_i - d_i)_+] \end{aligned} \tag{140}$$

holds for all $d \in \left(F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1)\right)$ such that $\sum_{i=1}^n d_i = d$. Hence, the smallest upper bound of the form $\sum_{i=1}^n \mathbb{E}[(X_i - d_i)_+]$ with $\sum_{i=1}^n d_i = d$ for the stop-loss premium $\mathbb{E}[(X_1 + X_2 + \ldots + X_n - d)_+]$ is the comonotonic upper bound.

Example Consider a random vector $(\alpha_1 X_1, \alpha_2 X_2, \ldots, \alpha_n X_n)$ of which the α_i are non-zero real numbers and the X_i are lognormally distributed: $\ln(X_i) \sim N\left(\mu_i, \sigma_i^2\right)$. We have that

$$\mathbb{E}\left[X_i\right] = e^{\mu_i + \frac{1}{2}\sigma_i^2}, \tag{141}$$

$$Var\left[X_i\right] = e^{2\mu_i + \sigma_i^2}\left(e^{\sigma_i^2} - 1\right). \tag{142}$$

As an example we can think of the situation where the α_i are deterministic payments at times i, and the X_i are the corresponding lognormally distributed discount factors. Then $(\alpha_1 X_1, \alpha_2 X_2, \ldots, \alpha_n X_n)$ is the vector of the stochastically discounted deterministic payments. From $\Phi^{-1}(1-p) = -\Phi^{-1}(p)$, and Theorem 8.3 ((a) for $\alpha_i > 0$ and (d) for $\alpha_i < 0$) we find that

$$F_{\alpha_i X_i}^{-1}(p) = \alpha_i\ e^{\mu_i + \mathrm{sgn}(\alpha_i)\ \sigma_i \Phi^{-1}(p)}, \qquad 0 < p < 1, \tag{143}$$

where $\mathrm{sgn}\left(\alpha_i\right)$ equals 1 if $\alpha_i > 0$ and -1 if $\alpha_i < 0$. In particular, we find that the product of n comonotonic lognormal random variables is again lognormal:

$$\Pi_{i=1}^n F_{X_i}^{-1}(U) =_d e^{\sum_{i=1}^n \mu_i + \sum_{i=1}^n \sigma_i \Phi^{-1}(U)}. \tag{144}$$

The stop-loss premiums of a lognormal distributed random variable are given by

$$\mathbb{E}[(X_i - d_i)_+] = e^{\mu_i + \frac{\sigma_i^2}{2}}\ \Phi(d_{i,1}) - d_i\ \Phi(d_{i,2}), \qquad d_i > 0. \tag{145}$$

where $d_{i,1}$ and $d_{i,2}$ are determined by

$$d_{i,1} = \frac{\mu_i + \sigma_i^2 - \ln\left(d_i\right)}{\sigma_i}, \quad d_{i,2} = d_{i,1} - \sigma_i. \tag{146}$$

This result can be proved as follows. Differentiating $\mathbb{E}[(X_i - d_i)_+]$ with respect to d_i using equation (97) we find that the derivative is given by $F_{X_i}(d_i) - 1$. Similarly, differentiation of $e^{\mu_i + \frac{\sigma_i^2}{2}}\ \Phi(d_{i,1}) - d_i\ \Phi(d_{i,2})$ with respect to d_i also gives $F_{X_i}(d_i) - 1$ (this requires quite some computation). Also, for $d_i \to \infty$, both sides in (145) go to zero. So both sides end up with the same value and have the same derivative everywhere. But then they should be equal everywhere.
For the lower tails we find

$$\mathbb{E}[(d_i - X_i)_+] = -e^{\mu_i + \frac{\sigma_i^2}{2}}\ \Phi(-d_{i,1}) + d_i\ \Phi(-d_{i,2}), \qquad d_i > 0. \tag{147}$$

As $\mathbb{E}[(\alpha_i(X_i - d_i))_+] = -\alpha_i\ \mathbb{E}[(d_i - X_i)_+]$ if α_i is negative, we find from (145) and (147)

$$\mathbb{E}[(\alpha_i(X_i - d_i))_+] = \alpha_i\ e^{\mu_i + \frac{\sigma_i^2}{2}}\ \Phi(\mathrm{sgn}(\alpha_i)\ \mathrm{d_{i,1}}) - \alpha_i\ \mathrm{d_i}\ \Phi(\mathrm{sgn}(\alpha_i)\ \mathrm{d_{i,2}}), \qquad \mathrm{d_i} > 0, \tag{148}$$

with $d_{i,1}$ and $d_{i,2}$ as defined above.

Now, let $S = \alpha_1 X_1 + \ldots + \alpha_n X_n$, and S^c its comonotonic counterpart: $S^c = F^{-1}_{\alpha_1 X_1}(U) + \ldots + F^{-1}_{\alpha_n X_n}(U)$. Then $S \leq_{cx} S^c$. As the marginal distribution functions are strictly increasing and continuous, we find by (124) that the distribution function $F_{S^c}(x)$ is implicitly defined by $F^{-1}_{S^c}(F_{S^c}(x)) = x$, or equivalently,

$$\sum_{i=1}^{n} \alpha_i \ e^{\mu_i + \mathrm{sgn}(\alpha_i) \ \sigma_i \Phi^{-1}(\mathrm{F_{S^c}(x)})} = x, \qquad F^{-1+}_{S^c}(0) < x < F^{-1}_{S^c}(1). \tag{149}$$

For $F^{-1+}_{S^c}(0) < d < F^{-1}_{S^c}(1)$, the stop-loss premium of S^c at level d follows from (132) and (143):

$$\begin{aligned}
\mathbb{E}[(S^c - d)_+] &= \sum_{i=1}^{n} \mathbb{E}[(\alpha_i X_i - F^{-1}_{\alpha_i X_i}(F_{S^c}(d)))_+] \\
&= \sum_{i=1}^{n} \mathbb{E}[\Big(\alpha_i \left(X_i - e^{\mu_i + \mathrm{sgn}(\alpha_i) \ \sigma_i \Phi^{-1}(\mathrm{F_{S^c}(d)})}\right)\Big)_+].
\end{aligned}$$

Now, from (146) we find that for $d_i = e^{\mu_i + \mathrm{sgn}(\alpha_i) \ \sigma_i \Phi^{-1}(\mathrm{F_{S^c}(d)})}$ we have

$$\begin{aligned}
d_{i,1} &= \frac{\mu_i + \sigma_i^2 - \ln(d_i)}{\sigma_i} \\
&= \frac{\mu_i + \sigma_i^2 - (\mu_i + \mathrm{sgn}(\alpha_i) \ \sigma_i \Phi^{-1}(\mathrm{F_{S^c}(d)}))}{\sigma_i} \\
&= \frac{\sigma_i^2 - \mathrm{sgn}(\alpha_i) \ \sigma_i \Phi^{-1}(\mathrm{F_{S^c}(d)})}{\sigma_i} = \sigma_i - \mathrm{sgn}(\alpha_i) \ \Phi^{-1}(\mathrm{F_{S^c}(d)}),
\end{aligned}$$

and $d_{i,2} = d_{i,1} - \sigma_i = -\mathrm{sgn}(\alpha_i) \ \Phi^{-1}(\mathrm{F_{S^c}(d)})$. Plugging this in (148) and using that for $d_i = e^{\mu_i + \mathrm{sgn}(\alpha_i) \ \sigma_i \Phi^{-1}(\mathrm{F_{S^c}(d)})}$ we have (using (149) for the last equality)

$$\begin{aligned}
\sum_{i=1}^{n} \alpha_i \ d_i \ \Phi(\mathrm{sgn}(\alpha_i) \ \mathrm{d_{i,2}}) &= \sum_{i=1}^{n} \alpha_i \ d_i \ \Phi(-\mathrm{sgn}(\alpha_i) \ \mathrm{sgn}(\alpha_i) \ \Phi^{-1}(\mathrm{F_{S^c}(d)})) \\
&= \sum_{i=1}^{n} \alpha_i \ d_i \ \Phi(- \ \Phi^{-1}(F_{S^c}(d))) \\
&= \sum_{i=1}^{n} \alpha_i \ d_i \ \Phi(\ \Phi^{-1}(1 - F_{S^c}(d))) \\
&= \sum_{i=1}^{n} \alpha_i \ e^{\mu_i + \mathrm{sgn}(\alpha_i) \ \sigma_i \Phi^{-1}(\mathrm{F_{S^c}(d)})} \ (1 - F_{S^c}(d)) \\
&= d(1 - F_{S^c}(d)),
\end{aligned}$$

we find

$$\mathbb{E}[(S^c - d)_+] = \sum_{i=1}^{n} \alpha_i \ e^{\mu_i + \frac{\sigma_i^2}{2}} \Phi\left(\mathrm{sgn}(\alpha_i) \ \sigma_i - \Phi^{-1}(\mathrm{F_{S^c}(d)})\right) - d \ (1 - F_{S^c}(d)). \tag{150}$$

Similarly, the lower tails are given by

$$\mathbb{E}[(d - S^c)_+] = -\sum_{i=1}^{n} \alpha_i \ e^{\mu_i + \frac{\sigma_i^2}{2}} \Phi\left(-\mathrm{sgn}(\alpha_i) \ \sigma_i + \Phi^{-1}(\mathrm{F_{S^c}(d)})\right) + d \ F_{S^c}(d). \tag{151}$$

8.5.2 Improved upper bounds for $\sum_{i=1}^{n} X_i$

If the only information available concerning the multivariate distribution function of the random vector $(X_1, \ldots, X_n)$ consists of the marginal distribution functions of the X_i, then the distribution function of $S^c = F_{X_1}^{-1}(U) + F_2^{-1}(U) + \ldots + F_n^{-1}(U)$ is a prudent choice for approximating the unknown distribution function of $S = X_1 + \ldots + X_n$. It is a supremum in terms of convex order, hence it is the best upper bound that can be derived under the given conditions.
Let us now assume that we have some additional information available concerning the stochastic nature of $(X_1, \ldots, X_n)$. More precisely, we assume that there exists some random variable Λ with a given distribution function, such that we know the conditional cdf's, given $\Lambda = \lambda$, of the random variables X_i, for all possible values of λ. We will show that in this case we can derive improved upper bounds in terms of convex order for S, which are smaller in convex order than the upper bound S^c. Essentially, the idea of this subsection is to determine comonotonic upper bounds for the sum S, conditionally given $\Lambda = \lambda$. Next, we mix the resulting distributions with weights $dF_\Lambda(\lambda)$. By this procedure, convex order is maintained. The upper bound obtained in this way turns out to be sharper than the comonotonic upper bound S^c because it still has the right marginal cdf's for its terms.

In the following theorem, we introduce the notation $F_{X_i|\Lambda}^{-1}(U)$ for the random variable $f_i(U, \Lambda)$, where the function f_i is defined by $f_i(u, \lambda) = F_{X_i|\Lambda=\lambda}^{-1}(u)$.

Theorem 8.12 *Let U be uniform(0,1), and independent of the random variable Λ. Then we have*

$$X_1 + X_2 + \ldots + X_n \leq_{cx} F_{X_1|\Lambda}^{-1}(U) + F_{X_2|\Lambda}^{-1}(U) + \ldots + F_{X_n|\Lambda}^{-1}(U). \tag{152}$$

Proof From Theorem 8.11, we get for any convex function v,

$$\begin{aligned}
\mathbb{E}\left[v\left(X_1 + \ldots + X_n\right)\right] &= \int_{-\infty}^{+\infty} \mathbb{E}\left[v\left(X_1 + \ldots + X_n\right) \mid \Lambda = \lambda\right] \; dF_\Lambda(\lambda) \\
&\leq \int_{-\infty}^{+\infty} \mathbb{E}\left[v\left(f_1(U,\lambda) + \ldots + f_n(U,\lambda)\right)\right] dF_\Lambda(\lambda) \\
&= \mathbb{E}\left[v\left(f_1(U,\Lambda) + \ldots + f_n(U,\Lambda)\right)\right].
\end{aligned}$$

From this the stated result follows, since one of the characterizations of convex order is that the above inequality holds for all convex functions v. ∎

Note that the random vector $\left(F_{X_1|\Lambda}^{-1}(U), F_{X_2|\Lambda}^{-1}(U), \ldots, F_{X_n|\Lambda}^{-1}(U)\right)$ has marginals $F_{X_1}, F_{X_2}, \ldots, F_{X_n}$, because

$$\begin{aligned}
F_{X_i}(x) &= \int_{-\infty}^{\infty} \mathbb{P}\left[X_i \leq x \mid \Lambda = \lambda\right] dF_\Lambda(\lambda) \\
&= \int_{-\infty}^{\infty} \mathbb{P}\left[F_{X_i|\Lambda=\lambda}^{-1}(U) \leq x\right] dF_\Lambda(\lambda) \\
&= \int_{-\infty}^{\infty} \mathbb{P}\left[f_i(U,\lambda) \leq x\right] dF_\Lambda(\lambda) \\
&= \mathbb{P}\left[f_i(U,\Lambda) \leq x\right].
\end{aligned}$$

As we have seen before, for a random vector with given marginal distributions we have that the comonotonic sum is the largest possible sum in the convex order sense. This implies

$$F_{X_1|\Lambda}^{-1}(U) + \ldots + F_{X_n|\Lambda}^{-1}(U) \leq_{cx} F_{X_1}^{-1}(U) + \ldots + F_{X_n}^{-1}(U), \tag{153}$$

which means that the upper bound derived in this subsection is indeed an improved upper bound.
If Λ is independent of all $X_1, X_2, \ldots, X_n$, then we actually do not have any extra information at all and the improved upper bound reduces to the comonotonic upper bound derived in Theorem 8.11.

Let S^u be defined by

$$S^u = F^{-1}_{X_1|\Lambda}(U) + F^{-1}_{X_2|\Lambda}(U) + \ldots + F^{-1}_{X_n|\Lambda}(U). \tag{154}$$

Then we are again interested in the distribution function of S^u and its stop-loss premiums.
In order to obtain the distribution function of S^u, observe that given the event $\Lambda = \lambda$, the random variable S^u is a sum of comonotonic random variables. Hence, from theorem 8.9 we have

$$F^{-1}_{S^u|\Lambda=\lambda}(p) = \sum_{i=1}^{n} F^{-1}_{X_i|\Lambda=\lambda}(p), \quad p \in (0,1). \tag{155}$$

If the marginal cdf's $F_{X_i|\Lambda=\lambda}$ are strictly increasing and continuous, then $F_{S^u|\Lambda=\lambda}(x)$ follows by solving

$$\sum_{i=1}^{n} F^{-1}_{X_i \mid \Lambda=\lambda}\left(F_{S^u \mid \Lambda=\lambda}(x)\right) = x, \qquad F^{-1+}_{S^u \mid \Lambda=\lambda}(0) < x < F^{-1}_{S^u \mid \Lambda=\lambda}(1), \tag{156}$$

see (124). In this case, we also find from (131) that for any $d \in \left(F^{-1+}_{S^u|\Lambda=\lambda}(0), F^{-1}_{S^u|\Lambda=\lambda}(1)\right)$:

$$\mathbb{E}\left[(S^u - d)_+ \mid \Lambda = \lambda\right] = \sum_{i=1}^{n} E\left[\left(X_i - F^{-1}_{X_i|\Lambda=\lambda}\left(F_{S^u|\Lambda=\lambda}(d)\right)\right)_+ \mid \Lambda = \lambda\right], \tag{157}$$

from which the stop-loss premium at level d of S^u can be determined.

8.5.3 Lower bounds for $\sum_{i=1}^{n} X_i$

Let $\vec{X} = (X_1, \ldots, X_n)$ be a random vector with given marginal cdf's $F_{X_1}, F_{X_2}, \ldots, F_{X_n}$. As in the previous subsection, we assume that there exists some random variable Λ with a given distribution function, such that we know the conditional cdf's, given $\Lambda = \lambda$, of the random variables X_i , for all possible values of λ. We will show how to obtain a lower bound, in the sense of convex order, for $S = X_1 + X_2 + \ldots + X_n$ by conditioning on this random variable. If we view S as a payment we have to make, this means that we are considering a more attractive random variable to pay than S. This will help to give an idea of the degree of overestimation of the risk involved by replacing S by the less attractive random variables S^u or S^c.
The idea of this subsection is to observe that the expectation of a random variable is always smaller than or equal in convex order than the random variable itself, and also that convex order is maintained under mixing.

Theorem 8.13 *For any random vector $\vec{X}$ and any random variable Λ, we have*

$$\mathbb{E}[X_1 \mid \Lambda] + \mathbb{E}[X_2 \mid \Lambda] + \ldots + \mathbb{E}[X_n \mid \Lambda] \leq_{cx} X_1 + X_2 + \ldots + X_n. \tag{158}$$

Proof By Jensen's inequality, we find that for any convex function v, the following inequality holds:

$$\begin{aligned}
\mathbb{E}\left[v\left(X_1 + X_2 + \ldots + X_n\right)\right] &= \mathbb{E}_\Lambda\ \mathbb{E}\left[v\left(X_1 + X_2 + \ldots + X_n\right) \mid \Lambda\right] \\
&\geq \mathbb{E}_\Lambda\ \left[v\left(\mathbb{E}\left[X_1 + X_2 + \ldots + X_n \mid \Lambda\right]\right)\right] \\
&= \mathbb{E}_\Lambda\ \left[v\left(\mathbb{E}\left[X_1 \mid \Lambda\right] + \ldots + \mathbb{E}\left[X_n \mid \Lambda\right]\right)\right].
\end{aligned}$$

From this the stated result follows, since one of the characterizations of convex order is that the above inequality holds for all convex functions v. ∎

Let S be defined as above, and let S^l be defined by

$$S^l = \mathbb{E}[S \mid \Lambda] \tag{159}$$

Note that if Λ and S are mutually independent, we find the trivial result

$$\mathbb{E}[S] \leq_{cx} S. \tag{160}$$

On the other hand, if Λ and S have a one-to-one relation (i.e. Λ completely determines S), the lower bound coincides with S. Note further that $\mathbb{E}\left[\mathbb{E}\left[X_i \mid \Lambda\right]\right] = \mathbb{E}\left[X_i\right]$ always holds, but $Var\left[\mathbb{E}\left[X_i \mid \Lambda\right]\right] < Var\left[X_i\right]$ unless $\mathbb{E}\left[Var\left[X_i \mid \Lambda\right]\right] = 0$. This follows from the identity

$$Var(X_i) = \mathbb{E}\left[Var\left[X_i \mid \Lambda\right]\right] + Var\left[\mathbb{E}\left[X_i \mid \Lambda\right]\right]$$

and this means that X_i, given $\Lambda = \lambda$, is degenerate for each λ. This implies that the random vector $\left(\mathbb{E}\left[X_1 \mid \Lambda\right], \mathbb{E}\left[X_2 \mid \Lambda\right], \ldots, \mathbb{E}\left[X_n \mid \Lambda\right]\right)$ will in general not have the same marginal distribution functions as X. But if we can find a conditioning random variable Λ with the property that all random variables $\mathbb{E}\left[X_i \mid \Lambda\right]$ are non-increasing functions of Λ (or all are non-decreasing functions of Λ), the lower bound S^l is a sum of n comonotonic random variables. The cdf of this sum can then be obtained by previous results.
To judge the quality of the stochastic lower bound $\mathbb{E}[S \mid \Lambda]$, we might look at its variance. To maximize it, i.e. to make it as close as possible to $Var\left[S\right]$, the average value of $Var[S \mid \Lambda = \lambda]$ should be minimized. In other words, to get the best lower bound, Λ and S should be as alike as possible.
Let's further assume that the random variable Λ is such that all $g_i\left(\lambda\right) \equiv \mathbb{E}\left[X_i \mid \Lambda = \lambda\right]$ are non-decreasing and continuous functions of λ. The quantiles of the lower bound S^l then follow from theorem 8.9 combined with an application of theorem 8.3(a):

$$\begin{aligned} F_{S^l}^{-1}(p) &= \sum_{i=1}^{n} F_{\mathbb{E}[X_i|\Lambda]}^{-1}(p) = \sum_{i=1}^{n} F_{g_i(\Lambda)}^{-1}(p) \\ &= \sum_{i=1}^{n} \mathbb{E}\left[X_i \mid \Lambda = F_\Lambda^{-1}(p)\right], \quad p \in (0,1). \end{aligned} \tag{161}$$

If we now additionally assume that the cdf's of the random variables $\mathbb{E}\left[X_i \mid \Lambda\right]$ are strictly increasing and continuous, then the cdf of S^l is also strictly increasing and continuous, and from (28) we get for all $x \in \left(F_{\mathbb{E}[S|\Lambda]}^{-1+}(0),\ F_{\mathbb{E}[S|\Lambda]}^{-1}(1)\right)$,

$$\sum_{i=1}^{n} F_{\mathbb{E}[X_i|\Lambda]}^{-1}\left(F_{S^l}(x)\right) = x,$$

or equivalently,

$$\sum_{i=1}^{n} \mathbb{E}\left[X_i \mid \Lambda = F_\Lambda^{-1}\left(F_{S^l}(x)\right)\right] = x, \tag{162}$$

which unambiguously determines the cdf of the convex order lower bound $S^l = \mathbb{E}\left[S \mid \Lambda\right]$ for S.

Under the same assumptions, the stop-loss premiums of S^l can be determined from (131):

$$\mathbb{E}\left[\left(S^l - d\right)_+\right] = \sum_{i=1}^{n} \mathbb{E}\left[\left(\mathbb{E}\left[X_i \mid \Lambda\right] - \mathbb{E}\left[X_i \mid \Lambda = F_\Lambda^{-1}\left(F_{S^l}(d)\right)\right]\right)_+\right], \tag{163}$$

which holds for all $d \in \left(F_{S^l}^{-1+}(0),\ F_{S^l}^{-1}(1)\right)$.

References

[1] Dhaene, J.; Denuit, M.; Goovaerts, M.J.; Kaas, R.; Vyncke, D. (2002a). "The concept of comonotonicity in actuarial science and finance: Theory.", *Insurance: mathematics and economics*, 31(1), 3-33.

[2] Dhaene, J.; Denuit, M.; Goovaerts, M.J.; Kaas, R.; Vyncke, D. (2002b). "The concept of comonotonicity in actuarial science and finance: Applications.", *Insurance: mathematics and economics*, 31(2), 133-161.

[3] Geman, H.; El Karoui, N.; Rochet, J.C. (1995). "Changes of Numeraire, changes of probability measure and option pricing.", *Journal of applied probability*, 32, 443-458.

[4] Harrison, J.M.; Kreps, D. (1979). "Martingales and arbitrage in multiperiod securities markets.", *Journal of economic theory*, 20, 381-408.

[5] Harrison, J.M.; Pliska, S. (1981). "Martingales and stochastic integrals in the theory of continuous trading.", *Stochastic processes and their applications*, 11, 215-260.

[6] Jamshidian, F. (1989). "An exact bond option formula.", *Journal of finance*, 44, 205-209.

[7] Kaas, R., Van Heerwaarden, A.E., Goovaerts, M.J. (1994). "Ordering of actuarial risks", Institute for Actuarial Science and Econometrics, University of Amsterdam, Amsterdam.

[8] Karatzas, I.; Shreve, S.E. (2000). "Brownian motion and stochastic calculus", Springer, second edition, pp. 191.

[9] Neftci, S. (2004). "Principles of financial engineering", Academic Press, pp. 133.

[10] Nielsen, J.A.; Sandmann, K. (2002). "Pricing of Asian exchange options under stochastic interest rates as a sum of options.", *Finance and stochastics*, 6, 355-370.

[11] Pelsser, A.A.J. (2003). "Mathematical foundation of convexity correction.", *Quantitative finance*, 3, 59-65.

[12] Plat, R.; Pelsser, A.A.J. (2009). "Analytical approximations for prices of swap rate dependent embedded options in insurance products", *Insurance: mathematics and economics*, 44, 124-134.

[13] Schrager, D.F.; Pelsser, A.A.J. (2004). "Pricing Rate of Return Guarantees in Regular Premium Unit Linked Insurance.", *Insurance: mathematics and economics*, 35, 369-398.

[14] Schrager, D.F.; Pelsser, A.A.J. (2006). "Pricing swaptions and coupon bond options in affine term structure models.", *Mathematical finance*, 16, 673-694.

[15] Shaked, M.; Shanthikumar, J.G. (1994). "Stochastic orders and their applications", Academic Press, pp. 545.