Faculteit Wetenschappen
Vakgroep Toegepaste Wiskunde en Informatica

Quadratic hedging in finance and insurance

Nele Vandaele

Promotor: Prof. dr. Michèle Vanmaele

Proefschrift tot het bekomen van de graad van
Doctor in de Wetenschappen: Wiskunde
Academiejaar 2009-2010

Dit werk kwam tot stand met een beurs van
het Fonds voor Wetenschappelijk Onderzoek Vlaanderen.

Contents

List of abbreviations

FS	Föllmer-Schweizer
GKW	Galtchouk-Kunita-Watanabe
HJM	Heath-Jarrow-Morton
LRM	locally risk-minimizing
MMM	minimal martingale measure
MVT	mean-variance tradeoff
NIG	normal-inverse Gaussion distribution
PII	process with independent increments
PIIS	process with stationary and independent increments
PIIAC	process with independent increments and absolutely continuous characteristics
QLC	quasi-left continuous
RM	risk-minimizing
SC	structure condition
VG	variance-gamma
VOMM	variance-optimal martingale measure

If C is a class of processes, then

C_0	the set of processes X with $X_0 = 0$
$C_{\mathbf{loc}}$	the localized class
$C_{0,\mathbf{loc}}$	the intersection of C_0 and C_{loc}

The following classes of processes contains

$\mathcal{A}$	processes with integrable variation (p. 10)
$\mathcal{A}^+$	integrable increasing processes (p. 10)
$\mathcal{H}^2$	square-integrable martingales (p. 9)
$\mathcal{L}$	local martingales starting at zero (p. 9)
$\mathcal{M}$	uniformly integrable martingales (p. 9)
$\mathcal{M}(\mathcal{E})$	$\mathcal{E}(N)$-martingales (p. 39)
$\mathcal{S}$	semimartingales (p. 10)
S_p	special semimartingales (p. 10)
$\mathcal{V}$	adapted processes with finite variation (p. 10)
$\mathcal{V}^+$	adapted, increasing processes (p. 10)

If we want to denote the measure explicitly, then we use e.g. $\mathcal{A}(P)$.

For the following symbols we refer to the page where the symbol is introduced.

Test all things; hold fast what is good.

The first epistle of Paul to the Thessalonians 5,21

1 Introduction

The last decennium complicated financial products like CDO's (Collateralized debt obligation) became very popular around the world. Tempted by the possible high profits, the products were bought without a good estimation of the risk involved or under the assumption that the products did not contain any risk at all. The recent financial crisis showed once more that these products do contain a lot of risk and that it is not sufficient to only determine a good buying/selling price of the product. After some time a product can become worthless or a very risky product. Hence it is recommended to follow up the total portfolio and use hedging strategies to reduce the risk if necessary. So besides pricing, it is also important to find good hedging strategies which minimize the risk as much as possible.

Maybe in the past, hedging strategies were not applied often enough, but one should also watch out not to use hedging strategies unnecessarily. Therefore it is always important to look if any risk is reduced by using a hedging strategy. If this is not the case, then there is no sense in applying a hedging strategy and one should only take in mind the possible risk related with the portfolio.

Of course there are various ways to hedge a portfolio, depending on the way the risk is measured. We give a short overview of the different types of (dynamic) hedging strategies that have been described in literature:

- **Delta-hedging**
  The delta at time t of an option or more general a derivative is the partial derivative with respect to the underlying of that option at time t. The delta is one of the so-called Greeks. These Greeks represent the sensitivity of a derivative with respect to various variables, such as the price of the underlying, the volatility, the time to maturity. For more information concerning the Greeks we refer to e.g. Shreve (2004). The delta-hedge of an option at each time goes short in as many underlyings as the delta Greek. The portfolio containing the option and which goes short in delta underlyings is hence delta-neutral, which means that it is invariant for small moves in the price of the underlying.
  This strategy is still very popular in practice due to its simplicity.
- **Superhedging**
  A strategy is a superhedging strategy if it is a self-financing strategy such that at maturity the value of the strategy is surely greater than the value of the derivative. The cost of superhedging is defined as the cost of the cheapest superhedging strategy. In case of jumps, the superhedging strategy is often too expensive or boils down to buy and hold strategies. For more details we refer to Section 10.2 of Cont and Tankov (2004).
- **Utility hedging**
  For every utility function U, which should be concave and increasing, we can determine the related utility hedge by maximizing the function:
  $$E[U(Z)]$$
  over different payoffs Z. Risk is therefore measured by expected utility. The function U should ideally not be symmetric and it is also possible to attain a sort of weights according to the size of the possible loss/gain.
  Popular choices are e.g. logarithmic utility: $U(x) = \ln(\alpha x)$ or exponential utility: $U^\alpha(x) = 1 - \exp(-\alpha x)$, with $\alpha > 0$.
  In the general case there are rarely explicit computations available and therefore in this thesis we will only consider a subclass consisting of quadratic hedging strategies.
- **Quadratic hedging**
  Quadratic hedging is a specific form of utility hedging, where the strategy minimizes the hedging error in mean square sense. Hence risk is in this case quantified as variance. One of the obvious drawbacks of quadratic hedging is that losses and gains are treated in the same way. On the other hand, this might be an advantage, in case you do not know whether you

deal with a buyer or a seller. Another advantage is that quadratic strategies related to different options can simply be added up as is also the case for delta-hedging strategies. In other words, quadratic hedging is a sort of linear hedging strategy. In this context we also refer to Kramkov and Sirbu (2007). They show that for a small number of contingent claims the linear approximation of utility-based hedging strategies is in fact a mean-variance hedging strategy under an appropriate numéraire and under the risk-neutral probability measure.
Mean-variance hedging strategy (also called variance-optimal hedging) is one of the two main quadratic hedging strategies we will discuss. The other one is the (locally) risk-minimizing hedging strategy.
In the mean-variance hedging theory the goal is to minimize the difference between the claim H at maturity T and the portfolio at that time, using a self-financing strategy. In the risk-minimizing hedging strategy, the goal is to minimize the variance of the cost process at any time t subject to the condition that the value of the portfolio at time T equals the claim H. In the latter case, it is only possible to find a self-financing portfolio when the claim is attainable[1]. The risk-minimizing hedging strategy only makes sense when the underlying is a martingale, the extension to semimartingales is called locally risk-minimization.

1.1 Outline of the thesis

This outline contains a brief summary of the content of this thesis. A thorough summary including a motivation of the studied subjects and references to related works can be found at the start of each chapter.
We focus mainly on quadratic hedging and especially on the locally risk-minimizing hedging strategy and related to it, the Föllmer-Schweizer decomposition. We do not only look at the pure financial market, but we also determine hedging strategies for the insurance market (Chapters 6-7), as well as for the interest rate derivatives market (Chapter 8) and the commodity market (Chapter 9). We work in continuous time and assume that there are no trading costs.
Concerning the processes describing the dynamics of the underlyings we use in Chapter 3 a general not necessarily quasi-left continuous semimartingale,

[1] A claim is attainable if it can be replicated with the available products on the market.

while in Chapter 4 we work with a quasi-left continuous semimartingale. Chapters 6 until 9 focus more on applications where the process of the underlying is either driven by a (geometric) Brownian motion or by a (geometric) (time-inhomogeneous) Lévy process.
The Brownian motion is unfortunately not capable to generate big jumps in the process, hence big jumps occurring in reality are not covered. This can be solved by allowing discontinuous processes.

After introducing some basic concepts concerning stochastic calculus in **Chapter 2**, we discuss in **Chapter 3** the relationship between the Föllmer-Schweizer decomposition under the original measure and the Galtchouk-Kunita-Watanabe decomposition under the minimal martingale measure. It is generally known that they coincide in case the underlying is continuous, but we prove with an explicit example that they are not always equal in the discontinuous case. Furthermore, we also provide a more explicit form of the Föllmer-Schweizer decomposition by using the predictable characteristics. This chapter is based on Choulli et al. (2010).
The use of the Föllmer-Schweizer decomposition for quadratic hedging becomes more clear when we explain the theory of the locally risk-minimizing (**Chapter 4**) and the mean-variance hedging strategy (**Chapter 5**). For both strategies we provide an overview containing the theoretical results as well as the applications. For the locally risk-minimizing hedging strategy we additionally give the extension to the multidimensional case in Chapter 4. The chapter concerning the locally risk-minimizing hedging strategy is a revised form of Vandaele and Vanmaele (2008a).
For the applications of the quadratic hedging strategies, we started with the determination of the risk-minimizing hedging strategy for unit-linked life insurance contracts with a surrender option. A unit-linked life insurance contract can be seen as a combination of an insurance and a mutual fund. The premium of the contract is invested in a number of *units* of the fund. Therefore in unit-linked life insurance contracts the benefits and sometimes also the premiums are random and they depend on the development of the mutual fund which is a prespecified reference asset or portfolio. The surrender option gives the opportunity to exit the contract before maturity against a pre-specified value. We assume that the surrender time is not a stopping time in the filtration generated by the financial market. This means that we make the realistic assumption that a policyholder does not only quit a contract due to the evolutions of the reference portfolio, but that often he/she has personal reasons to surrender before maturity. The results of **Chapter 6** are published in Vandaele and Vanmaele

(2009).
A second application given in **Chapter 7** is the determination of the locally risk-minimizing hedging strategy for unit-linked life insurance contracts when the underlying risky asset is driven by a Lévy process. Due to the discontinuity of the underlying, it is no longer possible to obtain the strategy by using the Galtchouk-Kunita-Watanabe decomposition under the minimal martingale measure and therefore we determine the Föllmer-Schweizer decomposition indirectly as is done in Vandaele and Vanmaele (2008b) or by using the more explicit form as described in Choulli et al. (2010). We do not assume a surrender option in this chapter and therefore we have stochastic independence between the financial market and the insurance market.
In the last two chapters of this dissertation we apply the quadratic hedging strategies in two specific frameworks. In these chapters we also discuss the implementation of the obtained formulas in such a way that we can compare the total costs related to different hedging strategies. In order to speed up the calculations we apply Fourier transformations to express the optimal numbers in terms of the characteristic function. This implementation is based on an extension of some intermediate results given in Hubalek et al. (2006).
In **Chapter 8** we determine and compare for a forward swaption the delta-hedge with the mean-variance hedge under the forward martingale measure linked with the maturity of the swaption. To model the interest rate derivatives market, we assume a Lévy extended Heath-Jarrow-Morton framework, where the driving process belongs to the class of normal inverse Gaussian processes. We first determine the price of the swaption and the delta-neutral hedge when one zero-coupon bond is used for hedging. In order to compare the hedge with the mean-variance hedging strategy, which uses two bonds, we also give the self-financing delta-hedge and the delta- and gamma-neutral hedge.
The mean-variance hedging strategy is always defined in terms of discounted assets, but because we cannot assume we have a risk-free interest rate we use as numéraire the zero-coupon bond with the same maturity as the swaption. Hence all discounted bonds are martingales under the forward measure and the mean-variance hedge follows from the Galtchouk-Kunita-Watanabe decomposition. The advantage of having a self-financing portfolio, as is the case in the mean-variance hedge, is that the optimal number of discounted assets equals the non-discounted number. The results of this chapter can also be found in Glau et al. (2010a, 2010b).
In **Chapter 9**, based on Leoni et al. (2010), we assume that options depend on several assets, while we can only invest in a weighted combination of these assets. This setting is inspired by actual problems faced by energy traders. They

mostly hedge by applying a volume-neutral or weight-adjusted delta-hedge. We compare these adjusted delta-hedges with the (adjusted) locally risk-minimizing hedging strategy. The simulations are restricted to two assets, but can easily be extended to more. As driving processes we use Brownian motions and a multivariate variance gamma process, both in a martingale as well as a semimartingale setting.
The presented simulations give us a good idea of the usefulness of the delta-hedge and the quadratic hedging strategies in the two settings we discuss.
At the end of this thesis we come to a general conclusion and discuss potential ideas for future research.

In mathematics you don't understand things. You just get used to them.

Johann Von Neumann
(1903-1957)

2 Basic concepts

2.1 Introduction

This chapter is intended for readers with a mathematical background who are not specialists in the field of stochastic processes. We restrict ourselves to concepts needed later on, we refer the interested reader to the books of e.g. Jacod (1979), Jacod and Shiryaev (2002), Protter (2005). We start with introducing some standard notations which will be used throughout the whole thesis. Most of the notations and the theorems in this chapter are based on the book of Jacod and Shiryaev (2002), unless it is stated otherwise.
A new contribution is Lemma 2.2.24 concerning the uniqueness of the representation theorem. Furthermore we characterize the class of equivalent measures which have the same minimal martingale measure as the original measure.
We assume as given a probability space $(\Omega, \mathcal{F}, P)$ and in addition a family of σ-algebras $(\mathcal{F}_t)_{0\leq t\leq\infty}$ which is increasing ($\mathcal{F}_s \subset \mathcal{F}_t$ if $s \leq t$) and where $\mathcal{F}_\infty$ is by convention equal to $\mathcal{F}$. We will call this family a **filtration** and denote it by $\mathbb{F}$. Often we will use the natural filtration linked with the process X and we mostly work on a finite time horizon $T \in [0, \infty)$. Furthermore we assume the filtration is right-continuous, i.e. $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$. Hence we have the following filtered

probability space also called the **stochastic basis** $(\Omega, \mathcal{F}, \mathbb{F}, P)$. We assume that the stochastic basis is complete as is frequently done in literature. This means that the σ-field $\mathcal{F}$ is P-complete and that every $\mathcal{F}_t$ contains all P-null sets of $\mathcal{F}$. Furthermore the assumption that $\mathcal{F}_0$ is trivial is frequently used.

Definition 2.1.1. A real-valued random variable X belongs to the set $L^p(\Omega, \mathcal{F}, P)$, for $p \in [1, \infty)$ if $|X|^p$ is **integrable**. Hence $E[|X|^p] < \infty$.

A process is a family $X = (X_t)_{t \in \mathbb{R}^+}$ of mappings from Ω into $\mathbb{R}^d$, $d \in \mathbb{N}^*$. This process can be 'continu à droite, limite à gauche' with short-hand notation **càdlàg** if all paths are right-continuous (i.e. for almost every $\omega \in \Omega$ the map $t \mapsto X_t(\omega)$ is right-continuous: $\lim_{s \to t, s > t} X_s = X_t$) and have left-hand limits ($\exists \lim_{s \to t, s < t} X_s =: X_{s-}$). Analogously the terms **càglàd** (i.e. left-continuous and right-hand limits), **càd** (i.e. right-continuous),... are used.
For a càdlàg process we define the jump at time t as $\Delta X_t := X_t - X_{t-}$.

Definition 2.1.2. A process X is called **adapted** if X_t is $\mathcal{F}_t$-measurable for all t.

Definition 2.1.3. A **stopping time** is a mapping $T : \Omega \to [0, \infty]$ such that $\{T \leq t\} \in \mathcal{F}_t$ for all $t \in \mathbb{R}^+$.

Hence T is a stopping time if it is possible to decide whether or not $\{T \leq t\}$ has occurred on the basis of the knowledge of $\mathcal{F}_t$. The process X^T is called the **stopped process** at time T, defined in the following way:

$$X_t^T = X_{T \wedge t}. \tag{2.1}$$

Definition 2.1.4. If S, T are two stopping times, then we can define four kinds of stochastic intervals: $[\![S, T]\!]$, $[\![S, T[\![$, $]\!]S, T]\!]$ and $]\!]S, T[\![$, where e.g.:

$$[\![S, T[\![= \{(\omega, t) : t \in [0, \infty), S(\omega) \leq t < T(w)\}.$$

On the set $\Omega \times [0, T]$, we define two σ-fields $\mathcal{O}$ and $\mathcal{P}$ generated by the adapted and càdlàg processes and the adapted and continuous processes respectively. On the set $\widetilde{\Omega} := \Omega \times [0, T] \times \mathbb{R}^d$, we consider the σ-field $\widetilde{\mathcal{O}} = \mathcal{O} \otimes \mathcal{B}(\mathbb{R}^d)$ (resp. $\widetilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$), where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ-field for $\mathbb{R}^d$.

Definition 2.1.5. A process or a random set is called a **optional** (resp. **predictable**) process (resp. random set) if the process (resp. random set) is $\mathcal{O}$ (resp. $\mathcal{P}$)-measurable.

From now on if the process g is e.g. $\widetilde{\mathcal{O}}$-measurable, we will denote this by $g \in \widetilde{\mathcal{O}}$.

Definition 2.1.6. A function W on $\widetilde{\Omega}$ that is $\widetilde{\mathcal{O}}$ (resp. $\widetilde{\mathcal{P}}$)-measurable is called an optional (resp. predictable) function.

Definition 2.1.7. A **predictable time** is a mapping $T : \Omega \to [0, \infty]$ such that the stochastic interval $[\![0, T[\![$ is predictable.

Definition 2.1.8. A càdlàg process X is **quasi-left-continuous** (QLC) if $\Delta X_T = 0$ a.s. on the set $\{T < \infty\}$ for every predictable time T.

2.2 (Semi)martingale

Definition 2.2.1. A **martingale** is an adapted process M on the basis $(\Omega, \mathcal{F}, \mathbb{F}, P)$ whose P-almost all paths are càdlàg, such that every M_t is integrable, and that for $s \leq t$:

$$M_s = E(M_t|\mathcal{F}_s).$$

A martingale M is called a uniformly integrable martingale if the family of random variables $(M_t)_{t \in \mathbb{R}^+}$ is uniformly integrable. $\mathcal{M}$ stands for the class of all uniformly integrable martingales. The class of all square-integrable martingales is denoted by $\mathcal{H}^2$ and contains all the martingales M such that $\sup_{t \in \mathbb{R}^+} E(M_t^2)$ $< \infty$.

Definition 2.2.2. A process M is a **local martingale** if there exists an increasing sequence (T_n) of stopping times such that $\lim_{(n)} T_n = \infty$ a.s. and such that each stopped process M^{T_n} is a martingale.

The class of local martingales is denoted by $\mathcal{M}_{\text{loc}}$ whereas $\mathcal{H}^2_{\text{loc}}$ stands for the class of locally square-integrable martingales. The class $\mathcal{L}$ contains all local martingales starting at zero.
If we denote by X^* the process $\sup_{s \leq \cdot} |X_s|$, then we have the following sufficient condition such that X is a martingale:

Theorem 2.2.3 (Protter (2005) Theorem I.51)**.** *Let X be a local martingale such that $E[X_t^*] < \infty$ for every $t \geq 0$. Then X is a martingale.*

The set of all real-valued processes A with $A_0 = 0$ that are càdlàg, adapted and for which each path $t \to A_t(\omega)$ has a finite variation over each finite interval $[0, t]$ is referred to as $\mathcal{V}$. The variation of A is given by $\int |dA_s|$. The set of processes which are non-decreasing instead of having finite variation is indicated by $\mathcal{V}^+$. The processes from this class which are also integrable (i.e. $E(A_\infty) < \infty$) are collected in $\mathcal{A}^+$. Analogously the subset of processes from $\mathcal{V}$ that have **integrable variation**: $E(\text{Var}(A)_\infty) < \infty$ is denoted by $\mathcal{A}$.
The concept of localizing used in the definition of local martingale can also be used for other class of processes e.g. $\mathcal{A}^+_{\text{loc}}$.

Definition 2.2.4. A **semimartingale** X is a process of the form

$$X = X_0 + M + B \tag{2.2}$$

with X_0 finite-valued and $\mathcal{F}_0$-measurable, $M \in \mathcal{L}$ and $B \in \mathcal{V}$.

The set of all semimartingales is denoted by $\mathcal{S}$.

Definition 2.2.5. A **special semimartingale** X is a semimartingale which admits a decomposition $X = X_0 + M + B$ with B predictable.

The space of all special semimartingales is denoted by $\mathcal{S}_p$. We remark that

Definition 2.2.6. If a semimartingale is special then the decomposition $X = X_0 + M + B$ with B predictable is unique and this decomposition is called the **canonical decomposition** of X.

Definition 2.2.7. For processes $X \in \mathcal{A}_{\text{loc}}$ we can define the unique (up to an evanescent set) process X^p, called the **compensator** under P, which is the predictable process in $\mathcal{A}_{\text{loc}}$ such that $X - X^p$ is a P-local martingale.

Hence the P-compensator of a semimartingale X will exist if it is P-locally integrable in the following sense:

Definition 2.2.8. A semimartingale X is P**-locally integrable** if X is a special semimartingale under P. From Jacod (1979) (2.14) we know that this is equivalent with the non-decreasing process $X^* = \sup_{s\leq \cdot} |X_s|$ belonging to $\mathcal{A}^+_{\text{loc}}$.

Before we introduce the notation for a stochastic integral we define the notion of random measure.

Definition 2.2.9. A **random measure** on $\mathbb{R}^+ \times \mathbb{R}^d$ is a family $\mu = (\mu(\omega; dt, dx : \omega \in \Omega))$ of non-negative measures on the Blackwell space $(\mathbb{R}^+ \times \mathbb{R}^d, \mathcal{R}^+ \otimes \mathcal{E})$ satisfying $\mu(\omega; \{0\} \times \mathbb{R}^d) = 0$.

Definition 2.2.10. By $\phi \cdot X$ we denote the **stochastic integral** of ϕ with respect to the process X. If X is d-dimensional then

$$\phi \cdot X = \sum_{i=1}^{n} \phi^{(i)} \cdot X^{(i)} = \int \phi' dX,$$

where by ϕ' we denote the **transpose** of ϕ.

If we calculate the integral over a random measure μ for an optional function W then we use the notation $\star$, e.g.

$$(W \star \mu)(\omega, t) = \int_0^t \int_{\mathbb{R}^d} W(\omega, s, x) \mu(\omega; ds, dx)$$

if $\int_0^t \int_{\mathbb{R}^d} |W(\omega, s, x)| \mu(\omega; ds, dx)$ is finite and $+\infty$ otherwise.
To be able to define the structure condition, we first need to introduce the (predictable) quadratic covariation of two semimartingales.

Definition 2.2.11. The **quadratic covariation** of two semimartingales X and Y is defined as

$$[X, Y] = XY - X_0 Y_0 - X_- \cdot Y - Y_- \cdot X. \tag{2.3}$$

Definition 2.2.12. The **predictable quadratic covariation** of two semimartingales X, Y is the compensator of the quadratic covariation $[X, Y]$. It is denoted by $\langle X, Y \rangle$ and therefore also called the angle brackets of X and Y. The short hand notation $\langle X \rangle$ will be used for the angle bracket $\langle X, X \rangle$.

We remark that this definition is an extension of the one given by Jacod and Shiryaev (2002). They only define the predictable quadratic covariation for locally square-integrable martingales as the compensator of XY, which equals the compensator of $[X, Y]$ in the martingale case. Furthermore the predictable quadratic covariation only exists if the compensator (see Definition 2.2.7) exists. Hence only if the quadratic covariation $[X, Y]$ is locally integrable as defined in

Definition 2.2.8. It is obvious that the predictable quadratic covariation is measure dependent, while the quadratic covariation is independent of the measure we work with. If we do not denote the measure in the notation of the predictable quadratic covariation, then we assume it is under the original measure P, in the other case we use the notation e.g. $\langle \cdot, \cdot \rangle^Q$.
In Jacod and Shiryaev (2002) the following theorem is described for a square-integrable martingale M:

Theorem 2.2.13. *If M is a square-integrable martingale then $\langle M, M \rangle$ is non-decreasing and it admits a continuous version if and only if M is quasi-left-continuous.*

For a semimartingale X we will often use the notations X^c (X^d) to denote the continuous (resp. discontinuous) local martingale part.
The proof of the following theorem can be found in Jacod and Shiryaev (2002) I.4.52:

Theorem 2.2.14. *If X, Y are semimartingales, then*

$$[X, Y]_t = \langle X^c, Y^c \rangle_t + \sum_{s \leq t} \Delta X_s \Delta Y_s. \tag{2.4}$$

From this theorem it follows that for continuous martingales X and Y the predictable quadratic covariation equals the quadratic covariation:

$$[X, Y]_t = \langle X, Y \rangle_t,$$

because the second term equals zero, $X = X^c$ and $Y = Y^c$.
The following properties will be very useful in later calculations:

Properties 2.2.15. $X \in \mathcal{S}$.

(1) If $Y \in \mathcal{V}$, then $[X, Y] = \Delta X \cdot Y$.

(2) If $Y \in \mathcal{V}$ and predictable, then $[X, Y] = \Delta Y \cdot X$. Hence if X is a local martingale, then also $[X, Y]$ is a local martingale.

(3) If $Y \in \mathcal{V}$ and X or Y is continuous, then $[X, Y] = 0$.

(4) If X and Y are local martingales, then $[X, Y] = 0$ whenever X is continuous and Y purely discontinuous.

(5) If Y is also a semimartingale, then $\Delta[X,Y] = \Delta X \Delta Y$.

An important concept is the orthogonality of two semimartingales.

Definition 2.2.16. Two P-semimartingales X and Y are called **orthogonal** under a measure P if $[X,Y]$ is a local martingale under P. Hence the angle bracket $\langle X, Y\rangle = 0$.

Assuming no-arbitrage (for sufficient conditions see Delbaen and Schachermayer (1994)) guarantees the structure condition (SC):

Definition 2.2.17. Assume the following canonical decomposition $X_0 + M + B$ for the semimartingale X. Then we say that the **structure condition** is satisfied if there exists a predictable process λ satisfying

$$dB_t = d\langle M\rangle_t \lambda_t \quad \text{and} \quad \int_0^T \lambda_u' d\langle M\rangle_u \lambda_u < +\infty \quad P\text{-a.s.}. \tag{2.5}$$

If the structure condition is satisfied, then we can define the mean-variance tradeoff process:

Definition 2.2.18. The **mean-variance tradeoff process** (MVT) K is defined as the increasing predictable process with:

$$K_t = \int_0^t \lambda_s' dB_s = \langle \int \lambda dM\rangle_t.$$

The predictable characteristics are an important concept to describe a semimartingale. They can be seen as an extension of the Lévy characteristics, namely the drift, the variance of the Gaussian part and the Lévy measure used to characterize the distribution of a process with independent increments.
Assume we have a d-dimensional semimartingale X with decomposition $X = X_0 + M + B$. From Jacod and Shiryaev (2002) Theorem I.4.18 we know that the local martingale M has a unique (up to indistinguishability) decomposition in a continuous local martingale M^c and a purely discontinuous local martingale M^d:

$$X = X_0 + M^c + M^d + B. \tag{2.6}$$

More generally, the continuous local martingale part of a semimartingale X is denoted by X^c and equals M^c, while the discontinuous local martingale part is

denoted by $X^d = M^d$.
If we introduce the notation μ for the random measure associated to the jumps of X and defined by

$$\mu(dt, dx) = \sum \mathbb{1}_{\{\Delta X_s \neq 0\}} \delta_{(s, \Delta X_s)}(dt, dx),$$

with $\mathbb{1}$ the indicatorfunction, δ_a the Dirac measure at point a and with compensator ν, then $M^d = x \star (\mu - \nu)$. Remark that we made here the additional assumption that X is special, hence from Jacod and Shiryaev (2002) we know that this holds if and only if $(|x|^2 \wedge |x|) \star \nu \in \mathcal{A}_{\text{loc}}$. Therefore as shown in Corollary II.2.38 of Jacod and Shiryaev (2002) the truncation function, which 'behaves like x' near the origin, in $M^d = h(x) \star (\mu - \nu)$ can be chosen equal to x.
Also in the following definition and proposition we restrict ourselves to the specific case that X is special. The more general form of this definition and proposition can be found in Jacod and Shiryaev (2002).

Definition 2.2.19. Assume that the canonical representation for the special semimartingale X is given by

$$X = X_0 + X^c + x \star (\mu - \nu) + B. \tag{2.7}$$

If C is the matrix with entries $C^{ij} := \langle X^{c,i}, X^{c,j} \rangle$, then the **predictable characteristics** of X are the triplet (B, C, ν).

Furthermore in Proposition II.2.9 Jacod and Shiryaev (2002) show:

Proposition 2.2.20. *One can find a version of the characteristics (B, C, ν) of X which is of the form:*

$$B = b \cdot A, \qquad C = c \cdot A, \qquad \nu(\omega; dt, dx) = dA_t(\omega) K_{\omega,t}(dx)$$

where:

1. *A is a predictable process in $\mathcal{A}^+_{loc}$, which may be chosen continuous if and only if X is quasi-left-continuous;*
2. *$b = (b^i)_{i \leq d}$ is a d-dimensional predictable process;*
3. *$c = (c^{ij})_{i,j \leq d}$ is a predictable process with values in the set of all symmetric non-negative and positive semidefinite $d \times d$-matrices;*

4. *$K_{\omega,t}(dx)$ satisfies the following properties:*

$$K_{\omega,t}(\{0\}) = 0 \qquad \int(|x|^2+1)K_{\omega,t}(dx) \le 1$$
$$\Delta A_t(\omega) > 0 \Rightarrow b_t(\omega) = \int xK_{\omega,t}(dx) \quad \Delta A_t(\omega)K_{\omega,t}(\mathbb{R}^d) \le 1.$$

Two other properties of the characteristics, described in Proposition 2.2.20 and which will be used later on, are

$$\Delta B_t = \int x\nu(\{t\}\times dx) \quad \text{and} \quad a_t := \nu(\{t\}\times\mathbb{R}^d) \le 1.$$

We will call the triplet (b, c, K) the **differential characteristics** of the semimartingale X and we denote them by ∂X.
We can now describe the characteristics of a stochastic integral. First we define the class $L(X)$ of predictable processes for which we can determine the stochastic integral with respect to a d-dimensional semimartingale X. It contains the processes H which belongs to $L^2_{\text{loc}}(M) \cap L^0(B)$. Hence they are an element of the intersection of the predictable processes H such that $H^2 \cdot \langle M, M\rangle$ are locally integrable and also the increasing process $|\sum_i H^i b^i| \cdot A$ is finite-valued.

Proposition 2.2.21 (See Kallsen (2006)). *Let X be an $\mathbb{R}^d$-valued semimartingale and H an $\mathbb{R}^{n\times d}$-valued predictable process with $H^{j\cdot} \in L(X)$, $j = 1, \ldots, n$. If $\partial X = (b, c, K)$, then the differential characteristics of the $\mathbb{R}^n$-valued integral process*

$$H \cdot X := (H^{j\cdot} \cdot X)_{j=1,\ldots,n}$$

equal $\partial(H \cdot X) = (\tilde{b}, \tilde{c}, \tilde{K})$, where

$$\begin{aligned}\tilde{b}_t &= H_t B_t,\\ \tilde{c}_t &= H_t c_t H_t',\\ \tilde{K}_t(G) &= \int_{\mathbb{R}^d} \mathbb{1}_G(H_t x)K_t(dx), \quad G \in \mathcal{B}^n.\end{aligned}$$

With any measurable function W on $\widetilde{\Omega} = \Omega \times [0, T] \times \mathbb{R}^d$ we associate the process

$$\hat{W}_t(\omega) = \int_{\mathbb{R}^d} W(\omega, t, x)\nu(\omega; \{t\}\times dx) \quad \text{if} \int_{\mathbb{R}^d} |W(\omega, t, x)|\nu(\omega, \{t\}\times dx) < \infty \tag{2.8}$$

and $+\infty$ in the other case.
For any process $g \in \widetilde{\mathcal{O}}$, we define $M_\mu^P(g \mid \widetilde{\mathcal{P}})$ as the unique $\widetilde{\mathcal{P}}$-measurable process, when it exists, such that for any bounded $W \in \widetilde{\mathcal{P}}$,

$$M_\mu^P(Wg) := E\left(\int_0^T \int_{\mathbb{R}^d} W(s,x)g(s,x)\mu(ds,dx)\right) = M_\mu^P\left(W M_\mu^P(g \mid \widetilde{\mathcal{P}})\right).$$

We denote by $G_{\text{loc}}(\mu)$ the set of all $\widetilde{\mathcal{P}}$-measurable real-valued functions W on $\widetilde{\Omega}$ such that the process $\widetilde{W}_t(\omega) = W(\omega, t, \Delta X_t(\omega))\mathbb{1}_{\{\Delta X_t(\omega)\neq 0\}}(\omega,t) - \hat{W}_t(\omega)$ satisfies

$$\left[\sum_{s\leq \cdot}(\widetilde{W}_s)^2\right]^{1/2} \in \mathcal{A}_{\text{loc}}^+.$$

We assumed here that the random measure μ is associated with the process X; for measures μ not necessarily linked with a process we refer to Jacod and Shiryaev (2002).

Definition 2.2.22. If $W \in G_{\text{loc}}(\mu)$, then any purely discontinuous martingale such that $\Delta(W \star (\mu - \nu))$ and $\widetilde{W}$ are indistinguishable is called the **stochastic integral of W with respect to** $\mu - \nu$ and is denoted by $W \star (\mu - \nu)$.

The following representation theorem will also be very useful later on. The theorem is an adaptation of the one given in Jacod (1979) (Theorem 3.75, page 103) and Jacod and Shiryaev (2002) (Lemma III.4.24).

Theorem 2.2.23. *Let $N \in \mathcal{M}_{0,loc}$. Then there exists a predictable and X^c-integrable process ϕ, $N^\perp \in \mathcal{M}_{0,loc}$ with $[N^\perp, X] = 0$ and functionals $f \in \widetilde{\mathcal{P}}$ and $g \in \widetilde{\mathcal{O}}$ such that*

$$\int_0^T \int_{\mathbb{R}^d\setminus\{0\}} (|f| \wedge |f|^2)\nu(dt,dx) < +\infty, \quad (\sum_{s=0}^t g(s,\Delta X_s)^2 \mathbb{1}_{\{\Delta X_s \neq 0\}})^{1/2} \in \mathcal{A}_{loc}^+,$$

$$M_\mu^P(g|\widetilde{\mathcal{P}}) = 0, \qquad W = f + \frac{\hat{f}}{1-a}\mathbb{1}_{\{a<1\}},$$

$$N = \phi \cdot X^c + W \star (\mu - \nu) + g \star \mu + N^\perp, \tag{2.9}$$

where f has a version such that $\{a = 1\} \subset \{\hat{f} = 0\}$. Moreover

$$\Delta N_t = (f_t(\Delta X_t) + g_t(\Delta X_t))\mathbb{1}_{\{\Delta X_t \neq 0\}} - \frac{\hat{f}_t}{1-a_t}\mathbb{1}_{\{\Delta X_t = 0\}} + \Delta N_t^\perp. \tag{2.10}$$

Proof. This theorem is a combination of the one given in Jacod (1979) and in Jacod and Shiryaev (2002). Hence we will only concentrate on the jump ΔN_t. We remark first that $\hat{W} = \frac{\hat{f}}{1-a}\mathbb{1}_{\{a<1\}}$ because

$$\begin{aligned}\hat{W} = &\int W\nu(\{t\} \times dx) = \int f\nu(\{t\}, dx) + \int \frac{\hat{f}}{1-a}\mathbb{1}_{\{a<1\}}\nu(\{t\}, dx)\\ =&\hat{f}(1 + \frac{a}{1-a})\mathbb{1}_{\{a<1\}} = \frac{\hat{f}}{1-a}\mathbb{1}_{\{a<1\}}.\end{aligned}$$

The jump is now calculated in the following way, where we use Definition 2.2.22:

$$\begin{aligned}\Delta N_t =&\Delta(\phi \cdot X^c)_t + \Delta(W \star (\mu - \nu))_t + \Delta(g \star \mu)_t + \Delta N_t^{\perp}\\ =&0 + W(\Delta X_t)\mathbb{1}_{\{\Delta X_t \neq 0\}} - \hat{W}(t) + g(\Delta X_t)\mathbb{1}_{\{\Delta X_t \neq 0\}} + \Delta N_t^{\perp}\\ =&(f_t(\Delta X_t) + g_t(\Delta X_t))\mathbb{1}_{\{\Delta X_t \neq 0\}} + \frac{\hat{f}}{1-a}\mathbb{1}_{\{a<1\}}\mathbb{1}_{\{\Delta X_t \neq 0\}}\\ &- \frac{\hat{f}}{1-a}\mathbb{1}_{\{a<1\}} + \Delta N_t^{\perp}\\ =&(f_t(\Delta X_t) + g_t(\Delta X_t))\mathbb{1}_{\{\Delta X_t \neq 0\}} - \frac{\hat{f}}{1-a}\mathbb{1}_{\{\Delta X_t = 0\}} + \Delta N_t^{\perp}.\end{aligned}$$

□

The uniqueness of the specific components in the decomposition (2.9) is proved in the following lemma.

Lemma 2.2.24. *The decomposition in* (2.9) *is unique (up to indistinguishability) in the following sense: if there exists a quadruplet* $(\phi, f, g, N^{\perp})$ *as in Theorem 2.2.23 satisfying*

$$0 = \phi \cdot X^c + W \star (\mu - \nu) + g \star \mu + N^{\perp}, \tag{2.11}$$

then

$$c\phi = 0 \quad P \otimes dA\text{-a.e.}, \quad f(x) = g(x) = 0 \quad \mu\text{-a.e.}, \quad N^{\perp} = 0.$$

Proof. From (2.6) we know that every local martingale has a unique decomposition in a continuous local martingale part and a discontinuous local martingale part, which in the case of (2.11) should both be zero. Hence, we deduce that

$$\phi \cdot X^c + (N^{\perp})^c \equiv 0. \tag{2.12}$$

Due to $[X, N^\perp] = 0$ and Properties 2.2.15(5) we obtain that for all t

$$\Delta[X, N^\perp]_t = \Delta X_t \Delta N_t^\perp = 0. \tag{2.13}$$

Therefore from Thereom 2.2.14 we know that $\langle X^c, (N^\perp)^c \rangle = 0$. Combining this with (2.12) leads to

$$0 = \langle \phi \cdot X^c + (N^\perp)^c, \phi \cdot X^c + (N^\perp)^c \rangle = \langle \phi \cdot X^c, \phi \cdot X^c \rangle + \langle (N^\perp)^c, (N^\perp)^c \rangle$$

Since for every semimartingale Y, $\langle Y, Y \rangle$ is a non-negative function this relation implies that

$$\phi \cdot X^c = 0, \quad \text{and} \quad (N^\perp)^c = 0, \tag{2.14}$$

Further, from $\phi \cdot X^c = 0$ it follows that also $\phi' c \phi \cdot A = 0$. Using the fact that c is positive semidefinite with $c = w'w$, we obtain that $(w\phi)'(w\phi) \cdot A = 0$. Furthermore if we choose $c = 0$, and hence also $w = 0$ on $\{\Delta A = 0\}$, we can conclude that $c\phi = 0$ $P \otimes dA$-a.e..
Formula (2.13) implies that on the set $\{\Delta X \neq 0\}$, $\Delta N^\perp = 0$. Hence (2.10) with $N = 0$ and therefore also $\Delta N = 0$, leads to $f(\Delta X) + g(\Delta X) = 0$, which is equivalent to

$$f(x) + g(x) = 0 \quad \mu\text{-}a.e.$$

By taking the conditional expectation under M_μ^P, and using $M_\mu^P(g|\widetilde{\mathcal{P}}) = 0$ we conclude that

$$f = g = 0 \quad M_\mu^P\text{-}a.e. \tag{2.15}$$

This implies that $\widehat{f} = 0$. So again by (2.10) we get $\Delta N^\perp = 0$. Combined with the second equation in (2.14) this leads to $N^\perp = 0$. This completes the proof of the lemma. □

The following definition is very useful in the context of change of measures.

Definition 2.2.25. The unknown càdlàg adapted process Y, which is the solution to

$$Y = 1 + Y_- \cdot X, \tag{2.16}$$

with X a given semimartingale, is denoted by $\mathcal{E}(X)$ and is called the **Doléans-Dade exponential**.

Hence it is obvious that Y is a local martingale if X is one.
For real-valued semimartingales X, the solution to (2.16) is given by

$$\mathcal{E}(X)_t = e^{X_t - X_0 - \frac{1}{2}\langle X^c, X^c \rangle_t} \prod_{s \le t} (1 + \Delta X_s) e^{-\Delta X_s}. \tag{2.17}$$

Furthermore it is possible to prove the following relationship for two semimartingales X and X':

$$\mathcal{E}(X)\mathcal{E}(X') = \mathcal{E}(X + X' + [X, X']), \tag{2.18}$$

by applying the well-known Itô's formula:

Theorem 2.2.26 (Itô's formula). *Let X be a d-dimensional semimartingale, and f a function of class C^2 on $\mathbb{R}^d$. Then $f(X)$ is a semimartingale and*

$$\begin{aligned} f(X_t) =& f(X_0) + \sum_{i\leq d} D_i f(X_-) \cdot X^i + \frac{1}{2}\sum_{i,j\leq d} D_{ij} f(X_-) \cdot \langle X^{i,c}, X^{j,c}\rangle \\ &+ \sum_{s\leq t}\left[f(X_s) - f(X_{s-}) - \sum_{i\leq d} D_i f(X_{s-})\Delta X^i_s \right]. \end{aligned}$$

In terms of the characteristic triplet (B, C, ν) of a semimartingale X, Itô's formula takes the following form:

$$\begin{aligned} f(X_t) =& f(X_0) + \sum_{i\leq d} D_i f(X_i) \cdot X^i + \frac{1}{2}\sum_{i,j\leq d} D_{ij} f(X_-) \cdot C \\ &+ \int_0^t \left[f(X_{s-} + x) - f(X_{s-}) - \sum_{i\leq d} D_i f(X_{s-})x \right] \mu_s(dx). \end{aligned} \tag{2.19}$$

2.3 Measures

2.3.1 Girsanov's theorems

In this section we will state some relevant theorems in order to derive the characteristics of a semimartingale X under a new measure P', knowing the characteristics of X under the original measure P. Assume the characteristics under P of X are (B, C, ν). Furthermore the density process of P' relative to P is denoted by Z and it is the martingale describing the change of measure from P to P' such that Z_t is the Radon-Nikodym derivative $\frac{dP'_{|\mathcal{F}_t}}{dP_{|\mathcal{F}_t}}$ of the restrictions of

P' and P to $(\Omega, \mathcal{F}_t)$ for every $t \in \mathbb{R}_+$. If the process $Z > 0$ then P' is a probability measure, in the other case we call P' a signed measure. If P' is **locally absolutely continuous** with respect to P, then we write $P' \overset{\text{loc}}{\ll} P$. A measure P' is **absolutely continuous** with respect to P, if $P'(A) = 0$ for every set A for which $P(A) = 0$. While $P' \sim P$ means that P' is **equivalent** with P, this means that $P' \ll P$ and $P \ll P'$.

Proposition 2.3.1 (See Jacod and Shiryaev (2002) III.3.8). *Assume that $P' \overset{loc}{\ll} P$ and let Z be the density process. Let M' be a càdlàg process.*

- *If $M'Z$ is a P-local martingale, then M' is a P'-local martingale.*
- *If M' is a P'-local martingale with a localizing sequence (T_n) having $P(\lim_n \uparrow T_n = \infty) = 1$, then $M'Z$ is a P-local martingale.*

Corollary 2.3.2. *If the process Y is a P-local martingale then it is also a local martingale under the new measure P' described by the Girsanov density $Z = \mathcal{E}(N)$ if and only if $[Y, N]$ is a P-local martingale.*

Proof. Y is a local martingale under the new measure if and only if YZ is a P-local martingale (remark we work here on a finite time horizon). Since

$$d(YZ) = Y_- dZ + Z_- dY + Z_- d[Y, N], \tag{2.20}$$

we see that this holds if and only if $[Y, N]$ is a P-local martingale. □

Remark 2.3.3. From (2.20) it also follows that if Y is a local martingale under P', then $Y + [Y, N]$ is a P-local martingale, since Z is a P-local martingale.

Theorem 2.3.4 ('classical' Girsanov's theorem). *Assume that $P' \overset{loc}{\ll} P$ and let Z be the density process. Let M be a P-local martingale such that $M_0 = 0$ and that the P-quadratic covariation $[M, Z]$ has P-locally integrable variation, and denote by $\langle M, Z \rangle$ its P-compensator. Then the process*

$$M' = M - \frac{1}{Z_-} \cdot \langle M, Z \rangle$$

is P'-a.s. well defined, and is a P'-local martingale.

The following theorem is an adaptation of Theorem III.3.24 of Jacod and Shiryaev (2002), in analogy with the one given in Kallsen (2006) for quasi-left-continuous processes.

Theorem 2.3.5 (Girsanov's theorem for semimartingales, see Kallsen (2006)). *Let X be an $\mathbb{R}^d$-valued semimartingale with differential characteristics $\partial X = (b, c, K)$. Suppose that $P' \overset{loc}{\ll} P$ with density process*

$$Z = \mathcal{E}(H \cdot X^c + W \star (\mu - \nu)) \tag{2.21}$$

for some $H \in L(X^c)$ and $W \in G_{loc}(\mu)$. Then the differential characteristics $(\tilde{b}, \tilde{c}, \tilde{K})$ of X relative to P' are given by

$$\begin{aligned}
\tilde{b}_t =& b_t + H_t' c_t + \int_{\mathbb{R}^d} (W(t,x) - \hat{W}(t)) h(x) K_t(dx), \\
\tilde{c}_t =& c_t, \\
\tilde{K}_t(G) =& \int_{\mathbb{R}^d} \mathbb{1}_G(x)(1 + W(t,x) - \hat{W}(t)) K_t(dx), \quad G \in \mathcal{B}^n.
\end{aligned}$$

2.3.2 Important measures

As we will describe in Section 2.5, we will mostly work with Lévy processes. This type of processes contains jumps and hence the market we will work in will rarely be complete.
A market is called **complete** if any $\mathcal{F}_T$-measurable claim H can be hedged, i.e. it can be replicated by a self-financing strategy V such that $P(H = V_T) = 1$. A strategy is **self-financing** if there are no external cash-flows after the start, so we may only rebalance the portfolio.
Using an incomplete market means that we have a set of martingale measures and there exists no longer a unique martingale measure. Hence also the concept of a unique arbitrage-free price for a product found through replication makes no sense in an incomplete market.
In this thesis we will concentrate on the hedging of products and especially on the quadratic hedging strategies. To determine these strategies two martingale measures are really important: the minimal martingale measure and the variance-optimal martingale measure. The goal of this section is to give more details about these measures and to explain how we can determine them.
We mainly use the following sets of local martingale densities associated with the measure P:

- The space $\mathcal{M}^s(P)$ which contains all signed measures $Q \ll P$ with $Q[\Omega] = 1$ and
$$E[\frac{dQ}{dP}Y] = 0 \quad \text{for all } Y \in \mathcal{V}.$$
An element of $\mathcal{M}^s$ is called a signed martingale measure.
- The space $\mathcal{M}^e(P)$ contains all probability measures $Q \in \mathcal{M}^s(P)$ such that Q is equivalent to P. An element of $\mathcal{M}^e$ is called an equivalent martingale measure.
- Related with these spaces we define the following set of densities:
$$\mathcal{D}^g := \left\{ \frac{dQ}{dP} \mid Q \in \mathcal{M}^g(P) \right\} \quad \text{with } g \in \{s, e\}. \tag{2.22}$$

2.3.2.1 Minimal martingale measure

The minimal martingale measure was introduced in Föllmer and Schweizer (1991). An extended version to local martingales of their definition is:

Definition 2.3.6. The **minimal martingale measure** $\widetilde{Q}$ related to a P-semimartingale X is the martingale measure such that any P-local martingale which is orthogonal to M, as defined in (2.2), under P remains a local martingale under $\widetilde{Q}$.

In the paper by Föllmer and Schweizer (1991) the uniqueness is shown and existence results are given for a one-dimensional continuous semimartingale. Furthermore also the orthogonality is preserved in this case, namely any square-integrable P-martingale, orthogonal to M under P is also orthogonal to X under the minimal martingale measure.
The terminology minimal does not come from the fact that this measure minimizes a certain criterion, but it is the measure that preserves the structure of the semimartingale X as much as possible under that change of measure which makes X a martingale.
Schweizer (1995a) gives three characterizations for the minimal martingale measure. Two only hold in the continuous case, while the third one also holds in the discontinuous case under the assumption that the structure condition holds and the MVT process, as defined in 2.2.18 is deterministic. This characterization is also crucial for the determination of the variance-optimal martingale

measure. Under the conditions mentioned above, the minimal martingale measure is the unique solution to

$$\min ||\frac{dQ}{dP} - 1||_{L^2(P)} = \sqrt{\text{Var}[\frac{dQ}{dP}]}$$

over all signed martingale measures Q for X with $\frac{dQ}{dP} \in L^2(P)$, with $|| \cdot ||_{L^2(P)}$ the norm of the space $L^2(P)$:

$$||A||_{L^2(P)} = E[|A|^2]^{1/2}.$$

In Example 1 of Choulli et al. (2007), we see that if $X = X_0 + M + B$ is a locally square-integrable semimartingale and the structure condition is satisfied, then the minimal martingale density is given by $\mathcal{E}(-\lambda \cdot M)$, with λ defined in (2.5). From Theorem 2.3.4, we deduce that if $\mathcal{E}(\widetilde{N}) > 0$ then the following process is a local martingale under the MMM:

$$X^{c,\widetilde{Q}} = X^c + c\lambda \cdot A \tag{2.23}$$

and the compensator of μ under the MMM is given by

$$\nu^{\widetilde{Q}}(dt, dx) = (1 - \lambda_t' x + \lambda_t' \Delta\langle M\rangle_t \lambda_t)\nu(dt, dx). \tag{2.24}$$

If $\mathcal{E}(\widetilde{N})$ is not a priori strictly positive, which is possible in discontinuous cases, we cannot link a probability measure with the density $\mathcal{E}(\widetilde{N})$. Hence it is not possible to define $\widetilde{Q}$-martingales and we need to use the concept of $\mathcal{E}(\widetilde{N})$-martingales, see Chapter 3 to overcome this problem. Definition 2.3.6 holds then in the following sense: for every P-local martingale L orthogonal to M under P, $\mathcal{E}(\widetilde{N})L$ should be a P-local martingale.
The following lemma will be useful later on:

Lemma 2.3.7. *The $\widetilde{Q}$-compensator of a finite variation process K coincides with the P-compensator of $(1 + \Delta\widetilde{N}) \cdot K$.*

Proof. Denote by $K^{\widetilde{Q}}$ the $\widetilde{Q}$-compensator of K. By Remark 2.3.3, $K - K^{\widetilde{Q}}$ is a $\widetilde{Q}$-local martingale if and only if $(K - K^{\widetilde{Q}}) + [K - K^{\widetilde{Q}}, \widetilde{N}]$ is a P-local martingale. Using the fact that K is a process of finite variation, the predictability and the finite variation property of the compensator $K^{\widetilde{Q}}$ and Properties 2.2.15, we rewrite this last expression as

$$\begin{aligned} K - K^{\widetilde{Q}} + [K - K^{\widetilde{Q}}, \widetilde{N}] &= K - K^{\widetilde{Q}} + \Delta\widetilde{N} \cdot K - \Delta K^{\widetilde{Q}} \cdot \widetilde{N} \\ &= (1 + \Delta\widetilde{N}) \cdot K - K^{\widetilde{Q}} - \Delta K^{\widetilde{Q}} \cdot \widetilde{N}. \end{aligned}$$

Due to the martingale property of $\widetilde{N}$, we conclude that $K - K^{\widetilde{Q}}$ is a $\widetilde{Q}$-local martingale if and only if $K^{\widetilde{Q}}$ is the P-compensator of $(1 + \Delta\widetilde{N}) \cdot K$. □

Characterizing the class of measures with $\widetilde{Q}$ as the MMM.

This paragraph is based on a remark given in Schweizer (1991):

> Furthermore, Theorem 3.2 tells us that this strategy is robust: it will again be optimal for a whole class of semimartingale models P, namely all those which admit $\widetilde{P}$ as their minimal equivalent martingale measure.

The goal is to find these measures with the same MMM as the original measure P. We know that P and the MMM are trivial elements of this class. We will assume that the process under consideration is quasi-left-continuous. This imposes no restriction because the strategy Schweizer is speaking of is the locally risk-minimizing hedging strategy, which is only defined in the case that the finite variation part is continuous, see Chapter 4.
As a first step we prove that any change of measure from the original measure P to an equivalent probability measure, can be described by the process Z as given in formula (2.21). This is a very useful result if we want to extend the proof of Lemma 2.3.9 to processes which are not necessarily quasi-left-continuous.

Theorem 2.3.8. *If Z describes a change of measure from P to the probability measure Q, $Q \overset{loc}{\ll} P$, where the filtration $\mathbb{F}$ is the natural filtration of the semimartingale X as defined in Definition 2.2.19, then Z is given by*

$$\mathcal{E}(H \cdot X^c + W \star (\mu - \nu)). \tag{2.25}$$

with $H \in L(X^c)$ and $W \in G_{loc}(\mu)$.

Proof. We remark that we were inspired by the proof of Lemma 5.1 of Kallsen (2004).
If Z describes a change of measure from P to Q, then Z is a P-local martingale.

From Theorem II.8.3 of Jacod and Shiryaev (2002) we know that the stochastic logarithm $N = \frac{1}{Z_-} \cdot Z$ of Z as the unique process such that $Z = \mathcal{E}(N)$. Hence N will also be a P-local martingale and according to Theorem 2.2.23 has the following representation:

$$N = H \cdot X^c + W \star (\mu - \nu) + g \star \mu + N^\perp. \tag{2.26}$$

We use Theorem III.3.24 of Jacod and Shiryaev (2002) to determine the characteristics of X under the new measure Q. First, we need to determine the process Y such that

$$YZ_- = M_\mu^P(Z|\widetilde{\mathcal{P}}). \tag{2.27}$$

From Jacod and Shiryaev (2002) Theorem III.3.17 b) we know that any non-negative version of $M_\mu^P(\frac{Z}{Z_-}\mathbb{1}_{\{Z_->0\}}|\widetilde{\mathcal{P}})$ satisfies condition (2.27). Using formula (2.17) and the continuity of $\langle X^c, X^c\rangle$ (see Theorem 2.2.13), we find

$$\begin{aligned} Z_t =& e^{N_t - N_{t-}} e^{N_{t-} - N_0 - \frac{1}{2}\langle N^c, N^c\rangle_t} \prod_{s<t} (1 + \Delta N_s) e^{-\Delta N_s} (1 + \Delta N_t) e^{-\Delta N_t} \\ =& Z_{t-}(1 + \Delta N_t), \end{aligned}$$

with ΔN_t as in formula (2.10). So Y is given by $M_\mu^P(1 + \Delta N|\widetilde{\mathcal{P}}) = W - \hat{W}$, because $M_\mu^P(g|\widetilde{\mathcal{P}}) = 0$ and $\Delta N^\perp = 0$ on the set $\{\Delta X \neq 0\}$ which follows from the fact that $[X, N^\perp] = 0$ as is shown in (2.13).
Second, we determine the process β, such that $\langle Z^c, X^c\rangle = (Z_- c\beta) \cdot A$:

$$\langle Z^c, X^c\rangle = Z_-\langle H \cdot X^c, X^c\rangle + Z_-\langle (N^\perp)^c, X^c\rangle = Z_- H' c \cdot A,$$

because we know again from the conclusion after (2.13) that $\langle X^c, (N^\perp)^c\rangle = 0$ for functions $N^\perp$ with $[X, N^\perp] = 0$. So β can be chosen equal to H. Hence comparison with Theorem 2.3.5 and Theorem III.3.24 of Jacod and Shiryaev (2002) leads to the conclusion that the characteristics of X under the measure Q are the same as those obtained after a change of measure described by (2.25). This means that without loss of generality the change of measure Z can be described by (2.25). □

This was in fact already proved by Jacod and Shiryaev (2002) in Theorem III.5.19 using the fact that if all local martingales satisfy the representation theorem, then the third and fourth term in formula (2.10) are trivial.

Theorem 2.3.9. *Assume $\widetilde{Q}$ is the MMM of the QLC, special semimartingale X with predictable characteristics as described in Definition 2.2.19 under the measure P.*

Then all measures $\widetilde{P}$ for which the Girsanov densities describing the change of measure from P to $\widetilde{P}$ are given by

$$Z = \mathcal{E}(\beta_P^{\widetilde{P}} \cdot X^c + W_P^{\widetilde{P}} \star (\mu - \nu)),$$

with

$$-\lambda_P' x = W_P^{\widetilde{P}} - (\lambda_P + \beta_P^{\widetilde{P}})' x (1 + W_P^{\widetilde{P}})$$

have also $\widetilde{Q}$ as MMM.

Proof. The change of measure from P to the MMM $\widetilde{Q}$ is denoted by

$$Z_P^{\widetilde{Q}} = \mathcal{E}(N_P^{\widetilde{Q}}) = \mathcal{E}(-\lambda_P \cdot M),$$

where we know from (2.5), Definition 2.2.19, Proposition 2.2.20 and because X is QLC that

$$\begin{aligned} M =& X^c + x \star (\mu - \nu) \quad \text{and} \quad B = b \cdot A & (2.28) \\ \langle M \rangle =& \langle X \rangle = \langle X^c \rangle + \langle x \star (\mu - \nu) \rangle = (c + \int x x' K(dx)) \cdot A & \\ \lambda_P =& (d\langle M \rangle)^{\text{inv}} dB = (c + \int_{\mathbb{R}^d} x x' K(dx))^{\text{inv}} b. & (2.29) \end{aligned}$$

We assume here that the inverse functions really exists and we do not use the extended Moore-Penrose pseudo-inverse.
We know from Theorem 2.3.8 that the following density process

$$Z_P^{\widetilde{P}} = \mathcal{E}(N_P^{\widetilde{P}}) = \mathcal{E}(\beta_P^{\widetilde{P}} \cdot X^c + W_P^{\widetilde{P}} \star (\mu - \nu)) \tag{2.30}$$

defines a change of measure from P to an equivalent measure $\widetilde{P}$. Remark we assume here that $Z_P^{\widetilde{P}} > 0$ and hence defines a true probability measure. It makes sense to have this assumption because we are only interested in finding equivalent measures which have the same MMM. On the other hand, if this assumption does not hold, the proof still remains true, but we are not always able to link a measure with the described density.
Applying Theorem 2.3.5 and using the QLC of X, the process X has the following representation under $\widetilde{P}$:

$$X = X_0 + X^{c,\widetilde{P}} + x \star (\mu - \nu^{\widetilde{P}}) + B^{\widetilde{P}},$$

with

$$X^{c,\widetilde{P}} = X^c - c\beta_P^{\widetilde{P}} \cdot A, \tag{2.31}$$

$$\nu^{\widetilde{P}} = (1 + W_P^{\widetilde{P}})K(dx) \cdot A := K^{\widetilde{P}}(dx) \cdot A, \tag{2.32}$$

$$B^{\widetilde{P}} = (b + c\beta_P^{\widetilde{P}} + \int_{\mathbb{R}^d} xW_P^{\widetilde{P}} K(dx)) \cdot A. \tag{2.33}$$

We now calculate the MMM Q^* of X under $\widetilde{P}$:

$$Z_{\widetilde{P}}^{Q^*} = \mathcal{E}(N_{\widetilde{P}}^{Q^*}) = \mathcal{E}(-\lambda_{\widetilde{P}} \cdot M^{\widetilde{P}}),$$

with

$$\lambda_{\widetilde{P}} = (c + \int_{\mathbb{R}^d} xx' K^{\widetilde{P}}(dx))^{\text{inv}}(b + c\beta_P^{\widetilde{P}} + \int_{\mathbb{R}^d} xW_P^{\widetilde{P}} K(dx)) \tag{2.34}$$

$$M^{\widetilde{P}} = X^c - c\beta_P^{\widetilde{P}} \cdot A + x \star (\mu - \nu^{\widetilde{P}}). \tag{2.35}$$

Next we determine $\beta_P^{\widetilde{P}}$ and $W_P^{\widetilde{P}}$ such that the MMM of P, namely $\widetilde{Q}$ is the same as the MMM of $\widetilde{P}$, namely Q^*. This will hold if and only if $Z_P^{\widetilde{Q}} = Z_P^{\widetilde{P}} Z_{\widetilde{P}}^{Q^*}$, by using formula (2.18):

$$\mathcal{E}(N_P^{\widetilde{Q}}) = \mathcal{E}(N_P^{\widetilde{P}})\mathcal{E}(N_{\widetilde{P}}^{Q^*}) = \mathcal{E}(N_P^{\widetilde{P}} + N_{\widetilde{P}}^{Q^*} + [N_P^{\widetilde{P}}, N_{\widetilde{P}}^{Q^*}]).$$

This leads to the following equation which should be satisfied, because as shown in Theorem II.8.3 of Jacod and Shiryaev (2002) the stochastic logarithm N of $Z = \mathcal{E}(N)$ is unique:

$$-\lambda_P \cdot M = \beta_P^{\widetilde{P}} \cdot X^c + W_P^{\widetilde{P}} \star (\mu - \nu) - \lambda_{\widetilde{P}} \cdot M^{\widetilde{P}} + [N_P^{\widetilde{P}}, N_{\widetilde{P}}^{Q^*}]. \tag{2.36}$$

We first calculate $[N_P^{\widetilde{P}}, N_{\widetilde{P}}^{Q^*}]$. Using the predictability and continuity of A and the fact that $M^{\widetilde{P}} = M + B - B^{\widetilde{P}} = M - (c\beta_P^{\widetilde{P}} + \int_{\mathbb{R}^d} xW_P^{\widetilde{P}} K(dx)) \cdot A$, we obtain by Properties 2.2.15(3) that $[N_P^{\widetilde{P}}, N_{\widetilde{P}}^{Q^*}] = [N_P^{\widetilde{P}}, -\lambda_{\widetilde{P}} \cdot M]$. Again using Properties 2.2.15, (2.28) and (2.30), we arrive at:

$$\begin{aligned}
[N_P^{\widetilde{P}}, N_{\widetilde{P}}^{Q^*}] &= [\beta_P^{\widetilde{P}} \cdot X^c, -\lambda_{\widetilde{P}} \cdot X^c] + [W_P^{\widetilde{P}} \star (\mu - \nu), -\lambda'_{\widetilde{P}} x \star (\mu - \nu)] \\
&= -(\beta_P^{\widetilde{P}})' c\lambda_{\widetilde{P}} \cdot A - \sum (W_P^{\widetilde{P}} \mathbb{1}_{\{\Delta X \neq 0\}})(\lambda'_{\widetilde{P}} x \mathbb{1}_{\{\Delta X \neq 0\}}) \\
&= -(\beta_P^{\widetilde{P}})' c\lambda_{\widetilde{P}} \cdot A - \lambda'_{\widetilde{P}} x W_P^{\widetilde{P}} \star \mu.
\end{aligned} \tag{2.37}$$

Hence substituting (2.37), (2.35) and (2.32) in (2.36), we get

$$-\lambda_P \cdot M = (\beta_P^{\widetilde{P}} - \lambda_{\widetilde{P}}) \cdot X^c + (W_P^{\widetilde{P}} - \lambda'_{\widetilde{P}} x - \lambda'_{\widetilde{P}} x W_P^{\widetilde{P}}) \star (\mu - \nu).$$

The uniqueness of the decomposition of a local martingale in a continuous part and a purely discontinuous part leads in view of (2.34) to the following non-linear system of equations in $\beta_P^{\widetilde{P}}$ and $W_P^{\widetilde{P}}$:

$$\begin{cases} -\lambda_P = \beta_P^{\widetilde{P}} - \lambda_{\widetilde{P}} \\ -\lambda'_P x = W_P^{\widetilde{P}} - \lambda'_{\widetilde{P}} x - \lambda'_{\widetilde{P}} x W_P^{\widetilde{P}}. \end{cases}$$

Solving the first equation for $\lambda_{\widetilde{P}}$ and substituting its expression in the second equation gives:

$$-\lambda'_P x = W_P^{\widetilde{P}} - (\lambda_P + \beta_P^{\widetilde{P}})' x (1 + W_P^{\widetilde{P}}). \tag{2.38}$$

Grouping the terms in λ_P, multiplying with x and integrating (2.38), we arrive at

$$\begin{aligned} & x' \lambda_P W_P^{\widetilde{P}} = W_P^{\widetilde{P}} - x' \beta_P^{\widetilde{P}} (1 + W_P^{\widetilde{P}}) \\ \Leftrightarrow & \int x x' \lambda_P W_P^{\widetilde{P}} K(dx) = \int x W_P^{\widetilde{P}} K(dx) - \int x x' \beta_P^{\widetilde{P}} (1 + W_P^{\widetilde{P}}) K(dx). \end{aligned} \tag{2.39}$$

Inserting (2.29) and (2.34) in the first equation of the system, using (2.32) leads to

$$\begin{aligned} \beta_P^{\widetilde{P}} =& \lambda_{\widetilde{P}} - \lambda_P \\ =& (c + \int x x' K(dx) + \int x x' W_P^{\widetilde{P}} K(dx))^{\text{inv}} (b + c \beta_P^{\widetilde{P}} + \int x W_P^{\widetilde{P}} K(dx)) \\ & - (c + \int x x' K(dx))^{\text{inv}} b. \end{aligned}$$

Multiplying both sides with $c + \int x x' (1 + W_P^{\widetilde{P}}) K(dx)$ (which is non-zero due to the assumption that the real inverse exists) gives

$$
\begin{aligned}
&(c + \int xx'(1 + W_P^{\widetilde{P}})K(dx))\beta_P^{\widetilde{P}} \\
&\quad = (b + c\beta_P^{\widetilde{P}} + \int xW_P^{\widetilde{P}}K(dx)) \\
&\qquad - (c + \int xx'K(dx) + \int xx'W_P^{\widetilde{P}}K(dx))(c + \int xx'K(dx))^{\text{inv}}b \\
&\Leftrightarrow \int (1 + W_P^{\widetilde{P}})xx'K(dx)\beta_P^{\widetilde{P}} \\
&\quad = \int W_P^{\widetilde{P}}xK(dx) - \int W_P^{\widetilde{P}}xx'K(dx)(c + \int xx'K(dx))^{\text{inv}}b. \qquad (2.40)
\end{aligned}
$$

Comparing (2.39) and (2.40) we see that the system of equations reduces to one single equation in view of (2.29). Therefore the only condition linking both parameters is formula (2.38). □

2.3.2.2 Variance-optimal martingale measure

The variance-optimal martingale measure (VOMM) searches for the element of $\mathcal{M}^s(X)$ with smallest L^2-norm:

Definition 2.3.10. The **variance-optimal martingale measure** is the unique measure $Q^* \in \mathcal{M}^s(X)$ such that $\frac{dQ^*}{dP}$ is in $L^2(P)$ and which minimizes

$$||\frac{dQ}{dP} - 1||_{L^2(P)}. \qquad (2.41)$$

over all $\frac{dQ}{dP}$ belonging to $\mathcal{D}^s \cap L^2(P)$.

From the previous section, we know that this is exactly the criterion which is minimized to determine the MMM if the MVT process is deterministic, hence we conclude that in this case the MMM will equal the VOMM. This is a valuable result, because the determination of the VOMM is not really straightforward, and we certainly do not have a unique closed-form solution as is the case for the MMM. Furthermore it is possible that the VOMM exists, but is not equivalent with the original measure P. In the continuous case the problem of equivalence is solved by Delbaen and Schachermayer (1996b):

Theorem 2.3.11 (See Delbaen and Schachermayer (1996b)). *Let X be a continuous, $\mathbb{R}^d$-valued semimartingale and suppose that $\mathcal{D}^e(X) \cap L^2(P) \neq \emptyset$, i.e. there is at least one equivalent local martingale measure with square-integrable density. Then the variance-optimal measure is a probability measure equivalent to P.*

We remark this is more generally proved for q-optimal martingale measures (this is the measure which minimize a criterion as defined in formula (2.41), but with the L^q-norm) in the continuous case by Grandits and Krawczyk (1998). Furthermore Delbaen and Schachermayer (1996b) also proved that in the continuous case the Radon-Nikodym derivative for the VOMM $\widetilde{Q}$ has the following form:

$$\frac{d\widetilde{Q}}{dP} = \frac{1 - \phi \cdot X_T}{E(1 - \phi \cdot X_T)}, \tag{2.42}$$

with ϕ an X-integrable and admissible process. By **admissible** we mean that the stochastic integral $\phi \cdot X$ is a uniformly integrable Q-martingale for any equivalent martingale measure Q with square-integrable density. The reciprocal of this statement is proved and discussed in more detail in Černý and Kallsen (2008b). A very useful result for the determination of the VOMM is given in Schweizer (1996). We first need to define another class of processes instead of the class $\mathcal{M}^s$ used in Definition 2.3.10 by introducing the following sets:

$$\Theta := \left\{ \theta \in L(X) \mid E\left[\int_0^T \theta'_u d\langle M \rangle_u \theta_u + \left(\int_0^T |\theta'_u dB_u| \right)^2 \right] < +\infty \right\}. \tag{2.43}$$

Hence the set Θ contains all the processes θ which belong to $L^2(M)$ and for which $\int_0^T |\theta_u dB_u| \in L^2(P)$. This ensures that $(\theta \cdot X)_{0 \leq t \leq T}$ is a semimartingale of class $\mathcal{S}^2(P)$ with norm defined as

$$||\theta \cdot X||_{\mathcal{S}^2} = E\left[\int_0^T \theta'_u d\langle M \rangle_u \theta_u + \left(\int_0^T |\theta'_u dB_u| \right)^2 \right] < \infty.$$

The space $G_T(\Theta)$ contains all the stochastic integrals $\theta \cdot X$, with $\theta \in \Theta$ and is a subspace of $L^2(X)$. Schweizer denotes by $P_X(\Theta)$ the set of all signed Θ-martingale measures.

Definition 2.3.12. A signed measure Q on $(\Omega, \mathcal{F})$ is called a **signed Θ-martingale measure** if $Q[\Omega] = 1$, $Q \ll P$ with $\frac{dQ}{dP} \in L^2(P)$ and $E[\frac{dQ}{dP} G_T(\theta)] = 0$ for all $\theta \in \Theta$.

The VOMM is then the measure with minimal $L^2(P)$-density over all signed Θ-martingale measures.

Lemma 2.3.13 (Schweizer (1996)). *$\widetilde{P} \in P_X(\Theta)$ is variance-optimal if and only if*

$$\frac{d\widetilde{P}}{dP} \in [1, \infty) + G_T(\Theta)^{\perp\perp}.$$

Note that $G_T(\Theta)^{\perp\perp} = \overline{G_T(\Theta)}$ and if $G_T(\Theta)$ is closed then it equals of course $G_T(\Theta)$.
Schweizer remarks that in many cases of interest the set $P_X(\Theta)$ coincides with the set of signed martingale measures for which the Girsanov density describing the change of measure is square integrable.
Lemma 4.1 of Schweizer (2001), which was already proved in Schweizer (1999), states that if G is a linear subspace of $L^2(P)$ and for which $\overline{G}$ does not contain the constant 1 then the variance-optimal signed G-martingale measure exists and is unique, where signed G-martingale measure is in the sense of Definition 2.3.12.
The extension of (2.42) to discontinuous processes is given by Černý and Kallsen (2007). We remark that they work with signed σ-martingale measures instead of with the signed Θ-martingale measure as Schweizer (1996) does. Furthermore the VOMM is obtained as the change of measure with minimal $L^2(P)$-density belonging to the class of signed σ-martingale measure. Therefore it is in fact the variance-optimal signed $\widetilde{\Theta}$-martingale measure as defined in Definition 2.3.2.2 with Θ replaced by $\widetilde{\Theta}$. The latter class is given by $\{\theta \in L(S) : \theta \text{ admissible}\}$, with admissible in the sense of Definition 2.2 of Černý and Kallsen (2007).
For the possible equivalence in the discontinuous case, we refer to Kohlmann et al. (2010). Under certain assumptions they proof that the equivalence is satisfied if and only if a certain backward semimartingale equation (BSE) has a solution. Unfortunately they do not really show how to find the solution to this BSE and they only give an example for the very specific case of a deterministic MVT.

Some properties, as e.g. the reverse Hölder inequality for the VOMM in the case of discontinuous semimartingales are proved in Arai (2005b).

2.4 Two decompositions

The Galtchouk-Kunita-Watanabe (GKW) decomposition and the Föllmer-Schweizer (FS) decomposition are two crucial decompositions in this work. We introduce here these two decompositions and describe the existence and uniqueness of the GKW decomposition. More details concerning the FS decomposition will be given in Chapter 3.

Definition 2.4.1. An $\mathcal{F}_T$-measurable random variable H has a **Galtchouk-Kunita-Watanabe decomposition** with respect to the local martingale M if there exist a constant H_0, a process $\xi \in L_{\text{loc}}(M)$ and a local martingale L, such that $[L, M]$ is a local martingale, and

$$H = H_0 + (\xi \cdot M)_T + L_T.$$

To ensure the existence the random variable H should be square-integrable and X should be a locally square-integrable martingale. These conditions are in fact too strong, because we know that the GKW decomposition of the local martingale N with respect to the local martingale M exists if the process $\langle N, M\rangle$ exists, if the process $\xi = \frac{d\langle N,M\rangle}{d\langle M,M\rangle}$ is M-integrable and if $\xi \cdot M$ is a local martingale. These conditions are all satisfied if M and N are both locally square integrable martingales.
In fact the GKW decomposition is a projection of $H - H_0$ on the subspace $G_T(\Theta)$ of $L^2(X)$, hence this decomposition exists for $H \in L^2$ if the space $G_T(\Theta)$ is closed. Due to the definition of stochastic integral of (more generally) a local martingale, the space is closed as the stochastic integral describes an isometry. A different way to prove the existence is given by Yor and de Sam Lazaro (1978). He proved that if Y^n and Y are uniformly integrable martingales such that Y^n_∞ converges weakly to Y_∞ in L^1, and if $Y^n_t = \int_0^t \phi^n_s dX_s$, then there exists a predictable process ϕ satisfying $Y_t = \int_0^t \phi_s dX_s$.
Ansel and Stricker (1993) were able to relax the L^2-conditions and still prove

the existence of the GKW decomposition in some particular cases e.g. if X is continuous then H can be arbitrary.

The FS decomposition is the extension of the GKW decomposition to a semimartingale $X = X_0 + M + B$:

Definition 2.4.2. An $\mathcal{F}_T$-measurable random variable H admits a **Föllmer-Schweizer** decomposition if there exist a constant H_0, a process $\xi^{\mathrm{FS}} \in \Theta$ and a local martingale L^{FS}, such that $[L^{\mathrm{FS}}, M]$ is a local martingale, and

$$H = H_0 + (\xi^{\mathrm{FS}} \cdot X)_T + L_T^{\mathrm{FS}}.$$

2.5 Lévy process

To illustrate and apply the obtained results we will often use Lévy processes. This class of processes contains e.g. the Brownian motion and the Poisson process. We will not go into all the details of Lévy processes, but we just repeat the interesting properties of this class of processes. For more facts we refer to Sato (1999) and Applebaum (2004).

Definition 2.5.1 (See Kallsen (2006))**.** An $\mathbb{R}^d$-valued semimartingale X, $X_0 = 0$, is a **Lévy process** if and only if it has a version (b, c, K) of the differential characteristic which does not depend on (ω, t).

Another type of definition frequently used to describe Lévy processes is given in e.g. Sato (1999).

Definition 2.5.2 (See Sato (1999))**.** An adapted and càdlàg process $X = (X_t)_{t \geq 0}$ with $X_0 = 0$ a.s. is a **Lévy process** if

1. X has increments independent of the past (i.e. $X_t - X_s$ is independent of $\mathcal{F}_s$, $0 \leq s < t < \infty$)
2. X has stationary increments (i.e. $X_t - X_s$ has the same distribution as X_{t-s}, $0 \leq s < t < \infty$)
3. X_t is continuous in probability or stochastically continuous (i.e. $\forall \varepsilon > 0$, $\lim\limits_{h \to 0} P(|X_{t+h} - X_t| > \varepsilon) = 0$).

As remarked in Cont and Tankov (2004), we can assume that X is càdlàg, because every Lévy process has a unique modification that is càdlàg, hence we assume the càdlàg property without loss of generality.
An additive process has all the properties of a Lévy process except the stationary increments.
A Brownian motion can now be defined in the following way:

Definition 2.5.3 (See Sato (1999))**.** A stochastic process X on $\mathbb{R}^d$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is a **Brownian motion**, if it is a continuous Lévy process and if, for $t > 0$, X_t has a Gaussian distribution with mean 0 and covariance matrix tI (I is the identity matrix).

Definition 2.5.4 (See Sato (1999))**.** A stochastic process X on $\mathbb{R}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is a **Poisson process** with parameter $c > 0$, if it is a Lévy process and for $t > 0$, X_t has Poisson distribution with mean ct.

Lévy processes are often characterized by the characteristic function $\Phi_X(t)$:

$$\Phi_X(t) := E[\exp(i\langle z, X_t\rangle)],$$

where the Euclidian scalar product on $\mathbb{R}^d$ is denoted by $\langle\cdot,\cdot\rangle$ and the respective norm by $|\cdot|$.
This characterization is called the Lévy-Khintchine representation which for a Lévy process X with characteristic triplet (B, C, ν) is given by

$$E[e^{i\langle z, X_t\rangle}] = e^{t\psi(z)},$$

with

$$\psi(z) = i\langle z, B\rangle - \frac{1}{2}\langle z, Cz\rangle + \int_{\mathbb{R}^d} (e^{i\langle z,x\rangle} - 1 - i\langle z, x\rangle \mathbb{1}_{\{|x|\geq 1\}})\nu(dx). \quad (2.44)$$

In Jacod and Shiryaev (2002) they do not use the terminology of Lévy processes, but instead they look at PII/PIIS processes:

Definition 2.5.5. A **process with independent increments** (PII) on $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is a càdlàg adapted $\mathbb{R}^d$-valued process X, such that $X_0 = 0$ and that for all $0 \leq s \leq t$ the variable $X_t - X_s$ is independent from the σ-field $\mathcal{F}_s$.

Hence PII satisfies is the first property of Definition 2.5.2.

Definition 2.5.6. A **process with stationary independent increments** (PIIS) on $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is a PII such that the distribution of the variables $X_t - X_s$ only depends on the difference $t - s$.

Hence this combines the first and the second property of Definition 2.5.2.
We remark that PII and PIIS are not necessarily semimartingales, see Jacod and Shiryaev (2002) Sectioin II.$4c$. The processes we will use are always semimartingales.
We also introduce an extension of the class of Lévy processes, namely the time-inhomogeneous (or non-homogeneous) Lévy processes, denoted by PIIAC, namely processes with independent increments and absolutely continuous characteristics. This class of processes is popular for the modeling of the interest-rate derivatives market, see Chapter 8.

Definition 2.5.7 (See Kluge (2005)). An adapted stochastic process $X = (X_t)_{0\leq t\leq T}$ with values in $\mathbb{R}^d$ is a **time-inhomogeneous Lévy process** if the following conditions hold:

1. X has independent increments;
2. For every $t \in [0,T]$, the law of X_t is characterized by the characteristic function

$$E[e^{i\langle u,X_t\rangle}] = \exp\int_0^t \left(i\langle u,b_s\rangle - \frac{1}{2}\langle u,c_s u\rangle + \int_{\mathbb{R}^d}(e^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle \mathbb{1}_{\{|x|\geq 1\}})F_s(dx)\right) ds;$$

with $b_s \in \mathbb{R}^d$, c_s a symmetric non-negative-definite $(d\times d)$-matrix and F_s a measure on $\mathbb{R}^d$ that integrates $(|x|^2 \wedge 1)$ and satisfies $F_s(\{0\}) = 0$. It is also assumed that

$$\int_0^T \left(|b_s| + ||c_s|| + \int_{\mathbb{R}^d}(|x|^2\wedge 1)F_s(dx)\right) ds < \infty, \tag{2.45}$$

where $||\cdot||$ denotes any norm on the set of $d\times d$ matrices.

The following properties of PIIAC are proved in Kluge (2005):

- The distribution of PIIAC is infinitely divisible;
- Every PIIAC is also an additive process;
- Every PIIAC is a semimartingale due to the condition (2.45).

An error does not become truth by reason of multiplied propagation, nor does truth become error because nobody sees it.
Mahatma Ghandi (1869-1948)

3 Föllmer-Schweizer decomposition

The locally risk-minimizing (LRM) hedging strategy, that will be studied in Chapter 4, heavily depends on the Galtchouk-Kunita-Watanabe (GKW) decomposition under the minimal martingale measure (MMM) and/or on the Föllmer-Schweizer (FS) decomposition under the original measure. In the present chapter we will discuss this FS decomposition in more detail.
We start with an overview of the results available in the literature. This will show that the uniqueness and existence of the FS decomposition is already proved. Next we will look at the possible preservation of the martingale and the orthogonality property when changing from the original measure to the MMM and vice versa. In Section 3.3, we will concentrate on the relationship between the FS decomposition under the original measure and the GKW decomposition under the minimal martingale measure $\widetilde{Q}$. Section 3.4 contains a practical counterexample by which we really prove that there exists examples where the two decompositions differ. An explicit formula in terms of the predictable characteristic triplet for the FS decomposition is given in Section 3.5.
In Section 2.3.2.1 on page 23 we recalled an example of Choulli et al. (2007) from which we know that $\mathcal{E}(-\lambda \cdot M)$ is the Girsanov density describing the change of measure to the possibly signed minimal martingale measure $\widetilde{Q}$ if

$\mathcal{E}(-\lambda \cdot M) \geq 0$. We denote this Girsanov density by $\mathcal{E}(\widetilde{N})$. This chapter is from the second section onward based on Choulli et al. (2010).
We assume we work in the probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$. The filtration is assumed to be right-continuous, complete and $\mathcal{F}_0$ is assumed to be trivial. The d-dimensional special semimartingale X has the usual canonical decomposition $X_0 + M + B$. Furthermore from Section 3.2 onwards we work under the following set of assumptions:

Assumptions 3.0.8. *The d-dimensional semimartingale X satisfies the structure condition, see Definition 2.2.17. Hence there exists a process λ such that $B = \lambda \cdot \langle M \rangle$. We also assume that the non-decreasing process $(\sup_{0 \leq s \leq t} |X_s|)_{0 \leq t \leq T}$ is locally square-integrable.*

It is important to notice that we do not assume quasi-left-continuity of the process X.

3.1 Definition and existence

We start this section with the definition of the FS decomposition. We use here a slightly different version than described in Definition 2.4.2 by restricting the contingent claims for which we determine the FS decomposition to the class of square-integrable contingent claims. Combining the square-integrability of H with (2.43) defining the space Θ immediately results in the stronger properties of L^{FS}.

Definition 3.1.1. An $\mathcal{F}_T$-measurable and square-integrable random variable H admits a **Föllmer-Schweizer** decomposition if there exist a constant H_0, a (X-integrable) process $\xi^{\text{FS}} \in \Theta$ and a square-integrable martingale L^{FS}, such that $[L^{\text{FS}}, M]$ is a local martingale, and

$$H = H_0 + (\xi^{\text{FS}} \cdot X)_T + L^{\text{FS}}_T.$$

The FS decomposition was introduced in Föllmer and Schweizer (1991), the same paper in which the MMM was introduced, but with the focus on the continuous case. The extension to the discontinuous case was given by Ansel and Stricker (1992), using a slightly different definition of the FS decomposition

than in Föllmer and Schweizer (1991). They show uniqueness and existence of the decomposition under the assumption that the one-dimensional semimartingale X is locally bounded and if certain conditions on the claim and the process, describing the change of measure to the MMM, are satisfied.
Necessary and sufficient conditions for the existence of the FS decomposition in the more-dimensional continuous case were shown in Schweizer (1995a). A simpler proof of the same result is given in Choulli and Stricker (1996), where in addition also the continuity of the uniform convergence in probability is shown. In the discrete-time case the existence of the FS decomposition for any square-integrable contingent claim was proved by Schweizer (1995b) and Schäl (1994) if the semimartingale X has a bounded MVT process.
For the continuous time case the existence of the FS decomposition for a square-integrable contingent claim is proved by Schweizer (1994) under the condition that the d-dimensional semimartingale X satisfies the SC and the MVT process is uniformly bounded and has jumps strictly bounded above by 1. Monat and Stricker (1994) obtained the same result without the condition on the jumps of the MVT process. The proof of the uniqueness and the continuity of the function mapping under the same conditions can be found in Monat and Stricker (1995).
The most general result concerning the existence and uniqueness of the FS decomposition is given by Choulli et al. (1998). They obtained necessary and sufficient conditions by generalizing the notion of martingale under a new measure Q to the concept of $\mathcal{E}(N)$-martingales (or also called $\mathcal{E}$-martingales).

Definition 3.1.2. An increasing sequence of stopping times $\mathbf{T_n}$ is defined by $T_0 = 0$ and $T_{n+1} = \inf\{t > T_n \mid {}^{T_n}\mathcal{E}_t = 0\} \wedge T$, where for any stopping time τ, ${}^{\tau}\mathcal{E}$ denotes the process $\mathcal{E}(N - N^{\tau})$ and N^{τ} is the stopped process as defined in (2.1).

Definition 3.1.3 (See Choulli et al. (1998)). If $N \in \mathcal{L}$, then a càdlàg process Y is an $\boldsymbol{\mathcal{E}}(\mathbf{N})$**-martingale**, if for any n,

$$E(|X_{T_n}{}^{T_n}\mathcal{E}_{T_{n+1}}|) < +\infty$$

and $({}^{T_n}X\ {}^{T_n}\mathcal{E})$ is a martingale.

The class of $\mathcal{E}(N)$-martingales is denoted by $\mathcal{M}(\mathcal{E})$.
The following two definitions are needed as necessary and sufficient conditions to have a FS decomposition:

Definition 3.1.4 (See Choulli et al. (1998)). Let $q \geq 1$. We say that $\mathcal{E}(N)$ satisfies the **reverse Hölder inequality** (R_q) if and only if there exists a constant

$C \geq 1$ such that for any t,

$$E(|^t\mathcal{E}(N)_T|^q|\mathcal{F}_t) \leq C.$$

Definition 3.1.5 (See Choulli et al. (1998))**.** We say that $\mathcal{E}(N)$ is **regular**, if for any n, ${}^{T_n}\mathcal{E}(N)$ is a martingale.

Choulli et al. (1998) also extended the concept of having a FS decomposition from square-integrable contingent claims to the underlying semimartingale X in the following way:

Definition 3.1.6. A semimartingale $X = X_0 + M + B$ **admits a FS decomposition** if there are unique continuous projections π_0, π_1, π_2 and π_3^n for $n \geq 1$: $L^2(\Omega, \mathcal{F}, P) \to L^2(\Omega, \mathcal{F}, P)$ such that every $H \in L^2(\Omega, \mathcal{F}_T, P)$ admits a FS decomposition:

$$\begin{aligned} H =& \pi_0(H) + \pi_1(H) + \pi_2(H) = H_0 + (\theta \cdot X)_T + L_T, \\ \pi_3^n(H) =& H_0 + (\theta \cdot X)_{T_n} + L_{T_n}, \end{aligned}$$

where $H_0 \in L^2(\Omega, \mathcal{F}_0, P)$, $\theta \in \Theta$ and $L \in \mathcal{H}_0^2$ with $\langle M, L \rangle = 0$.

They proved the following theorem concerning the existence and uniqueness of the FS decomposition:

Theorem 3.1.7. *$X = X_0 + M + B$ admits a Föllmer-Schweizer decomposition if and only if $\mathcal{E}(\widetilde{N}) = \mathcal{E}(-\lambda \cdot M)$ (with λ as in Definition 2.2.17) is regular and satisfies (R_2).*

An important underlying condition guaranteeing the existence of the Föllmer-Schweizer decomposition, which is now hidden beneath other conditions, is the closedness of $G_T(\Theta)$. We refer to Choulli et al. (1998) for its proof when $\mathcal{E}(\widetilde{N})$ is regular and satisfies (R_2).

3.2 Martingales under $\widetilde{Q}$ versus P-martingales

In this section we concentrate on the possible preservation of the martingale property and the orthogonality property when changing from the original measure P to minimal martingale measure $\widetilde{Q}$ and vice versa. On the basis of this

information we will detail the difference between the FS under the original measure and the GKW decomposition under the MMM in the following section. The GKW decomposition under the MMM $\widetilde{Q}$ will be the same as the FS decomposition if the martingale property and the orthogonality property are preserved.
We assume for this section that $\mathcal{E}(\widetilde{N}) > 0$, see page 23, hence the MMM $\widetilde{Q}$ exists as a true probability measure.
The first two results are obvious using the definition of minimal martingale measure and assuming sufficient integrability conditions:

Proposition 3.2.1. *If L is a P-local martingale and a $\widetilde{Q}$-local martingale then L is automatically also orthogonal to M under P.*

Proof. From Corollary 2.3.2, we see that if L is a $\widetilde{Q}$-local martingale, then also $[L, N]$ is a P-local martingale. Hence, $[L, N] = -\lambda \cdot [L, M]$ should be a P-local martingale and therefore $\langle L, M\rangle = 0$. □

Proposition 3.2.2. *If L is a P-local martingale orthogonal to M under P, then L is also a $\widetilde{Q}$-local martingale.*

Proof. This follows from Definition 2.3.6 for the minimal martingale measure. □

Therefore the two first propositions can be combined to:

Proposition 3.2.3. *If L is a P-local martingale then L is also a $\widetilde{Q}$-local martingale if and only if L is orthogonal to M under P.*

The preservation of the orthogonality from P to $\widetilde{Q}$ is proved by Föllmer and Schweizer (1991) for the continuous case. For the discontinuous case it is impossible to prove the preservation of the orthogonality, because L is orthogonal to X under $\widetilde{Q}$ if and only if $[L, X]$ is a $\widetilde{Q}$-martingale. Using the original orthogonality, we know that $[L, M] = [L, X]$ is already a P-martingale. Hence from Proposition 3.2.3 we conclude that $[L, X]$ is also a $\widetilde{Q}$-martingale if and only if $\langle M, [L, X]\rangle$ equals zero. In Section 3.3 we will show that an adaptation of the term $\langle M, [L, X]\rangle$ exactly expresses the difference between ξ and ξ^{FS}, while in Section 3.4, we will give an explicit example for which this is non-zero.
Note that the preservation of the orthogonality in fact reduces to the preservation of the martingale property by Definition 2.2.16 of orthogonality. Indeed, L is orthogonal to X under a certain measure $\widetilde{P}$ if and only if $[L, X]$ is a martingale under this measure. Hence investigating the possible preservations from

the minimal martingale measure $\widetilde{Q}$ to the original measure P, we will only discuss the martingale property. Exactly this preservation (or non-preservation) is crucial, because in literature many people want to deduce the FS decomposition from the GKW decomposition under the MMM and not the other way around.
In Proposition 3.2.5 we give two specific types of $\widetilde{Q}$-local martingales which are automatically P-local martingales and furthermore we characterize the $\widetilde{Q}$-local martingales in order to be also P-local martingales. From this proposition we see once more that there is an extra condition on the $\widetilde{Q}$-local martingale Z, namely (3.1), which differs from the orthogonality condition $\langle Z, X\rangle^{\widetilde{Q}} = 0$ to guarantee that Z is also a P-local martingale. Furthermore in the continuous case the condition will be satisfied by the assumption of orthogonality to X under the MMM.
We remarked already that the preservation of the orthogonality reduces to the preservation of the martingale property but due to the specificity of the measure $\widetilde{Q}$, we can also show that the preservation of the orthogonality implies the preservation of the martingale property.

Proposition 3.2.4. *Let L be a $\widetilde{Q}$-local martingale. Then, L is P-locally integrable and is P-orthogonal to M if and only if L is a P-local martingale that is orthogonal to M.*

Proof. According to Definition 2.2.8 L is P-locally integrable if L is a special semimartingale. If L is a P-local martingale, then L is of course also a special semimartingale and hence L is P-locally integrable. This proves the only if part.
Again by Definition 2.2.8 and by Definition 2.2.5 there exist a P-local martingale $\overline{L}$, and a predictable process $\overline{B}$ with finite variation such that

$$L = \overline{L} + \overline{B}.$$

By Properties 2.2.15(2), we deduce that then $\langle L, M\rangle = \langle \overline{L}, M\rangle$, and thus L is P-orthogonal to M if and only if $\overline{L}$ is P-orthogonal to M on one hand.
On the other hand, since L is a $\widetilde{Q}$-local martingale, LZ is a P-local martingale due to Proposition 2.3.1:

$$d(LZ) = Z_- dL + L_- dZ + Z_- d[L, \widetilde{N}] = Z_- d\overline{L} + L_- dZ + Z_- d([\overline{L}, \widetilde{N}] + \overline{B}).$$

Therefore $\overline{B} = -\langle \overline{L}, \widetilde{N}\rangle$, because then also the third term is a P-local martingale.
So, L is a $\widetilde{Q}$-local martingale if $\overline{B} = -\langle \overline{L}, M\rangle \cdot \lambda$. Thus we deduce that if L is a $\widetilde{Q}$-local martingale and is P-orthogonal to M, then $\overline{B} = 0$. This ends the proof. □

In the following we elaborate the main result of this subsection.

Proposition 3.2.5. *The following assertions hold:*

1. *Let S and Y be two $\widetilde{Q}$-local martingales such that $[X,S]=0$, and there exists an $\widetilde{\mathcal{O}}$-measurable functional, g, such that $Y=g\star\mu$ with $M^{\widetilde{Q}}_{\mu}(g|\widetilde{\mathcal{P}})=0$. Then S (resp. Y) is a P-local martingale if and only if S (resp. Y) is P-locally integrable.*

2. *Let Z be a $\widetilde{Q}$-local martingale whose decomposition through Theorem 2.2.23 is given by*

$$\begin{aligned} Z =& Z_0+\beta\cdot X^{c,\widetilde{Q}}+W\star(\mu-\nu^{\widetilde{Q}})+g\star\mu+Z^{\perp} \\ W_t(x) :=& f_t(x)+\frac{\int f_t(y)\nu^{\widetilde{Q}}(\{t\},dy)}{1-\nu^{\widetilde{Q}}(\{t\},\mathbb{R}^d)}. \end{aligned}$$

   Then Z is a P-local martingale if and only if
   (a) The processes $(|f|\wedge|f|^2)\star\mu$, $g\star\mu$ and $Z^{\perp}$ are P-locally integrable, and
   (b) For $P(d\omega)dA_t(\omega)$-almost all (t,ω), we have

$$\lambda'_t c_t\beta_t+\int\left[\lambda'_t x-\lambda'_t\Delta\langle M\rangle_t\lambda_t\right]W_t(x)F_t(dx)=0. \tag{3.1}$$

Proof. 1. It is again trivial that if S (resp. Y) is a P-local martingale then S (resp. Y) is P-locally integrable.
Suppose that S and Y are P-locally integrable. If we denote by $\widetilde{Z}=(\mathcal{E}(\widetilde{N}))^{-1}=\mathcal{E}(\widetilde{X})$ the Girsanov density describing the change of measure from the MMM $\widetilde{Q}$ to the original measure P. Then $\widetilde{X}$ is the solution to the following equation due to (2.18)

$$1=\widetilde{Z}Z=\mathcal{E}(\widetilde{N})\mathcal{E}(\widetilde{N})^{-1}=\mathcal{E}(\widetilde{N})\mathcal{E}(\widetilde{X})=\mathcal{E}(\widetilde{N}+\widetilde{X}+[\widetilde{N},\widetilde{X}]).$$

Therefore

$$\widetilde{N}+\widetilde{X}+[\widetilde{N},\widetilde{X}]=0\Rightarrow\widetilde{X}=-\widetilde{N}-[\widetilde{N},\widetilde{X}]. \tag{3.2}$$

Substituting this result for $\widetilde{X}$ in $[\widetilde{N},\widetilde{X}]$ leads to

$$[\widetilde{N},\widetilde{X}]=-[\widetilde{N},\widetilde{N}]-[\widetilde{N},[\widetilde{N},\widetilde{X}]]=-[\widetilde{N},\widetilde{N}]-\Delta\widetilde{N}\cdot[\widetilde{N},\widetilde{X}],$$

while using Properties 2.2.15(1) and the fact that the process $[X,Y]$ belongs to the set $\mathcal{V}$, which is proven in Theorem I.4.47 of Jacod and Shiryaev (2002). In this way we get

$$[\widetilde{N},\widetilde{X}] = -\frac{1}{1+\Delta\widetilde{N}}\cdot[\widetilde{N},\widetilde{N}]$$

and hence relation (3.2) for $\widetilde{X}$ becomes

$$\widetilde{X} = -\widetilde{N} + \frac{1}{1+\Delta\widetilde{N}}\cdot[\widetilde{N},\widetilde{N}].$$

From the fact that the bracket process has finite variation combined with Lemma I.4.14 b) of Jacod and Shiryaev (2002), which states that a local martingale belonging to $\mathcal{V}$ is purely discontinuous, we get that for any X, Y belonging to $\mathcal{S}$: $[X,Y]^c \equiv 0$ and therefore by (2.4) we obtain that

$$\begin{aligned}[K,[\widetilde{N},\widetilde{N}]] =& \sum \Delta K\Delta[\widetilde{N},\widetilde{N}] = \sum \Delta K \Delta\widetilde{N}\Delta\widetilde{N}\\ =&[\widetilde{N},[K,\widetilde{N}]] = \Delta\widetilde{N}\cdot[K,\widetilde{N}],\end{aligned} \tag{3.3}$$

with $K\in\mathcal{S}$. Furthermore from Corollary 2.3.2, we know that any $\widetilde{Q}$-local martingale L, is also a P-local martingale if and only if $[L,\widetilde{X}]$ is a $\widetilde{Q}$-local martingale. Hence for any semimartingale K, we calculate using (3.3)

$$\begin{aligned}[K,\widetilde{X}] =&[K,-\widetilde{N}+\frac{1}{1+\Delta\widetilde{N}}\cdot[\widetilde{N},\widetilde{N}]]\\ =&\frac{1}{1+\Delta\widetilde{N}}\cdot\Big(-(1+\Delta\widetilde{N})\cdot[K,\widetilde{N}]+\Delta\widetilde{N}\cdot[K,\widetilde{N}]\Big)\\ =&-\Big(1+\Delta\widetilde{N}\Big)^{-1}\cdot[K,\widetilde{N}].\end{aligned} \tag{3.4}$$

Using the equation $\widetilde{N} = -\lambda\cdot M = -\lambda\cdot(X-\lambda\cdot\langle M,M\rangle)$ in (3.4) gives

$$\begin{aligned}[K,\widetilde{X}] =&\frac{\lambda}{1+\Delta\widetilde{N}}\cdot[K,X] - \frac{1}{1+\Delta\widetilde{N}}\cdot[K,\langle\lambda\cdot M\rangle]\\ =&\frac{\lambda}{1+\Delta\widetilde{N}}\cdot[K,X] - \frac{\lambda'\Delta\langle M\rangle\lambda}{1-\lambda'\Delta X+\lambda'\Delta\langle M\rangle\lambda}\cdot K,\end{aligned} \tag{3.5}$$

due to Properties 2.2.15(2) and the fact that the angle bracket process is predictable and has finite variation. Now suppose that K satisfies $[K,X]=0$, then (3.5) becomes

$$[K,\widetilde{X}] = -\frac{\lambda'\Delta\langle M\rangle\lambda}{1-\lambda'\Delta X+\lambda'\Delta\langle M\rangle\lambda}\cdot K.$$

Hence the process $[K, \widetilde{X}]$ is a $\widetilde{Q}$-local martingale when K is a $\widetilde{Q}$-local martingale with $[K, X] = 0$.
Now suppose that $K = g \star \mu$ with $M_\mu^{\widetilde{Q}}\left(g \mid \widetilde{\mathcal{P}}\right) = 0$. Then we get from (3.5)

$$\begin{aligned}[K, \widetilde{X}] &= \sum g(\Delta X) \frac{\lambda' \Delta X - \lambda' \Delta \langle M \rangle \lambda}{1 - \lambda' \Delta X + \lambda' \Delta \langle M \rangle \lambda} \mathbb{1}_{\{\Delta X \neq 0\}} = G \star \mu, \\ G(x) :&= g(x) \frac{\lambda' x - \lambda' \Delta \langle M \rangle \lambda}{1 - \lambda' x + \lambda' \Delta \langle M \rangle \lambda}.\end{aligned}$$

Since $\dfrac{-\lambda' x + \lambda' \Delta \langle M \rangle \lambda}{1 - \lambda' x + \lambda' \Delta \langle M \rangle \lambda}$ is bounded, we obviously get that

$$M_\mu^{\widetilde{Q}}\left(G \mid \widetilde{\mathcal{P}}\right)(t,x) = \frac{\lambda' x - \lambda' \Delta \langle M \rangle \lambda}{1 - \lambda' x + \lambda' \Delta \langle M \rangle \lambda} M_\mu^{\widetilde{Q}}\left(g \mid \widetilde{\mathcal{P}}\right)(t,x) = 0.$$

Thus, $K = g \star \mu$ is a P-local martingale.
2. The proof of this assertion will be outlined in two steps. The first step (parts 1), 2) and 3) below) will show that Z is P-locally integrable if and only if the assertion 2(a) holds, while the second step (part 4)) will prove that under the P-local integrability of Z, the P-compensator of Z is zero if and only if the assertion 2-(b) is satisfied.
1) We start by noticing that $(|f| \wedge |f|^2) \star \mu$ is a process with finite variation, since its $\widetilde{Q}$-compensator exists (see Theorem 2.2.23: $(|f| \wedge |f|^2) \star \nu_T^{\widetilde{Q}} < +\infty$, P-a.s.). Therefore, $(|f| \wedge |f|^2) \star \mu$ is P-locally integrable if and only if $|f| \mathbb{1}_{\{|f|>1\}} \star \mu$ is P-locally integrable, since the process $|f|^2 \mathbb{1}_{\{|f| \leq 1\}} \star \mu$ is a locally bounded process. We also recall a result that is crucial to prove this first step, namely Theorem VII.25 of Dellacherie and Meyer (1980). Thanks to this theorem, a semimartingale K is P-locally integrable if and only if the non-decreasing process $\sup_{s \leq \cdot} |\Delta K_s|$ is P-locally integrable (i.e. it belongs to $\mathcal{A}_{\text{loc}}^+(P)$). This is also equivalent to the fact that both processes $\sup\limits_{s \leq \cdot} \left[|\Delta K_s| \mathbb{1}_{\{\Delta X_s \neq 0\}}\right]$ and $\sup\limits_{s \leq \cdot} \left[|\Delta K_s| \mathbb{1}_{\{\Delta X_s = 0\}}\right]$ are P-locally integrable.
2) Due to (2.10) and the fact that $[Z^\perp, X] = 0$, the process

$$\begin{aligned}\sup_{s \leq t} \Big[|\Delta Z_s| \mathbb{1}_{\{\Delta X_s \neq 0\}}\Big] &= \sup_{s \leq t} \Big[|f_s(\Delta X_s) + g_s(\Delta X_s)| \mathbb{1}_{\{\Delta X_s \neq 0\}}\Big] \\ &= \sup_{s \leq t} |f_s(\Delta X_s) + g_s(\Delta X_s)| \mathbb{1}_{\{\Delta X_s \neq 0\}},\end{aligned}$$

is P-locally integrable if and only if the two processes

$$\sup_{s\leq t}|f_s(\Delta X_s)+g_s(\Delta X_s)|\mathbb{1}_{\{|f_s(\Delta X_s)|>1,\Delta X_s\neq 0\}},$$
$$\sup_{s\leq t}|f_s(\Delta X_s)+g_s(\Delta X_s)|\mathbb{1}_{\{|f_s(\Delta X_s)|\leq 1,\Delta X_s\neq 0\}}$$

are P-locally integrable.
It is obvious that $\sup_{s\leq t}|f_s(\Delta X_s)+g_s(\Delta X_s)|\mathbb{1}_{\{|f_s(\Delta X_s)|>1,\Delta X_s\neq 0\}}$ is P-locally integrable if and only if the process, $\sup_{s\leq t}|g_s(\Delta X_s)|\mathbb{1}_{\{|f_s(\Delta X_s)|>1,\Delta X_s\neq 0\}}$, is P-locally integrable or equivalently $g\mathbb{1}_{\{|f|\leq 1\}}\star\mu$ is P-locally integrable, since the latter process exists as semimartingale. Since the two processes $f\mathbb{1}_{\{|f|>1\}}\star\mu$ and $g\mathbb{1}_{\{|f|>1\}}\star\mu$ exist again as semimartingales, we deduce that

$$\sup_{s\leq t}|f_s(\Delta X_s)+g_s(\Delta X_s)|\mathbb{1}_{\{|f_s(\Delta X_s)|>1,\Delta X_s\neq 0\}}$$

is P-locally integrable if and only if $(f+g)\mathbb{1}_{\{|f|>1\}}\star\mu$ is P-locally integrable. From Lemma 2.3.7, we know that the P-compensator of the finite variation process $K:=(W+g)\star\mu$ coincides with the $\widetilde{Q}$-compensator of

$$\left(1+\Delta\widetilde{N}\right)^{-1}\cdot K=\frac{f+g}{1-\lambda' x+\lambda'\Delta\langle M\rangle\lambda}\mathbb{1}_{\{|f|>1\}}\star\mu,$$

which is given by

$$\begin{aligned}&M_\mu^{\widetilde{Q}}\left(\frac{f+g}{1-\lambda' x+\lambda'\Delta\langle M\rangle\lambda}\mid\widetilde{\mathcal{P}}\right)\mathbb{1}_{\{|f|>1\}}\star\nu^{\widetilde{Q}}\\&=\frac{W}{1-\lambda' x+\lambda'\Delta\langle M\rangle\lambda}\mathbb{1}_{\{|f|>1\}}\star\nu^{\widetilde{Q}}=f\mathbb{1}_{\{|f|>1\}}\star\nu,\end{aligned}$$

because from Theorem 2.2.23 we know that $M_\mu^{\widetilde{Q}}(g|\widetilde{\mathcal{P}})=0$ and we also used the relationship (2.24).
As a result, this proves that $(f+g)\star\mu$ is P-locally integrable if and only if both $|f|\mathbb{1}_{\{|f|>1\}}\star\mu$ and $g\mathbb{1}_{\{|f|>1\}}\star\mu$ are P-locally integrable. By combining all these conclusions we deduce that $|f|\mathbb{1}_{\{|f|>1\}}\star\mu$ and $g\star\mu$ should be P-locally integrable.
3) Now consider the following process

$$\sup_{s\leq t}\Big[|\Delta Z_s|\mathbb{1}_{\{\Delta X_s=0\}}\Big]=\sup_{s\leq t}\Big[|-\hat{W}_s^{\widetilde{Q}}+\Delta Z_s^{\perp}|\mathbb{1}_{\{\Delta X_s=0\}}\Big]. \tag{3.6}$$

Thanks to VIII.11 of Dellacherie and Meyer (1980), the process $\sup_{s\le t}|\hat{W}_s^{\widetilde{Q}}|$ is locally bounded, and hence the P-local integrability of $\sup_{s\le t}[|\hat{W}_s^{\widetilde{Q}}|1\!\!1_{\{\Delta X_s=0\}}]$ follows. This implies that the process in (3.6) is P-locally integrable if and only if $\sup\limits_{s\le t}|\Delta Z_s^{\perp}|$ is P-locally integrable, or equivalently $Z^{\perp}$ is P-locally integrable. By combining all these, we conclude that the first step of our proof for assertion 2. is achieved.

Thanks to assertion 1. and the first step, we deduce that –under assertion 2-(a)– Z is a P-local martingale if and only if

$$Z^{(1)} := \beta\cdot X^{c,\widetilde{Q}} + W\star(\mu-\nu^{\widetilde{Q}}),$$

has a null P-compensator. As a consequence the process $Z^{(1)}$ is P-locally integrable or equivalently the process $W\star(\nu-\nu^{\widetilde{Q}})$ makes sense. Hence since β is X^c-integrable (in the semimartingale sense), we obtain by using Theorem 2.3.4 and Theorem 2.3.5 that

$$W\star(\mu-\nu^{\widetilde{Q}}) = W\star(\mu-\nu)+W\star(\nu-\nu^{\widetilde{Q}}),\quad \beta\cdot X^{c,\widetilde{Q}} = \beta\cdot X^c+\lambda' c\beta\cdot A,$$

with $\nu^{\widetilde{Q}}$ given in (2.24). Then, these equations imply that $Z^{(1)}$ is a P-local martingale if and only if

$$0 = W\star(\nu-\nu^{\widetilde{Q}})+\lambda' c\beta\cdot A = [\lambda' x-\lambda'\Delta\langle M\rangle\lambda]\,W\star\nu+\lambda' c\beta\cdot A.$$

Therefore, (3.1) follows. This ends the proof of the proposition. □

3.3 The FS decomposition versus the GKW decomposition

Assumptions 3.3.1. *We assume that $\mathcal{E}(\widetilde{N}) = \mathcal{E}(-\lambda\cdot M) > 0$, and there exists a constant $C>0$ such that for any stopping time σ,*

$$E\left[\left(\mathcal{E}\left(\widetilde{N}-\widetilde{N}^{\sigma}\right)_T\right)^2 \mid \mathcal{F}_\sigma\right] \le C,\quad P\text{-a.s.} \tag{3.7}$$

This assumption, Doob's inequality and Theorem 2.2.3 imply that $\mathcal{E}(\widetilde{N})$ is a true martingale. In fact from (3.7) we can even deduce in the same way that for any

n, ${}^{T_n}\mathcal{E}(\widetilde{N}) := \mathcal{E}\left(\widetilde{N} - \widetilde{N}^{T_n}\right)$ is a true martingale, where

$$T_0 = 0, \quad T_{n+1} := \inf\{t > T_n \mid \Delta\widetilde{N}_t = -1\} \wedge T, \quad n \geq 0.$$

This is exactly the regularity property defined in Definition 3.1.5, since by (2.17) this sequence coincides with the sequence from Definition 3.1.2. The regularity is proved without using the assumption that $\mathcal{E}(\widetilde{N}) > 0$. Under this assumption the sequence equals $T_0 = 0$ and $T_n = T$ for $n \geq 1$.
By applying Jensen's inequality, see, e.g., Theorem I.19 Protter (2005), we obtain that

$$E\left[\left(\mathcal{E}\left(\widetilde{N} - \widetilde{N}^{\sigma}\right)_T\right)^2 \mid \mathcal{F}_\sigma\right] \geq E\left[\mathcal{E}\left(\widetilde{N} - \widetilde{N}^{\sigma}\right)_T \mid \mathcal{F}_\sigma\right]^2 = 1,$$

hence the constant C is ≥ 1 and the reverse Hölder inequality of order 2 is satisfied.
Therefore under Assumptions 3.3.1 we can assume that for any square-integrable $\mathcal{F}_T$-measurable claim the FS decomposition exists, see Theorem 3.1.7. Furthermore the MMM $\widetilde{Q}$ really exists and is defined by

$$\widetilde{Q} := \mathcal{E}(\widetilde{N})_T \cdot P. \tag{3.8}$$

We now determine the number of risky assets deduced from the GKW (resp. the FS) decomposition, denoted by ξ (resp. ξ^{FS}).
It is generally known that the number ξ for a claim H under the (minimal) martingale measure $\widetilde{Q}$ for X, with $\widetilde{V}_t = E^{\widetilde{Q}}[H|\mathcal{F}_t]$ is determined by

$$\xi_t = (d\langle X, X\rangle_t^{\widetilde{Q}})^{\mathrm{inv}} d\langle \widetilde{V}, X\rangle_t^{\widetilde{Q}}, \tag{3.9}$$

where we denote by inv the Moore-Penrose pseudoinverse, see Remarks 3.3.3 (1) for more details. This follows easily from Definition 2.4.1 of the GKW decomposition:

$$\begin{aligned}\widetilde{V}_t =& E^{\widetilde{Q}}[H|\mathcal{F}_t] = H_0 + E^{\widetilde{Q}}[(\xi \cdot X)_T|\mathcal{F}_t] + E^{\widetilde{Q}}[L_T|\mathcal{F}_t] \\ =& H_0 + (\xi \cdot X)_t + L_t.\end{aligned} \tag{3.10}$$

Now taking the angle bracket under $\widetilde{Q}$ of $\widetilde{V}$ with respect to X and using the orthogonality between L and X gives:

$$d\langle \widetilde{V}, X\rangle^{\widetilde{Q}} = \xi d\langle X, X\rangle^{\widetilde{Q}}.$$

Solving for ξ gives the result (3.9).
Analogously the number ξ^{FS} is found under the original measure P by considering the FS decomposition (Definition 3.1.1) for H and taking the bracket with respect to M:

$$\langle \widetilde{V}, M\rangle = \xi^{\text{FS}} \cdot \langle X, M\rangle = \xi^{\text{FS}} \cdot \langle M, M\rangle, \tag{3.11}$$

where we also used Properties 2.2.15(2). Therefore

$$\xi_t^{\text{FS}} = (d\langle M, M\rangle_t)^{\text{inv}} d\langle \widetilde{V}, M\rangle_t. \tag{3.12}$$

We remark that again by Properties 2.2.15(2), we only need the P-martingale part I of the $\widetilde{Q}$-martingale $\widetilde{V}$ in (3.12).
In equation (4.9) of Černý and Kallsen (2007) formula (3.12) was already obtained for the number of risky assets under the opportunity neutral measure, see Chapter 5, with $\widetilde{V}$ replaced by I. Due to this special measure they can prove the equality with

$$\xi_t^{\text{FS}} = (d\langle X, X\rangle_t)^{\text{inv}} d\langle \widetilde{V}, X\rangle_t.$$

Combining (3.11) with the angle bracket process under P of $\widetilde{V}$ in (3.10) with M, we can already deduce a simple relationship between ξ and ξ^{FS}:

$$\xi^{\text{FS}} \cdot \langle M, M\rangle = \langle I, M\rangle = \langle \widetilde{V}, M\rangle = \xi \cdot \langle X, M\rangle + \langle L, M\rangle,$$

and therefore

$$\xi^{\text{FS}} = \xi + (d\langle M, M\rangle)^{\text{inv}} d\langle L, M\rangle. \tag{3.13}$$

It is important to remark that splitting up $\widetilde{V}$ in the two terms $\xi \cdot X$ and L implies that $[L, M]$ exists as a P-special semimartingale, which guarantees the existence of the angle bracket $\langle L, M\rangle$. Unfortunately this will not always hold, see also Remark (5) under Theorem 3.3.2.
We illustrate (3.13) on a simple setting which is inspired by an example given in Tankov (2009).

Example: Let N^1 and N^2 be two independent Poisson processes with intensity 1 under the measure P. Assume $X_t = \gamma t + 2N_t^1 + N_t^2 - 3t$, with $\gamma \neq \frac{25}{9}$ and the claim H for which we want to determine the GKW and the FS decomposition equals $5N_T^1$. We first determine the MMM $\widetilde{Q}$ given by $\mathcal{E}(-\lambda \cdot M)$, which is a square-integrable and positive martingale with λ given by (2.5):

$$\begin{aligned} M =& 2(N_t^1 - t) + (N_t^2 - t), && (3.14)\\ [M,M]_t =& [2(N_t^1 - t) + (N_t^2 - t), 2(N_t^1 - t) + (N_t^2 - t)] = 4N_t^1 + N_t^2, \\ \langle M \rangle_t =& 5t, \quad \text{and} \quad \lambda_t = \frac{dB_t}{d\langle M \rangle_t} = \frac{\gamma t}{5t} = \frac{\gamma}{5}. && (3.15) \end{aligned}$$

Hence according to Theorem 2.3.5 under the MMM the following Poisson processes are $\widetilde{Q}$-martingales:

$$N_t^1 - (1 - \frac{2\gamma}{5})t \quad \text{and} \quad N_t^2 - (1 - \frac{\gamma}{5})t$$

and

$$X_t = 2[N_t^1 - (1 - \frac{2\gamma}{5})t] + N_t^2 - (1 - \frac{\gamma}{5})t, \tag{3.16}$$

while $[X,X]_t = 4N_t^1 + N_t^2$ and its compensator is given by

$$\langle X \rangle_t^{\widetilde{Q}} = 4(1 - \frac{2\gamma}{5})t + (1 - \frac{\gamma}{5})t = \frac{25 - 9\gamma}{5}t.$$

The $\widetilde{Q}$-martingale $\widetilde{V}$ is described by

$$\widetilde{V}_t = E^{\widetilde{Q}}[H|\mathcal{F}_t] = E^{\widetilde{Q}}[5N_T^1|\mathcal{F}_t] = 5(N_t^1 - t + \frac{2\gamma}{5}t) + 5(1 - \frac{2\gamma}{5})T \tag{3.17}$$

and the P-martingale part I of $\widetilde{V}$ is given by $5(N_t^1 - t)$. Now we can easily calculate ξ and ξ^{FS} using (3.9) and (3.12):

$$\begin{aligned} \xi_t^{\mathrm{FS}} =& \frac{d\langle I, M \rangle_t}{d\langle M \rangle_t} = \frac{d\langle 5(N_t^1 - t), 2N_t^1 + N_t^2 - 3t \rangle}{d5t} = \frac{10dt}{5dt} = 2, && (3.18)\\ \xi_t =& \frac{d\langle \widetilde{V}, X \rangle_t^{\widetilde{Q}}}{d\langle X \rangle_t^{\widetilde{Q}}} = \frac{d10(1 - \frac{2\gamma}{5})t}{d\frac{25-9\gamma}{5}t} = \frac{50 - 20\gamma}{25 - 9\gamma}. && (3.19) \end{aligned}$$

Hence we obviously see that ξ^{FS} differs from ξ unless $\frac{50 - 20\gamma}{25 - 9\gamma} = 2$, which only holds if $\gamma = 0$, thus if X is already a martingale under the original measure.

From (3.10), (3.17) and (3.16), we deduce that the process L equals

$$\begin{aligned} L_t =& 5N_t^1 - 5t + 2\gamma t - \frac{50-20\gamma}{25-9\gamma}(2N_t^1 - 2t + \frac{4\gamma}{5}t + N_t^2 - t + \frac{\gamma}{5}t) \\ =& (\frac{125-45\gamma}{25-9\gamma} - \frac{100-40\gamma}{25-9\gamma})N_t^1 - (\frac{50-20\gamma}{25-9\gamma})N_t^2 \\ &+ (-5+2\gamma - \frac{50-20\gamma}{25-9\gamma}(-3+\gamma))t. \end{aligned}$$

Using (3.14) and (3.15), we obtain that

$$\frac{d\langle L, M\rangle_t}{d\langle M, M\rangle_t} = \frac{\frac{25-5\gamma}{25-9\gamma}2t - \frac{50-20\gamma}{25-9\gamma}t}{5t} = \frac{\frac{10\gamma t}{25-9\gamma}}{5t} = \frac{2\gamma}{25-9\gamma}.$$

Hence adding this amount to (3.19), we should obtain the number ξ^{FS} according to (3.13)

$$\xi_t + \frac{d\langle L, M\rangle_t}{d\langle M, M\rangle_t} = \frac{50-20\gamma}{25-9\gamma} + \frac{2\gamma}{25-9\gamma} = \frac{50-18\gamma}{25-9\gamma} = 2,$$

as we already deduced in (3.18).

The goal of this section and Section 3.5 is to find more explicit formulas for the number of risky assets ξ^{FS} and to characterize the term $(d\langle M, M\rangle)^{\mathrm{inv}} d\langle L, M\rangle$ in terms of the components of the GKW decomposition. This is more logical because the FS decomposition is determined using the GKW decomposition and not the other way around.
The more explicit characterization of the FS decomposition using the predictable characteristics is very useful taking in mind the following remarks:

(1) Through the use of the predictable characteristics, the variation of the FS decomposition with additional jumps and/or uncertainty will be easy to handle. Furthermore, this illustration using the predictable characteristics is helpful in avoiding pitfalls and misleading generalizations of results such as those of Riesner (2006a) and Section 10.4 of Cont and Tankov (2004), see Chapter 7 for more details. Many practical market models are described using the predictable characteristics such as Barndorff-Nielsen-Shephard models, see Benth and Meyer-Brandis (2005) and Rheinländer and Steiger (2006) and the references therein about these models and related subjects. Hence, we think that this description of the FS decomposition will be useful for those models.

(2) Recently, the more explicitly characterized optimal martingale measures in the literature are expressed in terms of the predictable characteristics, see Choulli and Stricker (2005, 2006) and Choulli et al. (2007) for the semimartingale framework, and Benth and Meyer-Brandis (2005), Fujiwara and Miyahara (2003), Jeanblanc et al. (2007) and Kassberger and Liebmann (2007) for models driven by Lévy processes. Thus, we believe that the current description of the FS decomposition is suitable for those contexts.

(3) Finally, as it will be illustrated in Chapter 7, the description generalizes the approach of Colwell and Elliott (1993) and Vandaele and Vanmaele (2008b) to the semimartingale context where the predictable martingale representation may be violated on one hand. On the other hand the predictable characteristics are the extension of Lévy characteristics for models driven by semimartingales.

When X is a continuous process, it is generally known that the GWK and the FS decompositions coincide, because as is proved in Föllmer and Schweizer (1991) the minimal martingale measure preserves orthogonality in the continuous case. However this fact is no longer true in the general framework due to the presence of jumps in X. The correct relationship between the two decompositions is completely determined in the following theorem.

Theorem 3.3.2. *Let H be an $\mathcal{F}_T$-measurable random variable with $E(H^2) < +\infty$ and whose FS decomposition components are denoted by $\left(H_0, \xi^{FS}, L^{FS}\right)$. Suppose that the $\widetilde{Q}$-martingale, $\widetilde{V}_t = E^{\widetilde{Q}}\left(H|\, \mathcal{F}_t\right)$, admits the Galtchouk-Kunita-Watanabe decomposition which is given by*

$$\widetilde{V} = \widetilde{V}_0 + \xi \cdot X + L, \tag{3.20}$$

where ξ is a predictable and X-integrable process such that $\xi \cdot X$ and L are $\widetilde{Q}$-local martingales and L is $\widetilde{Q}$-orthogonal to X. Then the following holds:

(1) If $(\widetilde{\beta}, \widetilde{f}, \widetilde{g}, L^{\perp})$ denotes the quadruplet associated with L under $\widetilde{Q}$ through Theorem 2.2.23, then

$$\widetilde{\Phi} := \Sigma^{inv} \int x\widetilde{f}(x)[-\lambda' x + \lambda'\Delta\langle M\rangle\lambda]K(dx), \tag{3.21}$$

is a well-defined predictable process, X-integrable, and satisfies

$$\xi^{FS} = \xi - \widetilde{\Phi} \qquad L^{FS} = L + \widetilde{\Phi} \cdot X. \tag{3.22}$$

Here, Σ^{inv} denotes the Moore-Penrose pseudoinverse of the square matrix Σ, given by

$$\Sigma := c + \int xx'K(dx) = \frac{d\langle X\rangle}{dA}. \tag{3.23}$$

(2) *If there exists a sequence of stopping times $(T_n)_n$ increasing stationarily to T such that $\xi \mathbb{1}_{[0,T_n]} \in \Theta$, then the process $\langle \widetilde{N}, [L, X]\rangle$ exists, and is absolutely continuous with respect to $\langle X\rangle$ of which the Radon-Nikodym derivative is a version of $\widetilde{\Phi}$. Furthermore $\widetilde{\Phi} \in \Theta$.*

Remarks 3.3.3.

(1) The Moore-Penrose pseudoinverse Σ^{inv} is chosen such that $\Sigma\Sigma^{\text{inv}}\Sigma = \Sigma$ and $\Sigma^{\text{inv}}\Sigma\Sigma^{\text{inv}} = \Sigma^{\text{inv}}$, for more details see Albert (1972).

(2) From the proof of Proposition II.2.17 of Jacod and Shiryaev (2002) we deduce that

$$\langle X\rangle = (c + \int xx'K(dx)) \cdot A.$$

From this equation (3.23) easily follows. Furthermore we also deduce from Theorem 2.2.14 that $\langle X\rangle = \langle M\rangle + \sum \Delta B \Delta B'$. Hence if the semimartingale X is not quasi-left-continuous then $\Delta B = \Delta\langle M\rangle\lambda \neq 0$ and $\langle X\rangle$ differs from $\langle M\rangle$ in the following way:

$$\begin{aligned}\langle X\rangle - [B,B] =& \langle X\rangle - \sum (\Delta\langle M\rangle\lambda)(\Delta\langle M\rangle\lambda)' \\ =& \langle X\rangle - \lambda'\Delta\langle M\rangle\lambda \cdot \langle M\rangle = \langle M\rangle,\end{aligned} \tag{3.24}$$

by Properties 2.2.15(1). We remark that in Definition II.2.16 of Jacod and Shiryaev (2002) the angle bracket $\langle M\rangle$ is called the modified second characteristic of X. The relationship between $\langle M\rangle$ and $\langle X\rangle$ was already given in Proposition I.2.17.

(3) Suppose that there exists a sequence of stopping times T_n increasing stationarily to T, and a sequence of positive numbers, δ_n, such that $1 \geq \delta_n > 0$ and

$$\delta_n \leq 1 + \Delta\widetilde{N}^{T_n} \leq \delta_n^{-1}. \tag{3.25}$$

Then, under Assumptions 3.3.1, for any P-square-integrable and $\mathcal{F}_T$-measurable claim H the process $\widetilde{V}_t = E^{\widetilde{Q}}[H|\mathcal{F}_t]$ admits the GKW decomposition under $\widetilde{Q}$ in (3.20) and there exists a sequence of stopping times (σ_n) increasing stationarily to T such that $\sigma_n \leq T_n$ and $\xi \mathbb{1}_{[0,\sigma_n]} \in \Theta$. In other words, the assumption in assertion (2) of Theorem 3.3.2 is fulfilled. To prove this fact, we proceed into two steps: in the first step we will prove that $\widetilde{V}$ is a $\widetilde{Q}$-locally square-integrable martingale, while the second step will deal with $\xi \mathbb{1}_{[0,\sigma_n]} \in \Theta$. Indeed, due to Assumptions 3.3.1 the reverse Hölder inequality of order 2 is satisfied and by Theorem 4.9 (iii) of Choulli et al. (1998), we deduce that for the $\mathcal{E}^\tau$-martingale $\widetilde{V}$: $E[\widetilde{V}, \widetilde{V}]_T \leq CE[\widetilde{V}_T^2] < +\infty$, due to the P-square integrability of $\widetilde{V}$ on one hand. On the other hand, due to the RHS inequality in (3.25), we get

$$E[(1 + \Delta\widetilde{N}) \cdot [\widetilde{V}, \widetilde{V}]_{T_n}] \leq \delta_n^{-1} E[\widetilde{V}, \widetilde{V}]_{T_n} < +\infty.$$

This proves that the compensator of $(1 + \Delta\widetilde{N}) \cdot [\widetilde{V}, \widetilde{V}]^{T_n}$, which coincides with the $\widetilde{Q}$-compensator of $[\widetilde{V}, \widetilde{V}]^{T_n}$, see Lemma 2.3.7, exists and is integrable. Therefore $\widetilde{V}^{T_n}$ is a $\widetilde{Q}$-square integrable martingale. This ends the first step. Knowing this, it is obvious that the GKW decomposition under $\widetilde{Q}$ for $\widetilde{V}$ described in (3.20) exists.
Furthermore, since $\widetilde{V}^{T_n}$ is a $\widetilde{Q}$-square integrable martingale, the process $\xi \cdot X^{T_n}$ is a $\widetilde{Q}$-square integrable martingale. This follows easily by using the orthogonality between X and L under $\widetilde{Q}$. Notice that according to Lemma 2.3.7 the $\widetilde{Q}$-compensator of $[\xi \cdot X, \xi \cdot X]$ coincides with the P-compensator of $(1 + \Delta\widetilde{N}) \cdot [\xi \cdot X, \xi \cdot X]$. Hence there exists a sequence of stopping times (τ_n) increasing stationarily to T, such that the latter process stopped at τ_n is P-integrable. Then, due to $\delta_n \leq 1 + \Delta\widetilde{N}^{T_n}$, we get

$$E[\xi \cdot X, \xi \cdot X]_{\tau_n \wedge T_n} \leq \delta_n^{-1} E\left[(1 + \Delta\widetilde{N}) \cdot [\xi \cdot X, \xi \cdot X]_{\tau_n \wedge T_n}\right] < +\infty.$$

By combining this with Theorem 4.9 (iii) of Choulli et al. (1998), we conclude that $\xi \mathbb{1}_{[0,\tau_n \wedge T_n]} \in \Theta$ and the proof of the claim is achieved.

(4) In this theorem we extended the definition of the GKW-decomposition, defined in Definition 2.4.1 at time T, to variable times $t \in [0, T]$, by taking the expectation under $\widetilde{Q}$ of the both sides.

(5) If we consider the condition $\delta_n \leq 1 + \Delta\widetilde{N}^{T_n}$ instead of (3.25), then the results in Remark (3) are still valid for P-square-integrable $\mathcal{F}_T$-measurable

claims that are $\widetilde{Q}$-square integrable. In fact, for a $\widetilde{Q}$-square integrable claim H, the process $\widetilde{V}$ is a $\widetilde{Q}$-square integrable martingale. Hence the remaining part of the proof follows exactly the second step in the proof of Remark (2).

(6) The integrability of ξ described in Theorem 3.3.2 assertion (1) is enough to achieve our goal and to prove the main idea of the theorem which lies in describing the difference between the two decompositions. We remark that we cannot prove that $\xi \in \Theta$, due to the fact that the process $[L, M]$ may not be a P-local martingale, nor even a special semimartingale under P.

Proof of Theorem 3.3.2. (1) A key tool in this proof is Theorem 2.2.23 applied under the probability measure $\widetilde{Q}$. We start with describing the representation of X under this measure. First, the compensator of the random measure μ under $\widetilde{Q}$ will be denoted by $\nu^{\widetilde{Q}}$ and by (2.24) is given by

$$\nu^{\widetilde{Q}}(dt, dx) = (1 - \lambda' x + \lambda' \Delta\langle M\rangle\lambda)\,\nu(dt, dx). \tag{3.26}$$

Then, the process X takes the following canonical decomposition under $\widetilde{Q}$,

$$X = X_0 + X^{c,\widetilde{Q}} + x \star (\mu - \nu^{\widetilde{Q}}), \quad X^{c,\widetilde{Q}} := X^c + c\lambda \cdot A, \tag{3.27}$$

where the second equality follows from (2.23). We remark that the P-local martingale L^{FS} is also a $\widetilde{Q}$-local martingale by definition of the minimal martingale measure, see Definition 2.3.6. Applying Theorem 2.2.23 to the $\widetilde{Q}$-local martingale L^{FS}, provides

$$\begin{aligned} &L^{\text{FS}} = \theta^{\text{FS}} \cdot X^{c,\widetilde{Q}} + W^{\text{FS}} \star (\mu - \nu^{\widetilde{Q}}) + g^{\text{FS}} \star \mu + L^{\text{FS},\perp}, \\ &[L^{\text{FS},\perp}, X] = 0, \quad M^{\widetilde{Q}}_{\mu}(g^{\text{FS}}|\widetilde{\mathcal{P}}) = 0, \end{aligned} \tag{3.28}$$

where $W^{\text{FS}}(x) = f^{\text{FS}}(x) + \left(1 - \nu^{\widetilde{Q}}(\{t\}, \mathbb{R}^d)\right)^{-1} \int f^{\text{FS}}(x)\nu^{\widetilde{Q}}(\{t\}, dx) = f^{\text{FS}}(x) + \left(1 - a^{\widetilde{Q}}_t\right)^{-1} \hat{f}^{\text{FS},\widetilde{Q}}_t$.

Analogously, we find for the $\widetilde{Q}$-local martingale L

$$\begin{aligned} &L = \widetilde{\theta} \cdot X^{c,\widetilde{Q}} + \widetilde{W} \star (\mu - \nu^{\widetilde{Q}}) + \widetilde{g} \star \mu + L^{\perp}, \\ &[L^{\perp}, X] = 0, \quad M^{\widetilde{Q}}_{\mu}(\widetilde{g}|\widetilde{\mathcal{P}}) = 0, \end{aligned} \tag{3.29}$$

with $\widetilde{W}(x) = \widetilde{f}(x) + \left(1 - \nu^{\widetilde{Q}}(\{t\},\mathbb{R}^d)\right)^{-1} \int \widetilde{f}(x)\nu^{\widetilde{Q}}(\{t\},dx)$. Due to the integrability conditions on ξ^{FS} and L^{FS} in Definition 3.1.1 and the assumption on X, we deduce that $\xi^{\text{FS}} \cdot X$ and L^{FS} are martingales under $\widetilde{Q}$, and

$$H_0 + \xi^{\text{FS}} \cdot X + L^{\text{FS}} = \widetilde{V}_0 + \xi \cdot X + L. \tag{3.30}$$

Notice that from (3.27), we get $\xi^{\text{FS}} \cdot X = \xi^{\text{FS}} \cdot X^{c,\widetilde{Q}} + x'\xi^{\text{FS}} \star (\mu - \nu^{\widetilde{Q}})$ and $\xi \cdot X = \xi \cdot X^{c,\widetilde{Q}} + x'\xi \star (\mu - \nu^{\widetilde{Q}})$. By plugging these two equations together with (3.28) and (3.29) into (3.30), we conclude that the two processes $H_0 + (\xi^{\text{FS}} + \theta^{\text{FS}}) \cdot X^{c,\widetilde{Q}} + (x'\xi^{\text{FS}} + W^{\text{FS}}) \star (\mu - \nu^{\widetilde{Q}}) + g^{\text{FS}} \star \mu + L^{H,\perp}$ and $\widetilde{V}_0 + (\xi + \widetilde{\theta}) \cdot X^{c,\widetilde{Q}} + (x'\xi + \widetilde{W}) \star (\mu - \nu^{\widetilde{Q}}) + \widetilde{g} \star \mu + L^{\perp}$ are identical. Therefore, due to the uniqueness of Jacod's decomposition (see Lemma 2.2.24), we derive $H_0 = \widetilde{V}_0$, $g^{\text{FS}}(x) = \widetilde{g}(x)$, $L^{\text{FS},\perp} = L^{\perp}$, and

$$\begin{aligned} c\xi + c\widetilde{\theta} &= c\xi^{\text{FS}} + c\theta^{\text{FS}} && P \otimes dA\text{-a.e.},\\ x'\xi + \widetilde{f}(x) &= x'\xi^{\text{FS}} + f^{\text{FS}}(x) && K(dx) \times dA\text{-a.e..} \end{aligned} \tag{3.31}$$

Since L^{FS} is a P-local martingale orthogonal to M, we know that $\langle L^{\text{FS}}, M\rangle = 0$. First we show that this implies that $(c\theta^{\text{FS}} + \int x f^{\text{FS}}(x)K(dx)) \cdot A = 0$. Note that by Properties 2.2.15(1) $\langle L^{\text{FS}}, M\rangle = \langle L^{\text{FS}}, X\rangle$ and that for a $\widetilde{Q}$-martingale X the process $\int_{\mathbb{R}^d} x\nu^{\widetilde{Q}}(\{t\} \times dx) = \Delta B^{\widetilde{Q}} = 0$, see formula (3.77) in Jacod (1979). We now further transform $[L^{\text{FS}}, X]$ to determine its compensator $\langle L^{\text{FS}}, X\rangle$.

$$\begin{aligned} [L^{\text{FS}}, X] &\overset{(3.28)}{=} [\theta^{\text{FS}} \cdot X^{c,\widetilde{Q}} + W^{\text{FS}} \star (\mu - \nu^{\widetilde{Q}}) + g^{\text{FS}} \star \mu + L^{\text{FS},\perp}, X]\\ &\overset{(3.28)}{=} \theta^{\text{FS}} \cdot [X^{c,\widetilde{Q}}, X^{c,\widetilde{Q}}] + [W^{\text{FS}} \star (\mu - \nu^{\widetilde{Q}}), x \star (\mu - \nu^{\widetilde{Q}})]\\ &\quad + [g^{\text{FS}} \star \mu, x \star (\mu - \nu^{\widetilde{Q}})]\\ &\overset{(2.4)}{=} c\theta^{\text{FS}} \cdot A + \sum \Delta(W^{\text{FS}} \star (\mu - \nu^{\widetilde{Q}}))\Delta(x \star (\mu - \nu^{\widetilde{Q}}))\\ &\quad + \sum \Delta(g^{\text{FS}} \star \mu)\Delta(x \star (\mu - \nu^{\widetilde{Q}}))\\ &\overset{(2.10)}{=} c\theta^{\text{FS}} \cdot A + \sum (f^{\text{FS}}\mathbb{1}_{\{\Delta X \neq 0\}} - \frac{\hat{f}^{\text{FS},\widetilde{Q}}}{1 - a^{\widetilde{Q}}}\mathbb{1}_{\{\Delta X = 0\}})(x\mathbb{1}_{\{\Delta X \neq 0\}})\\ &\quad + \sum (g^{\text{FS}}\mathbb{1}_{\{\Delta X \neq 0\}})(x\mathbb{1}_{\{\Delta X \neq 0\}})\\ &= c\theta^{\text{FS}} \cdot A + f^{\text{FS}}x \star \mu + g^{\text{FS}}x \star \mu, \end{aligned} \tag{3.32}$$

where $g^{\text{FS}}x \star \mu$ is a $\widetilde{Q}$-local martingale because $M_{\mu}^{\widetilde{Q}}(g^{\text{FS}}|\widetilde{\mathcal{P}}) = 0$ and where in the second step we also used Properties 2.2.15(3). By a same reasoning as done for

$\widetilde{V}$ in Remarks (2) on page 54 the P-local integrability of $g^{\mathrm{FS}}x \star \mu$ follows, hence we obtain from Proposition 3.2.5 (i) that $g^{\mathrm{FS}}x \star \mu$ is also a P-martingale. Thus

$$\langle L^{\mathrm{FS}}, X\rangle = (c\theta^{\mathrm{FS}} + \int_{\mathbb{R}^d} x f^{\mathrm{FS}}(x)K(dx)) \cdot A = 0. \tag{3.33}$$

In an analogous way we calculate the $\widetilde{Q}$-compensator of $[L, X]$, using (3.26) to conclude that dA-a.e.:

$$c\theta^{\mathrm{FS}} + \int x f^{\mathrm{FS}}(x)K(dx) = 0, \quad c\widetilde{\theta} + \int x\widetilde{f}(x)[1 - \lambda' x + \lambda'\Delta\langle M\rangle\lambda]K(dx) = 0. \tag{3.34}$$

The second equation in (3.31) leads to

$$\int xx'\xi K(dx) + \int x\widetilde{f}(x)K(dx) = \int xx'\xi^{\mathrm{FS}}K(dx) + \int x f^{\mathrm{FS}}(x)K(dx).$$

By adding this to the first equation of (3.31), taking (3.34) into account, and putting $\Sigma_t := c_t + \int xx' F_t(dx)$, we obtain

$$\Sigma\xi^{\mathrm{FS}} = \Sigma\xi + \int x\widetilde{f}(x)\,[\lambda' x - \lambda'\Delta\langle M\rangle\lambda]\,K(dx).$$

Therefore we conclude that the process $\widetilde{\Phi}$ defined in (3.21) is a well-defined predictable process. It is also X-integrable, since ξ^{FS} and ξ are X-integrable, and it satisfies the first equation in (3.22). The second equation of (3.22) follows from inserting the first equation of (3.22) in (3.30). This ends the proof of assertion (1).
(2) Since $\xi 1\!\!1_{[0,T_n]} \in \Theta$ and $\sup_{t\leq\cdot} |\widetilde{V}_t|^2 \in \mathcal{A}^+(P)$ by Proposition 2.8 and Theorem 4.9 of Choulli et al. (1998), we deduce that the process $\sup_{s\leq\cdot} |L_s|^2 \in \mathcal{A}^+_{\mathrm{loc}}$. Applying the Kunita-Watanabe inequality, see Theorem II.25 in Protter (2005) , to $[L, X]$ and since $0 \leq (A - B)^2$ we get

$$[L, X] \leq [L, L]^{\frac{1}{2}}[X, X]^{\frac{1}{2}} \leq [L, L] + [X, X].$$

Thus it follows that the process $[L, X]$ has P-locally integrable variation. Furthermore due to the orthogonality between L and X, we know that $[L, X]$ is a $\widetilde{Q}$-local martingale. This implies by Remark 2.3.3 that $[L, X] + [\widetilde{N}, [L, X]]$ is a P-local martingale and hence the process $[\widetilde{N}, [L, X]]$ is a P-semimartingale and has P-locally integrable variation. Therefore its compensator $\langle \widetilde{N}, [L,X]\rangle$ exists. Invoking (3.3), we obtain:

$$\begin{aligned}[\widetilde{N},[L,X]] =&\sum \Delta\widetilde{N}\Delta L\Delta X\\ =&\sum -\lambda'(\Delta X-\Delta B)(\widetilde{f}(\Delta X)+\widetilde{g}(\Delta X))\Delta X\mathbb{1}_{\{\Delta X\neq 0\}}\\ =&\sum \lambda'(-\Delta X+\Delta\langle M\rangle\lambda)(\widetilde{f}(\Delta X)+\widetilde{g}(\Delta X))\Delta X\mathbb{1}_{\{\Delta X\neq 0\}}\\ =&\lambda'(-x+\Delta\langle M\rangle\lambda)(\widetilde{f}(x)+\widetilde{g}(x))x\star\mu.\end{aligned}$$

Again from Theorem 3.2.5(i), we conclude that using (3.23)

$$\begin{aligned}\langle\widetilde{N},[L,X]\rangle =&\left\{\int x\widetilde{f}(x)\Big[-\lambda'x+\lambda'\Delta\langle M\rangle\lambda\Big]K(dx)\right\}\cdot A \qquad (3.35)\\ =&\Sigma^{\text{inv}}\left\{\int x\widetilde{f}(x)\Big[-\lambda'x+\lambda'\Delta\langle M\rangle\lambda\Big]K(dx)\right\}\cdot\langle X\rangle.\end{aligned}$$

Thus $\widetilde{\Phi}$ is a version of the Radon-Nikodym derivative $\widetilde{\Psi}$ of $\langle\widetilde{N},[L,X]\rangle$ with respect to $\langle X\rangle$. By version we mean $\Sigma\widetilde{\Psi}=\Sigma\widetilde{\Phi}$ or equivalently $\widetilde{\Phi}-\widetilde{\Psi}\in$ kernel(Σ). This completes the proof of the theorem. □

Remarks 3.3.4. (1) From (3.35) we conclude that the process $\widetilde{\Phi}$ can also be explained as the Radon-Nikodym derivative of $\Sigma^{\text{inv}}d\langle[L,X],\widetilde{N}\rangle$ with respect to dA, that is

$$\widetilde{\Phi}=\Sigma^{\text{inv}}\frac{d\langle[L,X],\widetilde{N}\rangle}{dA}.$$

(2) Through Theorem 3.3.2, we can easily claim that the two decompositions – FS decomposition and GKW decomposition – are equivalent when X is a continuous process. Indeed, in this case, both processes $[L,X]$ and $\widetilde{\Phi}$ vanish, implying that $L=L^{\text{FS}}$ and $\xi=\xi^{\text{FS}}$.

(3) This theorem also allows us to decide whether the two decompositions coincide or differ for any $\mathcal{F}_T$-measurable random variable and market model through the following statement: The two decompositions coincide if and only if

$$E\left[\int_0^T \mathbb{1}_{\left\{(\omega,t)\,:\,\Sigma_t^{\text{inv}}(\omega)\Lambda_t(\omega)\notin\text{ kernel}\left(\Sigma_t(\omega)\right)\right\}}dA_t\right]=0,$$

with $\Lambda := \int x\widetilde{f}(x)\Big[-\lambda' x + \lambda'\Delta\langle M\rangle\lambda\Big]K(dx)$.

(4) Combining the result of Theorem 3.3.2 with the results of Theorem 3.2.5 we can shown that $\widetilde{\Phi}$ is a null process if and only if L (the martingale component in the GKW decomposition of $\widetilde{V}$ under $\widetilde{Q}$) is P-orthogonal to M. Indeed, notice that $[L, X]$ is a $\widetilde{Q}$-local martingale if and only if $[L, M]$ is a $\widetilde{Q}$-local martingale if and only if

$$0 = \langle L, M\rangle + \langle \widetilde{N}, [\widetilde{L}, M]\rangle. \tag{3.36}$$

This follows from Remark 2.3.3 and the fact that $[L, M] - \langle L, M\rangle$ is by definition of the angle bracket a P-martingale.
Now we calculate

$$\begin{aligned}\widetilde{\Phi}\cdot\langle X\rangle =&\langle \widetilde{N}, [L, X]\rangle = \langle \widetilde{N}, [L, M]\rangle - \langle \lambda'\cdot M, \Delta B\cdot L\rangle\\ =&\langle \widetilde{N}, [L, M]\rangle - \lambda'\Delta\langle M\rangle\lambda\cdot\langle M, L\rangle.\end{aligned}$$

Therefore by inserting this equation into (3.36), we obtain

$$0 = (1 + \lambda'\Delta\langle M\rangle\lambda)\cdot\langle L, M\rangle + \widetilde{\Phi}\cdot\langle X\rangle.$$

Thus, $\widetilde{\Phi}$ is a null process if and only if $\langle L, M\rangle \equiv 0$. This ends the proof of the claim.

3.4 A practical counterexample

In this section, we construct an example for which we can really prove that the GKW and the FS decomposition differ. Hence, this proves that the results in Riesner (2006a) – which are based on the fact that the FS decomposition and the GKW decomposition under the minimal martingale measure coincide – are wrong. Also in Section 10.4 of the book Cont and Tankov (2004) the same error appears, but this will be corrected in the forthcoming second edition of the book. In Chapter 7 we will look in more detail to these two problems.

Consider the following one-dimensional discounted process

$$X_t := X_0\mathcal{E}(\overline{X})_t, \quad \overline{X}_t := \sigma W_t + \gamma\widetilde{p}_t + \mu t, \quad 0 \le t \le T, \tag{3.37}$$

where $(p_t)_{t\geq 0}$ is the standard Poisson process with intensity 1, $\widetilde{p}_t = p_t - t$ is the compensated Poisson process, W_t is the standard Brownian motion, $X_0 > 0$, $\sigma > 0$, and γ and μ are real numbers such that

$$\gamma > -1 \qquad \text{and} \qquad 0 \neq \mu\gamma < \sigma^2 + \gamma^2. \tag{3.38}$$

Hence the process X belongs to the class of Lévy processes, see Section 2.5. The process X represents the discounted stock price process that constitutes the market model. Then, the processes M, B, and A (defined in (2.2) and in Proposition 2.2.20 respectively) for this model are given by

$$dM_t = X_{t-}\left(\sigma dW_t + \gamma d\widetilde{p}_t\right), \quad dB_t = \mu X_{t-}dt, \quad A_t = t.$$

Hence, we deduce from formula (2.5) that

$$\lambda_t = \frac{1}{X_{t-}}\frac{\mu}{\sigma^2+\gamma^2}, \quad \widetilde{N}_t = \sigma_1 W_t + \gamma_1 \widetilde{p}_t, \quad \sigma_1 := \frac{-\mu\sigma}{\sigma^2+\gamma^2}, \quad \gamma_1 := \frac{-\mu\gamma}{\sigma^2+\gamma^2}.$$

Thus for the model described in (3.37) and under the assumptions (3.38), we deduce by using (2.17) that $\mathcal{E}(\widetilde{N})$ is a square-integrable and positive martingale. Hence the minimal martingale measure exists and is given by $\widetilde{Q} := \mathcal{E}(\widetilde{N})_T \cdot P$.

Now consider the European put option with strike price K whose payoff is given by $H = (K - X_T)^+$. In the following we will calculate the processes $\widetilde{V}$, ξ, $\widetilde{\Phi}$ and L. Due to the independent increments of X, see Definition 2.5.2, we deduce that $\widetilde{V}_t = f(t, X_t)$, where

$$f(t,x) = E^{\widetilde{Q}}\left[\left(K - x\frac{X_T}{X_t}\right)^+\right]. \tag{3.39}$$

To calculate the distribution function of X, we first determine the process X more explicitly by using formula (2.17):

$$\begin{aligned} X_t =& X_0\mathcal{E}(\overline{X})_t = S_0\mathcal{E}(\sigma W + \gamma\widetilde{p} + \mu t)_t \\ =& X_0 e^{\sigma W_t + \gamma\widetilde{p}_t + \mu t - \frac{1}{2}\sigma^2 t}\prod_{s\leq t}(1+\gamma\Delta\widetilde{p}_t)e^{-\gamma\Delta\widetilde{p}_t}. \end{aligned} \tag{3.40}$$

Furthermore the jumps in $\widetilde{p}$ equal the jumps in p: $\Delta\widetilde{p} = \Delta p - \Delta t = \Delta p$ and because p is a Poisson process with intensity one, we obtain that $(1+\gamma\Delta\widetilde{p}_t)e^{-\gamma\Delta\widetilde{p}_t}$ is 1 if there is no jump and $(1+\gamma)e^{-\gamma}$ if there is a jump. Therefore

$$\prod_{s\leq t}(1+\gamma\Delta\widetilde{p}_t)e^{-\gamma\Delta\widetilde{p}_t} = [(1+\gamma)e^{-\gamma}]^{\sum \Delta p_s} = e^{p_t(\log(1+\gamma)-\gamma)}.$$

Inserting this result in formula (3.40) leads to

$$\begin{aligned}X_t =& X_0 e^{\sigma W_t + \gamma \widetilde{p}_t + \mu t - \frac{1}{2}\sigma^2 t + (\widetilde{p}_t + t)[\log(1+\gamma)-\gamma]}\\ =& X_0 e^{\sigma W_t + \widetilde{p}_t(\log(1+\gamma)) + (\mu - \frac{1}{2}\sigma^2 + \log(1+\gamma)-\gamma)t}.\end{aligned}$$

So the strictly increasing – in the variable y– distribution function is given by

$$\begin{aligned}F(s, y = \log(x)) :=& \widetilde{Q}(\frac{X_s}{X_0} \leq x) = \widetilde{Q}(\log(X_s) - \log(X_0) \leq \log(x) = y)\\ =& \widetilde{Q}\left(\sigma W_s + \log(1+\gamma)\widetilde{p}_s + \overline{\mu} s \leq y\right),\end{aligned} \tag{3.41}$$

with

$$\overline{\mu} := \mu - \frac{1}{2}\sigma^2 + \log(1+\gamma) - \gamma, \quad y \in \mathbb{R}, \;\; s \in [0, T].$$

Thanks to the stationary property of X, see Definition 2.5.2 and the notation in (3.41), the function $f(t, x)$ in (3.39) takes the following form for every $x > 0$

$$\begin{aligned}f(t, x) =& E^{\widetilde{Q}}[(K - x\frac{X_T}{X_t})^+] = xE^{\widetilde{Q}}[(\frac{K}{x} - \frac{\mathcal{E}(\overline{X})_T}{\mathcal{E}(\overline{X})_t})^+]\\ =& xE^{\widetilde{Q}}[(\frac{K}{x} - \mathcal{E}(\overline{X})_{T-t})^+] = x\int_{-\infty}^{\log\frac{K}{x}} (\frac{K}{x} - e^y)dF(T-t, y)\\ =& \frac{xK}{x}[F(T-t, \log\frac{K}{x}) - F(T-t, -\infty)] - x\int_{-\infty}^{\log\frac{K}{x}} e^y F_y(T-t, y)dy\\ =& KF\Big(T-t, \log\frac{K}{x})\Big) - x\int_{-\infty}^{\log(\frac{K}{x})} e^y F_y(T-t, y)dy, \quad t \in [0, T].\end{aligned} \tag{3.42}$$

As a result, $f(t, x) \in C^{1,2}((0, T) \times (0, +\infty))$, and by applying Itô's formula, see Theorem 2.2.26, to $f(t, X_t)$ we derive

$$\begin{aligned}\widetilde{V}_t = \widetilde{V}_0 + \int_0^t f_t(u, X_{u-})du + \int_0^t f_x(u, X_{u-})dX_u + \frac{1}{2}\int_0^t f_{xx}(u, X_{u-})X_u^2\sigma^2 du\\ + \sum_{0<u\leq t}\left[f(u, X_u) - f(u, X_{u-}) - f_x(u, X_u)\Delta X_u\right].\end{aligned}$$

Remark that

$$
\begin{aligned}
&\sum_{0<u\le t}\left[f(u,X_u)-f(u,X_{u-})-f_x(u,X_u)\Delta X_u\right]\\
&=\sum_{0<u\le t}\left[f(u,X_{u-}(1+\gamma\Delta p_u))-f(u,X_{u-})-f_x(u,X_u)\gamma X_{u-}\Delta p_u\right]\\
&=\sum_{0<u\le t}\left[f(u,X_{u-}(1+\gamma))-f(u,X_{u-})-f_x(u,X_u)\gamma X_{u-}\right]\Delta p_u\\
&=\Gamma\cdot p,
\end{aligned}
$$

where

$$\Gamma_u := f(u,X_{u-}(1+\gamma))-f(u,X_{u-})-f_x(u,X_u)\gamma X_{u-}. \tag{3.43}$$

Thus, since $\widetilde{V}$ is a $\widetilde{Q}$-martingale, we deduce that the function $f(t,x)$ satisfies a PDE equation (a fact that can be verified directly since the function $f(t,x)$ is explicitly calculated in (3.42)), and

$$\widetilde{V}_t=\widetilde{V}_0+\int_0^t f_x(u,X_{u-})dX_u+\left(\Gamma\cdot\widetilde{p}^{\widetilde{Q}}\right)_t,\quad p_t^{\widetilde{Q}}:=p_t-(1+\gamma_1)t. \tag{3.44}$$

Here $\widetilde{p}^{\widetilde{Q}}$ is the compensated Poisson process under $\widetilde{Q}$ which is determined by Theorem 2.3.4. Now we will focus on calculating ξ using (3.9) and (3.44):

$$\xi_t=\frac{d\langle\widetilde{V},X\rangle_t^{\widetilde{Q}}}{d\langle X,X\rangle_t^{\widetilde{Q}}}=f_x\frac{d\langle X,X\rangle_t^{\widetilde{Q}}}{d\langle X,X\rangle_t^{\widetilde{Q}}}+\Gamma_t\frac{d\langle\widetilde{p}^{\widetilde{Q}},X\rangle_t^{\widetilde{Q}}}{d\langle X,X\rangle_t^{\widetilde{Q}}}$$

Furthermore by (3.40)

$$d\langle X,X\rangle_t^{\widetilde{Q}}=X_{t-}^2(\sigma^2+\gamma^2(1+\gamma_1))dt.$$

and recalling that the compensator of p_t under $\widetilde{Q}$ coincides with $(1+\gamma_1)t$, we obtain

$$
\begin{aligned}
[\widetilde{p}^{\widetilde{Q}},X] &=\sum_{s\le\cdot}\Delta\widetilde{p}_s^{\widetilde{Q}}\Delta X_s=X_-\gamma\cdot p\\
d\langle\widetilde{p}^{\widetilde{Q}},X\rangle_t^{\widetilde{Q}} &=X_{t-}\gamma(1+\gamma_1)dt.
\end{aligned}
$$

Therefore we derive the components of the GKW decomposition under $\widetilde{Q}$ for $\widetilde{V}$ as follows using (3.44):

$$\xi_t = f_x(t, X_{t-}) + \frac{\Gamma_t \gamma(1+\gamma_1)}{X_{t-}[\sigma^2 + \gamma^2(1+\gamma_1)]}, \quad L = \Gamma \cdot \widetilde{p}^{\widetilde{Q}} - \frac{\gamma\Gamma(1+\gamma_1)}{X_-[\sigma^2+\gamma^2(1+\gamma_1)]} \cdot X. \tag{3.45}$$

This allows us to state the following.

Corollary 3.4.1. *Consider the model described by (3.37)-(3.38). Then the following assertions hold:*

(1) The GKW decomposition of $\widetilde{V}$ under $\widetilde{Q}$ is given by

$$\widetilde{V} = \widetilde{V}_0 + \xi \cdot X + L,$$

where ξ and L are given by (3.45).

(2) The FS decomposition of H and the GKW decomposition under $\widetilde{Q}$ for $\widetilde{V}$ differ.

Proof. The first assertion is already proved, while the second assertion will follow after proving that the process $\widetilde{\Phi}$ defined in Theorem 3.3.2 for this model never vanishes. The calculation of this process requires the calculation of $[L, X]$ and $[[L, X], \widetilde{N}]$. Due to (3.45), these processes are given by

$$\begin{aligned} d[L,X]_t =& d[\Gamma \cdot \widetilde{p}^{\widetilde{Q}} - \frac{\gamma\Gamma(1+\gamma_1)}{X_-[\sigma^2+\gamma^2(1+\gamma_1)]} \cdot X, X]_t \\ =& \Gamma_t \gamma X_{t-} dp_t - \frac{\gamma\Gamma_t(1+\gamma_1)}{X_{t-}[\sigma^2+\gamma^2(1+\gamma_1)]} X_{t-}^2(\sigma^2 dt + \gamma^2 dp_t) \\ =& (\Gamma_t\gamma X_{t-} - \frac{\gamma\Gamma_t(1+\gamma_1)}{\sigma^2+\gamma^2(1+\gamma_1)} X_{t-}\gamma^2) dp_t - \frac{\gamma\Gamma_t(1+\gamma_1)}{\sigma^2+\gamma^2(1+\gamma_1)} \sigma^2 X_{t-} dt \\ =& \frac{\gamma\Gamma_t\sigma^2 X_{t-}}{\sigma^2+\gamma^2(1+\gamma_1)} dp_t - \frac{\gamma\Gamma_t\sigma^2(1+\gamma_1)X_{t-}}{\sigma^2+\gamma^2(1+\gamma_1)} dt, \end{aligned}$$

and so using Properties 2.2.15(3) and (4):

$$\begin{aligned} [[L,X],\widetilde{N}] =& [\frac{\gamma\Gamma\sigma^2 X_-}{\sigma^2+\gamma^2(1+\gamma_1)} \cdot p - \frac{\gamma\Gamma\sigma^2(1+\gamma_1)X_-}{\sigma^2+\gamma^2(1+\gamma_1)} \cdot t, \gamma_1 \cdot p] \\ =& \frac{\gamma_1\gamma\Gamma\sigma^2 X_-}{\sigma^2+\gamma^2(1+\gamma_1)} \cdot p. \end{aligned}$$

As a result, we derive

$$\widetilde{\Phi}_t = \frac{d\langle \widetilde{N}, [L,X]\rangle_t}{d\langle X,X\rangle_t} = \frac{-\mu\gamma^2\sigma^2}{(\gamma^2+\sigma^2)^2(\sigma^2+\gamma^2(1+\gamma_1))}\frac{\Gamma_t}{X_{t-}}. \tag{3.46}$$

Then, by putting $s_1(t) := \log\left(\frac{K}{X_{t-}}\right)$ and $s_2(t) := \log\left(\frac{K}{X_{t-}(1+\gamma)}\right)$, and using

$$\begin{aligned} f_x(t,x) =& KF_y(T-t,\log(K/x))\frac{\partial \log(K/x)}{\partial x} - \int_{-\infty}^{\log(K/x)} e^y F_y(T-t,y)dy \\ & - xe^{\log(K/x)}F_y(T-t,\log(K/x))\frac{\partial \log(K/x)}{\partial x} \\ =& -\int_{-\infty}^{\log(K/x)} e^y F_y(T-t,y)dy, \end{aligned} \tag{3.47}$$

we obtain the following for the process Γ defined in (3.43) by inserting (3.42) and (3.47):

$$\begin{aligned} \Gamma_t =& f(t,X_{t-}(1+\gamma)) - f(t,X_{t-}) - f_x(t,X_{t-})X_{t-}\gamma \\ =& KF(T-t,s_2(t)) - KF(T-t,s_1(t)) - X_{t-}(1+\gamma)\int_{-\infty}^{s_2(t)} e^y F_y(T-t,y)dy \\ & + X_{t-}\int_{-\infty}^{s_1(t)} e^y F_y(T-t,y)dy + X_{t-}\gamma\int_{-\infty}^{s_1(t)} e^y F_y(T-t,y)dy \\ =& \int_{s_1(t)}^{s_2(t)} \Big[K - X_{t-}(1+\gamma)e^y\Big] F_y(T-t,y)dy. \end{aligned}$$

If $(-1<)\gamma<0$, then $s_1<s_2$ and $\Big[K-S_{t-}(1+\gamma)e^y\Big]F_y(T-t,y)>0$, on the other hand when $\gamma>0$ then $s_2<s_1$ and

$$\begin{aligned} &\int_{s_1(t)}^{s_2(t)} \Big[K - S_{t-}(1+\gamma)e^y\Big] F_y(T-t,y)dy \\ =& \int_{s_2(t)}^{s_1(t)} \Big[S_{t-}(1+\gamma)e^y - K\Big] F_y(T-t,y)dy > 0. \end{aligned}$$

This proves that Γ is a positive process if $\gamma \neq 0$. By (3.46) the process $\widetilde{\Phi}$ then also has a constant sign and never vanishes under the conditions (3.38). Therefore, ξ and ξ^{FS} (see Definition 3.1.1) never coincide and hence the FS decomposition and the GKW decomposition under $\widetilde{Q}$ differ for this model. □

3.5 Determination of the Föllmer-Schweizer decomposition

This section proposes a description of the FS decomposition – under some integrability conditions that guarantee the existence of this decomposition – in terms of the predictable characteristics of X. The following assumptions hold throughout the whole section.

Assumptions 3.5.1. *We assume that there exists a constant $C > 0$ such that (3.7) holds.*

Remark 3.5.2. It is obvious that Assumptions 3.5.1 are weaker than Assumptions 3.3.1. That is in this section, the minimal martingale measure may not exist as a measure, and/or its density may vanish. This is an interesting generalization, especially when one is working with models that involve jumps such as Lévy market models. In our view, the integrability condition of (3.7) is less restrictive than the positivity of $\mathcal{E}(\widetilde{N})$. In many models considered in the literature the authors (see for instance Biagini and Cretarola (2006, 2009)) assume that $\int_0^T \lambda'_s d\langle M\rangle_s \lambda_s$ is bounded. Thanks to Proposition 3.7 in Choulli et al. (1998), this assumed condition implies (3.7).
Furthermore, as shown on page 48 this assumption implies the regularity and the reverse Hölder inequality of order 2. Therefore from Theorem 5.5 of Choulli et al. (1998) the existence of the FS decomposition for any P-square-integrable $\mathcal{F}_T$-measurable H follows.

Throughout this section, for any P-square-integrable $\mathcal{F}_T$-measurable H we denote

$$\widetilde{V}_t^H := \left[{}^{T_n}\mathcal{E}(\widetilde{N})_t\right]^{-1} E\left({}^{T_n}\mathcal{E}(\widetilde{N})_T H \mid \mathcal{F}_t\right), \quad T_n \le t < T_{n+1}, \tag{3.48}$$

with T_n the sequence defined in Definition 3.1.2.

Proposition 3.5.3. *The following assertions hold:*

(1) The process

$$\widetilde{K}_t := \widetilde{V}_t^H - \widetilde{V}_0^H + \langle \widetilde{V}^H, \widetilde{N}\rangle_t, \tag{3.49}$$

is a P-local martingale.

(2) *If* (H_0, ξ^{FS}, L^{FS}) *are the FS decomposition components of* H*, then*

$$\widetilde{V}_t^H = H_0 + \left(\xi^{FS} \cdot X\right)_t + L_t^{FS}. \tag{3.50}$$

Proof.

(1) From the proof of Proposition 3.12-(iii) of Choulli et al. (1998), we deduce that the process $\widetilde{V}_t^H$, defined in (3.48) is a $\mathcal{M}(\mathcal{E})$-martingale with $V_T = H$. By application of Corollary 3.16 of Choulli et al. (1998) we deduce that $[\widetilde{V}^H, \widetilde{N}]$ is P-locally integrable and that $\widetilde{V}^H + \langle \widetilde{V}^H, \widetilde{N} \rangle_t$ is a P-local martingale. Hence $\widetilde{K}$ is a P-local martingale with $K_0 = 0$.

(2) We prove first that $X + [X, \widetilde{N}]$ is a P-local martingale using Properties 2.2.15(2):

$$\begin{aligned} X + [X, \widetilde{N}] =& X_0 + M + B - [B, \lambda \cdot M] - [M, \lambda \cdot M] \\ =& X_0 + M + \lambda \cdot \langle M, M \rangle - \lambda' \Delta B \cdot M - \lambda \cdot [M, M] \\ =& X_0 + M - \lambda \cdot ([M, M] - \langle M, M \rangle) - \lambda' \Delta B \cdot M. \end{aligned}$$

Furthermore also L^{FS} and $[L^{\text{FS}}, \widetilde{N}] = -\lambda \cdot [L^{\text{FS}}, M]$ are P-local martingales due to Definition 3.1.1. From Proposition 3.15 of Choulli et al. (1998) we can conclude that therefore X and L^{FS} are $\mathcal{E}(\widetilde{N})$-local martingales. Which according to the definition of $\mathcal{E}(\widetilde{N})$-local martingales exactly means that for any $n \geq 0$, the processes ${}^{T_n}\mathcal{E}(\widetilde{N}) \left[(\xi^{\text{FS}} \cdot X) - (\xi^{\text{FS}} \cdot X)^{T_n} \right]$ and ${}^{T_n}\mathcal{E}(\widetilde{N}) \left[L^{\text{FS}} - L^{\text{FS}}_{T_n \wedge \cdot} \right]$ are P-local martingales. Furthermore, these processes are uniformly integrable due to (3.7) and the integrability of $\xi^{\text{FS}} \cdot X$ and L^{FS}. Then, for $t \geq T_n$ we derive

$$\begin{aligned} & E\left[{}^{T_n}\mathcal{E}(\widetilde{N})_T \Big((\xi^{\text{FS}} \cdot X)_T - (\xi^{\text{FS}} \cdot X)_{T_n} \Big) | \mathcal{F}_t \right] \\ & = {}^{T_n}\mathcal{E}(\widetilde{N})_t \Big((\xi^{\text{FS}} \cdot X)_t - (\xi^{\text{FS}} \cdot X)_{T_n} \Big), \\ & E\left[{}^{T_n}\mathcal{E}(\widetilde{N})_T \Big(L_T^{\text{FS}} - L_{T_n}^{\text{FS}} \Big) | \mathcal{F}_t \right] = {}^{T_n}\mathcal{E}(\widetilde{N})_t \Big(L_t^{\text{FS}} - L_{T_n}^{\text{FS}} \Big). \end{aligned}$$

As a result, due to ${}^{T_n}\mathcal{E}(\widetilde{N})_t \neq 0$ on $\{T_n \leq t < T_{n+1}\}$, we deduce that

$$\widetilde{V}_t^H = H_0 + (\xi^{\text{FS}} \cdot X)_t + L_t^{\text{FS}}.$$

This ends the proof of the second assertion. □

Now we will state the main result in this section.

Theorem 3.5.4. *Consider a square-integrable $\mathcal{F}_T$-measurable random variable H, and denote by $\left(H_0, \xi^{FS}, L^{FS}\right)$ its FS decomposition components. Then the following holds*

$$\xi^{FS} = \Sigma^{inv}\left\{c\widetilde{\phi} + \int x\widetilde{f}(x)K(dx)\right\} \quad \text{and} \quad L^{FS} = \widetilde{V}^H - \xi^{FS}\cdot X. \tag{3.51}$$

Here $(\widetilde{\phi}, \widetilde{f}, \widetilde{g}, \widetilde{K}^{\perp})$ is the quadruplet associated with $\widetilde{K}$ through Theorem 2.2.23, and Σ is a random symmetric matrix given by

$$\Sigma := c + \int xx'K(dx). \tag{3.52}$$

Proof. By applying Jacod's Theorem (Theorem 2.2.23) to the P-local martingale $\widetilde{K}$, we obtain

$$\begin{aligned} \widetilde{K} &= \widetilde{\phi}\cdot X^c + \widetilde{W}\star(\mu-\nu) + \widetilde{g}\star\mu + \widetilde{K}^{\perp}, \\ \widetilde{W}_t(x) &:= \widetilde{f}_t(x) + \frac{1}{1-a_t}\int \widetilde{f}_t(y)\nu(\{t\},dy). \end{aligned} \tag{3.53}$$

Another application of Theorem 2.2.23 now to L^{FS} leads to

$$\begin{aligned} L^{\text{FS}} &= \phi^{\text{FS}}\cdot X^c + W^{\text{FS}}\star(\mu-\nu) + g^{\text{FS}}\star\mu + L^{\text{FS},\perp}, \\ W^{\text{FS}}_t(x) &:= f^{\text{FS}}_t(x) + \frac{1}{1-a_t}\int f^{\text{FS}}_t(y)\nu(\{t\},dy). \end{aligned} \tag{3.54}$$

Since $\langle \widetilde{V}^H, \widetilde{N}\rangle = -\lambda\cdot\langle \widetilde{V}^H, M\rangle$ and using (3.49), Properties 2.2.15(2) on the predictable angle brackets and on the predictable process B with M and $\widetilde{K}$ both P-local martingales, we obtain that

$$\langle \widetilde{V}^H, M\rangle = \langle \widetilde{K}, M\rangle = \langle \widetilde{K}, X\rangle = \left\{c\widetilde{\phi} + \int x\widetilde{f}(x)K(dx)\right\}\cdot A,$$

where the last expression is obtained analogous to (3.33).
By plugging the resulting quantity into (3.49) while taking into account (3.53), we get

$$\begin{aligned} &\widetilde{V}^H \\ &= \widetilde{V}^H_0 + \widetilde{\phi}\cdot X^c + \widetilde{W}\star(\mu-\nu) + \widetilde{g}\star\mu + \widetilde{K}^{\perp} + \left(\lambda' c\widetilde{\phi} + \int \lambda' x\widetilde{f}(x)K(dx)\right)\cdot A. \end{aligned}$$

Furthermore, in view of (3.24) and (3.23) we find

$$\xi^{\text{FS}} \cdot B = \lambda' \xi^{\text{FS}} \cdot \langle M \rangle = \frac{\lambda' \xi^{\text{FS}}}{1 + \lambda' \Delta \langle M \rangle \lambda} \cdot \langle X \rangle = \frac{\lambda' c \xi^{\text{FS}} + \int \lambda' x x' \xi^{\text{FS}} K(dx)}{1 + \lambda' \Delta \langle M \rangle \lambda} \cdot A.$$

Hence, using this relationship, (3.50) and (3.54), we obtain

$$\begin{aligned}\widetilde{V}^H =& H_0 + \left(\xi^{\text{FS}} + \phi^{\text{FS}}\right) \cdot X^c + \left(W^{\text{FS}} + x' \xi^{\text{FS}}\right) \star (\mu - \nu) + g^{\text{FS}} \star \mu + L^{\text{FS},\perp} \\ &+ \frac{\lambda' c \xi^{\text{FS}} + \int \lambda' x x' \xi^{\text{FS}} K(dx)}{1 + \lambda' \Delta \langle M \rangle \lambda} \cdot A.\end{aligned}$$

Therefore, due to the uniqueness of Jacod's decomposition (Lemma 2.2.24) and of the canonical decomposition, see Definition 2.2.6, and due to the predictability of A, we conclude that

$$c\widetilde{\phi} = c\xi^{\text{FS}} + c\phi^{\text{FS}}, \quad \widetilde{f}(x) = x'\xi^{\text{FS}} + f^{\text{FS}}(x) \quad \widetilde{g}(x) = g^{\text{FS}}(x), \quad L^{\text{FS},\perp} = \widetilde{K}^{\perp}. \tag{3.55}$$

Thus by transforming the first two equations above, we derive

$$\begin{aligned}c\widetilde{\phi} + \int x \widetilde{f}(x) K(dx) =& c\xi^{\text{FS}} + c\phi^{\text{FS}} + \int x x' \xi^{\text{FS}} K(dx) + \int x f^{\text{FS}}(x) K(dx) \\ =& \Sigma \xi^{\text{FS}} + c\phi^{\text{FS}} + \int x f^{\text{FS}}(x) K(dx).\end{aligned} \tag{3.56}$$

Since L^{FS} satisfies

$$\langle L^{\text{FS}}, M \rangle = \langle L^{\text{FS}}, X \rangle = \left(c\phi^{\text{FS}} + \int x f^{\text{FS}}(x) K(dx) \right) \cdot A = 0,$$

see (3.33), equation (3.56) reduces to

$$\Sigma \xi^{\text{FS}} = c\widetilde{\phi} + \int x \widetilde{f}(x) K(dx),$$

and the first equation in (3.51) follows immediately. The second equation follows from (3.50). □

Even a correct decision is wrong when it was taken too late.

Lee Iacocca (1924-)

4 (Local) risk-minimization

In this chapter we discuss the (locally) risk-minimizing (LRM) hedging strategy. This strategy and also the mean-variance hedging strategy which will be discussed in Chapter 5 are quadratic hedging strategies. Hence they both minimize the hedging error in mean-square sense. Of course this has as drawback that losses and gains are treated in the same way. On the other hand having a symmetric criterion can be an advantage when you do not know if you deal with a buyer or a seller what is quiet natural when one is hedging and pricing options.

We start with explaining what a locally risk-minimizing hedging strategy is and how it extends the risk-minimizing hedging strategy. Afterwards we give an overview of the literature on LRM hedging strategies. In Section 4.3 we extend the theory of locally risk-minimization to the multidimensional setting. By independent research Schweizer (2008) also proved the multidimensional LRM hedging strategy under less restrictive conditions. In Section 4.4 we give a concrete procedure how to determine the LRM hedging strategy. We end this chapter with an illustration of this procedure on a simple Brownian motion setting with stochastic volatility. This example was already discussed in Poulsen et al. (2009), but the procedure they used is limited to continuous processes, while the procedure proposed here also allows for discontinuous processes.

This chapter is mainly based on Vandaele and Vanmaele (2008a). Two other, but older, overview articles which treat not only locally risk-minimization, but also the mean-variance hedging strategy are Pham (2000) and Schweizer (2001). Once we have described the LRM it will be obvious why the equivalence (or non-equivalence) between the FS and the GKW decomposition plays an important role, for more details we refer to Section 4.2.

4.1 (Local) risk-minimization

We work again under the assumptions defined in Section 2.1. The process X denotes here the one-dimensional discounted risky asset. We also assume there exists at least one change of measure to a martingale measure.
The locally risk-minimizing hedging strategy originates from the hedging strategy described by Harrison and Kreps (1979) for complete markets. This strategy was extended by Föllmer and Sondermann (1986) for incomplete markets if the underlying risky asset is still a martingale and is called the risk-minimizing hedging strategy. The extension to semimartingales was described for the first time by Schweizer (1988) and published in a series of papers: Schweizer (1990), Föllmer and Schweizer (1991) and Schweizer (1991).

4.1.1 Hedging in complete markets

Harrison and Kreps (1979) described how the hedging strategy for contingent claims can be found in a complete market. A **contingent claim** is a random variable that represents the time T payoff, denoted by H. In a complete market every contingent claim is redundant. **Redundant** means that the claim can be written as a sum of an initial cost, denoted by C_0, and a stochastic integral of the process X, which represents the discounted risky asset. Hence, the risk of the claim can be reduced to zero and the product does not add anything to the market containing the risky asset X.
A trading strategy φ is of the form (ξ, η), with $\xi = (\xi_t)_{0 \leq t \leq T}$ the number of risky assets and with $\eta = (\eta_t)_{0 \leq t \leq T}$ the amount invested in the riskless asset. The value of the discounted portfolio at time t is then given by $V_t = \xi_t X_t + \eta_t$.
Assume we want to hedge the claim H at time T. We perform a change of measure from the original measure P to the unique equivalent martingale measure

P^*. Due to the completeness of the market, every claim H can be decomposed as follows:

$$H = H_0 + \int_0^T \xi_u^* dX_u.$$

Then, the contingent claim can be reproduced at time T with the initial investment H_0 and the following strategy φ at time t:

$$(\xi_t^*, H_0 + \int_0^t \xi_u^* dX_u - \xi_t^* X_t).$$

4.1.2 Risk-minimization

The goal of Föllmer and Sondermann (1986) was to extend the hedging theory for redundant claims to contingent claims which are non-redundant. They searched for admissible strategies which minimize the risk in sequential sense and were able to prove that there exists a unique solution to this problem if the risky asset underlying the claim is a martingale.

Definition 4.1.1 (see Föllmer and Sondermann (1986))**.** A strategy is called **admissible** with respect to H or H**-admissible** if its value process has terminal value H.

In the theory of risk-minimization, Föllmer and Sondermann assumed that the discounted risky asset is a square-integrable martingale under the original measure P. From Section 2.2 we know this means that $E[X_t^2] < \infty$ and $E[X_T|\mathcal{F}_t] = X_t$, $0 \leq t \leq T$. They defined the risk-minimizing hedging strategy only in this setting. Schweizer (2001) proved that the conditions on X are too strong and that it is sufficient if X is only a local martingale under P, which do not even need to be locally square-integrable.

Definition 4.1.2. Assume X is a local martingale under the measure P. A couple $\varphi = (\xi, \eta)$ is called a **trading strategy** if

- ξ is a predictable process,
- $\xi \in L^2(X)$, with $L^2(X)$ the space of all $\mathbb{R}$-valued predictable processes ξ

such that

$$||\xi||_{L^2(X)} := (E[\int_0^T \xi_u^2 d[X,X]_u])^{1/2} < \infty,$$

- η is adapted,
- The value process $V = \xi X + \eta$ of the strategy φ has right-continuous paths and $E[V_t^2] < \infty$ for every $t \in [0,T]$ (i.e. $V_t \in L^2(P)$ for every $t \in [0,T]$).

Definition 4.1.3. The **cost process** is the difference between the value of the portfolio at time t and the gains/losses made from trading in the financial market up to time t:

$$C_t = V_t - \int_0^t \xi_u dX_u. \tag{4.1}$$

While the **risk process** is the conditional mean squared error process of the cost process:

$$R_t(\varphi) := E[(C_T(\varphi) - C_t(\varphi))^2 | \mathcal{F}_t], \quad 0 \le t \le T. \tag{4.2}$$

Definition 4.1.4. A strategy is called **self-financing** if the cost process, defined in (4.1), has constant paths, while a strategy is called **mean-self-financing** if the cost process is a square-integrable martingale.

We want to hedge a contingent claim $H \in L^2(P)$ due at time T. By searching for a hedging strategy for which the discounted portfolio has terminal value H, we find an H-admissible strategy. The martingale property of the risky asset process allows to show that the expected value of the terminal cost does not depend on the choice of the strategy: $E[C_T] = E[V_T - \int_0^T \xi_u dX_u] = E[H]$.
At any time t we have to minimize in mean-square sense the remaining cost $C_T - C_t$. Hence the risk process is minimized at any time t.

Definition 4.1.5. A strategy φ is called **risk-minimizing** if $R_t(\varphi) \le R_t(\hat{\varphi})$ P-almost surely for every admissible continuation $\hat{\varphi}$ of φ at time t. An **admissible continuation** of the strategy φ from t on is a strategy which coincides with φ for all times smaller than t and which also has terminal value H. In the following sense: $\hat{\varphi} = (\hat{\xi}, \hat{\eta})$ is an admissible continuation from t on of the strategy $\varphi = (\xi, \eta)$ if

$$\hat{\xi}_s = \xi_s, \quad \forall s \le t \quad \text{and} \quad \hat{\eta}_s = \eta_s, \quad \forall s < t$$

and $V_T(\hat{\varphi}) = V_T(\varphi)$ P-a.s..

Föllmer and Sondermann showed that the solution can be found using the Galtchouk-Kunita-Watanabe decomposition of the contingent claim H:

$$E[H|\mathcal{F}_t] = E[H] + \int_0^t \xi_u^* dX_u + N_t^*,$$

with $\xi^* \in L^2(X)$, N^* a square-integrable martingale, $N_0^* = 0$ and P-orthogonal to X.
The unique admissible strategy which is risk-minimizing is then given by $\varphi_t^* = (\xi_t^*, E[H|\mathcal{F}_t] - \xi_t^* X_t)$ at time t and the remaining risk equals $E[(N_T^* - N_t^*)^2|\mathcal{F}_t]$. This shows that all the risk related to the underlying is hedged away, while the only remaining risk is orthogonal to the fluctuations of the underlying and therefore cannot be hedged away.
We remark that in this case the number of risky assets invested using the mean-variance hedging strategy coincides with the number invested using the risk-minimizing hedging strategy. The amount invested in the riskless asset is different.

4.1.3 Local risk-minimization

On page 16 of Schweizer (1988) it is proved with an explicit counterexample that if X is not a P-local martingale, then a T-contingent claim H in general does not admit a risk-minimizing hedging strategy φ with $V_T(\varphi) = H$. Therefore Schweizer extended the theory of risk-minimization to locally risk-minimization.
The foundation for the locally risk-minimizing hedging strategy is described in Schweizer (1990), where the equivalence between the orthogonality of martingales and the risk-minimality under small perturbations is proved. In Schweizer (1991) the concept of locally risk-minimizing hedging strategies is introduced to be able to hedge claims when the underlying risky asset X is only a semimartingale under the original measure. The types of semimartingales for which the locally risk-minimizing hedging strategy is described have to be of the following form:

$$X = X_0 + M + B, \tag{4.3}$$

with M a square-integrable martingale for which $M_0 = 0$, and with B a predictable process belonging to the class $\mathcal{V}$.

Schweizer (1990) introduces a finite measure $P_M := P \times \langle M \rangle$ on the product space $\Omega \times [0,T]$ associated with the angle bracket process $\langle M \rangle$. We link an expectation E_M with this measure, defined in the following way:

$$E_M[\alpha] = \frac{E[(\alpha \cdot \langle M \rangle)_T]}{E[\langle M \rangle_T]}.$$

One also needs the following assumptions:

(A1) For P-almost all ω, the measure on $[0,T]$ induced by $\langle M \rangle(\omega)$ has the whole interval $[0,T]$ as its support. This means $\langle M \rangle$ should be P-almost surely strictly increasing on the whole interval $[0,T]$.

(A2) B is continuous.

(A3) B is absolutely continuous with respect to $\langle M \rangle$ with a density λ satisfying

$$E_M[|\lambda| \log^+ |\lambda|] < \infty.$$

A sufficient condition is that $E[\langle \int \lambda dM \rangle] < \infty$.

Therefore for a special quasi-left-continuous semimartingale satisfying the structure condition, whose MVT belongs to the set $L^1(P)$ and which satisfies assumption (A1), we can determine the LRM hedging strategy. Note that if $E[K_T] < +\infty$, then $E[\int_0^T |\lambda|_u^2 d\langle M \rangle_u] < +\infty$ and hence assumption (A3) is satisfied.
We remark that in Schweizer (1991) there was an extra condition: X is continuous at T P-a.s.. This condition is removed in Schweizer (2001).
The definition for the trading strategy is adjusted in this case:

Definition 4.1.6. Assume $X = X_0 + M + B$ is a semimartingale under the measure P. A couple $\varphi = (\xi, \eta)$ is called a **trading strategy** if

- ξ is a predictable process,
- ξ belongs to the space Θ, see (2.43).
- η is adapted,
- $V = \xi X + \eta$ has right-continuous paths and $E[V_t^2] < \infty$ for every $t \in [0,T]$.

In order to define the notion of locally risk-minimizing hedging strategies, we first explain what is meant by a small perturbation:

Definition 4.1.7. A trading strategy $\Delta = (\delta, \varepsilon)$ is called a **small perturbation** if it satisfies the following conditions:

- δ is bounded,
- $\int_0^T |\delta_u dB_u|$ is bounded,
- $\delta_T = \varepsilon_T = 0$.

For any subinterval $(s,t]$ of $[0,T]$, we define the small perturbation

$$\Delta|_{(s,t]} := (\delta \mathbb{1}_{(s,t]}, \varepsilon \mathbb{1}_{[s,t)}).$$

Next we define partitions $\tau = (t_i)_{0 \le i \le N}$ of the interval $[0,T]$. A **partition** of $[0,T]$ is a finite set $\tau = \{t_0, t_1, \dots, t_k\}$ of time points with $0 = t_0 < t_1 < \dots < t_k = T$ and the **mesh size** of τ is $|\tau| := \max_{t_i, t_{i+1} \in \tau} (t_{i+1} - t_i)$. A sequence $(\tau_n)_{n \in \mathbb{N}}$ is called increasing if $\tau_n \subseteq \tau_{n+1}$ for all n and it tends to the identity if $\lim_{n\to\infty} |\tau_n| = 0$.

Definition 4.1.8. For a trading strategy φ, a small perturbation Δ and a partition τ of $[0,T]$, the **risk quotient** $r^\tau[\varphi, \Delta]$ is defined as follows:

$$r^\tau[\varphi, \Delta](\omega, t) := \sum_{t_i, t_{i+1} \in \tau} \frac{R_{t_i}(\varphi + \Delta|_{(t_i, t_{i+1}]}) - R_{t_i}(\varphi)}{E[\langle M \rangle_{t_{i+1}} - \langle M \rangle_{t_i} | \mathcal{F}_{t_i}]} \mathbb{1}_{(t_i, t_{i+1}]}(t).$$

A trading strategy φ is called **locally risk-minimizing** if $\liminf_{n\to\infty} r^{\tau_n}(\varphi, \Delta) \ge 0$ P_M-a.e. on $\Omega \times [0,T]$ for every small perturbation Δ and every increasing sequence (τ_n) of partitions of $[0,T]$ tending to the identity.

This means that by LRM the riskiness of the cost process is measured locally in time.
The following theorems were originally proved by Schweizer (1991) if M belongs to the class $\mathcal{H}^2$ and X is one-dimensional. The extension to M only being a locally square-integrable martingale and to the multidimensional case is given in Schweizer (2008).

Lemma 4.1.9 (See Schweizer (1991) Lemma 2.1 and Schweizer (2008) Proposition 2.1)**.** *Assume that the special semimartingale X satisfies (A1). If a trading strategy is locally risk-minimizing, then it is also mean-self-financing.*

Theorem 4.1.10 (See Schweizer (1991) Proposition 2.3 and Schweizer (2008) Theorem 1.6)**.** *Assume that the special semimartingale X satisfies all conditions (A1)-(A3). Let the contingent claim H belong to $L^2(P)$ and let φ be an H-admissible trading strategy. Then φ is a locally risk-minimizing strategy if and only if φ is mean-self-financing and the martingale $C(\varphi)$, defined in* (4.1)*, is orthogonal to the martingale part M of the semimartingale X.*

Definition 4.1.11. A strategy φ is called **pseudo locally risk-minimizing** or, equivalently, **pseudo optimal risk-minimizing** if the associated cost process $C(\varphi)$, defined in (4.1), is a martingale under P and orthogonal to M.

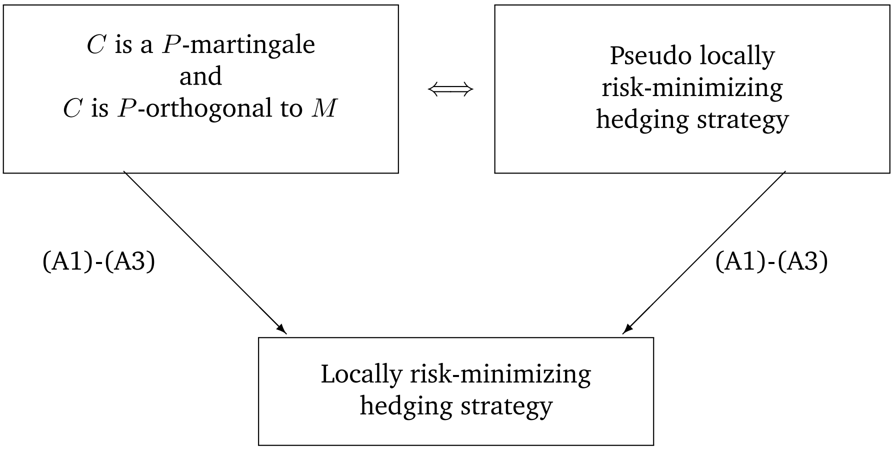


Figure 4.1.1: Equivalent conditions to determine the locally risk-minimizing hedging strategy under the original measure P.

In view of Definition 3.1.1, we can directly obtain the pseudo locally risk-minimizing hedging strategy φ from the FS decomposition:

$$\varphi_t = (\xi_t^{\mathrm{FS}}, E^{\widetilde{Q}}[H|\mathcal{F}_t] - \xi_t^{\mathrm{FS}} X_t),$$

where $E^{\widetilde{Q}}[H|\mathcal{F}_t] = H_0 + \int_0^t \xi_u^{\mathrm{FS}} dX_u + L_t^{\mathrm{FS}}$ see (3.50) and the fact that the MMM $\widetilde{Q}$ is here assumed to be a true martingale measure.
Using Proposition 4.1.10, we know that this pseudo locally risk-minimizing hedging strategy is the locally risk-minimizing strategy if the assumptions (A1)-(A3) are satisfied (see also Figure 4.1.1). In Chapter 3 we saw that if the process X is continuous, we can easily determine the Föllmer-Schweizer decomposition by performing a change of measure. Hence also the LRM strategy is quickly determined, as we will show along the lines of Föllmer and Schweizer (1991). We remark that under the continuity assumption of X the conditions (A1)-(A3) are trivially satisfied and it is easy to prove that the minimal martingale measure

preserves orthogonality. This means that any square-integrable P-martingale orthogonal to M is also orthogonal to X under $\widetilde{Q}$.
We emphasize once more that the preservation of orthogonality is not included in the definition of the minimal martingale measure, but that it is a consequence in some special cases.
Combining the previous results, we have the following proposition:

Proposition 4.1.12. *If X is continuous, the locally risk-minimizing strategy is determined by the Galtchouk-Kunita-Watanabe decomposition under the minimal martingale measure.*

Proof. Föllmer and Schweizer (1991) proved that the minimal martingale measure preserves orthogonality if X is continuous. In this case the Galtchouk-Kunita-Watanabe decomposition under the minimal martingale measure directly implies the Föllmer-Schweizer decomposition under the original measure. This already gives us the pseudo locally risk-minimizing hedging strategy. Since by the continuity of X, (A1)-(A3) are satisfied, we know from Proposition 4.1.10 that this strategy is locally risk-minimizing under the original measure. □

As a result of Proposition 4.1.12, we can determine the locally risk-minimizing hedging strategy in case of a continuous risky asset following the scheme in Figure 4.1.2.
In case the semimartingale is discontinuous, the only way to obtain the LRM hedging strategy is by calculating directly the FS decomposition as given in Theorem 3.5.4 or by adjusting the obtained numbers of risky assets from the GKW decomposition. In fact when searching the LRM hedging strategy, we will assume that the finite variation part is continuous and therefore we can find the number of risky assets in an easier way. A more detailed procedure is given in Section 4.4.

4.2 Applications in literature

This section presents the most important applications of the locally risk-minimizing hedging strategy. First of all, it is important to pay attention to the different notions of locally risk-minimizing hedging strategy in discrete and continuous time. In discrete time, the concept of locally risk-minimization is often used

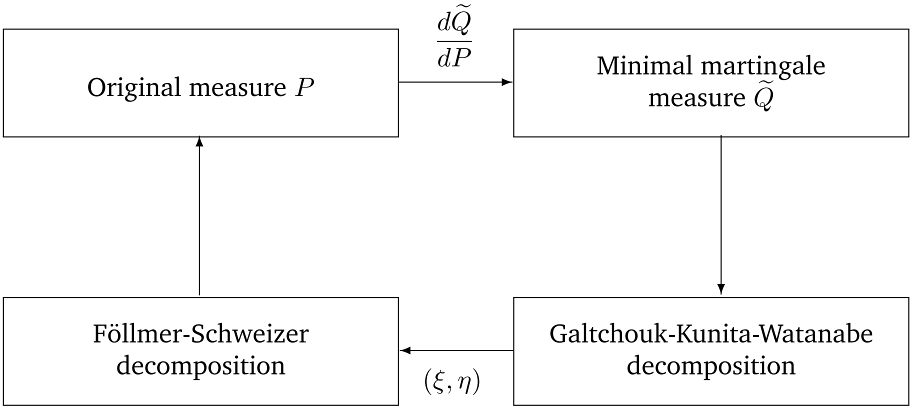


Figure 4.1.2: Scheme to determine the locally risk-minimizing hedging strategy in the case of a continuous underlying.

for the strategy which minimizes the difference between the cost process in every subsequent time interval, e.g. Coleman et al. (2006) and Coleman et al. (2007). This definition allows us to determine the hedging strategy by backward induction and furthermore in Coleman et al. (2007) they demonstrate that using a Black-Scholes model the discrete risk-minimizing hedging strategy performs better than the delta hedging to hedge variable annuities. Schäl (1994) proved equivalence in discrete time between the local risk-minimization and the risk-minimization used here, under certain conditions. In continuous time, the former definition is rarely used. The most important application of the theory of Föllmer and Sondermann (1986) is given in Møller (1998). He was the first to apply the risk-minimizing hedging strategy to unit-linked life insurance contracts. He considered a risky asset that followed a Brownian motion and a single premium contract with only a pure endowment and with a term insurance for which all the payments were deferred to maturity of the contract. To be able to hedge intermediate payments, Møller (2001) extended the theory of risk-minimizing hedging strategies. He determined the hedging strategy in the concrete case of a geometric Brownian motion, but the proof is more general because it holds for every risky asset which is a local martingale.
In fact in Møller (1998), the Brownian motion is only a semimartingale under the original measure, but we know already that it suffices to determine the

risk-minimizing hedging strategy under the minimal martingale measure when the price process of the risky asset is continuous. Indeed, for the continuous case the locally risk-minimizing hedging strategy under the original measure is equivalent to the risk-minimizing hedging strategy under the minimal martingale measure.

Riesner (2006a) wanted to extend the theory of Møller (1998) to the case of a geometric Lévy process. Unfortunately, the equivalence does no longer hold because of the discontinuity of the Lévy process. In Vandaele and Vanmaele (2008b) a correction is given based on Colwell and Elliott (1993) (see Chapter 7), who were the first to show how one can determine the Föllmer-Schweizer decomposition and the related locally risk-minimizing hedging strategy for a contingent claim when the underlying asset follows a Markov diffusion process with jumps.

Furthermore, the locally risk-minimizing hedging strategy for payment processes is defined in Schweizer (2008) and Riesner (2007). In Chapter 7 we will also discuss this last article. We remark that a more detailed version of the articles of Riesner can be found in Riesner (2006b).

Colwell et al. (2007) applied the locally risk-minimizing hedging strategy for index tracking. In Di Masi et al. (1994) various applications of locally risk-minimizing hedging strategies are given for stochastic volatility models. Using the same model for the underlying risky asset, Heath et al. (2001a and 2001b) compared the mean-variance hedging to the locally risk-minimizing strategy theoretically and numerically.

The risk-minimizing hedging strategy under restricted information is investigated by Di Masi et al. (1995) and Schweizer (1993). The application to stochastic volatility models is studied by Fisher et al. (1999). The extension of this application to the case of semimartingales is investigated by Frey and Runggaldier (1999).

Biagini and Pratelli (1999) proved the invariance under a change of measure of the locally risk-minimizing hedging strategy. Biagini and Cretarola (2007, 2006, 2009) determine the locally risk-minimizing hedging strategy for defaultable claims when the risky asset follows a Brownian motion, see also Cretarola (2007). Becchere and Mulinacci (1999) searched for the locally risk-minimizing hedging strategy to hedge American options in Merton's model.

In Mercurio and Vorst (1997) and in Lamberton et al. (1998), the locally risk-minimizing hedging strategy is defined in discrete time when transaction costs are taken into account.

Concerning risk-minimization Barbarin (2008a) studies the asset allocation problem of pure endowment and annuity portfolios when some longevity bonds are available for trading. In a non-life insurance setting with inflation and interest rate risk the risk-minimizing hedging strategy is calculated in Barbarin (2009). In Barbarin (2008b) the risk-minimizing hedging strategy is determined for life insurance contracts with a surrender option. We will make the obtained processes more concrete in Chapter 6.

4.3 Multidimensional local risk-minimization

In Schweizer (2001) it is mentioned that the extension of the locally risk-minimizing hedging strategy to the multidimensional case will be presented elsewhere. But finally it was never published until Schweizer (2008). In the meantime we had worked out the extension as part of this PhD research. In Schweizer (2008) the assumptions made in (A1)-(A3) are relaxed and the results are extended to payment streams. At the end of this section we will come back to the results of Schweizer. First we will give the proof we constructed based on Schweizer (1990, 1991).
Assume X is a d-dimensional special semimartingale having the decomposition $X = X_0 + M + B$, with M a d-dimensional square-integrable martingale with $M_0 = 0$ and B a d-dimensional predictable process of finite variation. Since M is assumed to be square-integrable, we can introduce the $d \times d$-dimensional variance process $\langle M \rangle$ with respect to P with components $\langle M^i, M^j \rangle$. The sum of all the components $\langle M^i, M^i \rangle$, $i = 1, \dots, d$ is given by

$$(M) = \sum_{i \in \{1,\dots,d\}} \langle M^i, M^i \rangle.$$

As was done in the one-dimensional case, we can link a product measure $P_M = P \times (M)$ and a related expectation E_M with the process (M) on the product space $\overline{\Omega} := \Omega \times [0, T]$.
We first give the extensions of the definitions of trading strategy, Definition 4.1.6, cost process, Definition 4.1.3, and small perturbation, Definition 4.1.7, to the multidimensional case. We remark that because the cost process is a one-dimensional process the formula (4.2) of the risk process remains unchanged.

Definition 4.3.1. A **trading strategy** φ contains $d+1$ processes $(\xi^1, \dots, \xi^d, \eta) = (\xi^1_t, \dots, \xi^d_t, \eta_t)_{0 \le t \le T}$ satisfying the following conditions:

- $\xi = (\xi^1, \ldots, \xi^d)$ is predictable,
- The process $\xi \cdot X$ is a semimartingale of class $\mathcal{S}^2(P)$, see page 30, meaning that
$$E[\int_0^T \xi'_u d\langle M\rangle_u \xi_u + (\int_0^T |\xi'_s dB_s|)^2] < \infty.$$
- η is adapted,
- $V = \xi' X + \eta$ has right-continuous paths and satisfies $V_t \in L^2(P)$, $0 \le t \le T$.

Definition 4.3.2. The **cost process** is the right-continuous square-integrable process $C(\varphi)$ defined by

$$C_t(\varphi) := V_t(\varphi) - \int_0^t \xi'_u dX_u. \tag{4.4}$$

Definition 4.3.3. A trading strategy $\Delta = (\delta^1, \ldots, \delta^d, \varepsilon)$ is called a **small perturbation** if it satisfies the following conditions:

- $\delta = (\delta^1, \ldots, \delta^d)$ is bounded,
- $\int_0^T |\delta'_u dB_u|$ is bounded, hence $\int_0^T |\delta'_u dB_u| \in L^\infty(P)$.
- $\langle \delta \cdot M \rangle$ is bounded.
- $\delta_T = \varepsilon_T = 0$.

It is important to understand that we will be able to leave out the condition that $E[K_T] < +\infty$ due to the third assumption. This means that by restricting the class of small perturbations, we can extend the class of semimartingales for which we are able to determine the LRM hedging strategy.
The partitions, the mesh and the small perturbation on the subinterval are completely analogously defined as in the one-dimensional case. Furthermore a partition τ induces the following σ-algebras:

$$\mathcal{B}^\tau := \sigma(\{D_0 \times \{0\}, D_i \times (t_{i-1}, t_i] \mid D_0 \in \mathcal{F}_0, t_i \in \tau, D_i \in \mathcal{F}_{t_i}\}), \tag{4.5}$$
$$\mathcal{P}^\tau := \sigma(\{D_0 \times \{0\}, D_{i-1} \times (t_{i-1}, t_i] \mid D_0 \in \mathcal{F}_0, t_i \in \tau, D_{i-1} \in \mathcal{F}_{t_{i-1}}\}). \tag{4.6}$$

Therefore
$$\mathcal{P} = \sigma(\cup_{n=1}^\infty \mathcal{P}^{\tau_n}). \tag{4.7}$$

Definition 4.3.4. Let φ be a trading strategy, Δ a small perturbation and τ a partition of $[0, T]$. The R-quotient is defined as

$$r^{\tau}[\varphi, \Delta](\omega, t) := \sum_{t_i, t_{i+1} \in \tau} \frac{R_{t_i}(\varphi + \Delta|_{(t_i, t_{i+1}]}) - R_{t_i}(\varphi)}{E[(M)_{t_{i+1}} - (M)_{t_i} | \mathcal{F}_{t_i}]}(\omega) \mathbb{1}_{(t_i, t_{i+1}]}(t). \quad (4.8)$$

The strategy φ is then **locally risk-minimizing** if $\liminf_{n \to \infty} r^{\tau_n}[\varphi, \Delta] \geq 0$ P_M-a.e. for every small perturbation Δ and every increasing 0-convergent sequence (τ_n) of partitions of $[0, T]$ tending to the identity.

We remark that where we used in the one-dimensional case the process $\langle M \rangle$ in the R-quotient, we now use the one-dimensional process (M) instead. To be sure the denominator in (4.8) is not zero, we need the following assumption

(A1') For P-almost all ω, the measure on $[0, T]$ induced by $(M)(\omega)$ has the whole interval $[0, T]$ as its support. This is equivalent with the assumption that $(M)(\omega)$ is strictly increasing P-a.s..

The lemmas and proofs given here are inspired by the proofs given in Schweizer (1990, 1991).

Lemma 4.3.5. *Assume that the special semimartingale X satisfies condition (A1'). If a trading strategy is locally risk-minimizing, then it is also mean-self-financing.*

Proof. The proof is given by a reductio ad absurdum. Assume that the trading strategy φ is not mean-self-financing, then we can construct a new mean-self-financing strategy $\hat{\varphi}$ with a risk process smaller than the risk process of the strategy φ in at least a small subinterval of [0,T] and with the same amount in the portfolio at time T. So this means that the trading strategy φ is not locally risk-minimizing.
The new strategy $\hat{\varphi} = (\hat{\xi}, \hat{\eta})$ is defined as follows for all $t \in [0, T]$:

$$\hat{\xi}_t = \xi_t, \quad (4.9)$$

$$\begin{aligned} \hat{\eta}_t &= E[V_T(\varphi) - \int_0^T \xi'_u dX_u | \mathcal{F}_t] + \int_0^t \xi'_u dX_u - \xi'_t X_t \\ &= E[C_T(\varphi) | \mathcal{F}_t] + \int_0^t \xi'_u dX_u - \xi'_t X_t. \quad (4.10) \end{aligned}$$

This means that

$$V_t(\hat{\varphi}) = \hat{\xi}_t' X_t + \hat{\eta}_t = E[C_T(\varphi)|\mathcal{F}_t] + \int_0^t \xi_u' dX_u. \tag{4.11}$$

The vector $\Delta = \hat{\varphi} - \varphi := (\delta, \varepsilon)$, linked with this new portfolio, is given by

$$(\hat{\xi}_t - \xi_t, \hat{\eta}_t - \eta_t) = (0, E[V_T(\varphi) - \int_0^T \xi_u' dX_u|\mathcal{F}_t] + \int_0^t \xi_u' dX_u - V_t(\varphi))$$

and so the conditions of a small perturbation are satisfied, because $\delta \equiv 0$ and $\varepsilon_T = 0$. Now we will prove that the strategy $\hat{\varphi}$ is mean-self-financing, admissible and that the risk process is smaller than the risk process of φ.

- $V_T(\hat{\varphi}) = V_T(\varphi)$
  It follows from (4.11) that

  $$V_T(\hat{\varphi}) = E[V_T(\varphi) - \int_0^T \xi_u' dX_u|\mathcal{F}_T] + \int_0^T \xi_u' dX_u = V_T(\varphi)$$

  and so also $C_T(\hat{\varphi}) = C_T(\varphi)$ in view of (4.4) and (4.9).
- The strategy $\hat{\varphi}$ is mean-self-financing. Using (4.11), (4.9) and the equality between $C_T(\hat{\varphi})$ and $C_T(\varphi)$:

  $$\begin{aligned} C_t(\hat{\varphi}) =& V_t(\hat{\varphi}) - \int_0^t \hat{\xi}_u' dX_u = E[C_T(\varphi)|\mathcal{F}_t] + \int_0^t \xi_u' dX_u - \int_0^t \xi_u' dX_u \\ =& E[C_T(\hat{\varphi})|\mathcal{F}_t]. \end{aligned}$$

- Using the martingale property of the cost process for $\hat{\varphi}$, we can easily prove that the risk process of the strategy $\hat{\varphi}$ is smaller or equal than the risk process of the strategy φ:

  $$\begin{aligned} R_s(\hat{\varphi}) =& E[(C_T(\hat{\varphi}) - C_s(\hat{\varphi}))^2|\mathcal{F}_s] \\ =& E[(C_T(\hat{\varphi}) - C_s(\varphi) + C_s(\varphi) - C_s(\hat{\varphi}))^2|\mathcal{F}_s] \\ =& E[(C_T(\varphi) - C_s(\varphi))^2|\mathcal{F}_s] + (C_s(\varphi) - C_s(\hat{\varphi}))^2 \\ &+ 2E[(C_T(\hat{\varphi}) - C_s(\varphi))|\mathcal{F}_s](C_s(\varphi) - C_s(\hat{\varphi})) \\ =& R_s(\varphi) + (C_s(\varphi) - C_s(\hat{\varphi}))^2 - 2(C_s(\varphi) - C_s(\hat{\varphi}))^2 \\ =& R_s(\varphi) - (C_s(\varphi) - C_s(\hat{\varphi}))^2 \\ \leq& R_s(\varphi). \end{aligned}$$

We assumed that the strategy φ is not mean-self-financing, so there exist a $\tilde{s} \in [0, T]$ and a set B of positive probability for which $C_{\tilde{s}}(\varphi)(\omega) \neq C_{\tilde{s}}(\hat{\varphi})(\omega)$ for all $\omega \in B$. Because of the right-continuity of the risk process, they will be different from each other in a closed interval $[\tilde{s}, d(\omega)]$ around $\tilde{s}$ for all $\omega \in B$ and so $R_k(\hat{\varphi})(\omega) < R_k(\varphi)(\omega)$ for all $k \in [\tilde{s}, d(\omega)]$ and $\omega \in B$.
This means that φ is not locally risk-minimizing because $\liminf_{n\to\infty} r^{\tau_n}[\varphi, \delta](\omega, t) < 0$ for any t in the $(\tilde{s}, d(\omega))$ and any $\omega \in B$ and therefore $\liminf_{n\to\infty} r^{\tau_n}[\varphi, \delta](\omega, t) \geq 0$ no longer holds P_M-a.e.. □

Therefore if we search an H-admissible locally risk-minimizing hedging strategy φ, it is certainly mean-self-financing. We show now that we find the locally risk-minimizing hedging strategy by only varying the ξ-component using the fact that this optimal strategy should be mean-self-financing.
Since the cost process $C(\varphi)$ is a martingale with terminal value

$$C_T(\varphi) = H - \int_0^T \xi'_u dX_u \quad P\text{-a.s.},$$

the cost process and hence also φ, is uniquely determined by ξ. This means we can write $C(\xi)$ and $R(\xi)$ instead of $C(\varphi)$ and $R(\varphi)$.
We now show that also for the perturbations it is sufficient to vary only the number of risky assets. We take a small perturbation $\Delta = (\delta, \varepsilon)$ and a partition τ of $[0, T]$. For a $t_i \in \tau$, we compare the H-admissible strategy $\varphi + \Delta|_{(t_i,t_{i+1}]}$ with the H-admissible mean-self-financing strategy associated with $\xi + \delta|_{(t_i,t_{i+1}]}$. The latter strategy can be constructed from the strategy $\varphi + \Delta_{(t_i,t_{i+1}]}$ as done in Lemma 4.3.5 and hence the amount of risky assets is the same as in the strategy $\varphi + \Delta|_{(t_i,t_{i+1}]}$.
The strategies $\varphi + \Delta|_{(t_i,t_{i+1}]}$ and $\xi + \delta|_{(t_i,t_{i+1}]}$ are both H-admissible and have the same ξ-component, so $C_T(\varphi + \Delta|_{(t_i,t_{i+1}]}) = C_T(\xi + \delta|_{(t_i,t_{i+1}]})$.
The cost process of the strategy $\varphi + \Delta|_{(t_i,t_{i+1}]}$ at time t_i is $C_{t_i}(\varphi) + \varepsilon_{t_i}$. To calculate the cost process at time t_i of $\xi + \delta|_{(t_i,t_{i+1}]}$, we use the fact that this strategy and the original strategy φ are mean-self-financing and that $M = X - X_0 - B$ is a P-martingale:

$$\begin{aligned} C_{t_i}(\xi + \delta|_{(t_i,t_{i+1}]}) =& E[C_T(\xi + \delta|_{(t_i,t_{i+1}]})|\mathcal{F}_{t_i}] \\ =& E[V_T(\xi + \delta|_{(t_i,t_{i+1}]}) - \int_0^T \xi'_u dX_u - \int_{t_i}^{t_{i+1}} \delta'_u dX_u|\mathcal{F}_{t_i}] \\ =& E[H - \int_0^T \xi'_u dX_u - \int_{t_i}^{t_{i+1}} \delta'_u dX_u|\mathcal{F}_{t_i}] \end{aligned}$$

$$=E[C_T(\varphi) - \int_{t_i}^{t_{i+1}} \delta'_u dX_u|\mathcal{F}_{t_i}]$$
$$=C_{t_i}(\varphi) - E[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}].$$

So we have the following equality between the cost processes:

$$C_T(\varphi + \Delta|_{(t_i,t_{i+1}]}) - C_{t_i}(\varphi + \Delta|_{(t_i,t_{i+1}]})$$
$$= C_T(\xi + \delta|_{(t_i,t_{i+1}]}) - C_{t_i}(\xi + \delta|_{(t_i,t_{i+1}]}) - (\varepsilon_{t_i} + E[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}]),$$

and for the associated risk process we find

$$R_{t_i}(\varphi + \Delta|_{(t_i,t_{i+1}]}) = R_{t_i}(\xi + \delta|_{(t_i,t_{i+1}]}) + (\varepsilon_{t_i} + E[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}])^2, \quad (4.12)$$

because the strategy $\xi + \delta|_{(t_i,t_{i+1}]}$ is mean-self-financing. Equation (4.12) holds for every t_i in the partition τ, so the risk quotient, defined in Definition 4.3.4 is

$$r^\tau[\varphi, \Delta] = r^\tau[\xi, \delta] + \sum_{t_i \in \tau} \frac{(\varepsilon_{t_i} + E[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}])^2}{E[(M)_{t_{i+1}} - (M)_{t_i}|\mathcal{F}_{t_i}]} \mathbb{1}_{(t_i,t_{i+1}]}. \quad (4.13)$$

To be able to define how a locally risk-minimizing hedging strategy will look like, we need to impose some extra conditions on the semimartingale X:

(A2') B is continuous.

(A3') B is absolutely continuous with respect to $\langle M \rangle$: there exists an $\mathbb{R}^d$-valued, predictable process λ such that for every $i = 1, \ldots, d$: $dB^i = \lambda \cdot d\langle M, M^i \rangle$. Furthermore we also assume that the MVT process K_T is finite. Hence the structure condition, see Definition 2.2.17, is satisfied.

From Proposition 2.2.20 it follows that there exists a predictable $d \times d$-matrix $\widetilde{\Sigma}$ such that $\langle M^i, M^j \rangle = \widetilde{\Sigma}^{ij} \cdot A$. This implies that

$$d(M) = \sum_{i=1}^{d} d\langle M^i \rangle = \sum_{i=1}^{d} \widetilde{\Sigma}^{ii} dA,$$

and $\langle M^i, M^j \rangle$ can be rewritten as

$$d\langle M^i, M^j \rangle = \widetilde{\Sigma}^{ij} dA = \frac{\widetilde{\Sigma}^{ij}}{\sum_{i=1}^{d} \widetilde{\Sigma}^{ii}} d(M) =: \widetilde{\sigma}^{ij} d(M). \quad (4.14)$$

Hence $\widetilde{\sigma} = \frac{\widetilde{\Sigma}}{\sum_{i=1}^{d} \widetilde{\Sigma}^{ii}}$ is a symmetric positive semi-definite $d \times d$-matrix.
We define the following processes as in Schweizer (1990):

$$Q_p[C, B, \tau](\omega, t) := \sum_{t_i \in \tau} \frac{|C_{t_i} - C_{t_{i-1}}|^p}{B_{t_i} - B_{t_{i-1}}}(\omega) \mathbb{1}_{(t_{i-1}, t_i]}(t), \tag{4.15}$$

$$\tilde{Q}_p[C, B, \tau](\omega, t) := \sum_{t_i \in \tau} \frac{E[|C_{t_i} - C_{t_{i-1}}|^p | \mathcal{F}_{t_{i-1}}]}{E[B_{t_i} - B_{t_{i-1}} | \mathcal{F}_{t_{i-1}}]}(\omega) \mathbb{1}_{(t_{i-1}, t_i]}(t). \tag{4.16}$$

These processes are used in Lemma 2.1 of the paper of Schweizer (1990), which we repeat here. We do not have to adapt the theorem to more dimensions, because it will only be used for one-dimensional processes.

Lemma 4.3.6 (See Schweizer (1990)). *Let $1 \leq r \leq p$ and assume that C is continuous and has integrable r-variation, in the sense that $E[(Var(C)_T)^r] < +\infty$. Then*

$$\lim_{n \to \infty} Q_p[C, B, \tau_n] = 0 \quad P_B\text{-a.e.}$$

If in addition

$$\sup_n Q_p[C, B, \tau_n] \in L^1(P_B)$$

and C is constant over any interval on which B is constant, then

$$\lim_{n \to \infty} \tilde{Q}_p[C, B, \tau_n] = 0 \quad P_B\text{-a.e.}$$

We need this result in the following proof.

Lemma 4.3.7.

$$\lim_{n \to \infty} \sum_{t_i \in \tau_n} \frac{E[(\int_{t_i}^{t_{i+1}} \delta'_u dB_u)^2 | \mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i} | \mathcal{F}_{t_i}]} \mathbb{1}_{(t_i, t_{i+1}]} = 0.$$

Proof. This means that using the process defined in formula (4.16), we have to prove that

$$\lim_{n \to \infty} \tilde{Q}_2[\delta \cdot B, (M), \tau_n] = 0 \quad P_M\text{-a.e.}.$$

We remark that $\delta \cdot B$ and (M) are one-dimensional processes and so we can use Lemma 4.3.6.

- Assumption (A2') guarantees that the process B is continuous and so $\delta \cdot B$ is also continuous.
- $\delta \cdot B$ is of integrable 1-variation, due to the definition of a small perturbation and the fact that an increasing process is of integrable variation if it is integrable.
- Assumption (A1') guarantees that (M) is strictly increasing and hence is never constant on the interval $[0, T]$.
- This means it remains to prove that

$$\sup_n Q_2[\delta \cdot B, (M), \tau_n] \in L^1(P_M). \tag{4.17}$$

In the next calculation we utilize in the first step that τ_n is a partition of $[0, T]$, while in the third step we use the fact that $B^j = \lambda \cdot \langle M, M^j \rangle$ by assumption (A3') and the relationship between $\langle M^i, M^j \rangle$ and (M) in (4.14).

$$\begin{aligned}
&Q_2[\delta \cdot B, (M), \tau_n] \\
&= Q_1[\delta \cdot B, (M), \tau_n] \sum_{t_i \in \tau_n} \Big| \int_{t_{i-1}}^{t_i} \delta'_u dB_u \Big| \mathbb{1}_{(t_{i-1}, t_i]} \\
&\le \sum_{t_i \in \tau_n} \frac{|\sum_{j=1}^d \int_{t_{i-1}}^{t_i} \delta^j_u dB^j_u|}{(M)_{t_i} - (M)_{t_{i-1}}} \mathbb{1}_{(t_{i-1}, t_i]} \int_0^T |\delta'_u dB_u| \\
&= \sum_{t_i \in \tau_n} \frac{|\sum_{k,j=1}^d \int_{t_{i-1}}^{t_i} \delta^j_u \lambda^k_u \widetilde{\sigma}^{kj}_u d(M)_u|}{(M)_{t_i} - (M)_{t_{i-1}}} \mathbb{1}_{(t_{i-1}, t_i]} \int_0^T |\delta'_u dB_u| \\
&\le E_M[|\delta' \widetilde{\sigma} \lambda| | \mathcal{B}^{\tau_n}] \int_0^T |\delta'_u dB_u|
\end{aligned} \tag{4.18}$$

because due to (4.5) we know:

$$\begin{aligned}
E_M[|\delta' \widetilde{\sigma} \lambda| | \mathcal{B}^{\tau_n}] &= \sum_{t_i \in \tau_n} \frac{E[\int_{t_{i-1}}^{t_i} |\delta'_u \widetilde{\sigma}_u \lambda_u| d(M)_u | \mathcal{F}_{t_i}]}{E[(M)_{t_i} - (M)_{t_{i-1}} | \mathcal{F}_{t_i}]} \mathbb{1}_{(t_{i-1}, t_i]} \\
&= \sum_{t_i \in \tau_n} \frac{\int_{t_{i-1}}^{t_i} |\delta'_u \widetilde{\sigma}_u \lambda_u| d(M)_u}{(M)_{t_i} - (M)_{t_{i-1}}} \mathbb{1}_{(t_{i-1}, t_i]}.
\end{aligned}$$

Combining the boundedness condition on

$$\int_0^T |\delta'_u dB_u| = \int_0^T |\delta'_u \widetilde{\sigma}_u \lambda_u| d(M)_u$$

assumed in Definition 4.3.3 of a small perturbation with the assumption that (M) is strictly increasing implies the boundedness of $|\delta'\widetilde{\sigma}\lambda|$. Therefore also $E_M[\delta'\widetilde{\sigma}\lambda|\mathcal{B}^{\tau_n}]$ will be bounded by a constant uniformly in n. The boundedness of the second factor follows again from Definition 4.3.3.
Therefore $Q_2[\delta \cdot B, (M), \tau_n]$ is uniformly bounded in n.
We can conclude that

$$\sup_n Q_2[\delta \cdot B, (M), \tau_n] \in L^1(P_M).$$

and by Lemma 4.3.6, (4.17) is satisfied. □

We remark that we no longer used that $E[K_T] < +\infty$ nor the condition of the form (A3) in the previous proofs, while these conditions were still necessary in Schweizer (2001). They are removed here by relying more on the conditions given by the definition of the small perturbation.

Lemma 4.3.8. *Assume that the special semimartingale X satisfies the assumptions (A1')-(A3'). Let the contingent claim H belong to the class $L^2(P)$ and $\varphi = (\xi, \eta)$ an H-admissible trading strategy. The strategy φ is locally risk-minimizing if and only if φ is mean-self-financing and $\liminf_{n\to\infty} r^{\tau_n}[\xi, \delta] \geq 0$ P_M-a.e. for every bounded predictable process δ satisfying the conditions of a small perturbation and for every increasing 0-convergent sequence (τ_n) of partitions of $[0, T]$.*

Proof. From Lemma 4.3.5, we know that φ is certainly mean-self-financing.
If $r^{\tau_n}[\xi, \delta] \geq 0$ P_M-almost everywhere for processes δ and sequences (τ_n) as defined in this lemma then from (4.13) we know $r^{\tau}[\varphi, \Delta] \geq 0$ and therefore the strategy is locally risk-minimizing.
If φ is locally risk-minimizing then the strategy is certainly mean-self-financing and we proved that it is sufficient to vary only ξ. Hence we can choose $\varepsilon \equiv 0$. Using Lemma 4.3.7, the second term in the right-hand side of equation (4.13) goes to zero when taking the limit and so $\liminf_{n\to\infty} r^{\tau_n}[\xi, \delta] \geq 0$ P_M-a.e. for every small perturbation δ and every increasing 0-convergent sequence τ_n of partitions of $[0, T]$. Thus $r^{\tau_n}[\xi, \delta] \geq 0$ P_M-almost everywhere. □

This means that the number of risky assets ξ is determined by Definition 4.3.4, which puts a condition on the risk quotient (4.8), while the amount invested in the riskless asset is determined using the fact that a locally risk-minimizing hedging strategy is mean-self-financing. We try now to give an other interpretation to the condition put on the risk quotient in terms of the cost process.

Lemma 4.3.9. *Assume that the special semimartingale X satisfies the assumptions (A1')-(A3'). Let H be a contingent claim and φ an H-admissible trading strategy.*
The strategy φ is locally risk-minimizing if and only if φ is mean-self-financing and the martingale $C(\varphi)$ is orthogonal to M.

Proof. Using Lemma 4.3.8, we only have to prove that the condition

$$\liminf_{n\to\infty} r^{\tau_n}[\xi,\delta] \geq 0$$

is equivalent with the fact that the cost process associated with ξ is orthogonal to the martingale part M of the risky asset, given the fact that the strategy is mean-self-financing. Hence we may assume that the cost process is a martingale which has a Galtchouk-Kunita-Watanabe decomposition:

$$C_T = C_0 + \int_0^T \mu'_u dM_u + L_T, \tag{4.19}$$

with L orthogonal to M. We start by looking at the consequences of perturbations on small intervals. The process $C_t(\delta,\tau_n,i)$ is the associated cost process of the mean-self-financing strategy $\varphi + \delta|_{(t_{i-1},t_i]}$.

$$\begin{aligned} C_t(\delta,\tau_n,i) =& E[C_T(\delta,\tau_n,i)|\mathcal{F}_t] = E[H - \int_0^T \xi'_u dX_u - \int_{t_{i-1}}^{t_i} \delta'_u dX_u|\mathcal{F}_t] \\ =& E\left[C_T - \int_{t_{i-1}}^{t_i} \delta'_u dX_u|\mathcal{F}_t\right] = C_t - E\left[\int_{t_{i-1}}^{t_i} \delta'_u dX_u|\mathcal{F}_t\right]. \end{aligned}$$

We now calculate the risk quotient associated with this perturbation:

$$\begin{aligned} C_T(\delta,\tau_n,i+1) - C_{t_i}(\delta,\tau_n,i+1) =& C_T - C_{t_i} - \int_{t_i}^{t_{i+1}} \delta'_u dM_u - \int_{t_i}^{t_{i+1}} \delta'_u dB_u \\ &+ E\left[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}\right]. \end{aligned}$$

Hence by raising to the square:

$$
\begin{aligned}
&(C_T(\delta,\tau_n,i+1)-C_{t_i}(\delta,\tau_n,i+1))^2-(C_T-C_{t_i})^2\\
&=\left(\int_{t_i}^{t_{i+1}}\delta'_u dM_u\right)^2+\left(\int_{t_i}^{t_{i+1}}\delta'_u dB_u\right)^2+\left(E\left[\int_{t_i}^{t_{i+1}}\delta'_u dB_u|\mathcal{F}_{t_i}\right]\right)^2\\
&\quad-2(C_T-C_{t_i})\left(\int_{t_i}^{t_{i+1}}\delta'_u dM_u+\int_{t_i}^{t_{i+1}}\delta'_u dB_u-E[\int_{t_i}^{t_{i+1}}\delta'_u dB_u|\mathcal{F}_{t_i}]\right)\\
&\quad+2\int_{t_i}^{t_{i+1}}\delta'_u dM_u\left(\int_{t_i}^{t_{i+1}}\delta'_u dB_u-E[\int_{t_i}^{t_{i+1}}\delta'_u dB_u|\mathcal{F}_{t_i}]\right)\\
&\quad-2\int_{t_i}^{t_{i+1}}\delta'_u dB_u E\left[\int_{t_i}^{t_{i+1}}\delta'_u dB_u|\mathcal{F}_{t_i}\right].
\end{aligned}
\tag{4.20}
$$

Assume that K is a martingale under P and that the integral $\int_{t_i}^{t_{i+1}} a_u dF_u$ is $\mathcal{F}_{t_{i+1}}$-measurable, then by the towerlaw we obtain that

$$
\begin{aligned}
E[(K_T-K_{t_i})\int_{t_i}^{t_{i+1}}a_u dH_u|\mathcal{F}_{t_i}]&=E[E[(K_T-K_{t_i})\int_{t_i}^{t_{i+1}}a_u dH_u|\mathcal{F}_{t_{i+1}}]|\mathcal{F}_{t_i}]\\
&=E[E[(K_T-K_{t_i})|\mathcal{F}_{t_{i+1}}]\int_{t_i}^{t_{i+1}}a_u dH_u|\mathcal{F}_{t_i}]\\
&=E[(K_{t_{i+1}}-K_{t_i})\int_{t_i}^{t_{i+1}}a_u dH_u|\mathcal{F}_{t_i}].
\end{aligned}
\tag{4.21}
$$

In the next calculation we use equation I.4.6 of Jacod and Shiryaev (2002) which states that $E(M_TN_T)=E(\langle M,N\rangle_T)+E(M_0N_0)$ for every M, $N\in\mathcal{H}^2$. Taking the expectation given the information until time t_i of (4.20) and using (4.21), the martingale property of the cost process of the strategy φ and of the martingale M, we get:

$$
\begin{aligned}
&R_{t_i}(\varphi+\delta|_{(t_i,t_{i+1})})-R_{t_i}(\varphi)\\
&=E\left[\int_{t_i}^{t_{i+1}}\delta'_u d\langle M\rangle_u\delta_u|\mathcal{F}_{t_i}\right]+E\left[\left(\int_{t_i}^{t_{i+1}}\delta'_u dB_u\right)^2|\mathcal{F}_{t_i}\right]\\
&\quad+\left(E\left[\int_{t_i}^{t_{i+1}}\delta'_u dB_u|\mathcal{F}_{t_i}\right]\right)^2\\
&\quad-2E\left[(C_T-C_{t_i})\left(\int_{t_i}^{t_{i+1}}\delta'_u dM_u+\int_{t_i}^{t_{i+1}}\delta'_u dB_u\right)|\mathcal{F}_{t_i}\right]
\end{aligned}
$$

$$
\begin{aligned}
&+ 2E[C_T - C_{t_i}|\mathcal{F}_{t_i}]E[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}] \\
&+ 2E\left[\int_{t_i}^{t_{i+1}} \delta'_u dM_u \int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}\right] \\
&- 2E\left[\int_{t_i}^{t_{i+1}} \delta'_u dM_u|\mathcal{F}_{t_i}\right] E\left[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}\right] - 2\left(E\left[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}\right]\right)^2 \\
= &\, E\left[\int_{t_i}^{t_{i+1}} \delta'_u d\langle M\rangle_u \delta_u|\mathcal{F}_{t_i}\right] + \text{Var}\left[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}\right] \\
&- 2E\left[(C_{t_{i+1}} - C_{t_i})\int_{t_i}^{t_{i+1}} \delta'_u dM_u|\mathcal{F}_{t_i}\right] - 2E\left[(C_{t_{i+1}} - C_{t_i})\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}\right] \\
&+ 2E\left[\int_{t_i}^{t_{i+1}} \delta'_u dM_u \int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}\right].
\end{aligned}
$$

Because of equation (4.19), we know that $C_{t_{i+1}} - C_{t_i} = \int_{t_i}^{t_{i+1}} \mu'_u dM_u + L_{t_{i+1}} - L_{t_i}$. Using the orthogonality of L and M, the difference of the risk processes can be written as

$$
\begin{aligned}
&R_{t_i}(\varphi + \delta|_{(t_i,t_{i+1})}) - R_{t_i}(\varphi) \\
&= E\left[\int_{t_i}^{t_{i+1}} \delta'_u d\langle M\rangle_u \delta_u|\mathcal{F}_{t_i}\right] + \text{Var}\left[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}\right] \\
&\quad - 2E\left[\int_{t_i}^{t_{i+1}} \delta'_u d\langle M\rangle \mu_u|\mathcal{F}_{t_i}\right] \\
&\quad + 2E\left[\left(\int_{t_i}^{t_{i+1}} \delta'_u dM_u - (C_{t_{i+1}} - C_{t_i})\right)\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}\right].
\end{aligned}
$$

So the risk quotient for this perturbation is

$$
\begin{aligned}
r^{\tau_n}[\xi,\delta] = &\sum_{t_i\in\tau_n} \frac{E[\int_{t_i}^{t_{i+1}} \delta'_u d\langle M\rangle_u \delta_u|\mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i}|\mathcal{F}_{t_i}]} \mathbb{1}_{(t_i,t_{i+1}]} \\
&- 2\sum_{t_i\in\tau_n} \frac{E[\int_{t_i}^{t_{i+1}} \delta'_u d\langle M\rangle \mu_u|\mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i}|\mathcal{F}_{t_i}]} \mathbb{1}_{(t_i,t_{i+1}]} \\
&+ \sum_{t_i\in\tau_n} \frac{\text{Var}[\int_{t_i}^{t_{i+1}} \delta'_u dB_u|\mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i}|\mathcal{F}_{t_i}]} \mathbb{1}_{(t_i,t_{i+1}]}
\end{aligned}
$$

$$+ 2 \sum_{t_i \in \tau_n} \frac{E\Big[\Big(\int_{t_i}^{t_{i+1}} \delta'_u dM_u - (C_{t_{i+1}} - C_{t_i})\Big) \int_{t_i}^{t_{i+1}} \delta'_u dB_u | \mathcal{F}_{t_i}\Big]}{E[(M)_{t_{i+1}} - (M)_{t_i} | \mathcal{F}_{t_i}]} \mathbb{1}_{(t_i, t_{i+1}]}. \tag{4.22}$$

From Lemma 4.3.7, we know that

$$\lim_{n\to\infty} \sum_{t_i \in \tau_n} \frac{\text{Var}[\int_{t_i}^{t_{i+1}} \delta'_u dB_u | \mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i} | \mathcal{F}_{t_i}]} \mathbb{1}_{(t_i, t_{i+1}]} = 0.$$

We rewrite the last term in equation (4.22), based on the martingale property of the cost process and the martingale part M of the risky asset X:

$$\begin{aligned}
&\sum_{t_i \in \tau_n} \frac{E[(\int_{t_i}^{t_{i+1}} \delta'_u dM_u - (C_{t_{i+1}} - C_{t_i})) \int_{t_i}^{t_{i+1}} \delta'_u dB_u) | \mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i} | \mathcal{F}_{t_i}]} \mathbb{1}_{(t_i, t_{i+1}]} \\
&= \sum_{t_i \in \tau_n} \frac{\text{Cov}[\int_{t_i}^{t_{i+1}} \delta'_u dM_u - (C_{t_{i+1}} - C_{t_i}), \int_{t_i}^{t_{i+1}} \delta'_u dB_u | \mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i} | \mathcal{F}_{t_i}]} \mathbb{1}_{(t_i, t_{i+1}]}.
\end{aligned}$$

Using the fact that $\text{Cov}(X, Y) \leq \text{Var}(X)^{1/2} \text{Var}(Y)^{1/2}$ by the Cauchy-Schwarz inequality, we find

$$\begin{aligned}
&\sum_{t_i \in \tau_n} \frac{E[(\int_{t_i}^{t_{i+1}} \delta'_u dM_u - (C_{t_{i+1}} - C_{t_i})) \int_{t_i}^{t_{i+1}} \delta'_u dB_u) | \mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i} | \mathcal{F}_{t_i}]} \mathbb{1}_{(t_i, t_{i+1}]} \\
&= \left(\sum_{t_i \in \tau_n} \frac{\text{Var}[\int_{t_i}^{t_{i+1}} \delta'_u dB_u | \mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i} | \mathcal{F}_{t_i}]} \mathbb{1}_{(t_i, t_{i+1}]} \right)^{1/2} \\
&\quad \cdot \left(\sum_{t_i \in \tau_n} \frac{\text{Var}[\int_{t_i}^{t_{i+1}} \delta'_u dM_u - (C_{t_{i+1}} - C_{t_i}) | \mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i} | \mathcal{F}_{t_i}]} \mathbb{1}_{(t_i, t_{i+1}]} \right)^{1/2}
\end{aligned} \tag{4.23}$$

The first factor will go to zero when n goes to infinity. It remains to prove that the second factor is bounded in n P_M-almost everywhere. Using the fact that

$-2E[AB] \leq E[A^2] + E[B^2]$, because $0 \leq E[(A+B)^2]$ we obtain

$$\begin{aligned}
&\text{Var}[\int_{t_i}^{t_{i+1}} \delta'_u dM_u - (C_{t_{i+1}} - C_{t_i})|\mathcal{F}_{t_i}] \\
&= E[(\int_{t_i}^{t_{i+1}} \delta'_u dM_u - (C_{t_{i+1}} - C_{t_i}))^2|\mathcal{F}_{t_i}] \\
&= E[\int_{t_i}^{t_{i+1}} \delta'_u d\langle M\rangle_u \delta_u + (C_{t_{i+1}} - C_{t_i})^2|\mathcal{F}_{t_i}] \\
&\quad - 2E[\int_{t_i}^{t_{i+1}} \delta'_u dM_u (C_{t_{i+1}} - C_{t_i})|\mathcal{F}_{t_i}] \\
&\leq 2E[\int_{t_i}^{t_{i+1}} \delta'_u d\langle M\rangle_u \delta_u + (C_{t_{i+1}} - C_{t_i})^2|\mathcal{F}_{t_i}] \\
&\leq 2E[\int_{t_i}^{t_{i+1}} d\langle \delta \cdot M\rangle_u + \int_{t_i}^{t_{i+1}} d\langle C\rangle_u|\mathcal{F}_{t_i}].
\end{aligned}$$

Due to Definition 4.3.3 of a small perturbation, the first term is bounded and because C is assumed to be a square-integrable martingale the second term is finite. When we divide each term by $E[(M)_{t_{i+1}} - (M)_{t_i}|\mathcal{F}_{t_i}]$ and take the sum over all t_i we obtain a non-negative $(P_M, \mathcal{P}^{\tau_n})$-supermartingale (a process X is a P-supermartingale on the filtration $\mathcal{F}$ if $X_s \geq E[X_t|\mathcal{F}_s]$ for every $s \leq t$) due to the fact that the angle bracket process is a concave function together with Jensen's inequality. Hence the second factor in (4.23) is bounded in n P_M-a.e.. So taking the limit for n to infinity, the equation (4.8) goes to:

$$\begin{aligned}
\lim_{n\to\infty} r^{\tau_n}[\xi,\delta] &= \lim_{n\to\infty} \sum_{t_i\in\tau_n} \frac{E[\int_{t_i}^{t_{i+1}} \delta'_u d\langle M\rangle_u \delta_u|\mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i}|\mathcal{F}_{t_i}]} \mathbb{1}_{(t_i,t_{i+1}]} \\
&\quad - 2\lim_{n\to\infty} \sum_{t_i\in\tau_n} \frac{E[\int_{t_i}^{t_{i+1}} \delta'_u d\langle M\rangle \mu_u|\mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i}|\mathcal{F}_{t_i}]} \mathbb{1}_{(t_i,t_{i+1}]} \\
&= \lim_{n\to\infty} \sum_{t_i\in\tau_n} \frac{E[\int_{t_i}^{t_{i+1}} \delta'_u \widetilde{\sigma}_u d(M)_u (\delta_u - 2\mu_u)|\mathcal{F}_{t_i}]}{E[(M)_{t_{i+1}} - (M)_{t_i}|\mathcal{F}_{t_i}]} \mathbb{1}_{(t_i,t_{i+1}]} \\
&= \lim_{n\to\infty} E_M[\delta'_u \widetilde{\sigma}_u \delta_u - 2\delta'_u \widetilde{\sigma}_u \mu_u|\mathcal{P}^{\tau_n}] \\
&= \delta'_u \widetilde{\sigma}_u \delta_u - 2\delta'_u \widetilde{\sigma}_u \mu_u, \qquad (4.24)
\end{aligned}$$

due to the predictability of $\delta'_u \widetilde{\sigma}_u \delta_u - 2\delta'_u \widetilde{\sigma}_u \mu_u$ and (4.7).
The cost process C is orthogonal to M if and only if $\mu \equiv 0$, then it is obvious

from (4.24) that $\liminf_{n\to\infty} r^{\tau_n}[\xi,\delta] \geq 0$.
For the reverse: choose $\delta^i = \varepsilon\ \text{sign}_{\{\sum_j \widetilde{\sigma}^{ij}\mu^j\}} \mathbb{1}_{\{(|\sum_j \sigma^{ij}\lambda^j|\cdot(M) \vee |\sum_{k,l}\sigma^{kl}|\cdot(M)) \leq K\}}$, because with this choice δ satisfies the conditions needed for a small perturbation. Furthermore the first term in (4.24) is certainly non-negative because $\widetilde{\sigma}$ is positive semi-definite and contains the factor ε^2. The second term gets the sign of ε.
Let $\varepsilon \to 0$ and $K \to \infty$, then $\sum_{i,j}\widetilde{\sigma}^{ij}\mu^j$ should be equal to zero otherwise $\liminf_{n\to\infty} r^{\tau_n}[\xi,\delta]$ would be negative. Hence also $\mu'\widetilde{\sigma}\mu = 0$ in the limit and again from the fact that $\widetilde{\sigma}$ is positive semi-definite, this means that $\mu \equiv 0$. Thus the cost process C is orthogonal to the martingale part M, see (4.19). □

Schweizer (2008) extended this result in several ways. First of all the condition that M should be a square-integrable martingale is weakened to a locally square-integrable martingale. Further, he developed the locally risk-minimizing hedging strategy for payment streams instead of contingent claims. Also the basic criterion to obtain a locally risk-minimizing hedging strategy is adjusted, such that the process $\langle M \rangle$ should no longer be strictly increasing. The new criterion which replaces (4.8) is:

$$r^\tau[\varphi,\Delta](\omega,t) := \sum_{t_i,t_{i+1}\in\tau} \frac{R_{t_i}(\varphi + \Delta|_{(t_i,t_{i+1}]}) - R_{t_i}(\varphi)}{E[D_{t_{i+1}} - D_{t_i}|\mathcal{F}_{t_i}]}(\omega)\mathbb{1}_{(t_i,t_{i+1}]}(t).$$

with D a bounded strictly increasing, predictable process null at 0 such that $\langle M^i, M^j\rangle \ll D$. Hence in Schweizer (2008) the following proposition is proved:

Proposition 4.3.10. *Suppose the $\mathbb{R}^d$-valued semimartingale X satisfies the structure condition (SC) and the finite variation process B is continuous. Then a payment stream H admits a LRM L^2-strategy φ if and only if H_T admits a FS decomposition with triplet $(H_0, \theta^{H_T}, L^{H_T})$. In that case $\varphi = (\xi,\eta)$ is given by*

$$\xi = \theta^{H_T},\ \eta = V^{H_T} - \theta^{H_T}\cdot X$$

with

$$V_t^{H_T} := H_T^{(0)} + \int_0^t \theta_s^{H_T} dX_s + L_t^{H_T} - H_t,\ 0 \leq t \leq T.$$

and $C_t(\varphi) = H_T^{(0)} + L_t^{H_T}$.

Combining this theorem with Theorem 3.1.7, we obtain the following existence result for the LRM hedging strategy:

Theorem 4.3.11. *Suppose X is an $\mathbb{R}^d$-valued special semimartingale satisfying the SC and which finite variation part B is continuous. If also $\mathcal{E}$ is regular and satisfies (R_2) then the LRM hedging strategy and the FS decomposition exist, and the L^2-strategy in which we have to invest can be derived from the FS decomposition.*

We remark that if K_T is assumed to be bounded, then by Proposition 3.7 of Choulli et al. (1998) $\mathcal{E}$ is regular and satisfies (R_2). Hence in literature this condition is often used in order to obtain easily sufficient conditions to guarantee the existence of the FS decomposition.

4.4 Determination of the LRM strategy

The procedure is based on the proof of Theorem 5.5 of Choulli et al. (1998). We denote by H the payment process for which we want to determine the hedging strategy.

(1) The value of the portfolio at time t is given by the $\mathcal{E}(\widetilde{N})$-martingale Y with $Y_T = H_T$. The existence and the form of this Y is given by Proposition 3.12 (iii) of Choulli et al. (1998):

$$Y_t := \frac{E[H^{T_n}\mathcal{E}_T|\mathcal{F}_t]}{{}^{T_n}\mathcal{E}_t} \text{ on } \{t \in [T_n, T_{n+1}[\},$$

with T_n as in Definition 3.1.2. If the Girsanov density Z describing the change of measure to the minimal martingale measure never vanishes then the sequence T_n reduces to the sequence $\{0, T\}$. This simplifies the expression of Y:

$$Y_t = \frac{E[HZ_T|\mathcal{F}_t]}{Z_t} = E^{\widetilde{Q}}[H|\mathcal{F}_t] \text{ on } \{t \in [0, T[\}. \tag{4.25}$$

Hence Y is not only an $\mathcal{E}(\widetilde{N})$-martingale but even a true $\widetilde{Q}$-martingale.

(2) Due to Remark 2.3.3 we know that we can decompose the $\widetilde{Q}$-martingale Y as follows $Y = Y_0 + I - \langle I, N\rangle = Y_0 + \xi \cdot X + L$, with $I = \xi \cdot M + L$ the GKW decomposition of the P-martingale I. Furthermore $\langle I, N\rangle$ is the finite variation part of the $\widetilde{Q}$-martingale Y under P. In the setting

described here the finite variation part is also continuous because $\langle I, N\rangle = \langle \xi \cdot M + L, N\rangle = -\lambda'\xi \cdot \langle M, M\rangle$. Therefore the decomposition of Y in a martingale part and a finite variation part is unique. This makes it easy to identify the P-martingale part I of Y.

(3) The problem of finding the FS decomposition of H is now reduced to the determination of the GKW decomposition of the P-martingale $I = \xi \cdot M + L$. Hence by taking the angle bracket with respect to M, we easily extract ξ:

$$\langle I, M\rangle^P = \xi \cdot \langle M, M\rangle^P.$$

Hence we can write ξ as a Radon-Nikodym derivative:

$$\xi = (d\langle M, M\rangle^P)^{\text{inv}} d\langle I, M\rangle^P \tag{4.26}$$

and the strategy at time t is given by (ξ_t, η_t) with $\eta_t = Y_t - \xi_t' X_t - H_t$. With the inverse we mean the pseudoinverse of Moore-Penrose, see Albert (1972).

4.5 Illustration on stochastic volatility models

The technique described in Section 4.4 to determine the LRM hedging strategy is here illustrated on the class of stochastic volatility models considered in Poulsen et al. (2009). They apply a three step procedure introduced by El Karoui et al. (1997): the market is completed, the hedging strategy is calculated in this completed market and then projected on the original market. In fact our procedure is similar, in that the determination of the MMM is part of completing the market. However our approach is more general because the approach of El Karoui et al. (1997) is limited to Brownian motions. Furthermore we also know that in the continuous setting the FS decomposition can easily be deducted from the GKW decomposition. This will be shown explicitly on the example applied in Poulsen et al. (2009).

The hedge is determined for European claims $H(S_T^*)$ when the underlying undiscounted risky asset follows a stochastic volatility model of the following form:

$$\frac{dS^*(t)}{S^*(t)} = \mu dt + S^*(t)^\gamma f(V(t))[\sqrt{1-\rho^2}dW^1(t) + \rho dW^2(t)],$$
$$\frac{dV(t)}{V(t)} = \beta(V(t))dt + g(V(t))dW^2(t),$$

with independent standard Brownian motions W^1 and W^2. We refer to Poulsen et al. (2009) for an overview of the models contained in this class and for more details concerning the functions/parameters β, γ, μ, ρ, g and f.
With S we denote the discounted dynamics:

$$\frac{dS(t)}{S(t)} = (\mu - r)dt + S^*(t)^\gamma f(V(t))[\sqrt{1-\rho^2}dW^1(t) + \rho dW^2(t)].$$

The notation M is used for the martingale part of the risky asset S. We are in a Markovian market model and we can easily determine the minimal martingale measure $\widetilde{Q}$: $Z_t = E[\frac{d\widetilde{Q}}{dP}|\mathcal{F}_t] = \mathcal{E}(-\lambda \cdot M)_t$ with

$$dM_t = S(t)S^*(t)^\gamma f(V(t))[\sqrt{1-\rho^2}dW^1(t) + \rho dW^2(t)],$$
$$d\langle M, M\rangle_t^P = S(t)^2 S^*(t)^{2\gamma} f^2(V(t))dt,$$
$$\lambda_t = \frac{\mu - r}{S(t)S^*(t)^{2\gamma} f^2(V(t))}.$$

Therefore $Z_t = \mathcal{E}[-\frac{\mu-r}{S^*(t)^\gamma f(V(t))}(\sqrt{1-\rho^2}dW^1(t) + \rho dW^2(t))]$. There exists a function $C(t, S^*(t), V(t))$ which equals $e^{-r(T-t)}E^{\widetilde{Q}}[H(S^*(T))|S^*(t), V(t)]$. Applying Itô's formula to this $\widetilde{Q}$-martingale gives

$$dC(t, S^*(t), V(t)) = C_{S^*} d(S^*)^{m,\widetilde{Q}}(t) + C_V V(t) g(V(t)) dW^{2,\widetilde{Q}}(t), \qquad (4.27)$$

with $W^{2,\widetilde{Q}}$ the $\widetilde{Q}$-Brownian motion originating from W^2 and with $(S^*)^{m,\widetilde{Q}}$ the $\widetilde{Q}$-martingale part of S^* with the following dynamics

$$d(S^*)^{m,\widetilde{Q}}(t) = S^*(t)^{\gamma+1} f(V(t))[\sqrt{1-\rho^2}dW^{1,\widetilde{Q}}(t) + \rho dW^{2,\widetilde{Q}}(t)]$$
$$= S^*(t)(\frac{dS(t)}{S(t)} - (\mu - r)dt),$$

with $W^{1,\widetilde{Q}}$ the $\widetilde{Q}$-Brownian motion originating from W^1. The angle bracket is given by

$$d\langle S, S\rangle_t^{\widetilde{Q}} = S(t)^2 S^*(t)^{2\gamma} f^2(V(t))dt.$$

The number of risky assets invested in S originating from the GKW decomposition is then given by

$$\xi_t^{\text{GKW}} = \frac{d\langle C, S\rangle_t^{\widetilde{Q}}}{d\langle S, S\rangle_t^{\widetilde{Q}}} = \frac{C_{S^*} d\langle (S^*)^{m,\widetilde{Q}}, S\rangle_t^{\widetilde{Q}} + C_V S^*(t)^\gamma S(t) g(V(t)) f(V(t)) V(t)\rho dt}{d\langle S, S\rangle_t^{\widetilde{Q}}}$$
$$= \left(C_{S^*} + \rho \frac{V(t)g(V(t))}{S^*(t)^{\gamma+1} f(V(t))} C_V \right) \frac{S^*(t)}{S(t)}. \tag{4.28}$$

Hence due to the deterministic short rate r the amount we have to invest in S^* equals:

$$C_{S^*} + \rho \frac{V(t)g(V(t))}{S^*(t)^{\gamma+1} f(V(t))} C_V,$$

which is exactly the amount given by Poulsen et al. (2009).
Next, we determine the number of risky assets invested in S under the original measure P. The dynamics of the P-martingale part I of (4.27) are

$$dI(t) = C_{S^*} d(S^*)^{m,P}(t) + C_V V(t) g(V(t)) dW^2(t),$$

where $d(S^*)^{m,P}(t) = \frac{S^*(t)}{S(t)} dM(t)$. Taking the P-angle bracket of M with I gives

$$d\langle I, M\rangle_t^P = C_{S^*} d\langle (S^*)^{m,P}, M\rangle_t^P + C_V V(t) g(V(t)) d\langle W^2, M\rangle_t^P,$$

with $d\langle (S^*)^{m,P}, M\rangle_t^P = \dfrac{S^*(t)}{S(t)} d\langle M, M\rangle_t^P$ and $d\langle W^2, M\rangle_t^P = S^*(t)^\gamma S(t) f(V(t))\rho dt$.
Therefore the units of stock for the LRM hedging strategy equal

$$\xi_t^{\text{FS}} = \frac{d\langle I, M\rangle_t^P}{d\langle M, M\rangle_t^P} = (C_{S^*} + \rho \frac{V(t)g(V(t))}{S^*(t)^{\gamma+1} f(V(t))} C_V) \frac{S^*(t)}{S(t)},$$

which is exactly the number found from the determination of the GKW decomposition, given in (4.28). This demonstrate again the equality between the GKW decomposition under the MMM and the FS decomposition under the original measure in the continuous case.
We remark that we can easily extend this model in several ways. We can add for example jumps or use a vector to drive the stochastic volatility. We will not do this explicitly for this setting, because all these stochastic volatility models can be seen as a special case of non-traded assets, which we will investigate in more detail in Chapter 9. Namely choose for $S^{(1)}$ the risky asset and for $(S^{(2)}, \ldots, S^{(d)})$ the (vector) of stochastic volatilities, furthermore as weights we take the vector $(1, 0, \ldots, 0)$.

There are risks and costs to a program of action. But they are far less than the long-range risks and costs of comfortable inaction.

John F. Kennedy (1917-1963)

5 Mean-variance hedging

In a complete market every contingent claim H is attainable and hence can be replicated by a self-financing strategy φ for which $V_T(\varphi) = H$. In an incomplete market there are contingent claims which are non-attainable. For these claims we can no longer insist on both conditions, but we need to relax at least one of them. The locally risk-minimizing hedging strategy, see Chapter 4, equals the contingent claim at maturity T, but is not self-financing, only mean-self-financing. The mean-variance hedging (MVH) strategy holds on to the self-financing condition and tries to minimize the difference between the contingent claim and the portfolio at maturity T. If a strategy is self-financing then intermediate costs or gains are avoided and only at the starting time 0 and at time T an investment is necessary. This type of strategies are of the following form at time $t \in [0, T]$:

$$V_0 + G_t(\widetilde{\Theta})$$

with $\widetilde{\Theta}$ e.g. the class Θ described in (2.43), but also other choices are possible. Often mean-variance hedging is called the global risk-minimizing hedging strategy. This property of minimizing the risk globally is of course a big advantage for this strategy, but it makes it also much harder to determine the strategy, as

will be shown in the next section. In this chapter we will only give an overview of the theory and applications concerning the MVH strategy. We will not add any new results in this chapter, but we will apply the theory to find the MVH strategy in the interest rate derivatives market, see Chapter 8.
The mean-variance hedging strategy is determined by projecting the square-integrable contingent claim on the space $\mathbb{R} + G_T(\widetilde{\Theta})$. This illustrate the importance of the possible closedness of e.g. the space $G_T(\Theta)$ in $L^2(P)$.
It was Bouleau and Lamberton (1989) who determined for the first time a self-financing portfolio minimizing a quadratic criterion in case the underlying process is a martingale and when it is a function of a Markov process. In a correlated Brownian motion setting, Duffie and Richardson (1991) searched for a mean-variance hedging strategy for a claim, depending on one Brownian motion, when hedging is only allowed in the correlated other. An extension to more general types of claims is given in Schweizer (1992).
We recall from Chapter 4 that the number of risky assets of the LRM hedging strategy in the martingale case will be the same as for the MVH strategy. Assume we have a square-integrable contingent claim H with GKW decomposition:

$$H = H_0 + (\xi \cdot X)_T + L_T, \tag{5.1}$$

then the number of risky assets at time t will equal ξ_t, while the amount of riskless assets follows from the self-financing assumption.
In the next section we will investigate the development of the determination of the MVH strategy. Afterwards we will give a short overview of the cases for which an explicit formula for the optimal number of risky assets is available.

5.1 Determination of the MVH strategy

In some papers there appear different formulations for the mean-variance hedging problem. To avoid any confusion we repeat the optimization problem we assume to work with: If H is a square-integrable and $\mathcal{F}_T$-measurable random variable, then the goal is to minimize over all $c \in \mathbb{R}$ and over all $\theta \in \widetilde{\Theta}$ the following expectation:

$$\min E[(H - c - \int_0^T \theta_s dX_s)^2]. \tag{5.2}$$

Remark that it is not always necessary to vary c, sometimes we will minimize over θ for a fixed starting amount c. Until Gourieroux et al. (1998) the space Θ as defined in (2.43) was always chosen for $\widetilde{\Theta}$.

5.1.1 Discrete time

We start with defining the discrete version of the MVT process:

$$\frac{(E[\Delta X_t|\mathcal{F}_{t-1}])^2}{\text{Var}[\Delta X_t|\mathcal{F}_{t-1}]}.$$

Schäl (1994) proved that when this ratio is deterministic then the optimization problem has a solution for every fixed starting amount $c \in \mathbb{R}$. Schweizer (1995b) extended this result to bounded MVT processes for one-dimensional processes, because if the MVT process is bounded then the space $G_T(\Theta)$ is closed in $L^2(P)$ and therefore it is possible to determine the MVH strategy. Schweizer also gives an explicit counterexample for which the MVT process is not bounded and there is no solution to the optimization problem. Another extension of Schäl (1994) to a non-deterministic ratio is given by Hipp (1993). An application of Schweizer (1995b) to affine stochastic volatility models by using Laplace transforms is given in Kallsen et al. (2009).
If we want to determine the optimal number c, then we first need to find the variance-optimal martingale measure (VOMM) Q^*, see Section 2.3.2.2, because

$$c = E^{Q^*}[H], \tag{5.3}$$

if the MVH problem has a solution, see Proposition 2 in Schweizer (1996). This result also holds for the continuous time case.
Another reference for the discrete time case is Melnikov and Nechaev (1999). We will discuss this no further, because we want to concentrate on the continuous time setting.

5.1.2 Continuous time

The first extensions to continuous time with $\widetilde{\Theta} = \Theta$ are due to Schweizer (1994) and Monat and Stricker (1994, 1995). They assumed that the underlying process is a special d-dimensional semimartingale for which the finite variation part of each component is absolutely continuous with respect to the angle bracket of the respectively chosen martingale part. Hence they obtain the first relationship of the structure condition (2.5). In Schweizer (1994) the extended MVT process defined as

$$\widetilde{K}_t = \int_0^t \frac{\lambda'_s}{1 + \lambda'_s \Delta\langle M\rangle_s \lambda_s} d\langle M\rangle_s \lambda_s$$

with λ as in (2.5), is assumed to be finite. This condition is called the extended structure condition. If $\widetilde{K}$ is also deterministic and if the FS decomposition exists then the MVH problem has the following solution:

$$\xi_t^{(c)} = \xi_t + \frac{\lambda_t}{1 + \lambda'_t \Delta\langle M\rangle_t \lambda_t}(V_{t-}^H - c - \int_0^{t-} \xi_s^{(c)} dX_s) \tag{5.4}$$

with V_t^H as defined in Chapter 4, namely equal to $E^{\widetilde{Q}}[H|\mathcal{F}_t]$ where $\widetilde{Q}$ is the MMM and with ξ the number of risky assets from the FS decomposition, see Definition 3.1.1, of the contingent claim H. We remark that in the case of QLC processes the extended MVT process equals the MVT process, see Definition 2.2.18.

The existence of the MVH strategy is closely linked with the closedness of the space $G_T(\Theta)$. If the space $G_T(\Theta)$ is closed then it is possible to uniquely project a contingent claim H, belonging to $L^2(P)$ on the closed subspace $G_T(\Theta)$. Monat and Stricker (1994, 1995) reduced the condition to the uniformly boundedness of the MVT process and no longer to the extended. This uniform boundedness is sufficient but not necessary as is shown with a counterexample in Monat and Stricker (1995).

In Section 2.3.2.2 we gave already some results concerning the existence of the VOMM. We will add results concerning the closedness of $G_T(\Theta)$. In later papers, the problem of the closedness was avoided by changing the class of processes $\widetilde{\Theta}$ to a class which is closed in $L^2(P)$ without any condition.

Delbaen et al. (1997) proved that in the case of a continuous semimartingale such that there exists a locally square-integrable equivalent martingale measure, the following assertions are equivalent:

- $G_T(\Theta)$ is closed in $L^2(\Omega, \mathcal{F}, P)$.
- There is an equivalent local martingale measure that satisfies the reverse Hölder inequality, see Definition 3.1.4 with $p = 2$.
- The variance-optimal local martingale measure is equivalent to P and satisfies the reverse Hölder inequality with $p = 2$.

A simpler proof for the closedness of $G_T(\Theta)$ in case the MVT process is bounded and continuous (hence X is QLC) is given in Pham et al. (1998).
For general semimartingales belonging to class of $\mathcal{E}$-martingales sufficient condition for $G_T(\Theta)$ to be closed in $L^2(P)$ are given in Choulli et al. (1998). Namely $\mathcal{E}(\widetilde{N}) = \mathcal{E}(-\lambda \cdot M)$ should be regular and has to satisfy the reverse Hölder inequality of order 2.
To determine the VOMM in the continuous time setting Schweizer (1996) defined the adjustment process β as:

Definition 5.1.1. A process $\beta \in L(X)$ is called an **adjustment process for** X if the process $\beta\mathcal{E}(\beta{\cdot}X)_-$ belongs to Θ and if the random variable $Z^* := \mathcal{E}(-\beta{\cdot}X)_T$ is in $G_T(\Theta)^\perp$. That is

$$E[Z^* G_T(\theta)] = 0 \quad \text{for all } \theta \in \Theta.$$

Using this definition the VOMM is defined in the following way:

Proposition 5.1.2 (See Schweizer (1996)). *Assume that the space $P_X(\Theta)$, see page 30, is non-empty. If β is an adjustment process for X, then Q^* defined by:*

$$\frac{dQ^*}{dP} := \frac{Z^*}{E[Z^*]}$$

is in $P_X(\Theta)$ and variance-optimal.

The process β is characterized in the following theorem:

Theorem 5.1.3 (See Schweizer (1996)). *Assume that the space $P_X(\Theta)$, see page 30, is non-empty. Then there exists an adjustment process β for X if and only if there exists a solution $(\beta, U) \in L(X) \times \mathcal{S}^2$ to the backward stochastic differential equation*

$$dU_t = -U_{t-}\beta_t dX_t, \quad U_T = \pi(1) \tag{5.5}$$

with U_0 deterministic and π the projection in $L^2(P)$ on $G_T(\Theta)^\perp$. More precisely, $\beta \in L(X)$ is an adjustment process for X if and only if $U := \mathcal{E}(\beta \cdot X)$ is in $\mathcal{S}^2$ and (β, U) solves (5.5).

The amount c is also in the continuous time case determined by (5.3), see Schweizer (1996). Concerning the optimization problem, Schweizer (1996) gives the following proposition:

Proposition 5.1.4 (See Schweizer (1996)). *Assume that there exists an adjustment process β for X. If $(\rho, Z) \in L(X) \times \mathcal{S}^2$ is a solution to*

$$dZ_t = \rho_t dX_t - Z_{t-}\beta_t dX_t, \quad Z_T = H - \pi(H)$$

with Z_0 deterministic, then

$$\xi_t^{(c)} = \rho_t - \beta_t(c + \int_0^{t-} \xi_s^{(c)} dX_s)$$

belongs to Θ for every $c \in \mathbb{R}$ and is the solution to the optimization problem (5.2).

5.1.2.1 The case of a continuous underlying

In Theorem 2.3.11 we remarked already that if there exists at least one equivalent local martingale measure with square-integrable density then in the continuous case the variance-optimal martingale measure is a probability measure equivalent with the original measure P.
If X is continuous and the VOMM equals the MMM, the equation in feedback form for the optimal number of risky assets is made more concrete by Pham et al. (1998) in case the contingent claim H belongs to $L^{2+\epsilon}(P)$ with $\epsilon > 0$ and the MVT process is bounded:

$$\xi_t^{(c)} = \xi_t - \frac{\varsigma_t}{\hat{Z}_t^*}(V_{t-} - c - \int_0^t \xi_s^{(c)} dX_s), \tag{5.6}$$

where

$$\hat{Z}_t^* := E^{Q^*}[Z_T^* | \mathcal{F}_t] = E[(Z_T^*)^2] + \int_0^t \varsigma_s dX_s \tag{5.7}$$

and

$$V_t = E^{Q^*}[H | \mathcal{F}_t] = H_0 + \int_0^t \xi_s dX_s + L_t^H \tag{5.8}$$

the GKW decomposition of H under the VOMM which equals the MMM.
Rheinländer and Schweizer (1997) no longer require that H belongs to $L^{2+\epsilon}(P)$ with $\epsilon > 0$ nor that the VOMM equals the MMM. They are still able to prove the same result as Pham et al. (1998), but they project on the space $G_T(\Theta)$ (hence without the constant c) under the following conditions:

- X is continuous
- $G_T(\Theta)$ is closed in $L^2(P)$
- $\mathcal{D}^s \cap L^2(P) \neq \emptyset$ with $\mathcal{D}^s$ as defined in (2.22)
- $H \in L^2(P)$.

Up till now the solution is obtained as a correction of the number of risky assets from the GKW decomposition of H under the VOMM. Another approach to find the solution to the optimization problem is given by Gourieroux et al. (1998). By a combination of a change of measure and a change of coordinates the original optimization problem is transformed into a problem which can be directly solved by the GKW decomposition. In fact a hedging numéraire is added as an asset to trade in. This hedging numéraire and the VOMM is determined by Laurent and Pham (1999) using dynamic programming methods. They illustrate it on stochastic volatility examples.
We remark that Gourieroux et al. (1998) work with a larger space than Θ, namely $\hat{\Theta}$ which contains all $\mathbb{R}^d$-valued predictable X-integrable processes θ such that $\theta \cdot X$ is a martingale under the VOMM and $(\theta \cdot X)_T$ belongs to $L^2(P)$. It was already proved by Delbaen and Schachermayer (1996a) that the space $G_T(\hat{\Theta})$ is closed in $L^2(P)$.
In Gourieroux et al. (1998) the interest rate may be stochastic. A discussion comparing the results of Rheinländer and Schweizer (1997) and Gourieroux et al. (1998) is included in Rheinländer and Schweizer (1997).
Schweizer (2001) assumed that the d-dimensional continuous process X is such that X^* is locally square-integrable under the original measure P and then proved that formula (5.6) still holds using the class $\hat{\Theta}$ if $\mathcal{D}^e \cap L^2(P) \neq \emptyset$ and with H a square-integrable claim. Furthermore he gives in that paper an overview of the main theorems and papers until 2000 for the mean-variance hedging strategy. Another paper which gives an overview until 2000 for the quadratic hedging strategies is Pham (2000). Also in this paper the focus to describe the VOMM and the MVH strategy is mainly on continuous processes.
An application to a specific setting of non-attainable claims, namely when the contingent claim depends on two correlated assets, but there is only one available for hedging, is given by Kohlmann and Peisl (2000). They use the concept of backward stochastic differential equations (BSDE) to find the optimal strategy. A generalization of the setting of Kohlmann and Peisl (2000) to allow for a stream of liabilities is given by Delong (2009). Lim (2004) uses BSDE, more precisely stochastic Riccati equations to find the MVH strategy in case of a multidimensional Brownian motion for the underlying process. Analogous results

to show how the mean-variance hedging problem can be seen and solved as a linear-quadratic stochastic control problem are given by Bobrovnytska and Schweizer (2004). Kohlmann and Xiong (2007) determines the strategy for a defaultable option by using BSDE's in a stochastic volatility model.
Biagini et al. (2000) determine the VOMM in case the underlying process follows a Brownian motion with stochastic volatility. They illustrate the technique on examples in which the volatility process is continuous or may contain jumps. For more general type of jump processes the VOMM and the MVH strategy is calculated in Biagini and Guasoni (2002).
A theoretical and numerical comparison between the MVH and the LRM strategy in the case of stochastic volatility models with continuous underlyings is given in Heath et al. (2001a, 2001b).
For the continuous case we can conclude that if $\mathcal{D}^e \cap L^2(P) \neq \emptyset$, then Delbaen and Schachermayer (1996b) showed that the VOMM is an equivalent probability measure and for every square-integrable contingent claim the MVH strategy is given by (5.6) with $c = E^{Q^*}[H]$ which is proved by Schweizer (2001).
The extension of the MVH strategy to the case that the underlying is discontinuous is reported in the next section.

5.1.2.2 More general processes for the underlying

Hubalek et al. (2006) is an extension of the preprint Hubalek and Krawczyk (1998). They describe the MVH strategy in discrete and continuous time in case the one-dimensional underlying process has stationary and independent increments. Under the assumption of a deterministic MVT process they write the contingent claim as a linear combination of exponential payoffs, which is possible by using the inverse Laplace transform and hence they obtain explicit strategies. We extend this approach to time-inhomogeneous Lévy processes in Chapter 8.
The MVH portfolio is determined in terms of Malliavin derivatives by Benth et al. (2003) in the case of a d-dimensional Lévy process with independent components.
Under additional assumptions, Arai (2005a) extended the results of Rheinländer and Schweizer (1997) and Gourieroux et al. (1998) to the discontinuous case. The additional assumptions are mainly related to the VOMM:

- VOMM exists as an equivalent probability measure;

- The density process Z^* linked with the change of measure to the VOMM satiesfies (R_2), see Definition 3.1.4;
- $\exists C$: $Z^*_- \leq CZ$.

Under these assumptions the optimal number of risky assets equals

$$\xi_t^{(c)} = \xi_t - \frac{\varsigma_t}{\hat{Z}_t^*}(V_{t-} - c - \int_0^{t-} \xi_u^{(c)} dX_u), \tag{5.9}$$

with $\hat{Z}^*$ as in (5.7) and V, ξ as in (5.8).
Arai (2005a) claims that the decomposition in (5.8) is not for sure a GKW decomposition. We agree on this point, but disagree with the way he proves it. He assumes that if $\hat{Z}^*XL$ is not a Q^*-local martingale then $[X, [\hat{Z}^*, L]]$ is not a Q^*-local martingale or equivalently $[X, [\hat{Z}^*, L]]$ is a Q^*-local martingale if $\hat{Z}^*XL$ is a Q^*-local martingale. By the product rule we find

$$\hat{Z}^*XL = \hat{Z}^*_- \cdot (XL) + (XL)_- \cdot \hat{Z}^* + [\hat{Z}^*, XL],$$

where

$$XL = X_- \cdot L + L_- \cdot X + [X, L] \tag{5.10}$$

$$\hat{Z}^* = \hat{Z}^*_0 + \widetilde{\zeta} \cdot X \tag{5.11}$$

and hence

$$\hat{Z}^*XL = \hat{Z}^*_- \cdot (XL) + (XL)_- \cdot \hat{Z}^* + X_- \cdot [\hat{Z}^*, L] + L_- \cdot [\hat{Z}^*, X] + [\hat{Z}^*, [X, L]]. \tag{5.12}$$

We know that $\hat{Z}^*$ and X are Q^*-local martingales. Since L is a R-local martingale (see Arai (2005a) for the definition of the measure R), $L\hat{Z}^*$ and in view of (5.11) also LX is a Q^*-local martingale. For LX this also follows from the fact that L is R-orthogonal to $Y = \dfrac{X}{\hat{Z}^*}$ and hence that $LX = LY\hat{Z}^*$ is a Q^*-local martingale. However this does not imply that $[\hat{Z}^*, L]$ is a Q^*-local martingale unless L is also a Q^*-local martingale. An assumption that is made in Arai (2005a) but which is not underpinned. According to us one can only prove that in his setting L is a Q^*-semimartingale. In the same way Arai (2005a) also uses that $[L, X]$ is a Q^*-local martingale which is equivalent to assuming that L is a Q^*-local martingale according to (5.10).

Let's assume that L or equivalently that $[L, X]$ or $[L, \hat{Z}^*]$, is a Q^*-local martingale then by (5.12) $\hat{Z}^* XL$ is a Q^*-local martingale if and only if $L_- \cdot [\hat{Z}^*, X] + [\hat{Z}^*, [X, L]]$ is a Q^*-local martingale. However $L_- \cdot [\hat{Z}^*, X]$ has finite variation and is not a Q^*-local martingale. Hence the Q^*-local martingale property of $\hat{Z}^* XL$ does not follow from the fact that $[X, [\hat{Z}^*, L]]$ is a Q^*-local martingale.
We prove the possible non-orthogonality of N to X under Q^* in the following way:

$$[X, N] = \hat{Z}^*_- \cdot [X, L] + [X, [\hat{Z}^*, L]] = \hat{Z}^*_- \cdot [X, L] + (\Delta X \cdot [\hat{Z}^*, L])$$

due to Properties 2.2.15(1). This proves that $[X, N]$ is a Q^*-local martingale if $[X, L]$ (and hence also $[\hat{Z}^*, L]$) is a Q^*-local martingale and this will only be true if and only if L is a Q^*-local martingale by (5.10).

Arai (2005b) looked in more detail for sufficient conditions to obtain the properties of the VOMM needed in (5.9) for discontinuous semimartingales. Lim (2004, 2005) and references therein extended the use of BSDE's to discontinuous processes whose characteristics are still adapted to a Brownian filtration.
Černý and Kallsen (2007) introduced a new measure, namely the opportunity neutral measure, whose minimal martingale measure equals the VOMM of the original measure. Hence once the opportunity neutral probability measure is determined, it is easy to find the VOMM. This opportunity neutral measure depends on the opportunity process L, which is $(0, 1]$-valued and whose characteristics are described in Theorem 3.25 of Černý and Kallsen (2007). We remark that they also fix the problem of choosing the correct space $\widetilde{\Theta}$, namely they chose the closure of the set of simple strategies as also Gourieroux et al. (1998) did in the continuous case. Both were inspired by Delbaen and Schachermayer (1996a).
They avoid the need to prove that the VOMM is equivalent or that it is a probability measure by using the concept of σ-martingales, whose properties resembles those of the $\mathcal{E}$-martingales as defined in Choulli et al. (1998) and hence the portfolio at time t is defined as

$$V_t = E(H\mathcal{E}(N - N^t)_T | \mathcal{F}_t)$$

with $\mathcal{E}(N)$ describing the change of measure from the original measure to the VOMM. This is rather similar to (3.48) without the appropriate sequence of stopping times.
The optimal number of risky assets is then given by

$$\xi^{(c)}_t = \xi_t - \widetilde{a}_t (c + \int_0^{t-} \xi^{(c)} \cdot X - V_{t-}) \tag{5.13}$$

with ξ the number of risky assets from the FS decomposition of the contingent claim under the opportunity neutral probability measure, $c = V_0$ and $\widetilde{a}$ is the adjustment process, which is also characterized in Theorem 3.25 of Černý and Kallsen (2007). We remark that the opportunity neutral measure is not necessarily a martingale measure. Hence the determination of the MVH strategy is reduced to the determination of the opportunity process L and the adjustment process $\widetilde{a}$.
Černý and Kallsen (2007) discuss also the connections to the literature of their results.
A first application of Černý and Kallsen (2007) is given in Černý and Kallsen (2008a). In case of a Heston model with correlation they calculate the VOMM using the opportunity neutral measure and derive formulas for the hedging strategy and the hedging error. Kallsen and Vierthauer (2009) applied the technique described by Černý and Kallsen to affine processes, including a whole range of stochastic volatility models in order to find the MVH strategy and the hedging error by applying Laplace transform techniques.
In the martingale case, Kallsen and Pauwels (2009a) determine semi-explicit formulas for the MVH strategy and the hedging error in case of affine stochastic volatility models. The formulas obtained there allow for a more efficient numerical computation due to the integral representation. Furthermore in Kallsen and Pauwels (2009b) they illustrate the obtained results on more concrete models.
Chan et al. (2009) can be seen as an extension of Biagini et al. (2000) in that they determine the VOMM for stochastic volatility models driven by Lévy processes, which may be correlated, unfortunately the explicit example is only given for the uncorrelated case.

5.2 Conclusion

In some cases we can easily determine the MVH strategy, we give an overview:

- If X is a martingale, then the MVH strategy is determined from the GKW decomposition of the claim under the original measure, see (5.1).
- If X is a continuous semimartingale, then we use (5.6) for which we first need to determine the VOMM and then the GKW decomposition of H under the VOMM.

- If the VOMM equals the MMM measure (e.g. in case the MVT process is deterministic), then the adjustment process $\widetilde{a} = 0$, because $L = 1$ and from (5.13) we see that the MVH strategy is found through the FS decomposition of the claim under the original measure.
- If we work with stochastic volatility models, then we can apply Černý and Kallsen (2007), because in the case of stochastic volatility it is more straightforward to obtain the positiveness of the process L. The optimal number of risky assets is then given by (5.13).
- If the VOMM is an equivalent martingale measure, then we can follow the approach of Arai (2005a), but his results are not easy to apply to concrete examples, because it is very hard to prove that the VOMM is equivalent. A possible way out is to solve the BSDE described by Kohlmann et al. (2010).

Remark that there remains a gap between the theory for the MVH strategy and its implementation. In Chapter 9, we will discuss the difficulties encountered during such implementation in the setting of non-traded assets.

Fun is like life insurance; the older you get, the more it costs.
Kin Hubbard (1868-1930)

6 Risk-minimization for unit-linked life insurance contracts with surrender option

6.1 Introduction

In a unit-linked life insurance contract the benefits and also the premiums may depend on the price of some specific traded stock or index. Hence such a contract can be seen as a combination of an insurance and a mutual fund, where the premium is invested in a number of *units* of the fund.
Møller (2001) describes the hedging strategy for unit-linked life insurance contracts with only a mortality option when the risky asset follows a geometric Brownian motion. Of course, in his setting it is logical to assume independence between the time of death and the financial market. We assume that the benefit at maturity is paid out also in case the policyholder is no longer alive. Hence we no longer have a mortality option, but instead we will add a surrender option to the life insurance contract. This option gives the owner of the insurance the right to relinquish the contract against the surrender value. Hence, it is no longer realistic to assume that this time is independent of the financial market,

in contrast to the mortality time. But complete dependence on the financial market is too restrictive since there can be more reasons to decide to surrender than just financial ones. In the research concerning credit risk, we find a good background to describe this partial dependence and handle this extra difficulty. An extensive overview of the credit risk theory can be found in the book of Bielecki and Rutkowski (2002).

Barbarin (2007) already applied some concepts from credit risk to describe the time of surrender, when the time is not a stopping time with respect to the filtration generated by the financial market. The goal here is to make his results more explicit and therefore extend the article of Møller (2001) and Riesner (2006a). We also refer to Biagini and Cretarola (2006, 2007, 2009) and Cretarola (2007), they determine quadratic hedging strategies for different types of defaultable claims when the underlying follows a geometric Brownian motion mainly by using the representation theorem. Unfortunately this does not give explicit results.

The goal of this chapter is to show how a risk-minimizing hedging strategy is determined for a unit-linked contract which combines financial and insurance risk. In this chapter the insurance risk is caused by the surrender event, while in Chapter 7 the insurance risk is due to the mortality risk. We will first determine the strategy for the case the underlying asset follows a geometric Brownian motion. In this setting the risk is completely diversifiable when the number of policyholders goes to infinity. In the second setting we use a geometric Lévy process and in this discontinuous case we will show that the risk is no longer diversifiable.

The basic contract that we will hedge contains fixed premiums, a continuous payment up to surrender and a payment at surrender or at maturity whichever comes first. We do not introduce a benefit in case of death, because then we would also need to know the dependence structure between mortality and surrender. Another way out would be to model the time of death independent of, but in the same way as, the surrender time. Hence the intensity λ^s, see page 118, is replaced by the sum of the intensity describing the surrender time and the intensity describing the time of death, as was done in Bacinello et al. (2009). Of course, the payment process, described in Section 6.2.3, will then be extended with a mortality benefit.

For the martingale process describing the one-dimensional risky asset we look at the two most popular choices: firstly, a geometric Brownian motion and secondly, a geometric Lévy process. In the Brownian motion case we extend the theory of Møller (2001), in that the random times of payment are no longer independent of the financial market. Riesner (2006a) determined the risk-

minimizing hedging strategy for life insurance contracts with a mortality option when the underlying follows a Lévy process. Therefore in the second case that we consider, the mortality option of the framework of Riesner (2006a) is replaced by a surrender option.
The quadratic hedging strategy, we apply here, is the risk-minimizing hedging strategy due to the martingale setting we assume to work with. In order to describe the number of risky assets and the amount of the riskless asset to be held in the hedging portfolio, we have to find the Galtchouk-Kunita-Watanabe decomposition of the claim, see Chapter 4 for more explanation. Hereto, we have to impose additional assumptions on the hazard process. This process describes the influences apart from the one of the financial market on the surrender time. A sufficient condition is the absolute continuity of the hazard process, which is not at all a strong condition from the point of view of practitioners.
Section 6.2 contains an overview of the setting and the assumptions made in this paper. The risk-minimizing hedging strategy for the geometric Brownian motion case is studied in Section 6.3. In that section, we first determine the portfolio for one policyholder, then extend this result to n policyholders and finally determine the risk process. In Section 6.4 we deal with the risk-minimizing hedging strategy when the risky asset is driven by a geometric Lévy process.
This chapter is based on Vandaele and Vanmaele (2009).

6.2 The theoretical setting

We assume that we work on the complete probability space $(\Omega, \mathcal{F}, Q)$ for the financial market with $\mathcal{F}$ the natural filtration generated by the risky asset X, whose dynamics are described in Section 6.2.1. We use the measure Q because we assume that the price processes are already martingales under the original measure.

6.2.1 Price process

We will first determine the risk-minimizing hedging strategy in case the dynamics of the underlying risky asset are described by a geometric Brownian motion and then extend the results to the case of a geometric Lévy process. We remark

that the risk-minimizing hedging strategy in the Lévy case can easily be further extended to the case of a general semimartingale as underlying.
The **geometric Brownian motion** was introduced to ensure the positiveness of the price process by using the Doléans-Dade exponential, namely $Y = \mathcal{E}(\alpha W)$ with W a standard Brownian motion or equivalently

$$dY_t = Y_t \alpha dW_t$$

and hence using (2.17) $Y_t = e^{\alpha W_t - \frac{\alpha^2}{2}t}$. By taking the Doléans-Dade exponential of a Lévy process, which is defined in Section 2.5, we extend this to **geometric Lévy processes**.
We will assume that the riskless asset B has the following dynamics under the measure Q:

$$dB_t = r(t, S_t)B_t dt,$$

with S the dynamics of the undiscounted risky asset.
As a first case we model the price process of the underlying asset by a geometric Brownian motion. We furthermore require that the discounted asset is a martingale under Q, hence the dynamics of S are given by

$$dS_t = r(t, S_t)S_t dt + \sigma(t, S_t)S_t dW_t,$$

as was also assumed in Møller (2001). The dynamics of the discounted risky asset, denoted by X, then are

$$dX_t = d(\frac{S_t}{B_t}) = \sigma(t, S_t)S_t B_t^{-1} dW_t = \sigma(t, S_t)X_t dW_t.$$

Due to the martingale property of the discounted risky asset, we know that the original measure is already a martingale measure and there is no need to perform a change of measure.
Secondly, we assume the risky asset is driven by a geometric càdlàg version of a Lévy process. Therefore the process of the discounted risky asset under the measure Q is given by

$$dX_t = \sigma_t X_{t-} d(cW_t + M_t) \quad \text{where } M_t = \int_0^t \int_{\mathbb{R}} x M(ds, dx) \tag{6.1}$$

is a square-integrable martingale under Q and $M(ds, dx)$ denotes the compensated Poisson random measure on $[0, \infty) \times \mathbb{R}\backslash\{0\}$.
It is still possible to apply the theory of Møller for general payment streams,

because his proofs aren't restricted to the case of the geometric Brownian motion, but are in fact holding for locally square integrable local martingales. We gave already in Chapter 4 a more extended form of the result of Møller (2001), namely the LRM hedging strategy in the case of payment streams obtained by Schweizer (2008) and repeated in Proposition 4.3.10. This result reduces to the risk-minimizing hedging strategy described below for a payment stream A_t, $t \in [0, T]$ if we work with martingales under the original measure Q. First of all, suppose the Galtchouk-Kunita-Watanabe decomposition of $E^Q[A_T|\mathcal{F}_t]$ is given by

$$V_t^* = E^Q[A_T|\mathcal{F}_t] = V_0^* + \int_0^t \xi_u^A dX_u + L_t^Q, \tag{6.2}$$

where L_t^Q is a square-integrable Q-martingale orthogonal to X. The risk-minimizing strategy is then the 0-admissible strategy $\varphi = (\xi, \eta)$ given by

$$(\xi_t, \eta_t) = (\xi_t^A, V_t^* - A_t - \xi_t^A X_t), \tag{6.3}$$

with 0-admissible in the sense of Definition 4.1.1. Compared with the risk-minimizing strategy for a contingent claim $H = A_T$, we see that the number of risky assets is exactly the same, while the amount invested in the riskless asset is adjusted for the payments A_t, namely those payments which are already paid out.

6.2.2 The theory of credit risk

The theory of credit/default risk is very useful when modeling a surrender option, because the surrender time has similarities with the default time. Both times depend on the information from the financial market, but not completely and it is exactly this partial dependency that we will model using concepts originally defined in the context of defaultable claims.
In fact in credit risk there are two main approaches: the *structural approach* and the *intensity-based approach*.
We will concentrate on the second approach, because in the structural approach the basic assumption is that the default time is a stopping time with respect to the information of the financial market. In the intensity-based, or also called the *reduced-form*, approach it is possible that the default time is not adapted to the filtration generated by the prices. The results given here are mostly based on the

following books/articles: Jeanblanc and Rutkowski (1999, 2000), Bielecki and Rutkowski (2002). For more background information concerning credit risk, we also refer to these references.
We will apply the results immediately to the surrender time, while in the original papers, they use of course the default time instead.

6.2.2.1 The hazard process

We assume that the surrender time T^s, which is a non-negative random variable, is not a $\mathbb{F}$-stopping time. Furthermore we assume that $Q(T^s = 0) = 0$ and $Q(T^s > t) > 0$, for all $t \in \mathbb{R}^+$. With the surrender time we associate the increasing and càdlàg process

$$H_t = \mathbb{1}_{\{T^s \leq t\}},$$

which is zero before default and which equals 1 after default. We denote by $\mathbb{H}$ the natural complete filtration $\mathcal{H}_t = \sigma(H_u, u \leq t)$. The filtration, combining the information from the financial market with the information of the surrender time, is denoted by $\mathbb{G} = (\mathcal{G}_t)_{0 \leq t \leq T}$ where $\mathcal{G}_t = \sigma(\mathcal{F}_t \vee \mathcal{H}_t)$. The filtration $\mathbb{G}$ is strictly larger than the filtration $\mathbb{F}$, because T^s is not an $\mathbb{F}$-stopping time.
The extension of the previous filtrations to the case of n policyholders is denoted by $\mathbb{G}^n$. If we denote by $H_t^i = \mathbb{1}_{\{T_i^s \leq t\}}$ and by $\mathcal{H}_t^i = \sigma(H_u^i, u \leq t)$, then the filtration $\mathbb{G}^n$ equals $(\mathcal{G}_t^n)_{0 \leq t \leq T}$, where $\mathcal{G}_t^n = \mathcal{F}_t \vee \mathcal{H}_t^1 \vee \cdots \vee \mathcal{H}_t^n$. For $n = 1$ we are back to the case of one policyholder and the filtration $\mathbb{G}^1 = \mathbb{G}$.
The results given here are formulated for the case of one policyholder. The surrender time T^s is of course a $\mathbb{G}$-stopping time, but not surely an $\mathbb{F}$-stopping time. Hence the process H is $\mathbb{G}$-adapted, but not necessarily $\mathbb{F}$-adapted. Therefore we introduce the process F:

$$F_t = Q(T^s \leq t | \mathcal{F}_t) = E^Q[\mathbb{1}_{\{T^s \leq t\}} | \mathcal{F}_t],$$

which is a bounded, non-negative $\mathbb{F}$-submartingale. We can easily prove that F is a submartingale by the towerlaw and by using that $\mathbb{1}_{\{\tau \leq t\}} \leq \mathbb{1}_{\{\tau \leq s\}}$ for every $t \leq s$:

$$E^Q[F_s | \mathcal{F}_t] = E^Q[\mathbb{1}_{\{T^s \leq s\}} | \mathcal{F}_t] \geq E^Q[\mathbb{1}_{\{T^s \leq t\}} | \mathcal{F}_t] = F_t.$$

According to Remark I.1.37 of Jacod and Shiryaev (2002) we can assume that we work with the right-continuous modification of F.

If $F_t < 1$ for every $t \in \mathbb{R}^+$ (this means T^s may not be an $\mathbb{F}$-stopping time), then we can define the $\mathbb{F}$**-hazard process** Γ, where

$$1 - F_t = e^{-\Gamma_t} \quad \text{or equivalently} \quad \Gamma_t = -\ln(1 - F_t) \quad \forall t \in \mathbb{R}^+.$$

We repeat Lemma 6.1 and Lemma 6.2 of Jeanblanc and Rutkowski (1999):

Lemma 6.2.1. *For any $\mathcal{G}$-measurable bounded random variable Y we have, for any $t \in \mathbb{R}^+$,*

$$E^Q[\mathbb{1}_{\{T^s>t\}} Y|\mathcal{G}_t] = \mathbb{1}_{\{T^s>t\}} \frac{E^Q[\mathbb{1}_{\{T^s>t\}} Y|\mathcal{F}_t]}{Q(T^s > t|\mathcal{F}_t)} = \mathbb{1}_{\{T^s>t\}} E^Q[\mathbb{1}_{\{T^s>t\}} e^{\Gamma_t} Y|\mathcal{F}_t],$$

for any $t \leq s$

$$Q(t < T^s \leq s|\mathcal{G}_t) = \mathbb{1}_{\{T^s>t\}} \frac{Q(t < T^s \leq s|\mathcal{F}_t)}{Q(T^s > t|\mathcal{F}_t)} = \mathbb{1}_{\{T^s>t\}} E^Q[1 - e^{\Gamma_t - \Gamma_s}|\mathcal{F}_t],$$

$$E^Q[\mathbb{1}_{\{T^s>s\}} Y|\mathcal{G}_t] = \mathbb{1}_{\{T^s>t\}} E^Q[Y \mathbb{1}_{\{T^s>s\}} e^{\Gamma_t}|\mathcal{F}_t]$$

and for any $\mathcal{F}_s$-measurable random variable X, we have

$$E^Q[\mathbb{1}_{\{T^s>s\}} X|\mathcal{G}_t] = \mathbb{1}_{\{T^s>t\}} E^Q[X e^{\Gamma_t - \Gamma_s}|\mathcal{F}_s]. \tag{6.4}$$

Especially (6.4) will be extremely useful. According to Lemma 6.3 and Proposition 6.2 of Jeanblanc and Rutkowski (1999), we can associate a martingale with the hazard process Γ:

Lemma 6.2.2. *The process $L_t = \mathbb{1}_{\{T^s>t\}} e^{\Gamma_t} = (1 - H_t) e^{\Gamma_t}$ follows a $\mathbb{G}$-martingale. If furthermore the $\mathbb{F}$-hazard process Γ follows a continuous process of bounded variation, then the process $\hat{M} = H - \Gamma_{\cdot \wedge T^s}$ follows a $\mathbb{G}$-martingale and L satisfies*

$$L_t = 1 - \int_{]0,t]} L_{u-} d\hat{M}_u.$$

In literature they frequently assume that the hazard process is absolutely continuous. We will also make this assumptions, and this implies that we can make the process $\hat{M}$ defined in Lemma 6.2.2 more explicit. If the hazard process is absolutely continuous then we can assume that there exists an $\mathbb{F}$-progressively measurable[1] process λ^s, called the **$\mathbb{F}$-intensity** of the random time T^s, such that $\Gamma_t = \int_0^t \lambda_u^s du$. The process $\hat{M}$ given by

$$\hat{M}_t = H_t - \int_0^{t\wedge T^s} \lambda_u^s du = H_t - \int_0^t \mathbb{1}_{\{T^s>u\}} \lambda_u^s du = H_t - \int_0^t \mathbb{1}_{\{T^s\geq u\}} \lambda_u^s du \quad (6.5)$$

follows a $\mathbb{G}$-martingale.
Furthermore we also assume, as in Azéma et al. (1993), one of the following conditions:

- Any $\mathbb{F}$-martingale is continuous.
- For any $\mathbb{F}$-stopping time θ, $P(T^s = \theta) = 0$.

These set of conditions are denoted by (C) by Blanchet-Scalliet and Jeanblanc (2004)

6.2.2.2 The (H)-hypothesis

An important hypothesis often assumed in literature is the (H)-hypothesis. This hypothesis is guaranteed if and only if the surrender time is modeled by a so-called Cox process. Denote by ξ an exponentially distributed random variable which is independent of the continuous hazard process Γ, then the surrender time is modeled by

$$T^s = \inf\{t \in \mathbb{R}^+ | \Gamma_t \geq \xi\}.$$

For a proof of this equivalence between the (H)-hypothesis and the modeling of the hazard process through a Cox process, we refer to Blanchet-Scalliet and Jeanblanc (2004).
It is not always realistic to assume the (H)-hypothesis, therefore recent literature as e.g. Jeanblanc and Le Cam (2009) try to avoid the use of the (H)-hypothesis by working with initial times. For these times the **(H')-hypothesis** will also hold. This means that every $\mathbb{F}$-semimartingale will be a $\mathbb{G}$-semimartingale.
The **(H)-hypothesis** is an extension of the (H')-hypothesis and demands that every square-integrable $\mathbb{F}$-martingale is a square-integrable $\mathbb{G}$-martingale. Furthermore it is equivalent with the condition that for every t, the σ-fields $\mathcal{F}_\infty$ and

[1] A process X is $\mathbb{F}$-progressively measurable if for each $t \in \mathbb{R}^+$ the mapping $(\omega, s) \to X(\omega, s)$ of $\Omega \times [0, T]$ into $\mathbb{R}$ is measurable with respect to $\mathcal{B}([0, T]) \otimes \mathcal{F}_t$, see Protter (2005).

$\mathcal{G}_t$ are conditionally independent given $\mathcal{F}_t$.
For the extension to n policyholders, we assume the (H)-hypothesis holds under Q between the filtration $\mathbb{G}^n$ and the subfiltration $\mathbb{F}$. This means that for every t, $\mathcal{F}_\infty$ and $\mathcal{G}_t^n$ are conditionally independent with respect to $\mathcal{F}_t$. This condition guarantees that the financial market is also arbitrage-free under Q with respect to $\mathbb{G}^n$.
It is possible to prove that the (H)-hypothesis will also hold between $\mathbb{F}$ and any filtration $\mathbb{K}$, such that $\mathbb{F} \subseteq \mathbb{K} \subseteq \mathbb{G}^n$.
We calculate the bracket process $[\hat{M}, \hat{M}]$ in case the hazard process is absolutely continuous, see (6.5), using Properties 2.2.15:

$$\begin{aligned}
[\hat{M}, \hat{M}]_t =& [H - \Gamma_{\cdot \wedge T^s}, H - \Gamma_{\cdot \wedge T^s}]_t = [H, H]_t = \sum_{k \le t} (\Delta H_k)^2 \\
=& \sum_{k \le t} \Delta H_k = \mathbb{1}_{\{T^s \le t\}} = H_t,
\end{aligned} \tag{6.6}$$

where we used that $\Delta H_k = +1$ or 0 and that the only possible jump will occur at surrender time.
To make calculations easier in the case of n policyholders, we assume that the surrender times are conditional independent with respect to the underlying filtration of the financial market. This means that

$$E^Q[\mathbb{1}_{\{T_1^s > t_1\}} \mathbb{1}_{\{T_2^s > t_2\}} \dots \mathbb{1}_{\{T_n^s > t_n\}} | \mathcal{F}_T] = \prod_{i=1}^n E^Q[\mathbb{1}_{\{T_i^s > t_i\}} | \mathcal{F}_T].$$

The main advantage of this assumption is that the surrender time for the i^{th} policyholder only depends on $\mathcal{F}$ and $\mathcal{H}^i$. Due to this assumption of $(\mathbb{F}, Q)$-independence it is possible to prove that $\hat{M}^i \hat{M}^j$ is a $\mathbb{G}$-martingale. This result is based on Proposition 5.1.3 of Bielecki and Rutkowski (2002):

Lemma 6.2.3. *If a bounded $\mathbb{F}$-martingale m is also a $\mathbb{G}$-martingale then the product $\hat{M}m$ is a $\mathbb{G}$-martingale.*

We can extend this lemma to the case of n policyholders, this is with $\mathbb{G} = \mathbb{G}^n$. Then, the $\mathbb{H}^j$-independence of T_i^s ensures that $\hat{M}m$ is a $\mathbb{G}^n$-martingale with $\hat{M} = \hat{M}^i$ and $m = \hat{M}^j$.

6.2.3 Payment process

The discounted payment process

$$A_t = \int_0^t \mathbb{1}_{\{u<T^s\}} B_u^{-1} g_u^c du + \int_0^t B_u^{-1} g_u^s dH_u^s + \mathbb{1}_{\{T<T^s\}} B_T^{-1} g_T^m \mathbb{1}_{\{t=T\}}$$
$$- \sum_{i=1}^{I} \mathbb{1}_{\{t_i<T^s\}} B_{t_i}^{-1} p_{t_i} \mathbb{1}_{\{t_i \le t\}}$$

consists of four parts, namely three benefits and the incoming payments for the insurance company:

- the payment up to surrender or until maturity at time T

$$A_t^c = \int_0^t \mathbb{1}_{\{u<T^s\}} B_u^{-1} g_u^c du, \tag{6.7}$$

- the payment at surrender

$$A_t^s = \int_0^t B_u^{-1} g_u^s dH_u^s, \tag{6.8}$$

- the payment at maturity

$$A_t^m = \mathbb{1}_{\{T<T^s\}} B_T^{-1} g_T^m \mathbb{1}_{\{t=T\}}, \tag{6.9}$$

- the premiums at fixed times t_i, $i = 1, \ldots, I$ with $0 \le t_i < T$

$$P_t = - \sum_{i=1}^{I} \mathbb{1}_{\{t_i<T^s\}} B_{t_i}^{-1} p_{t_i} \mathbb{1}_{\{t_i \le t\}}. \tag{6.10}$$

It is assumed that the functions $g^c(u, S_u)$, $g^s(u, S_u)$, $g^m(T, S_T)$ and $p(t_i, S_{t_i})$ are $\mathbb{F}$-adapted functions and that

$$\sup_{u \in [0,T]} E^Q[(B_u^{-1} g^c(u, S_u))^2] < \infty, \qquad \sup_{u \in [0,T]} E^Q[(B_u^{-1} g^s(u, S_u))^2] < \infty.$$

6.3 The portfolio in case the underlying risky asset follows a Brownian Motion

We work this out for the case of a Brownian motion, but remark that the Brownian motion can be replaced by any other continuous martingale. In Section 6.4 we will redo the calculations in case the underlying asset is driven by a Lévy process. In both sections we calculate first the portfolio for one policyholder and then we extend the obtained results to a portfolio of policyholders for which we also calculate the risk process. In the continuous case we will show that the relative risk process goes to zero if the number of policyholders increases, while in the discontinuous case the risk is no longer diversifiable.

6.3.1 The portfolio for one policyholder

6.3.1.1 The payment up to surrender

We derive the risk-minimizing hedging strategy for the payment up to surrender A_t^c given by (6.7). For the value of the discounted portfolio at time t holds that

$$\begin{aligned} V_t^c = A_t^c + E^Q[A_T^c - A_t^c|\mathcal{G}_t] &= A_t^c + E^Q[\int_t^T \mathbb{1}_{\{T^s>u\}} B_u^{-1} g_u^c du|\mathcal{G}_t] \\ &= A_t^c + B_t^{-1}\mathbb{1}_{\{T^s>t\}} e^{\Gamma_t} \int_t^T F^c(t,S_t,u)du \end{aligned} \tag{6.11}$$

where we used (6.7), (6.4) and the notation $F^c(t,S_t,u)$ for $E^Q[B_tB_u^{-1}e^{-\Gamma_u}g_u^c|\mathcal{F}_t]$.
We assume that the function $(t,s,u)\mapsto F^c(t,s,u)$ is continuously differentiable with respect to t and twice continuously differentiable with respect to s. Furthermore, we assume that the first order partial derivative with respect to s is uniformly bounded. Using Itô's formula, see Theorem 2.2.26, and the fact that $B_t^{-1}F^c(t,S_t,u)$ is an $(\mathbb{F},Q)$-martingale in the complete financial market, we find as dynamics of $B_t^{-1}F^c(t,S_t,u)$ for $t\leq u$:

$$d(B_t^{-1}F^c(t,S_t,u)) = F_s^c(t,S_t,u)dX_t. \tag{6.12}$$

Define

$$W^c(t,S_t) := e^{\Gamma_t} \int_t^T F^c(t,S_t,u)du. \tag{6.13}$$

Then inserting this in (6.11) provides

$$V_t^c = A_t^c + B_t^{-1} \mathbb{1}_{\{T^s>t\}} W^c(t,S_t). \tag{6.14}$$

Our goal is to obtain a decomposition of the form

$$B_t^{-1} W^c(t,S_t) = W^c(0,S_0) + \int_0^t \ldots dX_u + \int_0^t \ldots du.$$

Therefore we introduce

$$Y_t^u := B_t^{-1} e^{\Gamma_t} F^c(t,S_t,u), \tag{6.15}$$

which hence satisfies

$$\int_t^T Y_t^u du = \int_t^T B_t^{-1} e^{\Gamma_t} F^c(t,S_t,u)du = B_t^{-1} W^c(t,S_t). \tag{6.16}$$

Due to the fact that all the processes are continuous in the Brownian motion case and that also the hazard process is assumed to be continuous, we find:

$$\begin{aligned} dY_t^u =& d(B_t^{-1} F^c(t,S_t,u) e^{\Gamma_t}) \\ =& B_t^{-1} F^c(t,S_t,u) e^{\Gamma_t} d\Gamma_t + e^{\Gamma_t} d(B_t^{-1} F^c(t,S_t,u)) + d[e^{\Gamma}, B^{-1} F^c(\cdot,S,u)]_t. \end{aligned} \tag{6.17}$$

The third term is equal to zero because e^{Γ} is a continuous, increasing process of finite variation, see Properties 2.2.15(3). Writing Γ in terms of the intensity function λ^s and using (6.12) we arrive at

$$\begin{aligned} dY_t^u &= B_t^{-1} F^c(t,S_t,u) e^{\Gamma_t} \lambda_t^s dt + e^{\Gamma_t} F_s^c(t,S_t,u) dX_t \\ &=: \alpha_t^u dt + \beta_t^u dX_t. \end{aligned} \tag{6.18}$$

Hence in integral form we find

$$Y_t^u = Y_0^u + \int_0^t \alpha_\tau^u d\tau + \int_0^t \beta_\tau^u dX_\tau.$$

Using this information we rewrite $B_t^{-1}W^c(t,S_t)$ (6.16) in the following way

$$\begin{aligned}
B_t^{-1}W^c(t,S_t) &= \int_t^T Y_t^u du = \int_t^T \left[Y_0^u + \int_0^t \alpha_\tau^u d\tau + \int_0^t \beta_\tau^u dX_\tau\right] du \\
&= \int_0^T \left[Y_0^u + \int_0^t \mathbb{1}_{\{\tau \le u\}}\alpha_\tau^u d\tau + \int_0^t \mathbb{1}_{\{\tau \le u\}}\beta_\tau^u dX_\tau\right] du \\
&\quad - \int_0^t \left[Y_0^u + \int_0^u \alpha_\tau^u d\tau + \int_0^u \beta_\tau^u dX_\tau\right] du \\
&= \int_0^T Y_0^u du - \int_0^t Y_u^u du + \int_0^T \int_0^t \mathbb{1}_{\{\tau \le u\}}\alpha_\tau^u d\tau du \\
&\quad + \int_0^T \int_0^t \mathbb{1}_{\{\tau \le u\}}\beta_\tau^u dX_\tau du. \qquad (6.19)
\end{aligned}$$

By relation (6.16) and by definition of Y_t^u (6.15) and $F^c(t,S_t,u)$, we know that

$$\begin{aligned}
&\int_0^T Y_0^u du = B_0^{-1}W^c(0,S_0) = W^c(0,S_0) \\
&\int_0^t Y_u^u du = \int_0^t B_u^{-1}F^c(u,S_u,u)e^{\Gamma_u}du = \int_0^t B_u^{-1}E^Q[B_u B_u^{-1}e^{-\Gamma_u}g_u^c|\mathcal{F}_u]e^{\Gamma_u}du \\
&\qquad = \int_0^t B_u^{-1}g_u^c du.
\end{aligned}$$

The standard Fubini theorem allows us to rewrite the third term in (6.19) as:

$$\begin{aligned}
\int_0^T \int_0^t \mathbb{1}_{\{\tau \le u\}}\alpha_\tau^u d\tau du &= \int_0^t \int_\tau^T \alpha_\tau^u du d\tau = \int_0^t \int_\tau^T B_\tau^{-1}F^c(\tau,S_\tau,u)e^{\Gamma_\tau}\lambda_\tau^s du d\tau \\
&= \int_0^t B_\tau^{-1}\lambda_\tau^s W^c(\tau,S_\tau)d\tau,
\end{aligned}$$

because the function $(\omega,t,u) \mapsto \alpha_t^u(\omega)$ introduced in (6.18) is $\mathcal{O}\otimes\mathcal{B}([0,T])$-measurable and $\int_0^T\int_0^t \mathbb{1}_{\{T^s \le u\}}|\alpha_\tau^u|d\tau du < \infty$ Q-almost surely, and where we used (6.13).

By the Fubini theorem for stochastic integrals, we rewrite the fourth term in (6.19) as:

$$\int_0^T \int_0^t \mathbb{1}_{\{\tau \le u\}}\beta_\tau^u dX_\tau du = \int_0^t \int_\tau^T \beta_\tau^u du dX_\tau := \int_0^t \chi_\tau dX_\tau \qquad (6.20)$$

because the function $(\omega, t, u) \mapsto \beta_t^u(\omega)$ is $\mathcal{P}\otimes\mathcal{B}([0,T])$-measurable and uniformly bounded due to the assumed uniform boundedness of F_s^c. So equation (6.19) has the desired decomposition:

$$\begin{aligned} &B_t^{-1}W^c(t,S_t) \\ &= W^c(0,S_0) - \int_0^t B_u^{-1}g_u^c du + \int_0^t B_u^{-1}W^c(u,S_u)\lambda_u^s du + \int_0^t \chi_u dX_u \end{aligned} \tag{6.21}$$

or, equivalently,

$$d(B_t^{-1}W^c(t,S_t)) = -B_t^{-1}g_t^c dt + \chi_t dX_t + B_t^{-1}W^c(t,S_t)d\Gamma_t. \tag{6.22}$$

Now it is very easy to calculate $dV^c(t,S_t)$ from (6.14):

$$\begin{aligned} dV^c(t,S_t) =& dA_t^c + \mathbb{1}_{\{T^s>t-\}}d(B_t^{-1}W^c(t,S_t)) + B_t^{-1}W^c(t,S_t)d\mathbb{1}_{\{T^s>t\}} \\ &+ d[\mathbb{1}_{\{T^s>\cdot\}}, B^{-1}W^c(\cdot,S)]_t. \end{aligned}$$

The fourth term is equal to zero by Properties 2.2.15(3), the continuity of $B_t^{-1}W^c(t,S_t)$ and the fact that $\mathbb{1}_{\{T^s>t\}}$ has finite variation. Using (6.5), (6.7) and (6.22) we find

$$\begin{aligned} dV^c(t,S_t) =& \mathbb{1}_{\{T^s>t\}}B_t^{-1}g_t^c dt + B_t^{-1}W^c(t,S_t)d(1-\mathbb{1}_{\{T^s\le t\}}) \\ &+ \mathbb{1}_{\{T^s>t-\}}(-B_t^{-1}g_t^c dt + \chi_t dX_t + B_t^{-1}W^c(t,S_t)d\Gamma_t) \\ =& \mathbb{1}_{\{T^s>t-\}}\chi_t dX_t + B_t^{-1}W^c(t,S_t)d\Gamma_{t\wedge T^s} - B_t^{-1}W^c(t,S_t)dH_t. \end{aligned} \tag{6.23}$$

To explain why dA_t^c cancels out with $-\mathbb{1}_{\{T^s>t-\}}B_t^{-1}g_t^c dt$, we look at the integral form of both:

$$\begin{aligned} \int_0^t dA_u^c - \int_0^t \mathbb{1}_{\{T^s>u-\}}B_u^{-1}g_u^c du &= \int_0^t \mathbb{1}_{\{T^s>u\}}B_u^{-1}g_u^c du - \int_0^t \mathbb{1}_{\{T^s>u-\}}B_u^{-1}g_u^c du \\ &= \int_0^t \mathbb{1}_{\{T^s>u\}}B_u^{-1}g_u^c du - \int_0^t \mathbb{1}_{\{T^s\ge u\}}B_u^{-1}g_u^c du \\ &= \int_0^t \mathbb{1}_{\{T^s=u\}}B_u^{-1}g_u^c du = 0 \end{aligned} \tag{6.24}$$

because it is an integral with measure zero.
Our aim is to find the Galtchouk-Kunita-Watanabe decomposition of the $(\mathbb{G}, Q)$-martingale $V^c(t,S_t)$. Therefore we need the $(\mathbb{G}, Q)$-compensator of the process H determined in (6.5). We rewrite (6.23), using the martingale $\hat{M}$:

$$dV^c(t,S_t) = \mathbb{1}_{\{T^s>t-\}}\chi_t dX_t - B_t^{-1}W^c(t,S_t)d\hat{M}_t. \tag{6.25}$$

Recalling definition (6.20) of χ and integrating (6.25) leads to the required decomposition of $V^c(t, S_t)$:

$$\begin{aligned} V^c(t, S_t) =& V^c(0, S_0) + \int_0^t \mathbb{1}_{\{T^s > \tau -\}} e^{\Gamma_\tau} \int_\tau^T F_s^c(\tau, S_\tau, u) du dX_\tau \\ & - \int_0^t B_\tau^{-1} W^c(\tau, S_\tau) d\hat{M}_\tau, \end{aligned} \tag{6.26}$$

since $\hat{M}$ is a martingale for which $[X, \hat{M}] = 0$ by the continuity of X and the fact that $\hat{M}$ has finite variation, see Properties 2.2.15(3). So the risk-minimizing strategy at time t invests in $\mathbb{1}_{\{T^s > t\}} e^{\Gamma_t} \int_t^T F_s^c(t, S_t, u) du$ risky assets and in view of (6.3) and (6.14) an amount

$$\mathbb{1}_{\{T^s > t\}} B_t^{-1} W^c(t, S_t) - X_t \mathbb{1}_{\{T^s > t\}} e^{\Gamma_t} \int_t^T F_s^c(t, S_t, u) du$$

in the riskless asset. We remark that $\int_t^T F_s^c(t, S_t, u) du$ is exactly the number we have to invest in the risky assets if there is no surrender option. Hence this is the amount coming from the purely financial risk, thus the life insurance without the surrender option.

6.3.1.2 The payment at surrender

We perform an analogous calculation as in Section 6.3.1.1 to find the optimal hedging strategy for the payment at surrender. Hereto, we need the following theorem, for which a proof can be found in Theorem III.19 of Protter (2005).

Theorem 6.3.1. *Let X be an increasing process of integrable variation, and let Y be an adapted process with càglàd paths such that $E^Q[\int_0^t Y_s dX_s] < \infty$. Then*

$$E^Q\left[\int_0^t Y_s dX_s\right] = E^Q\left[\int_0^t Y_s d\tilde{X}_s\right],$$

with $\tilde{X}$ the compensator of X.

Using successively the process (6.8) for the payment at surrender, Theorem 6.3.1 combined with (6.5), the continuity of all the processes except the indicator function in the integral, result (6.4), the absolute continuity of the process

Γ and the notation $F^s(t, S_t, u) = E^Q[B_t B_u^{-1} e^{-\Gamma_u} \lambda_u^s g_u^s | \mathcal{F}_t]$, the discounted portfolio at time t is given by:

$$\begin{aligned}
V_t^s &= A_t^s + E^Q[A_T^s - A_t^s | \mathcal{G}_t] = A_t^s + E^Q \left[\int_t^T B_u^{-1} g_u^s dH_u | \mathcal{G}_t \right] \\
&= A_t^s + B_t^{-1} E^Q \left[\int_t^T B_t B_u^{-1} g_u^s \mathbb{1}_{\{T^s > u\}} d\Gamma_u | \mathcal{G}_t \right] \\
&= A_t^s + B_t^{-1} \mathbb{1}_{\{T^s > t\}} e^{\Gamma_t} \int_t^T E^Q[B_t B_u^{-1} g_u^s e^{-\Gamma_u} \lambda_u^s | \mathcal{F}_t] du \\
&= A_t^s + B_t^{-1} \mathbb{1}_{\{T^s > t\}} e^{\Gamma_t} \int_t^T F^s(t, S_t, u) du.
\end{aligned}$$

Note that thanks to the assumption of absolute continuity of the hazard process Γ it was possible to interchange the expectation sign E^Q and the integral sign in the one but last step; and hence to obtain an expression for which we can easily find the Galtchouk-Kunita-Watanabe decomposition. The further calculations are analogous to those for the payment up to surrender, if we make the same assumptions but with respectively $F^c(t, S_t, u) = E^Q[B_t B_u^{-1} e^{-\Gamma_u} g_u^c | \mathcal{F}_t]$ and $W^c(t, S_t)$ (6.13) replaced by

$$F^s(t, S_t, u) = E^Q[B_t B_u^{-1} e^{-\Gamma_u} \lambda_u^s g_u^s | \mathcal{F}_t]$$

and

$$W^s(t, S_t) = e^{\Gamma_t} \int_t^T F^s(t, S_t, u) du.$$

An important difference is that dA_t^s no longer cancels out with the term $-\mathbb{1}_{\{T^s > t-\}} B_t^{-1} g_t^s \lambda_t^s dt$, but that these two terms joined together lead to an extra risk $B_t^{-1} g_t^s d\hat{M}_t$. Hence the dynamics of V^s are given by

$$dV^s(t, S_t) = B_t^{-1}(g_t^s - W^s(t, S_t)) d\hat{M}_t + \mathbb{1}_{\{T^s > t-\}} \int_t^T e^{\Gamma_t} F_s^s(t, S_t, u) dX_t.$$

So the risk-minimizing strategy invests in $\mathbb{1}_{\{T^s > t-\}} e^{\Gamma_t} \int_t^T F_s^s(t, S_t, u) du$ risky assets and an amount $\mathbb{1}_{\{T^s > t\}} B_t^{-1} W^s(t, S_t) - X_t \mathbb{1}_{\{T^s > t-\}} e^{\Gamma_t} \int_t^T F_s^s(t, S_t, u) du$ in the riskless asset.

6.3.1.3 The payment at maturity and the premiums

We determine the strategy for the payment at maturity together with the strategy for the premiums, because it is easily seen that the portfolios can be calculated in the same way by putting $p(T, S_T) = g_T^m$. The value of the discounted portfolio at time t associated with a payment $p(t_i, S_{t_i})$ at time t_i is

$$V_t^{p_{t_i}} = E^Q[B_{t_i}^{-1} \mathbb{1}_{\{t<t_i\}} \mathbb{1}_{\{T^s>t_i\}} p_{t_i} | \mathcal{F}_t] = \mathbb{1}_{\{t<t_i\}} B_t^{-1} \mathbb{1}_{\{T^s>t\}} e^{\Gamma_t} F^{p_{t_i}}(t, S_t, t_i),$$

where we invoked relation (6.4) and denoted $F^{p_{t_i}}(t, S_t, t_i) = E^Q[B_t B_{t_i}^{-1} e^{-\Gamma_{t_i}} p_{t_i} | \mathcal{F}_t]$. Assuming that the function $F^{p_{t_i}}$ has the same properties as the function F^c, we can adapt the differential of $B_t^{-1} F^c(t, S_t, t_i)$ in (6.12) to find the differential of $B_t^{-1} F^{p_{t_i}}(t, S_t, t_i)$. Now it is very easy to calculate $dV^{p_{t_i}}$ by the product rule:

$$\begin{aligned} &dV_t^{p_{t_i}} \\ &= \mathbb{1}_{\{t<t_i\}} (\mathbb{1}_{\{T^s>t-\}} e^{\Gamma_t} d(B_t^{-1} F^{p_{t_i}}(t, S_t, t_i)) + B_t^{-1} e^{\Gamma_t} F^{p_{t_i}}(t, S_t, t_i) d\mathbb{1}_{\{T^s>t\}}) \\ &\quad + \mathbb{1}_{\{t<t_i\}} (B_t^{-1} \mathbb{1}_{\{T^s>t-\}} e^{\Gamma_t} F^{p_{t_i}}(t, S_t, t_i) d\Gamma_t) \\ &= \mathbb{1}_{\{t<t_i\}} (\mathbb{1}_{\{T^s>t-\}} e^{\Gamma_t} F_s^{p_{t_i}}(t, S_t, t_i) dX_t - B_t^{-1} e^{\Gamma_t} F^{p_{t_i}}(t, S_t, t_i) d\hat{M}_t), \end{aligned} \quad (6.27)$$

where the mixed terms are all zero because Γ is an increasing, continuous process, $1_{\{T^s>t\}}$ has finite variation and $B_t^{-1} F^{p_{t_i}}(t, S_t, t_i)$ is a continuous martingale.

This means that for every premium $p(t_i, S_{t_i})$ with $i \in 1, \ldots, I$, we may reduce the hedging strategy at time t with $\mathbb{1}_{\{t<t_i\}} \mathbb{1}_{\{T^s>t-\}} e^{\Gamma_t} F_s^{p_{t_i}}(t, S_t, t_i)$ risky assets and the amount of the riskless asset with $\mathbb{1}_{\{t<t_i\}} \mathbb{1}_{\{T^s>t\}} B_t^{-1} e^{\Gamma_t} F^{p_{t_i}}(t, S_t, t_i)$ $-\mathbb{1}_{\{t<t_i\}} X_t \mathbb{1}_{\{T^s>t-\}} e^{\Gamma_t} F_s^{p_{t_i}}(t, S_t, t_i)$.

For the payment at maturity we set $p(T, S_T) = g_T^m$ and increase the strategy with the obtained amounts (see also (6.28) and (6.29) in Section 6.3.2).

6.3.2 The portfolio for n policyholders

In this section we combine the results from Sections 6.3.1.1-6.3.1.3 to find the total risk-minimizing portfolio at time t of a unit-linked life insurance contract with a surrender option, a continuous payment until maturity or until surrender,

whatever comes first, a payment at maturity and I premiums for n policyholders.
We assume that the n hazard processes are homogeneous and use the linearity of the Galtchouk-Kunita-Watanabe decomposition. Denoting $n_t = \sum_{i=1}^n \mathbb{1}_{\{T_i^s \leq t\}}$, the total portfolio for n policyholders at time t contains

$$\begin{aligned}\hat{\psi}_t =&(n-n_{t-})e^{\Gamma_t}[\int_t^T (F_s^c(t,S_t,u)+F_s^s(t,S_t,u))du + F_s^m(t,S_t,T)] \\ &-(n-n_{t-})e^{\Gamma_t}\sum_{i=1}^I \mathbb{1}_{\{t<t_i\}}F_s^{p_{t_i}}(t,S_t,t_i) \end{aligned} \tag{6.28}$$

risky assets while the riskless asset amounts to

$$\begin{aligned}&(n-n_t)B_t^{-1}e^{\Gamma_t}[\int_t^T (F^c(t,S_t,u)+F^s(t,S_t,u))du + F^m(t,S_t,T)] \\ &-(n-n_t)B_t^{-1}e^{\Gamma_t}\sum_{i=1}^I \mathbb{1}_{\{t<t_i\}}F^{p_{t_i}}(t,S_t,t_i) - \hat{\psi}_t X_t, \end{aligned} \tag{6.29}$$

with

$$\begin{aligned} F^c(t,S_t,u) &= E^Q[B_tB_u^{-1}e^{-\Gamma_u}g_u^c|\mathcal{F}_t], \quad F^s(t,S_t,u) = E^Q[B_tB_u^{-1}e^{-\Gamma_u}\lambda_u^s g_u^s|\mathcal{F}_t], \\ F^m(t,S_t,T) &= E^Q[B_tB_T^{-1}e^{-\Gamma_T}g_T^m|\mathcal{F}_t], \quad F^{p_{t_i}}(t,S_t,t_i) = E^Q[B_tB_{t_i}^{-1}e^{-\Gamma_{t_i}}p_{t_i}|\mathcal{F}_t]. \end{aligned}$$

6.3.3 The risk process

From Chapter 4, we know that when the Galtchouk-Kunita-Watanabe decomposition of the $(\mathbb{F},Q)$-martingale $V_t^* = E^Q[A_T|\mathcal{F}_t]$ is given by $V_0^* + \int_0^t \xi_u^A dX_u$ $+L_t^Q$, the total risk process at time 0 equals $E^Q[(L_T^Q)^2]$. We will first calculate the risk process induced by the payment up to surrender. Once we have found the structure of one risk process, it is easy to determine the total risk process.

6.3.3.1 The payment up to surrender

Using formula (6.26), the assumed homogeneity of the hazard processes and the linearity of the Galtchouk-Kunita-Watanabe decomposition we find that L_T^Q

for n policyholders is equal to

$$-\sum_{i=1}^{n}\int_0^T B_u^{-1}W^c(u,S_u)d\hat{M}_u^i.$$

The risk process at time zero for the payment up to surrender for n policyholders is then given by

$$R^c = E^Q\left[\left(\sum_{i=1}^{n}\int_0^T B_t^{-1}W^c(t,S_t)d\hat{M}_t^i\right)^2\right],$$

or equivalently, due to the orthogonality between the $\hat{M}^i$'s:

$$R^c = \sum_{i=1}^{n} E^Q\left[\left(\int_0^T B_t^{-1}W^c(t,S_t)d\hat{M}_t^i\right)^2\right].$$

The compensator of $[\hat{M}^i,\hat{M}^i]_t$ equals $\Gamma^i_{t\wedge T_i^s}$ since by relation (6.6) $[\hat{M}^i,\hat{M}^i]_t = H_t^i$. So $d[\hat{M}^i,\hat{M}^i]_t = \mathbb{1}_{\{T_i^s>t-\}}d\Gamma_t^i = \mathbb{1}_{\{T_i^s>t-\}}\lambda_t^s dt$. In this way the risk process can be transformed further into:

$$R^c = \int_0^T (n-n_{t-})E^Q[B_t^{-2}(W^c(t,S_t))^2]\lambda_t^s dt. \tag{6.30}$$

Looking at $\frac{\sqrt{R^c}}{n}$, we may conclude that the risk originating from the payment up to surrender is completely diversifiable. Analogously we can calculate the risk processes at time 0 for the surrender option, the payment at maturity and the premiums which all turn out to be completely diversifiable.

6.3.3.2 The total risk process

When we calculate the total risk process, we have to use the term L_T^Q in the Galtchouk-Kunita-Watanabe decomposition of the discounted total portfolio and we cannot simply add up all the risk processes. This L_T^Q is now given by

$$L_T^Q = -\sum_{i=1}^{n} \int_0^T B_u^{-1} \Bigg(W^c(u, S_u) + W^s(u, S_u) - g_u^s + e^{\Gamma_u} F^m(u, S_u, T) - \sum_{i=1}^{I} \mathbb{1}_{\{u<t_i\}} e^{\Gamma_u} F^{p_{t_i}}(u, S_u, t_i) \Bigg) d\hat{M}_u^i$$

and performing an analogous calculation as in Section 6.3.3.1 we find

$$R^{\text{Tot}} = \int_0^T (n - n_{t-}) E^Q[B_t^{-2}(W^c(t, S_t) + W^s(t, S_t) - g_t^s + e^{\Gamma_t} F^m(t, S_t, T) - \sum_{i=1}^{I} \mathbb{1}_{\{t<t_i\}} e^{\Gamma_t} F^{p_{t_i}}(t, S_t, t_i))^2] \lambda_t^s dt.$$

So the relative risk process $\frac{\sqrt{R^{\text{Tot}}}}{n}$ goes to zero if n increases and hence the total risk is completely diversifiable.

6.4 The portfolio in case the underlying risky asset follows a Lévy process

In this section we assume that the risky asset is driven by a geometric Lévy process, as described in (6.1).

6.4.1 The portfolio for one policyholder

In the article of Riesner (2006a), $d(B_t^{-1} F(t, S_t, u))$ with $F(t, S_t, u) = E^Q[B_t B_u^{-1} g(u, S_u)|\mathcal{F}_t]$ is calculated. We can adapt these results to derive that, for example in the case of the continuous payment where

$$F^c(t, S_t, u) = E^Q[B_t B_u^{-1} e^{-\Gamma_u} g_u^c | \mathcal{F}_t],$$

the Galtchouk-Kunita-Watanabe decomposition of $B_t^{-1}F^c(t,S_t,u)$ for $0 \le t < u \le T$ is given by

$$B_t^{-1}F^c(t,S_t,u) = F(0,S_0,u) + \int_0^t \zeta^c(s,u)dX_s + K^c(t,u) \tag{6.31}$$

with $v = \int_{\mathbb{R}} x^2\nu(dx)$, $\kappa = c^2 + v$ and

$$\begin{cases} \zeta^c(t,u) &=& \frac{c^2}{\kappa}F_s^c(t,S_{t-},u) + \frac{1}{\sigma_t X_{t-}\kappa}\int_{\mathbb{R}} xJ^c(t,x,u)\nu(dx), \\ K^c(t,u) &=& \int_0^t \theta^c(s,u)dW_s + \int_0^t\int_{\mathbb{R}} \varkappa^c(s,y,u)M(ds,dy) \\ \theta^c(t,u) &=& c\sigma_t X_{t-}(F_s^c(t,S_{t-},u) - \zeta^c(t,u)), \\ \varkappa^c(t,y,u) &=& J^c(t,y,u) - y\sigma_t X_{t-}\zeta^c(t,u), \\ J^c(t,x,u) &=& B_t^{-1}[F^c(t,S_{t-}+\sigma_t S_{t-}x,u) - F^c(t,S_{t-},u)]. \end{cases} \tag{6.32}$$

Note: We do not completely agree with the calculations of Riesner, but we can adapt his result to use it here. He considers a risky asset which is not a martingale under the original measure. Hence he performs a change of measure to the minimal martingale measure. Under this new measure he determines the Galtchouk-Kunita-Watanabe decomposition and says that this is also the Föllmer-Schweizer decomposition under the original measure. This would be true if the underlying risky asset is continuous. For more details we refer to Chapter 7.

We adapt the calculations for the Brownian motion to this case. An important difference is that now $F^c(t,S_{t-},u)$ is different from $F^c(t,S_t,u)$ because of the possible jumps in the risky asset. So in this case, we have that equation (6.17) is equal to

$$\begin{aligned} dY_t^u =& B_t^{-1}F^c(t,S_{t-},u)e^{\Gamma_t}\lambda_t^s dt + e^{\Gamma_t}(\zeta^c(t,u)dX_t + dK^c(t,u)) \\ &+ d[e^{\Gamma}, B^{-1}F^c(\cdot,S,u)]_t \\ =& B_t^{-1}F^c(t,S_{t-},u)e^{\Gamma_t}\lambda_t^s dt + e^{\Gamma_t}\zeta^c(t,u)dX_t \\ &+ e^{\Gamma_t}(\theta^c(t,u)dW_t + \int_{\mathbb{R}} \varkappa^c(t,y,u)M(dt,dy)), \end{aligned} \tag{6.33}$$

where $[e^{\Gamma}, B^{-1}F^c(\cdot,S,u)] = 0$ because e^{Γ_t} is continuous and has finite variation. We know that $\zeta^c(t,u)$, $\theta^c(t,u)$ and $\varkappa^c(t,y,u)$ will be bounded Q-a.s. because the first order partial derivative of F^c with respect to the risky asset is assumed to be bounded, X is square integrable and $\sup_{t\in[0,T]}\int_{\mathbb{R}} x^2\nu_t(dx) < \infty$. So we can conclude that the above integrals are well defined. Adapting the calculations in (6.19)-(6.21) to the Lévy case and using (6.33), we find:

$$\begin{aligned}
&B_t^{-1}W^c(t,S_t)\\
&= W^c(0,S_0) - \int_0^t B_u^{-1}g_u^c du + \int_0^t\int_\tau^T B_\tau^{-1}F^c(\tau,S_{\tau-},u)e^{\Gamma_\tau}\lambda_\tau^s du d\tau\\
&\quad + \int_0^t\int_\tau^T e^{\Gamma_\tau}\zeta^c(\tau,u)du dX_\tau + \int_0^t\int_\tau^T e^{\Gamma_\tau}\theta^c(\tau,u)du dW_\tau\\
&\quad + \int_0^t\int_{\mathbb{R}}\int_\tau^T e^{\Gamma_\tau}\varkappa^c(\tau,y,u)du M(d\tau,dy).
\end{aligned} \tag{6.34}$$

Hereto, we also applied the standard Fubini theorem and the Fubini theorem for stochastic integrals. This is possible because the functions

$$(\omega,t,u)\mapsto B_t^{-1}F^c(t,S_{t-}(\omega),u)e^{\Gamma_t}\lambda_t^s$$

are $\mathcal{O}\otimes\mathcal{B}([0,T])$-measurable, the functions

$$(\omega,t,u)\mapsto e^{\Gamma_t}\zeta^c(t,u,\omega) \qquad \text{and} \qquad (\omega,t,u)\mapsto e^{\Gamma_t}\theta^c(t,u,\omega)$$

are $\mathcal{P}\otimes\mathcal{B}([0,T])$-measurable and the functions $(\omega,t,y,u)\mapsto e^{\Gamma_t}\varkappa^c(t,y,u,\omega)$ are $\mathcal{P}\otimes\mathcal{B}([0,T])\otimes\mathcal{B}(\mathbb{R})$-measurable for all $u\in[0,T]$. Furthermore all these functions are also uniformly bounded.

The dynamics of the discounted portfolio for the continuous payment process are in the Lévy setting:

$$\begin{aligned}
dV^c(t,S_t) =& dA_t^c + \mathbb{1}_{\{T^s>t-\}}d(B_t^{-1}W^c(t,S_t)) - B_t^{-1}W^c(t,S_{t-})dH_t\\
&- d[H,B^{-1}W^c(\cdot,S)]_t.
\end{aligned}$$

Due to Theorem 2.2.14 and the finite variation of H, we can rewrite the fourth term in the following way:

$$\begin{aligned}
[H,B_\cdot^{-1}W^c(\cdot,S_\cdot)]_t =& \sum_{s\le t}\Delta H_s\Delta(B_s^{-1}W^c(s,S_s))\\
=& \begin{cases}\Delta(B_{T^s}^{-1}W^c(T^s,S_{T^s})) & \text{if } T^s\le t\\ 0 & \text{if } T^s>t.\end{cases}
\end{aligned}$$

From Protter (2005) we know, as a consequence of Meyers's theorem, that Lévy processes only jump at totally inaccessible stopping times. Combining this with condition (C), see page 118, we know that the Lévy process cannot jump at the surrender time and hence $[H, B_{\cdot}^{-1}W^c(\cdot, S_{\cdot})] \equiv 0$.
Carrying out analogous calculations as in Section 6.3.1.1, we find using (6.7), (6.34), and a reasoning as in (6.24) that the dynamics of the discounted portfolio equal

$$\begin{aligned} dV^c(t,S_t) =& dA_t^c + \mathbb{1}_{\{T^s>t-\}} d(B_t^{-1}W^c(t,S_t)) - B_t^{-1}W^c(t,S_{t-})dH_t^s \\ =& \mathbb{1}_{\{T^s>t-\}} \int_t^T e^{\Gamma_t}\zeta^c(t,u)du dX_t - B_t^{-1}W^c(t,S_{t-})d\hat{M}_t \\ &+ \mathbb{1}_{\{T^s>t-\}} \int_t^T e^{\Gamma_t}\theta^c(t,u)du dW_t \\ &+ \mathbb{1}_{\{T^s>t-\}} \int_{\mathbb{R}}\int_t^T e^{\Gamma_t}\varkappa^c(t,y,u)du M(dt,dy). \end{aligned}$$

We remark again that all integrands are well-defined because $\mathbb{1}_{\{T^s>t-\}}e^{\Gamma_t}$ is bounded. The term L_t^Q in (6.2) is here equal to

$$\begin{aligned} L_t^Q =& -\int_0^t B_\tau^{-1}W^c(\tau,S_{\tau-})d\hat{M}_\tau + \int_0^t \mathbb{1}_{\{T^s>\tau-\}}\int_\tau^T e^{\Gamma_\tau}\theta^c(\tau,u)du dW_\tau \\ &+ \int_0^t \mathbb{1}_{\{T^s>\tau-\}}\int_{\mathbb{R}}\int_\tau^T e^{\Gamma_\tau}\varkappa^c(\tau,y,u)du M(d\tau,dy). \end{aligned} \tag{6.35}$$

The orthogonality between L^Q and the risky asset follows directly from the orthogonality between X and $K^c(u)$ and the orthogonality between X and $\hat{M}$ due to Lemma 6.2.3.
From analogous calculations we learn that for the payment at surrender the dynamics of the discounted portfolio V^s are given by

$$\begin{aligned} dV^s(t,S_t) =& \mathbb{1}_{\{T^s>t-\}} \int_t^T e^{\Gamma_t}\zeta^s(t,u)du dX_t + B_t^{-1}(g_t^s - W^s(t,S_{t-}))d\hat{M}_t \\ &+ \mathbb{1}_{\{T^s>t-\}} \int_t^T e^{\Gamma_t}\theta^s(t,u)du dW_t \\ &+ \mathbb{1}_{\{T^s>t-\}} \int_{\mathbb{R}}\int_t^T e^{\Gamma_t}\varkappa^s(t,y,u)du M(dt,dy), \end{aligned}$$

with ζ^s, θ^s, $\varkappa^s$ the functions in (6.32) with F^c replaced by F^s and superscripts c replaced by s. The dynamics of the discounted portfolio for the premiums and the payment at maturity are in an analogous way found to be

$$dV_t^{p_{t_i}} = \mathbb{1}_{\{t<t_i\}}[\mathbb{1}_{\{T^s>t-\}}e^{\Gamma_t}\zeta^{p_{t_i}}(t,t_i)dX_t - B_t^{-1}e^{\Gamma_t}F^{p_{t_i}}(t,S_t,t_i)d\hat{M}_t + \mathbb{1}_{\{T^s>t-\}}e^{\Gamma_t}(\theta^{p_{t_i}}(t,t_i)dW_t + \int_{\mathbb{R}}\varkappa^{p_{t_i}}(t,y,t_i)M(dt,dy))].$$

The structure of the portfolio for the payment at maturity ($p(T,S_T) = g_T^m$) resembles very much the pure endowment portfolio of Riesner (2006a). The mortality related factor $(l_x - \hat{M}_{s-}^I)_{T-s}p_{x+s}$ is replaced here by the surrender factor $\mathbb{1}_{\{T^s>t-\}}e^{\Gamma_t}$. The second term contains the most important difference, because when modeling the surrender time we can no longer assume independence of the financial market as Riesner could when dealing with mortality.

6.4.2 The portfolio for n policyholders

As in Section 6.3 we can again determine the portfolio for n policyholders by assuming homogeneity of the hazard processes. The total portfolio for n policyholders at time t contains an investment in

$$\tilde{\phi}_t = (n-n_{t-})e^{\Gamma_t}[\int_t^T(\zeta^c(t,u)+\zeta^s(t,u))du + \zeta^m(t,T) - \sum_{i=1}^I \mathbb{1}_{\{t<t_i\}}\zeta^{p_{t_i}}(t,t_i)]$$

risky assets and an amount

$$(n-n_t)B_t^{-1}e^{\Gamma_t}[\int_t^T(F^c(t,S_t,u)+F^s(t,S_t,u))du + F^m(t,S_t,T) - \sum_{i=1}^I \mathbb{1}_{\{t<t_i\}}F^{p_{t_i}}(t,S_t,t_i)] - \tilde{\phi}_t S_t^*$$

in the riskless asset. So we easily see the difference with the Brownian motion case (6.28) where we had only the first order derivative with respect to the risky asset while here we have

$$\zeta^a(t,u) = \frac{c^2}{\kappa}F_s^a(t,S_{t-},u) + \frac{1}{\sigma_t S_{t-}^*\kappa}\int_{\mathbb{R}} xJ^a(t,x,u)\nu(dx) \quad \text{with } a\in\{c,s,m,p_{t_i}\};$$

of course this also affects the amount invested in the riskless asset. When there are no jumps ($\nu(dx) \equiv 0$) this Lévy case obviously reduces to the Brownian motion case.

6.4.3 The risk process

When we compare the risk process with the risk process in the Brownian motion case, we get two extra terms caused by the incompleteness of the financial market. One is driven by the Brownian motion while the other is driven by the jumps of the Lévy process.

6.4.3.1 The payment up to surrender

Using relation (6.35) we calculate the risk process R^c at time zero as in Section 6.3.3 for n policyholders ($n \geq 1$):

$$R^c = E^Q[[\sum_{i=1}^{n}(-\int_0^T B_\tau^{-1} W^c(\tau, S_{\tau-}) d\hat{M}_\tau^i + \int_0^T \mathbb{1}_{\{T_i^s > \tau-\}} \int_\tau^T e^{\Gamma_\tau} \theta^c(\tau, u) du dW_\tau$$
$$+ \int_0^T \mathbb{1}_{\{T_i^s > \tau-\}} \int_{\mathbb{R}} \int_\tau^T e^{\Gamma_\tau} \varkappa^c(\tau, y, u) du M(d\tau, dy))]^2].$$

First of all we know that $d\langle W, W\rangle_t = dt$ and $\langle M(dt, dy), M(dt, dy)\rangle = \nu(dy)dt$. By means of Lemma 6.2.3, we can prove that W and M are both orthogonal to $\hat{M}^i$ for each $i = 1, \ldots, n$. Also W and M are orthogonal, see Properties 2.2.15(4). Using the orthogonality between the $\hat{M}^i$'s and the fact that $[\hat{M}^i, \hat{M}^i] = \mathbb{1}_{\{T_i^s > t-\}} \lambda_t^s dt$, the risk process becomes

$$R^c = \int_0^T (n - n_{t-}) E^Q[B_t^{-2} (W^c(t, S_{t-}))^2] \lambda_t^s dt$$
$$+ \int_0^T (n - n_{t-})^2 E^Q[(\int_t^T e^{\Gamma_t} \theta^c(t, u) du)^2] dt$$
$$+ \int_0^T (n - n_{t-})^2 E^Q[(\int_{\mathbb{R}} \int_t^T e^{\Gamma_t} \varkappa^c(t, y, u) du)^2] \nu(dy) dt.$$

Compared with the risk process (6.30) in the Brownian motion case for the payment up to surrender, we see that the first term is the same, but we have two extra terms which enlarge the risk. The first one originates from the Brownian motion while the second one is coming from the jumps of the Lévy process. The most important remark is that the risk is no longer completely diversifiable due to the genuine market risk caused by the incompleteness of the financial market.
The other risk processes are calculated in a similar way.

6.4.3.2 The total risk process

As in Section 6.3.3.2, we compute the total risk process at time zero by calculating $E^Q[(L_T^Q)^2]$ with L^Q the term in the Galtchouk-Kunita-Watanabe decomposition of the total portfolio:

$$
\begin{aligned}
R^{\mathrm{Tot}} = \int_0^T & (n - n_{t-}) E^Q[B_t^{-2}(W^c(t, S_{t-}) - g_t^s + W^s(t, S_{t-}) \\
& \qquad + e^{\Gamma_t}(F^m(t, S_{t-}, T) - \sum_{k=1}^{I} F^{p_{t_i}}(t, S_t, t_i)))^2]\lambda_t^s dt \\
& + \int_0^T (n - n_{t-})^2 E^Q[e^{2\Gamma_t}(\int_t^T (\theta^c(t,u) + \theta^s(t,u))du + \theta^m(t,T) \\
& \qquad - \sum_{k=1}^{I} \theta^{p_{t_i}}(t, t_i))^2]dt \\
& + \int_0^T (n - n_{t-})^2 E^Q[(e^{2\Gamma_t} \int_{\mathbb{R}} \int_t^T (\varkappa^c(t,y,u) + \varkappa^s(t,y,u))du \\
& \qquad + \varkappa^m(t,y,T) - \sum_{i=1}^{I} \varkappa^{p_{t_i}}(t,y,t_i))^2]\nu(dy)dt.
\end{aligned}
$$

We conclude that the total risk process is no longer completely diversifiable. If we assume that the number of policyholders is big enough, then the last two terms will cause the risk.

Have no fear of perfection you'll never reach it.

Salvador Dali (1904-1989)

7 Local risk-minimization applied to unit-linked life insurance contracts

In Chapter 4 we mentioned already the article of Møller (1998). He constructed the risk-minimizing hedging portfolio for unit-linked life insurance contracts with a pure endowment and a term insurance with single premium in the complete Black-Scholes market.

Riesner (2006a) combined the theory of locally risk-minimizing hedging strategies with the results of Møller (1998) but for a risky asset of which the price process is discontinuous as it follows a geometric Lévy process. However, determining the risk-minimizing hedging strategy under the minimal martingale measure does not provide the locally risk-minimizing hedging strategy in the Lévy case, as already discussed in Chapter 3.

We will derive the correct locally risk-minimizing hedging strategy by a direct construction of the Föllmer-Schweizer decomposition. First of all we will use an approach based on the article by Colwell and Elliott (1993). We applied this approach before we determined the general form of the Föllmer-Schweizer decomposition based on the characteristics as is described in Chapter 3. Therefore we will check here the results obtained in the first case with the form given in Chapter 3. This will highlight also the usefulness of the results given in Chapter

3.
In Section 7.1 we look in more detail at the mistakes made in Riesner (2006a, 2007) and in Section 10.4 of Cont and Tankov (2004). After recalling the setting of Riesner (2006a) in Section 7.2, we then show in Section 7.3 that in Riesner (2006a) the risk-minimizing hedging strategy under the new measure was found, but that this strategy is not the locally risk-minimizing one under the original measure. In Section 7.4 we show how to determine the correct locally risk-minimizing hedging strategy under the original measure and we calculate the associated risk process. In Section 7.5 we adapt the results of Riesner (2006a) for unit-linked life insurance contracts with a pure endowment and a term insurance. We end with a practical example in which we determine the FS decomposition using the results of Chapter 3. Afterwards we apply this result to the setting of Riesner (2006a) and this obviously shows that it is much easier to use the results of Chapter 3 than to utilize Colwell and Elliott (1993).
This chapter is based on Vandaele and Vanmaele (2008b).

7.1 Literature

Section 10.4 of Cont and Tankov (2004)

Although Cont and Tankov do not restrict the dynamics of the underlying to continuous processes, they say that the measure under which the martingale property and the orthogonality are preserved, is called a minimal martingale measure. Furthermore, they give a general procedure to find the locally risk-minimizing hedging strategy. This procedure describes that one first need to determine the MMM and the dynamics of the underlying risky asset under this MMM. Then the GKW decomposition of the claim H is determined under the MMM. From this decomposition they claim that the number of risky assets easily follows, without a correction term. As is shown in Theorem 3.3.2, this is incorrect. In the example where the underlying risky asset follows a jump-diffusion model, see Cont and Tankov (2004) page 341, they calculate the Galtchouk-Kunita-Watanabe decomposition to determine the locally risk-minimizing hedging strategy.
We remark that Cont and Tankov will correct this mistake in a second edition of their book.

Riesner (2006a)

It is assumed that finding the risk-minimizing hedging strategy under the minimal martingale measure directly gives the locally risk-minimizing hedging strategy under the original measure. The proof of Lemma 8 is based on Proposition 10.5 of Cont and Tankov (2004). The wrong reasoning Riesner (2006a) applies is also obvious from the remark on page 602 in which Riesner states that:

> Föllmer and Sondermann (1986) consider only the case where the discounted stock price is a martingale under the original measure P. Schweizer (1991, 2001) showed that their approach fails if $\hat{S}$ is just a P-semimartingale and develops the theory of local risk-minimization: under the Föllmer-Schweizer measure one applies Föllmer and Sondermann (1986) and finds a risk-minimizing strategy which is locally risk-minimizing under P.

Riesner (2007)

The goal of Riesner (2007) is to determine the locally risk-minimizing hedging strategy for payment streams. Unfortunately a similar mistake as in the previous paper is made, but more covered up. Riesner extends the LRM hedging theory to payment streams by using the optimality equation as defined in Schweizer (1991). Schweizer (1991) assumed that there exists a square-integrable P-martingale N, orthogonal to the martingale part M under P, such that every contingent claim H can be decomposed in the sum of a constant, a stochastic integral over M and a stochastic integral over N. If this holds then M and N form a P-basis of $L^2(P)$. Furthermore he also demanded that there exists an equivalent probability measure such that X and N form a basis under this equivalent measure. In the continuous case these assumptions are realistic and it is rather easy to check that for the MMM the last assumption will also be satisfied, as is shown in the later paper of Föllmer and Schweizer (1991). The problem is that Riesner (2007) claims to work under this set of assumptions, without checking if it is possible that they hold in case the underlying risky asset is no longer continuous. It is true that at the time Schweizer (1991) wrote his paper it was less clear under which conditions the assumptions hold. Therefore minimal martingale measure was then still described as the measure

which left the structure of the model, apart from changing X into a martingale intact. So also the orthogonality relations should remain preserved, while in later papers, Schweizer showed that this preservation of orthogonality holds in case the underlying asset is continuous, but he no longer demands it as a basic condition for the minimal martingale measure in a more general setting.
In the conclusion at the end of the appendix in Riesner (2007) is stated explicitly how he determines the LRM hedging strategy:

> First, one removes the drift of the semimartingale X performing an equivalent change of measure to the Föllmer-Schweizer measure $\mathbb{Q}$ and second, one applies the theory of Møller (2001) to find a risk-minimizing hedging strategy under $\mathbb{Q}$. **This strategy** is then locally risk-minimizing with respect to the original measure $\mathbb{P}$.

7.2 The model of Riesner (2006a)

We assume that the contract ends at the finite time T. We work in the probability space $(\Omega, \mathcal{G}, \mathbb{G}, P)$ which is the product space of two independent probability spaces. The first one is used to model the Lévy process, the second one to describe the insurance portfolio. The probability space is assumed to satisfy the usual conditions of right continuity and completeness.

The financial market

The probability space describing the financial market is given by $(\Omega_1, \mathcal{F}, \mathbb{F}, P_1)$, where $(\mathcal{F}_t)_{0\le t\le T}$ is the natural filtration of the process L, with $(L_t)_{0\le t\le T}$ and $L_0 = 0$, which is the càdlàg version of a Lévy process. Assume also that L^*, see page 9, is locally bounded.
Furthermore the compensator of the jump process μ is assumed to satisfy:

$$\int_{\{|x|>1\}} |x|^3 \nu(dx) < \infty;$$

for more details we refer to Chan (1999). The process L can be decomposed as follows:

$$L_t = cW_t + \widetilde{M}_t + at, \quad 0 \le t \le T,$$

with $(W_t)_{0 \le t \le T}$ a Brownian motion, $c > 0$, $a = E[L_1]$ and

$$\widetilde{M}_t = \int_0^t \int_{\mathbb{R}} x\widetilde{M}(ds, dx) = \int_0^t \int_{\mathbb{R}} x(N(ds, dx) - \nu(dx)ds), \quad 0 \le t \le T,$$

a square-integrable martingale. The compensated Poisson random measure on $[0, \infty) \times \mathbb{R}\backslash\{0\}$ corresponding to the jumps of L is thus given by $\widetilde{M}(ds, dx)$. The financial market itself consists of a risky asset $S = (S_t)_{\{0 \le t \le T\}}$ and a bond $B = (B_t)_{\{0 \le t \le T\}}$ with dynamics

$$dS_t = b_t S_{t-} dt + \sigma_t S_{t-} dL_t = (b_t + a\sigma_t) S_{t-} dt + \sigma_t S_{t-} d(cW_t + \widetilde{M}_t), \quad S_0 > 0, \tag{7.1}$$

$$dB_t = r_t B_t dt, \quad B_0 = 1. \tag{7.2}$$

The drift b_t, the volatility $\sigma_t > 0$ and the risk-free interest rate r_t are assumed to be continuous, deterministic functions. To be sure that the value of the risky asset is always positive, we need the assumption $\sigma_t \Delta L_t > -1$ a.s. for all $t \in [0, T]$. The process of the discounted risky asset $X = S/B$ with martingale part M and which is assumed to be a special semimartingale is easily calculated from (7.1) and (7.2):

$$dX_t = (b_t + a\sigma_t - r_t) X_{t-} dt + \sigma_t X_{t-} d(cW_t + \widetilde{M}_t). \tag{7.3}$$

If $(b - r) \cdot t + \sigma \cdot L$ is special then X is also special due to formula (2.51) in Jacod (1979). It is obvious that X is only a martingale under the measure P_1 in the very special case where $b_t + a\sigma_t - r_t = 0$ for all $t \in [0, T]$. Chan showed, before the general formula, see page 23, for the Girsanov density describing the change of measure to the MMM was known, that the Girsanov parameter related with a change of measure to the minimal martingale measure $\widetilde{Q}$ is given by

$$G_t = \frac{r_t - b_t - a\sigma_t}{\sigma_t(c^2 + v)}, \quad \text{with} \quad v = \int_{\mathbb{R}} x^2 \nu(dx). \tag{7.4}$$

Therefore the density describing the change of measure from P_1 to the MMM is given by

$$dZ_t = Z_{t-} G_t (cdW_t + d\widetilde{M}_t).$$

In fact as shown in Chapter 2, the MMM $\widetilde{Q}$ can also be determined using $Z = \mathcal{E}(-\lambda \cdot M)$. For the sum in the denominator we introduce the notation

$$\kappa = \frac{d\langle X\rangle_t}{\sigma_t^2 X_{t-}^2 dt} = \frac{d\langle M\rangle_t}{\sigma_t^2 X_{t-}^2 dt} = c^2 + v = c^2 + \int_{\mathbb{R}} x^2 \nu(dx). \tag{7.5}$$

Now we assume that (3.25) holds for all $t \in [0, T]$, hence $Z > 0$. We made here the conditions stronger than Riesner (2006a) did in order to derive the existence of the GKW decomposition from the conditions which guarantees the existence of the FS decomposition.
Under the new measure $\widetilde{Q}$ the discounted risky asset in (7.3) is a locally square-integrable martingale:

$$dX_t = \sigma_t X_{t-} d(cW_t^{\widetilde{Q}} + \widetilde{M}_t^{\widetilde{Q}}),$$

with

$$W_t^{\widetilde{Q}} = W_t - \int_0^t cG_u du \tag{7.6}$$

a standard Brownian motion under $\widetilde{Q}$ and with

$$\widetilde{M}_t^{\widetilde{Q}} = \int_{\mathbb{R}} x\widetilde{M}^{\widetilde{Q}}(dt, dx) = \int_{\mathbb{R}} x[N(dt, dx) - (1 + G_t x)\nu(dx)dt] \tag{7.7}$$

a locally square-integrable $\widetilde{Q}$-martingale. The compensator of $N(dt, dx)$ under the measure $\widetilde{Q}$ is

$$\nu_t^{\widetilde{Q}}(dx) = (1 + G_t x)\nu(dx). \tag{7.8}$$

The insurance market

The insurance market is described on the probability space $(\Omega_2, \mathcal{H}, \mathbb{H}_t, P_2)$, where the filtration $(\mathcal{H}_t)_{\{0 \le t \le T\}}$ is the natural filtration generated by $\mathbb{1}_{\{T_i \le t\}}$ with $i = 1, \ldots, N$ and where P_2 is the risk-neutral measure of the insurance market, meaning that the insurer is assumed to be risk-neutral towards insurance risks. The number N denotes the number of individuals, all of equal age

x and with i.i.d. non-negative lifetimes $T_1, \ldots, T_N$. Their survival probabilities are given by

$$_tp_x = P_2(T_1 > t + x | T_1 > x) = \exp(-\int_0^t \mu_{x+\tau} d\tau)$$

with μ_{x+t} the hazard rate at age $x + t$.
The number of deaths until time t, $0 \le t \le T$ is denoted by $N_t^I = \sum_{i=1}^N \mathbb{1}_{\{T_i \le t\}}$. Furthermore the P_2-martingale $M^I = (M_t^I)_{0 \le t \le T}$ is defined as

$$M_t^I = N_t^I - \int_0^t \lambda_u du,$$

with $\lambda_t = (N - N_{t-}^I)\mu_{x+t}$.

The combined model

The equivalent martingale measure $\hat{Q}$ for the combined model is just the product measure of the minimal martingale measure $\widetilde{Q}$ and the risk-neutral measure P_2 of the insurance market. We then work in the space $(\Omega, \mathcal{G}, \mathbb{G}, \hat{Q}) = (\Omega_1 \times \Omega_2, \mathcal{F} \otimes \mathcal{H}, \mathbb{F} \otimes \mathbb{H}, \widetilde{Q} \times P_2)$.

7.3 The strategy proposed by Riesner (2006a)

Suppose $g(u, S_u)$ is an $\mathcal{F}_u$-measurable claim for which

$$\sup_{t \in [0,T]} E[(B_t^{-1} g(t, S_t))^2] < \infty.$$

We impose here an integrability condition under the original measure, while Riesner (2006a) assumed an analogous condition but under the MMM. We think it makes more sense to specify the condition under the original measure and to derive from the extra assumption (3.25) the existence of the GKW decomposition of $g(u, S_u)$ under $\widetilde{Q}$. According to risk-neutral valuation, the arbitrage-free price $F(t, S_t, u)$ of the claim is given by $E^{\widetilde{Q}}[B_t B_u^{-1} g(u, S_u) | \mathcal{F}_t]$ if

$0 \leq t < u \leq T$ and by $B_t B_u^{-1} g(u, S_u)$ if $0 \leq u \leq t \leq T$. It is assumed that $F(\cdot, \cdot, u) \in C^{1,2}([0,T] \times [0, \infty])$ for any u. Furthermore we assume that the first derivative of F with respect to the second variable is bounded $\widetilde{Q}$-almost surely. We denote by F_t the derivative with respect to the first variable and by F_x the derivative with respect to the second variable.

In Riesner (2006a) is stated that the Föllmer-Schweizer decomposition of $B_t^{-1} F(t, S_t, u)$, with $0 \leq t < u \leq T$ is given by

$$B_t^{-1} F(t, S_t, u) = F(0, S_0, u) + \int_0^t \xi(s, u) dX_s + K(t, u), \tag{7.9}$$

with

$$\begin{aligned}
v_t^{\widetilde{Q}} &= \int_{\mathbb{R}} x^2 \nu_t^{\widetilde{Q}}(dx), \quad \kappa_t^{\widetilde{Q}} = c^2 + v_t^{\widetilde{Q}}, \\
\xi(t, u) &= \frac{c^2}{\kappa_t^{\widetilde{Q}}} F_x(t, S_{t-}, u) + \frac{1}{\sigma_t X_{t-} \kappa_t^{\widetilde{Q}}} \int_{\mathbb{R}} x J(t, x, u) \nu_t^{\widetilde{Q}}(dx), && (7.10) \\
J(t, x, u) &= B_t^{-1} [F(t, S_{t-}(1 + \sigma_t x), u) - F(t, S_{t-}, u)], && (7.11) \\
K(t, u) &= \int_0^t \varsigma^{(1)}(s, u) dW_s^{\widetilde{Q}} + \int_0^t \int_{\mathbb{R}} \varsigma^{(2)}(s, y, u) \widetilde{M}^{\widetilde{Q}}(ds, dy), && (7.12) \\
\varsigma^{(1)}(t, u) &= c \sigma_t X_{t-} (F_x(t, S_{t-}, u) - \xi(t, u)), && (7.13) \\
\varsigma^{(2)}(t, y, u) &= J(t, y, u) - y \sigma_t X_{t-} \xi(t, u). && (7.14)
\end{aligned}$$

In the article it is not mentioned explicitly under which measure this decomposition is true. In fact under the MMM this decomposition is a special Föllmer-Schweizer decomposition, namely the Galtchouk-Kunita-Watanabe decomposition, but it is not necessarily the Föllmer-Schweizer decomposition under the original measure as Riesner claims. We can only prove that $K(\cdot, u)$ is a P-semimartingale and without extra assumptions on the claim it is not surely a P-martingale, which is a necessary condition to have a Föllmer-Schweizer decomposition under the measure P. By applying Itô's formula on $B_t^{-1} F(t, S_t, u)$ (see the proof of Theorem 7.4.1 for more details concerning this calculation) and solving (7.9) to K, we find:

$$\begin{aligned}
K(t, u) = &(F_x(t, S_{t-}, u) - \xi(t, u)) \cdot X \\
&+ \int_{\mathbb{R}} [J(t, y, u) - F_x(t, S_{t-}, u) \sigma_t X_{t-} y] y \widetilde{M}^{\widetilde{Q}}(ds, dy). && (7.15)
\end{aligned}$$

Therefore by using (7.3), (7.4) and (7.7), the predictable finite variation part of K under P equals

$$-\int_0^t G_s c^2 \sigma_s X_{s-} F_x(s, S_{s-}, u) ds - \int_0^t \int_{\mathbb{R}} G_s J(s, y, u) y \nu(dx) ds + \int_0^t G_s \sigma_s X_{s-} \kappa \xi(s, u) ds. \quad (7.16)$$

If $K(\cdot, u)$ is a P-local martingale, then the predictable finite variation part should be zero. From (7.10) and (7.8), we obtain that in case $G_t \neq 0$ $K(\cdot, u)$ is a P-local martingale if and only if for any $t \in [0, T]$ the following expression equals zero:

$$\begin{aligned}
&c^2 \sigma_t X_{t-} F_x(t, S_{t-}, u) + \int_{\mathbb{R}} J(t, y, u) y \nu(dy) - \xi(t, u) \sigma_t X_{t-} \kappa \\
&= c^2 \sigma_t X_{t-} F_x(t, S_{t-}, u) + \int_{\mathbb{R}} J(t, y, u) y \nu(dy) \\
&\quad - \left[\frac{c^2}{\kappa_t^{\widetilde{Q}}} F_x(t, S_{t-}, u) + \frac{1}{\sigma_t X_{t-} \kappa_t^{\widetilde{Q}}} \int_{\mathbb{R}} J(t, y, u) y \nu_t^{\widetilde{Q}}(dy) \right] \sigma_t X_{t-} \kappa \\
&= c^2 \sigma_t X_{t-} \left[1 - \frac{\kappa}{\kappa_t^{\widetilde{Q}}} \right] F_x(t, S_{t-}, u) + \int_{\mathbb{R}} J(t, y, u) y \left[1 - \frac{\kappa}{\kappa_s^{\widetilde{Q}}} (1 + y G_t) \right] \nu(dy).
\end{aligned}$$

Concerning the orthogonality between $K(\cdot, u)$, (7.15), and M, we know by using Properties 2.2.15, (7.4) and the continuity of the finite variation part that

$$\begin{aligned}
\langle K, M \rangle_t =& \langle K^c + K^d, M \rangle_t \\
=& \langle (F_x(\cdot, S_-) - \xi(s, u)) \cdot \langle M, M \rangle_t \\
&+ \int_0^{\cdot} \int_{\mathbb{R}} (J(s, y, u) - F_x(s, S_{s-}) \sigma_s X_{s-} y) \widetilde{M}(ds, y), M \rangle_t \\
=& \int_0^t (F_x(s, S_{s-}) - \xi(s, u)) \sigma_s^2 X_{s-}^2 \kappa ds \\
&+ \int_0^t \int_{\mathbb{R}} (J(s, y, u) - F_x(s, S_{s-}) \sigma_s X_{s-} y) \sigma_s X_{s-} y \nu(dy) ds \\
=& \int_0^t F_x(s, S_{s-}) \sigma_s^2 X_{s-}^2 c^2 ds - \int_0^t \xi(s, u) \sigma_s^2 X_{s-}^2 \kappa ds \\
&+ \int_0^t \int_{\mathbb{R}} J(s, y, u) \sigma_s X_{s-} y \nu(dy) ds.
\end{aligned}$$

Comparing this with (7.16) we see that K is orthogonal to M under P if and only if K is a P-local martingale, but as shown with an explicit example in Chapter 3 this is not true in general. Note that the necessary condition is claim dependent, hence it is possible that for a specific claim the condition will be satisfied, but it will not hold in general.
On the other hand, it is true that the $\widetilde{Q}$-local martingale $K(\cdot,u)$, (7.15), is orthogonal to X under the minimal martingale measure $\widetilde{Q}$:

$$
\begin{aligned}
&\langle K, X\rangle_t^{\widetilde{Q}} \\
&= \langle (F_x(\cdot, S_-) - \xi(s,u)) \cdot X, X\rangle_t^{\widetilde{Q}} \\
&\quad + \langle \int_0^{\cdot} \int_{\mathbb{R}} (J(s,y,u) - F_x(s,S_{s-})\sigma_s X_{s-} y) \widetilde{M}^{\widetilde{Q}}(ds,y), X\rangle_t^{\widetilde{Q}} \\
&= \int_0^t (F_x(s,S_{s-}) - \xi(s,u))\sigma_s^2 X_{s-}^2 \kappa_s^{\widetilde{Q}} ds \\
&\quad + \int_0^t \int_{\mathbb{R}} (J(s,y,u) - F_x(s,S_{s-})\sigma_s X_{s-} y)\sigma_s X_{s-} y \nu^{\widetilde{Q}}(dy) ds \\
&= \int_0^t c^2 \sigma_s X_{s-} F_x(s,S_{s-},u) + \int_{\mathbb{R}} J(s,y,u) y \nu_s^{\widetilde{Q}}(dy)\sigma_s X_{s-} ds \\
&\quad - \int_0^t \xi(s,u)\kappa_s^{\widetilde{Q}} \sigma_s^2 X_{s-}^2 ds \\
&= \int_0^t \left[c^2 \sigma_s X_{s-} F_x(s,S_{s-},u) + \int_{\mathbb{R}} J(s,y,u) y \nu_s^{\widetilde{Q}}(dy) \right] \sigma_s X_{s-} ds \\
&\quad - \int_0^t \left[\frac{c^2}{\kappa_s^{\widetilde{Q}}} F_x(s,S_{s-},u) + \frac{1}{\sigma_s X_{s-} \kappa_s^{\widetilde{Q}}} \int_{\mathbb{R}} J(s,x,y) y \nu_s^{\widetilde{Q}}(dy) \right] \kappa_s^{\widetilde{Q}} \sigma_s^2 (X_{s-})^2 ds \\
&= 0.
\end{aligned}
$$

This calculation shows that the compensator of $K(\cdot,u)X$ is equal to zero, which means that $K(\cdot,u)X$ is a local martingale under $\widetilde{Q}$ and thus that $K(\cdot,u)$ is orthogonal to X under $\widetilde{Q}$.

7.4 The correct locally risk-minimizing strategy

We assume the same setting as in Riesner (2006a) for the insurance and the financial market. We adapt here the proofs of Colwell and Elliott (1993) to our setting. Of course this calculation is rather cumbersome and unnecessary long, because at the time the article was written the theory concerning locally risk-minimization was less developed e.g. the general form of the MMM was still not known. Although this calculation is no longer useful in practice, we repeat it here to show how the MMM comes out naturally as the only correct choice from the set of martingale measures. This makes of course sense in the continuous case where we can use the fact that the martingale and the orthogonality property are preserved, while in the discontinuous case this is less obvious without going into the details of the proofs concerning locally risk-minimization, see Chapter 4.
We start by summarizing the reasoning followed in this section.
We want to determine the Föllmer-Schweizer decomposition of the portfolio under the original measure P. The value of this portfolio at maturity should equal the necessary amount and should be a martingale under an equivalent martingale measure which will be specified later on.
To obtain such a portfolio, we will impose all the necessary conditions (see (i)-(v) in the proof of Theorem 7.4.1 on page 150) on the unknown decomposition. In this way, we will not only derive the desired decomposition, but we will also obtain the MMM as the equivalent measure.
More concrete we will start by describing all the equivalent probability measures by the Girsanov density (7.17). We will assume that the functions j and h in (7.17) are such that X is a local martingale under the equivalent measure $\breve{Q}$. This will lead to a first relation (7.21) for j and h. Then we will derive the dynamics (7.23) of the portfolio V, which is a $\breve{Q}$-local martingale. Finally, we will propose a decomposition of V under P, which is a Föllmer-Schweizer decomposition and determine the unknowns $\phi(\cdot, u)$ and $\Gamma(\cdot, u)$ (see (ii) in the proof of Theorem 7.4.1 on page 150) in this decomposition. At the same time, we will get the explicit expressions for j and h (see Theorem 7.4.2), in other words the $\breve{Q}$ measure will be completely defined. We will also show in Corollary 7.4.3 that this equivalent measure $\breve{Q}$ is the minimal martingale measure.
Let us start with the first step: From Theorem 2.3.8, we know that the set of equivalent probability measures is described by the following Girsanov density

for processes $j \in L(X^c)$ and $h - 1 \in G_{\text{loc}}(N)$:

$$\begin{aligned} D_t =&1 + \int_0^t D_{s-}\, j(s, X_{s-}) dW_s \\ &+ \int_0^t \int_{\mathbb{R}} D_{s-}[h(s, X_{s-}, y) - 1](N(ds, dy) - \nu(dy)ds). \end{aligned} \tag{7.17}$$

We assume that j and h are such that X is a martingale under this equivalent martingale measure, which we denote by $\breve{Q}$. Under the measure $\breve{Q}$, $W_t - \int_0^t j(s, X_{s-})ds$ is a standard Wiener process and the compensator of $N(dt, dy)$ is $h(t, X_{t-}, y)\nu(dy)dt$. We recall that under the original measure P_1 the process of the discounted risky asset is given by:

$$\begin{aligned} dX_t =&(b_t + a\sigma_t - r_t)X_{t-}dt + \sigma_t X_{t-} d(cW_t + \widetilde{M}_t) \\ =&(b_t + a\sigma_t - r_t)X_{t-}dt + \sigma_t X_{t-} c dW_t + \int_{\mathbb{R}} \sigma_t X_{t-} y(N(dt, dy) - \nu(dy)dt). \end{aligned} \tag{7.18}$$

Thus, X has the decomposition $X_t = X_0 + A_t + M_t$ with $A_t = \int_0^t (b_u + a\sigma_u - r_u)X_{u-}du$ the finite variation part and with

$$M_t = \int_0^t \sigma_u X_{u-} d(cW_u + \widetilde{M}_u) \tag{7.19}$$

the square-integrable martingale part for which $M_0 = 0$. We rewrite the process of the risky asset as follows:

$$\begin{aligned} dX_t =&\sigma_t X_{t-} c[dW_t - j(t, X_{t-})dt] + \int_{\mathbb{R}} \sigma_t X_{t-} y[N(dt, dy) - h(t, X_{t-}, y)\nu(dy)dt] \\ &+ (b_t + a\sigma_t - r_t)X_{t-}dt \\ &+ \left[\sigma_t X_{t-} c j(t, X_{t-}) + \int_{\mathbb{R}} \sigma_t X_{t-} y[h(t, X_{t-}, y) - 1]\nu(dy)\right] dt. \end{aligned} \tag{7.20}$$

We know from Theorem 2.3.5 that $N(dt, dy) - h(t, X_{t-}, y)\nu(dy)dt$ and $dW_t - j(t, X_{t-})dt$ are local martingales under the new measure $\breve{Q}$. Hence, the process for the discounted risky asset is a local martingale under the measure $\breve{Q}$ if and only if

$$(b_t + a\sigma_t - r_t) + \sigma_t c\, j(t, X_{t-}) + \int_{\mathbb{R}} \sigma_t[h(t, X_{t-}, y) - 1] y\nu(dy) = 0. \tag{7.21}$$

Theorem 7.4.1. *The $\breve{Q}$-local martingale $V(t,u)$ given by*

$$V(t,u) := E^{\breve{Q}}[B_u^{-1} g(u,S_u)|\mathcal{F}_t] = B_t^{-1} F(t,S_t,u) \quad \textit{for} \quad 0 \le t < u \le T \quad (7.22)$$

has the following process under the new measure $\breve{Q}$:

$$\begin{aligned} V(t,u) =& V(0,u) + \int_0^t F_x(s,S_{s-},u)\sigma_s X_{s-} c[dW_s - j(s,X_{s-})ds] \\ &+ \int_0^t \int_{\mathbb{R}} J(s,y,u)[N(ds,dy) - h(s,X_{s-},y)\nu(dy)ds]. \end{aligned} \quad (7.23)$$

Proof. We first derive the relationship between the non-discounted risky asset S and the discounted risky asset X:

$$dX_t = d\left(\frac{S_t}{B_t}\right) = B_t^{-1} dS_t - S_t B_t^{-2} dB_t = B_t^{-1} dS_t - X_t r_t dt.$$

Therefore $B_t^{-1} dS_t = dX_t + X_t r_t dt$. We apply Itô's formula for semimartingales to (7.22) and find the dynamics of $V(t,u)$ by using (7.11), but now under $\breve{Q}$ instead of $\widetilde{Q}$. These two measures will of course appear to be the same, but we cannot presume this at this stage. The notation $J(t,x,u)$, see (7.11), will also be used here.

$$\begin{aligned} & dV(t,u) \\ &= -r_t B_t^{-1} F(t,S_{t-},u)dt \\ &\quad + B_t^{-1}\left[F_t(t,S_{t-},u)dt + F_x(t,S_{t-},u)dS_t + \frac{1}{2}F_{xx}(t,S_{t-},u)d[S,S]_t^c\right] \\ &\quad + \int_{\mathbb{R}} B_t^{-1}[F(t,S_{t-} + \sigma_t S_{t-}\, y,u) - F(t,S_{t-},u)]N(dt,dy) \\ &\quad - \int_{\mathbb{R}} B_t^{-1} F_x(t,S_{t-},u)\sigma_t S_{t-}\, yN(dt,dy) \\ &= \left(-r_t B_t^{-1} F(t,S_{t-},u) + B_t^{-1} F_t(t,S_{t-},u) + X_t r_t F_x(t,S_{t-},u)\right) dt \\ &\quad + \frac{1}{2}F_{xx}(t,S_{t-},u)\sigma_t^2 B_t^{-1} S_{t-}^2 c^2 dt \\ &\quad + \int_{\mathbb{R}} J(t,y,u)h(t,X_{t-},y)\nu(dy)dt \end{aligned}$$

$$
\begin{aligned}
&- \int_{\mathbb{R}} B_t^{-1} F_x(t, S_{t-}, u)\sigma_t S_{t-}\, y h(t, X_{t-}, y)\nu(dy)dt + F_x(t, S_{t-}, u)dX_t \\
&+ \int_{\mathbb{R}} J(t, y, u)[N(dt, dy) - h(t, X_{t-}, y)\nu(dy)dt] \\
&- \int_{\mathbb{R}} B_t^{-1} F_x(t, S_{t-}, u)\sigma_t S_{t-}\, y[N(dt, dy) - h(t, X_{t-}, y)\nu(dy)dt].
\end{aligned}
$$

This is a $\breve{Q}$-local martingale and thus the drift term is zero. Next, substituting equation (7.20) for the discounted risky asset, we arrive at:

$$
\begin{aligned}
&dV(t, u) \\
&= F_x(t, S_{t-}, u)\sigma_t X_{t-} c[dW_t - j(t, X_{t-})dt] \\
&\quad + F_x(t, S_{t-}, u) \int_{\mathbb{R}} \sigma_t X_{t-}\, y[N(dt, dy) - h(t, X_{t-}, y)\nu(dy)dt] \\
&\quad + \int_{\mathbb{R}} [J(t, y, u) - F_x(t, S_{t-}, u)\sigma_t X_{t-}\, y][N(dt, dy) - h(t, X_{t-}, y)\nu(dy)dt] \\
&= F_x(t, S_{t-}, u)\sigma_t X_{t-} c[dW_t - j(t, X_{t-})dt] \\
&\quad + \int_{\mathbb{R}} J(t, y, u)[N(dt, dy) - h(t, X_{t-}, y)\nu(dy)dt].
\end{aligned}
$$

Integrating this last equation gives the required result. □

We now come to a next step in the procedure. From Theorem 4.3.11 and due to the assumptions we made, we know that the FS decomposition gives the locally risk-minimizing hedging strategy. In order to find this FS decomposition, we impose the following conditions:

(i) $V_{u,u}(\varphi) = B_u^{-1} g(u, S_u)$,
(ii) $V_{t,u}(\varphi) = V_{0,u}(\varphi) + \int_0^t \phi(s, u)dX_s + \Gamma(t, u)$,
(iii) $\Gamma(t, u)$ is a martingale under P, and $\Gamma(t, u)$ is orthogonal to the martingale part M of the discounted risky asset under P,

where the functions $\phi(s, u)$ and $\Gamma(t, u)$ in the Föllmer-Schweizer decomposition (ii) of the portfolio are unknown.
The earlier made requirement:

(iv) the chosen equivalent probability measure $\breve{Q}$ is such that X is a local martingale under $\breve{Q}$, as well as the condition

(v) $V_{t,u}(\varphi)$ is a local martingale under $\breve{Q}$,

help us to find the Föllmer-Schweizer decomposition (ii) explicitly.
Remark that in this section we do not always write the dependence on u explicitly as e.g. in the number of risky assets ϕ_t.

We now state our result.

Theorem 7.4.2. *The locally risk-minimizing hedging strategy for the claim $g(u, S_u)$ can be found by performing a change of measure as described in formula (7.17), with*

$$h(t, X_{t-}, y) = 1 - \frac{(b_t + a\sigma_t - r_t)y}{\sigma_t[c^2 + \int_{\mathbb{R}} x^2 \nu(dx)]}, \tag{7.24}$$

$$j(t, X_{t-}) = - \frac{(b_t + a\sigma_t - r_t)c}{\sigma_t[c^2 + \int_{\mathbb{R}} x^2 \nu(dx)]} \tag{7.25}$$

for all $0 \leq t \leq u \leq T$, i.e.

$$\begin{aligned} D_t =& 1 - \int_0^t D_{s-} \frac{(b_s + a\sigma_s - r_s)c}{\sigma_s[c^2 + \int_{\mathbb{R}} x^2 \nu(dx)]} dW_s \\ &- \int_0^t \int_{\mathbb{R}} D_{s-} \frac{(b_s + a\sigma_s - r_s)y}{\sigma_s[c^2 + \int_{\mathbb{R}} x^2 \nu(dx)]} (N(ds, dy) - \nu(dy)ds). \end{aligned} \tag{7.26}$$

The Föllmer-Schweizer decomposition under P of the associated portfolio, satisfying the conditions (i)-(v), is:

$$\begin{aligned} V_{t,u}(\varphi) =& V_{0,u}(\varphi) + \int_0^t \phi(s, u) dX_s + \int_0^t \varsigma^{(a)}(s, u) dW_s \\ &+ \int_0^t \int_{\mathbb{R}} \varsigma^{(b)}(s, y, u) \widetilde{M}(ds, dy), \end{aligned} \tag{7.27}$$

with

$$\phi(t, u) = \frac{F_x(t, S_{t-}, u)\sigma_t X_{t-} c^2 + \int_{\mathbb{R}} J(t, y, u) y \nu(dy)}{\sigma_t X_{t-}[c^2 + \int_{\mathbb{R}} x^2 \nu(dx)]} \tag{7.28}$$

$$\varsigma^{(a)}(t, u) = c\sigma_t X_{t-}[F_x(t, S_{t-}, u) - \phi(t, u)] \tag{7.29}$$

$$\varsigma^{(b)}(t, y, u) = J(t, y, u) - y\sigma_t X_{t-} \phi(t, u). \tag{7.30}$$

The strategy at time t, $0 \leq t \leq u$, is $\varphi(t,u) = (\phi(t,u), \eta(t,u))$ with $\eta(t,u) = V_{t,u}(\varphi) - \phi(t,u)X_t$.

Proof. Out of conditions (v), (i) and (7.22), we find that

$$V_{t,u}(\varphi) = E^{\breve{Q}}[V_{u,u}(\varphi)|\mathcal{F}_t] = E^{\breve{Q}}[B_u^{-1}g(u,S_u)|\mathcal{F}_t] = V(t,u),$$

in particular $V_{0,u}(\varphi) = V(0,u)$. Using the required form described in (ii), and the decomposition (7.23) of $V(t,u)$, we see that $\Gamma(t,u)$ can be expressed as follows:

$$\begin{aligned}
&\Gamma(t,u)\\
&= \int_0^t F_x(s,S_{s-},u)\sigma_s X_{s-}c[dW_s - j(s,X_{s-})ds]\\
&\quad + \int_0^t \int_{\mathbb{R}} J(s,y,u)[N(ds,dy) - h(s,X_{s-},y)\nu(dy)ds] - \int_0^t \phi(s,u)dX_s\\
&= \int_0^t F_x(s,S_{s-},u)\sigma_s X_{s-}cdW_s + \int_0^t \int_{\mathbb{R}} J(s,y,u)[N(ds,dy) - \nu(dy)ds]\\
&\quad - \int_0^t F_x(s,S_{s-},u)\sigma_s X_{s-}cj(s,X_{s-})ds\\
&\quad + \int_0^t \int_{\mathbb{R}} J(s,y,u)[1 - h(s,X_{s-},y)]\nu(dy)ds - \int_0^t \phi(s,u)(b_s + a\sigma_s - r_s)X_{s-}ds\\
&\quad - \int_0^t \phi(s,u)\sigma_s X_{s-}cdW_s - \int_0^t \int_{\mathbb{R}} \phi(s,u)\sigma_s X_{s-}y[N(ds,dy) - \nu(dy)ds]\\
&= \int_0^t [F_x(s,S_{s-},u) - \phi(s,u)]\sigma_s X_{s-}cdW_s - \int_0^t \phi(s,u)(b_s + a\sigma_s - r_s)X_{s-}ds\\
&\quad + \int_0^t \int_{\mathbb{R}} [J(s,y,u) - \sigma_s X_{s-}\phi(s,u)y][N(ds,dy) - \nu(dy)ds]\\
&\quad - \int_0^t F_x(s,S_{s-},u)\sigma_s X_{s-}cj(s,X_{s-})ds\\
&\quad + \int_0^t \int_{\mathbb{R}} J(s,y,u)[1 - h(s,X_{s-},y)]\nu(dy)ds.
\end{aligned}$$

We used here formula (7.18) and not (7.20) to describe the process of the discounted risky asset because we want the dynamics of $\Gamma(\cdot,u)$ under the original measure P. We now easily see that if we want $\Gamma(\cdot,u)$ to be a local martingale under P, then the drift term in the previous relation should be zero. Hence,

$$F_x(t, S_{t-}, t)\sigma_t X_{t-} cj(t, X_{t-}) \\ = \int_{\mathbb{R}} J(t, y, u)[1 - h(t, X_{t-}, y)]\nu(dy) - \phi(t, u)(b_t + a\sigma_t - r_t)X_{t-}, \quad (7.31)$$

for all $t \in [0, u]$ P-almost surely. Furthermore, $\Gamma(\cdot, u)$ and M, the martingale part of the risky asset, should be orthogonal under P. This will be the case if $\Gamma(\cdot, u)M$ is a local martingale under P. The process $\Gamma(\cdot, u)$ is thus given by

$$\begin{aligned} \Gamma(t, u) &= \int_0^t [F_x(s, S_{s-}, u) - \phi(s, u)]\sigma_s X_{s-} cdW_s \\ &\quad + \int_0^t \int_{\mathbb{R}} [J(s, y, u) - \sigma_s X_{s-}\phi(s, u)y][N(ds, dy) - \nu(dy)ds] \quad (7.32) \\ &= \int_0^t \varsigma^{(a)}(s, u)dW_s + \int_0^t \int_{\mathbb{R}} \varsigma^{(b)}(s, y, u)[N(ds, dy) - \nu(dy)ds], \end{aligned}$$

where $\varsigma^{(a)}(t, u)$ and $\varsigma^{(b)}(t, u)$ are introduced in (7.29) and (7.30) and the process of M follows from (7.19)

$$M_t = \int_0^t \sigma_s X_{s-} cdW_s + \int_0^t \int_{\mathbb{R}} \sigma_s X_{s-} y[N(ds, dy) - \nu(dy)ds].$$

Thus the product $\Gamma(\cdot, u)M$ is a local martingale under P if and only if the drift term in its dynamics is equal to zero. This leads to the following condition for all $t \in [0, u]$ P-almost surely:

$$[F_x(t, S_{t-}, u) - \phi(t, u)]\sigma_t^2 X_{t-}^2 c^2 \\ = -\int_{\mathbb{R}} [J(t, y, u) - \sigma_t X_{t-}\phi(t, u)y]\sigma_t X_{t-} y\nu(dy) = 0. \quad (7.33)$$

By solving this equation for $\phi(t, u)$ we obtain (7.28). We now recall condition (7.21), which insures that X is a martingale under the measure $\breve{Q}$, and which multiplied with $F_x(t, S_{t-}, u)X_{t-}$ equals:

$$F_x(t, S_{t-}, u)X_{t-} \\ \times \left[(b_t + a\sigma_t - r_t) + \sigma_t cj(t, X_{t-}) + \int_{\mathbb{R}} \sigma_t [h(t, X_{t-}, y) - 1]y\nu(dy)\right] = 0. \quad (7.34)$$

We substitute the value (7.28) for $\phi(t,u)$ in the condition (7.31) which insures that $\Gamma(\cdot,u)$ is a local martingale under P:

$$\begin{aligned}&-F_x(t,S_{t-},u)\sigma_t X_{t-} c j(t,X_{t-}) + \int_{\mathbb{R}} J(t,y,u)[1-h(t,X_{t-},y)]\nu(dy)\\ &= \frac{F_x(t,S_{t-},u)\sigma_t X_{t-} c^2 + \int_{\mathbb{R}} J(t,y,u) y \nu(dy)}{\sigma_t[c^2 + \int_{\mathbb{R}} x^2 \nu(dx)]}(b_t + a\sigma_t - r_t). \end{aligned} \tag{7.35}$$

Eliminating j from the equations (7.34) and (7.35) leads to:

$$\begin{aligned}&\int_{\mathbb{R}} [\sigma_t F_x(t,S_{t-},u) X_{t-} y - J(t,y,u)]\\ &\qquad \times \left[h(t,X_{t-},y) - 1 + \frac{(b_t + a\sigma_t - r_t)y}{\sigma_t(c^2 + \int_{\mathbb{R}} x^2\nu(dx))}\right]\nu(dy) = 0. \end{aligned} \tag{7.36}$$

A possible solution which is independent of the claim to be hedged, is $h(t,X_{t-},y)$ given by (7.24).
We remark that it is possible that equation (7.36) has not a unique solution. This is the case when $\sigma_t F_x(t,S_{t-},u)X_{t-}y - J(t,y,u) \equiv 0$. It is important to note that this equation is however claim-dependent and therefore it will not hold for all possible claims.
We now calculate j from equation (7.21)

$$\begin{aligned} j(t,S_{t-}) =& \frac{1}{\sigma_t c}\left[-(b_t + a\sigma_t - r_t) + \int_{\mathbb{R}} \sigma_t [\frac{(b_t + a\sigma_t - r_t)y}{\sigma_t(c^2 + \int_{\mathbb{R}} x^2\nu(dx))}] y\nu(dy)\right]\\ =& \frac{(b_t + a\sigma_t - r_t)}{\sigma_t c[c^2 + \int_{\mathbb{R}} x^2\nu(dx)]}\left[-c^2 - \int_{\mathbb{R}} x^2\nu(dx) + \int_{\mathbb{R}} y^2\nu(dy)\right]\\ =& -\frac{(b_t + a\sigma_t - r_t)c}{\sigma_t[c^2 + \int_{\mathbb{R}} x^2\nu(dx)]}, \end{aligned}$$

which is precisely (7.25). □

Corollary 7.4.3. *The equivalent martingale measure $\breve{Q}$ obtained through the change of measure* (7.26) *is precisely the minimal martingale measure $\widetilde{Q}$.*

Proof. In view of relation (7.4), the functions h (7.24) and j (7.25) satisfy the following relationship with the change of measure parameter G of Chan (1999):

$$\begin{aligned} h(t,X_{t-},y) =& 1 + G_t y\\ j(t,X_{t-}) =& cG_t. \end{aligned}$$

Hence the Girsanov density (7.17), describing the change of measure from P to $\breve{Q}$, can be rewritten as:

$$D_t = 1 + \int_0^t D_{s-}G_s c dW_s + \int_0^t \int_{\mathbb{R}} D_{s-}G_s y(N(ds,dy) - \nu(dy)ds).$$

From Section 2.3.2.1 we know this is exactly the MMM. We remark that this result is not a surprise at all, because we are searching for a martingale measure under which $V_{0,u}(\varphi) + \int_0^t \phi(s,u)dX_s + \Gamma(t,u)$ is a local martingale. The first two terms are obviously local martingales under a martingale measure and the only restriction comes from the third term. Therefore, the only claim-independent measure will be the one that demands that every local martingale orthogonal to the martingale part of the discounted risky asset under the original measure is also a local martingale under the new measure. This is exactly the description of the definition of the minimal martingale measure. □

We now compare the number of risky assets proposed in Riesner (2006a) to the one we derived for the claim $g(u, S_u)$. The number of risky assets found in Riesner (2006a) is given by $\xi(t,u)$ in (7.10):

$$\xi(t,u) = \frac{F_x(t, S_{t-}, u)\sigma_t X_{t-} c^2 + \int_{\mathbb{R}} J(t,y,u) y \nu_t^{\widetilde{Q}}(dy)}{\sigma_t X_{t-}[c^2 + \int_{\mathbb{R}} x^2 \nu^{\widetilde{Q}}(dx)]},$$

while the number of risky assets to get the correct locally risk-minimizing strategy is given by $\phi(t,u)$ in (7.28). We see that the structure of the formula for both numbers of risky assets is the same, but that we have to calculate it under the right measure. The first one is under the minimal martingale measure $\widetilde{Q}$, the second one under the original measure P. Invoking (7.8) and (7.5), we establish the following relation between $\xi(t,u)$ (7.10) and $\phi_t(u)$ (7.28):

$$\begin{aligned}\phi(t,u) &= \frac{F_x(t, S_{t-}, u)\sigma_t X_{t-} c^2 + \int_{\mathbb{R}} J(t,y,u) y \nu(dy)}{\sigma_t X_{t-}[c^2 + \int_{\mathbb{R}} x^2 \nu(dx)]} \\ &= \frac{\kappa + \int_{\mathbb{R}} x^3 G_t \nu(dx)}{\kappa} \xi(t,u) - \frac{1}{\kappa \sigma_t X_{t-}} \int_{\mathbb{R}} y^2 J(t,y,u) G_t \nu(dy).\end{aligned}$$

Hence,

$$\xi(t,u) - \phi(t,u) = -\frac{\int_{\mathbb{R}} x^3 G_t \nu(dx)}{\kappa} \xi(t,u) + \frac{1}{\kappa \sigma_t X_{t-}} \int_{\mathbb{R}} y^2 J(t,y,u) G_t \nu(dy). \tag{7.37}$$

Clearly $\xi(t,u)$ is equal to $\phi(t,u)$ if $G_t \equiv 0$ for all $0 \le t \le u$ or if $\nu(dy) \equiv 0$, which are both trivial cases. The former corresponds to the case $P_1 = \widetilde{Q}$ meaning that the risky asset is already a martingale under the original measure P. The latter refers to a Lévy process which has no jumps and hence, we are in the Black-Scholes setting where orthogonality is preserved by the continuity of the risky asset. In general the equation (7.37) will be non-zero, see Section 3.4 for an explicit example.

The associated risk process is not risk-minimizing in the present setting, but only locally risk-minimizing. However we are still able to calculate the risk and although we don't have a global risk-minimizing strategy, we are able to calculate locally how much risk remains. We introduce the notation ρ_t for

$$\rho_t := \int_{\mathbb{R}} J(t,y,u) y \nu(dy).$$

For the risk process we find by using successively (4.2), (4.1), condition (ii), (7.32), equation I.4.6 of Jacod and Shiryaev (2002) (see page 90), (7.28) and (7.5) that:

$$\begin{aligned}
&R_{t,u}(\varphi)\\
&= E^P[(C_T(\varphi) - C_t(\varphi))^2 | \mathcal{F}_t]\\
&= E^P[(V_T(\varphi) - \int_0^T \phi(s,u) dX_s - V_{t,u}(\varphi) + \int_0^t \phi(s,u) dX_s)^2 | \mathcal{F}_t]\\
&= E^P[(\Gamma(T) - \Gamma(t))^2 | \mathcal{F}_t]\\
&= E^P \Bigg[\Bigg(\int_t^T [F_x(s, S_{s-}, u) - \phi(s)] \sigma_s X_{s-} c dW_s\\
&\qquad + \int_t^T \int_{\mathbb{R}} [J(s,y,u) - \sigma_s X_{s-} \phi(s) y][N(ds,dy) - \nu(dy)ds] \Bigg)^2 | \mathcal{F}_t \Bigg]\\
&= \int_t^T E^P \left[[F_x(s, S_{s-}, u) - \phi(s)]^2 \sigma_s^2 X_{s-}^2 c^2 | \mathcal{F}_t \right] ds\\
&\quad + \int_t^T \int_{\mathbb{R}} E^P \left[[J(s,y,u) - \sigma_s X_{s-} \phi(s) y]^2 | \mathcal{F}_t \right] \nu(dy) ds\\
&= \int_t^T E^P \left[\left[F_x(s, S_{s-}, u) - \frac{F_x(s, S_{s-}, u) c^2}{\kappa} - \frac{\rho_s}{\sigma_s X_{s-} \kappa} \right]^2 \sigma_s^2 X_{s-}^2 c^2 | \mathcal{F}_t \right] ds
\end{aligned}$$

$$+\int_t^T\int_{\mathbb{R}} E^P\left[\left[J(s,y,u)-\frac{F_x(s,S_{s-},u)c^2\sigma_s X_{s-}y}{\kappa}-\frac{\rho_s y}{\kappa}\right]^2|\mathcal{F}_t\right]\nu(dy)ds$$

$$=\int_t^T E^P\left[\left[F_x(s,S_{s-},u)\frac{\kappa-c^2}{\kappa}-\frac{\rho_s}{\sigma_s X_{s-}\kappa}\right]^2\sigma_s^2X_{s-}^2c^2|\mathcal{F}_t\right]ds$$

$$+\int_t^T\int_{\mathbb{R}} E^P\left[\left[J^2(s,y,u)+\frac{F_x^2(s,S_{s-},u)c^4\sigma_s^2(X_{s-})^2y^2}{\kappa^2}+\frac{\rho_s^2y^2}{\kappa^2}\right.\right.$$

$$-2F_x(s,X_{s-},u)\sigma_sX_{s-}J(s,y,u)y\frac{c^2}{\kappa}-2J(s,y,u)\frac{y}{\kappa}\rho_s$$

$$\left.\left.+2F_x(s,X_{s-},u)\sigma_sX_{s-}\frac{c^2}{\kappa^2}y^2\rho_s\right]|\mathcal{F}_t\right]\nu(dy)ds$$

$$=\int_t^T\frac{c^2}{\kappa^2}E^P\left[[F_x^2(s,S_{s-},u)v^2\sigma_s^2(X_{s-})^2-2F_x(s,S_{s-},u)\sigma_sX_{s-}v\rho_s+\rho_s^2]|\mathcal{F}_t\right]ds$$

$$+\int_t^T\int_{\mathbb{R}} E^P\left[J^2(s,y,u)\nu(dy)+\frac{F_x^2(s,S_{s-},u)c^4\sigma_s^2(X_{s-})^2}{\kappa^2}v+\frac{\rho_s^2}{\kappa^2}v-2\frac{\rho_s^2}{\kappa^2}\kappa\right.$$

$$\left.-2F_x(s,X_{s-},u)\sigma_sX_{s-}\rho_s\frac{c^2\kappa}{\kappa^2}+2F_x(s,X_{s-},u)\sigma_sX_{s-}\frac{c^2}{\kappa^2}\rho_sv|\mathcal{F}_t\right]ds$$

$$=\int_t^T\frac{c^2}{\kappa^2}E^P\left[F_x^2(s,S_{s-},u)v^2\sigma_s^2(X_{s-})^2-2F_x(s,S_{s-},u)\sigma_sX_{s-}v\rho_s+\rho_s^2|\mathcal{F}_t\right]ds$$

$$+\int_t^T E^P\left[\int_{\mathbb{R}} J^2(s,y,u)\nu(dy)+\frac{F_x^2(s,S_{s-},u)c^4\sigma_s^2(X_{s-})^2}{\kappa^2}v\right.$$

$$\left.-\frac{\rho_s^2}{\kappa^2}(c^2+\kappa)-2F_x(s,X_{s-},u)\sigma_sX_{s-}\rho_s\frac{c^4}{\kappa^2}|\mathcal{F}_t\right]ds.$$

This risk process is equivalent with the one given in Riesner (2006a). All the processes are there taken under the measure $\widetilde{Q}$ while we work under the original measure P except for $F(t,x,u)$, which is still defined under the measure $\widetilde{Q}$ as $E^{\widetilde{Q}}[B_tB_T^{-1}g(T,S_T)|\mathcal{F}_t]$.

7.5 The locally risk-minimizing hedging strategy for unit-linked contracts

We report here the results given in Riesner (2006a) for the pure endowment and the term insurance, but with the correct decomposition for $B_t^{-1}F(t,S_t,T)$ given in formula (7.27).
We note that Corollary 4 and Corollary 6 in Riesner (2006a) are correct because he only claims he found the risk-minimizing hedging strategy under the MMM. His remark under Corollary 4 shows again his belief that he actually found not only the risk-minimizing under the MMM but also the locally risk-minimizing hedging strategy under the original measure.

7.5.1 The pure endowment

The total claim for N pure endowment contracts is

$$H = B_T^{-1} g(S_T) \sum_{i=1}^{N} \mathbb{1}_{\{T_i > T\}} = B_T^{-1} g(S_T)(N - N_T^I).$$

We remark that u is set here equal to the fixed time of maturity T.
The Föllmer-Schweizer decomposition for the claim H is given by:

$$\begin{aligned}
\hat{V}_{t,T} =& \hat{V}_{0,T} + \int_0^t (N - N_{s-}^I)\, {}_{T-s}p_{x+s}\phi(s,T)dX_s + K^H(t,T) \quad 0 \le t \le T, \\
K^H(t,T) =& \int_0^t (N - N_{s-}^I)\, {}_{T-s}p_{x+s}\varsigma^{(a)}(s,T)dW_s \\
&+ \int_0^t \int_{\mathbb{R}} (N - N_{s-}^I)\, {}_{T-s}p_{x+s}\varsigma^{(b)}(s,y,T)\widetilde{M}(ds,dy) \\
&- \int_0^t B_s^{-1}F(s,S_{s-},T)\, {}_{T-s}p_{x+s}dM_s^I.
\end{aligned}$$

The optimal portfolio invests $\phi^*(t,T) = (N - N_{t-}^I)\, {}_{T-t}p_{x+t}\phi(t,T)$ in the risky asset and $(N - N_t^I)\, {}_{T-t}p_{x+t}B_t^{-1}F(t,S_t,T) - \phi^*(t,T)X_t$ in the riskless asset for

$0 \leq t \leq T$. The associated risk process is then given by

$$
\begin{aligned}
&R_{t,T}(\varphi^*) \\
&= \int_t^T {}_{T-s}p_{x+s}^2 E[(N-N_s^I)^2] E\left[(\varsigma^{(a)}(s,T))^2 + \int_{\mathbb{R}} (\varsigma^{(b)}(s,y,T))^2 \nu(dy) | \mathcal{F}_t \right] ds \\
&\quad + (N-N_t^I)\, {}_{T-t}p_{x+t} \int_t^T {}_{T-s}p_{x+s} \mu_{x+s} B_s^{-2} E[F^2(s,S_s,T)|\mathcal{F}_t] ds.
\end{aligned}
$$

7.5.2 The term insurance

The payment $g(u, S_u)$ is time-dependent but we assume that the insurance company only pays out at time T. Thus the claim for a portfolio of N term insurance contracts is

$$
H = B_T^{-1} \sum_{i=1}^{N} B_T B_{T_i}^{-1} g(T_i, S_{T_i}) \mathbb{1}_{\{T_i \leq T\}} = \int_0^T B_u^{-1} g(u, S_u) dN_u^I.
$$

The Föllmer-Schweizer decomposition for the claim H is given by:

$$
\begin{aligned}
\hat{V}_{t,T} =& \hat{V}_{0,T} + \int_0^t (N - N_{s-}^I) \phi^*(s,T) dX_s + K^H(t,T) \quad 0 \leq t \leq T, \\
\phi^*(t,T) =& \int_t^T {}_{u-t}p_{x+t}\, \mu_{x+u} \phi(t,u) du, \\
K^H(t,T) =& \int_0^t (N - N_{s-}^I) \chi^{(a)}(s) dW_s \\
&+ \int_0^t \int_{\mathbb{R}} (N - N_{s-}^I) \chi^{(b)}(s,y) \widetilde{M}(ds,dy) + \int_0^t \theta(s) dM_s^I,
\end{aligned}
$$

where

$$
\begin{aligned}
\chi^{(a)}(t,T) :=& \int_t^T {}_{u-t}p_{x+t}\, \mu_{x+u} \varsigma^{(a)}(t,u) du, \\
\chi^{(b)}(t,y,T) :=& \int_t^T {}_{u-t}p_{x+t}\, \mu_{x+t} \varsigma^{(b)}(t,y,u) du, \\
\theta(t,T) =& B_t^{-1} g(t,S_t) - \int_t^T B_t^{-1} F(t, S_{t-}, u)\, {}_{u-t}p_{x+t} \mu_{x+u} du.
\end{aligned}
$$

The optimal portfolio at time $0 \leq t \leq T$ invests $\xi^*(t,T) = (N - N^I_{t-})\phi^*(t,T)$ in the risky asset and

$$\int_0^t B_u^{-1} g(u, S_u) dN_u^I + (N - N_t^I) \int_t^T B_t^{-1} F(t, S_t, u)\, {}_{u-t}p_{x+t} \mu_{x+u} du - \xi^*(t,T) X_t$$

in the riskless asset. The associated risk process is then

$$\begin{aligned} R_{t,T}(\varphi^*) = &\int_t^T E[(N - N_t^I)^2] E\left[(\chi^{(a)}(s,T))^2 + \int_{\mathbb{R}} (\chi^{(b)}(s,y,T))^2 \nu(dy) | \mathcal{F}_t \right] ds \\ &+ (N - N_t^I) \int_t^T E[\theta_s^2(T)|\mathcal{F}_t]\, {}_{s-t}p_{x+t}\, \mu_{x+s} ds. \end{aligned}$$

7.6 Practical example

In this section we apply the theory described in Chapter 3 on a rather general, but quasi-left continuous semimartingale with $A_t = t$. The assumption of QLC is no restriction when one wants to determine the LRM hedging strategy, because continuity of the predictable finite variation part is a necessary condition. The example will generalize the one of Riesner (2006a) and we will explicitly show that we obtain the same result. We end with the determination of the correction term $\widetilde{\Phi}$, described in (3.21) for the example of Riesner (2006a).
Consider a market model for which $A_t = t$ and Assumption 3.5.1 holds. Let H be an $\mathcal{F}_T$-measurable random variable such that the process $\widetilde{V}^H$ satisfies

$$\widetilde{V}_t^H = E^{\widetilde{Q}}[H|\mathcal{F}_t] = f(t, X_t),$$

where $f(t,x)$ is a $C^{1,2}((0,T) \times \mathbb{R}^d)$-function. This case generalizes the examples that are frequently used in the literature, such as those treated in Colwell and Elliott (1993) and in Theorem 7.4.2. By applying Itô's formula, see Theorem 2.2.26, we find

$$\begin{aligned} \widetilde{V}_t^H = &\widetilde{V}_0^H + \int_0^t f_x(s, X_{s-}) dX_s + \int_0^t [f_t(s, X_s) + \frac{1}{2} c_s f_{xx}(s, X_{s-})] ds \\ &+ \sum\nolimits_{0<s\leq t} [f(s,X_s) - f(s,X_{s-}) - f_x(s,X_{s-})\Delta X_s] \\ = &\widetilde{V}_0^H + \int_0^t f_x(s, X_{s-}) dX_s^c + \int_0^t [f_t(s, X_s) + \frac{1}{2} c_s f_{xx}(s, X_{s-})] ds \end{aligned}$$

$$+\int_0^t \int_{\mathbb{R}^d} [f(s, X_{s-} + x) - f(s, X_{s-})]\mu(ds, dx).$$

Since $\widetilde{V}^H$ is a special semimartingale, then by compensating the last term in the RHS of the above equation and simplifying the resulting equation, we obtain

$$\widetilde{V}^H = \widetilde{V}_0^H + f_x(\cdot, X_-) \cdot X^c + \Big[f(\cdot, X_- + x) - f(\cdot, X_-)\Big] \star (\mu - \nu) + \widetilde{B},$$

where $\widetilde{B}$ is a predictable process with finite variation. Therefore, this leads to the description of the process $\widetilde{K}$ defined in (3.49), and hence to the FS decomposition of H as follows.

Corollary 7.6.1. *The following assertions hold:*

(1) The process $\widetilde{K}$ is given by

$$\widetilde{K} = f_x(\cdot, X_-) \cdot X^c + \Big[f(\cdot, X_- + x) - f(\cdot, X_-)\Big] \star (\mu - \nu).$$

(2) The FS decomposition (H_0, ξ^H, L^H) of H is given by

$$\xi^H = \Sigma^{inv} \left[c f_x(\cdot, X_-) + \int_{\mathbb{R}^d} x[f(\cdot, X_- + x) - f(\cdot, X_-)]K(dx) \right], \quad (7.38)$$

$$L^H = \widetilde{V}^H - \widetilde{V}_0^H - \xi^H \cdot X. \quad (7.39)$$

Proof. The proof of the first assertion is obvious from the previous calculation, while the second assertion is an immediate application of Theorem 2.2.23 and the fact that the quadruplet of $\widetilde{L}$ through Theorem 3.5.4 is

$$(f_x(\cdot, X_-), [f(\cdot, X_- + x) - f(\cdot, X_-)], 0, 0)\,.$$

□

We apply (7.38) to the setting of Riesner (2006a) and take into account that we work here with $X = \sigma X_- \cdot (b \cdot t + cW_t + \widetilde{M}_t)$. We refer to Kallsen and Shiryaev (2002) for more details concerning a process of the form $\Theta \cdot X$. Inserting the characteristic c from Chapter 3, which equals $\frac{d\langle X^c, X^c\rangle}{dt} = c^2\sigma^2 X_-^2$ and $F(t, S_{t-}, u)$ instead of $f(t, X_t)$ in (7.38), we obtain:

$$\begin{aligned}
\xi_t^H =& \frac{1}{\sigma_t^2 X_{t-}^2 (c^2 + \int x^2 \nu(dx))} [c^2\sigma^2 X_{t-}^2 F_x(t, S_{t-}, u) + \int_{\mathbb{R}} J(t, y, u) x \sigma_t X_{t-} \nu(dx)] \\
=& \frac{F_x(t, S_{t-}, u) c^2}{c^2 + \int_{\mathbb{R}} x^2 \nu(dx)} + \frac{\int_{\mathbb{R}} J(t, y, u) x \nu(dx)}{\sigma_t X_{t-}(c^2 + \int x^2 \nu(dx))},
\end{aligned}$$

which equals (7.28).
We remark that in this simplified setting of QLC processes, the number of risky assets also follows from

$$\xi^H = \frac{d\langle \widetilde{K}, M\rangle}{d\langle M\rangle},$$

see (3.12). Secondly we check (3.21), which expresses the difference between the optimal number of risky assets of the GKW decomposition under the MMM and the optimal number of risky assets from the Föllmer-Schweizer decomposition. Formula (3.21) applied to our setting, namely $-\lambda = \frac{G}{\sigma X_-}$ gives

$$\widetilde{\Phi} = \frac{1}{\sigma^2 X_-^2 (c^2 + \int x^2 \nu(dx))} \int \sigma X_- x \widetilde{f}(x) G x \nu(dx)$$

with

$$\widetilde{f}(x) = \varsigma^{(2)}(s, y, u) = J(t, y, u) - y\sigma_s X_{s-}\xi(s, u)$$

because the local martingale orthogonal to X under the MMM is given by $K(t, u)$ (7.12):

$$K(t, u) = \int_0^t \varsigma^{(1)}(s, u) dW_s^{\widetilde{Q}} + \int_0^t \int_{\mathbb{R}} \varsigma^{(2)}(s, y, u) \widetilde{M}^{\widetilde{Q}}(ds, dy)$$

with $\varsigma^{(2)}(s, y, u)$ given in (7.14). Therefore

$$\begin{aligned}
&\xi(t, u) - \phi(t, u) = \widetilde{\Phi}_t \\
&= \frac{1}{\sigma_t X_{t-}\kappa} \int_{\mathbb{R}} (J(t, x, u) - x\sigma_t X_{t-}\xi(t, u)) G_t x^2 \nu(dx) \\
&= -\frac{1}{\kappa} \int_{\mathbb{R}} x^3 G_t \nu(dx) \xi(t, u) + \frac{1}{\kappa \sigma_t X_{t-}} \int_{\mathbb{R}} x^2 J(t, x, u) G_t \nu(dx).
\end{aligned}$$

We obtain the same result as derived in (7.37) and in this way we checked our results obtained by using Colwell and Elliott (1993). Furthermore we also showed how Chapter 3 can be very useful to determine the LRM hedging strategy in practice.

Derivatives are financial weapons of mass destruction.

Warren Buffett (1930-)

8 Hedging in the extended Heath-Jarrow-Morton model

8.1 Introduction

Interest rate derivatives are important tools for the risk-management of insurance companies. Life insurance companies are for example exposed to liabilities which will mature in at most sixty years, while the bonds offered on the market usually have maturities ranging from one to thirty years. Therefore interest rate derivatives are useful to manage the risk related with these life insurance contracts. Frequently traded instruments are e.g. swap contracts, where two parties agree on exchanging different types of interest rates, typically a floating rate against a fixed rate. Options on these swaps are called swaptions. We calculate in this chapter the price of swaptions and look at different hedging strategies when the underlying default-free zero-coupon bonds are used for hedging. When we mention zero-coupon bonds later on in this chapter, then we always mean the default-free zero-coupon bonds.
Nowadays models for the term structure of interest rates that are driven by the Brownian motion are widely used in practice. However serious shortcomings of those models, in particular concerning the smile effect, are well known. Therefore an extension of the Heath-Jarrow-Morton model developed in Heath et al.

(1992) with a time-inhomogeneous Lévy process as driving process was introduced in Eberlein and Kluge (2006). In this model the instantaneous forward rate is given by a time-inhomogeneous Lévy process. Kluge (2005) showed that such a model allows to reproduce the so-called smile surface.
Most of the literature about interest rate derivatives, see Section 8.2, concentrates on models driven by the Brownian motion and therefore assumes continuity of the price processes. Within the framework of the Lévy driven Heath-Jarrow-Morton model, pricing formulas based on Fourier transforms are known for the most liquid interest rate derivatives, namely caps, floors and swaptions, whereas hedging strategies have not been discussed yet.
The goal of this chapter is threefold. Besides the pricing of the interest rate derivatives, which can also be found in e.g. Eberlein and Kluge (2006), we focus on the hedging with zero-coupon bonds. Our contribution on the pricing side is by giving a compact representation by using the Jamshidian decomposition. Finally, we give a numerical implementation and results.
A nice property of the one-dimensional model that we consider is that it has a unique martingale measure, see Eberlein et al. (2005). According to the equivalence between uniqueness of the martingale measure and completeness of the market, a perfect hedge for any product in this market is possible. Unfortunately, for a perfect hedge infinitely many bonds should be available on the market. This is obviously unrealistic. The market does only provide a strip of bonds with different maturities and a trader can only invest in a small number of different products.
Hence although in our underlying model a unique martingale measure exists, it is only theoretically complete and for the development of hedging strategies one has to use the concepts that are designed for incomplete markets.
As hedging strategy we will examine the delta-hedge and a quadratic hedging strategy, namely the mean-variance hedging. A delta-hedging strategy makes the portfolio risk-neutral for changes in the underlying and is often also made self-financing by investments in the riskless asset. Unfortunately, in the Heath-Jarrow-Morton market we do not really have a real riskless asset and therefore a priori we will calculate the delta-neutral hedging strategy without assuming that the strategy is also self-financing. In this way we avoid the use of a riskless asset. On the other hand the zero-coupon bond with maturity equal to the maturity of the swaption will not contain any risk at maturity of the swaption and can therefore be seen as replacement of the riskless asset. Hence also the self-financing delta-hedge will be formulated using an arbitrary zero-coupon bond and the zero-coupon bond with maturity equal to the maturity of the swaption.
We will show that it is sufficient to hedge with one bond to make the product

neutral for changes in all bonds available on the market. In a delta-hedge one constructs a portfolio such that by the investments Δ in the underlying $B(\cdot, T_j)$:

$$\frac{\partial(-\mathrm{PS} + \Delta B(\cdot, T_j))}{\partial B(\cdot, T_j)} = 0,$$

where PS stands for the price of the payer swaption, which is the interest rate derivative we price and hedge.

Under the assumption that we work under the forward martingale measure linked with the settlement date T_0, we know that all the zero-coupon bonds discounted by the numéraire $B(\cdot, T_0)$ are martingales. As a starting point we derive the mean-variance hedging strategy related to the forward measure $\mathbb{P}_{T_0}$. In other words we calculate the self-financing strategy whose quadratic distance under $\mathbb{P}_{T_0}$ to the payoff of the swaption at maturity is minimal. This strategy follows from the Galtchouk-Kunita-Watanabe decomposition, see Chapter 5. Of course one could ask why we determine the hedging strategy under the measure $\mathbb{P}_{T_0}$ and not under the measure $\mathbb{P}$, from which we started. The best would be to find the strategy under the original measure, but what is the original measure in this model? This is ongoing research.

We only determined the mean-variance hedging strategy, because this strategy is self-financing. Hence, as we show in Section 8.3 the number of non-discounted zero-coupon bonds we need to invest equals the discounted number. This one-to-one correspondence does not hold for the locally risk-minimizing hedging strategy. The latter will be applied in Chapter 9 in a semimartingale framework, under the assumption that the interest rate equals zero. Hence in that setting the discounted prices equals the non-discounted ones.

This chapter is based on Glau et al. (2010a,b). We start with a short overview of the literature concerning hedging in the interest rate derivatives market. Next we show explicitly that for the mean-variance hedging strategy the optimal number of non-discounted assets equals the discounted ones.

In Section 8.4 we introduce in detail the model we use and describe the assumptions under which we work. The Jamshidian decomposition and the Fourier transformation, also explained in Section 8.4, are crucial steps in our calculation. The price and the hedging strategies for the payer swaption are determined in Section 8.5. The numerical results, containing a comparison of the total cost and its variability of the delta-hedge versus those of the mean-variance hedge, are given in Section 8.6.

8.2 Literature concerning hedging in the interest rate derivatives market

In this section we give an overview of the papers we are aware of concerning the hedging of products in the interest rate derivatives market.
Most of these papers concentrate on the delta-hedge in the original Heath-Jarrow-Morton model which is a continuous setting without jumps.

- Brace et al. (2001) determine analytical formulas containing the approximate prices and delta-hedging strategies for swaptions in the LIBOR[1] market model.
- Dun et al. (2001) assume first that the forward swap rate is lognormal under the forward swap measure and hedge the swaption with a combination of the underlying swap and a portfolio of zero-coupon bonds or by only using zero-coupon bonds. In a second case they hedge swaptions in the lognormal forward LIBOR model as done in Brace et al. (2001) and by numerical evaluation of the partial derivative.
- Henrard (2003) states the delta-hedge for a bond option and a swaption in the Heath-Jarrow-Morton model.
- Akume et al. (2003) price and hedge swaptions in the Black model, see Musiela and Rutkowski (2004). They determine not only the delta-hedge, but by adding another swaption they also obtain a delta- and gamma-neutral strategy.
- Piterbarg (2003, 2004) compute the deltas, the prices and the delta-hedges for different callable LIBOR exotics in the forward LIBOR model.
- Barbarin (2008a) determines the risk-minimizing hedging strategy for pure endowment and annuity portfolios when longevity bonds are modeled in the Heath-Jarrow-Morton framework.
- Driessen et al. (2003) compare the pricing and the delta-hedging of caps and swaptions for different term structure models all driven by a continuous underlying.

Up till now only the hedging of continuous models is discussed and they rarely discuss other hedging methods than delta-hedging. An exception is the PhD-

[1] Londen Interbank Offered Rate

thesis and related publications of Biagini (2001). In the original Heath-Jarrow-Morton framework she determined not only the mean-variance hedging strategy for an European option with maturity T_0, but also the locally risk-minimizing hedging strategy for a call option on a T_1-bond. To end she reformulated the mean-variance hedging strategy to hedge also assets with dividends. Using this extension to dividends she determined the hedging strategy for futures contracts. Her research remained restricted to the development of the formulas, no implementation was carried out.
We also want to point out the growing interest in literature for numerical comparisons between the delta-hedge and the quadratic hedging strategies, see Altmann et al. (2008), Denkl et al. (2009) and Brodén and Tankov (2010).

8.3 The self-financing property

We first look at the locally risk-minimizing hedging strategy. To obtain this strategy the square of the changes in the so-called discounted cost process is minimized. The terminology discounted cost process is somewhat misleading because the discounted cost process is not the discounted version of the non-discounted cost process, but it is the cost process when we hedge with the discounted assets. The questions that arise are: which criterion are we minimizing when we go back to the non-discounted assets and does this still make sense?
Elementary calculations show that the difference between the real discounted cost process $\overline{C}$ and the so-called discounted cost process $\widetilde{C}$ has dynamics with terms only depending on the numéraire $\widetilde{S}^0$. This means that $d(\overline{C}-\widetilde{C}) = \ldots d\widetilde{S}^0$.
Furthermore knowing the optimal number of discounted risky assets does not help us to determine the amount we need to invest in the non-discounted assets for the locally risk-minimizing hedging strategy, because the riskless asset used for discounting is stochastic. For the mean-variance hedging strategy we will show that it makes sense to invest in the non-discounted what was calculated for the discounted amount. This is done without using the fact that the numéraire should have special properties like having finite variation or being continuous, but by using the self-financing property of the mean-variance hedging strategy.
This is also the reason why we only determine the mean-variance hedging strategy in this chapter, because we do not know how to interpret the optimal number of discounted assets in terms of the non-discounted ones for the locally

risk-minimizing hedging strategy.

Proposition 8.3.1. *Let us denote by V the non-discounted self-financing portfolio with value at time t:*

$$V_t = \phi_t^0 S_t^0 + \phi_t^1 S_t^1,$$

where S^0 is the asset which will be used as numéraire and where S^1 can be a vector. Then the process of the discounted portfolio $\widetilde{V}_t = \frac{V_t}{S_t^0}$ is given by

$$\widetilde{V}_t = \widetilde{V}_0 + (\phi^1 \cdot \widetilde{S}^1)_t, \tag{8.1}$$

with $\widetilde{S}_t^1 = \frac{S_t^1}{S_t^0}$ the discounted asset.

Proof. Since V is self-financing its dynamics are

$$dV_t = \phi_t^0 dS_t^0 + \phi_t^1 dS_t^1, \tag{8.2}$$

this means that changes in the portfolio are only due to changes in the underlyings. The derivative of the discounted portfolio (8.1) can be written as

$$\begin{aligned} d\widetilde{V}_t =& d\frac{V_t}{S_t^0} = \frac{1}{S_{t-}^0} dV_t + V_{t-} d(\frac{1}{S_{t-}^0}) + d[V, \frac{1}{S^0}]_t \\ =& \frac{1}{S_{t-}^0}\phi_t^0 dS_t^0 + \frac{1}{S_{t-}^0}\phi_t^1 dS_t^1 + \phi_t^0 S_{t-}^0 d(\frac{1}{S_{t-}^0}) \\ &+ \phi_t^1 S_{t-}^1 d(\frac{1}{S_{t-}^0}) + d[\phi^0 \cdot S^0, \frac{1}{S^0}]_t + d[\phi^1 \cdot S^1, \frac{1}{S^0}]_t, \end{aligned} \tag{8.3}$$

where we used the fact that V is self-financing (8.2) and the predictability of ϕ^0 and ϕ^1 in the last step. Furthermore, we know by Itô's formula that

$$\frac{1}{S_{t-}^0}\phi_t^0 dS_t^0 + \phi_t^0 S_{t-}^0 d(\frac{1}{S_t^0}) + \phi_t^0 d[S^0, \frac{1}{S^0}] = \phi_t^0 d(S_{t-}^0 \frac{1}{S_{t-}^0}) = \phi_t^0 d(1) = 0.$$

Therefore (8.3) reduces to

$$d\widetilde{V}_t = \frac{\phi_t^1}{S_{t-}^0} dS_t^1 + \phi_t^1 S_{t-}^1 d(\frac{1}{S_{t-}^0}) + d[\phi^1 \cdot S^1, \frac{1}{S^0}]_t. \tag{8.4}$$

The right-hand side equals the derivative:

$$\phi_t^1 d\widetilde{S}_t^1 = \phi_t^1 d(\frac{S_t^1}{S_t^0}) = \frac{\phi_t^1}{S_{t-}^0} dS_t^1 + \phi_t^1 S_{t-}^1 d(\frac{1}{S_{t-}^0}) + d[\phi^1 \cdot S^1, \frac{1}{S^0}]_t. \tag{8.5}$$

The integral form of (8.4)-(8.5) is exactly (8.1), which proves our claim. □

8.4 Setting

8.4.1 The Lévy driven Heath-Jarrow-Morton model

We assume that the stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is given. There are different ways to model the interest rate derivatives market e.g. the Heath-Jarrow-Morton model, the Libor model and the swap market model. We will work with an extension of the Heath-Jarrow-Morton model. We give a short explanation of the setting we will use. For more details we refer to the following surveys which contain the most important facts concerning the interest rate derivatives market: Musiela and Rutkowski (2004), Björk (2004), Brigo and Mercurio (2001) and Pelsser (2000). These books concentrate on describing the models and on pricing interest rate products but rarely look at the hedging strategies.
We follow mainly the notation of Musiela and Rutkowski (2004) to introduce the original Heath-Jarrow-Morton model.
Denote by $T^* > 0$ the fixed horizon date, this means that trading will only happen during the interval $[0, T^*]$. The value at time t of a zero-coupon bond paying 1 unit at maturity T is given by $B(t,T)$ and, of course, $B(T,T)$ equals 1. The forward interest rate at date $t \leq T$, $f(t,T)$, is the instantaneous risk-free interest rate for borrowing or lending at time T seen from time t. When we have a family of forward interest rates $f(t,T)$, we easily derive the prices of the zero-coupon bonds:

$$B(t,T) = \exp\left(-\int_t^T f(t,u)du\right).$$

In general the converse does not hold except when the family of bond prices is sufficiently smooth with respect to the maturity T:

$$f(t,T) = -\frac{\partial \ln B(t,T)}{\partial T}.$$

In case of the one-to-one correspondence, it is sufficient to model either the zero-coupon bond prices or the forward interest rates. Further, the short-term interest rate r_t is described by $f(t,t)$ and $B_t = \exp(\int_0^t r_u du)$ represents the savings account.
The Heath-Jarrow-Morton model was introduced in Heath et al. (1992). Assume $\mathbb{F}$ is the natural filtration of the d-dimensional Brownian motion W under $\mathbb{P}$, which is independent of the time to maturity T, then the dynamics of the forward interest rates are given by

$$df(t,T) = \alpha(t,T)dt + \sigma'(t,T)dW_t,$$

with α and σ adapted stochastic processes in $\mathbb{R}$, respectively $\mathbb{R}^d$ and with $'$ denoting the transpose. Hence the dynamics of the zero-coupon bonds are described by

$$dB(t,T) = B(t,T)(a(t,T)dt - \sigma^{*\prime}(t,T)dW_t), \tag{8.6}$$

with

$$\begin{aligned} a(t,T) =& f(t,t) - \alpha^*(t,T) + \frac{1}{2}|\sigma^*(t,T)|^2 \\ \alpha^*(t,T) =& \int_t^T \alpha(t,u)du \\ \sigma^*(t,T) =& \int_t^T \sigma(t,u)du. \end{aligned}$$

Remark that formula (8.6) guarantees the positiveness of the zero-coupon bonds.

Definition 8.4.1. The spot martingale measure $\mathbb{P}^*$ is the measure under which the discounted zero-coupon bonds are martingales.

The dynamics of the bonds and the forward rates under $\mathbb{P}^*$ are

$$\begin{aligned} dB(t,T) =& B(t,T)(r_t dt - \sigma^{*\prime}(t,T)dW_t^*) \\ df(t,T) =& \sigma'(t,T)\sigma^*(t,T)dt + \sigma'(t,T)dW_t^*. \end{aligned}$$

The $\mathbb{P}^*$-Brownian motion W^* is defined as

$$W_t^* = W_t - \int_0^t \lambda_u du \quad \forall t \in [0,T],$$

where λ is an adapted $\mathbb{R}^d$-valued process such that

$$E^{\mathbb{P}}[\mathcal{E}(\lambda \cdot W)_{T^*}] = 1.$$

Furthermore for any maturity $T \leq T^*$:

$$\alpha^*(t,T) = \frac{1}{2}|\sigma^*(t,T)|^2 - \sigma^{*'}(t,T)\lambda_t.$$

We recall that we will consider an extension of this model namely the Lévy driven Heath-Jarrow-Morton model. This extension introduced by Eberlein and Raible (1999) replaces the Brownian motion in the original model by a Lévy process. The use of time-inhomogeneous Lévy processes instead of time-homogeneous Lévy processes is due to Eberlein and Kluge (2006). The advantage of using these extensions is the better accuracy that can be obtained when the calibration is performed.

Denote by $\mathbb{F}$ the natural filtration generated by the d-dimensional time-(in)homogeneous Lévy process L with characteristics (b_s, c_s, F_s). Then, the dynamics of the forward interest rates and the zero-coupon bonds under the measure $\mathbb{P}$ are given by

$$\begin{aligned} df(t,T) =& \alpha(t,T)dt - \sigma'(t,T)dL_t \\ B(t,T) =& B(0,T) \exp\left(\int_0^t (r_s - A(s,T))ds + \int_0^t \Sigma'(s,T)dL_s\right), \end{aligned}$$

with α, σ adapted stochastic processes in $\mathbb{R}$, respectively $\mathbb{R}^d$, and

$$A(s,T) = \int_{s\wedge T}^T \alpha(s,u)du \tag{8.7}$$

$$\Sigma(s,T) = \int_{s\wedge T}^T \sigma(s,u)du, \tag{8.8}$$

where $s \wedge T = \min(s,T)$ and $s \in [0,T^*]$.

In this chapter, we will restrict ourselves to a one-dimensional time-inhomogeneous Lévy process L with characteristics (b_s, c_s, F_s) and such that

$$\int_0^{T^*} \left(|b_s| + |c_s| + \int_{\mathbb{R}} (x^2 \wedge 1)F_s(dx)\right) ds < \infty.$$

We assume the following integrability condition on the measures F_s to ensure in particular that L is an exponential special semimartingale:

Assumptions 8.4.2 (EM). *There are constants $M, \epsilon > 0$ such that for every $u \in [-(1+\epsilon)M, (1+\epsilon)M]$:*

$$\int_0^{T^*} \int_{\{|x|>1\}} \exp(ux) F_s(dx) ds < \infty.$$

The law of L_t is characterized by the characteristic function

$$E[e^{izL_t}] = e^{\int_0^t \theta_s(iz)ds}, \quad \forall t \in [0, T^*] \tag{8.9}$$

with θ_s the cumulant associated with L by the Lévy-Khintchine triplet (b_s, c_s, F_s):

$$\theta_s(z) := b_s z + \frac{1}{2} c_s z^2 + \int_{\mathbb{R}} (e^{xz} - 1 - xz) F_s(dx). \tag{8.10}$$

Remark that due to Assumption 8.4.2 we do not need to truncate the large jumps in the cumulant as was necessary in (2.44).
For the applications we rewrite the dynamics of the zero-coupon bonds in the following way:

$$B(t,T) = \frac{B(0,T)}{B(0,t)} \exp\left(-\int_0^t A(s,t,T) ds + \int_0^t \Sigma(s,t,T) dL_s \right), \tag{8.11}$$

with $A(s,t,T) = A(s,T) - A(s,t)$ and $\Sigma(s,t,T) = \Sigma(s,T) - \Sigma(s,t)$.
The discounted bond prices are given by

$$\frac{B(t,T)}{B_t} = B(0,T) \exp\left(-\int_0^t A(s,T) ds + \int_0^t \Sigma(s,T) dL_s \right). \tag{8.12}$$

In the one-dimensional Lévy driven Heath-Jarrow-Morton model, there is a unique martingale measure $\mathbb{P}^*$, see Eberlein et al. (2005). If

$$A(s,T) := \theta_s(\Sigma(s,T)) \quad \forall T \in [0, T^*], \tag{8.13}$$

then the discounted bond prices are martingales and hence the model is described under the unique martingale measure $\mathbb{P}^*$, also called the spot martingale measure.
Concerning the volatility structure we make the following additional assumptions.

Assumptions 8.4.3 (DET)**.** *The volatility structure σ is bounded and deterministic. Furthermore for $0 \leq s$ and $T \leq T^*$, we assume that*

$$0 \leq \Sigma(s,T) \leq M' < M,$$

with M the constant defined in Assumption 8.4.2 and Σ given by (8.8).

To price the products we make an additional assumption, which is often met in practice, on the volatility σ:

Assumptions 8.4.4 (VOL)**.** *For all $T \in [0, T^*]$ we assume that $\sigma(\cdot\,, T) \not\equiv 0$ and*

$$\sigma(s,T) = \sigma_1(s)\sigma_2(T) \quad 0 \leq s \leq T,$$

where $\sigma_1 : [0,T^*] \to \mathbb{R}^+$ and $\sigma_2 : [0,T^*] \to \mathbb{R}^+$ are continuously differentiable. Furthermore we assume that $\inf_{s\in[0,T^*]} \sigma_1(s) \geq \underline{\sigma}_1 > 0$.

Under these assumptions the short rate r_t is Markovian, see Eberlein and Raible (1999).
An important measure for pricing and hedging interest rate derivative products is the forward martingale measure.

Definition 8.4.5. The forward measure is linked with a settlement date T, such that the forward price of any financial asset (in our case any zero-coupon bond) is a (local) martingale. The forward price at time t of an asset S is given by $S_t/B(t,T)$.

The change of measure from the spot martingale measure $\mathbb{P}^*$, which equals $\mathbb{P}$ in our setting, to the forward martingale measure linked with the settlement date T is according to (8.12)

$$\frac{d\mathbb{P}_T}{d\mathbb{P}} = \frac{1}{B_T B(0,T)} = \exp\left(-\int_0^T A(s,T)ds + \int_0^T \Sigma(s,T)dL_s\right)$$

and

$$\left.\frac{d\mathbb{P}_T}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \frac{B(t,T)}{B_t B(0,T)} = \exp\left(-\int_0^t A(s,T)ds + \int_0^t \Sigma(s,T)dL_s\right). \tag{8.14}$$

From Proposition 10 and Lemma 11 of Eberlein and Kluge (2006) we conclude that L is also a time-inhomogeneous Lévy process under the forward measure $\mathbb{P}_T$ and that L is again special under this measure. The characteristics

$(b_s^{\mathbb{P}_T}, c_s^{\mathbb{P}_T}, F_s^{\mathbb{P}_T})$ of L under $\mathbb{P}_T$ can be expressed in terms of those under $\mathbb{P}$:

$$\begin{aligned} b_s^{\mathbb{P}_T} &= b_s + c_s\Sigma(s,T) + \int_{\mathbb{R}} x(e^{\Sigma(s,T)x} - 1)F_s(dx), \\ c_s^{\mathbb{P}_T} &= c_s, \\ F_s^{\mathbb{P}_T}(dx) &= e^{\Sigma(s,T)x}F_s(dx), \end{aligned} \tag{8.15}$$

where we inserted the truncation function $h(x) = x$ see Proposition 10 of Eberlein and Kluge (2006).
The following proposition also due to Eberlein and Kluge (2006) is very useful for option pricing.

Proposition 8.4.6 (Eberlein and Kluge (2006) Proposition 8)**.** *Suppose that* $f : \mathbb{R}_+ \to \mathbb{C}^d$ *is a continuous function such that for all* $i \in \{1, \ldots, d\}$ *and* $x \in \mathbb{R}_+$ *the absolute value of the real part of* $f^i(x)$ *is smaller than or equal to* M*, then*

$$E[\exp(\int_t^T f(s)dL_s)] = \exp\int_t^T \theta_s(f(s))ds.$$

Note that f only depends on the time and hence should also be a deterministic function.

8.4.2 Jamshidian decomposition

The Jamshidian decomposition was introduced for an option written on a coupon-bearing bond expressed as a linear combination of zero-coupon bonds when the interest rate is modeled by a Vasiček model in Jamshidian (1989). In Brigo and Mercurio (2001) this decomposition was generalized to allow also for other models.
Jamshidian constructs a closed-form expression for the price of an European option on a coupon-bearing bond in terms of the European option prices on the individual zero-coupon bonds. He denotes by $P(r,t,s)$ the price at time t of a pure discount bond maturing at time s, given that $r(t) = r$ and $R_{r,t,s}$ is a normal random variable. To arrive at his results he proved for the option payoff at maturity T that

$$\max\{0, \sum a_j P(R_{r,t,T}, T, s_j) - K\} = \sum a_j \max\{0, P(R_{r,t,T}, T, s_j) - K_j\},$$

where $K_j = P(r^*, T, s_j)$ and r^* is the solution to the equation

$$\sum a_j P(r^*, T, s_j) = K.$$

According to Brigo and Mercurio (2001) this still holds for other short rate models as long as the price of the zero-coupon bond is a decreasing function of the interest rate. An extension of this property is obtained by Annaert et al. (2007) using only the more general concept of comonotonicity, see Kaas et al. (2000), for the zero-coupon bonds.

8.4.3 Fourier transformation

To price and hedge the interest rate derivatives we frequently apply Fourier transforms which allow to calculate derivative prices fast. We will concentrate on cases where the payoff functions are continuous and hence we repeat Theorem 2.2 of Eberlein et al. (2009). An analogous theorem for a discontinuous payoff function is Theorem 2.7 of Eberlein et al. (2009).
We denote by f the payoff function and by S the asset price process, modeled as an exponential semimartingale process: $S_t = S_0 e^{H_t}$, $0 \leq t \leq T$. Let X be the underlying process of the option, which possibly depends on the full history of H, namely:

$$X_t := \Psi(H_s, 0 \leq s \leq t) \quad t \in [0, T]$$

and where Ψ is a measurable function. Furthermore in Eberlein et al. (2009) they only consider those options where the path-dependence on the asset price process S can be incorporated into the driving process H. Such as options on the supremum or the trivial European vanilla options.

Theorem 8.4.7 (See Eberlein et al. (2009)). *If the following conditions are satisfied:*

(C1) the dampened function $g = e^{-Rx} f(x)$ is a bounded, continuous function in $L^1(\mathbb{R})$ (i.e. $\int_{\mathbb{R}} |g(x)| dx < +\infty$).

(C2) The moment generating function $M_{X_T}(R)$ of the random variable X_T exists.

(C3) The (extended) Fourier transform $\hat{g}$ belongs to $L^1(\mathbb{R})$,

then the time zero price function $\mathbb{V}_f(X; s = -\log S_0)$ *of an option on* $S = (S_t)_{0\le t\le T}$ *with payoff function* f *is given by*

$$\mathbb{V}_f(X; s = -\log S_0) = \frac{e^{-Rs}}{2\pi} \int_{\mathbb{R}} e^{ius} \varphi_{X_T}(-u - iR) \hat{f}(u + iR) du,$$

with φ_{X_T} *the (extended) characteristic function of the random variable* X_T.

As remarked in Eberlein et al. (2009) Theorem 8.4.7 still holds when the first and the third condition are replaced by

(C1') The dampened function g belongs to $L^1(\mathbb{R})$.

(C3') $\widehat{e^{Rx} P_{X_T}}$ belongs to $L^1(\mathbb{R})$, with P_{X_T} denoting the law of X_T.

8.5 Payer swaption

A (plain vanilla) interest rate swap is a contract to exchange a fixed interest rate against a floating reference rate, like the Libor. Both rates are based on the same notional amount and for the same period of time. In the case of a payer swap the investor pays the fixed rate and receives the floating rate. In a usual swap contract the fixed rate is chosen such that the contract is worth zero at the initial date.
A forward swap is an agreement to enter into a swap at a future date T_0 with a pre-specified fixed rate κ, while a payer swaption gives the owner the right to enter the forward payer swap at T_0. Musiela and Rutkowski (2004) showed that the payer swaption can be seen as a put option with strike price 1 on a coupon-bearing bond. Therefore we can write the payer swaption's payoff as

$$(1 - \sum_{j=1}^{n} c_j B(T_0, T_j))^+,$$

where $T_1 < T_2 < \ldots < T_n$ are the payment dates of the swap with $T_1 > T_0$. We denote the length of the accrual periods $[T_{j-1}, T_j]$, $j = 1, \ldots, n$ by $\delta_j := T_j - T_{j-1}$. The coupons c_i equal $c_i = \kappa\delta_i$ for $i = 1, \ldots, n-1$ and $c_n = 1 + \kappa\delta_n$ where κ is the fixed interest rate of the swap.

8.5.1 Pricing of the payer swaption

The price of the receiver swaption in the Lévy driven Heath-Jarrow-Morton framework is determined in Eberlein and Kluge (2006). The pricing of the payer swaption can be done in a very similar way. We present the derivation for the payer swaption in a slightly different way to make the application of the so called Jamshidian trick more visible. This allows for another interpretation of the payer swaption, namely it can be seen as a weighted sum of put options with different strikes on bonds with different maturities. This was in fact already noticed by Annaert et al. (2007) in a continuous setting for a general interest rate model were the zero-coupon bond prices are comonotonic.
The fair price of the payer swaption is given by

$$\mathrm{PS}_t = B_t E[\frac{1}{B_{T_0}}(1-\sum_{j=1}^{n} c_j B(T_0,T_j))^+|\mathcal{F}_t] \qquad t \in [0,T_0],$$

where the expectation is taken under the risk-neutral measure $\mathbb{P}$. We change to the forward measure $\mathbb{P}_{T_0}$ eliminating the instantaneous interest rate B_{T_0} under the expectation in this way:

$$\begin{aligned}\mathrm{PS}_t =& B(t,T_0)E^{\mathbb{P}_{T_0}}[(1-\sum_{j=1}^{n} c_j B(T_0,T_j))^+|\mathcal{F}_t]\\ =& B(t,T_0)E^{\mathbb{P}_{T_0}}[(1-\sum_{j=1}^{n} c_j \widetilde{D}_{T_0}^{T_j} e^{\widetilde{\Sigma}_{T_0}^{T_j} X_{T_0}})^+|\mathcal{F}_t]\end{aligned} \tag{8.16}$$

with according to (8.8), (8.11), (8.13) and Assumption 8.4.4

$$\begin{aligned}\widetilde{D}_{T_0}^{T_j} =& \frac{B(0,T_j)}{B(0,T_0)} \exp\left(\int_0^{T_0} [\theta_s(\Sigma(s,T_0)) - \theta_s(\Sigma(s,T_j))]\, ds\right)\\ \widetilde{\Sigma}_{T_0}^{T_j} =& \int_{T_0}^{T_j} \sigma_2(u)du \quad \text{and} \quad X_{T_0} = \int_0^{T_0} \sigma_1(s) dL_s.\end{aligned} \tag{8.17}$$

We denote by $g(s,t,x)$ the function

$$g(s,t,x) = \widetilde{D}_s^t e^{\widetilde{\Sigma}_s^t x} \qquad \forall 0 \le s \le t \le T^*, \tag{8.18}$$

then the price at time s of the zero-coupon bond with maturity t is given by

$$g(s,t,X_s) = B(s,t) \qquad \forall 0 \le s \le t \le T^*. \tag{8.19}$$

Notice that under Assumption 8.4.4 the functions $x \mapsto g(T_0, T_j, x)$ are non-decreasing functions for $j = 1, \dots, n$. This allows us to apply the decomposition of Jamshidian to (8.16), i.e.

$$\mathrm{PS}_t = B(t, T_0) \sum_{j=1}^{n} c_j E^{\mathbb{P}_{T_0}}[(b_j - g(T_0, T_j, X_{T_0}))^+ | \mathcal{F}_t] \tag{8.20}$$

with b_j such that

$$\widetilde{D}_{T_0}^{T_j} e^{\widetilde{\Sigma}_{T_0}^{T_j} z^*} = g(T_0, T_j, z^*) = b_j \tag{8.21}$$

and z^* is the solution to the equation

$$\sum_{j=1}^{n} c_j g(T_0, T_j, z^*) = 1. \tag{8.22}$$

Let us insert the expression for $g(T_0, T_j, X_{T_0})$ in (8.20) to obtain the representation of the price as the weighted sum of put options on bonds,

$$\mathrm{PS}_t = B(t, T_0) \sum_{j=1}^{n} c_j E^{\mathbb{P}_{T_0}}[(b_j - B(T_0, T_j))^+ | \mathcal{F}_t]. \tag{8.23}$$

We will prove the following theorem concerning the price of the payer swaption by applying Theorem 8.4.7.

Theorem 8.5.1. *Under the Assumptions 8.4.2, 8.4.3, 8.4.4 and if $|\sigma_1| < \overline{\sigma}_1$ for a certain $\overline{\sigma}_1 \in \mathbb{R}$, the price at time t of a forward payer swaption is given by a weighted sum of put options on bonds*

$$\begin{aligned} PS_t =& B(t, T_0) \sum_{j=1}^{n} c_j E^{\mathbb{P}_{T_0}}[(b_j - B(T_0, T_j))^+ | \mathcal{F}_t] \\ =& B(t, T_0) \sum_{j=1}^{n} c_j \frac{e^{-RX_t}}{2\pi} \int_{\mathbb{R}} e^{iuX_t} \varphi^{\mathbb{P}_{T_0}}_{X_{T_0} - X_t}(u + iR) \hat{v}^j(-u - iR) du, \end{aligned} \tag{8.24}$$

where $\varphi^{\mathbb{P}_{T_0}}_{X_{T_0} - X_t}$ is the characteristic function of $X_{T_0} - X_t$ under the measure $\mathbb{P}_{T_0}$ given by

$$\varphi^{\mathbb{P}_{T_0}}_{X_{T_0} - X_t}(z) = \exp \int_t^{T_0} \theta_s^{\mathbb{P}_{T_0}}(iz\sigma_1(s)) ds \tag{8.25}$$

$$= \exp \int_t^{T_0} [\theta_s(\Sigma(s, T_0) + iz\sigma_1(s)) - \theta_s(\Sigma(s, T_0))] ds \tag{8.26}$$

and where

$$\hat{v}^j(-u-iR) = \frac{b_j e^{(-iu+R)z^*\widetilde{\Sigma}_{T_0}^{T_j}}}{(-iu+R)(-iu+\widetilde{\Sigma}_{T_0}^{T_j}+R)}$$

for R in $]0, \frac{M}{\sigma_1}]$ and b_j such that $g(T_0, T_j, z^*) = b_j$, where z^* is the solution to the equation $\sum_{j=1}^n c_j g(T_0, T_j, z^*) = 1$ with g the non-decreasing function defined as $g(s,t,x) = \widetilde{D}_s^t e^{\widetilde{\Sigma}_s^t x}$.

Proof. We start with rewriting the price of the payer swaption as

$$\begin{aligned}\text{PS}_t =& B(t,T_0)\sum_{j=1}^n c_j E^{\mathbb{P}_{T_0}}[v^j(X_{T_0})|\mathcal{F}_t] \\ =& B(t,T_0)\sum_{j=1}^n c_j E^{\mathbb{P}_{T_0}}[v^j(X_{T_0}-X_t+y)]\,|_{y=X_t},\end{aligned} \tag{8.27}$$

with $v^j(x) := (b_j - \widetilde{D}_{T_0}^{T_j} e^{\widetilde{\Sigma}_{T_0}^{T_j} x})^+$. To apply a Fourier transformation on this price we first prove that the three alternative conditions of Theorem 8.4.7 are satisfied for $R \in]0, \frac{M}{\sigma_1}]$.

- The functions v^j, $j \in \{1, \ldots, n\}$, are not integrable over $\mathbb{R}$, but since $\widetilde{\Sigma}_{T_0}^{T_j} > 0$ we can define the dampened function $g(x) = e^{Rx}v^j(x)$, which is integrable for any $R > 0$ independent of j.
- Secondly, we need that the moment generating function of the random variable $X_{T_0} - X$ under the measure $\mathbb{P}_{T_0}$

  $$M_{X_{T_0}-X_t}^{\mathbb{P}_{T_0}}(-R) = E^{\mathbb{P}_{T_0}}[\exp(-R(X_{T_0}-X_t))]$$

  is finite for a certain $R > 0$. We wish to emphasize that the process $X_{T_0} - X$ is a time-inhomogeneous Lévy process under $\mathbb{P}$ as well as under $\mathbb{P}_{T_0}$. Hence, from the independence of the increments of X, we get that

  $$M_{X_{T_0}-X_t}^{\mathbb{P}_{T_0}}(z) = \frac{E^{\mathbb{P}_{T_0}}[e^{z(X_{T_0}-X_t)}]}{E^{\mathbb{P}_{T_0}}[e^{zX_t}]}E^{\mathbb{P}_{T_0}}[e^{zX_t}] = \frac{E^{\mathbb{P}_{T_0}}[e^{zX_{T_0}}]}{E^{\mathbb{P}_{T_0}}[e^{zX_t}]} \tag{8.28}$$

  is finite if $M_{X_{T_0}}^{\mathbb{P}_{T_0}}(z) = e^{\int_0^{T_0}\theta_s^{\mathbb{P}_{T_0}}(z\sigma_1(s))ds} < +\infty$, with cfr. (8.10)

  $$\theta_s^{\mathbb{P}_{T_0}}(x) = b_s^{\mathbb{P}_{T_0}}x + \frac{1}{2}c_s^{\mathbb{P}_{T_0}}x^2 + \int_{\mathbb{R}}(e^{xy}-1-xy)F_s^{\mathbb{P}_{T_0}}(dy), \tag{8.29}$$

the cumulant associated with L by the triplet $(b_s^{\mathbb{P}_{T_0}}, c_s^{\mathbb{P}_{T_0}}, F_s^{\mathbb{P}_{T_0}})$ under the measure $\mathbb{P}_{T_0}$ described in (8.15). Note that in the previous calculations we used Proposition 8.4.6 to say that $M_{X_{T_0}}^{\mathbb{P}_{T_0}}(z) = e^{\int_0^{T_0} \theta_s^{\mathbb{P}_{T_0}}(z\sigma_1(s))ds}$. Substituting iz for z in the moment generating function we get the characteristic function (8.25) of the process $X_{T_0} - X$ under the measure $\mathbb{P}_{T_0}$.
We conclude that $M_{X_{T_0}}^{\mathbb{P}_{T_0}}(z)$ is finite if and only if

$$\begin{aligned}&\int_0^{T_0} \int_{\{|y|>1\}} e^{z\sigma_1(s)y} F_s^{\mathbb{P}_{T_0}}(dy)ds \\ &= \int_0^{T_0} \int_{\{|y|>1\}} e^{(z\sigma_1(s)+\Sigma(s,T_0))y} F_s(dy)ds < +\infty. \end{aligned} \tag{8.30}$$

Due to Assumption 8.4.2, it is sufficient to choose R such that

$$-M \leq -R\sigma_1(s) + \Sigma(s, T_0) \leq M.$$

From Assumption 8.4.3, we know that $\Sigma(s, T_0) \leq M' < M$ and hence the inequality $-R\sigma_1(s) + \Sigma(s, T_0) \leq M$ is trivially satisfied. Under the additional assumption of the boundedness of σ_1, $|\sigma_1| < \overline{\sigma}_1 < +\infty$ and again using Assumption 8.4.3, we find the interval $]0, \frac{M}{\overline{\sigma}_1}]$ for R, indeed $-R\sigma_1(s) + \Sigma(s, T_0) \geq -R\sigma_1(s) \geq -R\overline{\sigma}_1 \geq -M$.

- To end, we check whether

$$\widehat{e^{-Rx} P}_{X_{T_0}-X_t}^{\mathbb{P}_{T_0}} \in L^1(\mathbb{R}),$$

where $\hat{\cdot}$ denotes the Fourier transform and $P_{X_{T_0}-X}^{\mathbb{P}_{T_0}}$ stands for the law of $X_{T_0} - X$ under the measure $\mathbb{P}_{T_0}$. Due to the exponential decay of $e^{-Rx} P_{X_{T_0}-X_t}^{\mathbb{P}_{T_0}}$, the integrability of the Fourier transform of this function is trivial.

Since the required conditions are satisfied, Theorem 8.4.7 implies that the price of the payer swaption can be rewritten as

$$\mathrm{PS}_t = B(t, T_0) \sum_{j=1}^n c_j \frac{e^{-RX_t}}{2\pi} \int_{\mathbb{R}} e^{iuX_t} \varphi_{X_{T_0}-X_t}^{\mathbb{P}_{T_0}}(u + iR)\hat{v}^j(-u - iR)du,$$

where $\varphi^{\mathbb{P}_{T_0}}_{X_{T_0}-X_t}$ is the characteristic function of $X_{T_0} - X_t$ under the measure $\mathbb{P}_{T_0}$. The expression $\hat{v}^j$ equals

$$
\begin{aligned}
&\hat{v}^j(-u - iR)\\
&= \hat{g}^j(-u) = \int_{\mathbb{R}} (b_j - \widetilde{D}^{T_j}_{T_0} e^{\widetilde{\Sigma}^{T_j}_{T_0} x})^+ e^{(-iu+R)x} dx\\
&= \int_{-\infty}^{z^*} (b_j - \widetilde{D}^{T_j}_{T_0} e^{\widetilde{\Sigma}^{T_j}_{T_0} x}) e^{(-iu+R)x} dx\\
&= b_j e^{(-iu+R)z^*} \int_{-\infty}^{0} e^{(-iu+R)x} dx - \widetilde{D}^{T_j}_{T_0} e^{(-iu+\widetilde{\Sigma}^{T_j}_{T_0}+R)z^*} \int_{-\infty}^{0} e^{(-iu+\widetilde{\Sigma}^{T_j}_{T_0}+R)x} dx\\
&= b_j e^{(-iu+R)z^*} \int_0^1 t^{-iu+R-1} dt - b_j e^{(-iu+R)z^*} \int_0^1 t^{-iu+\widetilde{\Sigma}^{T_j}_{T_0}+R-1} dt, \qquad (8.31)
\end{aligned}
$$

where we used in the second step that $(b_j - \widetilde{D}^{T_j}_{T_0} e^{\widetilde{\Sigma}^{T_j}_{T_0} x})$ equals zero for z^*, see (8.21)-(8.22), and the fact that the zero-coupon bonds are non-decreasing functions of x. In the third step x was replaced by $x + z^*$ and in the fourth t was substituted for e^x. The integrals of (8.31) can be evaluated using the well-known result about the Beta function, namely

$$
\int_0^1 t^{x-1}(1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},
$$

for complex x and y whose real parts are positive and where the Beta function is already replaced in terms of the Gamma function defined as $\Gamma(z+1) = z\Gamma(z)$ and $\Gamma(1) = 1$. Therefore $\hat{v}^j(-u - iR)$ equals

$$
\begin{aligned}
\hat{v}^j(-u - iR) =& b_j e^{(-iu+R)z^*} \left(\frac{1}{-iu+R} - \frac{1}{-iu + \widetilde{\Sigma}^{T_j}_{T_0} + R} \right)\\
=& \frac{b_j e^{(-iu+R)z^*} \widetilde{\Sigma}^{T_j}_{T_0}}{(-iu+R)(-iu + \widetilde{\Sigma}^{T_j}_{T_0} + R)}.
\end{aligned}
$$

Now we prove the remaining equation (8.26). Starting from (8.29) and invoking the relations (8.15) for $T = T_0$, we express the cumulant generating function for L under $\mathbb{P}_{T_0}$ in terms of the cumulant generating function of L

under $\mathbb{P}$:

$$
\begin{aligned}
&\theta_s^{\mathbb{P}_{T_0}}(x) \\
&= b_s^{\mathbb{P}_{T_0}} x + \frac{1}{2} c_s^{\mathbb{P}_{T_0}} x^2 + \int_{\mathbb{R}} (e^{xy} - 1 - xy) F_s^{\mathbb{P}_{T_0}}(dy) \\
&= [b_s + c_s \Sigma(s, T_0) + \int_{\mathbb{R}} y(e^{\Sigma(s,T_0)y} - 1) F_s(dy)] x + \frac{1}{2} c_s x^2 \\
&\quad + \int_{\mathbb{R}} (e^{xy} - 1 - xy) e^{\Sigma(s,T_0)y} F_s(dy) \\
&= b_s x + \frac{1}{2} c_s x^2 + c_s \Sigma(s, T_0) x + \int_{\mathbb{R}} [-xy + e^{xy} e^{\Sigma(s,T_0)y} - e^{\Sigma(s,T_0)y}] F_s(dy) \\
&= b_s x + \frac{1}{2} c_s x^2 + c_s \Sigma(s, T_0) x + \int_{\mathbb{R}} [e^{[x+\Sigma(s,T_0)]y} - 1 - [x + \Sigma(s, T_0)] y] F_s(dy) \\
&\quad - \int_{\mathbb{R}} [e^{\Sigma(s,T_0)y} - 1 - \Sigma(s, T_0) y] F_s(dy) \\
&= b_s [x + \Sigma(s, T_0)] + \frac{1}{2} c_s [x + \Sigma(s, T_0)]^2 \\
&\quad + \int_{\mathbb{R}} [e^{[x+\Sigma(s,T_0)]y} - 1 - [x + \Sigma(s, T_0)] y] F_s(dy) \\
&\quad - b_s \Sigma(s, T_0) - \frac{1}{2} c_s \Sigma(s, T_0)^2 - \int_{\mathbb{R}} [e^{\Sigma(s,T_0)y} - 1 - \Sigma(s, T_0) y] F_s(dy) \\
&= \theta_s(x + \Sigma(s, T_0)) - \theta_s(\Sigma(s, T_0)). \qquad \square
\end{aligned}
$$

8.5.2 Delta-hedging of the payer swaption

We start with the determination of the delta-hedge for a short position in the payer swaption when one zero-coupon bond is used for hedging.

Delta-neutral hedge

Theorem 8.5.2. *Under the Assumptions of Theorem 8.5.1 and if $|u| \cdot |\varphi^{\mathbb{P}_{T_0}}_{X_{T_0} - X_t}(u + iR)|$ is integrable then the optimal amount, denoted by Δ_t^j, to invest in the zero-coupon bond with maturity T_j to delta-hedge a short position in the forward payer* *swaption is given by:*

$$\Delta_t^j = \frac{B(t,T_0)}{B(t,T_j)\widetilde{\Sigma}_t^{T_j}} \sum_{k=1}^n c_k(\widetilde{\Sigma}_t^{T_0} H^k(t,X_t) + \frac{\partial}{\partial X_t} H^k(t,X_t)),$$

with

$$H^k(t,X_t) = \frac{e^{-RX_t}}{2\pi} \int_{\mathbb{R}} e^{iuX_t} \varphi_{X_{T_0}-X_t}^{\mathbb{P}_{T_0}}(u+iR)\hat{v}^k(-u-iR)du. \tag{8.32}$$

$$\frac{\partial H^k(t,X_t)}{\partial X_t} = \frac{1}{2\pi} \int_{\mathbb{R}} e^{(-R+iu)X_t} \varphi_{X_{T_0}-X_t}^{\mathbb{P}_{T_0}}(u+iR)\hat{v}^k(-u-iR)(-R+iu)du. \tag{8.33}$$

Proof. According to Theorem 8.5.1 and (8.19) the price of the payer swaption is given by

$$\mathrm{PS}_t = B(t,T_0)\sum_{j=1}^n c_j E^{\mathbb{P}_{T_0}}[(b_j - B(T_0,T_j))^+|\mathcal{F}_t] = B(t,T_0)\sum_{j=1}^n c_j H^j(t,X_t),$$

with $H^j(t,X_t)$ as defined in (8.32). To determine the delta-hedge at time t when investing in the bond $B(\cdot,T_j)$ as a hedge, we have to take the derivative of the price of the payer swaption, PS_t, with respect to this bond:

$$\frac{\partial \mathrm{PS}_t}{\partial B(t,T_j)} = \frac{\partial \mathrm{PS}_t}{\partial X_t}(\frac{\partial B(t,T_j)}{\partial X_t})^{-1} \tag{8.34}$$

with

$$\frac{\partial \mathrm{PS}_t}{\partial X_t} = \sum_{k=1}^n c_k H^k(t,X_t)\frac{\partial B(t,T_0)}{\partial X_t} + B(t,T_0)\sum_{k=1}^n c_k \frac{\partial H^k(t,X_t)}{\partial X_t}, \tag{8.35}$$

and according to (8.18)-(8.19),

$$\frac{\partial B(t,T_i)}{\partial X_t} = \frac{\partial(\widetilde{D}_t^{T_i} e^{\widetilde{\Sigma}_t^{T_i} X_t})}{\partial X_t} = B(t,T_i)\widetilde{\Sigma}_t^{T_i} \qquad \forall i \in \{0,\dots,n\}. \tag{8.36}$$

Inserting these results in (8.34) we arrive at

$$\frac{\partial \mathrm{PS}_t}{\partial B(t,T_j)} = \frac{B(t,T_0)}{B(t,T_j)\widetilde{\Sigma}_t^{T_j}} \sum_{k=1}^n c_k(\widetilde{\Sigma}_t^{T_0} H^k(t,X_t) + \frac{\partial}{\partial X_t} H^k(t,X_t)). \tag{8.37}$$

The optimal number we should hold in a portfolio to delta-hedge a short position in the payer swaption using the bond $B(\cdot, T_j)$ equals $\Delta_t^j = \frac{\partial \mathrm{PS}_t}{\partial B(t,T_j)}$. Indeed, the portfolio with price process V going short in the payer swaption and Δ^j bonds $B(\cdot, T_j)$ is delta-neutral:

$$\frac{\partial V_t}{\partial B(t,T_j)} = -\frac{\partial \mathrm{PS}_t}{\partial B(t,T_j)} + \Delta_t^j = 0.$$

Taking the partial derivative of H^j (8.32) with respect to X, we get the more explicit expression (8.33). Hereto we interchanged the integral and the derivative based on Eberlein et al. (2009) since

(a) $|u| \cdot |\varphi^{\mathbb{P}_{T_0}}_{X_{T_0}-X_t}(u+iR)|$ is integrable

(b) $\hat{v}^k(\cdot - iR)$ is bounded for $k \in \{1, \ldots, n\}$.

Indeed, condition (b) is satisfied because the dampened functions are integrable, hence their Fourier transforms are bounded. Condition (a) is satisfied by assumption. □

Remark: A natural question now is which bond we should use in the hedging strategy without attributing more importance to one bond compared to the others. Recall that each bond depends on the same function X and hence making the product Δ-neutral with respect to one bond, implies that it is Δ-neutral with respect to all zero-coupon bonds:

$$\frac{\partial(-\mathrm{PS}_t + \Delta_t^k B(t,T_k))}{\partial B(t,T_l)} = \frac{\partial B(t,T_k)}{\partial B(t,T_l)} \left(-\frac{\partial \mathrm{PS}_t}{\partial B(t,T_k)} + \Delta_t^k \right) = 0.$$

A possible optimal choice is the bond which minimizes the hedging costs, meaning that we minimize the amount invested in the delta-hedge:

$$\min_k |B(t,T_k) \frac{\partial \mathrm{PS}_t}{\partial B(t,T_k)}| = \min_k |\frac{\partial \mathrm{PS}_t}{\partial X_t} \frac{1}{\widetilde{\Sigma}_t^{T_k}}|.$$

Due to Assumption 8.4.4 $\widetilde{\Sigma}_t^{T_k}$ is a strictly increasing function in T_k, hence the optimal choice is the zero-coupon bond with the longest time to maturity, namely $B(\cdot, T_n)$.

The mean-variance hedging strategy, which will be discussed in Section 8.5.3, uses the zero-coupon bond with maturity T_0 and a second zero-coupon bond with maturity different from T_0. Remark that the zero-coupon bond with maturity T_0 can serve as replacement for the cash account, because at maturity of the option, we know with certainty the value of this cash account.
Hence to compare the mean-variance hedging strategy with the delta-hedge we look at two different possibilities: or we use our so-called cash account to make the portfolio self-financing, or we use this second zero-coupon bond to make the portfolio not only delta- but also gamma-neutral. For the latter, we look at a portfolio consisting of a short position in the swaption, Δ^1 zero-coupon bonds with maturity T_k and Δ^2 zero-coupon bonds with maturity $T_l \neq T_k$. By the previous remark, we know that by making the portfolio delta-neutral with respect to one bond, it is automatically delta-neutral with respect to X and all the other bonds. Hence we can use the second zero-coupon bond to make the portfolio also gamma-neutral. The easiest way to find the optimal amount is by taking the first and second derivative of the portfolio with respect to X, because in the previous remark we showed already that by making the derivative with respect to X zero, then the derivative is also zero when it is taken with respect to any other zero-coupon bond.
We impose again Assumptions 8.4.2-8.4.4, $|\sigma_1| < \overline{\sigma}_1$ and additionally we assume that

(a') the function $u \mapsto |u|^2 \cdot |\varphi^{\mathbb{P}_{T_0}}_{X_{T_0}-X_t}(u+iR)|$ is integrable.

The portfolio is then delta- and gamma-neutral if and only if the following system of equations is satisfied:

$$\begin{cases} \dfrac{\partial V}{\partial X} &= -\dfrac{\partial \text{PS}}{\partial X} + \Delta^1 \dfrac{\partial B(\cdot, T_k)}{\partial X} + \Delta^2 \dfrac{\partial B(\cdot, T_l)}{\partial X} \\ &= -\dfrac{\partial \text{PS}}{\partial X} + \Delta^1 B(\cdot, T_k)\widetilde{\Sigma}^{T_k} + \Delta^2 B(\cdot, T_l)\widetilde{\Sigma}^{T_l} = 0 \\ \dfrac{\partial^2 V}{\partial X^2} &= -\dfrac{\partial^2 \text{PS}}{\partial X^2} + \Delta^1 \dfrac{\partial^2 B(\cdot, T_k)}{\partial X^2} + \Delta^2 \dfrac{\partial^2 B(\cdot, T_l)}{\partial X^2} \\ &= -\dfrac{\partial^2 \text{PS}}{\partial X^2} + \Delta^1 B(\cdot, T_k)(\widetilde{\Sigma}^{T_k})^2 + \Delta^2 B(\cdot, T_l)(\widetilde{\Sigma}^{T_l})^2 = 0 \end{cases}$$

which can be uniquely solved since $\widetilde{\Sigma}_t^{\cdot}$ is strictly increasing by Assumption 8.4.4.

The optimal numbers are given by $\Delta_t^1 = \Delta_t(T_k, T_l)$ and $\Delta_t^2 = \Delta_t(T_l, T_k)$ with

$$\Delta_t(r,v) = \frac{\widetilde{\Sigma}_t^v \frac{\partial \mathrm{PS}_t}{\partial X_t} - \frac{\partial^2 \mathrm{PS}_t}{\partial X_t^2}}{B(t,r)\widetilde{\Sigma}_t^r(\widetilde{\Sigma}_t^v - \widetilde{\Sigma}_t^r)}.$$

We rewrite the derivatives of PS in terms of the price of the payer swaption PS_t and the first and second order partial derivative of H^j with respect to X, by using (8.35) and (8.36) from which we obtain that

$$\begin{aligned}\frac{\partial \mathrm{PS}_t}{\partial X_t} =& \widetilde{\Sigma}_t^{T_0}\mathrm{PS}_t + B(t,T_0)\sum_{j=1}^n c_j \frac{\partial H^j(t,X_t)}{\partial X_t},\\ \frac{\partial^2 \mathrm{PS}_t}{\partial X_t^2} =& (\widetilde{\Sigma}_t^{T_0})^2\mathrm{PS}_t + 2B(t,T_0)\widetilde{\Sigma}_t^{T_0}\sum_{j=1}^n c_j \frac{\partial H^j(t,X_t)}{\partial X_t}\\ &+ B(t,T_0)\sum_{j=1}^n c_j \frac{\partial^2 H^j(t,X_t)}{\partial X_t^2}.\end{aligned}$$

Hence the numbers Δ^1 and Δ^2 can be expressed as values of the function:

$$\begin{aligned}&\Delta_t(r,v)\\ &= \frac{\widetilde{\Sigma}_t^{T_0}(\widetilde{\Sigma}_t^v - \widetilde{\Sigma}_t^r)\mathrm{PS}_t + B(t,T_0)\left[(\widetilde{\Sigma}_t^v - 2\widetilde{\Sigma}_t^{T_0})\sum_{j=1}^n c_j \frac{\partial H_t^j}{\partial X_t} - \sum_{j=1}^n c_j \frac{\partial^2 H^j}{\partial X_t^2}\right]}{B(t,r)\widetilde{\Sigma}_t^r(\widetilde{\Sigma}_t^v - \widetilde{\Sigma}_t^r)},\end{aligned}$$

for $(r,v) = (T_k, T_l)$ respectively $(r,v) = (T_l, T_k)$, where the second derivative of H^j with respect to X follows from (8.33)

$$\frac{\partial^2 H^j(t,X_t)}{\partial X_t^2} = \frac{1}{2\pi}\int_{\mathbb{R}} e^{(iu-R)X_t}\varphi^{\mathbb{P}_{T_0}}_{X_{T_0}-X_t}(u+iR)\hat{v}^k(-u-iR)(-R+iu)^2 du. \tag{8.38}$$

by assumption (a').
Note that we will implement this for $u = T_0$, then the amount $\Delta_t(r, T_0)$ simplifies.

Self-financing delta-hedge

In Theorem 8.5.2 we determined the optimal amount we need to invest in the zero-coupon bond with maturity T_j according to the delta-hedge when only

one zero-coupon bond is used for hedging. To compare the delta-hedge with the MVH strategy we also determined the delta- and gamma-neutral hedge, which uses two zero-coupon bonds, but a more logical choice would be to use the extra zero-coupon bond again as a sort of cash account, as we did for the MVH strategy. In this way we construct a self-financing delta-hedge, where the optimal amount we invest in the zero-coupon bond $B(\cdot, T_j)$ differs from the one determined in Theorem 8.5.2, because the cash account depends also on the underlying variable X. In view of the numerical results we want to obtain, we determine the self-financing delta-hedge directly for the case of discrete hedging. For more information concerning trading in discrete time, we refer to Angelini and Herzel (2009).
In fact at every instantaneous time point we solve the following system of equations:

$$\begin{cases} \dfrac{\partial V_t}{\partial X_t} = -\dfrac{\partial \text{PS}_t}{\partial X_t} + \Delta_t^j \dfrac{\partial B(t, T_j)}{\partial X_t} + \Delta_t^0 \dfrac{\partial B(t, T_0)}{\partial X_t} = 0 \\ \\ (\Delta_t^j - \Delta_{t-1}^j) B(t, T_j) + (\Delta_t^0 - \Delta_{t-1}^0) B(t, T_0) = 0, \end{cases}$$

where at time zero the last equation is replaced by

$$\Delta_0^j B(0, T_j) + \Delta_0^0 B(0, T_0) = 0.$$

Therefore the optimal amount to invest in the cash account in terms of Δ^j equals

$$\Delta_t^0 = \Delta_{t-1}^0 + (\Delta_{t-1}^j - \Delta_t^j) B(t, t_j) / B(t, T_0).$$

Inserting this equation in the first one and using (8.36) leads to

$$\begin{aligned} \frac{\partial \text{PS}_t}{\partial X_t} =& \Delta_t^j B(t, T_j) \widetilde{\Sigma}_t^{T_j} - \Delta_t^j \frac{B(t, T_j)}{B(t, T_0)} B(t, T_0) \widetilde{\Sigma}_t^{T_0} \\ &+ \Delta_{t-1}^j \frac{B(t, T_j)}{B(t, T_0)} B(t, T_0) \widetilde{\Sigma}_t^{T_0} + \Delta_{t-1}^0 B(t, t_0) \widetilde{\Sigma}_t^{T_0} \end{aligned}$$

and hence, since for $T_j \neq T_0$ $\widetilde{\Sigma}_t^{T_j} \neq \widetilde{\Sigma}_t^{T_0}$,

$$\Delta_t^j = \frac{1}{B(t, T_j)(\widetilde{\Sigma}_t^{T_j} - \widetilde{\Sigma}_t^{T_0})} \left(\frac{\partial \text{PS}_t}{\partial X_t} - \widetilde{\Sigma}_t^{T_0} (\Delta_{t-1}^j B(t, T_j) + \Delta_{t-1}^0 B(t, T_0)) \right).$$

8.5.3 Mean-variance hedging strategy for the payer swaption

Quadratic hedging strategies are defined in terms of the discounted portfolio, the discounted underlyings,.... Discounting happens mostly with the risk-free interest rate, but the theory is in fact described more general for any numéraire whose price process is strictly positive. In the interest rate derivatives market it is unrealistic to hedge with the risk-free interest rate product in contrast to modeling in the stock market. Therefore we choose the bond $B(\cdot, T_0)$ as numéraire to develop a hedging strategy for the payer swaption.

The quadratic hedging strategies are always defined in terms of the assets discounted by the numéraire, which of course are not available on the market. From a practical point of view one has to translate the amount to invest in a discounted asset into the corresponding amount to invest in the non-discounted one in order to have a meaningful strategy. In the mean-variance hedging strategy the portfolio is self-financing. This property ensures that the optimal amount of non-discounted assets equals the optimal amount of discounted assets. This motivates our choice to study the mean-variance hedging strategy and not the locally risk-minimizing hedging strategy, which is not self-financing.

We determine the mean-variance hedging strategy for the payer swaption under the forward measure $\mathbb{P}_{T_0}$ using the numéraire $B(\cdot, T_0)$. For the explicit determination of the strategy we use ideas of Hubalek et al. (2006) adapted to our setting. They determine the variance-optimal hedging strategy for an exponential Lévy process, which is not necessarily a martingale, while we work under the forward measure $\mathbb{P}_{T_0}$ which ensures us that the discounted asset $B(\cdot, T_j)/B(\cdot, T_0)$ is a martingale. Due to this martingale measure we only need to find the Galtchouk-Kunita-Watanabe decomposition of the claim H. On the other hand Hubalek et al. only use time-homogeneous processes, while we are more interested in time-inhomogeneous processes.

We start with repeating the intermediate results of Hubalek et al. for the determination of the Föllmer-Schweizer decomposition. They denote by $S_t = S_0 \exp(X_t)$ the discounted price process of $S = S_0 + Z + B$, where Z denotes the martingale part and B the finite variation part. The T-contingent claim H is written as a function of the underlying S: $H = f(S_T)$ with $f : (0, \infty) \to \mathbb{R}$ and having the form

$$f(s) = \int s^z \Pi(dz) \tag{8.39}$$

for some finite complex measure Π on a strip $\{z \in \mathbb{C} : R' \leq \text{Re}(z) \leq R\}$ where R and $R' \in \mathbb{R}$ are chosen such that $E(e^{2R'X_1}) < \infty$ and $E(e^{2RX_1}) < \infty$. Furthermore they exclude the case that S is deterministic by assuming that $\kappa(2) - 2\kappa(1) \neq 0$, with κ the cumulant generating function.

Lemma 8.5.3 (See Hubalek et al. (2006) Lemma 3.8). *Let $z \in \mathbb{C}$ with $S_T^z \in L^2(P)$. Then $H(z) = S_T^z$ admits a Föllmer-Schweizer decomposition $H(z) = H_0(z) + \int_0^T \xi_t(z) dS_t + L_T(z)$, where*

$$\begin{aligned} H_t(z) :=& e^{\eta(z)(T-t)} S_t^z, \\ \xi_t(z) :=& \gamma_t(z) e^{\eta(z)(T-t)} S_{t-}^{z-1}, \\ L_t(z) :=& H_t(z) - H_0(z) - \int_0^t \xi_u(z) dS_u \end{aligned} \tag{8.40}$$

and the processes γ and η are defined as

$$\begin{aligned} \gamma(z) :=& \frac{\kappa(z+1) - \kappa(z) - \kappa(1)}{\kappa(2) - 2\kappa(1)}, \\ \eta(z) :=& \kappa(z) - \kappa(1)\gamma(z). \end{aligned} \tag{8.41}$$

Moreover, Z is a square-integrable martingale and hence $L(z)Z$ is a martingale.

Remark that it looks as if in Hubalek et al. (2006) the process for the claim $H(z)$ is found by trial and error. In fact imposing that the process is of the form $H(z) = e^{\eta(z)(T-\cdot)} S^z$ we can determine $\eta(z)$ from the following relation, (4.26), for the optimal number of risky assets

$$\xi_t(z) = \frac{d\langle (H(z))^m, Z \rangle_t}{d\langle Z, Z \rangle_t},$$

where $(H(z))^m$ denotes the martingale part of the process $H(z)$ under the original measure. After inserting the process for $(H(z))^m$, we rewrite the optimal number as

$$\xi_t(z) = \frac{e^{\eta(z)(T-t)} d\langle Z(z), Z \rangle_t}{d\langle Z, Z \rangle_t},$$

with $Z(z)$ the martingale part of the process S^z. Invoking formulas (13) and (19) of Hubalek et al. (2006) leads to the required result (8.40). The formulas

(8.41) are then determined by imposing the martingale property on the process L.

Proposition 8.5.4 (See Hubalek et al. (2006) Proposition 3.10). *Any contingent claim $H = f(S_T)$, with f as described in* (8.39) *admits a Föllmer-Schweizer decomposition $H = H_0 + \int_0^T \xi_t dS_t + L_T$. Using the notations of Lemma 8.5.3, it is given by*

$$\begin{aligned} H_t &= \int H_t(z)\Pi(dz), \\ \xi_t &= \int \xi_t(z)\Pi(dz), \\ L_t &= \int L_t(z)\Pi(dz) = H_t - H_0 - \int_0^t \xi_u dS_u. \end{aligned}$$

Moreover the processes H, ξ and L are real-valued.

As a second step, we apply this result to our setting of time-inhomogeneous processes whose discounted price processes are martingales under the measure $\mathbb{P}_{T_0}$.
Determining the mean-variance hedging strategy for a payer swaption means finding a self-financing strategy that minimizes

$$E^{\mathbb{P}_{T_0}}[(\text{PS}_{T_0} - \widetilde{V}_{T_0})^2]$$

with $\text{PS}_{T_0} = \frac{\text{PS}_{T_0}}{B(T_0,T_0)}$ the (discounted) price at time T_0 of the payer swaption and $\widetilde{V}$ the (discounted) portfolio value process which equals $\frac{V}{B(\cdot,T_0)}$. The value of the self-financing portfolio V, containing ξ^0 zero-coupon bonds with maturity T_0 and ξ^j zero-coupon bonds with maturity T_j, is given by

$$V_t = \xi_t^0 B(t,T_0) + \xi_t^j B(t,T_j) = V_0 + (\xi^0 \cdot B(\cdot,T_0))_t + (\xi^j \cdot B(\cdot,T_j))_t.$$

Due to the self-financing condition the discounted value is given by

$$\widetilde{V}_t = \frac{V_t}{B(t,T_0)} = \widetilde{V}_0 + \int_0^t \xi_s^j d\widetilde{B}(s,T_j),$$

where $\widetilde{B}(s,T_j) = \frac{B(s,T_j)}{B(s,T_0)}$. Therefore the mean-variance hedging condition is to minimize

$$E^{\mathbb{P}_{T_0}}[(\mathrm{PS}_{T_0} - (\widetilde{V}_0 + \int_0^{T_0} \xi_u^j d\widetilde{B}(u,T_j)))^2]$$

Note that the discounted bond $\widetilde{B}(\cdot, T_j)$ is a martingale under the forward measure $\mathbb{P}_{T_0}$. Hence the determination of the MVH strategy reduces to finding the Galtchouk-Kunita-Watanabe decomposition of PS_{T_0} under this measure, see Chapter 5. Furthermore from Section 2.4, we know that the existence of the Galtchouk-Kunita-Watanabe decomposition is guaranteed if the claim and the underlying are both locally square-integrable local martingales. Furthermore, the number of discounted risky assets is given by

$$\xi_t^j = \frac{d\langle \overline{V}, \widetilde{B}(\cdot,T_j)\rangle_t^{\mathbb{P}_{T_0}}}{d\langle \widetilde{B}(\cdot,T_j), \widetilde{B}(\cdot,T_j)\rangle_t^{\mathbb{P}_{T_0}}}, \tag{8.42}$$

with $\overline{V}_t = E^{\mathbb{P}_{T_0}}[\mathrm{PS}_{T_0}|\mathcal{F}_t]$ and the angle brackets $\langle X, Y\rangle^{\mathbb{P}_{T_0}}$ denoting the $\mathbb{P}_{T_0}$-compensator of $[X,Y]$.
To express the right-hand side of formula (8.42) more explicit in terms of the portfolio characteristics we proceed as Hubalek et al. (2006). We calculate the strategy for special types of claims, namely $H(z) = \widetilde{B}(T_0,T_j)^z$ for a $z \in \mathbb{C}$, satisfying $E^{\mathbb{P}_{T_0}}[\widetilde{B}(T_0,T_j)^{2z}] < +\infty$. By the square-integrability of the claim the existence of the Galtchouk-Kunita-Watanabe decomposition is guaranteed.
In Lemma 8.5.5 we rewrite the price of the payer swaption at maturity T_0 in the form (8.39).

Lemma 8.5.5. *The price of the payer swaption at maturity* T_0

$$PS_{T_0} = \sum_{k=1}^{n} c_k \frac{1}{2\pi} \int_{\mathbb{R}} e^{(iu-R)X_{T_0}} \hat{v}^k(-u - iR)du$$

can be expressed as

$$PS_{T_0} = \int_{\mathbb{R}} \widetilde{B}(T_0,T_j)^{\frac{iu-R}{\tilde{\Sigma}_{T_0}^{T_j}}} \Pi(du),$$

with

$$\Pi(du) = \sum_{k=1}^{n} \frac{c_k}{2\pi} (f_{T_0}^j)^{\frac{iu-R}{\tilde{\Sigma}_{T_0}^{T_j}}} \hat{v}^k(-u - iR)du, \tag{8.43}$$

$$f_{T_0}^j = \frac{B(0,T_0)}{B(0,T_j)} \exp(\int_0^{T_0} [\theta_s(\Sigma(s,T_j)) - \theta_s(\Sigma(s,T_0))]ds). \tag{8.44}$$

Proof. We first derive the dynamics of the discounted zero-coupon bond used for hedging. Combining (8.12) and (8.13) we get

$$\begin{aligned}\widetilde{B}(t,T_j) =&\frac{B(t,T_j)}{B(t,T_0)}\\ =&\frac{B(0,T_j)}{B(0,T_0)}\exp\Big(-\int_0^t[\theta_s(\Sigma(s,T_j))-\theta_s(\Sigma(s,T_0))]ds\Big)\\ &\times\exp\Big(\int_0^t[\Sigma(s,T_j)-\Sigma(s,T_0)]dL_s\Big)\\ :=&\widetilde{B}(0,T_j)\exp(\widetilde{X}_t^j)\end{aligned} \tag{8.45}$$

where, with an analogous reasoning and notation as to arrive at (8.16),

$$\begin{aligned}\widetilde{X}_t^j =&-\int_0^t[\theta_s(\Sigma(s,T_j))-\theta_s(\Sigma(s,T_0))]ds+\int_0^t[\Sigma(s,T_j)-\Sigma(s,T_0)]dL_s\\ =&-\int_0^t[\theta_s(\Sigma(s,T_j))-\theta_s(\Sigma(s,T_0))]ds+\widetilde{\Sigma}_{T_0}^{T_j}X_t.\end{aligned} \tag{8.46}$$

Note that the process $\widetilde{X}^j$ is a time-inhomogeneous Lévy process with local characteristics $(b_s^{\widetilde{X}^j}, c_s^{\widetilde{X}^j}, F_s^{\widetilde{X}^j})_{s\in[0,T^*]}$ under the measure $\mathbb{P}_{T_0}$.
By combining (8.45), (8.46) and (8.44) we deduce that

$$e^{X_t}=(\widetilde{B}(t,T_j)f_t^j)^{\frac{1}{\widetilde{\Sigma}_{T_0}^{T_j}}}. \tag{8.47}$$

By means of (8.47) we rewrite the price at time T_0 in terms of the discounted zero-coupon bond with maturity T_j:

$$\begin{aligned}\frac{\mathrm{PS}_{T_0}}{B(T_0,T_0)} =&\sum_{k=1}^n c_k\frac{1}{2\pi}\int_{\mathbb{R}}e^{(iu-R)X_{T_0}}\hat{v}^k(-u-iR)du\\ =&\sum_{k=1}^n c_k\frac{1}{2\pi}\int_{\mathbb{R}}(\widetilde{B}(T_0,T_j)f_{T_0}^j)^{\frac{iu-R}{\widetilde{\Sigma}_{T_0}^{T_j}}}\hat{v}^k(-u-iR)du\\ =&\int_{\mathbb{R}}\widetilde{B}(T_0,T_j)^{\frac{iu-R}{\widetilde{\Sigma}_{T_0}^{T_j}}}\Pi(du),\end{aligned}$$

where $\Pi(du)$ as in (8.43) is a deterministic function. Furthermore by substituting z for $\frac{iu-R}{\widetilde{\Sigma}_{T_0}^{T_j}}$ we can conclude that Π is a finite complex measure defined on a strip as described by Hubalek et al. (2006). □

Using Lemma 8.5.5 we state in the next theorem the optimal number of bonds in which we should invest according to the mean-variance hedging strategy.

Theorem 8.5.6. *Under the assumptions of Theorem 8.5.1 and if additionally* $3M' \leq M$ *and if* R *is chosen in the interval* $]0, \frac{M}{2\sigma_1}]$ *then the Galtchouk-Kunita-Watanabe decomposition of the forward payer swaption* (8.24) *exists. The optimal number* ξ_t^j *to invest in the zero-coupon bond with maturity* T_j *is according to the mean-variance hedging strategy given by*

$$\int_{\mathbb{R}} e^{\int_t^{T_0} \kappa_s^{\widetilde{X}^j}(\frac{iu-R}{\widetilde{\Sigma}_{T_0}^{T_j}})ds} \widetilde{B}(t-,T_j)^{\frac{iu-R}{\widetilde{\Sigma}_{T_0}^{T_j}}-1} \frac{\kappa_t^{\widetilde{X}^j}(\frac{iu-R}{\widetilde{\Sigma}_{T_0}^{T_j}}+1) - \kappa_t^{\widetilde{X}^j}(\frac{iu-R}{\widetilde{\Sigma}_{T_0}^{T_j}})}{\kappa_t^{\widetilde{X}^j}(2)} \Pi(du),$$

with

$$\kappa_s^{\widetilde{X}^j}(w) = \theta_s(w\Sigma(s,T_j) + (1-w)\Sigma(s,T_0)) - w\theta_s(\Sigma(s,T_j)) - (1-w)\theta_s(\Sigma(s,T_0)), \tag{8.48}$$

and $\Pi(du)$ *as in* (8.43).

Proof. We proceed in several steps.

- First, we calculate the moment generating function of the process $\widetilde{X}$ under the measure $\mathbb{P}_{T_0}$ using (8.13) and (8.14) with $T = T_0$:

$$\begin{aligned}
&E^{\mathbb{P}_{T_0}}[\exp(w\widetilde{X}_t^j)] \\
&= E[\exp(w\widetilde{X}_t^j) \left.\frac{d\mathbb{P}_{T_0}}{d\mathbb{P}}\right|_{\mathcal{F}_t}] \\
&= E[\exp(w\widetilde{X}_t^j) \exp(-\int_0^t \theta_s(\Sigma(s,T_0))ds + \int_0^t \Sigma(s,T_0)dL_s)] \\
&= \exp(-w\int_0^t [\theta_s(\Sigma(s,T_j)) - \theta_s(\Sigma(s,T_0))]ds - \int_0^t \theta_s(\Sigma(s,T_0))ds) \\
&\quad \times E[\exp(w\int_0^t [\Sigma(s,T_j) - \Sigma(s,T_0)]dL_s + \int_0^t \Sigma(s,T_0)dL_s)]
\end{aligned}$$

$$= \exp(\int_0^t \theta_s(w\Sigma(s,T_j) + (1-w)\Sigma(s,T_0))ds)$$
$$\times \exp(\int_0^t [-w\theta_s(\Sigma(s,T_j)) - (1-w)\theta_s(\Sigma(s,T_0))]ds)$$
$$:= \exp(\int_0^t \kappa_s^{\widetilde{X}^j}(w)ds).$$

where $\kappa^{\widetilde{X}^j}$ denotes the cumulant function of the process $\widetilde{X}^j$ under the measure $\mathbb{P}_{T_0}$ given by (8.48). Then, processes of the form

$$N^z := \widetilde{B}(\cdot,T_j)^z e^{-\int_0^\cdot \kappa_u^{\widetilde{X}^j}(z)du}$$

are martingales under the measure $\mathbb{P}_{T_0}$ due to the independent increments of the process $\widetilde{X}^j$. Indeed it holds that

$$E^{\mathbb{P}_{T_0}}[N_T^z|\mathcal{F}_t] = e^{-\int_0^T \kappa_u^{\widetilde{X}^j}(z)du}\widetilde{B}(0,T_j)^z e^{z\widetilde{X}_t^j} E^{\mathbb{P}_{T_0}}[e^{z(\widetilde{X}_T^j - \widetilde{X}_t^j)}]$$
$$= e^{-\int_0^t \kappa_u^{\widetilde{X}^j}(z)du}\widetilde{B}(t,T_j)^z = N_t^z. \tag{8.49}$$

- As a second step we calculate the process

$$(H_t(z))_{t\in[0,T_0]} := E^{\mathbb{P}_{T_0}}[\widetilde{B}(T_0,T_j)^z|\mathcal{F}_t].$$

According to (8.49) we immediately have $H_t(z)$ in terms of the cumulant function of the process $\widetilde{X}^j$

$$H_t(z) = E^{\mathbb{P}_{T_0}}[\widetilde{B}(T_0,T_j)^z|\mathcal{F}_t] = \widetilde{B}(t,T_j)^z e^{\int_t^{T_0} \kappa_u^{\widetilde{X}^j}(z)du}. \tag{8.50}$$

- Next we calculate the angle brackets

$$\langle H(z), \widetilde{B}(\cdot,T_j)\rangle_t^{\mathbb{P}_{T_0}} \qquad \text{and} \qquad \langle \widetilde{B}(\cdot,T_j), \widetilde{B}(\cdot,T_j)\rangle_t^{\mathbb{P}_{T_0}}$$

where the latter equals $\langle H(1), \widetilde{B}(\cdot,T_j)\rangle_t^{\mathbb{P}_{T_0}}$. This follows from the fact that $H_t(1) = \widetilde{B}(t,T_j)$ due to (8.50) and since according to (8.48) $\kappa^{\widetilde{X}^j}(1) = 0$. Note that for quasi-left-continuous processes $\langle X,Y\rangle^{\mathbb{P}_{T_0}} = \langle X^m, Y^m\rangle^{\mathbb{P}_{T_0}}$, where we denote by $\cdot^m$ the $\mathbb{P}_{T_0}$-martingale part. Thus we only need the $\mathbb{P}_{T_0}$-martingale part of $H(z)$ for the calculation of the angle brackets. From Itô's formula we deduce that

$$dH_t(z)^{\mathrm{m}} = e^{\int_t^{T_0} \kappa_u^{\widetilde{X}^j}(z)du}\widetilde{B}(0,T_j)^z d(e^{z\widetilde{X}_t^j})^{\mathrm{m}}.$$

We rewrite $e^{z\widetilde{X}_t^j}$ as the Doléans-Dade exponential $\mathcal{E}(\widetilde{Y}^j(z))$ with, see Lemma 2.6 of Kallsen and Shiryaev (2002)

$$\widetilde{Y}^j(z) = z\widetilde{X}^j + \int_\cdot^{T_0} \kappa_u^{\widetilde{X}^j}(z)du + \frac{z^2}{2}\langle(\widetilde{X}^j)^c\rangle^{\mathbb{P}_{T_0}} + (e^{zx} - 1 - zx) * \mu^{\widetilde{X}^j},$$

with $\mathbb{P}_{T_0}$-martingale part $(\widetilde{Y}^j)^{\mathrm{m}}$ equal to

$$(\widetilde{Y}^j)^{\mathrm{m}} = z(\widetilde{X}^j)^c + (e^{zx} - 1) * (\mu^{\widetilde{X}^j} - F^{\widetilde{X}^j}) \tag{8.51}$$

and where we used the characteristics $(b_s^{\widetilde{X}^j}, c_s^{\widetilde{X}^j}, F_s^{\widetilde{X}^j})$ of $\widetilde{X}^j$ under $\mathbb{P}_{T_0}$ and we denoted the measure associated to the jumps of $\widetilde{X}^j$ by $\mu^{\widetilde{X}^j}$. Recall that $H_t(1) = \widetilde{B}(t, T_j) = \widetilde{B}(0, T_j)e^{\widetilde{X}_t^j}$, hence

$$d\langle H(z), \widetilde{B}(\cdot, T_j)\rangle_t^{\mathbb{P}_{T_0}} = e^{\int_t^{T_0} \kappa_u^{\widetilde{X}^j}(z)du}\widetilde{B}(t-, T_j)^{z+1}d\langle\widetilde{Y}^j(z), \widetilde{Y}^j(1)\rangle_t^{\mathbb{P}_{T_0}}. \tag{8.52}$$

Taking into account that $\kappa^{\widetilde{X}^j}$ is given by

$$b_s^{\widetilde{X}^j}z + \frac{1}{2}c_s^{\widetilde{X}^j}z^2 + \int_{\mathbb{R}}(e^{xz} - 1 - xz)F_s^{\widetilde{X}^j}(dx),$$

we derive from (8.51) by elementary calculations that

$$\begin{aligned}\langle\widetilde{Y}^j(z), \widetilde{Y}^j(1)\rangle_t^{\mathbb{P}_{T_0}} =& z\int_0^t c_u^{\widetilde{X}^j}du + \int_0^t\int_{\mathbb{R}}(e^{zx} - 1)(e^x - 1)F_u^{\widetilde{X}^j}(dx)du \\ =& \int_0^t [\kappa_u^{\widetilde{X}^j}(z+1) - \kappa_u^{\widetilde{X}^j}(z) - \kappa_u^{\widetilde{X}^j}(1)]du \\ =& \int_0^t [\kappa_u^{\widetilde{X}^j}(z+1) - \kappa_u^{\widetilde{X}^j}(z)]du.\end{aligned} \tag{8.53}$$

- Combining (8.52) and (8.53) in (8.42), the optimal number of risky assets related to the claim $H_{T_0}(z)$ is for every $t \in [0, T_0]$ given by

$$\begin{aligned}\xi_t^j(z) =& \frac{d\langle H(z), \widetilde{B}(\cdot, T_j)\rangle_t^{\mathbb{P}_{T_0}}}{d\langle\widetilde{B}(\cdot, T_j), \widetilde{B}(\cdot, T_j)\rangle_t^{\mathbb{P}_{T_0}}} \\ =& e^{\int_t^{T_0}\kappa_u^{\widetilde{X}^j}(z)du}\widetilde{B}(t-, T_j)^{z-1}\frac{\kappa_t^{\widetilde{X}^j}(z+1) - \kappa_t^{\widetilde{X}^j}(z)}{\kappa_t^{\widetilde{X}^j}(2)}.\end{aligned}$$

- According to Proposition 8.5.4 and by using the linearity of the Galtchouk-Kunita-Watanabe decomposition, we need to invest the amount

$$\xi_t^j = \int_{\mathbb{R}} \xi_t^j \left(\frac{iu - R}{\widetilde{\Sigma}_{T_0}^{T_j}}\right) \Pi(du), \tag{8.54}$$

with Π as defined in Lemma 8.5.5, in the bond $B(\cdot, T_j)$ to hedge the forward payer swaption if the decomposition exists. Hence we will discuss the adequateness of the assumptions to ensure the existence of the decomposition. First of all, every claim $H(u) = \widetilde{B}(t, T_j)^{\frac{iu-R}{\widetilde{\Sigma}_{T_0}^{T_j}}}$, $u \in \mathbb{R}$, should be square-integrable. From (8.45) and by an analogous reasoning as on page 180, where the finiteness of $M_{X_{T_0}}^{\mathbb{P}_{T_0}}(z)$ was proved, we deduce here that every claim $H(u)$, $u \in \mathbb{R}$, is square-integrable if and only if

$$-M \leq 2 \frac{-R}{\widetilde{\Sigma}_{T_0}^{T_j}} \widetilde{\Sigma}_{T_0}^{T_j} \sigma_1(s) + \Sigma(s, T_0) \leq M.$$

Therefore this leads in a similar way as on page 180 to $R \leq \frac{M}{2\overline{\sigma}_1}$. Secondly, we obtain by a same reasoning that the discounted zero-coupon bond $\widetilde{B}(\cdot, T_j)$ is square-integrable under the measure $\mathbb{P}_{T_0}$ if and only if

$$-M \leq 2 \widetilde{\Sigma}_{T_0}^{T_j} \sigma_1(s) + \Sigma(s, T_0) \leq M.$$

From (8.8), (8.17) and Assumption 8.4.4 follows that $\widetilde{\Sigma}_{T_0}^{T_j} \sigma_1(s) \leq \Sigma(s, T^*)$. By Assumption 8.4.3 we obtain that therefore the discounted zero-coupon bond will be square-integrable if $3M' \leq M$. Under these two extra assumptions the square-integrability of the claim and the underlying is guaranteed. Hence according to Ansel and Stricker (1993), see Section 2.4, the Galtchouk-Kunita-Watanabe decomposition of the forward payer swaption under the measure $\mathbb{P}_{T_0}$ exists. Combined with the linearity of the decomposition we immediately obtain (8.54) and we do not need to check explicitly all the conditions as is done in Hubalek et al. (2006) Proposition 3.1. □

8.6 Numerical results

In this section we give a first insight in the performance of the delta-hedge and the mean-variance hedge in the typical setting of interest rate derivatives markets. These experiments are still preliminary and only give us a rough idea of the usefulness.

To simulate the bond prices we need to fix the time-inhomogeneous process L and the volatility structure. Our choice was inspired by Kluge (2005) who calibrated a set of at-the-money swaptions when L follows a normal inverse Gaussian (NIG) process and the Vasiček model is used to describe the volatility structure. We did not simulate the time-inhomogeneous NIG process, because as long as we keep maturity fixed, there is no real necessity in allowing time-inhomogeneity. When payer swaptions with different maturities would be hedged, then it is advised to introduce also time-inhomogeneity in the simulations.

The NIG model with parameters $\alpha > 0$, $-\alpha < \beta < \alpha$ and $\delta > 0$ has as characteristic function $\phi(z)$:

$$\phi(z) = \exp(-\delta(\sqrt{\alpha^2 - (\beta + iz)^2} - \sqrt{\alpha^2 - \beta^2})),$$

see Schoutens (2003).

The Vasiček volatility structure is given by $\sigma(s,T) = \hat{\sigma}e^{-a(T-s)}$, for real constants $\hat{\sigma} > 0$ and $a \neq 0$. Furthermore we can add the multiplicative constant $\hat{\sigma}$ to the process L and hence without loss of generality we can choose $\hat{\sigma}$ equal to 1. As model parameters we were inspired by the parameters Kluge (2005) obtained by calibration of receiver swaptions. As a first choice, we work with

$$\begin{aligned} a &= 0.02 \\ \alpha = 2730.651, \quad \beta &= -230.663, \quad \delta = 0.1. \end{aligned}$$

Furthermore $\sigma(s,T)$ is given by $e^{-a(T-s)}$, then obviously $\sigma_1(s) = e^{as}$ and $\sigma_2(T) = e^{-aT}$. The processes $\widetilde{\Sigma}$ and Σ equal

$$\begin{aligned} \widetilde{\Sigma}_t^{T_j} &= \int_t^{T_j} \sigma_2(u)du = \int_t^{T_j} e^{-au}du = \frac{1}{a}[e^{-at} - e^{-aT_j}], \\ \Sigma(s,T) &= \int_s^T \sigma(s,u)du = \int_s^T e^{-a(u-s)}du = \frac{1 - e^{-a(T-s)}}{a}. \end{aligned}$$

We assume a swaption with maturity T_0 in 10 years and another 10 years as tenor of the underlying swap. As starting bond prices, we use the prices of Kluge (2005) which gives twice a year the price of the bond for the next 20 years. We hedge during the first 10 years using zero-coupon bonds from the set $\{B(\cdot, T_0), \dots, B(\cdot, T_{20})\}$, with $T_n = 10 + n/2$ in years, which are the same bonds on which the swaption depends.
In the previous sections we investigated the payer swaption, which is a weighted sum of put options on zero-coupon bonds. Hence the risk related with this product is really limited.
Therefore in this section we price and hedge receiver swaptions, which are weighted sums of call options on zero-coupon bonds. The formulas obtained above remain the same and only the range of R, which should be negative now, has to be changed, see also Example 5.1 of Eberlein et al. (2009).
To speed up the computations we apply fast Fourier transform as described by Carr and Madan (1999), but we do not calculate the prices and the hedges for a range of strike prices, but for a range of values of the underlying X. The changes in this X can be rather small and hence, we adjusted the parameters suggested by Carr and Madan (1999) to $N = 32768$ and $\eta = 0.125$. We rehedge once a week, more is not really necessary in an interest rate derivatives market.

Comparison of the hedges for the receiver swaption

To compare the performance of the different hedges, namely the MVH strategy (Section 8.5.3), the self-financing delta-hedge and the non-self-financing delta- and gamma-neutral hedge (Section 8.5.2), we calculate the mean and the standard deviation of the total cost using 12 500 simulation paths. The standard deviation is given between brackets in the table. The total cost is defined as the sum of the initial cost and the final deviation of the hedging portfolio from the swaption in case of a self-financing portfolio. Hence the L^2 hedging error, which is often calculated in literature, equals the sum of the variance of the total cost with the square of the difference between the mean of total cost and the initial cost. For a non-self-financing portfolio the total cost is defined as the sum of all the costs on every rebalance date.
We first investigate which zero-coupon bond is the most optimal to use as hedging bond when the portfolio is weekly rebalanced. Knowing that we use already the bond $B(\cdot, T_0)$ as sort of cash account, the most logical choice would be to use the zero-coupon bond which differs most from the cash-account, namely

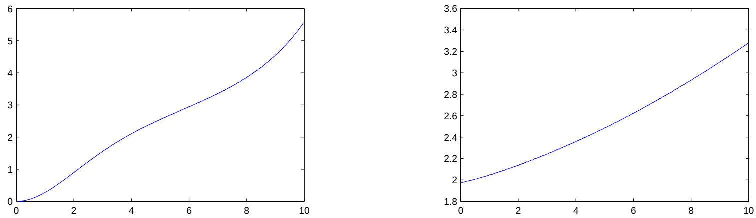
	$B(\cdot, T_1)$	$B(\cdot, T_{10})$	$B(\cdot, T_{20})$
Delta	$9.51\ (0.77)$	$3.02\ (0.24)$	$-\ 2.30\ (0.22)$
Delta-gamma	$87.93\ (5.78)$	$35.19\ (2.63)$	$30.01\ (2.64)$
MVH	$4.36\ (0.40)$	$3.88\ (0.39)$	$3.28\ (0.38)$

Table 8.1: Hedging cost and standard deviation in case of a receiver swaption.

$B(\cdot, T_{20})$ with $T_{20} = 20$. Based on the results given in Table 8.1, we see that all the proposed strategies perform the best when the last zero-coupon bond is used, because with the last bond all strategies not only have a lower expected total cost, but also a lower variance.

From now on we focus on the hedging when the last bond is used. The different intentions of the hedging strategies are obvious: a MVH strategy tries to minimize the square of the cost, and does this in a cautious way. While a delta-hedge is following possible changes in the price at every time and by using the last bond the delta-hedge is even winning money. We think that the delta-hedge performs in this setting better than the MVH strategy, due to the rather smooth behaviour of the underlying processes, which does not have extreme jumps and due to the natural limitations of the zero-coupon bonds, which values are always between zero and one. In Figure 8.6.1, we plotted the average value of the delta-hedge strategy and the MVH strategy. The starting value of the delta-hedge is zero, while the MVH strategy starts at the expected value of the swaption. At maturity the value of the swaption is on average equal to 3.29. Hence at maturity the delta-hedge portfolio is worth more than what is needed for the swaption, while the MVH strategy contains just enough. The delta-hedge is even making money, see Figure 8.6.1, at every instant because roughly speaking the bond prices are increasing when approaching maturity. The mean-variance hedging strategy is also making profit from this increase, but just enough to have the expected payoff at maturity, see Figure 8.6.1 and is not making any gains, as requested from the quadratic conditions. We believe that if there would be a serious crash on the market, the MVH strategy would outperform the delta-hedge and that in such case the real power of quadratic hedging strategies would manifest, see also Chapter 9.

The delta- and gamma-neutral hedge is overreacting and is protecting the strategy too much, because this strategy also tries to cover possible changes in the second derivative. From these results we should say that the cost of a delta-gamma hedge is unreasonable high for this setting, but this strongly depends



Delta-hedge Mean-variance hedge

Figure 8.6.1: Average value of the hedging portfolio when $B(\cdot, T_{20})$ is used to hedge a receiver swaption.

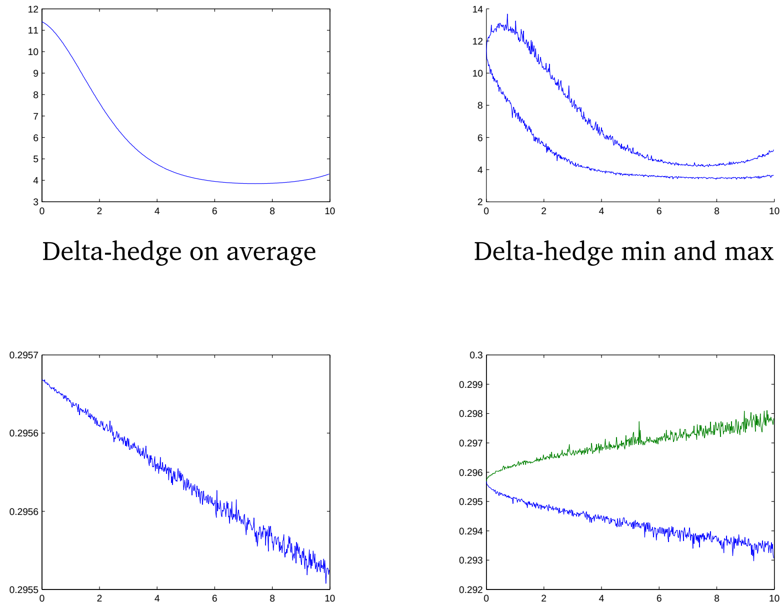
	$a = 0.02$	$a = 0.06$
$\delta = 0.1$	30.01 (2.64)	20.92 (1.80)
$\delta = 0.06$	17.68 (1.53)	12.32 (1.07)

Table 8.2: Hedging cost and standard deviation for the delta- and gamma-hedging of a receiver swaption.

on the parameter δ of the NIG model and the parameter a of the Vasiček model. Decreasing δ or increasing a leads to a better performance of this hedge, see Table 8.2.

In Figure 8.6.2 we investigate in more detail the amounts we have to invest according to the MVH strategy and the self-financing delta-hedge. In this figure we plotted at the left side the optimal amount we have to invest on average in $B(\cdot, T_{20})$ according to the delta-hedge respectively the mean-variance hedge, while on the right side we show at every time the minimum and maximum amount taken in the set of all simulations to invest in $B(\cdot, T_{20})$ needed for the delta-hedging, respectively the MVH strategy. Note also the different scales of the plots, especially the difference between the delta-hedge and the MVH strategy. The closer we get to maturity, the less risk there is and hence both strategies decrease the amount they invest in the zero-coupon bond $B(\cdot, T_{20})$. Moreover from comparison of the left plot of the MVH strategy with the right one, we learn that there will only be few paths in the maximum range, because on average the amount is decreasing through whole the hedging period.

Next, we also plotted the amount we invest in the bond $B(\cdot, T_0)$ for the delta-hedge and the MVH strategy. This highlights again the different aim of both



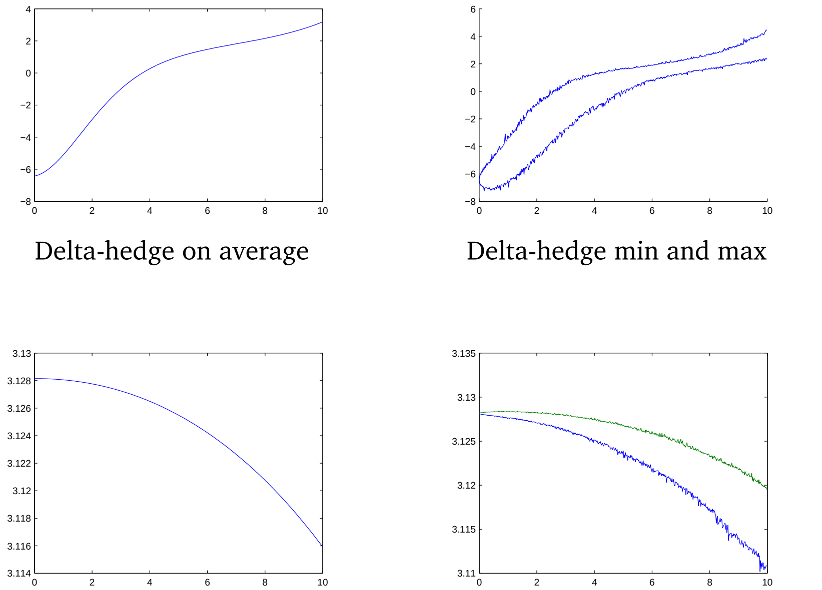
Delta-hedge on average — Delta-hedge min and max

Mean-variance hedge on average — Mean-variance hedge min and max

Figure 8.6.2: Optimal number to invest in $B(\cdot, T_{20})$ to hedge a receiver swaption.

strategies. We deduce from the amount we invest in $B(\cdot, T_0)$, see figure 8.6.3, that the delta-hedge is making profit, while the amount invested according to the MVH strategy is reduced over the lifetime of the option. Furthermore we also see that none of the sample paths have a completely different range from the others. This illustrates again the good behaviour of the market and the little risk which is involved.

We discuss the cost in more detail. In the MVH strategy there is only an input of money at the start and possibly also at maturity of the option. The starting cost of the MVH strategy is defined as the value of the option under the martingale measure $\mathbb{P}_{T_0}$ at the starting time, which equals in our setting 1.97 on average. Due to the self-financing property of the delta-hedge and the mean-variance hedge, all the other costs are due at maturity of the contract. Hence in order to compare the cost we divide the original cost by the bond $B(\cdot, T_0)$ to obtain the



Delta-hedge on average — Delta-hedge min and max

Mean-variance hedge on average — Mean-variance hedge min and max

Figure 8.6.3: Optimal number to invest in $B(\cdot, T_0)$ to hedge a receiver swaption.

cost at time T_0. This leads to a cost of 3.29 for the MVH strategy and the only way to reduce this cost is making gains by trading on the market. Unfortunately on average the market is almost not moving differently and therefore the cost of the MVH strategy is almost equal to the original cost. Combining the self-financing condition, with the decreasing movements in Figure 8.6.2 and Figure 8.6.3, we obviously see that the portfolio is not gaining money.

If we look at the price of the option on average at maturity then this equals 3.26, hence it became on average even cheaper than expected at start of the contract. The standard deviation equals 0.41 for the case we follow no hedging strategy at all. Therefore we learn that the mean-variance hedging strategy can hardly beat the strategy of doing nothing. A possibility would be to simulate extreme market conditions, but in contrast to the stock market the risk is always limited, because a zero-coupon bond will never be worth more than 1. This means that for the swaption we are sure the price at maturity will certainly be smaller than 5.

These simulations show that when we apply hedging strategies, we should never forget to look at the original risk of the product we try to hedge and hence compare if by hedging the risk is really reduced. Furthermore we also deduce that the delta-hedge in an almost riskless, predictable market performs better than the mean-variance hedge, whose goal is to avoid losses as well as gains.

In the setting simulated here, the mean-variance hedging strategy is not causing more problems but the strategy is also not reducing the risk enormously, because there is almost no risk to take away. We also illustrated clearly the different goals of the delta-hedge compared to the mean-variance hedge. In the next chapter we will illustrate how in a really risky market the quadratic hedging strategies perform excellent.

Benjamin Franklin may have discovered electricity, but it was the man who invented the meter who made the money.

Earl Wilson (1906-1990)

9 Hedging strategies applied to non-traded assets

In this chapter we concentrate on so-called non-traded assets in the setting of the energy market. We call it non-traded because we assume that the derivative that we want to hedge depends on several assets $S^{(1)}, \ldots, S^{(d)}$, while we can only use a fixed combination S of the assets to hedge. For more explanation concerning this non-traded assets and why they are important in the energy market, we refer to Section 9.1.
In the literature, see Section 9.2, there are only few articles concerning non-traded assets and often the setting is different from the one we assume here.
It will become clear that in the proposed setting it is not possible to determine a classical delta-hedge, see Section 9.3. Therefore in practice adjusted delta-hedges are often applied, while we suggest to use quadratic hedging strategies instead.
The theoretical determination of the locally risk-minimizing hedging strategy, see Chapter 4, in this context is rather straightforward once one understands properly the existence conditions and how it is determined. An important concept for LRM hedging strategies is the minimal martingale measure (MMM), see Section 2.3.2.1. Concentrating on our setting, it is important to understand that we always need to take the martingale measure with respect to the underlying asset which is used for hedging and NOT the martingale measure with respect

to the assets on which the claim depends. More concrete we need to determine the MMM of S and not of $S^{(1)}, \ldots, S^{(d)}$. This obviously demonstrates the incompleteness of the market we work in, even in the case with two driving Brownian motions (there is only one martingale measure for the vector $(S^{(1)}, S^{(2)})$, but infinitely many for S).

Assuming exponential Lévy processes to generate the underlying assets leads to a stochastic mean-variance tradeoff, see Definition 2.2.18. Due to this stochastic mean-variance tradeoff process we can no longer determine the number of risky assets in terms of the cumulant function if we assume the processes are not martingales, but semimartingales, see Section 9.4.2. Therefore we consider specific processes in Section 9.4.3 to compute the optimal number explicitly. In this way we can apply LRM hedging strategies to the setting of non-traded assets. However because we are not able to express the optimal number in terms of the cumulant function, we need to solve PIDE's and therefore the simulations will take more time. We will circumvent this issue by implementing an adjusted locally risk-minimizing hedging strategy, without solving PIDE's.

Another option would be to apply the mean-variance hedging strategy, but unfortunately in the setting of non-traded assets this becomes complicated even in the simple case that the underlyings are driven by continuous processes as we show in Section 9.5.

In Section 9.6 we give the results of our preliminary experiments. We compare the different current market practices to the newly obtained results and show that the (adjusted) locally risk-minimizing hedging strategies outperform, even in the simplest cases. This can also be found in Leoni et al. (2010). We make the extra assumption that the interest rate equals zero, hence the discounted asset price equals the non-discounted price. This is important because the locally risk-minimizing hedging strategy defines the hedging strategy in terms of the discounted assets. Moreover the LRM hedging strategy is not self-financing such that the knowledge of the optimal number of the discounted assets not automatically implies knowledge of the optimal number of the non-discounted ones. Furthermore this assumption is not really restrictive, in view of the low interest rates on the market nowadays.

In this chapter we clearly illustrate that although theoretically the quadratic hedging strategy for a specific and at first sight simple setting may exist, the concrete implementation can be really involved.

9.1 Non-traded assets

Hedging under restrictions is a problem of great practical importance. It can become relevant in all financial markets but it is extremely important in energy markets where liquidity can be poor. We will apply the theoretical results concerning local risk-minimization to a setting that is a good representation of every-day practice in an energy market. The obtained locally risk-minimizing hedging strategy is compared to a few common practices and shown to outperform significantly.
Since energy such as electricity or gas are non storable commodities, the trading market has been organised around futures and forwards. Those provide an agreement between the two transacting parties to deliver the commodity over a fixed period of time rather than ensuring an instantaneous delivery. This means that the variety of delivery periods is enormous and although the correlation between those is not always strong, usually only a few contracts are liquid enough to execute trading strategies.
Since the underlying asset is a commodity that gets delivered physically in a certain volume, it is natural that the prices are denominated in currency per unit of volume per unit of time. Over the years, energy markets have attained a specific structure suitable for handling this flow nature. One of the unique features of the forward curve is its decomposition, or bucketing, into different granularities. Far ahead into the future, the only forward contracts traded are forwards for delivery of power over a complete calendar year. Once the calendar year approaches, these contracts gradually break down into quarterly contracts in a 'cascading' process. Closer to delivery, these quarterly contracts will break up, into monthly, weekly and even daily forward contracts.
We discuss this setting in more detail for the electricity market and for the gas market separately.

(1) Electricity market:

Besides the delivery period during which European Power or electricity gets delivered, one often distinguishes three different products: base, peak and off-peak. Those are best explained by means of an example. A CAL-11 peak product is a contract that will deliver electricity during the entire calendar year 2011, but only during the peak hours of the day. That means that during the weekend no power will be delivered and during the weekdays, the delivery only takes place during the day (peak hours). A base contract ensures delivery of power during every single hour of the delivery period, without exception. The specific definition of which hour is a peak hour or off-peak hour, depends on the market in question.

As indicated, prices are denominated in currency per unit of volume per unit of time. For example a CAL-11 base product for a volume of 10MW (megawatt) could have a price of €50/MWh. This means that per delivered hour and per MW, the price is €50. There are 8760 hours in the year 2011, so the total premium that is paid for the delivered power is €$50 \times 8760 \times 10 =$ €$4\ 380\ 000$.

A peak contract for the same delivery period 2011, could have a price of €80/MWh, which is higher per active unit, but since the number of active hours is much less, the total premium is still lower compared to the base contract. Roughly speaking the number of peak hours is one third of the amount of base hours. This means that the premium is €$80 \times 8760/3 \times 10 =$ €$2\ 336\ 000$.

It is clear that there is a relationship between the peak, off-peak and base price in the market. If one buys electricity for delivery during peak hours and at the same time buys a contract that ensures delivery during the off-peak hours, it is obvious that the power is delivered without interruption and this is equivalent to a base contract.

If we denote the price of a peak contract by $S^{(p)}$ and the price of an off-peak contract as $S^{(o)}$, then the price of the base contract $S^{(b)}$ is given by

$$S^{(b)} = \omega^{(p)} S^{(p)} + \omega^{(o)} S^{(o)}$$

where the weights $\omega^{(p)}$ and $\omega^{(o)}$ depend on the number of peak and off-peak hours of the market. Since all prices are normalised to one unit of power and one hour, the typical weights are $\omega^{(p)} = 1/3$ and $\omega^{(o)} = 2/3$, where we should never forget that the actual cash flows will take into account the number of hours.

In terms of liquidity, peak or off-peak contracts are not as liquid as base, especially far ahead in the future. This has its implications for the writer of an option on a peak forward contract. Although it is one of the very basic assumptions in the derivatives theory, energy market traders often find themselves into a situation where they sell options although the underlying contract is not liquidly traded. In the application of the theory we will assume that a claim on peak power is transacted. It could be that at the time of this transaction, the value of this peak contract is known but that the spread between the bid price and the offer price is too big to efficiently delta-hedge this position.

For this reason, it is common practice to hedge with a base product rather than a peak product until closer to maturity of the option, the strategy is swapped to a strategy in the peak product when the liquidity has increased. We will call the peak product non-tradable for reasons of illiquidity and in Section 9.6 we will study the effect of different hedging strategies.

(2) Gas market:

The gas market is highly seasonal with summer prices usually substantially cheaper than winter prices. Because of this, it is easy to understand that the most liquid products are forward contracts for the delivery of gas during the winter or during the summer. Gas by itself is more storable than power because the pressure difference in the network allow for the variations in demand during the day. Because of this, there is no peak and off-peak market for gas.
The interest in option contracts in the European gas market (e.g. UK or NBP market) is increasing since these kind of contracts provide an interesting way of hedging the portfolios of big gas players in the market. However, for historical reasons, the most liquid option contract is a so-called seasonal option, that actually consists of a strip of 6 monthly options. So there are 6 underlying levels $F^{(1)}, F^{(2)}, \ldots, F^{(6)}$ that are relevant for the pricing and hedging of such a contract.
However, at the time when the options are written, not all of these monthly forward prices are known and the hedging has to be done by means of the season forward, which is in fact given by

$$S = \sum_{i=1}^{6} \omega^{(i)} F^{(i)}.$$

For the remaining of the chapter, the fixed combination in which we can invest, will be denoted by

$$S = S_0 + M + B = \sum_{i=1}^{d} w^{(i)} S^{(i)},$$

with martingale part M and finite variation part B and $S^{(i)} = S_0^{(i)} + M^{(i)} + B^{(i)}$ the discounted assets. We will denote by $F(t, \mathbf{S}_t)$ the value of the claim at time t under the minimal martingale measure. We remark that the filtration $\mathbb{F}$ in this section contains the information of the non-traded assets and not only the information of the traded asset. In the sense that the filtration will be generated by the underlying driving processes.

9.2 Literature

We found only few articles dealing with non-traded assets. The setting used in these articles differs from the one we work with, because they start from a different underlying problem. Furthermore, they all concentrate on the continuous setting, while we also allow discontinuous processes.
In literature, see e.g. Davis (2006), the term basis risk is often used for the non-hedgeable risk which remains and cannot be hedged away due to the fact that the asset on which the option is written is not available for hedging. Hedging in this case can only be done by using some closely related asset. Sometimes the underlying asset is available for hedging but is too expensive due to transactions costs. Hence in their setting the determination of the quadratic hedging strategies is less complicated, due to the possible deterministic mean-variance tradeoff process.
We mention here some important papers, the interested reader can also look at the references in those papers.

- Davis (2006) assumes that the underlying asset cannot be traded but is observable. Instead 'a closely related' asset, with a continuous price process, is traded. This closely related asset is assumed to follow a Brownian motion which is correlated with the underlying risky asset. The optimal hedging strategy is determined using exponential utility as a criterion. Numerical results for this case are derived in Monoyios (2004).
- Henderson (2002) and Henderson and Hobson (2002) work in the same setting. Futhermore they also determine the power utility and give numerical results.
- Hobson (2005) gives an upper bound for the utility indifference price of a contingent claim on a non-traded asset again in the setting described by Davis (2006).
- Ankirchner et al. (2010) also calculate the exponential utility-based indifference prices and corresponding hedges in a continuous setting. Their

results are obtained in terms of solutions of forward-backward stochastic differential equations. Hence the optimal hedging strategies are described in terms of the indifference price gradient and the correlation coefficients. Furthermore the hedge can be seen as a generalization of the 'delta-hedge' in complete markets.

- Horst et al. (2010) concentrate on transferring non-financial risk, as for example depending on the temperature, to the capital markets. They give numerical results of equilibirum prices and optimal utilities in a continuous framework.

9.3 Strategy derived from the delta-hedge

In practice, non-traded assets are often hedged using a strategy based on the delta-hedge. We will use the intuitively obtained hedging strategies to compare them with the locally risk-minimizing hedging strategy.
For this section we restrict to the two-dimensional case but we can easily extend the obtained strategies to more dimensions.
In the standard two-dimensional case, we trade $\xi^{(i)}$ assets $S^{(i)}$, $i = 1,2$ such that the risk originating from the rate of change of the claim price with respect to the asset prices equals

$$\big[\frac{\partial F}{\partial S^{(1)}} - \xi^{(1)}\big]dS^{(1)} + \big[\frac{\partial F}{\partial S^{(2)}} - \xi^{(2)}\big]dS^{(2)}. \tag{9.1}$$

This risk can be completely eliminated by choosing $\xi^{(i)}$ equal to $\frac{\partial F}{\partial S^{(i)}}$, $i = 1, 2$.
For non-traded assets we can only invest in ξ assets S. It is impossible to eliminate the risk exposure completely because the following equations should be satisfied by ξ

$$\xi^{(1)} = w^{(1)}\xi \quad \text{and} \quad \xi^{(2)} = w^{(2)}\xi$$

and so we have to search for the most optimal ξ. We give some intuitively based solutions for ξ:

- Volume-neutral ξ:

$$\xi = w^{(1)}\frac{\partial F}{\partial S^{(1)}} + w^{(2)}\frac{\partial F}{\partial S^{(2)}}.$$

- Price-adjusted ξ:

$$\xi = \frac{w^{(1)}S^{(1)}\frac{\partial F}{\partial S^{(1)}} + w^{(2)}S^{(2)}\frac{\partial F}{\partial S^{(2)}}}{w^{(1)}S^{(1)} + w^{(2)}S^{(2)}}.$$

- Delta hedging with minimal risk exposure:
  Restricting (9.1) to the setting of non-traded assets and calculating the differential of the portfolio consisting of the claim and ξ assets S, we find the following:

$$\Big[\frac{\partial F}{\partial S^{(1)}} - \xi w^{(1)}\Big]dS^{(1)} + \Big[\frac{\partial F}{\partial S^{(2)}} - \xi w^{(2)}\Big]dS^{(2)}.$$

  Remark we left out the dt-part because this is not the risky part.
  The variance of this remaining risk is in vector notation:

$$\mathrm{var}(\xi) = \Big[\frac{\partial F}{\partial \mathbf{S}} - \xi \mathbf{w}\Big]' d\langle \mathbf{S}, \mathbf{S}\rangle^P \Big[\frac{\partial F}{\partial \mathbf{S}} - \xi \mathbf{w}\Big],$$

  where $\frac{\partial F}{\partial \mathbf{S}}$ is the gradient of F, $\mathbf{w}$ is the vector containing the weights and $\mathbf{S} = (S^{(1)}, S^{(2)})$. We minimize this variance to obtain the optimal ξ:

$$\frac{d\mathrm{var}(\xi)}{d\xi} = -\mathbf{w}' d\langle \mathbf{S}, \mathbf{S}\rangle^P [\frac{\partial F}{\partial \mathbf{S}} - \xi \mathbf{w}] - [\frac{\partial F}{\partial \mathbf{S}} - \xi \mathbf{w}]' d\langle \mathbf{S}, \mathbf{S}\rangle^P \mathbf{w} = 0.$$

  Solving this equation for ξ, gives the following result

$$\xi = \frac{\frac{\partial F}{\partial \mathbf{S}}' d\langle \mathbf{S}, \mathbf{S}\rangle^P \mathbf{w}}{\mathbf{w}' d\langle \mathbf{S}, \mathbf{S}\rangle^P \mathbf{w}}.$$

  This is exactly the result we will obtain when we apply the LRM hedging theory. So we achieved here an intuitive explanation for the rather complicated theory of local risk-minimization. We remark that we cannot follow blindly this intuitive approach, because in some case e.g. if the finite variation part is no longer continuous, it makes no longer sense to try to minimize the risk involved.
  We note that independently Poulsen et al. (2009) made an analogous conclusion.

These intuitively based solutions will be used for comparison with the solutions to the local risk-minimization.

9.4 Locally risk-minimizing hedging strategy

9.4.1 Setting

In this section we list all the assumptions under which we assume to work.
We work on the probability space $(\Omega, \mathcal{F}, P)$. The filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ satisfies all the usual conditions and $T \in [0, +\infty)$ is the fixed time horizon.

Assumptions 9.4.1. *We assume the filtration is Markovian.*

The process of the riskless asset is given by B and $B_t = 1$. Therefore the discounted asset prices equal the undiscounted prices.

Assumptions 9.4.2. *We assume that the $\mathbb{R}^d$-valued process S is a special square-integrable semimartingale under the original measure P and there exists a predictable process λ such that*

$$dB_t = \lambda_t d\langle M, M\rangle_t^P, \quad t \in [0, T].$$

Furthermore, we will also assume that $E[\int_0^T |\lambda_u|' d\langle M\rangle_u |\lambda_u|] < \infty$.

The following assumption is needed in order that the locally risk-minimizing hedging strategy exists, see Chapter 4.

Assumptions 9.4.3. *We assume that the finite variation part is continuous.*

Hence the semimartingale S is quasi-left continuous, this (restrictive) assumption was not made in Choulli et al. (2010), because the interest of that paper was the determination of the FS decomposition and not the LRM hedging strategy.

Assumptions 9.4.4. *We assume that $\mathcal{E}(-\lambda \cdot M)$ is a strictly positive square-integrable martingale. Under this assumption the minimal martingale measure $\widetilde{Q}$ is an equivalent probability measure.*

Due to the results described in Chapter 4 and under the previous assumptions, the determination of the LRM hedging strategy for non-traded assets is reduced to a straightforward application of the formulas given there.
We distinguish two different cases: the continuous case and the discontinuous case.

- In the continuous case the number of risky assets for the LRM hedging strategy equals the one of the RM hedging strategy under the minimal martingale measure $\widetilde{Q}$. Hence

$$\xi = \frac{d\langle F, S\rangle^{\widetilde{Q}}}{d\langle S, S\rangle^{\widetilde{Q}}} = \frac{\frac{\partial F}{\partial \mathbf{S}}' d\langle \mathbf{S}, \mathbf{S}\rangle^{\widetilde{Q}}\mathbf{w}}{\mathbf{w}' d\langle \mathbf{S}, \mathbf{S}\rangle^{\widetilde{Q}}\mathbf{w}}.$$

  Hereto, apply Itô's formula to F and rely on the continuity and martingale property of F.

- In the discontinuous case we apply formula (4.26) and hence the optimal number of risky assets is given by

$$\xi = \frac{d\langle I, M\rangle^P}{d\langle M, M\rangle^P} = \frac{d\langle I, \mathbf{S}^m\rangle^P \mathbf{w}}{\mathbf{w}' d\langle \mathbf{S}^m, \mathbf{S}^m\rangle^P \mathbf{w}}, \tag{9.2}$$

  with I the P-martingale part of F and where M and $\mathbf{S}^m$ stand for the P-martingale part of S, respectively $\mathbf{S}$.

In the next section we try to obtain more explicit results for the LRM hedging strategy in terms of the cumulant functions. We show that due to the specific setting we work in, this is not possible. Hence in Section 9.4.3 we make the amount more explicit by filling in the dynamics of the processes. This means that the price under the minimal martingale measure will be determined by solving a partial integro differential equation (PIDE).

9.4.2 LRM hedging strategy in terms of the cumulant function

In Chapter 8 we calculated the optimal amount of risky assets in terms of the cumulant function extending ideas of Hubalek et al. (2006). The use of cumulant functions allows to calculate the optimal number really fast by means of Fourier transformation. Also in the setting of non-traded assets we would like to express the optimal number in terms of the cumulant function of the underlying. We will show we cannot apply the approach of Hubalek et al. (2006), as we did in Chapter 8, due to the specific setting in this chapter. We emphasize that this does not mean we are not able to determine the locally risk-minimizing hedging strategy in the setting of non-traded assets, but we can only do it in a

longer computational time.
In the setting of Hubalek et al. (2006) exponential Lévy processes are used to model the underlyings, because for these type of processes martingales are easily found by compensating the process with the cumulant function. For simplicity we assume we only have two underlying assets, namely $S^{(1)}$ and $S^{(2)}$:

$$S_t^{(i)} = S_0^{(i)} \exp(X_t^{(i)}).$$

With S we denote the combination $w^{(1)}S^{(1)} + w^{(2)}S^{(2)}$ in which we can invest and with canonical decomposition $S = S_0 + M + B$. Kallsen and Shiryaev (2002) extended the use of cumulants to the class of semimartingales, called (modified) Laplace cumulant process.

Theorem 9.4.5 (Kallsen and Shiryaev (2002) Theorem 2.19)**.** *Let $\theta \in L(X)$ such that $\theta \cdot X$ is exponentially special. Then $K^X(\theta)$ is the exponential compensator of $\theta \cdot X$. More specifically,*

$$Z := \exp(\theta \cdot X - K^X(\theta)) \in \mathcal{M}_{loc}.$$

Furthermore if X is quasi-left-continuous and the process X has characteristics (b_t, c_t, F_t), then from Theorem 2.18 of Kallsen and Shiryaev (2002) the cumulant function can be written more explicitly as

$$K^X(\theta) = \tilde{\kappa}(\theta) \cdot A$$

with $\tilde{\kappa}(\theta)_t := \theta_t' b_t + \frac{1}{2}\theta_t' c_t \theta_t + \int (e^{\theta_t' x} - 1 - \theta_t' h(x)) F_t(dx)$. We denote by $\kappa(z_1, z_2)t$ the cumulant generating function of the joint distribution under the original measure:

$$E[e^{z_1 X_t^{(1)} + z_2 X_t^{(2)}}] := e^{\kappa(z_1, z_2)t}.$$

For a concrete example in which κ for a multivariate variance gamma model is calculated, we refer to Leoni and Schoutens (2008). As a first step we show we can easily obtain $d\langle M, M\rangle$, needed to determine ξ, in terms of the cumulant functions. Using κ and by assuming stationarity and independence of increments for $X^{(i)}$, $i = 1, 2$ we obtain that the process

$$N_t(z_1, z_2) := e^{-\kappa(z_1, z_2)t} (S_t^{(1)})^{z_1} (S_t^{(2)})^{z_2}$$

is a martingale under the original measure:

$$\begin{aligned}
&E[N_T(z_1, z_2)|\mathcal{F}_t] \\
&= E[e^{-\kappa(z_1,z_2)T}(S_T^{(1)})^{z_1}(S_T^{(2)})^{z_2}|\mathcal{F}_t] \\
&= e^{-\kappa(z_1,z_2)T}(S_0^{(1)})^{z_1}e^{z_1X_t^{(1)}}(S_0^{(2)})^{z_2}e^{z_2X_t^{(2)}}E[e^{z_1(X_T^{(1)}-X_t^{(1)})+z_2(X_T^{(2)}-X_t^{(2)})}] \\
&= e^{-\kappa(z_1,z_2)T}(S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2}e^{\kappa(z_1,z_2)(T-t)} \\
&= e^{-\kappa(z_1,z_2)t}(S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2} = N_t(z_1, z_2).
\end{aligned}$$

For certain z_1, z_2 belonging to $\mathbb{C}$, the dynamics of the process $(S^{(1)})^{z_1}(S^{(2)})^{z_2}$ equal

$$\begin{aligned}
d((S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2}) =& d(e^{\kappa(z_1,z_2)t}N_t(z_1, z_2)) \\
=& e^{\kappa(z_1,z_2)t}dN_t(z_1, z_2) + (S_{t-}^{(1)})^{z_1}(S_{t-}^{(2)})^{z_2}\kappa(z_1, z_2)dt,
\end{aligned}$$

where the term $[e^{\kappa(z_1,z_2)\cdot}, N(z_1, z_2)]$ is zero due to the fact that the process $e^{\kappa(z_1,z_2)\cdot}$ is continuous and has finite variation. This implies the following canonical decomposition:

$$(S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2} = (S_0^{(1)})^{z_1}(S_0^{(2)})^{z_2} + M_t(z_1, z_2) + B_t(z_1, z_2),$$

with

$$M_t(z_1, z_2) = \int_0^t e^{\kappa(z_1,z_2)u}dN_u(z_1, z_2), \tag{9.3}$$

$$B_t(z_1, z_2) = \kappa(z_1, z_2)\int_0^t (S_{u-}^{(1)})^{z_1}(S_{u-}^{(2)})^{z_2}du. \tag{9.4}$$

This result will simplify the calculation of the angle brackets:

$$d\langle(S^{(1)})^{z_1}(S^{(2)})^{z_2}, S^{(1)}\rangle \qquad \text{and} \qquad d\langle(S^{(1)})^{z_1}(S^{(2)})^{z_2}, S^{(2)}\rangle.$$

By Itô's formula we rewrite the bracket as follows

$$\begin{aligned}
&[(S^{(1)})^{z_1}(S^{(2)})^{z_2}, S^{(1)}]_t \\
&= (S_t^{(1)})^{z_1+1}(S_t^{(2)})^{z_2} - (S_0^{(1)})^{z_1+1}(S_0^{(2)})^{z_2} \\
&\quad - \int_0^t (S_{u-}^{(1)})^{z_1}(S_{u-}^{(2)})^{z_2}dS_u^{(1)} - \int_0^t S_{u-}^{(1)}d((S_u^{(1)})^{z_1}(S_u^{(2)})^{z_2})
\end{aligned}$$

$$= \int_0^t e^{\kappa(z_1+1,z_2)u} dN_u(z_1+1,z_2) + \int_0^t (S_u^{(1)})^{z_1+1}(S_u^{(2)})^{z_2}\kappa(z_1+1,z_2)du$$
$$- \int_0^t (S_u^{(1)})^{z_1}(S_u^{(2)})^{z_2} e^{\kappa(1,0)u} dN_u^{(1)}(1) - \int_0^t (S_u^{(1)})^{z_1+1}(S_u^{(2)})^{z_2}\kappa(1,0)du$$
$$- \int_0^t S_{u-}^{(1)} e^{\kappa(z_1,z_2)u} dN_u(z_1,z_2) - \int_0^t (S_{u-}^{(1)})^{z_1+1}(S_{u-}^{(2)})^{z_2}\kappa(z_1,z_2)du.$$

Hence by definition of the angle bracket, we find that

$$\frac{d\langle (S^{(1)})^{z_1}(S^2)^{z_2}, S^{(1)}\rangle}{dt} = (S_{t-}^{(1)})^{z_1+1}(S_{t-}^{(2)})^{z_2}(\kappa(z_1+1,z_2) - \kappa(z_1,z_2) - \kappa(1,0)), \tag{9.5}$$

and by symmetry

$$\frac{d\langle (S^{(1)})^{z_1}(S^2)^{z_2}, S^{(2)}\rangle}{dt} = (S_{t-}^{(1)})^{z_1}(S_{t-}^{(2)})^{z_2+1}(\kappa(z_1,z_2+1) - \kappa(z_1,z_2) - \kappa(0,1)). \tag{9.6}$$

Due to Assumption 9.4.3 we know that $\langle M, M\rangle = \langle S, S\rangle$ and by using (9.5) and (9.6) we obtain:

$$\begin{aligned} d\langle M, M\rangle_t/dt =& (w^{(1)}S_{t-}^{(1)})^2(\kappa(2,0) - 2\kappa(1,0)) \\ &+ 2w^{(1)}w^{(2)}S_{t-}^{(1)}S_{t-}^{(2)}(\kappa(1,1) - \kappa(1,0) - \kappa(0,1)) \\ &+ (w^{(2)}S_{t-}^{(2)})^2(\kappa(0,2) - 2\kappa(0,1)). \end{aligned} \tag{9.7}$$

This means that we found an expression for the denominator of (9.2) in terms of the cumulant functions and furthermore the process $\lambda = \frac{dB}{d\langle M,M\rangle}$ equals

$$\begin{aligned} \lambda_t =& \frac{w^{(1)}B_t(1,0) + w^{(2)}B_t(0,1)}{d\langle M, M\rangle_t} \\ =& \frac{w^{(1)}S_{t-}^{(1)}\kappa(1,0) + w^{(2)}S_{t-}^{(2)}\kappa(0,1)}{d\langle M, M\rangle_t}dt, \end{aligned} \tag{9.8}$$

where we used (9.4) and with $d\langle M, M\rangle_t/dt$ as in (9.7). The main idea in Hubalek et al. (2006) was to rewrite the T-contingent claim H as a Fourier transform $\int S_T^z \Pi(dz)$, then they determined the Föllmer-Schweizer decomposition for the components S_T^z with $z \in \mathbb{C}$ and such that $S_T^z \in L^2(P)$. The

Föllmer-Schweizer decomposition of the total claim H is then found by integrating over Π.
In our setting each claim H can depend on $S^{(1)}$ and $S^{(2)}$, hence our claim will be of the form

$$\int (S_T^{(1)})^{z_1}(S_T^{(2)})^{z_2}\Pi(dz_1, dz_2).$$

For more details concerning option valuation for an option on multiple assets by using Fourier transformation, we refer to Theorem 3.2 of Eberlein et al. (2009).
We have to search the Föllmer-Schweizer decomposition for the individual components in the integral: $(S_T^{(1)})^{z_1}(S_T^{(2)})^{z_2}$. As described in (4.25), we first need to determine the process $H_t(z_1, z_2) = E^{\widetilde{Q}}[H_T(z_1, z_2)|\mathcal{F}_t]$, where $\widetilde{Q}$ is the minimal martingale measure for which the Girsanov density describing the change of measure from P to $\widetilde{Q}$ is given by $\mathcal{E}(-\lambda\cdot M)$, see Section 2.3.2.1, and $H_T(z_1, z_2) = (S_T^{(1)})^{z_1}(S_T^{(2)})^{z_2}$.
We show that in the present case it is not possible to determine the process $H(z_1, z_2)$ explicitly in terms of the cumulant function under the minimal martingale measure, due to the fact that now this cumulant function is not deterministic but stochastic.
From Kallsen and Shiryaev (2002) it follows that if the process $z^{(1)}X^{(1)} + z^{(2)}X^{(2)}$ is exponentially special then the cumulant function exists and furthermore it is predictable and of finite variation. By Assumption 9.4.3 we also know that the cumulant function is continuous and of the form $\tilde{\kappa}^{\widetilde{Q}}(z_1, z_2)\cdot t$. Utilizing the properties of the Föllmer-Schweizer decomposition, (9.2) and (2.5), we deduce that

$$dH^{\text{FV}}(z_1, z_2) = \xi dB = \frac{d\langle H^m(z_1, z_2), M\rangle}{d\langle M, M\rangle}dB = d\langle H^m(z_1, z_2), M\rangle\lambda, \tag{9.9}$$

where $H^{\text{FV}}(z_1, z_2)$ is the finite variation part and $H^m(z_1, z_2)$ is the martingale part of the process $H(z_1, z_2)$ under P. Furthermore we make the process $H(z_1, z_2)$ more explicit

$$\begin{aligned} H_t(z_1, z_2) =& E^{\widetilde{Q}}[(S_T^{(1)})^{z_1}(S_T^{(2)})^{z_2}|\mathcal{F}_t] \\ =& (S_0^{(1)})^{z_1}(S_0^{(2)})^{z_2}E^{\widetilde{Q}}[\exp(z_1X_T^{(1)} + z_2X_T^{(2)})|\mathcal{F}_t] \\ =& (S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2}E^{\widetilde{Q}}[\exp(z_1(X_T^{(1)} - X_t^{(1)}) + z_2(X_T^{(2)} - X_t^{(2)}))|\mathcal{F}_t] \\ :=& (S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2}E^{\widetilde{Q}}[e^{\int_t^T \tilde{\kappa}_u^{\widetilde{Q}}(z_1,z_2)du}|\mathcal{F}_t] \end{aligned} \tag{9.10}$$

If $\tilde{\kappa}^{\widetilde{Q}}$ would be deterministic then

$$H_t(z_1,z_2) = (S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2} e^{\int_t^T \tilde{\kappa}_u^{\widetilde{Q}}(z_1,z_2)du} \tag{9.11}$$

and in view of (9.3) and (9.4)

$$dH_t^{\mathrm{FV}}(z_1,z_2) = (S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2} e^{\int_t^T \tilde{\kappa}_u^{\widetilde{Q}}(z_1,z_2)du}(\kappa(z_1,z_2) - \tilde{\kappa}_t^{\widetilde{Q}}(z_1,z_2))dt, \tag{9.12}$$

$$dH_t^m(z_1,z_2) = e^{\int_t^T \tilde{\kappa}_u^{\widetilde{Q}}(z_1,z_2)du} dM_t(z_1,z_2). \tag{9.13}$$

Inserting (9.12) and (9.13) in (9.9) leads to

$$\begin{aligned}
&(S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2}(\kappa(z_1,z_2) - \tilde{\kappa}_t^{\widetilde{Q}}(z_1,z_2))dt \\
&= d\langle M(z_1,z_2), M\rangle_t \lambda_t \\
&= (S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2} \frac{(w^{(1)}S_t^{(1)}\kappa(1,0) + w^{(2)}S_t^{(2)}\kappa(0,1))dt}{d\langle M, M\rangle_t} \\
&\times \Big(w^{(1)}S_t^{(1)}(\kappa(z_1+1,z_2) - \kappa(z_1,z_2) - \kappa(1,0)) \\
&\quad + w^{(2)}S_t^{(2)}(\kappa(z_1,z_2+1) - \kappa(z_1,z_2) - \kappa(0,1)) \Big) dt,
\end{aligned}$$

where the last equality follows from (9.5), (9.6) and (9.8). Therefore

$$\begin{aligned}
\tilde{\kappa}_t^{\widetilde{Q}}(z_1,z_2) =& \kappa(z_1,z_2) - \Big(w^{(1)}S_t^{(1)}(\kappa(z_1+1,z_2) - \kappa(z_1,z_2) - \kappa(1,0)) \\
&+ w^{(2)}S_t^{(2)}(\kappa(z_1,z_2+1) - \kappa(z_1,z_2) - \kappa(0,1)) \Big) \\
&\times \frac{(w^{(1)}S_t^{(1)}\kappa(1,0) + w^{(2)}S_t^{(2)}\kappa(0,1))dt}{d\langle M, M\rangle_t},
\end{aligned}$$

which is not a deterministic function. This means that (9.11) does not hold and we cannot make the process $H(z_1,z_2)$ more explicit than expression (9.10). A different way to proceed would be to make the change of measure explicit. After some calculations the process $H(z_1,z_2)$ equals

$$\begin{aligned}
H_t(z_1,z_2) =& (S_t^{(1)})^{z_1}(S_t^{(2)})^{z_2} E\Bigg[\exp\Bigg(\int_t^T (z_1 - \lambda_u S_{u-}^{(1)} w^{(1)}) dX_u^{(1)} \\
&+ \int_t^T (z_2 - \lambda_u S_{u-}^{(2)} w^{(2)}) dX_u^{(2)} \Bigg) |\mathcal{F}_t \Bigg].
\end{aligned}$$

The integrandum is a stochastic function and we cannot apply Proposition 8.4.6. This clearly shows that in the setting of non-traded we are not capable to express the LRM hedging strategy in terms of the cumulant function.

9.4.3 LRM hedging strategy using PIDE's

For the dynamics of the underlying processes we assume a linear combination of independent Brownian motions for the continuous martingale part. For the implementation of the multidimensional measure, we look at two extreme cases. Firstly, we assume every risky asset has the same jump part. Secondly, we assume complete independent jump parts. Both approaches are meaningful for practice. In the first case, we can say that the jump part models extreme market conditions, such that all the risky assets experience the same movement. This assumption is, especially in the commodity market, acceptable, because almost all prices are mainly influenced by the oil price. In the case of complete independence, we assume that the joint movements are included in the Brownian motion part and if there is a jump, that this jump will only influence one asset at the time.
The last model we will look at is a combination of the two previous ones. So there is a jump part influencing all the assets at the same time and independent of this first jump, we have for each risky asset a Lévy measure independent of the other measures. This method combines the advantages of the two other approaches.
Note that we avoid here possible problems with the dependence structure of Lévy processes by assuming complete independence. A different approach would be to model it by using copulas, see Tankov (2004).

9.4.3.1 Lévy processes with the same jump

We assume the following dynamics for the d-dimensional discounted risky asset:

$$\begin{aligned}\mathbf{S}_t =& \mathbf{S}_0 + \mathbf{B}_t + \int_0^t \operatorname{diag}(\boldsymbol{\sigma}_u)\mathbf{c}_u \operatorname{diag}(\mathbf{S}_u)\boldsymbol{\Theta} d\mathbf{W}_u \\ &+ \int_0^t \operatorname{diag}(\boldsymbol{\sigma}_u)\mathbf{S}_{u-} \int_{\mathbb{R}} x[N(du,dx) - \nu(dx)du],\end{aligned}$$

with

$$\mathbf{S}_t = \begin{pmatrix} S_t^{(1)} & \cdots & S_t^{(d)} \end{pmatrix}'$$
$$\mathbf{B}_t = \begin{pmatrix} B_t^{(1)} & \cdots & B_t^{(d)} \end{pmatrix}' \quad \text{with } B_t^{(i)} \text{ the finite variation part of } S_t^{(i)}$$
$$\boldsymbol{\sigma}_t = \begin{pmatrix} \sigma_t^{(1)} & \cdots & \sigma_t^{(d)} \end{pmatrix}'$$
$$\mathbf{c}_t = \text{diagonal matrix with } (\mathbf{c}_t)_{ii} = c_t^{(i)},\ i = 1, \ldots, d$$
$$\boldsymbol{\Theta}\boldsymbol{\Theta}' = \text{variance-covariance matrix}$$
$$\mathbf{W}_t = \begin{pmatrix} W_t^{(1)} & \cdots & W_t^{(d)} \end{pmatrix}' \quad d \text{ independent standard Brownian motions}$$

and with $N(dt, dx) - \nu(dx)dt$ a one-dimensional compensated Lévy measure. This means that the i^{th} component of $\mathbf{S}$ has dynamics

$$dS_t^{(i)} = dB_t^{(i)} + \sigma_t^{(i)} c_t^{(i)} S_t^{(i)} \sum_{l=1}^{d} \Theta_{il} dW_t^{(l)} + \sigma_t^{(i)} S_{t-}^{(i)} \int_{\mathbb{R}} x[N(dt, dx) - \nu(dx)dt].$$

By $M^{(i)}$ we denote the martingale part of the asset $S^{(i)}$:

$$M_t^{(i)} = \int_0^t \sigma_u^{(i)} c_u^{(i)} S_u^{(i)} \sum_{l=1}^{d} \Theta_{il} dW_u^{(l)} + \int_0^t \sigma_u^{(i)} S_{u-}^{(i)} \int_{\mathbb{R}} x[N(du, dx) - \nu(dx)du]$$

constituting the vector $\mathbf{M}$.
To calculate the optimal number to invest in the risky asset S we proceed in several steps:

- Using the Markov property of the risky assets, we obtain

  $$Y_t = E^{\widetilde{Q}}[H(\mathbf{S}_T)|\mathcal{F}_t] = F(t, \mathbf{S}_t). \tag{9.14}$$

  We introduce the notation $\tilde{J}$ for

  $$\tilde{J}(t, x) = F(t, \mathbf{S}_{t-} + \text{diag}(\boldsymbol{\sigma}_t)\mathbf{S}_{t-}x) - F(t, \mathbf{S}_{t-}).$$

- Determination of the dynamics under the MMM
  According to Assumption 9.4.2 and Definition 2.3.6, the Girsanov density describing the change of measure from the original measure to the

minimal martingale measure linked with the process S equals

$$\mathcal{E}(-\lambda \cdot M) = \mathcal{E}(-\frac{\mathbf{w}'\mathbf{B}}{\mathbf{w}' d\langle \mathbf{M}, \mathbf{M}\rangle \mathbf{w}} \mathbf{w} \cdot \mathbf{M}),$$

with

$$\frac{d\langle M^{(i)}, M^{(j)}\rangle_t}{dt} = \sigma_t^{(i)} S_t^{(i)} \sigma_t^{(j)} S_t^{(j)} ((\Theta\Theta')_{ij} c_t^{(i)} c_t^{(j)} + \int_{\mathbb{R}} x^2 \nu(dx)), \quad 1 \leq i, j \leq d. \quad (9.15)$$

Under the minimal martingale measure and by using the 'classical' Girsanov's theorem (see Theorem 2.3.4), we deduce that for every P-local martingale Y the process $Y' = Y + \lambda \cdot \langle Y, M\rangle$ is a $\widetilde{Q}$-local martingale. Therefore the following Brownian motions are martingales:

$$\begin{aligned}\mathbf{W}^{\widetilde{Q}} =& \mathbf{W} + \lambda \cdot \langle M, \mathbf{W}\rangle = \mathbf{W} + \lambda \sum_{i,l=1}^{d} w^{(i)} \sigma^{(i)} c^{(i)} S^{(i)} \Theta_{il} \cdot \langle W^{(l)}, \mathbf{W}\rangle \\ =& \mathbf{W} + \lambda \mathbf{\Theta}' \mathrm{diag}(\boldsymbol{\sigma}) \mathbf{c} \mathrm{diag}(\mathbf{S}) \mathbf{w} \cdot \mathbb{1}_{d\times d}, \qquad (9.16)\end{aligned}$$

with $\mathbb{1}_{d\times d}$ the $d \times d$-matrix with the identical function on the diagonal. The compensator of the Lévy measure equals

$$\begin{aligned}\nu^{\widetilde{Q}}(dx)dt = \nu(dx)dt - \lambda_t \langle N(d\cdot, dx) - \nu(dx)d\cdot,& \\ \mathbf{w}' \int_0 \mathrm{diag}(\boldsymbol{\sigma}_u) \mathbf{S}_u \int_{\mathbb{R}} x[N(du, dx) - \nu(dx)du]\rangle_t& \\ = (1 - \lambda_t \mathbf{w}' \mathrm{diag}(\boldsymbol{\sigma}_t) \mathbf{S}_t x) \nu(dx) dt.& \qquad (9.17)\end{aligned}$$

The dynamics of the processes $S^{(i)}$ under the minimal martingale measure are

$$dS_t^{(i)} = \sigma_t^{(i)} c_t^{(i)} S_t^{(i)} \sum_{l=1}^{d} \Theta_{il} dW_t^{(l),\widetilde{Q}} + \sigma_t^{(i)} S_{t-}^{(i)} \int_{\mathbb{R}} x[N(dt, dx) - \nu^{\widetilde{Q}}(dx)dt].$$

- Next we apply Itô's formula to the $\widetilde{Q}$-martingale F:

$$
\begin{aligned}
F(t,\mathbf{S}_t) = & \int_0^t F_t(u,\mathbf{S}_u)du \\
& + (F_{\mathbf{s}}(\cdot,\mathbf{S}_-)\cdot \mathbf{S})_t + \frac{1}{2}\sum_{i,j=1}^{d}(F_{s^{(i)}s^{(j)}}(\cdot,\mathbf{S})\cdot[S^{(i)},S^{(j)}]^c)_t \quad (9.18)\\
& + \int_0^t \int_{\mathbb{R}} \{\tilde{J}(u,x) - F_{\mathbf{s}}(u,\mathbf{S}_{u-})\boldsymbol{\sigma}'_u \mathbf{S}_{u-}x\}N(du,dx),
\end{aligned}
$$

where $F_{\mathbf{s}}$ is the gradient of F and $F_{\mathbf{ss}}$ is the Hessian matrix of F with respect to $\mathbf{S}$.
We rewrite this $\widetilde{Q}$-martingale, by denoting with Θ_k the k^{th} column of the matrix $\boldsymbol{\Theta}$, as

$$
\begin{aligned}
& dF(t,\mathbf{S}_t) \\
& = F_t(t,\mathbf{S}_t)dt \\
& + \frac{1}{2}\sum_{k=1}^{d}(\text{diag}(\boldsymbol{\sigma}_t)\mathbf{c}_t\mathbf{S}_t)'(\text{diag}(\Theta_k)F_{\mathbf{ss}}(t,\mathbf{S}_{t-})\text{diag}(\Theta_k))(\text{diag}(\boldsymbol{\sigma}_t)\mathbf{c}_t\mathbf{S}_t)dt \\
& + \boldsymbol{\sigma}'_t\mathbf{S}_{t-}\int_{\mathbb{R}}\tilde{J}(t,x)\nu^{\widetilde{Q}}(dx)dt \\
& + F_{\mathbf{s}}(t,\mathbf{S}_t)d(\mathbf{S}_t)^{c,\widetilde{Q}} + \boldsymbol{\sigma}'_t\mathbf{S}_{t-}\int_{\mathbb{R}}\tilde{J}(t,x)[N(dt,dx)-\nu^{\widetilde{Q}}(dx)dt].
\end{aligned}
$$

This means that we need to solve the following PIDE problem to obtain the price of the claim under the minimal martingale measure

$$
\begin{cases}
F_t(t,\mathbf{s}) + \frac{1}{2}\sum_{k=1}^{d}(\text{diag}(\boldsymbol{\sigma}_t)\mathbf{c}_t\mathbf{s})'(\text{diag}(\Theta_k)F_{\mathbf{ss}}(t,\mathbf{s})\text{diag}(\Theta_k))(\text{diag}(\boldsymbol{\sigma}_t)\mathbf{c}_t\mathbf{s}) \\
+\boldsymbol{\sigma}'_t\mathbf{s}\int_{\mathbb{R}}\tilde{J}(t,x)\nu^{\widetilde{Q}}(dx) = 0 \\
F(T,\mathbf{s}) = H(\mathbf{s}).
\end{cases}
$$

From (9.18) it follows that the P-martingale part I of the process Y (9.14) has the following dynamics:

$$
I_t = (F_{\mathbf{s}}\cdot\mathbf{S}^c)_t + \int_0^t \boldsymbol{\sigma}'_u\mathbf{S}_u\int_{\mathbb{R}}\tilde{J}(u,x)(N(du,dx)-\nu(dx)du).
$$

We need to search the Galtchouk-Kunita-Watanabe decomposition of this process with respect to the one-dimensional process M, the martingale part of S:

$$dI_t = \theta_t dM_t + dL_t,$$

where L is a P-martingale orthogonal to the P-martingale M. According to (9.2) the optimal number of risky assets θ is given by

$$\theta_t = \sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\sigma_t^{(i)} S_t^{(i)} \sigma_t^{(j)} S_t^{(j)} w^{(j)}}{d\langle M, M\rangle_t} [F_{s^{(i)}}(t, \mathbf{S}_t)(\mathbf{\Theta}\mathbf{\Theta}')_{ij} c_t^{(i)} c_t^{(j)} + \int_{\mathbb{R}} \tilde{J}(t,x) x\nu(dx)] dt,$$

with $d\langle M, M\rangle = \mathbf{w}' d\langle \mathbf{M}, \mathbf{M}\rangle \mathbf{w}$ and with matrix $d\langle \mathbf{M}, \mathbf{M}\rangle$ as in (9.15).

9.4.3.2 Independent jumps in the Lévy process

This section is very similar to the previous one and therefore we only repeat the parts which change due to the different dynamics for the process $\mathbf{S}$. We assume the following dynamics for the d-dimensional discounted risky asset:

$$\begin{aligned}\mathbf{S}_t =& \mathbf{S}_0 + \mathbf{B}_t + \int_0^t \mathrm{diag}(\boldsymbol{\sigma}_u)\mathbf{c}_u \mathrm{diag}(\mathbf{S}_u)\mathbf{\Theta} d\mathbf{W}_u \\ &+ \int_0^t \mathrm{diag}(\boldsymbol{\sigma}_u)\mathrm{diag}(\mathbf{S}_{u-}) \int_{\mathbb{R}} x[\mathbf{N}(du, dx) - \boldsymbol{\nu}(dx)du],\end{aligned}$$

with $\mathbf{N}(dt, dx) - \boldsymbol{\nu}(dx)dt$ a d-dimensional vector consisting of one-dimensional independent Lévy measures. This means that the i^{th} component of $\mathbf{S}$ is

$$dS_t^{(i)} = dB_t^{(i)} + \sigma_t^{(i)} c_t^{(i)} S_t^{(i)} \sum_{l=1}^{d} \Theta_{il} dW_t^{(l)} + \sigma_t^{(i)} S_{t-}^{(i)} \int_{\mathbb{R}} x[N^{(i)}(dt, dx) - \nu^{(i)}(dx)dt].$$

We start with the determination of the dynamics of the process $F(t, \mathbf{S}_t)$ as defined in (9.14). In the calculations we will use the notation $\overline{J}$ for

$$\overline{J}^{(i)}(t,x) = F\left(t, \mathbf{S}_{t-} + \sigma_t^{(i)} S_{t-}^{(i)} x \mathbf{e}_i\right) - F(t, \mathbf{S}_{t-}),$$

with $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ having 1 on the i^{th} place.
The matrix $(d\langle \mathbf{M}, \mathbf{M}\rangle/dt)$ has components

$$\frac{d\langle M^{(i)}, M^{(j)}\rangle_t}{dt} = = \sigma_t^{(i)} S_t^{(i)} \sigma_t^{(j)} S_t^{(j)} \left((\Theta\Theta')_{ij} c_t^{(i)} c_t^{(j)} + \int_{\mathbb{R}} x^2 \delta_{ij} \nu^{(i)}(dx) \right),$$

with δ_{ij} the Kronecker delta. The processes (9.16) are again Brownian motions under the minimal martingale measure, but the compensators of the Lévy measures under the MMM are now equal to

$$\begin{aligned} &\nu^{(i),\widetilde{Q}}(dx)dt \\ &= \nu^{(i)}(dx)dt - \lambda_t \langle N^{(i)}(d\cdot, dx), \\ &\qquad \mathbf{w}' \int_0 \operatorname{diag}(\boldsymbol{\sigma}_u)\operatorname{diag}(\mathbf{S}_u) \int_{\mathbb{R}} x[\mathbf{N}(du, dx) - \boldsymbol{\nu}(dx)du]\rangle_t \\ &= (1 - \lambda_t w^{(i)} \sigma_t^{(i)} S_t^{(i)} x)\nu^{(i)}(dx)dt, \ i = 1, \ldots, d. \end{aligned} \tag{9.19}$$

The PIDE problem we need to solve becomes

$$\begin{cases} F_t(t, \mathbf{s}) + \dfrac{1}{2}\displaystyle\sum_{k=1}^d (\operatorname{diag}(\boldsymbol{\sigma}_t)\mathbf{c}_t\mathbf{s})'(\operatorname{diag}(\Theta_k)F_{\mathbf{ss}}(t, \mathbf{s})\operatorname{diag}(\Theta_k))(\operatorname{diag}(\boldsymbol{\sigma}_t)\mathbf{c}_t\mathbf{s}) \\ + \displaystyle\sum_{i=1}^d \sigma_t^{(i)} s^{(i)} \int_{\mathbb{R}} \overline{J}^{(i)}(t, x)\nu^{(i),\widetilde{Q}}(dx) = 0 \\ F(T, \mathbf{s}) = H(\mathbf{s}). \end{cases}$$

The P-martingale part I of the process F equals

$$I_t = (F_{\mathbf{s}} \cdot \mathbf{S}^c)_t + \sum_{i=1}^d \int_0^t \int_{\mathbb{R}} \overline{J}^{(i)}(u, x)\sigma_u^{(i)} S_{u-}^{(i)} (N^{(i)}(du, dx) - \nu^{(i)}(dx)du).$$

Therefore the optimal number of risky assets is given by

$$\begin{aligned} \theta_t = \sum_{i=1}^d \sum_{j=1}^d \frac{\sigma_t^{(i)} S_t^{(i)} \sigma_t^{(j)} S_t^{(j)} w^{(j)}}{d\langle M, M\rangle_t} &[F_{s^{(i)}}(t, \mathbf{S}_t)(\Theta\Theta')_{ij} c_t^{(i)} c_t^{(j)} \\ &+ \int_{\mathbb{R}} \overline{J}^{(i)}(t, x) x \delta_{ij} \nu^{(i)}(dx)]dt. \end{aligned}$$

9.4.3.3 Combination of the two approaches

We assume the following dynamics for the d-dimensional discounted risky asset:

$$\begin{aligned}\mathbf{S}_t =&\mathbf{S}_0 + \mathbf{B}_t + \int_0^t \text{diag}(\boldsymbol{\sigma}_u)\mathbf{c}_u \text{diag}(\mathbf{S}_u)\boldsymbol{\Theta} d\mathbf{W}_u \\ &+ \int_0^t \text{diag}(\boldsymbol{\sigma}_u)\mathbf{S}_{u-} \int_{\mathbb{R}} x[N(du,dx) - \nu(dx)du] \\ &+ \int_0^t \text{diag}(\boldsymbol{\sigma}_u)\mathbf{g}_u \text{diag}(\mathbf{S}_{u-}) \int_{\mathbb{R}} x[\mathbf{N}(du,dx) - \boldsymbol{\nu}(dx)du].\end{aligned}$$

We used the same notations as in the previous two sections and added a process described by a diagonal matrix $\mathbf{g}$ with $(g)_{ii} = g^{(i)}$ to allow for more variation between the two independent Lévy measures N and $\mathbf{N}$, that by assumption do not jump at the same time. The notation $\overline{J}$ will here be used for

$$\overline{J}^{(i)}(t,x) =F(t, \mathbf{S}_{t-} + \sigma_t^{(i)} g_t^{(i)} S_{t-}^{(i)} x\mathbf{e}_i) - F(t, \mathbf{S}_{t-}).$$

The price is obtained by solving the following PIDE problem:

$$\begin{cases} F_t(t,\mathbf{s}) + \dfrac{1}{2}\displaystyle\sum_{k=1}^{d}(\text{diag}(\boldsymbol{\sigma}_t)\mathbf{c}_t\mathbf{s})'(\text{diag}(\Theta_k)F_{\mathbf{ss}}(t,\mathbf{s})\text{diag}(\Theta_k))(\text{diag}(\boldsymbol{\sigma}_t)\mathbf{c}_t\mathbf{s}) \\ +\boldsymbol{\sigma}_t'\mathbf{s}\displaystyle\int_{\mathbb{R}} \tilde{J}(t,x)\nu^{\tilde{Q}}(dx) + \sum_{i=1}^{d}\sigma_t^{(i)} s^{(i)} g_t^{(i)} \int_{\mathbb{R}} \overline{J}^{(i)}(t,x)\nu^{(i),\tilde{Q}}(dx) = 0 \\ F(T,\mathbf{s}) = H(\mathbf{s}), \end{cases}$$

with $\nu^{\tilde{Q}}$, $\nu^{(i),\tilde{Q}}$ described in (9.17) respectively (9.19), where the function λ equals

$$\lambda = \frac{\mathbf{w}'\mathbf{B}}{\mathbf{w}'d\langle\mathbf{M},\mathbf{M}\rangle\mathbf{w}}$$

with

$$\begin{aligned}&\frac{d\langle M^{(i)}, M^{(j)}\rangle_t}{dt} \\ &= \sigma_t^{(i)} S_t^{(i)} \sigma_t^{(j)} S_t^{(j)}((\Theta\Theta')_{ij} c_t^{(i)} c_t^{(j)} + \int_{\mathbb{R}} x^2 \nu(dx) + \int_{\mathbb{R}} (g_t^{(i)} x)^2 \delta_{ij} \nu^{(i)}(dx)).\end{aligned}$$

For the optimal number of risky assets we obtain

$$\theta_t = \frac{1}{d\langle M, M\rangle_t} \sum_{i=1}^{d} \sum_{j=1}^{d} \sigma_t^{(i)} S_t^{(i)} \sigma_t^{(j)} S_t^{(j)} w^{(j)} [F_{s^{(i)}}(t, \mathbf{S}_t)(\Theta\Theta')_{ij} c_t^{(i)} c_t^{(j)} + \int_{\mathbb{R}} \tilde{J}(t,x) x \nu(dx) + g_t^{(i)} g_t^{(j)} \int_{\mathbb{R}} \overline{J}^{(i)}(t,x) x \delta_{ij} \nu^{(i)}(dx)] dt.$$

9.5 Mean-variance hedging strategy

Due to Assumption 9.4.4, the convex set of equivalent local martingale measures with square integrable density, denoted by $\mathbb{P}_e^2$, is not empty. Hence the variance-optimal martingale measure exists and is also unique. From Černý and Kallsen (2007) it follows that by choosing the right space of admissible strategies the existence of the MVH strategy is guaranteed.
From Chapter 5 we know that the key to determine the mean-variance hedging is the variance-optimal martingale measure and the expectation under the VOMM of the Girsanov density describing the change of measure from the original measure to the VOMM. Once Z^*, $\hat{Z}^*$ and ς (5.7) are found, we can e.g. use formula (5.9) to obtain the optimal number of risky assets.
Hence this means we first look for the variance-optimal martingale measure. In Section 2.3.2.2 we observed that the VOMM equals the MMM if the mean-variance tradeoff process K, see Definition 2.2.18, is deterministic. Unfortunately, we rarely have a deterministic mean-variance tradeoff in our setting of non-traded assets. As an example we give the mean-variance tradeoff process in the case of two underlyings driven by two independent Brownian motion:

$$\begin{aligned} dS_t^{(1)} &= S_t^{(1)}(b_t^{(1)} dt + \sigma_t^{(1)} dW_t^{(1)}) \\ dS_t^{(2)} &= S_t^{(2)}(b_t^{(2)} dt + \sigma_t^{(2)}(\rho dW_t^{(1)} + \sqrt{1-\rho^2} dW_t^{(2)})). \end{aligned} \tag{9.20}$$

Hence the filtration $\mathbb{F}$ will be generated by the two Brownian motions $W^{(1)}$ and $W^{(2)}$.

For later calculations, we will use the following shorthand notations:

$$\begin{aligned}dS_t =&w^{(1)}dS_t^{(1)} + w^{(2)}dS_t^{(2)}\\ =&(w^{(1)}S_t^{(1)}b_t^{(1)} + w^{(2)}S_t^{(2)}b_t^{(2)})dt + (w^{(1)}\sigma_t^{(1)}S_t^{(1)} + w^{(2)}\sigma^{(2)}S_t^{(2)}\rho)dW_t^{(1)}\\ &+ w^{(2)}S_t^{(2)}\sigma^{(2)}\sqrt{1-\rho^2}dW_t^{(2)}\\ :=&\alpha_t dt + \beta_t^{(1)}dW_t^{(1)} + \beta_t^{(2)}dW_t^{(2)}.\end{aligned}$$

Using (2.5) where B and M are in the sense of Definition 2.2.5, we can easily calculate the mean-variance tradeoff process K, see Definition 2.2.18, with

$$\lambda = \frac{dB}{d\langle M\rangle} = \frac{\alpha}{(\beta^{(1)})^2 + (\beta^{(2)})^2} \tag{9.21}$$

and therefore

$$K_t = \left(\frac{(w^{(1)}S_u^{(1)}b_u^{(1)} + w^{(2)}S_u^{(2)}b_u^{(2)})^2}{(w^{(1)}S_u^{(1)}\sigma_u^{(1)})^2 + 2\rho w^{(1)}w^{(2)}S_u^{(1)}S_u^{(2)}\sigma_u^{(1)}\sigma_u^{(2)} + (w^{(2)}S_u^{(2)}\sigma_u^{(2)})^2} \cdot u\right)_t.$$

Given this stochastic mean-variance tradeoff process, a first attempt to determine the VOMM would be to use the technique described by Černý and Kallsen (2007). They perform first a change of measure to the opportunity neutral measure, which can be seen as the measure neutralizing the effect of the stochastic mean-variance tradeoff process. Then the variance-optimal martingale measure is simply the minimal martingale measure with respect to the opportunity neutral measure. The simplest way to find the opportunity neutral measure is through Theorem 3.25 of Černý and Kallsen (2007). The drawback in our case is that it is really hard to find an L, which is always smaller or equal than 1. Hence we try another approach based on Biagini et al. (2000) and Chan et al. (2009). We determine the variance-optimal martingale measure Q^* using the fact that it should be a martingale measure with $\frac{dQ^*}{dP}$ of the form $c + \gamma \cdot S$ with c a constant $\in [1, \infty[$, see Lemma 2.3.13. Using the continuity of the process S, Assumption 9.4.4 and Theorem 2.3.11, we can rewrite $c + (\gamma \cdot S)_T$ as a Doléans-Dade exponential because it is strictly positive

$$\overline{Z}_T = c\mathcal{E}(\widetilde{\gamma} \cdot S)_T \tag{9.22}$$

with $\widetilde{\gamma}\overline{Z}_- = \gamma$. The Girsanov densities describing the class of equivalent measures $\breve{Q}$ is given by

$$\frac{d\breve{Q}}{dP} = \mathcal{E}(-k \cdot W^{(1)} - l \cdot W^{(2)})_T, \tag{9.23}$$

with functions k and l chosen in such a way that $\frac{d\breve{Q}}{dP}$ is square-integrable. These $\breve{Q}$ are martingale measures if and only if

$$\alpha = k\beta^{(1)} + l\beta^{(2)}. \tag{9.24}$$

Equating the two conditions (9.22) and (9.23), with k and l satisfying (9.24) gives

$$\mathcal{E}(-k \cdot W^{(1)} - l \cdot W^{(2)})_T = c\mathcal{E}(\widetilde{\gamma} \cdot S)_T. \tag{9.25}$$

Theorem 9.5.1. *The density describing the change of measure to the VOMM is given by $\mathcal{E}(-\lambda \cdot S)$, which means that the VOMM equals the MMM.*

Proof. We remark first that when c has to be a positive constant, then the only possible form it can have is $c^{-1} = E[\mathcal{E}(\widetilde{\gamma} \cdot S)_T]$, because only with this choice $E[c\mathcal{E}(\widetilde{\gamma} \cdot S)_T] = 1$, which is a necessary condition for the Girsanov density describing a change of measure.
We define the function $R(t, S_t) := E[\mathcal{E}(\widetilde{\gamma} \cdot S)_T / \mathcal{E}(\widetilde{\gamma} \cdot S)_t | \mathcal{F}_t]$ and we assume that $R(t, S_t)$ is an element of the class $C^{1,2}$. The derivative with respect to the second variable $\frac{\partial R(t,S_t)}{\partial S}$ will be denoted by $R_S(t, S_t)$. Furthermore we also introduce the P-martingale

$$D_t = E[\mathcal{E}(\widetilde{\gamma} \cdot S)_T | \mathcal{F}_t] = \mathcal{E}(\widetilde{\gamma} \cdot S)_t R(t, S_t). \tag{9.26}$$

Using the martingale representation property, we know there exist functions $\psi^{(1)}$ and $\psi^{(2)}$ such that

$$D_t = D_0 + (\psi^{(1)} \cdot W^{(1)})_t + (\psi^{(2)} \cdot W^{(2)})_t. \tag{9.27}$$

Applying Itô's formula to D, (9.26), and using the martingale property and the continuity of all the processes, we obtain:

$$\begin{aligned} dD_t =& \mathcal{E}(\widetilde{\gamma} \cdot S)_t (R(t, S_t)\widetilde{\gamma} + R_S(t, S_t)\beta_t^{(1)}) dW_t^{(1)} \\ &+ \mathcal{E}(\widetilde{\gamma} \cdot S)_t (R(t, S_t)\widetilde{\gamma} + R_S(t, S_t)\beta_t^{(2)}) dW_t^{(2)}. \end{aligned} \tag{9.28}$$

Hence combining (9.28) with (9.27), we get:

$$\frac{D_T}{D_0} = \mathcal{E}\left(\int_0^{\cdot} \frac{1}{D_u}\psi_u^{(1)} dW_u^{(1)} + \int_0^{\cdot} \frac{1}{D_u}\psi_u^{(2)} dW_u^{(2)}\right)_T \tag{9.29}$$

with

$$\begin{aligned}\psi_t^{(1)} =&\mathcal{E}(\widetilde{\gamma}\cdot S)_t(R(t,S_t)\widetilde{\gamma}+R_S(t,S_t))\beta_t^{(1)}\\ \psi_t^{(2)} =&\mathcal{E}(\widetilde{\gamma}\cdot S)_t(R(t,S_t)\widetilde{\gamma}+R_S(t,S_t))\beta_t^{(2)}.\end{aligned} \tag{9.30}$$

From the definition of D, we see that

$$\frac{D_T}{D_0} = \frac{\mathcal{E}(\widetilde{\gamma}\cdot S)_T R(T,S_T)}{\mathcal{E}(\widetilde{\gamma}\cdot S)_0 R(0,S_0)} = \frac{\mathcal{E}(\widetilde{\gamma}\cdot S)_T}{E[\mathcal{E}(\widetilde{\gamma}\cdot S)_T]} = c\mathcal{E}(\widetilde{\gamma}\cdot S)_T. \tag{9.31}$$

So equating (9.29) and (9.31) we obtain k and l from (9.25) using the relations (9.26) and (9.30):

$$k_t = -\left(\widetilde{\gamma}_t + \frac{R_S(t,S_t)}{R(t,S_t)}\right)\beta_t^{(1)} \quad \text{and} \quad l_t = -\left(\widetilde{\gamma}_t + \frac{R_S(t,S_t)}{R(t,S_t)}\right)\beta_t^{(2)}. \tag{9.32}$$

The unknown $\widetilde{\gamma}$ then follows from condition (9.24):

$$\alpha_t + \left(\widetilde{\gamma}_t + \frac{R_S(t,S_t)}{R(t,S_t)}\right)((\beta_t^{(1)})^2 + (\beta_t^{(2)})^2) = 0.$$

By inserting the expression for λ (9.21) in the previous equation we obtain:

$$-\widetilde{\gamma}_t - \frac{R_S(t,S_t)}{R(t,S_t)} = \lambda_t. \tag{9.33}$$

Taking this relation into account, we find for (9.32) that $k_t = \lambda_t\beta_t^{(1)}$ and $l_t = \lambda_t\beta_t^{(2)}$, which means that the VOMM is exactly the MMM. □

We remark that our setting is one of the examples where, although the mean-variance tradeoff process is not deterministic, the VOMM still equals the MMM. The next step would be to apply formula (5.9), but then we need to determine $\widetilde{\gamma}$ explicitly. Solving equation (9.33) is precisely the hard part in the implementation of the MVH strategy. It is also not possible to use formula (5.4) instead because as shown in Example 1 of Schweizer (1996) the deterministic MVT is a necessary condition to obtain this formula. Another way out would be to calculate the expectation of the change of measure under the VOMM, but then we have a similar problem as under the LRM hedging strategy, where we need to find the expected value of the claim under the MMM. This clearly illustrate that by loosing deterministic properties of e.g. the mean-variance tradeoff and the

cumulants, it is much harder to determine the strategies.
In a discontinuous case and under the assumption that the VOMM is strictly positive one can show that $\widetilde{\gamma}$ has to solve the following equation:

$$\alpha_t + \left(\widetilde{\gamma}_t + \frac{R_S(t,S_t)}{R(t,S_t)}\right)((\beta_t^{(1)})^2 + (\beta_t^{(2)})^2) + \int_{\mathbb{R}^2} \frac{R(t,S_{t-}+\omega(\mathbf{x})) - R(t,S_{t-})}{R(t,S_{t-})}(w^{(1)}x^{(1)} + w^{(2)}x^{(2)})\nu(d\mathbf{x}) = 0, \quad (9.34)$$

when the processes of the underlying assets have as dynamics

$$dS_t^{(1)} = S_t^{(1)}(b_t^{(1)}dt + \sigma_t^{(1)}dW_t^{(1)}) + \int_{\mathbb{R}^2} x^{(1)}(N(dt,d\mathbf{x}) - \nu(d\mathbf{x})dt)$$

$$dS_t^{(2)} = S_t^{(2)}(b_t^{(2)}dt + \sigma_t^{(2)}(\rho dW_t^{(1)} + \sqrt{1-\rho^2}dW_t^{(2)})) \quad (9.35)$$

$$+ \int_{\mathbb{R}^2} x^{(2)}(N(dt,d\mathbf{x}) - \nu(d\mathbf{x})dt) \quad (9.36)$$

and the combination $S = w^{(1)}S^{(1)} + w^{(2)}S^{(2)}$. Therefore in a discontinuous case the VOMM will not be equal to the MMM.

9.6 Numerical results

We restrict ourselves to a setting in which the claim is depending on $S^{(1)}$ but the hedging will be done with $S = w^{(1)}S^{(1)} + w^{(2)}S^{(2)}$. We investigate the hedging for an at-the-money call on $S^{(1)}$ as this is the most challenging example. It is well-known that hedging of far in-the-money or out-of-the-money options is relatively easy.
Note that both $S^{(1)}$ and $S^{(2)}$ are contracts for delivery over different periods of time. The delivery period of S itself is hence larger and total premiums should always be adjusted to the delivery period. We take the example of the base/peak problem, see Section 9.1. This means that the weights are roughly speaking $w^{(1)} = 1/3$ for peak and $w^{(2)} = 2/3$ for off-peak power. The cashflow corresponding to a purchase or sale of such a contract is $w^{(i)} \times S^{(i)}$ to adjust correctly for the delivery period.
Hence, it is clear that buying the base contract S delivers power during the peak and off-peak hours, corresponding to a position in both assets $S^{(1)}$ and

$S^{(2)}$. This means that there are two intuitively choices for hedging the claim on peak power $S^{(1)}$ by using base. One could try and focus on the volume risk or on the price risk, see Section 9.3.

9.6.1 Setup

We introduce the following notations: $C(S^{(1)})$ stands for the price of the claim while $\Delta^{(1)} = \frac{\partial C}{\partial S^{(1)}}$ and $\Delta^{(2)} = \frac{\partial C}{\partial S^{(2)}}$ represent its partial derivatives with respect to the peak and off-peak contract prices. Note that in our example the option only depends on $S^{(1)}$ and thus $\Delta^{(2)} = 0$. For convenience we will assume that the interest rate $r = 0$. The amount of risky assets, that are used as a hedge for the claim, is denoted by ξ.

Given the specific nature of this problem, we assume a lifetime of the claim of $T = 3$ years, where for the first $T1 = 0.5$ year, a strategy in the base asset is followed. This is inspired by the fact that at some point, liquidity grows in the peak contract. We call the time $T1$ the roll-over point. We assume that after this time, the claim can be hedged further with a classical delta-hedge or any other hedging strategy on the asset $S^{(1)}$ itself. In a Brownian setting from this point onwards we will hedge perfectly and there is no need in an analysis beyond this point.

The price of the claim at time zero is such that the expected total cost of the strategy is zero, where the price of the option at roll-over time is determined under the unique martingale measure in the Brownian motion case, while in the discontinuous case the mean-correcting martingale measure is used. Due to the zero interest rate, we can restrict ourselves to observing the total cost, see page 198, over the lifetime at the roll-over time. This cost of hedging will be neutralised by the initial premium of the claim. We will show that the uncertainty over the outcome of the different strategies is quite large and therefore we will also study the standard deviation of the hedging cost in those different strategies. The one with the lowest variance is clearly to be preferred in practice.

Both the peak and the off-peak contracts are assumed to follow a geometric Brownian motion:

$$S_t^{(i)} = S_0^{(i)} \exp(\mu^{(i)} t + \sigma^{(i)} W_t^{(i)}), \qquad i = 1, 2$$

where the correlation between the Brownian motions is given by $d \left\langle W^{(1)}, W^{(2)} \right\rangle_t = \rho dt$.

As parameters we choose $\sigma^{(1)} = 40\%$, $\sigma^{(2)} = 30\%$ and $\rho = 75\%$. For the drift we look at two different situations. The first and most easy one is where we assume both assets to be martingales. Hence $\mu^{(i)} = r - 0.5(\sigma^{(i)})^2$ and we will call this the martingale case. In a second example, the semimartingale case, we introduce a drift by setting $\mu^{(1)} = 0.07$ and $\mu^{(2)} = 0.05$. For both cases we will look at the performance of the different strategies.
As starting levels for the prices, we assume that $S^{(1)} =$ €90/MWh and $S^{(2)} =$ €60/MWh, and hence the base asset is worth $S =$ €70/MWh. If we normalize the time of the base contract to one, the cash flows would be given by €70 for baseload of which €30 is coming from the peak contract and €40 from the off-peak. Note that although the price for off-peak is lower, the total cashflow is higher compared to the peak contract because the amount of delivered hours during off-peak is higher.

9.6.2 Different strategies

In this section we repeat the strategies described in Section 9.3 and Section 9.4, but adjusted to the setting described here, namely where the claim only depends on $S^{(1)}$. The two first strategies are extra. The first one we describe is the control strategy, which we cannot follow in practice, while the second one makes more sense due to the specific setting.

9.6.2.1 Control strategy

In order to verify our results, we calculate the classical strategy. This means that we are hedging the claim on $S^{(1)}$ by effectively taking positions in this asset.

9.6.2.2 Total volume-neutral strategy

The number ξ is here equal to

$$\xi = \Delta^{(1)} + \Delta^{(2)} = \Delta^{(1)}.$$

We basically focus on the total volume of the peak contract. If the derivative of the claim C with respect to $S^{(1)}$ requires a certain amount in $S^{(1)}$, this same amount is taken in S, ensuring that the volume during the peak hours is correct. However, the residual risk that comes into the picture, is the volume taken in the off-peak asset.
In fact the volatility of the off-peak asset is lower, and therefore ignoring this asset is safe. Clearly, the risk in this strategy is coming mostly from the second risky asset $S^{(2)}$.

9.6.2.3 Volume-neutral strategy

The power market has two natural units of volume since the commodity is delivered in a certain magnitude over a period of time. The magnitude is expressed in MW and the time in hours. So instead of focussing on the MW position, one could also focus on the total volume, taking into account the length of the delivery period:

$$\xi = w^{(1)}\Delta^{(1)} + w^{(2)}\Delta^{(2)} = w^{(1)}\Delta^{(1)}.$$

In this strategy it is assumed that if we need 3MW of peak power, one can replace this by 1MW of base power, because the total amount of power over the delivery period is then roughly the same. Or in other words, it is assumed that we can replace volume in the peak hours by volume in the off-peak hours.
It will become clear that this is the worst strategy.

9.6.2.4 Price-adjusted strategy

If we take

$$\xi = \frac{w^{(1)}\Delta^{(1)}S^{(1)} + w^{(2)}\Delta^{(2)}S^{(2)}}{w^{(1)}S^{(1)} + w^{(2)}S^{(2)}} = \frac{w^{(1)}\Delta^{(1)}S^{(1)}}{w^{(1)}S^{(1)} + w^{(2)}S^{(2)}},$$

the value of the hedge in S and the value of the (theoretical) hedge in $S^{(1)}$ are equal. This ensures that the cash-flows during the hedging strategy are the ones one would have from the hedging strategy in $S^{(1)}$.

9.6.2.5 Risky strategy

If one wants to fully understand the concept of hedging, one should always be prepared to take one step back and ask oneself if the riskyness really decreased by setting up a strategy. Therefore, we compare the strategies to the strategy of doing nothing and waiting until the roll-over time before starting to hedge the claim. In this case, the full risk is taken and $\xi = 0$.

9.6.2.6 Adjusted locally risk-minimizing hedging strategy

In none of the above strategies, the volatility or correlation between $S^{(1)}$ and $S^{(2)}$ played a role. It is however very natural that this should have an effect on the strategy one should follow. The LRM strategy captures this completely, see Section 9.4.3 and in fact outperforms all of the above strategies.
In the Brownian case the optimal amount invested in the risk asset is given by:

$$\xi = \frac{w^{(1)}\left(\Delta^{(1)}\left(\sigma^{(1)}\right)^2 + \Delta^{(2)}\rho\sigma^{(1)}\sigma^{(2)}\right) + w^{(2)}\left(\Delta^{(1)}\rho\sigma^{(1)}\sigma^{(2)} + \Delta^{(2)}\left(\sigma^{(2)}\right)^2\right)}{w^{(1)}\left(\sigma^{(1)}\right)^2 + 2w^{(1)}w^{(2)}\rho\sigma^{(1)}\sigma^{(2)} + w^{(2)}\left(\sigma^{(2)}\right)^2}, \tag{9.37}$$

where the amounts Δ are calculated under the martingale measure for $S^{(1)}$ and $S^{(2)}$ separately, which is also the measure under which we price. In fact they should be calculated under the minimal martingale measure related with S in the semimartingale case. Still we find that this makes sense in our setting, because at roll-over we anyhow arrive at this market in the continuous case and as we will discuss later on finding the correct drift is almost impossible, hence finding the correct martingale measure is equally well unsecure.

9.6.3 Results

To obtain the results we simulated $25\,000$ paths and rebalance twice a week.

Strategy	Martingale case	Semimartingale case
Control	24.39 (0.40)	24.39 (0.40)
Total Volume-neutral	24.42 (8.34)	26.23 (9.40)
Volume-neutral	24.46 (14.09)	28.10 (15.98)
Price-adjusted	24.45 (13.08)	27.82 (14.76)
(Adjusted) LRM	**24.41** (**7.30**)	**25.78** (**8.18**)
Full risk	24.48 (17.33)	29.04 (19.63)

Table 9.1: Hedging cost and standard deviation in the case of Brownian motions.

9.6.3.1 Hedging cost

For the various strategies, we determine the hedging cost up to the roll-over time. We look at this cost both for the martingale case as well as for the semimartingale case. Table 9.1 contains for each case the expected cost and the standard deviation between brackets. The larger this standard deviation, the more uncertainty and hence the more risk remains in the hedging procedure. Let us focus first on the martingale case. It is obvious that the LRM strategy outperforms the current market practices. Compared to doing nothing, hence this is the full risk case, the improvement is very good. The reason that the total volume-neutral strategy works well is because a big part of the risk is concentrated in the peak price since this contract has the highest volatility.
In practice the hedging cost is considered as the fair value price of the option. From Table 9.1 we deduce that the average cost of hedging is almost identical across all the strategies, hence each strategy indicates the same fair value price for the option at the start.
In the semimartingale case, we can observe that for the control strategy, there is no effect. This is natural as we already know that pricing is always done under a risk-neutral martingale measure, which is unique in the continuous case. For all the other strategies we see that the cost of the strategy is changing. At the same time, the uncertainty grows as well. However, once again the LRM behaves better than any of the others.
Remark that in practice, the estimation of the drift term is virtually impossible. Knowing the drift would mean, knowing where the prices would go and often it might be possible to distinguish trends in the short or extremely long run, but the deviations from these kind of trends make it very hard to even estimate the

drift term correctly. Since the energy market is a forward market, we can assume that the market prices everything correctly, and hence that the quantities are indeed martingales. If later, it turns out that there was a systematic drift, we then hope that the margin taken at inception in the option premium is sufficient to cover this.
We conclude that the LRM hedging strategy outperforms the more intuitive approaches and even in case the assets are only semimartingales, the method still works well.
In fact we could even go one step further. We want to calculate the cost of hedging in case the underlyings follow a discontinuous price process, but in a very fast way and hence by avoiding again the use of PIDE's as described in Section 9.4.3. Therefore we use the amount ξ described in (9.37), where the price of the claim and the Δ's are calculated under the mean-correcting martingale measure linked with the two processes $S^{(1)}$ and $S^{(2)}$. For this purpose we assume that both peak and off-peak can be written as exponential variance gamma processes, where the characteristic function of a variance gamma function equals

$$\phi(z) = (1 - iu\theta\nu + \frac{1}{2}\sigma^2\nu u^2)^{-1/\nu},$$

see Schoutens (2003). As in Leoni and Schoutens (2008), we assume that the Gamma clock is equal for both assets and hence they jump at the same time. We will take the following parameters:

$\mu^{(1)}_{\text{mart}} = -0.0179$	$\mu^{(2)}_{\text{mart}} = -0.005$
$\mu^{(1)}_{\text{semi}} = 0.07$	$\mu^{(2)}_{\text{semi}} = 0.05$
$\sigma^{(1)} = 40.50\%$	$\sigma^{(2)} = 30\%$
$\theta^{(1)} = -0.10$	$\theta^{(2)} = -0.05$
$\nu = 0.25$	
$\rho = 74.80\%$	

These numbers ensure us that the option price, calculated under the mean-correcting measure for peak, leads to the same price as we had in the Brownian case. The correlation between the Brownian components has been adjusted downwards such that the linear correlation coefficient between the logreturns of the assets remained at 75% as earlier. Furthermore, $\mu^{(i)}$ in the martingale case is determined by the following equation:

$$\mu^{(i)}_{\text{mart}} = r - \log(1 - 0.5\nu(\sigma^{(i)})^2 - \theta^{(i)}\nu)/\nu, \qquad i = 1, 2.$$

Strategy	Martingale case	Semimartingale case
Control	24.39 (2.47)	24.39 (2.43)
Volume MW	23.86 (9.03)	24.84 (9.68)
Volume MWh	23.56 (14.66)	25.61 (15.82)
Price-adjusted	23.61 (13.66)	25.49 (14.68)
(Adjusted) LRM	**23.93** (**7.97**)	**24.64** (**8.50**)
Full risk	23.41 (17.85)	26.00 (19.28)

Table 9.2: Hedging cost and standard deviation in the case of a multivariate variance gamma process.

Within this setup we obtain the results reported in Table 9.2. In the martingale case, all the strategies have a lower cost of hedging, but with a greater uncertainty than in the Brownian motion case. This can be explained by the fact that within a VG model, there are only small changes in the prices of the assets until a significant jump is noticeable. The fat-tailed distribution (compared to the normal distribution) favours smaller moves most of the time and some extreme jumps once in a while.

The interesting aspect of this analysis is that the control strategy becomes less good in the sense that the uncertainty becomes much bigger. Whereas for the other strategies the uncertainty only increased slightly.

When we turn to the semimartingale case, we can deduce similar results as before. The cost of hedging depends on the actual drift of the process and in general this is not a nice feature of a strategy because this drift is extremely hard to measure or estimate. However, it becomes clear in this case as well, that the LRM strategy is rather robust, making it the most suitable candidate for real hedging of claims on non-tradable assets.

In fact one could also argue that we do not know what the real locally risk-minimizing hedging strategy would do and if this would not behave even better. We think the real strategy will surely not perform better than in the continuous martingale case, and our adjusted LRM hedging strategy only performs slightly worse than the exact LRM hedging strategy used in the continuous martingale case. Hence we believe that the possible increase in accuracy of the results will not outweigh for the loss in computational speed. Furthermore the reason why this adjusted strategy works well, even in the discontinuous case, is because it accurately captures the real correlation between the two assets, without using the exact angle bracket processes.

Conclusion

We focused in this thesis on the class of quadratic hedging strategies. We did not restrict ourselves to the pure theory, but also looked at applications in different markets such as the insurance market, the energy market and the interest rate derivatives market. The most important contribution is to the theory of local risk-minimization and the related Föllmer-Schweizer decomposition. Some long time unanswered questions are solved. Secondly, in literature there are only few implementations of the quadratic hedging strategies. Moreover, there is a growing interest for comparison to the most common hedging strategy in practice, namely the delta-hedge, see Altmann et al. (2008), Denkl et al. (2009) and Brodén and Tankov (2010). Such a practical implementation and comparison is given in Chapters 8 and 9 for an example in the interest rate derivatives market respectively the energy market.

The main results we obtained are:

- The Föllmer-Schweizer decomposition is determined for the general class of semimartingales in terms of the predictable characteristics.
- A concrete example is given and proved for which the Föllmer-Schweizer decomposition under the minimal martingale measure is different from the Galtchouk-Kunita-Watanabe decomposition under the original measure.
- The locally risk-minimizing hedging strategy is extended to the multidimensional case.
- An easy procedure to determine the locally risk-minimizing hedging strategy is given.

- Risk-minimization is applied to the framework of unit-linked life insurance with surrender option.
- The locally risk-minimizing hedging strategy for unit-linked life insurance contracts is determined in case the underlying can contain jumps.
- We calculated and implemented the delta-hedging and the mean-variance hedging strategy for the forward swaption under the forward martingale measure linked with the maturity of the swaption and we compared the obtained results with the case of following no hedging strategy at all.
- Adjusted delta-hedges are compared with the (adjusted) locally risk-minimizing hedging strategy in a setting typical for the energy market.
- The adjusted locally risk-minimizing hedging strategy we propose seems to perform really well in the setting we described for the energy market.
- The usefulness of quadratic hedging in risky markets is shown, but we also noticed the almost equality in performance to doing nothing in less risky markets.
- We revealed the gap between the theory of quadratic hedging and the implementation in the semimartingale case.

Research is a never ending story. Possible future research topics include:

- Closing the gap between the theory of mean-variance hedging and the implementation for general semimartingales.
- Finding conditions under which the (H)-hypothesis remains valid after a change of measure.
- Related with the previous point is the extension of the risk-minimization of unit-linked life insurance contracts with surrender option to semimartingales.
- Determination of the quadratic hedging strategy for unit-linked life insurance contracts containing mortality and surrender risk.
- Determination of hedging strategies for interest rate derivatives under the 'original' measure.
- Extending the mean-variance hedging strategy for the swaption when more than two zero-coupon bonds are used for hedging.
- Hedging other interest rate derivatives.
- Hedging of interest rate derivatives in the Libor market model with its typical tenor structure.

- Determination of the cumulant process under the minimal martingale measure in case the mean-variance tradeoff process is stochastic.
- Testing the proposed adjusted locally risk-minimizing hedging strategy to different situations and comparing it with the real locally risk-minimizing hedging strategy.
- ...

Samenvatting

De eerste betekenis van het Engelse werkwoord hedge is heggen maken. Zoekt men naar de figuurlijke zin, dan betekent hedge zich indekken. Meer formeel beschreven is hedging in de financiële markt dus het (geheel of gedeeltelijk) indekken tegen een financieel risico van een investering door middel van een andere investering. Er zijn natuurlijk veel verschillende mogelijkheden om te hedgen nl. deltahedge, superhedge, kwadratische hedge,. . . .
De deltahedge is populair in de praktijk omdat ze snel bepaald kan worden. Men moet namelijk enkel de afgeleide bepalen van het te hedgen product naar het product dat gebruikt wordt om het risico weg te nemen. Superhedging daarentegen zorgt ervoor dat je met 100% zekerheid ingedekt bent tegen alle risico's. De kostprijs van een dergelijke (over)bescherming is vaak zeer groot in vergelijking met het originele risico. De focus in deze thesis ligt op de bepaling en het toepassen van de kwadratische hedgingstrategieën.
Hedgingstrategieën zijn kwadratisch als ze het kwadraat van de hedgingfout minimaliseren. De definitie van deze hedgingfout varieert naargelang de specifieke strategie. De twee meest gebruikte zijn de lokale risico-minimaliserende hedgingstrategie (locally risk-minimizing hedging strategy) en de gemiddelde variantie hedgingstrategie (mean-variance hedging strategy). Bij de eerste is de hedgingfout gelijk aan het kostproces, terwijl men bij de tweede het verschil tussen de waarde van de hedgingportefeuille en van het te hedgen product minimaliseert bij afloop van het te hedgen product. Deze afloop wordt de maturiteit genoemd. Bovendien is de gemiddelde variantie strategie zelffinancierend, dit betekent dat men enkel bij de start van de hedge en bij de maturiteit een mogelijke kost zal hebben. Tussendoor wordt er dus geen geld toegevoegd of weggehaald uit de portefeuille. Hierdoor is het eenvoudig de totale kost van deze hedge uit te rekenen.

Doordat men een kwadratisch criterium minimaliseert worden winsten evenveel afgestraft als verliezen. Het voordeel van deze strategieën is echter dat de strategie voor een som van risicovolle financiële producten gelijk is aan de som van de strategieën behorend bij de afzonderlijke producten en dat de oplossing vaak expliciet gegeven kan worden.

De theoretische bijdrage van deze thesis situeert zich vooral in het domein van de lokale risico-minimaliserende hedgingstrategie en in de bepaling van de Föllmer-Schweizerdecompositie. Deze decompositie is cruciaal voor het opstellen van de lokale risico-minimaliserende hedgingstrategie.
De strategieën worden niet enkel berekend in de zuiver financiële markt, maar we bestuderen ook specifieke producten uit de verzekeringsmarkt (Hoofdstuk 6-7), de markt van de rentevoetderivaten (Hoofdstuk 8) en de energiemarkt (Hoofdstuk 9).
Om het prijsverloop van de gebruikte producten te beschrijven, moet men veronderstellingen maken betreffende de gevolgde dynamiek. In Hoofdstuk 3 werken we met algemene semimartingalen, die zelfs niet noodzakelijk quasi-links continu moeten zijn. In de context van de lokale risico-minimaliserende hedgingstrategie is quasi-links continu een basisvoorwaarde om de strategie te kunnen definiëren, bijgevolg beperken we ons in Hoofdstuk 4 tot quasi-links continue semimartingalen. In de Hoofdstukken 6 tot 9 kijken we eerder naar de toepassingen in de verschillende markten. Hier is het dus ook belangrijk de processen meer concreet te maken. Meestal veronderstellen we dan ook dat de processen ofwel een (geometrische) Brownse beweging of een (geometrisch) (tijdsinhomogeen) Lévyproces volgen.
Vaak starten we met het eenvoudigere geval van een Brownse beweging om vervolgens door toevoeging van sprongen de strategie ook te bepalen voor Lévyprocessen. Deze Lévyprocessen hebben het voordeel dat ze (grote) sprongen, die regelmatig voorkomen op de financiële markt, ook kunnen bevatten.
Nadat we enkele basisbegrippen uit de stochastische calculus ingevoerd hebben in **Hoofdstuk 2**, bespreken we in **Hoofdstuk 3** de relatie tussen de Föllmer-Schweizerdecompositie onder de originele maat en de Galtchouk-Kunita-Watanabedecompositie onder de minimale martingaalmaat. Het is algemeen gekend dat deze decomposities gelijk zijn in het geval de prijs van het onderliggend product een continu proces volgt, we bewijzen met een expliciet voorbeeld dat ze niet altijd gelijk zijn in het discontinue geval. Bovendien, geven we ook een meer algemene vorm van de Föllmer-Schweizerdecompositie door gebruik te maken van de zogenaamde voorspelbare karakteristieken.
Het nut van de Föllmer-Schweizerdecompositie voor kwadratische hedgingstra-

tegieën wordt duidelijk wanneer we de theorie aangaande de lokale risico-minimaliserende (**Hoofdstuk 4**) en de gemiddelde variantie hedgingstrategie (**Hoofdstuk 5**) bespreken. Voor beide strategieën geven we een overzicht dat niet enkel de theoretische resultaten bevat maar ook toepassingen. In Hoofdstuk 4 bepalen we bovendien de uitbreiding van de lokale risico-minimaliserende strategie naar het meerdimensionaal geval.

Voor de toepassingen van de kwadratische hedgingstrategieën starten we met de bepaling van de risico-minimaliserende hedgingstrategie voor unit-linked[1] levensverzekeringscontracten met een afkoopoptie in **Hoofdstuk 6**. De uitkeringen en eventueel ook de premies kunnen dus afhangen van het verloop van een vooraf vastgelegd referentie-aandeel of een portefeuille. De afkoopoptie geeft de mogelijkheid het contract vóór maturiteit voor een bepaalde waarde op te zeggen. We maken de veronderstelling dat de afkooptijd geen stoptijd is in de filtratie gegenereerd door de financiële markt. Dit betekent dat we de realistische veronderstelling maken dat de verzekeringshouder het contract niet enkel opzegt door de evoluties in de referentieportefeuille, maar dat hij/zij vaak persoonlijke redenen heeft om voor de afloopdatum het contract op te zeggen.

Een tweede toepassing beschreven in **Hoofdstuk 7** is de bepaling van de lokale risico-minimaliserende hedgingstrategie voor unit-linked levensverzekeringscontracten wanneer de prijs van het onderliggend risicovol product gedreven is door een Lévyproces. Door de mogelijke sprongen in de onderliggende is het niet langer mogelijk de strategie te bepalen aan de hand van de Galtchouk-Kunita-Watanabedecompositie onder de minimale martingaalmaat en dus bepalen we de Föllmer-Schweizerdecompositie indirect zoals gebeurd is in Vandaele and Vanmaele (2008b) of door gebruik te maken van de meer expliciete vorm beschreven in Hoofdstuk 3. We veronderstellen geen afkoopoptie in dit hoofdstuk en bijgevolg hebben we stochastische onafhankelijkheid tussen de financiële en de verzekeringsmarkt.

In de laatste twee hoofdstukken van dit doctoraat passen we de kwadratische hedgingstrategieën toe in twee specifieke settings. In deze hoofdstukken implementeren we ook de bekomen resultaten en vergelijken we de totale kosten bij gebruik van de verschillende hedgingstrategieën. Om snel resultaten te bekomen passen we, indien mogelijk, Fouriertransformaties toe zodat we de hedgingstrategie kunnen uitdrukken in termen van de cumulatieve functie van het prijsproces van het onderliggend aandeel.

[1]Verzekering waarbij er gespaard wordt door te beleggen in beleggingsfondsen. Het spaardeel wordt gebruikt om 'units' (beleggingseenheden) aan te kopen (cfr. hypotheekvisie.nl).

In **Hoofdstuk 8** bepalen en vergelijken we voor de forward swaption[2] de deltahedge met de gemiddelde variantie hedge onder de voorwaartse martingaalmaat gelinkt aan de maturiteit van de swaption. Om de markt van de interestvoetderivaten te modelleren veronderstellen we dat deze markt gegenereerd wordt door een Lévy uitgebreid Heath-Jarrow-Morton model, waarbij het drijvend proces tot de klasse van de 'normal inverse Gaussian' processen behoort. We starten met de bepaling van de prijs van de swaption en de deltahedge wanneer één enkele nul-coupon obligatie gebruikt wordt. Om de deltahedge te kunnen vergelijken met de gemiddelde variantie hedge geven we ook de delta- en gammaneutrale hedge en de deltahedge met als risicovrij aandeel de nul-coupon obligatie met dezelfde looptijd als de swaption. Beide hedges gebruiken, net zoals voor de gemiddelde variantie hedge, twee verschillende nul-coupon obligaties.
De kwadratische hedges zijn altijd gedefinieerd in termen van de verdisconteerde assets, maar in de markt van de interestvoetafgeleiden is het onrealistisch te veronderstellen dat de risicovrije interestvoet beschikbaar is. Bijgevolg gebruiken we als numéraire de nul-coupon obligatie met dezelfde maturiteit als de swaption. Onder de voorwaartse maat zijn dan ook alle verdisconteerde obligaties martingalen en dus kan de gemiddelde variantie hedge afgeleid worden uit de Galtchouk-Kunita-Watanabedecompositie.
In **Hoofdstuk 9** veronderstellen we dat de opties die we willen hedgen afhankelijk zijn van meerdere assets, terwijl we enkel kunnen investeren in een gewogen combinatie van deze assets. Deze setting is uit de praktijk gegrepen. Om deze producten te hedgen wordt er vaak een volume-neutrale of gewichtsaangepaste deltahedge gebruikt. We vergelijken deze aangepaste deltahedges met de (aangepaste) lokale risico-minimaliserende hedgingstrategie. De simulaties zijn beperkt tot twee onderliggenden, maar kunnen eenvoudig uitgebreid worden tot meerdere. Als drijvende processen gebruiken we Brownse bewegingen en een multivariaat 'variance gamma model', beide zowel in het martingaal geval als het semimartingaal geval.

Tot slot geven we een overzicht van de bekomen resultaten:

- De Föllmer-Schweizerdecompositie is bepaald voor de algemene klasse van semimartingalen door gebruik te maken van de voorspelbare karakteristieken.

[2]Dit is een optie op een swap, waarbij vaak een vaste interestvoet geruild wordt voor een variabele interestvoet.

- Met een voorbeeld werd expliciet aangetoond dat de Föllmer-Schweizer decompositie onder de originele maat kan verschillen van de Galtchouk-Kunita-Watanabe decompositie onder de minimale martingaalmaat.
- De theorie voor de lokale risico-minimaliserende hedgingstrategie is uitgebreid naar het multidimensionaal geval.
- Een eenvoudige procedure voor de bepaling van de lokale risico-minimaliserende hedgingstrategie is beschreven.
- De risico-minimaliserende hedgingstrategie voor unit-linked levensverzekeringscontracten met een afkoopoptie is bepaald.
- De lokale risico-minimaliserende hedgingstrategie voor unit-linked levensverzekeringscontracten in het geval de onderliggende sprongen kan bevatten wordt gegeven.
- De deltahedge en de gemiddelde variantie hedgingstrategie werden opgesteld en onderling vergeleken voor de swaption onder de voorwaartse martingaalmaat gelinkt met de maturiteit van de swaption en werden ook vergeleken met het geval dat er helemaal geen strategie toegepast wordt.
- We vergeleken aangepaste deltahedges met de (aangepaste) lokale risico-minimaliserende hedgingstrategie in een typische setting voorkomend in de energiemarkt.
- De aangepaste lokale risico-minimaliserende hedgingstrategie, die we introduceerden, blijkt zeer goed te werken in de voorgestelde setting voor de energiemarkt.
- Aan de hand van simulaties, toonden we het nut aan van kwadratische hedgingstrategieën in risicovolle markten. Anderzijds merkten we ook de beperkte invloed op van kwadratische hedgingstrategieën in markten met weinig risico.
- We toonden het hiaat aan tussen de gemiddelde variantietheorie en de implementatie hiervan in het semimartingaal geval.

Deze thesis beperkt zich dus niet enkel tot de theoretische uitwerking en de toepassingen van de kwadratische hedgingstrategieën, maar we vonden het ook belangrijk deze eerder theoretische strategieën op concrete voorbeelden te vergelijken met de in de praktijk populaire deltahedge. Vooral omdat er tot op heden in de literatuur weinig implementaties gebeurd zijn van de kwadratische hedgingstrategieën. De groeiende interesse in een vergelijkende studie van de kwadratische hedgingstrategieën met de deltahedge is ook merkbaar in de literatuur, zie Altmann et al. (2008), Denkl et al. (2009) and Brodén and Tankov (2010).

Dankwoord

Als afsluiter van mijn doctoraat rest me alleen nog het schrijven van mijn dankwoord. Het zou eenvoudig zijn snel een kort woordje van dank te richten aan mijn promotor, co-auteurs en het Fonds voor Wetenschappelijk Onderzoek-Vlaanderen voor de financiële steun, maar er zijn nog zoveel meer mensen die mij bijgestaan hebben de afgelopen vier jaar en die ik daarom expliciet wil bedanken. Bovendien is het dankwoord vaak het meest gelezen en vooral het best begrepen deel van een doctoraat. Bijgevolg wil ik er toch mijn werk van maken zoveel mogelijk mensen uitgebreid te bedanken in de hoop dat ik niemand vergeet.

Het is misschien te sterk uitgedrukt om het schrijven van een doctoraat te vergelijken met het bouwen van de Chinese muur, daarom laat ik het maar afzwakken naar het beklimmen ervan onder een fel schijnende middagzon. Het einde lijkt soms nog zo ver af en vooral zo onbereikbaar. Gelukkig heeft mijn promotor Prof. Vanmaele mij op deze ogenblikken (en ja ook bij de echte beklimming) altijd geholpen. Michèle, talloze malen ben ik binnengewandeld met de melding dat het één en al chaos was in mijn hoofd en altijd stond je klaar om te luisteren en voor opheldering te zorgen. Bedankt voor alle hulp en voor het ontelbare keren nalezen van mijn abstracts, artikels, presentaties en uiteindelijk dit doctoraat.
I also want to thank my co-authors: Prof. Choulli, Kathrin Glau and Peter Leoni for the articles we wrote together. In het bijzonder Peter: bedankt voor de leuke samenwerking, voor de opbeurende mailtjes en voor al je vrije tijd die je besteed hebt aan de implementaties. De samenwerking met Peter was er nooit gekomen

zonder Michel, die ons aan elkaar voorgesteld heeft. Trouwens, Michel, je hebt nog altijd een etentje te goed van mij als dank om mij samen met mijn broer door je lievelingsstad Wenen te gidsen. Xinliang guided me and the Leuven guys through the city of Beijing and the beautiful surroundings, including a visit to a typical karaoke. Thanks Xinliang for being our perfect guide.
The list of people I have met on conferences is very long, but still I want to thank some of them explicitly for making my time abroad pleasant: Koen, Xinliang, Zhaoning, Oriol, Gregory, Markus, Kevin, Barbara and Norbert.
Also special thanks to the Leuven guys for the nice times we had together in Belek, Piraeus, Dalian, Istanbul and in June in Toronto.
Daarnaast wil ik ook Prof. Deelstra, Prof. Dhaene en Prof. Vyncke bedanken voor de tijd die we samen doorgebracht hebben op conferenties en om lid te willen zijn van mijn examencommissie. Reeds vier jaar geleden waren jullie bij de verdediging van mijn thesis aanwezig, ik ben blij dat ik nu 4 jaar later mijn doctoraat kan voorstellen.
Alle leden van de vakgroep TWI: bedankt voor de vele koffie- en middagpauzes en de toffe avondactiviteiten die we samen beleefd hebben. In het bijzonder wil ik Heide bedanken: sinds mijn eerste bezoekje aan de S9 zo'n 8 jaar geleden hebben we ontelbare uren samen doorgebracht, eerst op de lesbanken, vervolgens in dezelfde vakgroep. Bedankt voor alle tips, adviezen en vooral het luisterend oor. Nog heel veel succes bij de afwerking van je doctoraat. Ik vind het jammer dat ik het niet meer van nabij zal kunnen beleven, maar dat zal je wel lukken hoor! Gilles, dit geldt ook voor jou! Je hebt het ontstaan van mijn doctoraat van héél dichtbij beleefd (en meegeleefd). Chapeau dat je het volgehouden hebt met mij als (op dat ogenblik enige) bureaugenoot en mijn excuses als ik je soms wat te veel van je werk afgehouden heb. Trouwens, mijn zus brengt straks de beloofde kussen. Ook mijn andere bureaugenoten: Timur en Klaas wil ik bedanken.
Patricia en Tom, we hebben op hetzelfde moment ons doctoraat afgewerkt en het was leuk om gedurende de laatste maanden al mijn grote (en kleine) problemen die ik onderweg ondervond met jullie te kunnen delen. Als laatste wil ik nog eens in het bijzonder Ann, Wouter, Katia en Herbert bedanken voor alle (niet-wetenschappelijke) hulp en de vele lange en kortere gesprekjes.

Voor de barnummertjes: ik hoop dat we elkaar blijven volgen. Het is super om te weten dat er een grote steun ter beschikking is in moeilijke tijden en er toffe mensen zijn om de leuke momenten mee te vieren. Ik hoop dat jullie de weg naar Bambrugge zullen weten te vinden. Trouwens, mijn valies voor het weekend staat al klaar!

Leen en Stijn, jullie blijven nog altijd mijn grote broer en zus waar ik toch wel naar opkijk (hoewel ik dit vroeger soms wou negeren door de bazige zus te spelen). Bedankt om het pad te effenen voor mij en mij op de juiste weg voor te gaan naar het "volwassen leven". Leen, hoewel je het zelf bij momenten ontzettend druk hebt, maak je toch nog tijd vrij voor je kleine zus. Ondertussen probeer ik zoveel mogelijk met mijn ogen te stelen hoe je werk, gezin en ontspanning evenwichtig blijft combineren. Stijn, gedurende onze 4 jaar samen op kot heb ik geprobeerd een beetje van je (en ook die van Pieter) onverwoestbare rust over te nemen. Jongens, bedankt voor de leuke tijd samen!
Mama en papa, bedankt voor alle kansen die jullie mij gegeven hebben. Zonder jullie onvoorwaardelijke steun was ik nooit in staat geweest mijn doctoraat af te leggen. Gedurende mijn blokperiodes hebben jullie mij altijd ongelooflijk vertroeteld, waardoor thuis blokken bijna een plezier werd. Jullie kakkernest is nu misschien wel groot geworden, maar ze wordt toch nog altijd graag vertroeteld. Trouwens papa, zijn er nog aardbeien?
Als laatste wil ik Philippe bedanken. Sinds mijn tweede kan heb jij mijn ups and downs verdragen. Ik weet dat ik niet altijd even genietbaar was, vooral niet in de laatste maanden van mijn doctoraat (en ja misschien ook niet in het tweede jaar van mijn doctoraat en ...). Gelukkig slaag jij er telkens weer opnieuw in een lach op mijn gezicht te toveren. De komende maanden zal ik na 8 jaar zowel Gent als de S9 verlaten, dat zal met een klein hartje zijn, maar samen met jou zal dat wel snel overwonnen worden.

Gent, mei 2010
Nele Vandaele

Bibliography

D. Akume, B. Luderer, and G.-W. Weber. Pricing and hedging of swaptions. *Journal of Computational Technologies*, 8(4):5–14, 2003.

A. Albert. *Regression and the Moore-Penrose pseudoinverse*, volume 94 of *Mathematics in Science and Engineering*. Academic Press, 1972.

T. Altmann, T. Schmidt, and W. Stute. A shot noise model for financial assets. *International Journal of Theoretical and Applied Finance*, 11(1):87–106, 2008.

F. Angelini and S. Herzel. Measuring the error of dynamic hedging: a Laplace transform approach. *Journal of Computational Finance*, 13(2):47–72, 2009.

S. Ankirchner, P. Imkeller, and G. Reis. Pricing and hedging of derivatives based on non-tradable underlyings. *Mathematical Finance*, 20(2):289–312, 2010.

J. Annaert, G. Deelstra, D. Heyman, and Vanmaele M. Risk management of a bond portfolio using options. *Insurance: Mathematics and Economics*, 41(3): 299–316, 2007.

J.-P. Ansel and C. Stricker. Lois de martingale, densités et décomposition de Föllmer-Schweizer. *Annales de l'institut Henri-Poincaré Probabilités et Statistiques*, 28(3):375–392, 1992.

J.P. Ansel and C. Stricker. Décomposition de Kunita-Watanabe. In J. Azema, P.A. Meyer, and M. Yor, editors, *Séminaire de Probabilités XXVII*, volume 1557 of *Lecture Notes in Mathematics*, pages 30–32. Springer, 1993.

D. Applebaum. *Lévy processes and stochastic calculus*. Number 93 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2004.

T. Arai. An extension of mean-variance hedging to the discontinuous case. *Finance and Stochastics*, 9(1):129–139, 2005a.

T. Arai. Some properties of the variance-optimal martingale measure for discontinuous semimartingales. *Statistics and Probability Letters*, 74:163–170, 2005b.

J. Azéma, T. Jeulin, F. Knight, and M. Yor. Le théorème d'arrêt en une fin d'ensemble prévisible. In *Séminaire de Probabilités XXVII*, volume 1557 of *Lecture Notes in Mathematics*, pages 133–158. Springer, 1993.

A.R. Bacinello, E. Biffis, and P. Millossovich. Pricing life insurance contracts with early exercise features. *Journal of Computational and Applied Mathematics*, 233:27–35, 2009.

J. Barbarin. Risk minimizing strategies for life insurance contracts with surrender option. Technical report, Université Catholique de Louvain, 2007.

J. Barbarin. Heath-Jarrow-Morton modelling of longevity bonds and the risk minimization of life insurance portfolios. *Insurance: Mathematics and Economics*, 43(1):41–55, 2008a.

J. Barbarin. *Valuation, hedging and the risk-management of insurance contracts*. PhD thesis, Université Catholique de Louvain, 2008b.

J. Barbarin. Risk-minimization with inflation and interest rate risk. Applications to non-life insurance. *Scandinavian Actuarial Journal*, 2009(2):119–151, 2009.

G. Becchere and S. Mulinacci. Hedging American options in Merton's model: a locally risk minimizing approach. *Asia-Pacific Financial Markets*, 6:153–170, 1999.

F. Benth and T. Meyer-Brandis. The density process of the minimal entropy martingale measure in a stochastic volatility model with jumps. *Finance and Stochastics*, 9:563–575, 2005.

F. Benth, G. Di Nunno, A. Løkka, B. Øksendal, and F. Proske. Explicit representation of the minimal variance portfolio in markets driven by Lévy processes. *Mathematical Finance*, 13(1):55–72, 2003.

F. Biagini. *Quadratic hedging approach for interest rate models with stochastic volatility*. PhD thesis, Scuola Normale Superiore of Pisa, 2001.

F. Biagini and A. Cretarola. Quadratic hedging methods for defaultable claims. *Applied Mathematics and Optimization*, 56(3):425–443, 2007.

F. Biagini and A. Cretarola. Local risk-minimization for defaultable markets. *Mathematical Finance*, 19(4):669–689, 2009.

F. Biagini and A. Cretarola. Local risk-minimization for defaultable claims with recovery process. Technical report, LMU University of München and University of Bologna, 2006.

F. Biagini and P. Guasoni. Mean-variance hedging with random volatility jumps. *Stochastic Analysis and Applications*, 20(3):471–494, 2002.

F. Biagini and M. Pratelli. Local risk-minimization and numéraire. *Journal of Applied Probability*, 36:1126–1139, 1999.

F. Biagini, P. Guasoni, and M. Pratelli. Mean-variance hedging for stochastic volatility models. *Mathematical Finance*, 10(2):109–123, 2000.

T. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging*. Springer Finance, Berlin, 2002.

T. Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, 2004.

C. Blanchet-Scalliet and M. Jeanblanc. Hazard rate for credit risk and hedging defaultable contingent claims. *Finance and Stochastics*, 8:145–159, 2004.

O. Bobrovnytska and M. Schweizer. Mean-variance hedging and stochastic control: beyond the Brownian settings. *IEEE Transactions on Automatic Control*, 49(3):396–408, 2004.

N. Bouleau and D. Lamberton. Residual risks and hedging strategies in Markovian markets. *Stochastic Processes and their Applications*, 33:131–150, 1989.

A. Brace, T. Dun, and G. Barton. Towards a central interest rate model. In E. Jouini, J. Cvitanić, and M. Musiela, editors, *Option Pricing, Interest Rates and Risk Management*, Handbooks in Mathematical Finance, pages 278–313. Cambridge, 2001.

D. Brigo and F. Mercurio. *Interest Rate Models: Theory and Practice*. Springer Finance, 2001.

M. Brodén and P. Tankov. Tracking errors from discrete hedging in exponential Lévy models. Technical report, Lund University and Ecole Polytechnique, 2010.

P. Carr and D.H. Madan. Option valuation using the fast Fourier transform. *The Journal of Computational Finance*, 2(4):61–73, 1999.

A. Černý and J. Kallsen. On the structure of general mean-variance hedging strategies. *Annals of Probability*, 35(4):1479–1531, 2007.

A. Černý and J. Kallsen. Mean-variance hedging and optimal investment in Heston's model with correlation. *Mathematical Finance*, 18:473–492, 2008a.

A. Černý and J. Kallsen. A counterexample concerning the variance-optimal martingale measure. *Mathematical Finance*, 18(2):305–316, 2008b.

T. Chan. Pricing contingent claims on stocks driven by Lévy processes. *Annals of Applied Probability*, 9(2):504–528, 1999.

T. Chan, J. Kollar, and A. Wiese. The variance-optimal martingale measure in Lévy models with stochastic volatility. Technical report, 2009.

T. Choulli and C. Stricker. Minimal entropy-Hellinger martingale measure in incomplete markets. *Mathematical Finance*, 15(3):465–490, 2005.

T. Choulli and C. Stricker. Deux applications de la décomposition Galtchouk-Kunita-Watanabe. In *Séminaire de Probabilités de Strasbourg XXX*, Lecture Notes in Mathematics, pages 22–23. Springer-Verlag, 1996.

T. Choulli and C. Stricker. More on minimal entropy-Hellinger martingale measure. *Mathematical Finance*, 16(1):1–19, 2006.

T. Choulli, L. Krawczyk, and C. Stricker. $\mathcal{E}$-martingales and their applications in mathematical finance. *Annals of Probability*, 26(2):853–876, 1998.

T. Choulli, C. Stricker, and J. Li. Minimal Hellinger martingale measure of order q. *Finance and Stochastics*, 11(3):399–427, 2007.

T. Choulli, N. Vandaele, and M. Vanmaele. The Föllmer-Schweizer decomposition: Comparison and description. *Stochastic Processes and their Applications*, doi:10.1016/j.spa.2010.02.004, 2010.

T. Coleman, Y. Li, and M.-C. Patron. Hedging guarantees in variable annuities under both equity and interest rate risks. *Insurance: Mathematics and Economics*, 38:215–228, 2006.

T. Coleman, Y. Kim, Y. Li, and M.-C. Patron. Robustly hedging variable annuities with guarantees under jump and volatility risks. *The Journal of Risk and Insurance*, 74(2):347–376, 2007.

D. Colwell and R. Elliott. Discontinuous asset prices and non-attainable contingent claims. *Mathematical Finance*, 3(3):295–308, 1993.

D. Colwell, N. El-Hassan, and O. Kwon. Hedging diffusion processes by local risk minimization with applications to index tracking. *Journal of Economic Dynamics & Control*, 31:2135–2151, 2007.

R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall, 2004.

A. Cretarola. *Local risk-minimization for defaultable markets*. PhD thesis, University of Bologna, 2007.

M. Davis. Optimal hedging with basis risk. In Y. Kabanov, R. Liptser, and J. Stoyanov, editors, *From Stochastic Calculus to Mathematical Finance*, Mathematics and Statistics, pages 169–187. Springer, 2006.

F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(3):463–520, 1994.

F. Delbaen and W. Schachermayer. Attainable claims with p-th moments. *Annales de l'Institut Henri Poincaré*, 36(6):743–763, 1996a.

F. Delbaen and W. Schachermayer. The variance-optimal martingale measure for continuous processes. *Bernoulli*, 2(1):81–105, 1996b.

F. Delbaen, P. Monat, C. Stricker, W. Schachermayer, and M. Schweizer. Weighted norm inequalities and hedging in incomplete markets. *Finance and Stochastics*, 1(3):181–227, 1997.

C. Dellacherie and P.-A. Meyer. *Théorie des Martingales (Chapters V-VIII)*. Probabilités et potentiel. Hermann, 1980.

Ł. Delong. An optimal investment strategy for a stream of liabilities driven by a step process in a financial market driven by a Lévy process. Technical report, Warsaw School of Economics, 2009.

S. Denkl, M. Goy, J. Kallsen, J. Muhle-Karbe, and A. Pauwels. On the performance of delta hedging strategies in exponential Lévy models. Technical report, Universität zu Kiel, 2009.

G. Di Masi, Y. Kabanov, and W. Runggaldier. Mean-variance hedging of options on stocks with Markov volatilities. *SIAM: Theory of Probability and its Applications*, 39:172–182, 1994.

G. Di Masi, E. Platen, and W. Runggaldier. Hedging of options under discrete observations on assets with stochastic volatilities. In E. Bolthausen, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications*, volume 36 of *Progress in Probability*, pages 359–364. Birkhäuser Verlag, 1995.

J. Driessen, P. Klaassen, and B. Melenberg. The performance of multi-factor term structure models for pricing and hedging caps and swaptions. *Journal of Financial and Quantitative Analysis*, 38(3):635–672, 2003.

D. Duffie and H.R. Richardson. Mean-variance hedging in continuous time. *Annals of Applied Probability*, 1:1–15, 1991.

T. Dun, E. Schlögl, and G. Barton. Simulated swaption delta-hedging in the lognormal forward LIBOR market. *International Journal of Theoretical and Applied Finance*, 4(4):677–709, 2001.

E. Eberlein and W. Kluge. Exact pricing formulae for caps and swaptions in a Lévy term structure model. *Journal of Computational Finance*, 9:99–125, 2006.

E. Eberlein and S. Raible. Term structure models driven by general Lévy processes. *Mathematical Finance*, 9(1):31–53, 1999.

E. Eberlein, J. Jacod, and S. Raible. Lévy term structure models: no-arbitrage and completeness. *Finance and Stochastics*, 9(1):67–88, 2005.

E. Eberlein, K. Glau, and A. Papapantoleon. Analysis of Fourier transform valuation formulas and applications. Technical report, 2009. To appear in Applied Mathematical Finance.

N. El Karoui, S. Peng, and M. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71, 1997.

P. Fisher, E. Platen, and W. Runggaldier. Risk-minimizing hedging strategies under partial observation. In R.C. Dalang, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications*, volume 45 of *Progress in Probability*, pages 175–188. Birkhäuser, 1999.

H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. In M.H.A. Davis and R.J. Elliot, editors, *Applied Stochastic Analysis*, volume 5 of *Stochastic Monographs*, pages 389–414. Gordon and Breach, 1991.

H. Föllmer and D. Sondermann. Hedging of non-redundant contingent claims. In W. Hildenbrand and A. Mas-Colell, editors, *Contributions to Mathematical Economics*, pages 205–223. North-Holland, Elsevier, 1986.

R. Frey and W. Runggaldier. Risk-minimizing hedging strategies under restricted information: The case of stochastic volatility models observable only at discrete random times. *Mathematical Methods of Operations Research*, 50:339–350, 1999.

T. Fujiwara and Y. Miyahara. The minimal entropy martingale measure for geometric Lévy processes. *Finance and Stochastics*, 7:509–531, 2003.

K. Glau, N. Vandaele, and M. Vanmaele. Hedging of swaptions in the Lévy driven Heath-Jarrow-Morton model. Technical report, University of Freiburg and Ghent University, 2010a.

K. Glau, N. Vandaele, and M. Vanmaele. Pricing and hedging of interest rate derivatives in a Lévy driven term structure model. In M. Vanmaele, G. Deelstra, A. De Schepper, J. Dhaene, W. Schoutens, and S. Vanduffel, editors, *Handelingen Contactforum Actuarial and Financial Mathematics Conference, February 4-5, 2010*. Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, 2010b.

C. Gourieroux, J.-P. Laurent, and H. Pham. Mean-variance hedging and numéraire. *Mathematical Finance*, 8(3):179–200, 1998.

P. Grandits and L. Krawczyk. Closedness of some spaces of stochastic integrals. In M. Yor, M. Emery, and J. Azema, editors, *Séminaire de Probabilités XXXII*, volume 1686 of *Lecture Notes in Mathematics*, pages 73–85. Springer, 1998.

J. Harrison and D. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation. *Econometrica*, 60:77–105, 1992.

D. Heath, E. Platen, and M. Schweizer. A comparison of two quadratic approaches to hedging in incomplete markets. *Mathematical Finance*, 11:385–413, 2001a.

D. Heath, E. Platen, and M. Schweizer. Numerical comparison of local risk-minimisation and mean-variance hedging. In E. Jouini, Cvitanić J., and M. Musiela, editors, *Option Pricing, Interest Rates and Risk Management*, pages 509–537. Cambridge University Press, 2001b.

V. Henderson. Valuation of claims on nontraded assets using utility maximization. *Mathematical Finance*, 12(4):351–373, 2002.

V. Henderson and D. Hobson. Real options with constant relative risk aversion. *Journal of Economic Dynamics and Control*, 27:329–355, 2002.

M. Henrard. Explicit bond option and swaption formula in Heath-Jarrow-Morton one factor model. *The International Journal of Theoretical and Applied Finance*, 6(1):57–72, 2003.

C. Hipp. Hedging general claims. In *Proceedings of the 3rd AFIR Colloquium*, number 2, pages 603–613, 1993.

D.G. Hobson. Bounds for the utility-difference prices of non-traded assets in incomplete markets. *Decisions in Economics and Finance*, 28:33–52, 2005.

U. Horst, T. Pirvu, and G. Reis. On securitization, market completion and equilibrium risk transfer. *Mathematics and Financial Economics*, 2(4):211–252, 2010.

F. Hubalek and L. Krawczyk. Simple explicit formulae for variance-optimal hedging for processes with stationary independent increments. Technical report, 1998.

F. Hubalek, J. Kallsen, and L. Krawczyk. Variance-optimal hedging for processes with stationary independent increments. *Annals of Applied Probability*, 16: 853–885, 2006.

J. Jacod. *Calcul stochastique et problèmes de martingales*, volume 714 of *Lecture Notes in Mathematics*. Springer-Verlag, 1979.

J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, 2002.

F. Jamshidian. An exact bond option formula. *The Journal of Finance*, 44(1): 205–209, 1989.

M. Jeanblanc and Y. Le Cam. Progressive enlargement of filtrations with initial times. *Stochastic Processes and their Applications*, 119(8):2523–2543, 2009.

M. Jeanblanc and M. Rutkowski. Modelling of default risk: An overview, 1999.

M. Jeanblanc and M. Rutkowski. Modelling of default risk: Mathematical tools, 2000.

M. Jeanblanc, S. Klöppel, and Y. Miyahara. Minimal f^q-martingale measures for exponential Lévy processes. *Annals of Applied Probability*, 17(5–6):1615–1638, 2007.

R. Kaas, J. Dhaene, and M. Goovaerts. Upper and lower bounds for sums of random variables. *Insurance: Mathematics and Economics*, 27(2):151–168, 2000.

J. Kallsen. σ-localization and σ-martingales. *Theory of Probability and its Applications*, 48:152–163, 2004.

J. Kallsen. A didactic note on affine stochastic volatility models. In Y. Kabanov, R. Liptser, and J. Stoyanov, editors, *From Stochastic Calculus to Mathematical Finance*, pages 343–368. Springer, Berlin, 2006.

J. Kallsen and A. Pauwels. Variance-optimal hedging in general affine stochastic volatility. Technical report, 2009a. To appear in Advances in Applied Probability.

J. Kallsen and A. Pauwels. Variance-optimal hedging for time-changed Lévy processes. Technical report, 2009b. To appear in Applied Mathematical Finance.

J. Kallsen and A. Shiryaev. The cumulant process and Esscher's change of measure. *Finance and Stochastics*, 6:397–428, 2002.

J. Kallsen and R. Vierthauer. Quadratic hedging in affine stochastic volatility models. *Review of Derivatives Research*, 12(1):3–27, 2009.

J. Kallsen, J. Muhle-Karbe, N. Shenkman, and R. Vierthauer. Discrete-time variance-optimal hedging in affine stochastic volatility models. In R. Kiesel, M. Scherer, and R. Zagst, editors, *Alternative Investments and Strategies*. World Scientific, 2009.

S. Kassberger and Th. Liebmann. q-optimal martingale measures for exponential Lévy processes. Technical report, University of Ulm, 2007.

W. Kluge. *Time-inhomogeneous Lévy processes in interest rate and credit risk models*. PhD thesis, University of Freiburg, 2005.

M. Kohlmann and B. Peisl. A note on mean-variance hedging of non-attainable claims. Technical report, 2000.

M. Kohlmann and D. Xiong. The mean-variance hedging of a defaultable option with partial information. *Stochastic Analysis and Applications*, 25:869–893, 2007.

M. Kohlmann, D. Xiong, and Y. Zhonxing. The mean-variance hedging in a general jump model. *Applied Mathematical Finance*, 17(1):29–57, 2010.

D. Kramkov and M. Sirbu. Asymptotic analysis of utility-based hedging strategies for small number of contingent claims. *Stochastic processes and their Applications*, 117:1606–1620, 2007.

D. Lamberton, H. Pham, and M. Schweizer. Local risk-minimization under transaction costs. *Mathematics of Operations Research*, 23:585–612, 1998.

J.-P. Laurent and H. Pham. Dynamic programming and mean-variance hedging. *Finance and Stochastics*, 3(1):83–110, 1999.

P. Leoni and W. Schoutens. Multivariate smiling. *Wilmott Magazine*, March 2008.

P. Leoni, N. Vandaele, and M. Vanmaele. Hedging strategies applied to non-traded assets. Technical report, Ghent University, 2010.

A. Lim. Quadratic hedging and mean-variance portfolio selection with random parameters in an incomplete market. *Mathematics of Operations Research*, 29 (1):132–161, 2004.

A. Lim. Mean-variance hedging when there are jumps. *SIAM Journal on Control and Optimization*, 44(5):1893–1922, 2005.

A.V. Melnikov and M.L. Nechaev. On the mean-variance hedging problem. *Theory of Probability and its Applications*, 43(4):588–603, 1999.

F. Mercurio and T. Vorst. Option pricing and hedging in discrete time with transaction costs and incomplete markets. In M. Dempster and S. Pliska, editors, *Mathematics of Derivative Securities*, pages 190–215. Cambridge University Press, 1997.

T. Møller. Risk-minimizing hedging strategies for unit-linked life insurance contracts. *Astin Bulletin*, 28(1):17–47, 1998.

T. Møller. Risk-minimizing hedging strategies for insurance payment processes. *Finance and Stochastics*, 5:419–446, 2001.

P. Monat and C. Stricker. Fermeture de $G_T(\theta)$ et de $L^2(\mathcal{F}_0) + G_T(\Theta)$. In *Séminaire de Probabilités XXVIII*, Lecture Notes in Mathematics, pages 189–194. Springer, 1994.

P. Monat and C. Stricker. Föllmer-Schweizer decomposition and mean-variance hedging for general claims. *Annals of Probability*, 23:605–628, 1995.

M. Monoyios. Performance of utility-based strategies for hedging basis risk. *Quantitative Finance*, 4:245–255, 2004.

M. Musiela and M. Rutkowski. *Martingale Methods in Financial Modelling*, volume 36 of *Stochastic Modelling and Applied Probability*. Springer, 2004.

A. Pelsser. *Efficient Methods for Valuing Interest Rate Derivatives*. Springer Finance, 2000.

H. Pham. On quadratic hedging in continuous time. *Mathematical Methods of Operations Research*, 51:315–339, 2000.

H. Pham, T. Rheinländer, and M. Schweizer. Mean-variance hedging for continuous processes: New proofs and examples. *Finance and Stochastics*, 2: 173–198, 1998.

V. Piterbarg. A practitioner's guide to pricing and hedging callable LIBOR exotics in forward LIBOR models. Technical report, 2003. doi:10.2139/ssrn.427084.

V. Piterbarg. Computing deltas of callable LIBOR exotics in forward LIBOR models. *Journal of Computational Finance*, 7(3):107–144, 2004.

R. Poulsen, K.R. Schenk-Hoppé, and C.-O. Ewald. Risk minimization in stochastic volatility models: Model risk and empirical performance. *Quantitative Finance*, 9(6):693–704, 2009.

P. Protter. *Stochastic integration and differential equations, Second Edition, Version* 2.1. Number 21 in Stochastic Modelling and Applied Probability. Springer, Berlin, 2005.

T. Rheinländer and M. Schweizer. On L^2-projections on a space of stochastic integrals. *Annals of Probability*, 25:1810–1831, 1997.

T. Rheinländer and G. Steiger. The minimal entropy martingale measure for general Barndorff-Nielsen/Shephard models. *Annals of Applied Probability*, 16(3):1319–1351, 2006.

M. Riesner. Hedging life insurance contracts in a Lévy process financial market. *Insurance: Mathematics and Economics*, 38:599–608, 2006a.

M. Riesner. *Unit-linked life insurance in Lévy process financial market*. PhD thesis, University of Ulm, 2006b.

M. Riesner. Locally risk-minimizing hedging of insurance payment streams. *Astin Bulletin*, 37(1):67–91, 2007.

K.-I. Sato. *Lévy processes and infinitely divisible distributions*. Number 68 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.

M. Schäl. On quadratic cost criteria for option hedging. *Mathematics of Operations Research*, 19(1):121–131, 1994.

W. Schoutens. *Lévy Processes in Finance*. Wiley Series in Probability and Statistics, 2003.

M. Schweizer. *Hedging of options in a general semimartingale model*. PhD thesis, Swiss Federal Institute of Technology Zurich, 1988.

M. Schweizer. Risk-minimality and orthogonality of martingales. *Stochastics and Stochastics Reports*, 30(1):123–131, 1990.

M. Schweizer. Option hedging for semimartingales. *Stochastic Processes and their Applications*, 37:339–363, 1991.

M. Schweizer. Mean-variance hedging for general claims. *Annals of Applied Probability*, 2:171–179, 1992.

M. Schweizer. Risk-minimizing hedging strategies under restricted information. *Mathematical Finance*, 4:327–342, 1993.

M. Schweizer. Approximating random variables by stochastic integrals. *Annals of Probability*, 22(3):1536–1575, 1994.

M. Schweizer. On the minimal martingale measure and the Föllmer-Schweizer decomposition. *Stochastic Analysis and Applications*, 13:573–599, 1995a.

M. Schweizer. Variance-optimal hedging in discrete time. *Mathematics of Operations Research*, 24:204–236, 1995b.

M. Schweizer. Approximation pricing and the variance-optimal martingale measure. *Annals of Probability*, 24(1):206–236, 1996.

M. Schweizer. Risky options simplified. *International Journal of Theoretical and Applied Finance*, 2:59–82, 1999.

M. Schweizer. A guided tour through quadratic hedging approaches. In E. Jouini, J. Cvitanić, and M. Musiela, editors, *Option Pricing, Interest Rates and Risk Management*, pages 538–574. Cambridge University Press, 2001.

M. Schweizer. Local risk-minimization for multidimensional assets and payment streams. *Banach Center Publications*, 83:213–229, 2008.

S. Shreve. *Stochastic Calculus for Finance II, Continuous-Time Models*. Springer Finance, 2004.

P. Tankov. *Lévy processes in finance: inverse problems and dependence modelling*. PhD thesis, Ecole Polytechnique, 2004.

P. Tankov. Pricing and hedging jump risk. Presentation notes of Spring School in Finance, Bologna, May 2009.

N. Vandaele and M. Vanmaele. Overview: (locally) risk-minimizing hedging strategies. In M. Vanmaele, G. Deelstra, A. De Schepper, J. Dhaene, H. Reynaerts, W. Schoutens, and P. Van Goethem, editors, *Handelingen Contactforum Actuarial and Financial Mathematics Conference, February 7-8, 2008*, pages 107–119. Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, 2008a.

N. Vandaele and M. Vanmaele. A locally risk-minimizing hedging strategy for unit-linked life insurance contracts in a Lévy process financial market. *Insurance: Mathematics and Economics*, 42(3):1128–1137, 2008b.

N. Vandaele and M. Vanmaele. Explicit portfolio for unit-linked life insurance contracts with surrender option. *Journal of Computational and Applied Mathematics*, 233:16–26, 2009.

M. Yor and J. de Sam Lazaro. Sous-espace denses dans L^1 ou H^1 et représentation des martingales. In M. Yor, M. Emery, and J. Azema, editors, *Séminaire de Probabilités de Strasbourg XII,* volume 1686 of *Lecture Notes in Mathematics*, pages 265–309. Springer-Verlag, 1978.