University of Alberta

DYNAMIC HEDGING: CVAR MINIMIZATION AND PATH-WISE COMPARISON

by

Ivan Smirnov

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mathematical Finance

Department of Mathematical and Statistical Sciences

Spring 2013
Edmonton, Alberta

Abstract

Imposing constraints on the class of the available self-financing strategies may eliminate the possibility of using replicating or superhedging strategies, which leads to the problem of partial hedging. In the present work, the partial hedging problem is investigated from the viewpoint of the contingent claim seller who aims to minimize the shortfall risk through dynamic hedging under the constraint on the initial capital. The shortfall risk is measured via conditional value-at-risk, a coherent quantile risk measure. Another problem consists in finding a strategy that minimizes hedging costs under a constraint on conditional value-at-risk of the hedging portfolio. In a complete market, an explicit algorithm for constructing the optimal hedging strategy in both problems is presented, along with a number of detailed illustrations. In the incomplete case, the optimal solution is no longer explicit, however a certain generalization of the Neyman-Pearson lemma may be used to deduce the general structure of the optimal strategy. Some of such generalizations assume weak compactness of the set of densities of equivalent sigma-martingale measures. We show that this requirement is in fact never satisfied in the incomplete market setting and provide detailed discussion of the matter in both the discrete- and continuous-time cases. Finally, we demonstrate how pathwise comparison can be used in the problem of approximate option hedging and pricing, and we illustrate the approach in the framework of the constant elasticity of variance model.

Acknowledgements

I want to express my sincere gratitude to my thesis advisor, Dr. Alexander Melnikov, for the continuous support of my Ph.D. study and research, for his immense knowledge and thoughtful guidance. I am also grateful to the other members of my thesis committee whose insightful suggestions and comments helped shape the final version of this thesis.

I would like to acknowledge the contribution of Alexey Kuznetsov, George Lowther and Vlad Krasin who I have had pleasure to collaborate with on the topics regarding the telegraph market model, compactness of the set of densities of martingale measures and the path-wise comparison theorem.

Last but not least, I would like to thank my parents and my beautiful girlfriend, Tanya, for their understanding, endless patience, love and encouragement when it was most needed.

Table of Contents

List of Figures

Chapter 1

Introduction

This thesis is composed of several relatively independent parts which represent the areas of research that I have been involved in during my stay at the University of Alberta.

The first part has been my primary focus of study and relates to the problem of partial hedging of contingent claims with conditional value-at-risk chosen as the performance criterion. The problem naturally arises in situations where the claim seller is unwilling to engage in a perfect hedging or a super-hedging strategy due to the high cost of the such. By limiting the initial capital from above, we restrain the class of the available self-financing strategies and we then search for a strategy that is optimal in the sense of minimizing conditional value-at-risk of the hedging portfolio at the terminal moment of time. It turns out that in a complete market model the optimal strategy can be deduced in a semi-explicit fashion in terms of randomized tests by utilizing the Neyman-Pearson lemma in conjunction with a certain representation theorem for conditional value-at-risk. This method is also applicable to the dual problem in which the required initial capital is minimized subject to a constraint on the conditional value-at-risk. In the incomplete market setting, the optimal strategy can no longer be found explicitly, however by using duality methods for state-dependent utility functions it can be shown that the solution still has the typical 0-1 structure.

The second part presents a joint research with Alexander Melnikov and Vladislav Krasin and is devoted to the problem of finding option price bounds in diffusion models via the path-wise comparison theorem. To illustrate the approach, we derive the upper price bound explicitly for a call option in the framework of the constant elasticity of variance model and compare precision of our estimate to the distribution-free approach. We show that our method provides highly precise price bounds for in-the-money options and can also be viewed as an alternative to the existing numerical methods of computing option prices in the constant elasticity of variance model for extremely short maturities.

Last but not least, this thesis questions weak L^1-compactness of the family of densities of equivalent sigma-martingale measures. To our knowledge, this topic hasn't been studied in detail in existing literature, yet it carries significant meaning for the problems of composite hypotheses testing that arise in partial hedging problems since certain duality methods require the aforesaid set of densities to be compact in L^1. We show that this set is never closed in an incomplete arbitrage-free market due to the violation of measure equivalence of the limit measure, both in discrete- and continuous-time settings. We also investigate the discrete-time case thoroughly and find other reasons for non-closedness and the lack of relative compactness of the set of densities of equivalent martingale measures.

1.1 Outline

The dissertation is divided into five chapters. Chapter 1 introduces the reader to the notion of conditional value-at-risk and its characteristic properties, the Neyman-Pearson lemma with some of its generalizations, and a certain version of the path-wise comparison theorem.

Chapter 2 contains our main theoretical results regarding the problem of CVaR-optimal partial hedging. We first pose the problem of conditional value-at-risk minimization under a capital constraint in a complete market setting and demonstrate how the optimal partial hedging strategies can be constructed semi-explicitly. Next, we consider the the dual problem of hedging costs minimization under the conditional value-at-risk constraint, where the solution can also be derived in a straightforward way. Finally, we briefly discuss the general structure of the optimal solution in the incomplete market setting.

In Chapter 3, we present several illustrations of our approach in complete market models: the Black-Scholes model, the telegraph market model, and the problem of hedging a unit-linked life insurance contract in the Black-Scholes setting.

The path-wise comparison theorem and its applicability to mathematical finance are discussed Chapter 4, where the explicit option price bounds and the approximate hedging strategy are derived in the framework of the constant elasticity of variance model, and performance of the suggested method is compared numerically to the distribution-free approach.

Finally, Chapter 5 contains full proofs of several auxiliary results, including the detailed investigation of the reasons for non-compactness of the family of densities of equivalent sigma-martingale measures.

1.2 Conditional value-at-risk and other quantile risk measures

The primary purpose of this introductory section is to familiarize the reader with the notion of conditional value-at-risk, outline its main characteristic features and point out its key distinctions from the related quantile risk measures (which in practice get often mixed up and are used interchangeably).

Before we start off with formal definitions, let us note that conditional value-at-risk (or, as the practitioners prefer to call it, the expected shortfall) has recently become the central topic of attention in the banking industry. In the consultative document by [Basel Committee on Banking Supervision, 2012] (the outline for "Basel 3.5"), it is suggested to move away from the VaR-based methodology of estimating the required regulatory capital and adopt expected shortfall as a unified risk metric in both the standardized and the internal-model approaches. Quoting the above document,

> "A number of weaknesses have been identified with using value-at-risk (VaR) for determining regulatory requirements, including its inability to capture tail risk. For this reason, the Committee has considered alternative risk metrics, in particular expected shortfall (ES). <...> Accordingly, the Committee is proposing the use of ES for the internal models-based approach and also indends to determine risk weights for the standardised approach using an ES methodology."

Concerning the aforementioned "weaknesses" of value-at-risk, we refer to [Sarykalin et al., 2008] for a detailed survey of properties of value-at-risk and conditional value-at-risk in risk management and optimization problems.

In this section, we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a choice-dependent $\mathcal{F}$-measurable random variable $L = L(x)$ characterizing the financial loss contingent upon the choice of strategy $x \in X$, where X is the set of all available strategies. We require that $\mathbb{E}\left(|L(x)|\right) < \infty$ for all $x \in X$.

Denote the lower and upper α-quantiles of $L(x)$ by $L_{(\alpha)}(x)$ and $L^{(\alpha)}(x)$ respectively:

$$L_{(\alpha)} = L_{(\alpha)}(x) = \inf\left\{t \in \mathbb{R} \ : \ \mathbb{P}\left(L \leqslant t\right) \geqslant \alpha\right\}, \tag{1.1}$$

$$L^{(\alpha)} = L^{(\alpha)}(x) = \inf\left\{t \in \mathbb{R} \ : \ \mathbb{P}\left(L \leqslant t\right) > \alpha\right\}. \tag{1.2}$$

For a given strategy $x \in X$ and a fixed confidence level $\alpha \in (0,1)$ with typical values close to 1, *value-at-risk* (VaR) at the confidence level α is defined as the upper α-quantile of the loss distribution:

$$\mathrm{VaR}^{\alpha}(L) = L^{(\alpha)}. \tag{1.3}$$

In other words, VaR at confidence level α is the smallest number λ such that the probability that the loss exceeds λ is not larger than $(1-\alpha)$.

Following [Acerbi and Tasche, 2002a], we define *conditional value-at-risk* (CVaR) at the confidence level α, also known as *expected shortfall* (ES) or *average value-at-risk* (AVaR), as

$$\mathrm{CVaR}^{\alpha}(L) = \frac{1}{1-\alpha}\left(\mathbb{E}\left(\mathbf{1}_{\left\{L \geqslant L^{(\alpha)}\right\}} L\right) + L^{(\alpha)}\left(1 - \alpha - \mathbb{P}\left(L \geqslant L^{(\alpha)}\right)\right)\right), \tag{1.4}$$

where $\mathbf{1}_A$ is the indicator of event $A \in \mathcal{F}$:

$$\mathbf{1}_A(\omega) = \mathbf{1}_A = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases} \tag{1.5}$$

Alternatively, CVaR can be defined as a mixture of VaRs evaluated at varying confidence levels (see the definition of AV@R in Föllmer and Schied [2011]):

$$\mathrm{CVaR}^{\alpha}(L) = \frac{1}{1-\alpha} \int_{\alpha}^{1} L^{(p)} dp = \frac{1}{1-\alpha} \int_{\alpha}^{1} \mathrm{VaR}^{p}(L) dp, \tag{1.6}$$

or as the mean of the α-tail of a suitably modified distribution (see the definition of CVaR in Rockafellar and Uryasev [2002]).

Conditional value-at-risk is closely related to the notions of tail conditional expectation (TCE) and worst conditional expectation (WCE). *Lower TCE*, also known as *conditional tail expectation* (CTE) or *tail value-at-risk* (TVaR), is the conditional expectation of loss above the upper α-quantile of the loss distribution:

$$\mathrm{TCE}_{\alpha}(L) = \mathbb{E}\left(L \mid L \geqslant L^{(\alpha)}\right) = \mathbb{E}\left(L \mid L \geqslant \mathrm{VaR}^{\alpha}(L)\right). \tag{1.7}$$

Similarly, *upper TCE* is the conditional expectation of loss above the lower α-quantile of the loss distribution:

$$\mathrm{TCE}^{\alpha}(L) = \mathbb{E}\left(L \mid L \geqslant L_{(\alpha)}\right). \tag{1.8}$$

Worst conditional expectation is defined as in [Artzner et al., 1999] (up to a discount factor):

$$\mathrm{WCE}_{\alpha}(L) = \sup\left\{\mathbb{E}\left(L \mid A\right) : A \in \mathcal{F}, \mathbb{P}(A) > 1-\alpha\right\}. \tag{1.9}$$

Proposition 1.1 (Acerbi and Tasche [2002b]):

$$\mathrm{TCE}^{\alpha}(L) \leqslant \mathrm{TCE}_{\alpha}(L) \leqslant \mathrm{CVaR}^{\alpha}(L), \tag{1.10}$$

$$\mathrm{TCE}^{\alpha}(L) \leqslant \mathrm{WCE}_{\alpha}(L) \leqslant \mathrm{CVaR}^{\alpha}(L). \tag{1.11}$$

Proposition 1.2 (Acerbi and Tasche [2002b]):

$$\mathrm{CVaR}^{\alpha}(L) = \mathrm{WCE}_{\alpha}(L) = \mathrm{TCE}_{\alpha}(L) = \mathrm{TCE}^{\alpha}(L) \tag{1.12}$$

if and only if

$$\mathbb{P}\left(L \geqslant L_{(\alpha)}\right) = 1-\alpha, \tag{1.13}$$

$$\mathbb{P}\left(L > L^{(\alpha)}\right) > 0 \tag{1.14}$$

or

$$\mathbb{P}\left(\left\{L \geqslant L_{(\alpha)}\right\} \cap \left\{L \neq L^{(\alpha)}\right\}\right) = 0. \tag{1.15}$$

Proposition 1.3 (Acerbi and Tasche [2002b]):

$$\mathrm{CVaR}^{\alpha}(L) = \mathrm{TCE}_{\alpha}(L) \tag{1.16}$$

if and only if

$$\mathbb{P}\left(L \geqslant L^{(\alpha)}\right) = 1-\alpha \tag{1.17}$$

or

$$\mathbb{P}\left(L > L^{(\alpha)}\right) = 0. \tag{1.18}$$

Propositions 1.2 and 1.3 imply that if L is a continuous random variable, CVaR coincides with TCE and therefore equals the conditional expectation of the α-tail of the loss distribution. However, the difference between CVaR and TCE becomes immediately apparent once we consider a discontinuous loss distribution. TCE and other tail risk measures like VaR and WCE are generally not continuous with respect to the confidence level α, hence changing the confidence level by a small amount may cause drastic jumps in their values. In contrast, CVaR is continuous with respect to α regardless of the underlying loss distribution.

Proposition 1.4 (Acerbi and Tasche [2002b]): *For any real-valued random variable L satisfying $\mathbb{E}(|L|) < \infty$, the mapping $\alpha \mapsto \mathrm{CVaR}^{\alpha}(L)$ is continuous on $(0, 1)$.*

Another prominent property of CVaR is that it is a *coherent risk measure* (in the sense of the definition given in Artzner et al. [1999]).

Proposition 1.5 (Acerbi and Tasche [2002b]): *Consider a fixed $\alpha \in (0, 1)$ and a set $\mathcal{V}$ of real-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}(|L|) < \infty$ for all $L \in \mathcal{V}$. Then $\rho : L \mapsto \mathbb{R}$ with $\rho(L) = \mathrm{CVaR}^{\alpha}(-L)$ is a coherent risk measure, i.e. it is:*

1. *monotonic*[1]*: $L_1, L_2 \in \mathcal{V}, L_1 \leqslant L_2 \Rightarrow \rho(L_1) \geqslant \rho(L_2)$;*

2. *sub-additive: $L_1, L_2 \in \mathcal{V} \Rightarrow \rho(L_1 + L_2) \leqslant \rho(L_1) + \rho(L_2)$;*

[1] Alternatively, the monotonicity property can be formulated as $L \in \mathcal{V}, L \geqslant 0 \Rightarrow \rho(L) \leqslant 0$.

3. positively homogeneous: $L \in \mathcal{V}, h > 0 \Rightarrow \rho(hL) = h\rho(L)$*;*

4. translation invariant: $L \in \mathcal{V}, a \in \mathbb{R} \Rightarrow \rho(L + a) = \rho(L) - a$*.*

Note that TCE and VaR fail to be coherent since they are not sub-additive. Violation of the sub-additivity property contradicts the diversification principle, hence the use of a non-sub-additive risk measure to quantify risk where the loss distribution is not absolutely continuous is questionable at best. WCE is coherent but it is initially defined in a way that makes it highly impractical to use it in applications.

It turns out that it is possible to compute both VaR and CVaR simultaneously by solving a certain one-dimensional convex optimization problem which is presented in the theorem below. For simplicity of notation, we shall use $\text{VaR}^\alpha(x)$, $\text{CVaR}^\alpha(x)$ and $\text{VaR}^\alpha(L(x))$, $\text{CVaR}^\alpha(L(x))$ interchangeably in what follows.

Theorem 1.1 (Rockafellar and Uryasev [2002]): *As a function of* z*, function*

$$F_\alpha(x, z) = z + \frac{1}{1-\alpha}\mathbb{E}\left((L(x) - z)^+\right), \tag{1.19}$$

where $y^+ = \max\{0, y\}$*, is finite and convex (hence continuous), and*

$$\text{CVaR}^\alpha(x) = \min_{z \in \mathbb{R}} F_\alpha(x, z), \tag{1.20}$$

$$\text{VaR}^\alpha(x) = \min\left\{y : y \in \text{argmin}_{z \in \mathbb{R}} F_\alpha(x, z)\right\}. \tag{1.21}$$

In particular, one always has

$$\mathrm{VaR}^\alpha(x) \in \mathrm{argmin}_{z\in\mathbb{R}} F_\alpha(x,z), \tag{1.22}$$

$$\mathrm{CVaR}^\alpha(x) = F_\alpha(x, \mathrm{VaR}^\alpha(x)). \tag{1.23}$$

An evident yet important corollary is that the problem of CVaR minimization over the set of available strategies may be expressed as a problem of $F_\alpha(x,z)$ minimization.

Corollary 1.1.1: Minimization of $\mathrm{CVaR}^\alpha(x)$ over the strategy set X is equivalent to minimization of $F_\alpha(x,z)$ over $X \times \mathbb{R}$:

$$\min_{x\in X} \mathrm{CVaR}^\alpha(x) = \min_{x\in X} \min_{z\in\mathbb{R}} F_\alpha(x,z). \tag{1.24}$$

1.3 The generalized Neyman-Pearson lemma

Consider probability measures $\mathbb{P}^*$ and $\mathbb{Q}$ on a measurable space $(\Omega, \mathcal{F})$ defining two distinct distributions. The problem in question is to discern between these distributions based on a single observed outcome $\omega \in \Omega$. Measure $\mathbb{P}^*$ shall be referred to as the (simple) *null hypothesis*, and $\mathbb{Q}$ shall correspond to the (simple) *alternative hypothesis*.

One of the ways of approaching this problem is to search for a solution in the form of a *pure test* — a rule that unambiguously defines the choice between the two measures for each elementary outcome. Such rule can be naturally expressed as a random variable $\varphi : \Omega \to \{0, 1\}$, which rejects $\mathbb{P}^*$ when $\varphi(\omega) = 1$. For a given test φ, the *probability of type-I error* (rejecting $\mathbb{P}^*$

when it is true) can be calculated as $\mathbb{P}^*(\varphi = 1)$, and the *probability of type-II error* (accepting $\mathbb{P}^*$ when it is false) can be calculated as $1-\mathbb{Q}(\varphi = 1)$. Since in general it's impossible to minimize both error probabilities at the same time, the usual method is to choose a significance level $\alpha \in (0,1)$ and minimize the probability of type-II error while ensuring that the probability of type-I error does not exceed α:

$$\begin{cases} \mathbb{Q}(\varphi = 1) \longrightarrow \max\limits_{\varphi \in \mathcal{R}_p}, \\ \mathbb{P}^*(\varphi = 1) \leqslant \alpha, \end{cases} \tag{1.25}$$

where $\mathcal{R}_p$ is the set of all pure tests on $(\Omega, \mathcal{F})$. Note that minimizing the probability of type-II error is equivalent to maximizing the *power of the test*, i.e. the probability of rejecting $\mathbb{P}^*$ when it is false.

Assume that there is a third measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ such that both $\mathbb{P}^*$ and $\mathbb{Q}$ are absolutely continuous with respect to $\mathbb{P}$, and denote the Radon-Nikodym derivates of $\mathbb{P}^*$ and $\mathbb{Q}$ by

$$Z_{\mathbb{P}^*} = \frac{d\mathbb{P}^*}{d\mathbb{P}}, \quad Z_{\mathbb{Q}} = \frac{d\mathbb{Q}}{d\mathbb{P}}. \tag{1.26}$$

The classic Neyman-Pearson theory provides the optimal likelihood-ratio test $\tilde{\varphi}$ in terms of $Z_{\mathbb{P}^*}$ and $Z_{\mathbb{Q}}$ for problem (1.25):

$$\tilde{\varphi} = \mathbf{1}_{\{\tilde{a} \cdot Z_{\mathbb{P}^*} < Z_{\mathbb{Q}}\}}, \tag{1.27}$$

provided that there exists such $\tilde{a} > 0$ that

$$\mathbb{P}^* \left(\tilde{a} \cdot Z_{\mathbb{P}^*} < Z_{\mathbb{Q}}\right) = \alpha. \tag{1.28}$$

In the case when the underlying distributions are absolutely continuous, it is always possible to find a positive $\tilde{a}$ that solves (1.28). However, in general the solution is not guaranteed to exist (in the form of a pure test) for any given α.

To ensure that the optimal test can be constructed for any significance level, the class of available tests has to be extended to include the randomized tests. A *randomized test* is defined by a random variable $\varphi : \Omega \to [0, 1]$, and for each observed outcome $\omega \in \Omega$ it rejects $\mathbb{P}^*$ with probability $\varphi(\omega)$. The probability of type-I error can then be expressed as

$$\mathbb{E}_{\mathbb{P}^*}(\varphi) = \int \varphi(\omega)\mathbb{P}^*(d\omega), \tag{1.29}$$

and the power of the test equals

$$\mathbb{E}_{\mathbb{Q}}(\varphi) = \int \varphi(\omega)\mathbb{Q}(d\omega). \tag{1.30}$$

Denote the set of all randomized tests on $(\Omega, \mathcal{F})$ by $\mathcal{R}$ (evidently, $\mathcal{R}_p \subset \mathcal{R}$). Similar to (1.25), the following problem is considered:

$$\begin{cases} \mathbb{E}_{\mathbb{Q}}(\varphi) \longrightarrow \max\limits_{\varphi \in \mathcal{R}}, \\ \mathbb{E}_{\mathbb{P}^*}(\varphi) \leqslant \alpha. \end{cases} \tag{1.31}$$

Theorem 1.2 (Cvitanić and Karatzas [2001]): *Problem (1.31) admits a solution for any $\alpha \in (0, 1)$, and the optimal randomized test $\tilde{\varphi}$ has the form*

$$\tilde{\varphi} = \mathbf{1}_{\{\tilde{a} \cdot Z_{\mathbb{P}^*} < Z_{\mathbb{Q}}\}} + \gamma \cdot \mathbf{1}_{\{\tilde{a} \cdot Z_{\mathbb{P}^*} = Z_{\mathbb{Q}}\}}, \tag{1.32}$$

where

$$\tilde{a} = \inf \left\{ a \geqslant 0 : \mathbb{P}^* \left(a \cdot Z_{\mathbb{P}^*} < Z_{\mathbb{Q}} \right) \leqslant \alpha \right\} \tag{1.33}$$

and

$$\gamma = \frac{\alpha - \mathbb{P}^* \left(\tilde{a} \cdot Z_{\mathbb{P}^*} < Z_{\mathbb{Q}} \right)}{\mathbb{P}^* \left(\tilde{a} \cdot Z_{\mathbb{P}^*} = Z_{\mathbb{Q}} \right)}. \tag{1.34}$$

The Neyman-Pearson lemma can also be generalized to account for compound hypotheses. Consider the problem of discriminating a family $\mathcal{P}^*$ of probability measures on $(\Omega, \mathcal{F})$ (*compound null hypothesis*) against another family $\mathcal{Q}$ of probability measures on $(\Omega, \mathcal{F})$ (*compound alternative hypothesis*). The objective is to minimize the probability of accepting $\mathcal{P}^*$ when it is false (type-II error) under the condition that the probability of rejecting $\mathcal{P}^*$ when it is false (type-I error) is less than a given significance level $\alpha \in (0, 1)$. Same as before, we search for a solution in the form of a randomized test $\varphi : \Omega \to [0, 1]$ which rejects $\mathcal{P}^*$ when $\varphi(\omega) = 1$, and we denote the class of such tests by $\mathcal{R}$. The problem is then to maximize the smallest power of a randomized test over all tests of size less or equal to α:

$$\begin{cases} \inf\limits_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbb{Q}}(\varphi) \longrightarrow \max\limits_{\varphi \in \mathcal{R}}, \\ \\ \sup\limits_{\mathbb{P}^* \in \mathcal{P}^*} \mathbb{E}_{\mathbb{P}^*}(\varphi) \leqslant \alpha. \end{cases} \tag{1.35}$$

Let all $\mathbb{P}^* \in \mathcal{P}^*$ and $\mathbb{Q} \in \mathcal{Q}$ be absolutely continuous with respect to a probability measure $\mathbb{P}$ and denote

$$Z_{\mathbb{P}^*} = \frac{d\mathbb{P}^*}{d\mathbb{P}}, \quad Z_{\mathbb{Q}} = \frac{d\mathbb{Q}}{d\mathbb{P}}, \quad Z_{\mathcal{P}^*} = \left\{ Z_{\mathbb{P}^*} : \mathbb{P}^* \in \mathcal{P}^* \right\}. \tag{1.36}$$

A solution to problem (1.35) when the set of densities $Z_{\mathcal{P}^*}$ is compact was suggested in [Rudloff, 2006] (a more recent published result can be found in Karatzas and Rudloff [2010]). Denote also the σ-algebra of all Borel sets of $Z_{\mathcal{P}^*}$ with $\mathcal{B}$, the set of all finite measures on $(Z_{\mathcal{P}^*}, \mathcal{B})$ with Λ_+ and the closure of the convex hull of densities $Z_{\mathbb{Q}}$ with respect to the norm topology in L^1 with $\overline{co}\mathcal{Q}$.

Theorem 1.3 (Rudloff [2006]): *Let $Z_{\mathcal{P}^*}$ be a compact set. Then problem (1.35) admits a solution for any $\alpha \in (0,1)$, and the optimal randomized test $\tilde{\varphi}$ has the form*

$$\tilde{\varphi} = \begin{cases} 0, & Z_{\tilde{\mathbb{Q}}} < \int_{\mathcal{P}^*} Z_{\mathbb{P}^*} d\tilde{\lambda}, \\ 1, & Z_{\tilde{\mathbb{Q}}} > \int_{\mathcal{P}^*} Z_{\mathbb{P}^*} d\tilde{\lambda}, \end{cases} \quad (\mathbb{P}\text{-}a.s.) \tag{1.37}$$

with

$$\mathbb{E}_{\mathbb{P}^*}(\tilde{\varphi}) = \alpha, \quad \tilde{\lambda}\text{-}a.s., \tag{1.38}$$

where $(\tilde{\mathbb{Q}}, \tilde{\lambda})$ is a solution to the dual problem

$$\mathbb{E}_{\mathbb{P}}\left(\left(Z_{\mathbb{Q}} - \int_{\mathcal{P}^*} Z_{\mathbb{P}^*} d\lambda\right)^+\right) + \alpha\lambda(Z_{\mathcal{P}^*}) \longrightarrow \min_{\mathbb{Q}\in\overline{co}\mathcal{Q}, \lambda\in\Lambda_+}. \tag{1.39}$$

Corollary 1.3.1: The optimal randomized test $\tilde{\varphi}$ in Theorem 1.3 can be expressed as

$$\tilde{\varphi} = \mathbf{1}_{\left\{Z_{\tilde{\mathbb{Q}}} > \int_{\mathcal{P}^*} Z_{\mathbb{P}^*} d\tilde{\lambda}\right\}} + \gamma \cdot \mathbf{1}_{\left\{Z_{\tilde{\mathbb{Q}}} = \int_{\mathcal{P}^*} Z_{\mathbb{P}^*} d\tilde{\lambda}\right\}}, \tag{1.40}$$

where random variable $\gamma \in \mathcal{R}$ is chosen to satisfy condition (1.38).

The above result plays a central role in the problems of composite hypotheses testing and partial hedging in incomplete markets in the papers [Rudloff, 2006], [Rudloff, 2005] and [Rudloff, 2009], with $\mathcal{P}^*$ being the set of equivalent sigma-martingale measures. However, the authors fail to notice the fact that in this case the set of densities $Z_{\mathcal{P}^*}$ is *never* compact, except when $\mathcal{P}^*$ is a singleton, which corresponds to the complete market case. The formal statement of this fact and the full proof can be found in Theorem 5.4 in Section 5.1. In a more recent paper by [Karatzas and Rudloff, 2010], the compactness assumption has been replaced with a weaker requirement of weak compactness. However, the set of densities of equivalent martingale measures is not weakly compact either (see Section 5.1), which makes this version of the Neyman-Pearson lemma hardly usable in mathematical finance in the incomplete market setting.

Therefore, we need to consider a more general version of the Neyman-Pearson lemma (or a result of the similar type) which does not require compactness. A corresponding theorem based on duality approach with respect to a state-dependent utility function shall be presented in Section 2.4 (Theorem 2.3) when we consider the problem of minimizing conditional value-at-risk in the incomplete market case.

1.4 The path-wise comparison theorem

In this section, we shall briefly quote a somewhat intuitive and a well-known statement in the theory of stochastic processes known as the path-wise comparison theorem, which we shall use in Chapter 4 to find an upper price bound an construct a corresponding approximate hedging strategy. Informally, the comparison theorem states that if two stochastic processes are driven by the same underlying process and have the same volatilities, the dominance of drifts leads to an almost-sure dominance of trajectories. As there exist many versions of the same result, we shall present one of the earliest ones suggested in [Yamada, 1973].

Theorem 1.4 (Yamada [1973]): *Consider the following two stochastic differential equations:*

$$dX_t = \mu_x(t, X_t)dt + \sigma(t, X_t)dW_t, \tag{1.41}$$

$$dY_t = \mu_y(t, Y_t)dt + \sigma(t, Y_t)dW_t, \tag{1.42}$$

where $\mu_x(t,x)$, $\mu_y(t,x)$ and $\sigma(t,x)$ are continuous in (t,x) on $\mathbb{R}_+ \times \mathbb{R}$, and there exists a positive increasing function $\rho(u)$, $u \in \mathbb{R}_+$ such that

$$|\sigma(t,x) - \sigma(t,y)| \leqslant \rho(|x-y|), \quad \forall x, y \in \mathbb{R} \tag{1.43}$$

and

$$\int_{\mathbb{R}_+} \rho^{-2}(u)du = \infty, \tag{1.44}$$

and, moreover,

$$\mu_x(t,x) \leqslant \mu_y(t,x), \quad \forall (t,x) \in \mathbb{R}_+ \times \mathbb{R}. \tag{1.45}$$

Under these conditions, $X_0 = Y_0$ implies $X_t \leqslant Y_t$ a.s. for every $t \in \mathbb{R}_+$.

Chapter 2

CVaR hedging: the theoretical results

In a complete financial market every contingent claim with payoff H delivered at time $t = T$ can be hedged perfectly: given the sufficient amount of the initial wealth, an agent who holds a short position in claim H can construct a hedging portfolio (V_0, ξ_t) that will replicate the liability without risk, that is, $V_T = H$, a.s. In an incomplete market not all claims are attainable and the equivalent risk-neutral measure is no longer unique, which makes the perfect hedging in general not possible; however, one can still stay on the safe side by employing the superhedging strategy (Föllmer and Leukert [1999]). Whatever the case, the cost of both perfect hedging and superhedging is commonly regarded as unreasonably high from a practical point of view.

The problem of partial hedging is to construct a portfolio that minimizes the risk of the difference $L = H - V_T$ subject to a constraint on the initial wealth. The efficiency and consistency of such an approach depend to a great extent on selecting a specific way of quantifying the risk. For instance, one of the most studied methods known as quadratic hedging suggests minimizing the quadratic error $\mathbb{E}\left(L^2\right)$. Despite its simplicity, this method has obvious disadvantages since the quadratic risk measure does not distinguish between loss and profit and equally penalizes both. Another method, known as quantile hedging, involves maximizing probability of a successful hedge $\mathbb{P}\left(L \leqslant 0\right)$ (see, for instance, Föllmer and Leukert [1999] or Cvitanić and Spivak [1999]). This approach was generalized in [Föllmer and Leukert, 2000] to address the problem of minimizing the expected shortfall $\mathbb{E}\left(L^+\right)$ and, more generally, $\mathbb{E}\left(l(L^+)\right)$ for an arbitrary loss function $l(\cdot)$.

We address the problem of partial hedging by minimizing conditional value-at-risk (CVaR) of loss at the terminal moment of time. Conditional value-at-risk is a quantile downside risk measure which is rapidly gaining popularity among professionals in both risk management and insurance and boasts many mathematically attractive features (it is a coherent quantile risk measure continuous with respect to the confidence level, see Section 1.2 for details). Our main objective in this section is to derive a dynamic hedging strategy which minimizes conditional value-at-risk associated with loss L subject to a constraint on the initial wealth; we also consider the dual problem: minimization of hedging costs subject to a constraint on CVaR. Note that a somewhat related problem was investigated in [Li and Xu, 2008] from the optimal investment point of view: minimization of CVaR when the returns are bounded; however, in the present work we focus mainly on derivatives hedging under capital constraints and the related applications.

In the following sections we suggest a method which can be used to construct CVaR-optimal hedging strategies explicitly in complete market models, which will be illustrated in the examples provided in Sections 3.1 and 3.2. We also provide the general structure of the solution for the problem of CVaR minimizaiton in the incomplete market case.

Note that the Sections 2.1, 2.2 and 2.3 are largely based on a recent article by [Melnikov and Smirnov, 2012].

2.1 Problem setup in the complete case

Let the discounted stock price[1] be governed by a stochastic process $X = (X_t)$ on a standard stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ with a fixed time horizon $T > 0$ and $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

A *self-financing strategy* is defined by initial wealth $V_0 > 0$ and a predictable process $\xi = (\xi_t)_{t\in[0,T]}$ which indicates the holding in the stock and essentially determines the portfolio dynamics. For each strategy (V_0, ξ) the corresponding *value process* V_t is defined as

$$V_t = V_0 + \int_0^t \xi_s dX_s, \quad \forall t \in [0, T]. \tag{2.1}$$

We shall call a strategy (V_0, ξ) *admissible* if it satisfies

$$V_t \geqslant 0, \quad \forall t \in [0, T], \quad \mathbb{P}\text{−a.s.}, \tag{2.2}$$

and we shall denote the set of all admissible self-financing strategies by $\mathcal{A}$.

Consider a contingent claim whose discounted payoff is an $\mathcal{F}_T$–measurable non-negative random variable $H \in L^1(\mathbb{P})$. In a complete market there exists

[1] It is implicitly assumed that there is a freely tradable risk-free asset in the market whose returns are known in advance; this asset's future prices determine the associated discounting factors. In the 'discounted' world where the stock price and all payoffs are discounted, the risk-free asset has zero return so it can be essentially viewed as cash.

a unique equivalent risk-neutral (martingale) measure $\mathbb{P}^* \approx \mathbb{P}$, under which the discounted stock price is a martingale, and construction of a perfect hedge is possible for any contingent claim. According to the risk-neutral valuation theory, the perfect hedging strategy requires allocating the initial wealth in the amount equal to the *fair price*

$$H_0 = \mathbb{E}^* (H), \tag{2.3}$$

where $\mathbb{E}^*$ denotes the expectation with respect to the risk-neutral measure $\mathbb{P}^*$.

The first problem that we shall investigate relates to the following question: what is the "best" hedge for H that can be achieved if the available amount of the initial wealth is limited from above by $\tilde{V}_0 \in (0, H_0)$? To specify the exact meaning of "best", we shall use the notion of conditional value-at-risk.

We define loss from the viewpoint of the claim seller who hedges a short position in H with portfolio (V_0, ξ), thus the realized loss at time T equals the claim value less the terminal value of the hedging portfolio:

$$L(V_0, \xi) = H - V_T = H - V_0 - \int_0^T \xi_s dX_s. \tag{2.4}$$

For a fixed confidence level $\alpha \in (0, 1)$, our problem is to find an admissible strategy (V_0, ξ) which minimizes conditional value-at-risk $\text{CVaR}^\alpha(V_0, \xi)$ associated with loss $L(V_0, \xi)$ while using no more initial wealth than $\tilde{V}_0$.

Another question relates to the dual problem: what is the least amount of the initial wealth that we have to allocate in order to keep CVaR of a given confidence level below a certain threshold?

2.2 Minimizing conditional value-at-risk

In this section we develop a method of constructing the optimal hedging strategy for the problem of CVaR minimization subject to a constraint on the initial wealth:

$$\begin{cases} \mathrm{CVaR}^{\alpha}(V_0, \xi) \longrightarrow \min\limits_{(V_0,\xi)\in\mathcal{A}}, \\ \\ V_0 \leqslant \tilde{V}_0. \end{cases} \tag{2.5}$$

Denote by $\mathcal{A}_{\tilde{V}_0}$ the set of all admissible strategies satisfying the wealth constraint:

$$\mathcal{A}_{\tilde{V}_0} = \left\{ (V_0, \xi) : (V_0, \xi) \in \mathcal{A}, V_0 \leqslant \tilde{V}_0 \right\}. \tag{2.6}$$

According to Corollary 1.1.1, problem (2.5) is equivalent to

$$F_{\alpha}\left((V_0, \xi), z\right) \longrightarrow \min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}} \min_{z\in\mathbb{R}}. \tag{2.7}$$

Recall that F_{α} is defined by (1.19) and the loss function for this problem is given by (2.4), so (2.7) becomes

$$z + \frac{1}{1-\alpha} \cdot \mathbb{E}\left((H - V_T - z)^{+}\right) \longrightarrow \min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}} \min_{z\in\mathbb{R}}. \tag{2.8}$$

Let us introduce an auxiliary real-valued function

$$c(z) = z + \frac{1}{1-\alpha} \cdot \min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}} \mathbb{E}\left((H - V_T - z)^{+}\right), \tag{2.9}$$

so that

$$\min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}} \mathrm{CVaR}^{\alpha}(V_0, \xi) = \min_{z\in\mathbb{R}} \; c(z). \tag{2.10}$$

Assume that for each $z \in \mathbb{R}$ the minimum in (2.9) is attained at $(\hat{V}_0(z), \hat{\xi}(z))$ and that $c(z)$ reaches its global minimum at point $\hat{z}$:

$$\min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}} \mathbb{E}\left((H - V_T - z)^+\right) = \mathbb{E}\left((H - \hat{V}_T(z) - z)^+\right), \tag{2.11}$$

$$\min_{z\in\mathbb{R}} c(z) = c\left(\hat{z}\right). \tag{2.12}$$

Then strategy $(\hat{V}_0, \hat{\xi}) = (\hat{V}_0\left(\hat{z}\right), \hat{\xi}\left(\hat{z}\right))$ is the optimal strategy for (2.5):

$$\min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}} \mathrm{CVaR}^\alpha(V_0, \xi) = \mathrm{CVaR}^\alpha(\hat{V}_0\left(\hat{z}\right), \hat{\xi}\left(\hat{z}\right)). \tag{2.13}$$

The definition (2.9) of function $c(z)$ is not immediately helpful as it contains expected value minimization. Deriving a closed-form expression for $(\hat{V}_0(z), \hat{\xi}(z))$ for each z or, equivalently, deriving an explicit expression for $c(z)$ would allow us to reduce the initial problem (2.5) to a problem of one-dimensional optimization with respect to z.

For each z, strategy $(\hat{V}_0(z), \hat{\xi}(z))$ is a solution for the following problem:

$$\mathbb{E}\left((H - V_T - z)^+\right) \longrightarrow \min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}} \tag{2.14}$$

Since H and V_T are both non-negative, (2.14) can be rewritten as

$$\mathbb{E}\left(((H - z)^+ - V_T)^+\right) \longrightarrow \min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}}. \tag{2.15}$$

Moreover,

$$\left((H - z)^+ - V_T\right)^+ = (H - z)^+ - V_T \wedge (H - z)^+ = (1 - \varphi(z))(H - z)^+, \tag{2.16}$$

where $\varphi(z) : \Omega \to [0,1]$ is such an $\mathcal{F}_T$-measurable random variable ("success ratio") that

$$\varphi(z)(H-z)^+ = V_T \wedge (H-z)^+. \tag{2.17}$$

It is shown in [Föllmer and Leukert, 2000] (Theorem 3.2) that problem (2.15) is equivalent to the problem of finding the optimal success ratio $\tilde{\varphi}(z)$:

$$\begin{cases} \mathbb{E}\left((1-\varphi(z))(H-z)^+\right) \longrightarrow \min\limits_{\varphi(z)\in\mathcal{R}}, \\ \mathbb{E}^*\left(\varphi(z)(H-z)^+\right) \leqslant \tilde{V}_0, \end{cases} \tag{2.18}$$

where $\mathcal{R}$ is the set of all $\mathcal{F}_T$-measurable random variables taking on values from 0 to 1. The optimal strategy $(\hat{V}_0(z), \hat{\xi}(z))$ can then be found as a perfect hedge to the modified claim $\tilde{H}(z) = (H-z)^+\tilde{\varphi}(z)$. Problem (2.18), in its turn, can be rearranged to match the Neyman-Pearson class of problems:

$$\begin{cases} \mathbb{E}_{\mathbb{Q}}\left(\varphi(z)\right) \longrightarrow \max\limits_{\varphi(z)\in\mathcal{R}}, \\ \mathbb{E}_{\mathbb{Q}^*}\left(\varphi(z)\right) \leqslant \dfrac{\tilde{V}_0}{\mathbb{E}^*\left((H-z)^+\right)}, \end{cases} \tag{2.19}$$

where measures $\mathbb{Q}$ and $\mathbb{Q}^*$ are defined via their Radon-Nikodym derivatives:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{(H-z)^+}{\mathbb{E}\left((H-z)^+\right)}, \quad \frac{d\mathbb{Q}^*}{d\mathbb{P}^*} = \frac{(H-z)^+}{\mathbb{E}^*\left((H-z)^+\right)}. \tag{2.20}$$

Applying Theorem 1.2 (the Neyman-Pearson lemma for simple hypotheses) to (2.19) and expressing the results in terms of measures $\mathbb{P}$ and $\mathbb{P}^*$ gives us the optimal randomized test in the form

$$\tilde{\varphi}(z) = \mathbf{1}_{\left\{\frac{d\mathbb{P}}{d\mathbb{P}^*} > \tilde{a}(z)\right\}} + \gamma(z) \cdot \mathbf{1}_{\left\{\frac{d\mathbb{P}}{d\mathbb{P}^*} = \tilde{a}(z)\right\}}, \tag{2.21}$$

where

$$\tilde{a}(z) = \inf\left\{a \geqslant 0 : \mathbb{E}^*\left((H-z)^+ \cdot \mathbf{1}_{\left\{\frac{d\mathbb{P}}{d\mathbb{P}^*} > a\right\}}\right) \leqslant \tilde{V}_0\right\}, \tag{2.22}$$

$$\gamma(z) = \frac{\tilde{V}_0 - \mathbb{E}^*\left((H-z)^+ \cdot \mathbf{1}_{\left\{\frac{d\mathbb{P}}{d\mathbb{P}^*} > \tilde{a}(z)\right\}}\right)}{\mathbb{E}^*\left((H-z)^+ \cdot \mathbf{1}_{\left\{\frac{d\mathbb{P}}{d\mathbb{P}^*} = \tilde{a}(z)\right\}}\right)}. \tag{2.23}$$

Note that we assume convention $0/0 = 0$, so that

$$\gamma(z) = 0, \quad \text{if } \mathbb{P}^*\left(\left\{\frac{d\mathbb{P}}{d\mathbb{P}^*} = \tilde{a}(z)\right\} \cap \{H > 0\}\right) = 0, \tag{2.24}$$

which implies that the optimal randomized test reduces to a pure test if the underlying distribution is atomless.

Let us collect the above results in the following theorem.

Theorem 2.1: *The optimal strategy $(\hat{V}_0, \hat{\xi})$ for the problem of CVaR minimization (2.5) is a perfect hedge for the modified contingent claim $\tilde{H}(\hat{z}) = (H-\hat{z})^+ \tilde{\varphi}(\hat{z})$:*

$$\mathbb{E}^*\left(\tilde{H}(z) \mid \mathcal{F}_t\right) = \hat{V}_0(z) + \int_0^t \hat{\xi}_s(z) dX_s, \quad \mathbb{P}\text{-}a.s., \quad \forall t \in [0,T], \tag{2.25}$$

where $\tilde{\varphi}(z)$ is defined by (2.21), $\hat{z}$ is the point of global minimum of function

$$c(z) = \begin{cases} z + \dfrac{1}{1-\alpha} \cdot \mathbb{E}\left((H-z)^+ (1-\tilde{\varphi}(z))\right), & \text{for } z < z^*, \\ z, & \text{for } z \geqslant z^* \end{cases} \tag{2.26}$$

on interval $z \in [0, z^*]$, and z^* is a real root of equation

$$\tilde{V}_0 = \mathbb{E}^*\left((H-z^*)^+\right). \tag{2.27}$$

Besides, one always has

$$\mathrm{CVaR}^{\alpha}\left(\hat{V}_0, \hat{\xi}\right) = c\left(\hat{z}\right), \tag{2.28}$$

$$\mathrm{VaR}^{\alpha}\left(\hat{V}_0, \hat{\xi}\right) = \min\left\{z : z \in \operatorname{argmin}_{z\in[0,z^*]}\ c(z)\right\}. \tag{2.29}$$

Note that it only makes sense to search for a non-trivial solution to problem (2.15) when $\tilde{V}_0 < \mathbb{E}^*\left((H-z)^+\right)$, otherwise a perfect hedge for $(H-z)^+$ can be used as the optimal strategy, providing zero expected shortfall. As a function of z,

$$h(z) = \mathbb{E}^*\left((H-z)^+\right) \tag{2.30}$$

is monotonically non-increasing, and

$$h(0) = H_0 > \tilde{V}_0, \tag{2.31}$$

$$\lim_{z\to\infty} h(z) = 0, \tag{2.32}$$

so there always exists such $z^* > 0$ that

$$\begin{cases} \mathbb{E}^*\left((H-z)^+\right) \geqslant \tilde{V}_0, & \text{for } z \leqslant z^*, \\ \mathbb{E}^*\left((H-z)^+\right) < \tilde{V}_0, & \text{for } z > z^*. \end{cases} \tag{2.33}$$

Therefore, when z is greater than z^*, a perfect hedge for $(H-z)^+$ is the optimal solution for (2.15), which explains why $c(z) = z$ for $z \geqslant z^*$.

According to Theorem 1.1, the leftmost point of the argminimum of $c(z)$ coincides with the value-at-risk of the CVaR-optimal hedge, so

$$\mathrm{VaR}^{\alpha}\left(\hat{V}_0, \hat{\xi}\right) \leqslant \hat{z}. \tag{2.34}$$

Besides, the loss function is always non-negative:

$$L(z) = H - \tilde{H}(z) = H - \tilde{\varphi}(z)(H - z)^{+} \geqslant 0, \tag{2.35}$$

therefore the value-at-risk associated with the optimal strategy would also be non-negative, which in conjunction with (2.34) implies $\hat{z} \geqslant 0$.

Finally, $c(z) = z$ for $z \geqslant z^{*}$ and $c(z)$ is increasing at $z = z^{*}$, hence the global minimum of $c(z)$ coincides with its minimum on $[0, z^{*}]$.

2.3 Minimizing hedging costs

In this section we shall focus on minimizing the initial wealth required for construction of a hedging strategy over all admissible strategies (V_0, ξ) under the condition that conditional value-at-risk of a given confidence level α does not exceed a predefined threshold $\tilde{C}$:

$$\begin{cases} V_0 \longrightarrow \min\limits_{(V_0,\xi)\in\mathcal{A}}, \\ \mathrm{CVaR}^{\alpha}(V_0, \xi) \leqslant \tilde{C}. \end{cases} \tag{2.36}$$

Let us rephrase the problem in terms of the terminal capital $V_T = V_0 + \int_0^T \xi_s dX_s$ (we can always derive a corresponding admissible trading strategy (V_0, ξ) by constructing a perfect hedge for V_T):

$$\begin{cases} \mathbb{E}^{*}\left(V_T\right) \longrightarrow \min\limits_{V_T\in\mathcal{V}^{+}}, \\ \mathrm{CVaR}^{\alpha}(V_T) \leqslant \tilde{C}, \end{cases} \tag{2.37}$$

where $\mathcal{V}^+$ is the class of all non-negative $\mathcal{F}_T$-measurable random variables. Recall that

$$\mathrm{CVaR}^\alpha(V_0,\xi) = \min_{z\in\mathbb{R}} \left(z + \frac{1}{1-\alpha}\mathbb{E}\left((H - V_T - z)^+\right)\right), \qquad (2.38)$$

and consider the following family of problems:

$$\begin{cases} \mathbb{E}^*(V_T) \longrightarrow \min\limits_{V_T\in\mathcal{V}^+}, \\ \mathbb{E}\left((H - V_T - z)^+\right) \leqslant \tilde{C}_z, \end{cases} \qquad (2.39)$$

where

$$\tilde{C}_z = (\tilde{C} - z)(1-\alpha). \qquad (2.40)$$

In order to establish the link between problems (2.37) and (2.39), we shall make use of the following lemma.

Lemma 2.1: *Let $\tilde{x}$ be a solution for*

$$\begin{cases} f(x) \longrightarrow \min\limits_{x\in\mathcal{X}}, \\ \min\limits_{z\in\mathbb{R}} \; g(x,z) \leqslant c. \end{cases} \qquad (2.41)$$

Then the following family of problems also admit solutions, denoted $\tilde{x}(z)$:

$$\begin{cases} f(x) \longrightarrow \min\limits_{x\in\mathcal{X}}, \\ g(x,z) \leqslant c. \end{cases} \qquad (2.42)$$

Besides, one always has

$$\tilde{x} = \tilde{x}(\tilde{z}), \qquad (2.43)$$

where z is a point of global minimum of $f(\tilde{x}(z))$:

$$\min_{z\in\mathbb{R}} \; f(\tilde{x}(z)) = f(\tilde{x}(\tilde{z})). \tag{2.44}$$

Denote

$$\mathcal{G}_z = \{x \;:\; g(x,z) \leqslant c\}\,, \tag{2.45}$$

$$\mathcal{G} = \left\{x \;:\; \min_{z\in\mathbb{R}} \; g(x,z) \leqslant c\right\}. \tag{2.46}$$

Then for each $z \in \mathbb{R}$

$$\bigcup_{z\in\mathbb{R}} \mathcal{G}_z = \mathcal{G}, \tag{2.47}$$

and

$$\bigcup_{z\in\mathbb{R}} (\mathcal{X} \cap \mathcal{G}_z) = \mathcal{X} \cap \mathcal{G}. \tag{2.48}$$

Therefore,

$$\min_{x\in\mathcal{X}\cap\mathcal{G}} \; f(x) = \min_{z\in\mathbb{R}} \left(\min_{x\in\mathcal{X}\cap\mathcal{G}_z} \; f(x) \right), \tag{2.49}$$

which proves the lemma. □

Denote the solution for the parametrized subproblem (2.39) for each $z \in \mathbb{R}$ by $\tilde{V}_T(z)$, then, according to Lemma 2.1, the solution for the original problem (2.37) can be expressed in the form

$$\tilde{V}_T = \tilde{V}_T(\tilde{z}), \tag{2.50}$$

where

$$\mathbb{E}^* \left(\tilde{V}_T(\tilde{z}) \right) = \min_{z\in\mathbb{R}} \; \mathbb{E}^* \left(\tilde{V}_T(z) \right). \tag{2.51}$$

We shall now focus on deriving $\tilde{V}_T(z)$ explicitly by solving (2.39). To begin with, note that in case $z > \tilde{C}$ the problem admits no solution since the left-hand side of the constraint in (2.39) is always non-negative.

In case $z \leqslant \tilde{C}$ we have

$$(H - V_T - z)^+ = ((H - z)^+ - V_T)^+ = (1 - \varphi(z))(H - z)^+, \tag{2.52}$$

where $\varphi(z)$ is a randomized test defined the same way as in the previous section. Problem (2.39) can then be rewritten in terms of $\varphi(z)$:

$$\begin{cases} \mathbb{E}\left((H - z)^+ \varphi(z)\right) \leqslant \tilde{C}_z, \\ \mathbb{E}^*\left((H - z)^+ \varphi(z)\right) \longrightarrow \max\limits_{\varphi(z) \in \mathcal{R}}. \end{cases} \tag{2.53}$$

With the help of the two auxiliary measures $\mathbb{Q}$ and $\mathbb{Q}^*$ defined in (2.20), we can transmute (2.53) into a classic Neyman-Pearson problem:

$$\begin{cases} \mathbb{E}_{\mathbb{Q}}\left(\varphi(z)\right) \leqslant \dfrac{\tilde{C}_z}{\mathbb{E}\left((H - z)^+\right)}, \\ \mathbb{E}_{\mathbb{Q}^*}\left(\varphi(z)\right) \longrightarrow \max\limits_{\varphi(z) \in \mathcal{R}}. \end{cases} \tag{2.54}$$

It it now straightforward to apply the Neyman-Pearson lemma (Theorem 1.2) to (2.54) to obtain the optimal solution. However, one thing should be noted before we do that: problem (2.54) admits a non-trivial solution only in the case when $\mathbb{E}\left((H - z)^+\right) > \tilde{C}_z$, since otherwise the optimal test has the form $\tilde{\varphi}(z) \equiv 1$ (and, consequently, $\tilde{V}_T(z) \equiv 0$).

Lemma 2.2: *Condition*

$$\mathbb{E}\left((H - z)^+\right) > \tilde{C}_z \tag{2.55}$$

is satisfied for all $z \leqslant \tilde{C}$ if and only if both of the following inequalities hold true:

$$\begin{cases} \mathbb{E}(H) > \tilde{C}_0, \\ \mathbb{E}\left((H-\tilde{C})^+\right) > 0. \end{cases} \tag{2.56}$$

Note that both the right- and the left-hand sides of (2.55) are monotonically non-increasing functions of z. In addition,

$$\frac{d}{dz}\mathbb{E}\left((H-z)^+\right) = -1, \qquad \text{for } z < 0, \tag{2.57}$$

and

$$\frac{d}{dz}\tilde{C}_z = -1 + \alpha, \tag{2.58}$$

so it is necessary and sufficient that (2.55) holds true at points $z = 0$ and $z = \tilde{C}$ only, which implies (2.56) and thus proves the lemma. □

Lemma 2.2 provides an immediate way to check whether condition (2.55) is satisfied for all $z \leqslant \tilde{C}$ or not. If it turns out that there exists such $z = z^* \leqslant \tilde{C}$ that it doesn't hold true, then $\tilde{V}_T(z^*) \equiv 0$ and, according to (2.50) and (2.51), the solution for the initial problem (2.37) would also have the form $\tilde{V}_T \equiv 0$, which can be interpreted as adopting the passive trading strategy. Indeed, if the first inequality in (2.56) is not satisfied, the target CVaR value is too high in comparison with the expected payoff on the contingent claim, so there is no need to hedge at all; if the second inequality is not satisfied, the payoff is bounded from above by a value less than $\tilde{C}$, so CVaR can never reach its threshold value no matter what hedging strategy is used.

We shall summarize the results of this section in the following theorem.

Theorem 2.2: *The optimal strategy* $(\hat{V}_0, \hat{\xi})$ *for the problem of hedging costs minimization (2.36) is*

a) *a perfect hedge for the contingent claim* $(H-\hat{z})^+(1-\tilde{\varphi}(\hat{z}))$ *if condition (2.56) holds true, where* $\tilde{\varphi}(z)$ *is defined by*

$$\tilde{\varphi}(z) = \mathbf{1}_{\left\{\frac{d\mathbb{P}^*}{d\mathbb{P}} > \tilde{a}(z)\right\}} + \gamma(z) \cdot \mathbf{1}_{\left\{\frac{d\mathbb{P}^*}{d\mathbb{P}} = \tilde{a}(z)\right\}}, \tag{2.59}$$

$$\tilde{a}(z) = \inf\left\{a \geqslant 0 : \mathbb{E}\left((H-z)^+ \cdot \mathbf{1}_{\left\{\frac{d\mathbb{P}^*}{d\mathbb{P}} > a\right\}}\right) \leqslant \tilde{C}_z\right\}, \tag{2.60}$$

$$\gamma(z) = \frac{\tilde{C}_z - \mathbb{E}\left((H-z)^+ \cdot \mathbf{1}_{\left\{\frac{d\mathbb{P}^*}{d\mathbb{P}} > \tilde{a}(z)\right\}}\right)}{\mathbb{E}\left((H-z)^+ \cdot \mathbf{1}_{\left\{\frac{d\mathbb{P}^*}{d\mathbb{P}} = \tilde{a}(z)\right\}}\right)}, \tag{2.61}$$

and $\hat{z}$ *is a point of minimum of function*

$$d(z) = \mathbb{E}^*\left((H-z)^+(1-\tilde{\varphi}(z))\right) \tag{2.62}$$

on interval $-\infty < z \leqslant \tilde{C}$*;*

b) *a passive trading strategy:*

$$\hat{V}_t = 0, \qquad t \in [0, T], \tag{2.63}$$

if condition (2.56) is not satisfied.

2.4 Minimizing CVaR in the incomplete case

In an incomplete financial market not every contingent claim is attainable and the set of equivalent martingale measures is no longer a singleton. In this case, the perfect hedging is no longer applicable, yet it is still possible to completely eliminate the shortfall risk by using a superhedging strategy (El Karoui and Quenez [1995], Föllmer and Leukert [1999]). However, the cost of superhedging (the superhedging price) which is equal to the supremum of expected values over all equivalent martingale measures is often considered unreasonably high from the practical point of view. We investigate the situation where the amount of available initial wealth is less than the superhedging price. We shall use the CVaR risk measure to quantify the shortfall risk and minimize it over the set of self-financing strategies subject to a constraint on the the initial wealth. Towards this end, we shall use the general ideas laid out in Sections 2.2 and 2.3 along with the duality approach presented in [Rudloff, 2006] and [Xu, 2004].

Note that the problem of minimizing expected linear shortfall in the incomplete market setting in discrete-time case was thoroughly investigated in [Schulmerich and Trautmann, 2003], where it was stated that finding the optimal modified claim can be reduced to a certain linear program due to the fact that $\mathcal{P}_\sigma^*$ is a convex polyhedron (e.g., in the multinomial market model). Using the techniques presented in this chapter, these results can be used to explicitly construct CVaR-optimal strategies in discrete-time models.

The discounted stock price process is assumed to be a semimartingale $X = (X_t)_{t\in[0,T]}$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ with a fixed time horizon $T > 0$ and $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Let $\mathcal{P}_\sigma^*$ denote the set of probability measures equivalent to $\mathbb{P}$ under which X is a sigma-martingale (a semimartingale with an integral representation, refer to Delbaen and Schachermayer [2006] for details). We assume that there are no arbitrage opportunities in the market in the sense that $\mathcal{P}_\sigma^*$ is non-empty[2]. Note that all equations and inequalities in this section involving random variables are understood as $\mathbb{P}$-a.s. unless explicitly stated otherwise.

Similar to Section 2.1, a self-financing strategy is defined by the initial wealth $V_0 > 0$ and a predictable process $\xi = (\xi_t)_{t\in[0,T]}$ so that the value process V_t has the form

$$V_t = V_0 + \int_0^t \xi_s dX_s, \quad \forall t \in [0,T]. \tag{2.64}$$

Denote by $\mathcal{A}$ the set of all admissible self-financing strategies (V_0, ξ) that satisfy

$$V_t \geqslant 0, \quad \forall t \in [0,T], \quad \mathbb{P}\text{−a.s.}, \tag{2.65}$$

which ensures that the value process is well-defined. Denote also

$$\mathcal{A}_{\tilde{V}_0} = \left\{ (V_0, \xi) : (V_0, \xi) \in \mathcal{A}, V_0 \leqslant \tilde{V}_0 \right\}. \tag{2.66}$$

Consider a contingent claim whose discounted payoff is an $\mathcal{F}_T$–measurable non-negative random variable $H \in L^1(\mathbb{P})$. The smallest amount H_0 of initial wealth for which there exists an admissible strategy satisfying $V_T \geqslant H$ is called the *superhedging price* (Delbaen and Schachermayer [2006]), which is assumed to be finite:

$$H_0 = \sup_{\mathbb{P}^* \in \mathcal{P}_\sigma^*} \mathbb{E}^*(H) < +\infty, \tag{2.67}$$

where $\mathbb{E}^*$ denotes the expectation with respect to measure $\mathbb{P}^*$. In a complete market, the superhedging price equals the unique arbitrage-free ("fair") price

[2]This type of no-arbitrage condition is known as "no free lunch with vanishing risk" (NFLVR), see e.g. Delbaen and Schachermayer [2006].

of the contingent claim, while in an incomplete market, it coincides with the upper bound of the arbitrage-free price interval.

We shall now investigate the problem of finding a hedge for H that minimizes the shortfall risk in the case when the available amount of initial capital is limited by $\tilde{V}_0 \in (0, H_0]$ and construction of the superhedging strategy is not possible. Same as before, the shortfall risk shall be measured by CVaR at a pre-defined confidence level $\alpha \in (0, 1)$, with the loss function defined as in (2.4). The resulting optimization problem thus has the following form:

$$\mathrm{CVaR}^{\alpha}(V_0, \xi) \longrightarrow \min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}} . \tag{2.68}$$

By following the same logic that we used in Section 2.2, the optimal strategy $(\hat{V}_0, \hat{\xi})$ for the problem of CVaR minimization (2.68) can be expressed as $(\hat{V}_0, \hat{\xi}) = (\hat{V}_0(\hat{z}), \hat{\xi}(\hat{z}))$, where strategy $(\hat{V}_0(z), \hat{\xi}(z))$ is the solution to

$$\mathbb{E}\left((H - V_T - z)^+\right) \longrightarrow \min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}} , \tag{2.69}$$

for each $z \in \mathbb{R}$, and $\hat{z}$ is the point of global minimum of function

$$c(z) = z + \frac{1}{1-\alpha} \cdot \min_{(V_0,\xi)\in\mathcal{A}_{\tilde{V}_0}} \mathbb{E}\left((H - V_T - z)^+\right) . \tag{2.70}$$

Problem (2.69) can be viewed as a problem of linear expected shortfall minimization where the contingent claim is parametrized by a real-valued parameter z. It is no longer possible to apply the classic Neyman-Pearson lemma to derive the explicit solution like we did in Section 2.2. However, it is still possible to show that the optimal test still has the typical 0-1 structure.

In [Föllmer and Leukert, 2000], it is shown that the problem can be broken into two independent subproblems — finding the optimal modified claim (the static part) and then constructing the optimal strategy as a superhedge for the modified claim (the dynamic part), however it is not specified how one can obtain the modified claim in the incomplete market setting. In [Rudloff, 2006], a Fenchel duality method is employed to deal with this problem when the market is incomplete under a somewhat vague assumption that the set of densities of equivalent martingale measures is compact. Through this approach, it is possible to express the optimal randomized test explicitly in the terms of a least favorable distribution defined on the set of equivalent martingale measures. However, it turns out that the set of densities of equivalent martingale measures is never compact unless the market is complete (see Section 5.1 for details), hence a more general approach is required that doesn't assume compactness. In what follows, we shall use the utility-based method suggested in [Rudloff, 2006] and [Xu, 2004], which is largely based on the duality results for state-dependent utility functions proposed in [Kramkov and Schachermayer, 1999].

Denote by $\mathcal{V}(x)$ the set of admissible self-financing value processes V starting at the initial wealth $x > 0$:

$$\mathcal{V}(x) = \{V : V_t = x + \int_0^t \xi_s dS_s \geqslant 0, \quad t \in [0, T]\}, \tag{2.71}$$

and the set of contingent claims that are super-replicable by some admissible self-financing strategy with the initial wealth $x > 0$:

$$\mathcal{C}(x) = \{g \in L^0 : 0 \leqslant g \leqslant V_T \text{ for some } V \in \mathcal{V}(x)\}, \tag{2.72}$$

where $L^0 = L^0(\Omega, \mathcal{F}, \mathbb{P})$ is the set of all random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.

For each $z \in \mathbb{R}$, define the state-dependent utility function $U_z : \mathbb{R}_+ \times \Omega \mapsto \mathbb{R}_+$ as

$$U_z(x, \omega) = (H(\omega) - z)^+ \wedge x, \tag{2.73}$$

then problem (2.69) (primal problem) becomes

$$\begin{aligned} u_z(\tilde{V}_0) &= \sup_{V \in \mathcal{V}(\tilde{V}_0)} \mathbb{E}\left(U_z(V_T(\omega), \omega)\right) \\ &= \sup_{g \in \mathcal{C}(\tilde{V}_0)} \mathbb{E}(U_z(g(\omega), \omega) \\ &= \sup_{g \in \mathcal{C}(\tilde{V}_0)} \mathbb{E}((H - z)^+ \wedge g. \end{aligned} \tag{2.74}$$

As in [Kramkov and Schachermayer, 1999], the dual space is defined as a set of processes Y such that

$$\mathcal{Y}(y) = \{Y \geqslant 0 : Y_0 = y \text{ and } VY \text{ is a } \mathbb{P}\text{-supermartingale } \forall V \in \mathcal{V}(1)\}, \tag{2.75}$$

and the dual extended set $\mathcal{D}\,(y)$ of random variables h is defined by

$$\mathcal{D}\,(y) = \{h \in L^0 : 0 \leqslant h \leqslant Y_T \text{ for some } Y \in \mathcal{Y}(y)\}. \tag{2.76}$$

Then the dual problem is to minimize the expected value of the conjugate function:

$$v(\tilde{y}(z)) = \mathbb{E}(V(Y_T) \to \inf_{Y \in \mathcal{Y}(y)}, \tag{2.77}$$

for a suitable $\tilde{y}(z) \geqslant 0$.

We shall now present a theorem which provides the structure of the optimal modified claim, which is an adapted version of Theorem 4.38 in [Rudloff, 2006] with $\mathbb{Q} = \mathbb{P}$.

Theorem 2.3: *The optimal strategy $(\hat{V}_0(z), \hat{\xi}(z))$ in problem (2.69) is a superhedging strategy for a modified claim $(H-z)^+\tilde{\varphi}(z)$, where for each $z \in \mathbb{R}$*

$$\tilde{\varphi}(z) = \begin{cases} 0, & \tilde{h}(z) > 1 \\ 1 & 0 \leqslant \tilde{h}(z) < 1, \end{cases} \quad \mathbb{P}\text{-}a.s. \tag{2.78}$$

with

$$\mathbb{E}\left(\tilde{\varphi}(z)(H-z)^+\tilde{h}(z)\right) = \tilde{V}_0\tilde{y}(z), \tag{2.79}$$

where $\tilde{y}(z) \in \partial u_z(\tilde{V}_0)$ is assumed to satisfy $\tilde{y}(z) \geqslant 0$ and $\tilde{h}(z) \in \mathcal{D}(\tilde{y}(z))$ solves

$$\mathbb{E}\left((1-h)^+(H-z)^+\right) \to \inf_{h \in \mathcal{D}(\tilde{y}(z))}. \tag{2.80}$$

Remark 2.1: The optional decomposition theorem (Föllmer and Kabanov [1998]) ensures the existence of the superhedging strategy along with an increasing optional process $\hat{C}_t(z)$ with $\hat{C}_0(z) = 0$ such that

$$\operatorname*{ess\,sup}_{\mathbb{P}^* \in \mathcal{P}^*_\sigma} \mathbb{E}^*\left(\tilde{\varphi}(z)(H-z)^+ \mid \mathcal{F}_t\right) = \hat{V}_0(z) + \int_0^t \hat{\xi}_s(z) dX_s - \hat{C}_t(z). \tag{2.81}$$

Let us now summarize the above results in the form of a theorem.

Theorem 2.4: *The optimal strategy $(\hat{V}_0, \hat{\xi})$ for the problem of CVaR minimization (2.68) is a superhedging strategy for the modified contingent claim $\tilde{H}(\hat{z}) = (H-\hat{z})^+ \tilde{\varphi}(\hat{z})$:*

$$\operatorname*{ess\,sup}_{\mathbb{P}^* \in \mathcal{P}^*_\sigma} \mathbb{E}^*\left(\tilde{H}(\hat{z}) \mid \mathcal{F}_t\right) = \hat{V}_0(\hat{z}) + \int_0^t \hat{\xi}_s(\hat{z}) dX_s - \hat{C}_t, \tag{2.82}$$

where $\tilde{\varphi}(z)$ is defined by (2.78), $\hat{z}$ is the point of global minimum of function

$$c(z) = \begin{cases} z + \dfrac{1}{1-\alpha} \cdot \mathbb{E}\left((H-z)^{+}\left(1-\tilde{\varphi}(z)\right)\right), & \textit{for } z < z^{*}, \\ z, & \textit{for } z \geqslant z^{*} \end{cases} \tag{2.83}$$

on interval $z \in [0, z^{*}]$, and z^{*} is a real root of equation

$$\tilde{V}_{0} = \sup_{\mathbb{P}^{*} \in \mathcal{P}_{\sigma}^{*}} \mathbb{E}^{*}\left((H-z^{*})^{+}\right). \tag{2.84}$$

Remark 2.2: The explanation of why we only need to consider the interval $[0, z^{*}]$ when minimizing $c(z)$ can be found in Section 2.2 following Theorem 2.1.

Chapter 3

CVaR hedging: applications and examples

In this chapter, we demonstrate how CVaR-optimal hedging strategies (and, alternatively, hedging strategies that minimize hedging costs) can be explicitly constructed in complete market models.

First, we consider the problem of conditional value-at-risk minimization in the framework of the classical Black-Scholes model. We provide numerical illustrations for various confidence levels and investigate the structure of the optimal hedging strategy. We also observe how the suggested hedging strategy performs in the case when a CVaR-based regulatory capital requirement is imposed. Then, we apply our technique to the telegraph market model, in which we suggest an efficient method of computing expected values with respect to both the historical and the risk-neutral measure. Finally, we use the results obtained for the Black-Scholes model to derive the relationship between the optimal age of the insured and the minimal conditional value-at-risk of the financial component of a pure endowment unit-linked insurance contract.

3.1 Hedging a call option in the Black-Scholes model

Consider a fixed time horizon $T > 0$. In the framework of the standard Black-Scholes model (Black and Scholes [1973]), the price of the underlying $S = (S_t)_{t\in[0,T]}$ and the bond price $B = (B_t)_{t\in[0,T]}$ follow

$$\begin{cases} B_t = \mathrm{e}^{rt}, \\ S_t = S_0 \exp(\sigma W_t + \mu t), \end{cases} \tag{3.1}$$

where r is the riskless interest rate, $\sigma > 0$ is the constant volatility, μ is the constant drift, S_0 is the initial stock price and $W = (W_t)_{t\in[0,T]}$ is a Wiener process under $\mathbb{P}$. We assume that there are no transaction costs and both instruments are freely tradable.

The stochastic differential equation (SDE) for the discounted price process $X_t = B_t^{-1} S_t$ has the following form:

$$\begin{cases} dX_t = X_t(\sigma dW_t + m dt), \\ X_0 = x_0, \end{cases} \tag{3.2}$$

where $m = \mu - r + \frac{1}{2}\sigma^2$.

The unique equivalent martingale measure $\mathbb{P}^*$ can be then derived with the help of the Girsanov theorem:

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \exp\left(-\frac{m}{\sigma} W_T - \frac{1}{2}\left(\frac{m}{\sigma}\right)^2 T\right). \tag{3.3}$$

Note that

$$X_T = x_0 \exp\left(\sigma W_T + (m - \tfrac{1}{2}\sigma^2)T\right), \tag{3.4}$$

so (3.3) can be expressed in the form

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \text{const} \cdot X_T^{-m/\sigma^2}. \tag{3.5}$$

In this section, we focus on hedging a plain vanilla call option with the strike price of K, i.e. a contingent claim whose payoff equals $(S_T - K)^+$. The discounted contingent claim is also a call option with respect to X_t, with the strike price of $K\mathrm{e}^{-rT}$:

$$H = (X_T - K\mathrm{e}^{-rT})^+, \tag{3.6}$$

The amount of the initial wealth H_0 required for a perfect hedge is provided by the Black-Scholes formula:

$$H_0 = \mathbb{E}^*(H) = x_0\Phi_+(K\mathrm{e}^{-rT}) - K\mathrm{e}^{-rT}\Phi_-(K\mathrm{e}^{-rT}), \tag{3.7}$$

where

$$\Phi_\pm(K) = \Phi\left(\frac{\ln x_0 - \ln K}{\sigma\sqrt{T}} \pm \frac{1}{2}\sigma\sqrt{T}\right), \tag{3.8}$$

and $\Phi(\cdot)$ is the cumulative distribution function for the standard normal distribution.

In the case when the amount of the initial wealth is bounded from above by $\tilde{V}_0 \in (0, H_0)$, we cannot construct a perfect hedge for the call option. Instead, we shall minimize CVaR over all admissible strategies with the initial wealth not exceeding $\tilde{V}_0$. The results of Section 2.2 shall be used to derive the explicit solution.

As stated in Theorem 2.1, the original problem can be reduced to a problem of minimizing an auxiliary function $c(z)$ on interval $[0, z^*]$, where

$$c(z) = \begin{cases} z + \dfrac{1}{1-\alpha}\mathbb{E}\left((H-z)^+\tilde{\varphi}(z)\right), \text{for } z < z^*, \\ z, \qquad\qquad\qquad\qquad\qquad\quad \text{for } z \geqslant z^*, \end{cases} \tag{3.9}$$

$\tilde{\varphi}(z)$ is defined by (2.21) and z^* is a real root of equation

$$\tilde{V}_0 = \mathbb{E}^*\left((H-z^*)^+\right). \tag{3.10}$$

The optimal hedging strategy is a perfect hedge for the modified contingent claim $(H-\hat{z})^+\tilde{\varphi}(\hat{z})$, where $\hat{z}$ is the point of minimum of $c(z)$.

Since we only consider $z \geqslant 0$,

$$(H-z)^+ = ((X_T - K\mathrm{e}^{-rT})^+ - z)^+ = (X_T - (K\mathrm{e}^{-rT} + z))^+. \tag{3.11}$$

For simplicity of notation, denote

$$H(z) = (X_T - K(z))^+, \tag{3.12}$$

$$K(z) = K\mathrm{e}^{-rT} + z, \tag{3.13}$$

$$\tilde{\Phi}_\pm(x) = \Phi_\pm\left(x\mathrm{e}^{-mT}\right), \tag{3.14}$$

$$\Lambda_\pm(x,y) = \Phi_\pm(x) - \Phi_\pm(y), \tag{3.15}$$

$$\tilde{\Lambda}_\pm(x,y) = \tilde{\Phi}_\pm(x) - \tilde{\Phi}_\pm(y). \tag{3.16}$$

Evidently, $H(z)$ is also a call option with respect to X, with the strike price of $K(z)$, hence the optimal strategy is a perfect hedging strategy for a knockout call option with a modified strike price $K(z) \geqslant K$. By minimizing $c(z)$, we simultaneously derive both the knockout threshold and the modified strike price.

By applying the Black-Scholes formula to (3.10), we obtain:

$$\tilde{V}_0 = x_0 \Phi_+(K(z^*)) - K(z^*)\Phi_-(K(z^*)). \tag{3.17}$$

In what follows, we shall refer to z^* as to the solution for (3.17).

We shall consider two distinct cases.

(a) $\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\sigma}^2 > \mathbf{r}$ $\quad(\mathbf{m} > \mathbf{0})$

The set $\left\{\frac{d\mathbb{P}}{d\mathbb{P}^*} > a\right\}$ has the following structure:

$$\left\{\frac{d\mathbb{P}}{d\mathbb{P}^*} > a\right\} = \left\{X_T^{m/\sigma^2} > \hat{b}\right\} = \left\{X_T > b\right\}, \tag{3.18}$$

and, moreover,

$$\mathbb{P}\left(\frac{d\mathbb{P}}{d\mathbb{P}^*} = a\right) = \mathbb{P}^*\left(\frac{d\mathbb{P}}{d\mathbb{P}^*} = a\right) = 0. \tag{3.19}$$

Consider the case when $z \in [0, z^*)$, since otherwise we have $c(z) = z$. By applying (3.18) and (3.19) to (2.21), (2.22) and (2.23), we obtain

$$\tilde{\varphi}(z) = \mathbf{1}_{\left\{X_T > \tilde{b}(z)\right\}}, \tag{3.20}$$

$$\tilde{b}(z) = \inf\left\{b \geqslant 0 : \mathbb{E}^*\left(H(z) \cdot \mathbf{1}_{\{X_T > b)\}}\right) \leqslant \tilde{V}_0\right\}, \tag{3.21}$$

$$\gamma(z) = 0. \tag{3.22}$$

Note that in this particular case the infimum in (3.21) is always attained since we deal with atomless measures. The expectation in (3.21) can be rewritten as

$$\mathbb{E}^* \left(H(z) \cdot \mathbf{1}_{\{X_T > b\}} \right) = \begin{cases} \mathbb{E}^* \left(H(z) \right), & \text{for } b < K(z), \\ x_0 \Phi_+(b) - K(z) \Phi_-(b), & \text{for } b \geqslant K(z). \end{cases} \tag{3.23}$$

Furthermore, (2.33) implies that

$$\begin{cases} \mathbb{E}^* \left(H(z) \right) \geqslant \tilde{V}_0, & \text{for } z \leqslant z^*, \\ \mathbb{E}^* \left(H(z) \right) < \tilde{V}_0, & \text{for } z > z^*, \end{cases} \tag{3.24}$$

therefore the minimum is not attained on the set $\{b \in \mathbb{R} : b < K(z)\}$, hence $\tilde{b}(z)$ is a solution to the following system:

$$\begin{cases} x_0 \Phi_+(b) - K(z) \Phi_-(b) = \tilde{V}_0, \\ b \geqslant K(z). \end{cases} \tag{3.25}$$

The constraint $b \geqslant K(z)$ in (3.25) is crucial since the equation may have multiple real roots. It is straightforward to show that for all $z \in [0, z^*)$ there exists a unique solution $\tilde{b}(z)$ for (3.25).

We are now able to write down the function $c(z)$:

$$c(z) = \begin{cases} z + \dfrac{1}{1-\alpha} \cdot \mathbb{E} \left(\mathbf{1}_{\left\{X_T \leqslant \tilde{b}(z)\right\}} \cdot H(z) \right), & z < z^*, \\ z, & z \geqslant z^*, \end{cases} \tag{3.26}$$

or, by evaluating the expectation,

$$c(z) = \begin{cases} z + \dfrac{1}{1-\alpha} \cdot \left(x_0 \mathrm{e}^{mT} \tilde{\Lambda}_+(K(z), \tilde{b}(z)) - K(z) \tilde{\Lambda}_-(K(z), \tilde{b}(z)) \right), & z < z^*, \\ z, & z \geqslant z^*, \end{cases} \tag{3.27}$$

where $\tilde{b}(z)$ is a solution for (3.25) when $z < z^*$.

According to Theorem 2.1, the optimal strategy $(\hat{V}_0, \hat{\xi})$ is a perfect hedge for the modified contingent claim

$$\tilde{H}(\hat{z}) = (H - \hat{z})^+ \cdot \mathbf{1}_{\left\{X_T > \tilde{b}(\hat{z})\right\}}, \tag{3.28}$$

where $\hat{z}$ is a point of minimum of $c(z)$ on interval $[0, z^*]$.

(b) $\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{\sigma}^2 < \mathbf{r}$ $\quad$ ($\mathbf{m} < \mathbf{0}$)

In this case the set $\left\{\frac{d\mathbb{P}}{d\mathbb{P}^*} > a\right\}$ has the form

$$\left\{\frac{d\mathbb{P}}{d\mathbb{P}^*} > a\right\} = \left\{X_T^{m/\sigma^2} > \hat{b}\right\} = \left\{X_T < b\right\}, \tag{3.29}$$

therefore Theorem 2.1 provides the following optimal randomized test in the case when $z \in [0, z^*)$:

$$\tilde{\varphi}(z) = \mathbf{1}_{\left\{X_T < \tilde{b}(z)\right\}}, \tag{3.30}$$

$$\tilde{b}(z) = \sup\left\{b \geqslant 0 : \mathbb{E}^* \left(H(z) \cdot \mathbf{1}_{\{X_T < b)\}}\right) \leqslant \tilde{V}_0\right\}, \tag{3.31}$$

$$\gamma(z) = 0. \tag{3.32}$$

Denote

$$\beta(b, z) = x_0 \Lambda_+(K(z), b) - K(z) \Lambda_-(K(z), b), \tag{3.33}$$

then

$$\mathbb{E}^*\left(H(z)\cdot \mathbf{1}_{\{X_T<b\}}\right) = \begin{cases} 0, & \text{for } b < K(z), \\ \beta(b,z), & \text{for } b \geqslant K(z). \end{cases} \tag{3.34}$$

Same as above, recall that $\mathbb{E}^*\left(H(z)\right) \geqslant \tilde{V}_0$ for $z \leqslant z^*$ and consider the properties of function $\beta(b,z)$:

$$\beta(K(z),z) = 0, \qquad \beta(+\infty,z) = \mathbb{E}^*\left(H(z)\right), \qquad \frac{\partial}{\partial b}\beta(b,z) \geqslant 0. \tag{3.35}$$

It follows from (3.35) that the supremum in (3.31) is attained on the set $\{b \in \mathbb{R} : b \geqslant K(z)\}$, hence $\tilde{b}(z)$ is a unique solution for the following system:

$$\begin{cases} x_0\Lambda_+(K(z),b) - K(z)\Lambda_-(K(z),b) = \tilde{V}_0, \\ b \geqslant K(z). \end{cases} \tag{3.36}$$

Function $c(z)$ then has the form

$$c(z) = \begin{cases} z + \dfrac{1}{1-\alpha}\cdot \mathbb{E}\left(\mathbf{1}_{\{X_T \geqslant \tilde{b}(z)\}}\cdot H(z)\right), & z < z^*, \\ z, & z \geqslant z^*, \end{cases} \tag{3.37}$$

or

$$c(z) = \begin{cases} z + \dfrac{1}{1-\alpha}\cdot \left(x_0 \mathrm{e}^{mT}\tilde{\Phi}_+(\tilde{b}(z)) - K(z)\tilde{\Phi}_-(\tilde{b}(z))\right), & z < z^*, \\ z, & z \geqslant z^*, \end{cases} \tag{3.38}$$

where $\tilde{b}(z)$ is a solution for (3.36) when $z \in [0, z^*)$, and the optimal strategy

$(\hat{V}_0, \hat{\xi})$ is a perfect hedge for the modified contingent claim

$$\tilde{H}(\hat{z}) = (H - \hat{z})^+ \cdot \mathbf{1}_{\left\{X_T < \tilde{b}(\hat{z})\right\}}, \tag{3.39}$$

where $\hat{z}$ is a point of minimum of $c(z)$ on interval $[0, z^*]$.

To illustrate the above method numerically, consider a financial market evolving in accordance with the Black-Scholes model with parameters $\sigma = 0.3$, $\mu = 0.09$, $r = 0.05$ and a European call option with the strike price of $K = 110$ and the time to maturity $T = 0.25$. Let the initial price of the underlying be equal to $S_0 = 100$. In this setting, we are interested in hedging strategies that minimize conditional value-at-risk at various confidence levels and various amounts of the initial wealth. The results of numeric computations can be observed in Figure 3.1.

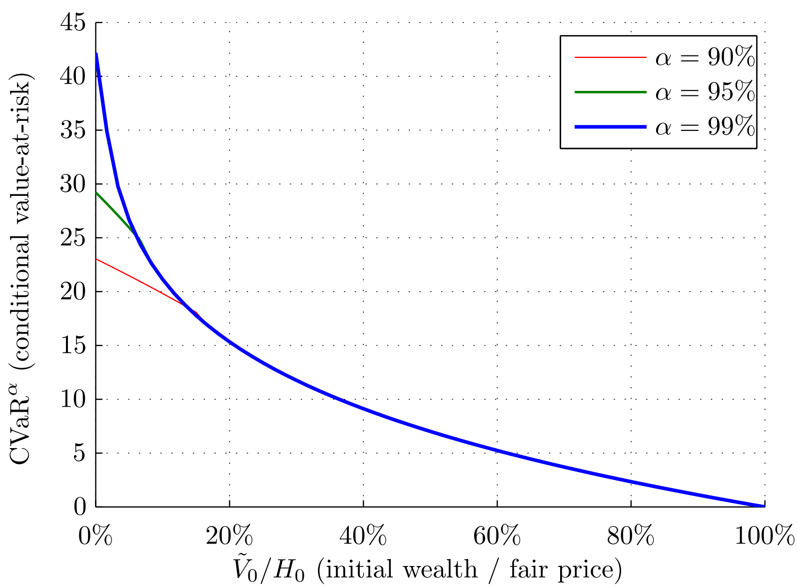

Figure 3.1: CVaR of the optimal hedging strategy at confidence levels of 90%, 95% and 99% for varying levels of initial wealth in the Black-Scholes model.

It can be seen in Figure 3.1 that the optimal strategy becomes insensitive to the confidence level as we increase the initial wealth. This is due to the fact that that the minimum of $c(z)$ for high values of $\tilde{V}_0$ is attained at z^*, thus $c(\hat{z}) = c(z^*) = z^*$, which does not depend on α. The optimal hedging strategy is a perfect hedge for a modified claim obtained by shifting the initial claim down by $z^* e^{rT}$ and then taking the positive part (e.g., if the original claim is a call option, the optimal strategy is a perfect hedge for a call option with a higher strike price). In this case, the tail risk is concentrated entirely at the α-quantile of the loss distribution, and value-at-risk coincides with conditional value-at-risk. Alternatively, when $\tilde{V}_0$ is low, the optimal strategy is a perfect hedge for a knockout option based on the modified claim described above, where the knockout threshold depends on the amount of the initial wealth. Both types of strategies are presented in Figures 3.2 and 3.3 below.

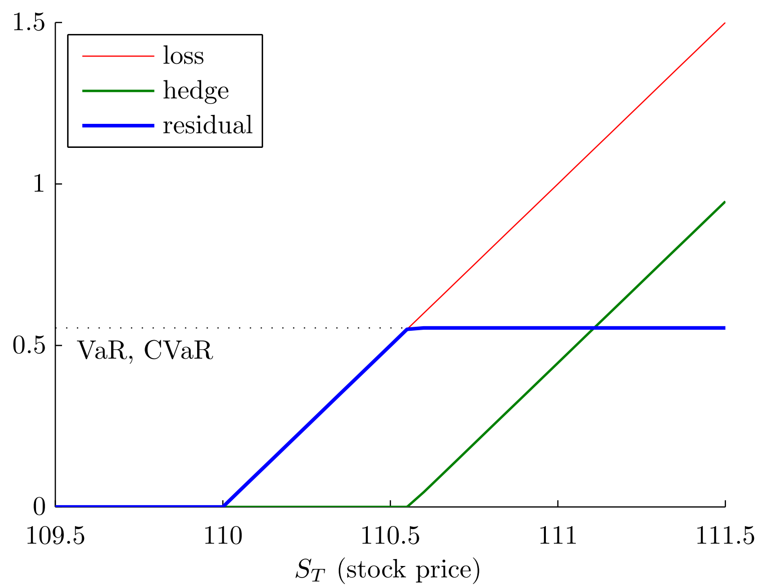

Figure 3.2: The structure of the CVaR-optimal hedging strategy at the 95% confidence level with the initial wealth $\tilde{V}_0 = 0.06 \cdot H_0$ in the Black-Scholes model.

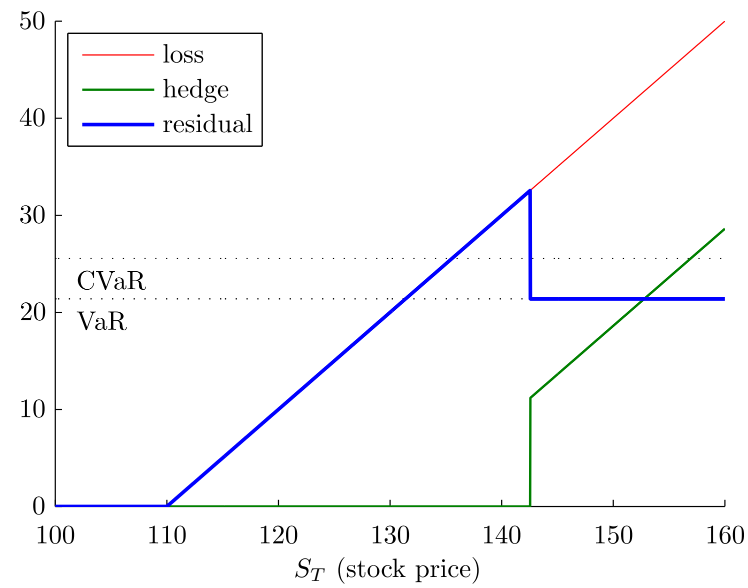

Figure 3.3: The structure of the CVaR-optimal hedging strategy at the 95% confidence level with the initial wealth $\tilde{V}_0 = 0.95 \cdot H_0$ in the Black-Scholes model.

We shall now consider a situation when a CVaR-based capital requirement is imposed by the regulator. This constitutes a relevant and important topic, as in the latest consultative document which was presented recently by [Basel Committee on Banking Supervision, 2012] (the so-called "Basel 3.5" outline) for discussion to the banking industry, "the Committee is proposing the use of ES for the internal models-based approach" as a replacement for the well-settled value-at-risk methodology. Assuming that the regulator requires to allocate $\beta > 0$ units of capital for each unit of expected shortfall, we shall take position of the option seller who is faced a choice of a partial hedging strategy and is concerned with the short-term capital allocation. Thus, the option seller compares $\rho_\beta(0) = \beta \cdot \mathrm{CVaR}^\alpha(0)$ against $\rho_\beta(\tilde{V}_0) = \beta \cdot \mathrm{CVaR}^\alpha(\tilde{V}_0) + \tilde{V}_0$, where $\mathrm{CVaR}^\alpha(v)$ denotes the CVaR at the confidence level α of the CVaR-minimizing hedging strategy which uses no more than v of the initial wealth.

(Note that although this particular way of measuring hedging performance is consistent with the short-term capital allocation objective, many other criteria may considered based on the option seller's preferences.) On Figure 3.4, the ratio $\rho_\beta(\tilde{V}_0)/\rho_\beta(0)$ can be observed for varying levels of β and $\tilde{V}_0$. The regions of the graphs where the aforesaid ratio is less than 1 correspond to the case where the option seller can reduce the required regulatory capital by a value exceeding the initial cost of hedging, while at the same time reducing the total risk exposure CVaR-wise.

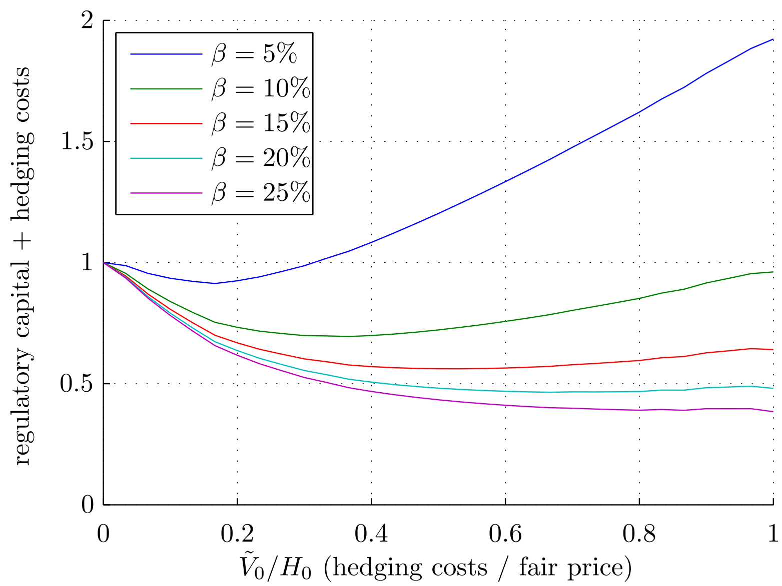

Figure 3.4: Performance of the CVaR-optimal hedging strategy at the 99% confidence level with varying levels of the initial wealth and the capital requirement ratio in the Black-Scholes model.

In the second half of this section, we shall use the results of Section 2.3 to explicitly construct hedging strategies minimizing the initial wealth with CVaR not exceeding a given target value $\tilde{C} > 0$.

According to Theorem 2.2, a passive trading strategy is optimal in the hedging costs minimization problem if at least one of the inequalities in (2.56) is violated. In the Black-Scholes setting, these inequalities take the form

$$\begin{cases} x_0 \mathrm{e}^{mT}\tilde{\Phi}_+(K) - K\tilde{\Phi}_-(K) - \tilde{C}(1-\alpha) > 0, \\ x_0 \mathrm{e}^{mT}\tilde{\Phi}_+(K+\tilde{C}) - (K+\tilde{C})\tilde{\Phi}_-(K+\tilde{C}) > 0. \end{cases} \tag{3.40}$$

In what follows, we assume a non-trivial case, i.e. both inequalities in (3.40) are assumed to be satisfied. Again, we consider two distinct cases.

(a) $\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{\sigma}^2>\mathbf{r} \quad (\mathbf{m}>\mathbf{0})$

In this case

$$\left\{\frac{d\mathbb{P}^*}{d\mathbb{P}} > a\right\} = \{X_T < b\}, \qquad \mathbb{P}\left(\frac{d\mathbb{P}}{d\mathbb{P}^*} = a\right) = \mathbb{P}^*\left(\frac{d\mathbb{P}}{d\mathbb{P}^*} = a\right) = 0, \tag{3.41}$$

so we have

$$\tilde{\varphi}(z) = \mathbf{1}_{\left\{X_T<\tilde{b}(z)\right\}}, \tag{3.42}$$

$$\tilde{b}(z) = \sup\left\{b \geqslant 0 : \mathbb{E}\left(H(z)\cdot \mathbf{1}_{\{X_T<b)\}}\right) \leqslant (\tilde{C}-z)(1-\alpha)\right\}, \tag{3.43}$$

$$\gamma(z) = 0. \tag{3.44}$$

Denote

$$\delta(b,z) = x_0\mathrm{e}^{mT}\tilde{\Lambda}_+(K(z),b) - K(z)\tilde{\Lambda}_-(K(z),b), \tag{3.45}$$

then

$$\mathbb{E}\left(H(z)\cdot \mathbf{1}_{\{X_T<b\}}\right) = \begin{cases} 0, & \text{for } b < K(z), \\ \delta(b,z), & \text{for } b \geqslant K(z), \end{cases} \tag{3.46}$$

$$\delta(K(z),z) = 0, \qquad \delta(+\infty,z) = \mathbb{E}\left(H(z)\right), \qquad \frac{\partial}{\partial b}\delta(b,z) \geqslant 0. \tag{3.47}$$

Assuming that inequalities (3.40) hold true and $z < \tilde{C}$, the supremum in (3.43) is attained on the set $\{b \in \mathbb{R} : b \geqslant K(z)\}$, hence $\tilde{b}(z)$ is a unique solution for

$$\begin{cases} x_0 e^{mT} \tilde{\Lambda}_+(K(z), b) - K(z)\tilde{\Lambda}_-(K(z), b) = (\tilde{C} - z)(1 - \alpha), \\ b \geqslant K(z), \end{cases} \quad \text{if } z < \tilde{C}, \tag{3.48}$$

and

$$\tilde{b}(z) = \sup\left\{b \geqslant 0 : x_0 e^{mT} \tilde{\Lambda}_+(K(z), b) - K(z)\tilde{\Lambda}_-(K(z), b) = 0\right\}, \quad \text{if } z = \tilde{C}. \tag{3.49}$$

The optimal strategy $(\hat{V}_0, \hat{\xi})$ in the hedging costs minimization problem is then a perfect hedge for the modified contingent claim $H(\hat{z})^+ \cdot \mathbf{1}_{\left\{X_T > \tilde{b}(\hat{z})\right\}}$, where $\tilde{b}(z)$ is defined by (3.48) and (3.49), and $\hat{z}$ is a point of minimum of function

$$d(z) = \begin{cases} x_0 \Phi_+(\tilde{b}(z)) - K(z)\Phi_-(\tilde{b}(z)), & \text{if } z < \tilde{C}, \\ x_0 \Phi_+(K(z)) - K(z)\Phi_-(K(z)), & \text{if } z = \tilde{C}, \end{cases} \tag{3.50}$$

on interval $z \in (-\infty, \tilde{C}]$.

(b) $\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\sigma}^2 < \mathbf{r}$ $\quad (\mathbf{m} < \mathbf{0})$

We have

$$\left\{\frac{d\mathbb{P}^*}{d\mathbb{P}} > a\right\} = \{X_T > b\}, \qquad \mathbb{P}\left(\frac{d\mathbb{P}}{d\mathbb{P}^*} = a\right) = \mathbb{P}^*\left(\frac{d\mathbb{P}}{d\mathbb{P}^*} = a\right) = 0, \tag{3.51}$$

hence

$$\tilde{\varphi}(z) = \mathbf{1}_{\left\{X_T > \tilde{b}(z)\right\}}, \tag{3.52}$$

$$\tilde{b}(z) = \inf\left\{b \geqslant 0 : \mathbb{E}\left(H(z)\cdot \mathbf{1}_{\{X_T > b)\}}\right) \leqslant (\tilde{C} - z)(1-\alpha)\right\}, \tag{3.53}$$

$$\gamma(z) = 0. \tag{3.54}$$

Denote

$$\zeta(b,z) = x_0\tilde{\Phi}_+(b) - K(z)\tilde{\Phi}_-(b), \tag{3.55}$$

then

$$\mathbb{E}\left(H(z)\cdot \mathbf{1}_{\{X_T > b\}}\right) = \begin{cases} \mathbb{E}\left(H(z)\right), \text{for } b < K(z), \\ \\ \zeta(b,z), \qquad \text{for } b \geqslant K(z), \end{cases} \tag{3.56}$$

and

$$\zeta(K(z),z) = \mathbb{E}\left(H(z)\right), \qquad \zeta(+\infty, z) = 0, \qquad \frac{\partial}{\partial b}\zeta(b,z) \leqslant 0. \tag{3.57}$$

When $z < \tilde{C}$, the supremum in (3.53) is attained on the set $\{b \in \mathbb{R} : b \geqslant K(z)\}$, thus $\tilde{b}(z)$ is a unique solution for the system

$$\begin{cases} x_0\tilde{\Phi}_+(b) - K(z)\tilde{\Phi}_-(b) = (\tilde{C} - z)(1-\alpha), \\ \\ b \geqslant K(z). \end{cases} \qquad \text{if } z < \tilde{C}, \tag{3.58}$$

and

$$\tilde{b}(z) = \inf\left\{b \geqslant 0 : x_0\tilde{\Phi}_+(b) - K(z)\tilde{\Phi}_-(b) = 0\right\}, \quad \text{if } z = \tilde{C}. \tag{3.59}$$

The optimal strategy $(\hat{V}_0, \hat{\xi})$ in the hedging costs minimization problem in this case is a perfect hedge for the modified contingent claim $H(\hat{z})^+ \cdot \mathbf{1}_{\{X_T < \tilde{b}(\hat{z})\}}$, where $\tilde{b}(z)$ is a solution for (3.58) and $\hat{z}$ is a point of minimum of function

$$d(z) = \begin{cases} x_0\Lambda_+(K(z), \tilde{b}(z)) - K(z)\Lambda_-(K(z), \tilde{b}(z)), & \text{if } z < \tilde{C}, \\ x_0\Phi_+(K(z)) - K(z)\Phi_-(K(z)), & \text{if } z = \tilde{C}, \end{cases} \tag{3.60}$$

on interval $z \in (-\infty, \tilde{C}]$.

We shall illustrate the above results numerically in the Black-Scholes model with the same parameters as earlier ($\sigma = 0.3$, $\mu = 0.09$, $r = 0.05$, European call option with the strike price of $K = 110$, time to maturity $T = 0.25$, initial price $S_0 = 100$). Figure 3.5 shows the minimum amount of the initial wealth that has to be invested in the hedging strategy so that conditional value-at-risk of the resulting portfolio does not exceed a specified threshold, for various confidence levels of CVaR.

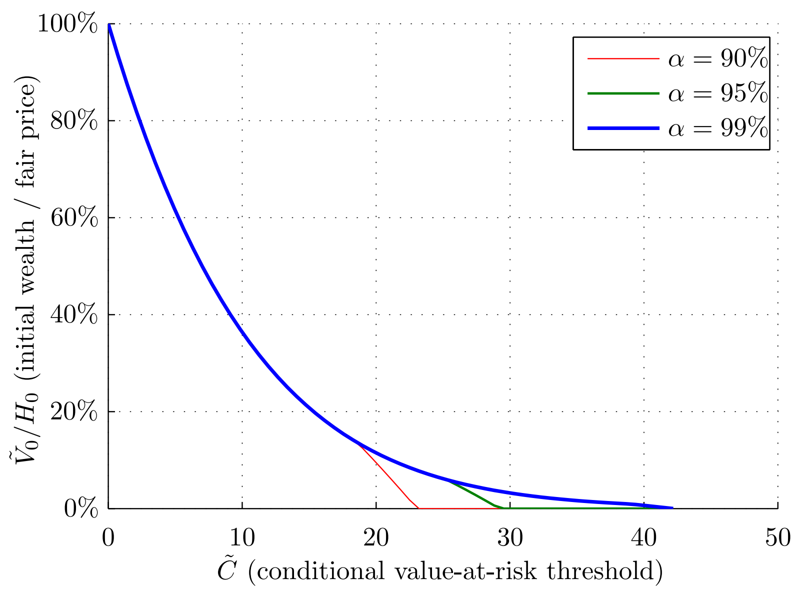

Figure 3.5: Initial wealth of the optimal hedging strategy for varying levels of CVaR threshold at confidence levels of 90%, 95% and 99% in the Black-Scholes model.

3.2 Hedging a call option in the telegraph market model

In this section, we shall demonstrate how a CVaR-optimal hedging strategy for a plain vanilla call option can be constructed in the framework of the telegraph market model suggested in [Melnikov and Ratanov, 2008] and [Melnikov and Ratanov, 2007][1].

The telegraph market model can be informally described as a complete market model in which the dynamics of the risky asset features jumps and regime switching. The model can be viewed as a generalization of the pure-jump version of the Merton's model that is arbitrage-free and preserves market completeness under some restrictions on model parameters. In addition, it can be shown that the telegraph market model converges to the Black-Scholes model in distribution under proper scaling (Melnikov and Ratanov [2008]), hence it can be used to approximate the lognormal stock price behavior.

For the sake of illustration, let us briefly compare the properties of the telegraph market model to those of the well-known Cox-Ross-Rubinstein binomial model. Both models feature jump dynamics, market completeness and convergence of the stock price process to a lognormal process in distribution. However, the key difference between the two is that in the binomial model the jumps have fixed timing and random size, while in the telegraph market model the size of the jumps is pre-determined but they occur at random.

Consider a simplified version of the Merton's model (Merton [1976]) with

[1] The author is grateful to Alexey Kuznetsov (*kuznetsov@mathstat.yorku.ca*) for his contribution to the results presented in this section.

no diffusion term and a constant jump impact:

$$\begin{cases} dB_t = rB_t dt, \\ dS_t = S_t(mdt + \sigma dN_t), \end{cases} \tag{3.61}$$

where S_t and B_t are the prices of the stock and the bond respectively, N_t is a Poisson process, $r \geqslant 0$ is the risk-free interest rate, $\sigma > 0$ is the constant jump impact and m is the constant drift. It is easy to verify that this model represents a complete arbitrage-free market; the risk comes only from the uncertainty in timing of jumps in the Poisson process, with the size of jumps known in advance.

How can this model be generalized without violating completeness? If we simply allow the jump size to be random, a new source of randomness will be introduced which will render the model incomplete. Instead, we can let the jump size change in a predictable fashion. Let the time between jumps be determined by a family $\{\tau_i\}_{i\in\mathbb{N}}$ of independent exponential random variables with intensities $\{\lambda_i\}_{i\in\mathbb{N}}$, and let $\{h_i\}_{i\in\mathbb{N}}$ be an arbitrary deterministic sequence specifying the size of jumps. Define the moment of the n-th jump as

$$T_n = \sum_{i=1}^{n} \tau_i, \tag{3.62}$$

let the log-price process $Y_t = \ln(S_t)$ follow

$$Y_t = mt + \sum_{i\in\mathbb{N}, T_i \leqslant t} h_i, \tag{3.63}$$

and consider the market

$$\begin{cases} dB_t = rB_t dt, \\ dS_t = S_t dY_t. \end{cases} \tag{3.64}$$

Similar to (3.61), the market model defined by (3.64) also contains a single source of randomness: the timing of the next jump is random while the size of the jump is known in advance. However, unless certain restrictions are imposed, this market model admits arbitrage. For the sake of illustration, we shall provide a simple example. Let

$$(h_1, h_2) = (-h, h), \quad (\lambda_1, \lambda_2) = (1, 1), \quad r = 0, \quad m = \text{const} > 0, \quad h > 0.$$

In this setting, consider the following strategy on $t \in [0, 1]$:

- if a negative jump occurs at time $t = t_0 < 1$, borrow S_{t_0} in cash and buy one unit of stock. Sell the stock at time $t = t_1 \wedge 1$ and repay the debt of S_{t_0}, where $t = t_1$ is the time of the next (positive) jump;
- if no negative jumps occur on $[0, 1]$, don't do anything.

Evidently, the strategy presented above is an arbitrage strategy: if there are no negative jumps on $[0, 1]$, the profit is zero, otherwise there is a positive gain since $S_{t_1 \wedge 1} > S_{t_0}$. In the case when the drift is negative or both jumps are negative it is straightforward to construct similar examples.

To ensure the no-arbitrage property in model (3.64) with alternating jump sizes, the drift should switch sign and/or magnitude at the time of the jump:

$$d(t) = \int_0^t \sum_{i \in \mathbb{N}} d_i \cdot \mathbf{1}_{[T_{i-1}, T_i)}(s) ds, \tag{3.65}$$

where d_i should be constrained in such way that the model does not allow arbitrage opportunities. Processes of this type are known as *telegraph processes*.

To summarize, the telegraph market model is a generalization of the Poisson market model obtained by introducing regime switching, where only the timing of regime switching is uncertain. In this section, we shall consider the simplest case of a telegraph process with two alternating states. This particular case is important since it inherits all of the characteristic features of a general telegraph model, yet at the same time it allows for quite simple semi-explicit solutions for the problem of option pricing and hedging.

Definition 3.1: **The two-state telegraph market model**.

1. Let $\sigma(t) \in \{1, 2\}$ be a continuous time Markov chain process with Markov generator
$$L_\sigma = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \\ \lambda_2 & -\lambda_2 \end{pmatrix}. \tag{3.66}$$
Process $\sigma(t)$ represents the current state in the telegraph market model. Without loss of generality, we assume that $\sigma(0) = 1$.

2. Define the telegraph process X_t and the jump process J_t:
$$X_t = \int_0^t c_{\sigma(s)} ds, \qquad J_t = \sum_0^{N_t} h_{\sigma(T_j-)}, \tag{3.67}$$
where $\mathbf{c} = (c_1, c_2)$ determines drift states, $\mathbf{h} = (h_1, h_2)$ determines jump size states, and N_t denotes the number of jumps of $\sigma(t)$ up to time t.

3. The risk-free asset is defined by

$$dB_t = r_{\sigma(t)} B_t dt \tag{3.68}$$

or, equivalently,

$$B_t = \exp\left(\int_0^t r_{\sigma(s)} ds \right), \tag{3.69}$$

where $\mathbf{r} = (r_1, r_2)$ determines the states of the (non-negative) risk-free rate.

4. The risky asset is defined as a telegraph market price process

$$dS_t = S_{t-} d(X_t + J_t). \tag{3.70}$$

Following [Melnikov and Ratanov, 2008], we can express S_t as

$$S_t = \mathcal{E}_t(X + J) = S_0 \mathrm{e}^{X_t} \kappa(N_t), \tag{3.71}$$

where $\mathcal{E}_t(\cdot)$ denotes the stochastic exponential and

$$\kappa(N_t) = \prod_{s \leqslant t} (1 + \Delta J_s), \tag{3.72}$$

which implies

$$\kappa(2k - 1) = (1 + h_1)^k (1 + h_2)^{k-1}, \tag{3.73}$$

$$\kappa(2k) = (1 + h_1)^k (1 + h_2)^k, \tag{3.74}$$

for all $k \in \mathbb{N}$.

One of the most important questions regarding the telegraph market model is the absence of arbitrage and market completeness. This topic has been investigated in detail in [Melnikov and Ratanov, 2008], and we shall reproduce the main results here in the form of a theorem.

Theorem 3.1 (Melnikov and Ratanov [2008]): *The two-state telegraph market model is arbitrage-free if and only if*

$$\frac{r_\sigma - c_\sigma}{h_\sigma} > 0, \qquad \sigma \in \{1, 2\}. \tag{3.75}$$

If the telegraph market model is arbitrage-free, then it is complete, and the unique equivalent martingale measure $\mathbb{P}^*$ is defined by

$$\left.\frac{d\mathbb{P}^*}{d\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t = \mathcal{E}_t(X^* + J^*) = \mathrm{e}^{X_t^*}\kappa^*(N_t), \tag{3.76}$$

where

$$\kappa^*(N_t) = \prod_{s \leqslant t}(1 + \Delta J_s^*),$$

X_t^* is a telegraph process with intensities λ_σ defined in (3.66) and drift

$$c_\sigma^* = \lambda_\sigma + \frac{c_\sigma - r_\sigma}{h_\sigma}, \qquad \sigma \in \{1, 2\}, \tag{3.77}$$

and J_t^* is a jump process with jumps

$$h_\sigma^* = \frac{r_\sigma - c_\sigma}{h_\sigma \lambda_\sigma} - 1, \qquad \sigma \in \{1, 2\}. \tag{3.78}$$

The discounted price process $B_t^{-1} S_t$ is a telegraph market process under $\mathbb{P}^*$

with drift c_σ, jumps h_σ and modified intensities

$$\lambda_\sigma^* = \frac{r_\sigma - c_\sigma}{h_\sigma} > 0. \tag{3.79}$$

We shall now turn to the problem of CVaR hedging in a complete telegraph market model. Consider a plain vanilla call option with maturity T and the strike price of K. The discounted payoff H at time $t = T$ has the form

$$H = B_T^{-1}(S_T - K)^+, \tag{3.80}$$

where B_t and S_t are defined by (3.68) and (3.70). Denote the unique fair price

$$H_0 = \mathbb{E}^*(H),$$

where $\mathbb{E}^*$ is the expectation with respect to the unique equivalent martingale measure $\mathbb{P}^*$ defined by (3.76), and assume that the amount of available initial wealth is bounded from above by $\tilde{V}_0 \in (0, H_0)$. Same as before, our objective is to find an admissible strategy that minimizes conditional value-at-risk while requiring no more than $\tilde{V}_0$ of the initial wealth.

According to Theorem 2.1, the problem of CVaR hedging can be reduced to a problem of minimizing a special function $c(z)$ over interval $[0, z^*]$, where

$$c(z) = \begin{cases} z + \dfrac{1}{1-\alpha}\mathbb{E}\left((H - z)^+ \tilde{\varphi}(z)\right), \text{ for } z \leqslant z^*, \\ z, \qquad\qquad\qquad\qquad\qquad\quad \text{for } z > z^*, \end{cases} \tag{3.81}$$

$\tilde{\varphi}(z)$ is defined by (2.21) and z^* is a real root of equation

$$\tilde{V}_0 = \mathbb{E}^*\left((H - z^*)^+\right). \tag{3.82}$$

Since $z \geqslant 0$, we have

$$(H - z)^+ = (B_T^{-1}(S_T - K)^+ - z)^+ = (B_T^{-1}S_T - (B_T^{-1}K + z))^+ \tag{3.83}$$

Explicit derivation of $\tilde{\varphi}(z)$ and $c(z)$ involves computing expectations

$$E_t(f, a) = \mathbb{E}\left(f(S_t, B_t) \cdot \mathbf{1}_{\{Z_t < a\}}\right) \tag{3.84}$$

and

$$E_t^*(f, a) = \mathbb{E}^*\left(f(S_t, B_t) \cdot \mathbf{1}_{\{Z_t < a\}}\right) \tag{3.85}$$

for arbitrary functions $f : [0, \infty) \times \mathbb{N}_0 \mapsto \mathbb{R}$ and $a \in \mathbb{R}$, where $\mathbb{N}_0$ is the set of natural numbers including zero. Towards this end, we shall first express processes B_t, S_t and Z_t in terms of X_t and N_t.

By the definition of B_t,

$$B_t = \exp\left(\int_0^t r_{\sigma(s)} ds\right) = \exp\left(\frac{r_2 - r_1}{c_2 - c_1} X_t + \frac{c_2 r_1 - c_1 r_2}{c_2 - c_1} t\right). \tag{3.86}$$

Process S_t can be related to X_t and N_t via (3.71):

$$S_t = S_0 \mathrm{e}^{X_t} \kappa(N_t). \tag{3.87}$$

Finally, according to Theorem 3.1,

$$Z_t = \mathrm{e}^{X_t^*}\kappa^*(N_t). \tag{3.88}$$

In order to express Z_t in terms of X_t and N_t, we shall use Lemma 2.2 from [Melnikov and Ratanov, 2008] which states that two telegraph processes based on the same process $\sigma(t)$ are linearly related. More specifically,

$$X_t^* = \frac{c_2^* - c_1^*}{c_2 - c_1} X_t + \frac{c_2 c_1^* - c_1 c_2^*}{c_2 - c_1} t, \tag{3.89}$$

which in conjunction with (3.88) yields

$$Z_t = \exp\left(\frac{c_2^* - c_1^*}{c_2 - c_1} X_t + \frac{c_2 c_1^* - c_1 c_2^*}{c_2 - c_1} t\right)\kappa^*(N_t). \tag{3.90}$$

Note also that

$$E_t^*(f, a) = E_t(f \cdot Z_t, a), \tag{3.91}$$

hence it is sufficient to only consider the expectation $E_t(f, a)$ under $\mathbb{P}$.

We can now substitute (3.86), (3.87) and (3.90) into (3.84):

$$E_t(f, a) = \mathbb{E}\left(f(S_t, B_t) \cdot \mathbf{1}_{\{Z_t < a\}}\right) = \mathbb{E}\left(g(X_t, N_t)\right). \tag{3.92}$$

By conditioning on $\{N_t = n\}$, we obtain

$$\mathbb{E}\left(g(X_t, N_t)\right) = \sum_{\mathbb{N}_0} \mathbb{E}\left(g(X_t, n) \cdot \mathbf{1}_{\{N_t = n\}}\right) = \sum_{\mathbb{N}_0} \int_{\mathbb{R}} g(x, n) p_n(t, x) dx, \tag{3.93}$$

where $p_n(t, x)$ is the corresponding conditional density

$$p_n(t, x) = \frac{d}{dx}\mathbb{P}\left(\{X_t \leqslant x\} \cap \{N_t = n\}\right). \tag{3.94}$$

Therefore, we can reduce evaluating expectations of the form (3.84) and (3.85) to a summation of one-dimensional integrals with respect to densities $p_n(t,x)$. Theorem 5.5 in Section 5.2 provides the recursive relationship for these conditional densities so that the expectation can be computed explicitly.

To illustrate how the conditional value-at-risk can be minimized numerically in this setting, we shall derive CVaR-optimal hedging strategy numerically for a European call option with the strike price of $K = 100$ and time to maturity $T = 0.25$ in the telegraph market model with parameters $\mathbf{c} = (-0.5, 0.5)$, $\boldsymbol{\lambda} = (5, 5)$, $\mathbf{r} = (0.07, 0.07)$, $\mathbf{h} = (0.5, -0.35)$, $S_0 = 100$.

The process of finding the CVaR-optimal strategy involves computing expected value of various functions a large number of times. One possible way of evaluating the integral in (3.93) efficiently is to consider a fixed grid

$$x_i = x_{\min} + \frac{i}{2N^{(x)}} \cdot (x_{\max} - x_{\min}), \quad i = 0, 1, \dots 2N^{(x)}, \tag{3.95}$$

$$n_j = j, \quad j = 0, 1, \dots N^{(n)}, \tag{3.96}$$

where $x_{\min} = c_1 T \wedge c_2 T$, $x_{\max} = c_1 T \vee c_2 T$, and approximate the expectation by partitioning the interval $(x_{\min}, x_{\max})$ into $N^{(x)}$ parts:

$$\sum_{\mathbb{N}_0} \int_{\mathbb{R}} g(x,n) p_n(t,x) dx \approx \sum_{i=0}^{2N^{(x)}} \sum_{j=0}^{N^{(n)}} g(x_i, n_j) p_{n_j}(T, x_i) \zeta_i, \tag{3.97}$$

where values $p_{n_j}(T, x_i)$ are computed in advance, and ζ_i are the Simpson's method weights: $\zeta_0 = \zeta_{N^{(x)}} = \frac{1}{3}$, $\zeta_{2k} = \frac{2}{3}$, $\zeta_{2k+1} = \frac{4}{3}$.

Figure 3.6 shows the minimal CVaR that can be attained by using the CVaR-optimal hedging strategy in the telegraph market model for various values of the initial wealth and confidence level. In Figure 3.7, we present

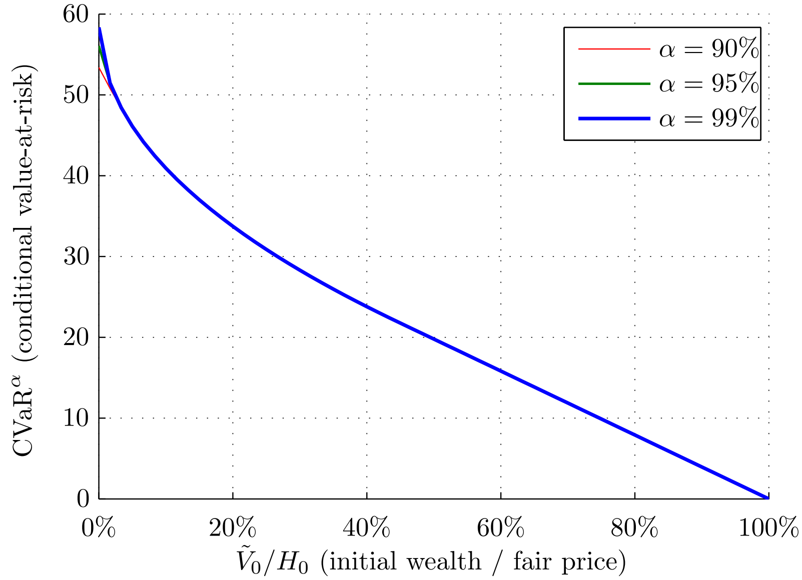

Figure 3.6: CVaR of the optimal hedging strategy at confidence levels of 90%, 95% and 99% for varying levels of initial wealth in the telegraph market model.

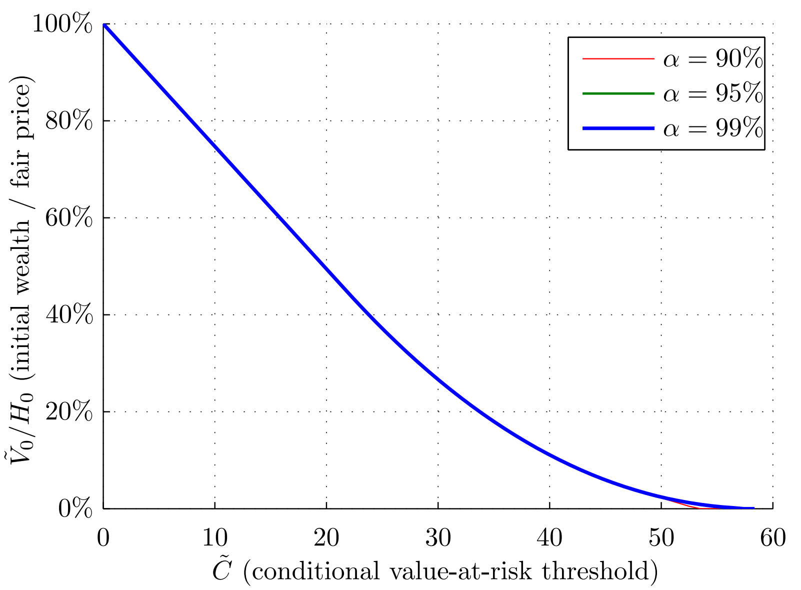

Figure 3.7: Initial wealth of the optimal hedging strategy for varying levels of CVaR threshold at confidence levels of 90%, 95% and 99% in the telegraph market model.

the numerical solution for the problem of hedging costs minimization in the telegraph market model.

3.3 Hedging a unit-linked life insurance contract

In this section, we shall use the quantile methodology which was proposed in [Melnikov and Skornyakova, 2005] and [Melnikov and Romaniuk, 2006] along with the techniques presented in Sections 2.2 and 2.3 to construct CVaR-optimal hedges for an embedded call option in an equity-linked life insurance contract.

In addition to the "financial" probability space $(\Omega, \mathcal{F}, \mathbb{P})$ considered earlier, let us consider the "actuarial" probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. Let a random variable $T(x)$ denote the remaining lifetime of a person aged x, and let ${}_Tp_x = \tilde{\mathbb{P}}[T(x) > T]$ be a survival probability for the next T years of the insured. We assume that $T(x)$ does not depend on the evolution of financial market, so we can treat $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ as independent.

Under an equity-linked pure endowment contract, the insurance company is obliged to pay the benefit in the amount of $\bar{H}$ (an $\mathcal{F}_T$-measurable random variable) to the insured provided the insured is alive at time T. Since the benefit is linked to the evolution of financial market, an insurance contract of this kind poses two independent kinds of risk to the insurance company: the mortality risk and the market risk.

The optimal price of the contract is traditionally calculated as an expected present value of cash flows under the risk-neutral probability. However, the insurance part of the contract doesn't need to be risk-adjusted since the mortality risk is essentially unsystematic. Denote the discounted benefit by $H = \bar{H}e^{-rT}$, then the price of the contract (known as the "Brennan-Schwartz price", see Brennan and Schwartz [1976]) equals

$$ {}_TU_x = \tilde{\mathbb{E}}\left(\mathbb{E}^*\left(H \cdot \mathbf{1}_{\{T(x)>T\}}\right)\right) = {}_Tp_x \cdot \mathbb{E}^*(H), \tag{3.98} $$

where $\tilde{\mathbb{E}}$ denotes the expectation with respect to $\tilde{\mathbb{P}}$.

The objective of the insurance company is to mitigate the financial component of risk and hedge $\bar{H}$ in the financial market. However,

$$ {}_TU_x < \mathbb{E}^*(H), \tag{3.99} $$

in other words, the insurance company is not able to hedge the benefit perfectly; instead, a partial hedging strategy may be used.

For a fixed client age x, denote the maximum amount of capital that is allocated for partial hedging of $\bar{H}$ by $\tilde{V}_0 = {}_Tp_x \cdot \mathbb{E}^*(H)$. We can now use the results of Theorem 2.1 to derive CVaR-optimal hedging strategy. Along with providing a way of hedging, this may be viewed as a possible way of estimating financial exposure of contracts for given values of age. Note that by applying Theorem 2.2 we can also address the inverse problem: given the financial claim and a fixed CVaR threshold, we can find the target survival probability (and hence the largest acceptable age) for the contract.

In the following example we investigate a pure endowment contract with a fixed guarantee which pays $\bar{H}$ at maturity given the insured is alive:

$$ \bar{H} = \max\{S_T, kS_0\}, \tag{3.100} $$

where S_t is the stock price process and k is a fixed percentage value. Since

$$\max\{S_T, kS_0\} = kS_0 + (S_T - kS_0)^+, \tag{3.101}$$

it is sufficient for our purposes to only consider the embedded call $(S_T - K)^+$, where $K = kS_0$.

Let the financial part of the model follow the Black-Scholes with parameters $\sigma = 0.3$, $\mu = 0.09$, $r = 0.05$ and let the embedded call option have the strike price of $K = 110$; the length of the contract T will vary in this example. Let the initial price of the stock be equal to $S_0 = 100$. In this example, we shall use the survival probabilities listed in mortality table *UP94 @2015* in [McGill et al., 2004] (*Uninsured Pensioner Mortality 1994 Table Projected to the Year 2015*). The objective is to construct partial hedging strategies that minimize CVaR for varying values of client age and time horizon. Please note that since we are dealing with the Black-Scholes, we can refer to Section 3.1 for the explicit calculation of the optimal CVaR for a given amount of initial wealth. The numeric results are presented in Figure 3.8.

We also consider the dual problem: for a fixed CVaR threshold $\tilde{C}$, specify the optimal client age for an equity-linked life insurance contract. In the Black-Scholes setting, we can employ the results of Section 2.3 to derive the optimal survival probability. Then it's just the matter of using the corresponding life table to find the optimal client age. (Note: depending on the life table, the client age may not be uniquely defined by the survival probability; in our example, we pick the highest possible value). The corresponding numeric results can be observed in Figure 3.9.

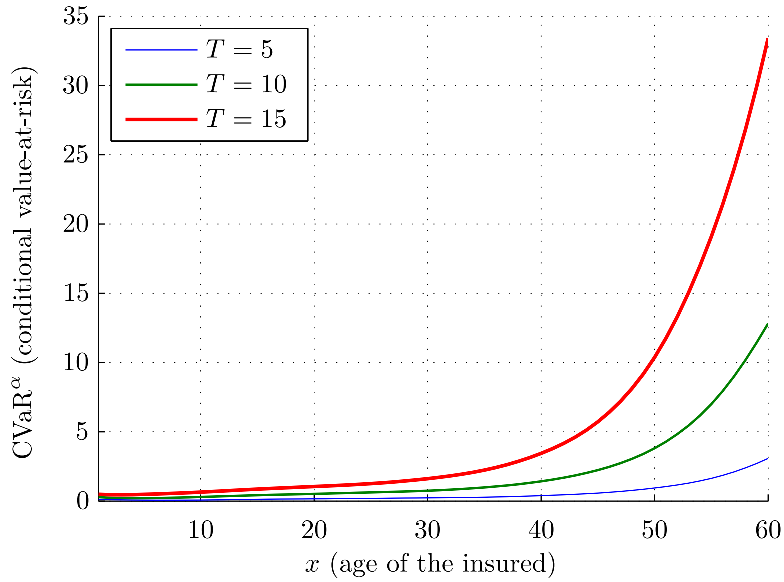

Figure 3.8: CVaR of the optimal hedging strategy at the 99% confidence level for varying age of the insured and varying length of the unit-linked life insurance contract.

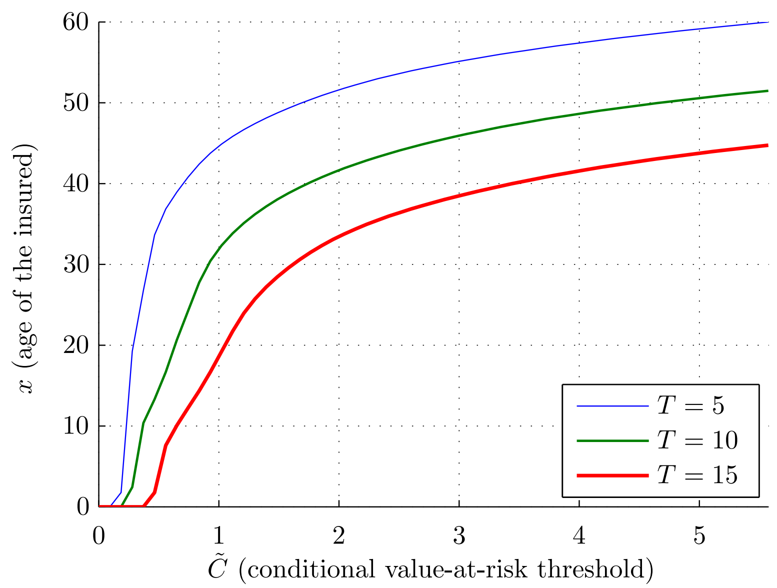

Figure 3.9: The optimal age of the insured for varying length of the unit-linked life insurance contract for varying levels of CVaR threshold at the 99% confidence level.

Chapter 4

Approximate hedging via path-wise comparison

This chapter explores an approach based upon a result from the theory of stochastic processes known as the path-wise comparison theorem (see Theorem 1.4 in the introductory chapter). Our main objective is to demonstrate how path-wise comparison can be used to derive closed-form option price bounds in diffusion models. The stock price process is dominated path-wise by a suitably chosen stochastic process with a known distribution, which allows to obtain explicit bounds for various financial quantities. This chapter is primarily focused on practical applications of the comparison theorem. Towards this end, we provide full analytical derivation of the option price bounds and the approximate hedging strategy in the constant elasticity of variance (CEV) model along with the supplementary numerical illustrations.[1]

Considering the growing importance of the comparison theorem and its

[1] This chapter represents a joint research of the author with Vladislav Krasin and Alexander Melnikov.

applicability to the problems arising in mathematical finance (Krasin [2010], Krasin and Melnikov [2010]), we shall briefly describe its history and point out its main differences from similar results (see, e.g., Hajek [1985]).The path-wise comparison theorem for stochastic differential equations (SDEs), which in this chapter will be referred to as the comparison theorem, was first introduced by [Skorokhod, 1961] who established that the solution for a stochastic diffusion equation with constant diffusion coefficient is a monotone function of the continuous drift coefficient. About a decade later, [Yamada, 1973] significantly weakened Skorokhod's conditions by proving the comparison theorem for SDEs with continuous drift coefficient and the diffusion coefficient satisfying the Hölder condition of order $\alpha \geqslant \frac{1}{2}$. In addition to generalizing the original comparison theorem, Yamada's method, which is essentially based on monotone smoothing of the absolute value function, enables the use of the classic version of the Itō's formula. This approach was further developed in papers by [Melnikov, 1979a] and [Galtchouk, 1982], where the comparison theorem was extended to handle SDEs involving continuous and non-continuous semimartingales, respectively. Since then, these results have been proved again by a number of authors under the same or similar conditions; note, for instance, the work of [Yan, 1986] where the Tanaka-Meyer's formula is used instead of the classic Itō's formula. The first version of the multi-dimensional comparison theorem was suggested in [Melnikov, 1983], along with the method of monotone approximations which allows proving uniqueness and existence of solutions for SDEs driven by semimartingales with non-Lipshitz coefficients (see also Melnikov [1979b] and Barlow and Perkins [1984]). Among the more recent research on the comparison theorem, we would like to mention the works by [Ding and Wu, 1998], [Peng and Zhu, 2006] and [Cohen et al., 2010], the latter of which illustrates the growing interest towards the comparison theorem in the context of the theory of backward SDEs. In addition to [Hajek, 1985], a compelling review of comparison methods in the theory of stochastic processes may be found in [Bergenthum and Ruschendorf, 2007].

Note that monotonicity of the option price with respect to the price of the underlying has been previously used by several authors to construct explicit option price bounds. In [Lèvy, 1985], the stochastic domination results are used to find option price bounds in discrete time markets. A distribution-independent result can be found in [Schepper and Heijnen, 2007], where option price bounds are based on several moments and the mode of the terminal value of the stock price.

The rest of the chapter will be organized into three sections. Section 4.1 contains our main results, where we demonstrate how the comparison theorem can be applied to the CEV model in order to derive the upper option price bound. Section 4.2 presents the approximate hedging strategy and the conditional value-at-risk upper bound. In Section 4.3, we conclude the paper with numerical illustrations in the CEV model and we also compare the relative precision of our approach with the distribution-free method.

4.1 Deriving option price bounds in the CEV model

In this section we are going to demonstrate how the comparison theorem can be applied in the constant elasticity of variance (CEV) model to derive the upper and lower price bounds for contingent claims with monotone payoff functions.

The CEV model which was first introduced in [Cox and Ross, 1976] suggests that the risk-neutral stock price process follows

$$\begin{cases} dS_t = rS_t dt + \sigma S_t^{\frac{\beta}{2}} dW_t, \\ S_0 = s, \end{cases} \tag{4.1}$$

where the initial stock price s, the risk-free interest rate r and the volatility σ are non-negative constants, $t > 0$ and $0 < \beta < 2$. Note that the solution for (4.1) is positive till random time ν_0, where S_t hits zero, and is zero afterwards.

The CEV model generally provides a better fit for the observable option prices relative to the Black-Scholes (see, e.g., MacBeth and Merville [1980]). However, the improved precision comes at a price: the CEV model does not admit a simple way of calculating option prices. [Cox and Ross, 1976] derived a general formula for the option price and [Schroder, 1989] expressed it in terms of the non-central chi-square distribution, but both approaches rely on numerical procedures in some form or another. Applying the comparison theorem to this model allows us to construct an explicit option price bound featuring a high level of precision.

Consider a smooth transform $X_t = F(S_t)$ of the discounted stock price process. According to the Itō's formula,

$$dX_t = \left(rS_t F'(S_t) + \frac{\sigma^2}{2} S_t^{\beta} F''(S_t) \right) dt + \sigma S_t^{\frac{\beta}{2}} F'(S_t) dW_t. \tag{4.2}$$

The key point in our approach is to find such transform $F(\cdot)$ that the SDE for X_t has a relatively simple form. In this regard, we may want the volatility term in (4.2) to become constant and equal to one:

$$\sigma S_t^{\frac{\beta}{2}} F'(S_t) = 1. \tag{4.3}$$

In order to simplify X_t even further, we would like it to take off from the origin, i.e.

$$X_0 = F(s) = 0. \tag{4.4}$$

By solving ordinary differential equation (4.3) with boundary value (4.4), we obtain

$$F(z) = \frac{1}{\sigma}\frac{2}{2-\beta}\left(z^{\frac{2-\beta}{2}} - s^{\frac{2-\beta}{2}}\right). \tag{4.5}$$

Taking derivatives yields

$$F'(z) = \frac{1}{\sigma} z^{-\frac{\beta}{2}}, \quad F''(z) = -\frac{1}{\sigma}\frac{\beta}{2} z^{-\frac{2+\beta}{2}}, \tag{4.6}$$

and therefore (4.2) on $[0, \nu_0)$ takes the form

$$dX_t = \left(\frac{r}{\sigma} S_t^{\frac{2-\beta}{2}} - \frac{\sigma\beta}{4} S_t^{-\frac{2-\beta}{2}}\right) dt + dW_t. \tag{4.7}$$

Since

$$S_t = F^{-1}(X_t) = \left(\sigma \frac{2-\beta}{2} X_t + s^{\frac{2-\beta}{2}}\right)^{\frac{2}{2-\beta}}, \tag{4.8}$$

we obtain the following SDE for X_t on $[0, \nu_0)$:

$$dX_t = \left(r\left(\frac{2-\beta}{2} X_t + \frac{1}{\sigma} s^{\frac{2-\beta}{2}}\right) - \frac{\sigma\beta}{4}\left(\sigma\frac{2-\beta}{2} X_t + s^{\frac{2-\beta}{2}}\right)^{-1}\right) dt + dW_t. \tag{4.9}$$

Let us now consider process Y_t defined by

$$\begin{cases} dY_t = r\left(\frac{2-\beta}{2}Y_t + \frac{1}{\sigma}s^{\frac{2-\beta}{2}}\right)dt + dW_t, \\ Y_0 = 0. \end{cases} \tag{4.10}$$

Equation (4.10) is inhomogeneous and linear in the narrow sense (see, for instance, Kloeden and Platen [1992]), hence its solution Y_t can be derived explicitly:

$$Y_t = \exp\left\{rt\frac{2-\beta}{2}\right\}\left(\frac{2s^{\frac{2-\beta}{2}}}{\sigma(2-\beta)}\left(1-\exp\left\{-rt\frac{2-\beta}{2}\right\}\right) + I_t\right), \tag{4.11}$$

where

$$I_t = \int_0^t \exp\left\{-r\tau\frac{2-\beta}{2}\right\}dW_\tau = \sqrt{\frac{1-\exp\{-rt(2-\beta)\}}{r(2-\beta)}}\xi, \tag{4.12}$$

with ξ being a standard normal random variable. Process Y_t is Gaussian and Markov, and for any fixed moment of time $t > 0$ its distribution is normal:

$$Y_t = \sqrt{\frac{\exp\{(2-\beta)rt\}-1}{r(2-\beta)}}\xi + \frac{2s^{\frac{2-\beta}{2}}}{\sigma(2-\beta)}\left(\exp\left\{\frac{2-\beta}{2}rt\right\}-1\right). \tag{4.13}$$

We have specifically constructed Y_t in such a way that the drift term in (4.10) dominates the drift term in (4.9); besides, $X_0 = Y_0 = 0$. Therefore, according to the comparison theorem, for $t \in [0, \nu_0]$ we have

$$X_t \leqslant Y_t, \quad \text{(a.s.)} \tag{4.14}$$

hence

$$X_t \leqslant |Y_t|, \quad \text{(a.s.)} \tag{4.15}$$

or, equivalently,

$$S_t \leqslant Z_t \quad \text{(a.s.)}, \tag{4.16}$$

where

$$Z_t = F^{-1}(|Y_t|). \tag{4.17}$$

Note that taking the absolute value in (4.17) is essential in the general case since $F^{-1}(Y_t)$ may be undefined for negative values of Y_t.

Consider a contingent claim of European type whose payoff at time $t = T$ is equal to $f(S_T)$. Its risk-neutral price at time $t = 0$ is equal to

$$f_0 = \mathbb{E}\left(\mathrm{e}^{-rT} f(S_T)\right). \tag{4.18}$$

If the payoff is a non-increasing (non-decreasing) function of the stock price, inequality (4.16) can be used to determine the upper (lower) price bound.

Theorem 4.1: *The price at time $t = 0$ of a contingent claim with payoff $f(S_T)$ is bounded from above (below) by*

$$C_0 = \mathbb{E}\left(\mathrm{e}^{-rT} f(Z_T)\right), \tag{4.19}$$

provided that f is a non-decreasing (non-increasing) function.

In our particular case Z_T is a function of a normally distributed random variable, hence the expectation in (4.19) can be computed explicitly.

For the purpose of illustration, consider a European call option with the strike price of K in the CEV model with $\beta = 1$. Note that we can replace $|Y_t|$

in (4.17) with Y_t since in the case when $\beta = 1$ the inverse function $F^{-1}(z)$ is defined for all real z. According to Theorem 4.1, the upper bound $C_0 = C(T, s)$ for the call price can be calculated as

$$C(T,s)=\mathrm{e}^{-rT}\mathbb{E}\left(\left(\frac{\sigma}{2}Y_T+\sqrt{s}\right)^2-K\right)^+. \tag{4.20}$$

Recalling that Y_T is defined by (4.13), we arrive at the following result.

Corollary 4.1.1: The price of a European call option with the strike price of K in the CEV model with $\beta = 1$ is bounded from above by

$$C(T,s)=\mathrm{e}^{-rT}\mathbb{E}\left(\left(\frac{1}{2}\sigma\sqrt{\frac{\mathrm{e}^{rT}-1}{r}}\xi+\sqrt{s\mathrm{e}^{rT}}\right)^2-K\right)^+, \tag{4.21}$$

where ξ is a standard normal random variable.

4.2 Construction of the approximate hedging strategy

The comparison theorem can also be used to construct an approximate hedging strategy. Let us show how this can be done in the CEV model. First, we will derive the SDE for Z_t by applying the Itō formula to (4.17):

$$\begin{aligned} dZ_t = &\left(r\left(\frac{2-\beta}{2}Y_t+\frac{1}{\sigma}s^{\frac{2-\beta}{2}}\right)\left(F^{-1}\right)'(Y_t)+\frac{1}{2}\left(F^{-1}\right)''(Y_t)\right)dt \\ &+\left(F^{-1}\right)'(Y_t)\,dW_t. \end{aligned} \tag{4.22}$$

Calculating the derivatives,

$$\left(F^{-1}\right)'(Y_t) = \sigma\left(\sigma\frac{2-\beta}{2}Y_t + s^{\frac{2-\beta}{2}}\right)^{\frac{\beta}{2-\beta}} = \sigma Z_t^{\frac{\beta}{2}}, \tag{4.23}$$

$$\left(F^{-1}\right)''(Y_t) = \frac{\beta\sigma^2}{2}\left(\sigma\frac{2-\beta}{2}Y_t + s^{\frac{2-\beta}{2}}\right)^{\frac{2\beta-2}{2-\beta}} = \frac{\beta\sigma^2}{2}Z_t^{\beta-1}, \tag{4.24}$$

we obtain

$$\begin{aligned} dZ_t &= \left(\frac{r}{\sigma}Z_t^{\frac{2-\beta}{2}}\sigma Z_t^{\frac{\beta}{2}} + \frac{\beta\sigma^2}{4}Z_t^{\beta-1}\right)dt + \sigma Z_t^{\frac{\beta}{2}}dW_t \\ &= \left(rZ_t + \frac{\beta\sigma^2}{4}Z_t^{\beta-1}\right)dt + \sigma Z_t^{\frac{\beta}{2}}dW_t. \end{aligned} \tag{4.25}$$

Consider a two-dimensonal degenerate diffusion process (S_t, Z_t) with S_t and Z_t being strong solutions to SDEs (4.1) and (4.25) respectively. This process is homogeneous and Markov, therefore inequality

$$S_t \leqslant Z_t \quad \text{(a.s.)}, \tag{4.26}$$

which is provided by the comparison theorem and holds true on time interval $[0, T]$ if the initial value of the process at time $t = 0$ is (s, s), will also be true on interval $[u, T]$ given the initial value of the process at time $t = u$ is (S_u, S_u). Define a Markov family of processes $Z_t^{(u)}$ as a family of solutions to SDE

$$\begin{cases} dZ_t^{(u)} = \left(rZ_t^{(u)} + \frac{\beta\sigma^2}{4}\left(Z_t^{(u)}\right)^{\beta-1}\right)dt + \sigma\left(Z_t^{(u)}\right)^{\frac{\beta}{2}}dW_t, \\ Z_u^{(u)} = S_u, \end{cases} \tag{4.27}$$

for each u, then for any fixed $t \in [u, T]$

$$S_t \leqslant Z_t^{(u)} \quad \text{(a.s.)}. \tag{4.28}$$

Same as before, $Z_t^{(u)}$ can be represented as a function of $Y_t^{(u)}$:

$$Z_t^{(u)} = F_{(u)}^{-1}\left(Y_t^{(u)}\right), \tag{4.29}$$

where

$$F_{(u)}(z) = \frac{1}{\sigma}\frac{2}{2-\beta}\left(z^{\frac{2-\beta}{2}} - S_u^{\frac{2-\beta}{2}}\right), \tag{4.30}$$

$$F_{(u)}^{-1}(z) = \left(\sigma\frac{2-\beta}{2}z + S_u^{\frac{2-\beta}{2}}\right)^{\frac{2}{2-\beta}}, \tag{4.31}$$

and $Y_t^{(u)}$ is a solution to

$$\begin{cases} dY_t^{(u)} = r\left(\frac{2-\beta}{2}Y_t^{(u)} + \frac{1}{\sigma}S_u^{\frac{2-\beta}{2}}\right)dt + dW_t, \\ Y_u^{(u)} = 0. \end{cases} \tag{4.32}$$

We can now use (4.27) to construct an option price bound at time u.

Theorem 4.2: *The price at time $t = u$ of a contingent claim with payoff $f(S_T)$ is bounded from above (below) by*

$$C_u = \mathbb{E}\left(\mathrm{e}^{-r(T-u)}f\left(Z_T^{(u)}\right) \mid \mathcal{F}_u\right), \tag{4.33}$$

provided that f is a non-decreasing (non-increasing) function.

For the remainder of this section, we assume that f is non-decreasing and focus on the upper price bound. The conditional distribution of $Y_t^{(u)}$ given S_u is normal, therefore the conditional expectation in (4.33) can be computed explicitly as a function of current stock price and time. This function can be used to estimate the option price from above at an arbitrary moment of time, and it also allows constructing a hedging strategy. In case $\beta = 1$ the upper price bound is equal to $C_u = C(T-u, S_u)$, where $C(t,x)$ is defined by (4.20).

It turns out that for for any $v \in [u, T]$

$$Z_t^{(v)} \leqslant Z_t^{(u)} \quad \text{(a.s.)}, \quad v \leqslant t \leqslant T. \tag{4.34}$$

To prove (4.34), let us rewrite it as

$$F_{(v)}^{-1}\left(Y_t^{(v)}\right) \leqslant F_{(u)}^{-1}\left(Y_t^{(u)}\right). \tag{4.35}$$

Applying increasing function $F_{(v)}$ to both sides, we obtain

$$Y_t^{(v)} \leqslant F_{(v)}\left(F_{(u)}^{-1}\left(Y_t^{(u)}\right)\right) = \frac{1}{\sigma}\frac{2}{2-\beta}\left(\sigma\frac{2-\beta}{2}Y_t^{(u)} + S_u^{\frac{2-\beta}{2}} - S_v^{\frac{2-\beta}{2}}\right). \tag{4.36}$$

Therefore, (4.34) is equivalent to

$$Y_t^{(v)} - Y_t^{(u)} \leqslant F_{(v)}(S_u) \quad \text{(a.s.)}, \quad v \leqslant t \leqslant T. \tag{4.37}$$

Recalling that $Y_t^{(u)}$ is a solution to (4.32),

$$d\left(Y_t^{(v)} - Y_t^{(u)}\right) = \frac{2-\beta}{2}\left((Y_t^{(v)} - Y_t^{(u)}) - F_v(S_u)\right)dt, \tag{4.38}$$

or, if we pin the initial value at time $t = v$,

$$Y_t^{(v)} - Y_t^{(u)} = F_v(S_u) + \left(Y_v^{(v)} - Y_v^{(u)} - F_v(S_u)\right) \exp\left\{r\frac{2-\beta}{2}(t-v)\right\}. \quad (4.39)$$

The path-wise comparison theorem provides

$$Z_v^{(v)} = S_v \leqslant Z_v^{(u)} \quad (4.40)$$

or, equivalently,

$$Y_v^{(v)} - Y_v^{(u)} \leqslant F_{(v)}(S_u), \quad (4.41)$$

hence (4.37) follows directly from (4.39), which concludes the proof.

The option price estimate at time $t = T$ is equal to

$$\mathbb{E}\left(\mathrm{e}^{-r(T-T)} f\left(Z_T^{(T)}\right) \mid \mathcal{F}_T\right) = \mathbb{E}\left(f(S_T) \mid \mathcal{F}_T\right), \quad (4.42)$$

therefore the proposed hedging strategy constitutes a perfect hedge. However, since different processes $Z_T^{(u)}$ are used as upper boundaries for S_T at different moments of time, the option price estimate is not a martingale, which implies that the strategy is not self-financing. Moreover, it can be shown that for a non-decreasing payoff function the proposed hedging strategy is a strategy with consumption.

Denote the discounted hedge value by L_t and consider two moments of time $v > u$:

$$L_u = \mathrm{e}^{-rT}\mathbb{E}\left(f\left(Z_T^{(u)}\right) \mid \mathcal{F}_u\right), \quad (4.43)$$

$$L_v = \mathrm{e}^{-rT}\mathbb{E}\left(f\left(Z_T^{(v)}\right) \mid \mathcal{F}_v\right). \quad (4.44)$$

Applying (4.34) yields

$$\mathbb{E}(L_v \mid \mathcal{F}_u) = \mathrm{e}^{-rT}\mathbb{E}\left(f\left(Z_T^{(v)}\right) \mid \mathcal{F}_u\right) \leqslant \mathrm{e}^{-rT}\mathbb{E}\left(f\left(Z_T^{(u)}\right) \mid \mathcal{F}_u\right) = L_u. \quad (4.45)$$

Thus, process $L = (L_t)_{t\in[0,T]}$ is a supermartingale and as such it can be represented as a sum of a martingale (which can be viewed as the discounted value of some self-financing strategy) and a decreasing process. The latter can be treated as a consumption process which continuously removes excess hedging capital from the strategy until the value of the hedge exactly matches the option's value at maturity. Note: for non-increasing payoff functions, the option price estimate is a submartingale and the hedge is a strategy with the inflow of capital.

Remark 4.1: In addition to providing a way of constructing option price bounds and the associated approximate hedging strategies, the path-wise comparison theorem may also be used to derive bounds for various risk measures. To demonstrate this we shall estimate conditional value-at-risk of a European call option in the CEV model from the option seller's point of view.

According to Theorem 1.1, CVaR associated with loss L at confidence level $\alpha \in (0,1)$ can be defined as

$$\mathrm{CVaR}_\alpha = \inf\left\{z + \frac{1}{1-\alpha}\tilde{\mathbb{E}}\left((L-z)^+\right) \;:\; z \in \mathbb{R}\right\}. \quad (4.46)$$

Note that the expectation in (4.46) is taken with respect to the historical measure $\tilde{\mathbb{P}}$, under which the stock price follows

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t^{\frac{\beta}{2}} d\tilde{W}_t, \\ S_0 = s, \end{cases} \tag{4.47}$$

where μ is a constant drift coefficient and $\tilde{W}_t$ is a $\tilde{\mathbb{P}}$–Brownian motion.

If $\beta = 1$ and $L = f(S_T) = (S_T - K)^+$, we can use (4.16) to estimate conditional value-at-risk from above in the following way:

$$\mathrm{CVaR}_\alpha \leqslant \min_{z \in \mathbb{R}} \left\{ z + \frac{1}{1-\alpha} \tilde{\mathbb{E}} \left(\left(\tilde{Z}_T^2 - K - z \right)^+ \right) \right\},$$

where

$$\tilde{Z}_T = \frac{1}{2} \sigma \sqrt{\frac{\mathrm{e}^{\mu T} - 1}{\mu}} \tilde{\xi} + \sqrt{s \mathrm{e}^{\mu T}},$$

and $\tilde{\xi}$ is a standard normal random variable.

4.3 Numerical illustration and comparison with the distribution-free method

In this section, we assume that the risk-neutral stock price follows the CEV model with $\beta = 1$:

$$\begin{cases} dS_t = r S_t dt + \sigma \sqrt{S_t} dW_t, \\ S_0 = s. \end{cases} \tag{4.48}$$

Our objective is to compare the true model price of a call option and the upper price bound provided by the comparison theorem. To evaluate the relative precision of the proposed comparison theorem price estimate, we shall also compute the so-called distribution-free upper price bound which was suggested in [Schepper and Heijnen, 2007] and which is largely based on the results presented in [Jansen et al., 1986].

Denote by μ_m the m-th moment of random variable S_T:

$$\mu_m = \mathbb{E}\left(S_T^m\right). \tag{4.49}$$

According to [Schepper and Heijnen, 2007], if three moments of S_T are known, the price of the call option satisfies the boundary condition

$$\mathbb{E}\left((S_T - K)^+\right) \leqslant F_4(K), \tag{4.50}$$

with

$$F_4(K) = \begin{cases} (\mu_1 - K) + \dfrac{\mu_2 - \mu_1^2}{\mu_2} K, & K \leqslant \dfrac{\mu_2}{2\mu_1}, \\ \dfrac{(\mu_1 - K) + \sqrt{(\mu_2 - \mu_1^2) + (\mu_1 - K)^2}}{2}, & \dfrac{\mu_2}{2\mu_1} \leqslant K \leqslant \dfrac{c_1 + c_2}{2}, \\ \dfrac{\mu_{3\mu 1} - \mu_2^2}{\mu_3 - 2c_2\mu_2 + c_2^2\mu_1} \cdot \dfrac{c_2 - K}{c_2}, & \dfrac{c_1 + c_2}{2} \leqslant K \leqslant \dfrac{2c_2^2}{3c_2 - c_1}, \\ \dfrac{\mu_3\mu_1 - \mu_2^2}{\mu_3 - 2c_3\mu_2 + c_3^2\mu_1} \cdot \dfrac{c_2 - K}{c_2}, & K \geqslant \dfrac{2c_2^2}{3c_2 - c_1}, \end{cases} \tag{4.51}$$

where $c_1 < c_2$ are the roots (which are guaranteed to exist and which are real and positive) of

$$\left(\mu_2 - \mu_1^2\right) x^2 + \left(\mu_1\mu_2 - \mu_3\right) x + \left(\mu_1\mu_3 - \mu_2^2\right) = 0, \tag{4.52}$$

and where c_3 is the unique root in the interval $[c_2, +\infty)$ of

$$2\mu_1 x^3 - (2\mu_2 + 3K\mu_1)\, x^2 + 4K\mu_2 x - K\mu_3 = 0. \tag{4.53}$$

In order to find the distribution-free upper price bound in the CEV model, we need to calculate the first three moments of S_T. It turns out that in the case $\beta = 1$ the m-th moment of the CEV stock price can be derived in an analytical way; this fact along with the full proof is presented in Lemma 5.5 in Section 5.3.

Applying Lemma 5.5 for $m \in \{1, 2, 3\}$ yields

$$\mu_1 = s\mathrm{e}^{rT}, \tag{4.54}$$

$$\mu_2 = s^2\mathrm{e}^{2rT} + 2\kappa^{-1}s\mathrm{e}^{rT}, \tag{4.55}$$

$$\mu_3 = s^3\mathrm{e}^{3rT} + 6\kappa^{-1}s^2\mathrm{e}^{2rT} + 6\kappa^{-2}s\mathrm{e}^{rT}, \tag{4.56}$$

where

$$\kappa = \frac{2r}{\sigma^2\left(\mathrm{e}^{rT} - 1\right)}. \tag{4.57}$$

For the numerical demonstration, we consider the CEV model with $\beta = 1$, $T = 1$, $r = 0.05$, $\sigma = 0.35$ and a European call option with the strike price of $K = 15$. The true model price of the option is calculated with the help of the exact pricing formula given in [Schroder, 1989]:

$$C_t = S_t Q(2y; 4, 2x) - K\mathrm{e}^{-rT}(1 - Q(2x; 2, 2y)), \tag{4.58}$$

with

$$x = \kappa S_t \mathrm{e}^{r\tau}, \quad y = \kappa K, \tag{4.59}$$

where κ is given in (4.57) and $Q(y; \nu, \lambda)$ is the survivor function at y for a non-central chi-squared random variable with ν degrees of freedom and non-centrality parameter λ.

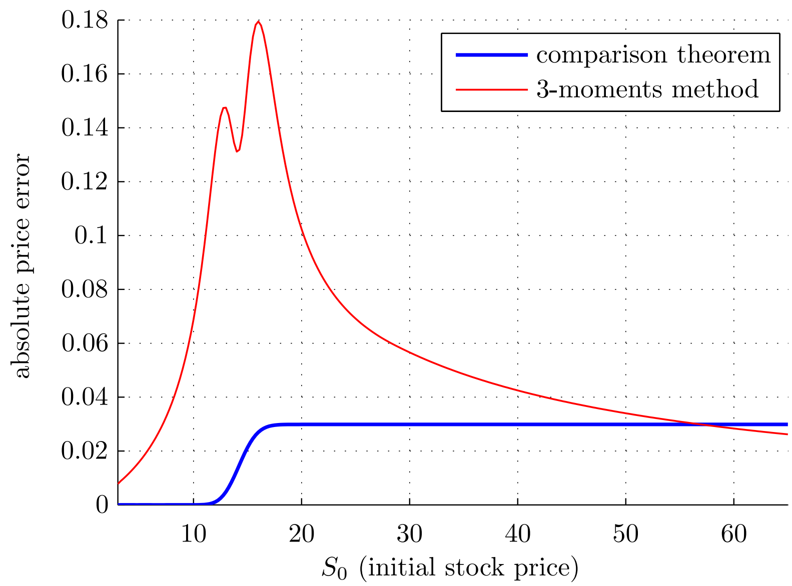

Figure 4.1: The absolute error of the upper price bounds based on the comparison theorem method and the 3-moments method for varying values of initial stock price in the CEV model.

As can be seen on Figure 4.1, the upper price bound based on the comparison theorem is far more accurate than the distribution-free bound for all values of initial stock price but the extreme high. In addition, the comparison theorem bound exhibits superb level of precision for out-of-the-money options.

Figure 4.2 can be viewed as an illustration of the fact that the comparison theorem bound converges to the true price as the option approaches its maturity, which essentially makes it reasonable to use the proposed bound as

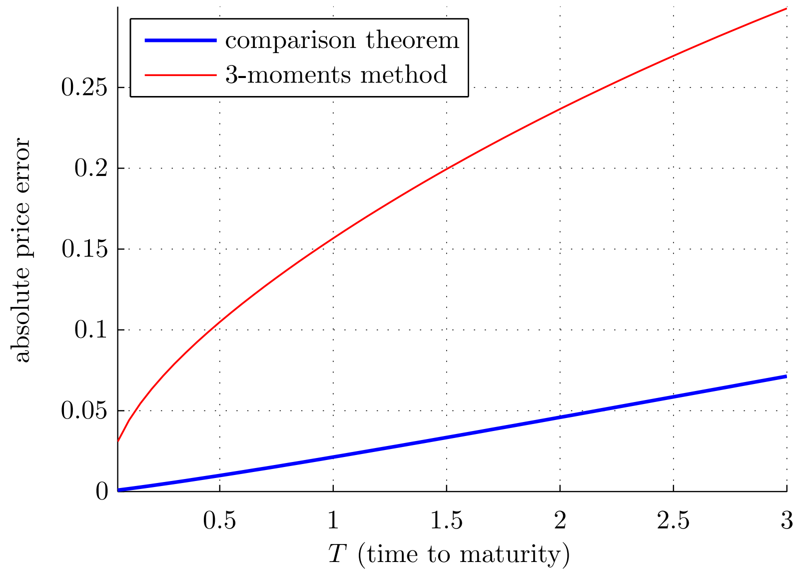

Figure 4.2: The absolute error of the upper price bounds based on the comparison theorem method and the 3-moments method for an at-the-money option with varying time to maturity in the CEV model.

an alternative to Schroder's formula for extremely short maturities. Indeed, as the time to maturity nears zero, computation of the Schroder's formula becomes increasingly slow and inaccurate (since both the quantile and the non-centrality parameter become infinite) while the upper price bound provided by the comparison theorem becomes increasingly precise. Essentially, this fact also makes it possible to construct an approximate hedging strategy as suggested in Section 4.2.

Chapter 5

Supplementary results

5.1 Weak compactness of the set of densities of equivalent sigma-martingale measures

In this section, we investigate why the assumption of weak compactness in L^1 of the set of densities of equivalent sigma-martingale measures is never satisfied in the incomplete market case.

We shall first discuss the weak closedness and the relative weak compactness in a discrete-time case where the results are more intuitive and we shall then move on to the proof of non-closedness in a more general continuous-time semimartingale setting. It turns out that in a discrete-time case the set of densities of equivalent martingale measures may be not compact for one of the following reasons, each of which we consider in detail below:

1. non-closedness due to non-equivalence of the limiting measure (Theorem 5.1);

2. non-closedness due to the violation of the martingale property when

the support sets of conditional distributions of price increments are not bounded (Theorem 5.2);

3. relative non-compactness in the case when the support sets of conditional distributions of price increments contain a limit point (Theorem 5.3).

In a discrete-time case with finite time horizon, consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathcal{T}}, \mathbb{P})$ with $\mathcal{T} = \{0, 1, \ldots T\}$, $T \in \mathbb{N}$, under the standard assumption that $\mathcal{F}_0 = \{\emptyset, \Omega\}$. On this probability space, consider an $\mathbb{R}^d$-valued $\mathcal{F}_t$-measurable discounted price process of the risky asset $X = (X_t)_{t\in\mathcal{T}}$, where the discounting is being done with respect to an $\mathcal{F}_{t-1}$-measurable price process of the risk-free asset $B = (B_t)_{t\in\mathcal{T}}$.

Throughout the section, we shall use the following notation:

- ri(A) is the relative interior of set A;
- conv(A) is the convex hull of set A;
- lin(A) is the linear hull of set A;
- aff(A) is the affine hull of set A;
- $\overline{\text{co}}(A)$ is the closure of the convex hull of set A;
- supp$(\mathbb{P})$ is the topological support of measure $\mathbb{P}$;
- $B(x, \varepsilon)$ is the closed ε-neighborhood of vector x;
- $A^c = \mathcal{S} \backslash A$ is the complement of set A with respect to space $\mathcal{S}$;
- $||x||$ is the Euclidean norm of vector x;
- $x \vee y = \min\{x, y\}$, $x \wedge y = \max\{x, y\}$ for $x, y \in \mathbb{R}$;

- $L^1(\mathbb{P})$ is the space of $\mathbb{P}$-integrable real-valued functions.

In addition, denote by $\mathbb{P}_t(\omega, \cdot)$ the regular conditional distributions of increments ΔX_t:

$$\mathbb{P}_t(\omega, \cdot) = \mathbb{P}(\Delta X_t \in \cdot \mid \mathcal{F}_{t-1})(\omega), \quad \text{for } t \in \mathcal{T}_+, \tag{5.1}$$

where $\mathcal{T}_+ = \{1, 2, \ldots T\}$.

Denote by $\mathcal{P}^*$ and $\mathcal{P}^*_{\text{loc}}$ the sets of probability measures equivalent to $\mathbb{P}$ under which X is a martingale or a local martingale, respectively. According to an extended version of the first fundamental theorem of asset pricing given in [Shiryaev, 1998], the following statements are equivalent:

(1a) the (B, X) market is arbitrage-free;

(1b) $\mathcal{P}^* \neq \emptyset$;

(1c) $\mathcal{P}^*_{\text{loc}} \neq \emptyset$;

(1d) $0 \in \text{ri}(\overline{\text{co}}(\text{supp}(\mathbb{P}_t(\omega, \cdot))$ for all $t \in \mathcal{T}_+$ and $\mathbb{P}$-a.s. for all $\omega \in \Omega$.

In addition, if the (B, X) market is arbitrage-free, then the following statements are equivalent (an extended version of the second fundamental theorem of asset pricing, Shiryaev [1998]):

(2a) the (B, X) market is complete;

(2b) $\mathcal{P}^*$ is a singleton;

(2c) $\mathcal{P}^*_{\text{loc}}$ is a singleton;

(2d) There exist $d+1$ predictable affinely independent $\mathbb{R}^d$-valued processes $(a_{1,t}, a_{2,t}, \dots a_{d+1,t})$, $t \in \mathcal{T}_+$, such that $\operatorname{supp} \mathbb{P}_t(\omega, \cdot) \subseteq \{a_{1,t}, a_{2,t}, \dots a_{d+1,t}\}$, for $t \in \mathcal{T}_+$, $\mathbb{P}$-a.s.

Proposition 5.1: *If the (B, X) market is not arbitrage-free or if it is complete, the set of densities of equivalent (local) martingale measures is compact.*

Indeed, if the market is not arbitrage-free, then the set of (local) martingale measures is empty and so is the set of corresponding densities. If the market is complete, then both sets are singletons. Either way, the set of densities is compact. We shall see below that this is, in fact, the only case when the set of densities of equivalent martingale measures is compact.

In a discrete-time case the terms sigma-martingale, generalized martingale[1] and local martingale are interchangeable (Jacod and Shiryaev [1998]). Moreover, every discrete-time local martingale bounded from below (above) with finite expectation at the initial moment of time is a proper martingale (Shiryaev [1998]), which in conjunction with non-negativity of X and the choice of $\mathcal{F}_0 = \{\emptyset, \Omega\}$ implies that $\mathcal{P}^*_{\text{loc}} = \mathcal{P}^*$. Therefore, it is sufficient to only consider the class of equivalent measures under which the discounted price process is a martingale.

Note also that convergence of densities in $L^1(\mathbb{P})$ is equivalent to convergence of the associated probability measures in total variation norm to a measure that is absolutely continuous with respect to $\mathbb{P}$. Choosing an alternative reference measure $\mathbb{P}^* \sim \mathbb{P}$ would not affect the L^1-convergence of densities (although it would obviously alter density values). Besides, the set $\mathcal{P}^*$ would also

[1] By following the terminology in [Jacod and Shiryaev, 1998], a generalized martingale is a process that is only required to satisfy the martingale property, whereas a (proper) martingale is also assumed to lie in L^1.

remain unchanged since it is only dependent on the reference measure up to measure equivalence. Therefore, we can replace the original reference measure $\mathbb{P}$ with an element from $\mathcal{P}^*$ without loss of generality.

Assumption 5.1: The reference probability measure is a martingale measure, i.e. $\mathbb{P} \in \mathcal{P}^*$.

For each $\mathbb{P}^* \in \mathcal{P}^*$, define the Radon-Nikodym densities

$$Z_{\mathbb{P}^*} = \frac{d\mathbb{P}^*}{d\mathbb{P}} \tag{5.2}$$

and the set of densities of equivalent martingale measures

$$Z_{\mathcal{P}^*} = \{Z_{\mathbb{P}^*} : \mathbb{P}^* \in \mathcal{P}^*\}. \tag{5.3}$$

Before we move on to discussing the closedness and the relative compactness of $Z_{\mathcal{P}^*}$, we shall introduce several auxiliary lemmas.

Lemma 5.1: *Let $\mathcal{P}^* \neq \emptyset$ be the set of equivalent martingale (local martingale, sigma-martingale) measures and $Z_{\mathcal{P}^*}$ be the set of the associated densities with respect to $\mathbb{P}$. Then, if there exists a probability measure $\tilde{\mathbb{Q}}$ such that $\tilde{\mathbb{Q}} \ll \mathbb{P}$, $\tilde{\mathbb{Q}} \nsim \mathbb{P}$ and X is a martingale (local martingale, sigma-martingale) under $\tilde{\mathbb{Q}}$, the set $Z_{\mathcal{P}^*}$ is not weakly closed in $L^1(\mathbb{P})$.*

Proof. Since $Z_{\mathcal{P}^*}$ is non-empty, we can select an equivalent measure $\mathbb{Q} \in \mathcal{P}^*$ and construct a family of measures

$$\mathbb{Q}_n = \frac{1}{n}\mathbb{Q} + \left(1 - \frac{1}{n}\right)\tilde{\mathbb{Q}}, \quad n \in \mathbb{N}. \tag{5.4}$$

The measures $\mathbb{Q}_n$ are all in $\mathcal{P}^*$, hence the densities

$$Z_{\mathbb{Q}_n} = \frac{1}{n} Z_{\mathbb{Q}} + \left(1 - \frac{1}{n}\right) Z_{\tilde{\mathbb{Q}}}, \quad n \in \mathbb{N}, \tag{5.5}$$

are in $Z_{\mathcal{P}^*}$. However,

$$\lim_{n \to \infty} Z_{\mathbb{Q}_n} = Z_{\tilde{\mathbb{Q}}} \notin Z_{\mathcal{P}^*}, \tag{5.6}$$

which implies that $Z_{\mathcal{P}^*}$ is not closed in $L^1(\mathbb{P})$.

Moreover, $Z_{\mathcal{P}^*}$ is not weakly closed since in a locally convex normed space a convex set is weakly closed if and only if it is strongly closed (see e.g. Akilov and Kantorovich [1984]). □

Lemma 5.2: *Let $(\mathbb{R}^d, \mathcal{B}, \mathbb{P})$ be a probability space with Borel σ-algebra $\mathcal{B}$, and $B \in \mathcal{B}$ be an almost sure event:*

$$\mathbb{P}(B) = 1. \tag{5.7}$$

Then

$$\int_{\mathbb{R}^d} x \; \mathbb{P}(dx) \in \operatorname{conv}(B). \tag{5.8}$$

Proof[2]**.** Denote

$$y = \int_{\mathbb{R}^d} x \; \mathbb{P}(dx) \tag{5.9}$$

and assume that $y \notin \operatorname{conv}(B)$. Let $V = \operatorname{aff}(\operatorname{supp}(\mathbb{P}))$, then V is the affine subspace of minimum dimension such that $\mathbb{P}(V) = 1$. It is easy to show that $y \in \overline{\operatorname{co}}(\operatorname{supp}(\mathbb{P}))$, thus $y \in V$ since the affine hull is closed in $\mathbb{R}^d$ and thus contains the closure of the convex hull. By the hyperplane separation theorem

[2] The proof of Lemma 5.2 belongs to George Lowther.

there exists a non-trivial affine map $L : V \mapsto \mathbb{R}$ such that

$$L(x) \geqslant 0, \quad \text{for } x \in \operatorname{conv}(B), \tag{5.10}$$

$$L(y) = 0. \tag{5.11}$$

Note that we only get proper separation and not strict separation since the closedness of the convex hull of B is not assumed. Since

$$L(y) = \int_V L(x) \; \mathbb{P}(dx) \tag{5.12}$$

and $L(x)$ is non-negative $\mathbb{P}$-almost everywhere due to (5.10), $L(y) = 0$ requires $L(x) = 0$, $\mathbb{P}$-almost everywhere. This implies

$$\mathbb{P}(\ker L) = 1, \qquad \dim \ker L = \dim V - 1, \tag{5.13}$$

which contradicts the minimality of V. Therefore, $y \in \operatorname{conv}(B)$. □

Remark 5.1: The statement of Lemma 5.2 only holds true for the finite-dimensional spaces; in an infinite-dimensional space, the expected value is only guaranteed to lie in the closure of the convex hull. As an illustration to this fact, consider the space ℓ^2 of square-summable sequences with an orthonormal basis $\{\mathbf{e}_k\}_{k\in\mathbb{N}}$ and a discrete probability measure $\mathbb{P}$ concentrated in points $a_k\mathbf{e}_k$ with $a_k \neq 0$, $\sum_{k=1}^{\infty} a_k^2 < \infty$ and $p_k = \mathbb{P}(\{a_k\mathbf{e}_k\}) > 0$, $k \in \mathbb{N}$. By evaluating the expected value y as a Gelfand-Pettis integral $y = \sum_{k=1}^{\infty} p_k a_k \mathbf{e}_k \in \ell^2$, we conclude that $y \in \overline{\operatorname{co}}(\operatorname{supp}(\mathbb{P}))$ but $y \notin \operatorname{conv}(\operatorname{supp}(\mathbb{P}))$ since the convex hull only consists of the finite convex combinations of vectors $a_k\mathbf{e}_k$.

Lemma 5.3 (Jacod and Shiryaev [1998]): *Let* $(H, \mathcal{H}, \mathbb{P})$ *be a probability space,* G *be a Polish*[3] *space with its Borel* σ*-field* $\mathcal{G}$*, and let* A *be an* $\mathcal{H} \otimes \mathcal{G}$*-measurable subset of* $H \times G$*, with* H*-projection* $\pi(A) = \{x \in H : \exists y \in G, (x, y) \in A\}$*. Then there exists a* G*-valued* $\mathcal{H}$*-measurable function* Y *such that* $(x, Y(x)) \in A$ *for* $\mathbb{P}$*-almost all* x *in* $\pi(A)$*.*

Note that Lemma 5.3 (the *measurable selection theorem*) is an adapted version of the complete section theorem whose full statement and proof may be found in [Dellacherie and Meyer, 1975] (Theorem 82).

Proposition 5.2 (Rogers [1994]): *Let* Π *denote the compact metric space of all* $d \times d$ *orthogonal projection matrices. Then there exists an* $\mathcal{F}_{t-1}$*-measurable mapping* $R : \Omega \to \Pi$ *such that for almost all* ω

$$\ker R(\omega) = \mathrm{lin}(\mathrm{supp}(\mathbb{P}_t(\omega, \cdot))). \tag{5.14}$$

Theorem 5.1: *In a discrete-time arbitrage-free incomplete market, the set* $Z_{\mathcal{P}^*}$ *is not weakly closed in* $L^1(\mathbb{P})$*.*

Proof. To prove the statement of the theorem, it is sufficient to construct a probability measure $\mathbb{Q}$ that is absolutely continuous with respect to $\mathbb{P}$ and not equivalent to $\mathbb{P}$, such that $0 \in \mathrm{ri}(\overline{\mathrm{co}}(\mathrm{supp}(\mathbb{Q}_t(\omega, \cdot)))$ almost surely for all $t \in \mathcal{T}_+$. In this case, we can consider an alternative market model in which the original reference measure is replaced with $\mathbb{Q}$. This market would still be arbitrage-free since $0 \in \mathrm{ri}(\overline{\mathrm{co}}(\mathrm{supp}(\mathbb{Q}_t(\omega, \cdot)))$. Therefore, by the first fundamental theorem there must exist a measure $\tilde{\mathbb{Q}}$ equivalent to $\mathbb{Q}$ such that X is a martingale under $\tilde{\mathbb{Q}}$. However, $\tilde{\mathbb{Q}} \ll \mathbb{P}$ and $\tilde{\mathbb{Q}} \nsim \mathbb{P}$, hence by Lemma 5.1 the set of densities $Z_{\mathcal{P}^*}$ would be not weakly closed.

[3] Separable, metrizable and topologically complete.

Since the (B, X) market is incomplete, there exists a moment of time $t = s \in \mathcal{T}_+$ and an $\mathcal{F}_{s-1}$-measurable set Θ with $\mathbb{P}(\Theta) > 0$ such that for each $\omega \in \Theta$ the set $F_s(\omega) = \mathrm{supp}(\mathbb{P}_s(\omega, \cdot))$ contains at least $d+2$ different points. Moreover, due to the absence of arbitrage, $0 \in \mathrm{ri}(\overline{\mathrm{co}}(F_s(\omega)))$. Let $\Gamma_k = \{\omega \in \Theta : |F_s(\omega)| \leqslant k\}$, then $\Theta = \Gamma^c_{d+1}$ and hence $\mathbb{P}(\Gamma^c_{d+1}) > 0$. Now consider two exhaustive cases.

Case 1: $\mathbb{P}(\Gamma_{2d} \backslash \Gamma_{d+1}) > 0$ and $d > 1$.

In this case it is straightforward to show that there exists an absolutely continuous martingale measure $\mathbb{Q}$ such that $\mathbb{Q}_s(\omega, \cdot)$ is concentrated on at most $d+1$ points (hence it is not equivalent to $\mathbb{P}$). Indeed, since $F_s(\omega)$ is finite, it is compact, therefore $\overline{\mathrm{co}}(F_s(\omega)) = \mathrm{conv}(F_s(\omega))$ and the no-arbitrage condition implies $0 \in \mathrm{ri}(\overline{\mathrm{co}}(F_s(\omega))) \subseteq \mathrm{conv}(F_s(\omega))$. According to the Caratheodory's theorem, zero can be represented as a convex combination of at most $d+1$ points from $F_s(\omega)$. Moreover, the choice of this convex combination can be done in a measurable way. Consider the space $G = \mathbb{R}^{d\times 2d} \times S_{2d}$ of vectors of the form $g = (g_1, g_2, \dots g_{2d}, s_1, s_2, \dots s_{2d})$, where $S_{2d} = \{(s_1, s_2, \dots s_{2d}) : \sum_{i=1}^{2d} s_i = 1, s_i \geqslant 0, i = 1, 2, \dots 2d\} \subseteq \mathbb{R}^{2d}$, $g_i \in \mathbb{R}^d$ are the elements from $F_s(\omega)$ and s_i are the associated probabilities. Let A be a (non-empty) subset of $\Omega \times G$ consisting of all such pairs (ω, g) that satisfy $\omega \in \Gamma_{2d} \backslash \Gamma_{d+1}$, $g_i \neq g_j$, for $i \neq j$, $g_i \in F_s(\omega)$, for $i = 1, 2, \dots 2d$, $(s_1, s_2, \dots s_{2d}) \in S_{2d}$, $\sum_{i=1}^{d+1} g_i s_i = 0$ and $\sum_{i=1}^{2d} \mathbf{1}_{\{s_i > 0\}} \leqslant d+1$. Note that condition $g_i \in F_s(\omega)$ does not violate joint measurability as it can be expressed in a measurable form. Concretely, since $\mathbb{R}^d$ is a second-countable space, its naturaly topology has a countable base $\mathcal{U} = \{U_k\}_{k=1}^{\infty}$. On the other hand, the complement of the topological support is the largest open set of zero measure, so we have $\{(\omega, g) : g_i \in F_s(\omega)\} = (\bigcup_{k=1}^{\infty} \{\omega : \mathbb{P}_s(\omega, U_k) = 0\} \times B_{i,k} \times S_{2d})^c$, where $B_{i,k} = B_1 \times B_2 \times \dots B_{2d} \subseteq \mathbb{R}^{2d\times d}$, $B_j = \mathbb{R}^d$ for $j \neq i$ and

$B_i = U_k$. Therefore, according to Lemma 5.3, there exists a G-valued $\mathcal{F}_{s-1}$-measurable selector $Y(\omega) = (x_1(\omega), x_2(\omega), \dots x_{2d}(\omega), q_1(\omega), q_2(\omega), \dots q_{2d}(\omega))$ which satisfies the above conditions for each $\omega \in \Gamma_{2d} \backslash \Gamma_{d+1}$.

Define measure $\mathbb{Q}$ via its density Z, that is, $d\mathbb{Q}(\omega) = Z(\omega) d\mathbb{P}(\omega)$:

$$Z(\omega) = \mathbf{1}_{(\Gamma_{2d} \backslash \Gamma_{d+1})^c} + \mathbf{1}_{\Gamma_{2d} \backslash \Gamma_{d+1}} \sum_{k=1}^{2d} \frac{q_k(\omega)}{p_k(\omega)} \mathbf{1}_{\{\Delta X_s(\omega) = x_k(\omega)\}}, \tag{5.15}$$

where $x_i(\omega)$ and $q_i(\omega)$ are the components of the measurable selector and $p_k(\omega) = \mathbb{P}_s(\omega, \{g_k(\omega)\})$. By construction, the total number of non-zero weights $q_i(\omega)$ does not exceed $d+1 < 2d$, hence measure $\mathbb{Q}$ is not equivalent to $\mathbb{P}$.

Case 2: $\mathbb{P}(\Gamma_{2d} \backslash \Gamma_{d+1}) = 0$.

Note that in this case the market incompleteness condition implies that $\mathbb{P}(\Gamma_{2d}^c) > 0$ and denote $r(\omega) = \operatorname{rank}(I - R(\omega)) \leqslant d$, where I is the identity matrix and $R(\omega)$ is the $\mathcal{F}_{s-1}$-measurable orthogonal projection matrix from Proposition 5.2. Since $\operatorname{ri}(\overline{\operatorname{co}}(F_s(\omega))) = \operatorname{ri}(\operatorname{conv}(F_s(\omega)))$ (see, for instance, [Rockafellar, 1970]), we can apply the Steinitz theorem (see [Steinitz, 1913], [Steinitz, 1914], [Steinitz, 1916]) which ensures that there exists a finite subset $K(\omega) \subset F_s(\omega)$ with $|K(\omega)| \leqslant 2r(\omega) \leqslant 2d$ such that $0 \in \operatorname{ri}(\operatorname{conv}(K(\omega)))$ on Γ_{2d}^c. Since $|K(\omega)| \leqslant 2d$, for each $\omega \in \Gamma_{2d}^c$ there exists at least one more point $g_{2d+1}(\omega)$ which belongs to $F_s(\omega)$ but does not belong to $K(\omega)$. Moreover, the set $K(\omega)$ and the point $g_{2d+1}(\omega)$ can be chosen in a jointly measurable fashion. Let $G = \mathbb{R}^{(2d+1)\times d}$ and let A be a (non-empty) subset of $\Omega \times G$ of all such pairs (ω, g) that $\omega \in \Gamma_{2d}^c$, $g = (g_1, g_2, \dots g_{2d+1})$, $g_i \in F_s(\omega)$ for $i = 1, 2, \dots 2d+1$, $0 \in \operatorname{ri}(\operatorname{conv}(\{g_1, g_2, \dots g_{2d}\}))$ and $g_{2d+1} \notin \{g_1, g_2, \dots g_{2d}\}$. Note that the condition $0 \in \operatorname{ri}(\operatorname{conv}(\{g_1, g_2, \dots g_{2d}\}))$ can be expressed in a measurable way. Indeed, denote $B = \{g_1, g_2, \dots g_{2d}\}$ and let $\Delta(B)$ be the distance from the origin to the boundary of conv(B). Then $0 \in \text{ri}(\text{conv}(B))$ if and only if $\Delta(B) > 0$. Besides, $\Delta(B) = \inf_{x:||x||=1} h_B(x)$, where h_B is the support function of B, and $h_B(x) = \bigvee_{i=1}^{2d} \langle x, g_i \rangle$, where $\langle \cdot, \cdot \rangle$ is the inner product. Therefore, $\{\Delta(B) > 0\} = \bigcup_{n=1}^{\infty} \bigcap_{x \in \mathcal{D}} \left\{ \bigvee_{k=1}^{2d} \langle x, g_i \rangle > \frac{1}{n} \right\}$, where $\mathcal{D}$ is an arbitrary countable dense subset of $\mathbb{R}^d$, hence Lemma 5.3 ensures the existence of a G-valued $\mathcal{F}_{s-1}$-measurable selector $Y(\omega) = (x_1(\omega), x_2(\omega), \ldots x_{2d+1}(\omega)))$ satisfying the above conditions for each $\omega \in \Gamma_{2d}^c$.

Choose an $\varepsilon(\omega)$-neighborhood of $x_{2d+1}(\omega)$, with $\varepsilon(\omega)$ being $\mathcal{F}_{s-1}$-measurable, so that $B(x_{2d+1}(\omega), \varepsilon(\omega)) \cap \{x_1(\omega), x_2(\omega), \ldots x_{2d}(\omega)\} = \emptyset$ (take, for instance, $\varepsilon(\omega) = \frac{1}{2} \bigwedge_{i=1}^{2d} ||x_{2d+1}(\omega) - x_i(\omega)||$). Note that $\mathbb{P}_s\left(\omega, B^c(x_{2d+1}(\omega), \varepsilon(\omega))\right) > 0$ and consider the following random variable:

$$Z(\omega) = \mathbf{1}_{\Gamma_{2d}}(\omega) + \mathbf{1}_{\Gamma_{2d}^c}(\omega) \frac{\mathbf{1}_{B^c(x_{2d+1}(\omega), \varepsilon(\omega))}(\Delta X_s(\omega))}{\mathbb{P}_s\left(\omega, B^c(x_{2d+1}(\omega), \varepsilon(\omega))\right)}. \tag{5.16}$$

Notice that $Z(\omega) \geqslant 0$ and $\int_\Omega Z(\omega)\mathbb{P}(d\omega) = 1$, therefore it is a proper density and we can consider a probability measure $\mathbb{Q}$ defined as $\mathbb{Q}(d\omega) = Z(\omega)\mathbb{P}(d\omega)$. Essentially, we "cut out" the $\varepsilon(\omega)$-neighborhood of $x_{2d+1}(\omega)$ from the conditional distribution $\mathbb{P}_s(\omega, \cdot)$ on the set Θ in a measurable way and then compensate for it by renormalizing the density. Note also that $\mathbb{P}(Z(\omega) = 0) > 0$, therefore $\mathbb{Q} \not\sim \mathbb{P}$. ☐

Denote by $\mathcal{Q}^*$ the set of all probability measures $\mathbb{Q}$ absolutely continuous with respect to $\mathbb{P}$ under which X is a supermartingale, and let $Z_{\mathcal{Q}^*}$ be the set of densities of measures $\mathbb{Q} \in \mathcal{Q}^*$ with respect to $\mathbb{P}$.

Proposition 5.3: *The set $Z_{\mathcal{Q}^*}$ is weakly closed in $L^1(\mathbb{P})$.*

Proof. Note first that $L^1(\mathbb{P})$ is a locally convex space, therefore convexity of $Z_{\mathcal{Q}^*}$ implies that it is sufficient to establish its strong closedness in order to prove that it is weakly closed (see, for instance, Akilov and Kantorovich [1984]).

Consider a sequence $\{Z_n\}_{n\in\mathbb{N}}$ of densities from $Z_{\mathcal{Q}^*}$ converging to Z in $L^1(\mathbb{P})$ as $n \to \infty$. Let $\mathbb{Q}_n(d\omega) = Z_n(\omega)\mathbb{P}(d\omega)$, $\mathbb{Q}(d\omega) = Z(\omega)\mathbb{P}(d\omega)$. If $Z_n \to Z$ in $L^1(\mathbb{P})$, then there exists a subsequence $\{Z_{n_k}\}_{k\in\mathbb{N}}$ that converges $\mathbb{P}$-almost surely. Choose two moments of time $s, t \in \mathcal{T}$, $s \leqslant t$, then, taking into account the non-negativity of X and the supermartingale property, we can apply the conditional Fatou's lemma:

$$\begin{aligned}\mathbb{E}(ZX_s \mid \mathcal{F}_t) &= \mathbb{E}(\liminf_{k\to\infty} Z_{n_k} X_s \mid \mathcal{F}_t)\\ &\leqslant \liminf_{k\to\infty} \mathbb{E}(Z_{n_k} X_s \mid \mathcal{F}_t)\\ &\leqslant \liminf_{k\to\infty} X_t \mathbb{E}(Z_{n_k} \mid \mathcal{F}_t)\\ &= X_t \mathbb{E}(Z \mid \mathcal{F}_t). \quad (\mathbb{P}\text{-a.s.})\end{aligned} \tag{5.17}$$

Therefore,

$$\mathbb{E}_{\mathbb{Q}}(X_s \mid \mathcal{F}_t) = \frac{\mathbb{E}(ZX_s \mid \mathcal{F}_t)}{\mathbb{E}(Z \mid \mathcal{F}_t)} \leqslant X_t, \quad (\mathbb{Q}\text{-a.s.}) \tag{5.18}$$

which ensures that X is a $\mathbb{Q}$-supermartingale and $Z \in Z_{\mathcal{Q}^*}$. Note that the denominator in (5.18) is $\mathbb{Q}$-almost surely non-zero and the ratio is well-defined since $\mathbf{1}_{\{\mathbb{E}(Z \mid \mathcal{F}_t)=0\}} \leqslant \mathbf{1}_{\{Z=0\}}$. □

Remark 5.2: Note that the proof of Proposition 5.3 is applicable to the continuous-time case as is.

We shall say that *the martingale property is preserved* for a sequence of densities $Z_n \in Z_{\mathcal{P}^*}$ which converges to a limiting density Z in $L^1(\mathbb{P})$ if X is a martingale under the limiting measure $\mathbb{Q}$, with $\mathbb{Q}(d\omega) = Z(\omega)\,\mathbb{P}(d\omega)$.

Theorem 5.2: *In a discrete-time arbitrage-free market, the necessary and sufficient condition for the martingale property to be preserved for every*

$L^1(\mathbb{P})$-converging sequence of densities $Z_n \in Z_{\mathcal{P}^*}$ is the boundedness (i.e. compactness) of the topological support of $\mathbb{P}_t(\omega, \cdot)$ for all $t \in \mathcal{T}_+$ and $\mathbb{P}$-a.s. for all $\omega \in \Omega$.

Proof. $\boxed{\Rightarrow}$ If there exists a moment of time $t = s \in \mathcal{T}_+$ and an $\mathcal{F}_{s-1}$-measurable set Θ with $\mathbb{P}(\Theta) > 0$ such that for each $\omega \in \Theta$, $\operatorname{supp}(\mathbb{P}_s(\omega, \cdot))$ is not bounded, then the market is incomplete. By the no-arbitrage condition and the Steinitz theorem, there exists a subset $\mathcal{X}(\omega) \subset \operatorname{supp}(\mathbb{P}_s(\omega, \cdot))$ such that $|\mathcal{X}(\omega)| \leqslant 2d$ and $0 \in \operatorname{ri}(\overline{\operatorname{co}}(\mathcal{X}(\omega)))$. For the sake of simplicity, we can safely assume that the set $\mathcal{X}(\omega)$ always contains exactly $2d$ points (indeed, for each $\omega \in \Theta$ there is an infinite amount of extra points to choose, and adding points to $\mathcal{X}(\omega)$ cannot shrink its convex hull). In addition, one can also choose an $\mathcal{F}_{s-1}$-measurable unbounded sequence $z_n(\omega) \in \operatorname{supp}(\mathbb{P}_s(\omega, \cdot))$, $n \in \mathbb{N}$, such that $||z_n(\omega)|| \to +\infty$ as $n \to \infty$, $z_n(\omega)/||z_n(\omega)||$ converges, $\{z_n(\omega)\} \cap \mathcal{X}(\omega) = \emptyset$ and $||z_n(\omega)|| > 2 \bigvee_{x \in \mathcal{X}(\omega)} ||x||$, for all $n \in \mathbb{N}$.

The set $\mathcal{X}(\omega)$ and the sequence $\{z_n(\omega)\}$ can be selected simultaneously for each ω in a jointly measurable way. Consider a countable product $(G, \mathcal{G}) = \otimes_{n=1}^{\infty} (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, where $\mathcal{B}(\mathbb{R}^d)$ is a Borel σ-algebra on $\mathbb{R}^d$, with elements $g = (g_1, g_2, \ldots)$, $g_i \in \mathbb{R}^d$, $i \in \mathbb{N}$. Let A be a subset of $\Omega \times G$ consisting of all such pairs (ω, g) that $\omega \in \Theta$, $g_i \neq g_j$ for $i \neq j$, $g_i \in \operatorname{supp}(\mathbb{P}_s(\omega, \cdot))$ for all i, $0 \in \operatorname{ri}(\overline{\operatorname{co}}(\{g_1, g_2, \ldots g_{2d}\}))$, $||g_i|| > 2 \bigvee_{j=1}^{2d} ||g_j(\omega)||$ for $i > 2d$, $||g_i|| \to \infty$ and $g_i/||g_i||$ converges as $i \to \infty$. Since a countable product of Polish spaces is Polish, we can apply Lemma 5.3, which ensures that there exists a G-valued $\mathcal{F}_{s-1}$-measurable infinite-dimensional selector $Y(\omega) = (y_1(\omega), y_2(\omega), \ldots)$, and for each $\omega \in \Theta$ we can take $\mathcal{X}(\omega) := \{y_1(\omega), y_2(\omega), \ldots y_{2d}(\omega)\}$ and $z_n(\omega) := y_{2d+n}(\omega)$ for $n \in \mathbb{N}$.

For $\omega \in \Theta$, consider an $\mathcal{F}_{s-1}$-measurable positive random variable $\varepsilon(\omega)$, then $\mathbb{P}_s(\omega, B(x_i(\omega), \varepsilon(\omega))) > 0$ and $\mathbb{P}_s(\omega, B(z_n(\omega), \varepsilon(\omega))) > 0$, for $i = 1, 2, \ldots 2d$ and $n \in \mathbb{N}$. Define

$$\tilde{x}_i(\omega, \varepsilon(\omega)) = \frac{1}{\mathbb{P}_s(\omega, B(x_i(\omega), \varepsilon(\omega)))} \int_{B(x_i(\omega), \varepsilon(\omega))} x \mathbb{P}_s(\omega, dx), \qquad (5.19)$$

$$\tilde{z}_n(\omega, \varepsilon(\omega)) = \frac{1}{\mathbb{P}_s(\omega, B(z_n(\omega), \varepsilon(\omega)))} \int_{B(z_n(\omega), \varepsilon(\omega))} x \mathbb{P}_s(\omega, dx), \qquad (5.20)$$

then $\tilde{x}_i(\omega, \varepsilon(\omega)) \in B(x_i(\omega), \varepsilon(\omega))$ and $\tilde{z}_n(\omega, \varepsilon(\omega)) \in B(z_n(\omega), \varepsilon(\omega))$. Denote $\tilde{\mathcal{X}}(\omega, \varepsilon) = \{\tilde{x}_1(\omega, \varepsilon), \tilde{x}_2(\omega, \varepsilon), \ldots \tilde{x}_{2d}(\omega, \varepsilon)\}$.

Let

$$\tilde{\varepsilon}\,(\omega) = \frac{1}{3}\left(\Delta(\mathcal{X}(\omega)) \wedge \bigwedge_{i \neq j} ||x_i(\omega) - x_j(\omega)||\right), \qquad (5.21)$$

where $\Delta(A)$ is the distance from the origin to the boundary of the closure of the convex hull of set A:

$$\Delta(A) = \sup\left\{\gamma > 0 \ : \ B(0, \gamma) \subseteq \overline{\mathrm{co}}(A)\right\}. \qquad (5.22)$$

Then, since $\bigwedge_{i \neq j} ||x_i(\omega) - x_j(\omega)|| \leqslant 2 \bigvee_{i=1}^{2d} ||x_i(\omega)||$, for each $n \in \mathbb{N}$ the neighborhoods $B(x_i(\omega), \tilde{\varepsilon}(\omega))$ are mutually disjoint for all $i = 1, 2, \ldots 2d$ and don't intersect with $B(z_n(\omega), \tilde{\varepsilon}(\omega))$. Moreover, it is easy to verify that our choice of (5.21) ensures that $B(0, \tilde{\varepsilon}(\omega)) \subset \mathrm{ri}(\overline{\mathrm{co}}(\tilde{\mathcal{X}}(\omega, \tilde{\varepsilon}(\omega))))$.

Take

$$\pi_n(\omega) = \frac{\tilde{\varepsilon}(\omega)}{\tilde{\varepsilon}(\omega) + ||\tilde{z}_n(\omega, \tilde{\varepsilon}(\omega))||} \qquad (5.23)$$

and

$$y_n(\omega) = \tilde{z}_n(\omega, \tilde{\varepsilon}(\omega)) \frac{\pi_n(\omega)}{1 - \pi_n(\omega)} = \tilde{\varepsilon}(\omega) \frac{\tilde{z}_n(\omega, \tilde{\varepsilon}(\omega))}{||\tilde{z}_n(\omega, \tilde{\varepsilon}(\omega))||}, \qquad (5.24)$$

then $||y_n(\omega)||= \tilde{\varepsilon}(\omega)$, hence $-y_n(\omega) \in \mathrm{ri}(\overline{\mathrm{co}}(\tilde{\mathcal{X}}(\omega, \tilde{\varepsilon}(\omega))))$. Therefore, for each $n \in \mathbb{N}$ there exists an $\mathcal{F}_{s-1}$-measurable set of weights[4] $\{p_i^{(n)}(\omega)\}$ such that

$$\sum_{i=1}^{2d} p_i^{(n)}(\omega)\tilde{x}_i(\omega, \tilde{\varepsilon}(\omega)) = -y_n(\omega), \tag{5.25}$$

or

$$\sum_{i=1}^{2d}(1 - \pi_n(\omega))p_i^{(n)}(\omega)\tilde{x}_i(\omega, \tilde{\varepsilon}(\omega)) + \pi_n(\omega)\tilde{z}_n(\omega, \tilde{\varepsilon}(\omega)) = 0, \tag{5.26}$$

with $\sum_{i=1}^{2d} p_i^{(n)}(\omega) = 1$ and $p_i^{(n)}(\omega) \geqslant 0$ for $i = 1, 2, \dots 2d$.

Since $p_i^{(n)}(\omega) \in [0, 1]$, we can select a subsequence $p_i^{(n_k)}(\omega)$ such that $p_i^{(n_k)}(\omega)$ converges weakly (due to uniform boundedness) to $p_i^*(\omega) \in [0, 1]$, for all $i = 1, 2, \dots 2d$. Consider the sequence of random variables

$$\begin{aligned} Z^{(k)}(\omega) = \mathbf{1}_{\Theta^c}(\omega) + \mathbf{1}_{\Theta}(\omega) \sum_{i=1}^{2d}(1 - \pi_{n_k}(\omega))p_i^{(n_k)}(\omega)\frac{\mathbf{1}_{B(x_i(\omega),\tilde{\varepsilon}(\omega))}(\Delta X_s(\omega))}{\mathbb{P}_s(\omega, B(x_i(\omega), \tilde{\varepsilon}(\omega)))} \\ +\mathbf{1}_{\Theta}(\omega)\pi_{n_k}(\omega)\frac{\mathbf{1}_{B(z_{n_k}(\omega),\tilde{\varepsilon}(\omega))}(\Delta X_s(\omega))}{\mathbb{P}_s(\omega, B(z_{n_k}(\omega), \tilde{\varepsilon}(\omega)))}, \end{aligned} \tag{5.27}$$

then $Z^{(k)}(\omega) \geqslant 0$ and $\int_\Omega Z^{(k)}(\omega)\mathbb{P}(d\omega) = 1$. Therefore, $Z^{(k)}$ are proper densities with respect to $\mathbb{P}$ and thus define a sequence of probability measures $\mathbb{Q}^{(k)}$ with $\mathbb{Q}^{(k)}(d\omega) = Z^{(k)}(\omega)\mathbb{P}(d\omega)$. Moreover, equality (5.26) ensures that $\mathbb{E}(Z^{(k)}\Delta X_s \mid \mathcal{F}_{s-1}) = 0$, hence every $\mathbb{Q}^{(k)}$ is a martingale measure. Besides, since $\pi_n(\omega) \to 0$ as $n \to \infty$, $Z^{(k)}(\omega)$ converges weakly to $Z^*(\omega)$, where

$$Z^*(\omega) = \mathbf{1}_{\Theta^c}(\omega) + \mathbf{1}_{\Theta}(\omega) \sum_{i=1}^{2d} p_i^*(\omega)\frac{\mathbf{1}_{B(x_i(\omega),\tilde{\varepsilon}(\omega))}(\Delta X_s(\omega))}{\mathbb{P}_s(\omega, B(x_i(\omega), \tilde{\varepsilon}(\omega)))} \tag{5.28}$$

[4]This set of weights may or may not be unique. In the latter case, a specific set of weights can be chosen in a measurable way by selecting the one that minimizes a strictly convex function of the weights, e.g. the sum of squares.

is a density for the limiting measure $\mathbb{Q}^*$.

Note that measures $\mathbb{Q}^{(k)}$ are absolutely continuous but not necessarily equivalent with respect to $\mathbb{P}$. To remedy this fact, consider a sequence of probability measures $\tilde{\mathbb{P}}^{(n)}$ defined as mixtures $\tilde{\mathbb{P}}^{(k)} = \frac{1}{2}\mathbb{P} + \frac{1}{2}\mathbb{Q}^{(k)}$, $k \in \mathbb{N}$. Clearly, $\tilde{\mathbb{P}}^{(k)} \sim \mathbb{P}$ and X is a $\mathbb{P}^{(k)}$-martingale for each k, therefore $\mathbb{P}^{(k)} \in \mathcal{P}^*$ and $Z_{\mathbb{P}^{(k)}} \in Z_{\mathcal{P}^*}$. According to the argument above, the sequence of measures $\mathbb{P}^{(k)}$ converges to some measure $\mathbb{P}^* {=} \frac{1}{2}\mathbb{P} + \frac{1}{2}\mathbb{Q}^*$ with density $Z_{\mathbb{P}^*}(\omega) = \frac{1}{2} + \frac{1}{2}Z^*(\omega)$.

However, due to our choice of $z_n(\omega)$ via construction of the measurable selector, $y_n(\omega) \to y^*(\omega)$, hence

$$\mathbb{E}(Z_{\mathbb{P}^*}\Delta X_s \mid \mathcal{F}_{s-1}) = -\frac{y^*(\omega)}{2}, \tag{5.29}$$

Note that all coordinates of vector $y^*(\omega)$ are non-negative since $y_n(\omega)$ is defined by (5.24), $||z_n(\omega)|| \to \infty$ and $z_n(\omega)$ is bounded from below $-X_{t-1}(\omega)$ due to the non-negativity of the price process. Moreover, at least one coordinate of $y^*(\omega)$ must be positive since $||y^*(\omega)|| = \tilde{\varepsilon}(\omega)$. Therefore, X is a supermartingale but not a martingale under $\mathbb{P}^*$, which implies that boundedness of the topological support of conditional distributions of ΔX_t is indeed a necessary condition.

$\boxed{\Leftarrow}$ Consider a sequence of probability measures $\mathbb{Q}^{(k)}$ defined by $\mathbb{Q}^{(k)}(d\omega) = Z^{(k)}(\omega)\mathbb{P}(d\omega)$, for $k \in \mathbb{N}$. Then $Z^{(k)} \to Z$ in $L^1(\mathbb{P})$ and, since $\mathbb{Q} \ll \mathbb{P}$, we have for all $k \in \mathbb{N}$ and $s \in \mathcal{T}_+$

$$\mathbb{E}(Z^{(k)} \mid \mathcal{F}_{s-1}) \cdot \mathbb{E}_{\mathbb{Q}^{(k)}}(\Delta X_s \mid \mathcal{F}_{s-1}) = \mathbb{E}(Z^{(k)}\Delta X_s \mid \mathcal{F}_{s-1}), \quad \mathbb{P}\text{-a.s.}, \tag{5.30}$$

$$\mathbb{E}(Z \mid \mathcal{F}_{s-1}) \cdot \mathbb{E}_{\mathbb{Q}}(\Delta X_s \mid \mathcal{F}_{s-1}) = \mathbb{E}(Z\Delta X_s \mid \mathcal{F}_{s-1}), \quad \mathbb{P}\text{-a.s.} \tag{5.31}$$

In particular, since X is a martingale under $\mathbb{Q}^{(k)}$,

$$\mathbb{E}(Z^{(k)}\Delta X_s \mid \mathcal{F}_{s-1}) = 0, \quad \mathbb{P}\text{-a.s.} \tag{5.32}$$

Boundedness of the support sets of conditional distributions of ΔX_s means that there exists an $\mathcal{F}_{s-1}$-measurable non-negative random variable Y_{s-1} such that $||\Delta X_t|| \leqslant Y_{s-1}$, $\mathbb{P}$-a.s. In conjunction with (5.32) this implies that

$$\begin{aligned} ||\mathbb{E}(Z\Delta X_s \mid \mathcal{F}_{s-1})|| &= ||\mathbb{E}((Z - Z^{(k)})\Delta X_s \mid \mathcal{F}_{s-1})|| \\ &\leqslant \mathbb{E}(|Z^{(k)} - Z| \cdot ||\Delta X_s|| \mid \mathcal{F}_{s-1}) \\ &\leqslant Y_{s-1} \cdot \mathbb{E}(|Z^{(k)} - Z| \mid \mathcal{F}_{s-1}). \end{aligned} \tag{5.33}$$

Evidently, $\mathbb{E}(|Z^{(k)} - Z| \mid \mathcal{F}_{s-1}) \to 0$ in $L^1(\mathbb{P})$, so we can choose such subsequence $Z^{(k_j)}$ that this conditional expectation converges $\mathbb{P}$-almost surely. Then, by applying (5.33) to $Z^{(k_j)}$, we have

$$\mathbb{E}(Z\Delta X_s \mid \mathcal{F}_{s-1}) = 0, \quad \mathbb{P}\text{-a.s.} \tag{5.34}$$

Note also that $\mathbb{E}(Z \mid \mathcal{F}_{s-1}) > 0$, $\mathbb{Q}$-a.s., therefore (5.31) along with (5.34) ensure that $\mathbb{E}_{\mathbb{Q}}(\Delta X_s \mid \mathcal{F}_{s-1}) = 0$, $\mathbb{Q}$-a.s., which concludes the proof. □

Theorem 5.3: *In a discrete-time arbitrage-free market, if there exists a moment of time $t = s \in \mathcal{T}_+$ and an $\mathcal{F}_{s-1}$-measurable set Θ of positive probability such that for each $\omega \in \Theta$ the support of $\mathbb{P}_s(\omega, \cdot)$ contains a limit point, then the set $Z_{\mathcal{P}^*}$ is not weakly relatively compact.*

Proof. If there exists a moment of time $t = s \in \mathcal{T}_+$ and an $\mathcal{F}_{s-1}$-measurable set Θ with $\mathbb{P}(\Theta) > 0$ such that $\text{supp}(\mathbb{P}_s(\omega, \cdot))$ contains a limit point for each $\omega \in \Theta$, then the market is incomplete. The no-arbitrage argument in conjunction with the Steinitz theorem implies that there exists a subset $\mathcal{X}(\omega) \subset \operatorname{supp}(\mathbb{P}_s(\omega, \cdot))$ such that $|\mathcal{X}(\omega)| \leqslant 2d$ and $0 \in \operatorname{ri}(\overline{\operatorname{co}}(\mathcal{X}(\omega)))$. Same as before, we shall assume without loss of generality that $|\mathcal{X}(\omega)| = 2d$. Moreover, we can assume that $z(\omega) \notin \mathcal{X}(\omega)$. If that were not true, we could use the fact that $z(\omega)$ is a limit point and choose a point $\tilde{z}(\omega) \in \operatorname{supp}(\mathbb{P}_s(\omega, \cdot))$, $\tilde{z}(\omega) \neq z(\omega)$, such that $||\tilde{z}(\omega) - z(\omega)|| < \Delta(\mathcal{X}(\omega))$, where $\Delta(\mathcal{X}(\omega))$ is the distance from the origin to the closure of the convex hull of $\mathcal{X}(\omega)$. We can then replace $\mathcal{X}(\omega)$ with the set $\tilde{\mathcal{X}}(\omega) := \mathcal{X}(\omega) \cup \{\tilde{z}(\omega)\} \backslash \{z(\omega)\}$ which consists of $2d$ points from $\operatorname{supp}(\mathbb{P}_s(\omega, \cdot))$ and satisfies $0 \in \operatorname{ri}(\overline{\operatorname{co}}(\tilde{\mathcal{X}}(\omega)))$, $z(\omega) \notin \tilde{\mathcal{X}}(\omega)$.

The set $\mathcal{X}(\omega)$ and the limit point $z(\omega)$ can be chosen in a jointly measurable fashion by using the similar arguments that we used previously and taking into account the fact that $\mathbb{R}^d$ is a second-countable space. Let $G = \mathbb{R}^{(2d+1)\times d}$ with elements $g = (g_1, g_2, \ldots g_{2d+1})$, $g_i \in \mathbb{R}^d$. Let A be a subset of $\Omega \times G$ consisting of all such pairs (ω, g) that $\omega \in \Theta$, $g_i \neq g_j$ for $i \neq j$, $g_i \in \operatorname{supp}(\mathbb{P}_s(\omega, \cdot))$ for $i = 1, 2, \ldots 2d+1$, $0 \in \operatorname{ri}(\overline{\operatorname{co}}(\{g_1, g_2, \ldots g_{2d}\}))$ and g_{2d+1} is a limit point. By Lemma 5.3, there exists a G-valued $\mathcal{F}_{s-1}$-measurable selector $Y(\omega) = (y_1(\omega), y_2(\omega), \ldots y_{2d+1}(\omega))$, and for each $\omega \in \Theta$ we can assign $\mathcal{X}(\omega) := \{y_1(\omega), y_2(\omega), \ldots y_{2d}(\omega)\}$ and $z(\omega) := y_{2d+1}(\omega)$. Constructively, we can first select an arbitrary limit point $z(\omega)$ and then select $\mathcal{X}(\omega)$ so that it satisfies the above conditions. This would be always possible since excluding a limit point from a set does not alter the closure of its convex hull.

For $\omega \in \Theta$, consider $\delta(\omega)$-neighborhoods of $x_i(\omega)$ and $\varepsilon_n(\omega)$-neighborhood of $z(\omega)$ with its center cut out, for some $\mathcal{F}_{s-1}$-measurable $\delta(\omega)$ and $\varepsilon_n(\omega) \to 0$. Since $\mathbb{P}(B(x_i(\omega), \delta(\omega))) > 0$ and $\mathbb{P}(B(z(\omega), \varepsilon_n(\omega)) \backslash \{z(\omega)) > 0$, for $i =$

$1, 2, \dots 2d$, we can define

$$\tilde{x}_i(\omega, \delta(\omega)) = \frac{1}{\mathbb{P}(B(x_i(\omega), \delta(\omega)))} \int_{B(x_i(\omega),\delta(\omega))} x\mathbb{P}_s(\omega, dx), \tag{5.35}$$

$$\tilde{z}(\omega, \varepsilon_n(\omega)) = \frac{1}{\mathbb{P}(B(z(\omega), \varepsilon_n(\omega))\backslash\{z(\omega)\})} \int_{B(z(\omega),\varepsilon_n(\omega))\backslash\{z(\omega)} x\mathbb{P}_s(\omega, dx). \tag{5.36}$$

Note that $\tilde{x}_i(\omega, \delta(\omega)) \in B(x_i(\omega), \delta(\omega))$ and $\tilde{z}(\omega, \varepsilon_n(\omega)) \in B(z(\omega), \varepsilon_n(\omega))$, and denote $\tilde{\mathcal{X}}(\omega, \delta(\omega)) = \{\tilde{x}_1(\omega, \delta(\omega)), \tilde{x}_2(\omega, \delta(\omega)), \dots \tilde{x}_{d+1}(\omega, \delta(\omega))\}$. We intend to show that one can choose such $\mathcal{F}_{s-1}$-measurable random variables $\delta(\omega) > 0$, $\varepsilon_n(\omega) > 0$ and $\pi(\omega) \in (0, 1)$ that all of the above neighborhoods are mutually disjoint for each $n \in \mathbb{N}$ and vector

$$y_n(\omega) = \tilde{z}(\omega, \varepsilon_n(\omega))\frac{\pi(\omega)}{1 - \pi(\omega)}$$

is sufficiently small in the sense that $-y_n(\omega) \in \mathrm{ri}(\overline{\mathrm{co}}(\tilde{\mathcal{X}}(\omega, \delta(\omega))))$, for all $\omega \in \Theta$ and $n \in \mathbb{N}$. Consider the distance $\Delta(A)$ from the origin to the boundary of the closure of the convex hull of set A, then $-y_n(\omega) \in \mathrm{ri}(\overline{\mathrm{co}}(\tilde{\mathcal{X}}(\omega, \delta(\omega))))$ if

$$\Delta(\tilde{\mathcal{X}}(\omega, \delta(\omega))) > ||y_n(\omega)||, \tag{5.37}$$

which in its turn is satisfied if

$$\Delta(\tilde{\mathcal{X}}(\omega, \delta(\omega))) > (||z(\omega)||+\varepsilon_n(\omega))\, \frac{\pi(\omega)}{1 - \pi(\omega)}, \tag{5.38}$$

or

$$\Delta(\mathcal{X}(\omega)) > (||z(\omega)||+\varepsilon_n(\omega))\, \frac{\pi(\omega)}{1 - \pi(\omega)} + \delta(\omega). \tag{5.39}$$

Our choice of the measurable selector guarantees that $\Delta(\mathcal{X}(\omega)) > 0$. Take

$$\tilde{\delta}(\omega) = \frac{1}{3}\left(\Delta(\mathcal{X}(\omega)) \wedge \bigwedge_{i \neq j} ||x_i(\omega) - x_j(\omega)|| \wedge \bigwedge_{i} ||z(\omega) - x_i(\omega)||\right) \tag{5.40}$$

and choose such $\tilde{\varepsilon}_n(\omega)$ that $\tilde{\varepsilon}_n(\omega) < \tilde{\delta}(\omega)$, then the neighborhoods $B(x_i(\omega), \tilde{\delta}(\omega))$, for $i = 1, 2, \ldots 2d$ and $B(z(\omega), \tilde{\varepsilon}_n(\omega))$ are mutually disjoint for each $n \in \mathbb{N}$. Moreover, by assigning

$$\tilde{\pi}(\omega) = \frac{\Delta(\mathcal{X}(\omega)) - \tilde{\delta}(\omega)}{\Delta(\mathcal{X}(\omega)) + ||z(\omega)||}, \tag{5.41}$$

we have $\tilde{\pi}(\omega) \in (0, 1)$ and, since $\tilde{\varepsilon}_n(\omega) < \tilde{\delta}(\omega)$, (5.39) is satisfied, hence

$$\tilde{y}_n(\omega) = \tilde{z}(\omega, \tilde{\varepsilon}_n(\omega)) \frac{\tilde{\pi}(\omega)}{1 - \tilde{\pi}(\omega)} \in \mathrm{ri}(\overline{\mathrm{co}}(\tilde{\mathcal{X}}(\omega, \tilde{\delta}(\omega)))). \tag{5.42}$$

Therefore, for each $\omega \in \Theta$ and $n \in \mathbb{N}$ there exists an $\mathcal{F}_{s-1}$-measurable set of weights $\{p_i^{(n)}(\omega)\}$ (which may not be unique but can still be selected measurably, cf. the proof of Theorem 5.2) such that

$$-\tilde{y}_n(\omega) = \sum_{i=1}^{d+1} p_i^{(n)}(\omega) \tilde{x}_i(\omega, \tilde{\delta}(\omega)) \tag{5.43}$$

or, equivalently,

$$\sum_{i=1}^{2d} (1 - \tilde{\pi}(\omega)) p_i^{(n)}(\omega) \tilde{x}_i(\omega, \tilde{\delta}(\omega)) + \tilde{\pi}(\omega) \tilde{z}(\omega, \tilde{\varepsilon}_n(\omega)) = 0, \tag{5.44}$$

with $\sum_{i=1}^{2d} p_i^{(n)}(\omega) = 1$ and $p_i^{(n)}(\omega) \geqslant 0$ for $i = 1, 2, \ldots 2d$.

Consider the sequence of random variables

$$Z^{(n)}(\omega) = \mathbf{1}_{\Theta^c}(\omega) \ +\mathbf{1}_{\Theta}(\omega) \left(\sum_{i=1}^{2d} (1-\tilde{\pi}(\omega)) p_i^{(n)}(\omega) \frac{\mathbf{1}_{B(x_i(\omega),\tilde{\delta}(\omega))}(\Delta X_s(\omega))}{\mathbb{P}_s\left(\omega, B(x_i(\omega),\tilde{\delta}(\omega))\right)} \right.$$
$$\left. + \frac{\mathbf{1}_{B(z(\omega),\tilde{\varepsilon}_n(\omega))\setminus\{z(\omega)\}}(\Delta X_s(\omega))}{\mathbb{P}_s\left(\omega, B(z(\omega),\tilde{\varepsilon}_n(\omega))\setminus\{z(\omega)\}\right)} \right), \tag{5.45}$$

then $Z^{(n)}(\omega) \geqslant 0$ and $\int_\Omega Z^{(n)}(\omega)\mathbb{P}(d\omega) = 1$, hence $Z^{(n)}(\omega)$ are proper densities with respect to $\mathbb{P}$ and thus define the associated sequence of measures $\mathbb{Q}^{(n)}$ with $\mathbb{Q}^{(n)}(d\omega) = Z^{(n)}(\omega)\mathbb{P}(d\omega)$. By construction, (5.43) implies that $\mathbb{E}(Z^{(n)}\Delta X_s \mid \mathcal{F}_{s-1}) = 0$, hence X is a martingale under $\mathbb{Q}^{(n)}$ for each $n \in \mathbb{N}$.

Consider the sequence of probability measures $\tilde{\mathbb{P}}^{(n)}$ specified as mixtures $\tilde{\mathbb{P}}^{(n)} = \frac{1}{2}\mathbb{P} + \frac{1}{2}\mathbb{Q}^{(n)}$, for $n \in \mathbb{N}$. Since $\mathbb{Q}^{(n)} \ll \mathbb{P}$ and both $\mathbb{P}$ and $\mathbb{Q}^{(n)}$ are martingale measures, $\tilde{\mathbb{P}}^{(n)} \sim \mathbb{P}$ and X is a martingale under $\mathbb{P}^{(n)}$, therefore $\mathbb{P}^{(n)} \in \mathcal{P}^*$ and $Z_{\mathbb{P}^{(n)}} \in Z_{\mathcal{P}^*}$. We shall now show that the sequence $Z_{\mathbb{P}^{(n)}}$ is not uniformly integrable.

Notice first that $Z_{\mathbb{P}^{(n)}} = \frac{1}{2} + \frac{1}{2}Z^{(n)}$ is uniformly integrable if and only if $Z^{(n)}$ is uniformly integrable. For simplicity of further notation, for all $\omega \in \Theta$ and $n \in \mathbb{N}$ let

$$\varphi_s^{(n)}(\omega, x) = \frac{\mathbf{1}_{B(z(\omega),\tilde{\varepsilon}_n(\omega))\setminus\{z(\omega)\}}(x)}{\mathbb{P}_s\left(\omega, B(z(\omega),\tilde{\varepsilon}_n(\omega))\setminus\{z(\omega)\}\right)}. \tag{5.46}$$

As a function of x, $\varphi_n^{(n)}(\omega, x)$ can be interpreted as a conditional density of ΔX_s given $\mathcal{F}_{s-1}$. Define an $\mathcal{F}_s$-measurable random variable

$$\tilde{\varphi}_n(\omega) = \varphi_s^{(n)}(\omega, \Delta X_s(\omega)), \tag{5.47}$$

then

$$Z^{(n)} \geqslant Y^{(n)} = \tilde{\pi}(\omega)\mathbf{1}_{\Theta}(\omega)\tilde{\varphi}_n(\omega) \geqslant 0, \tag{5.48}$$

hence the uniform integrability of $Z^{(n)}$ implies the uniform integrability of $Y^{(n)}$. For an arbitrary C, consider the following expectation:

$$\mathbb{E}\left(Y^{(n)}\mathbf{1}_{\{Y^{(n)}\geqslant C\}}\right) = \mathbb{E}\left(\tilde{\pi}(\omega)\mathbf{1}_{\Theta}(\omega)\tilde{\varphi}_n(\omega)\mathbf{1}_{\{\tilde{\pi}(\omega)\tilde{\varphi}_n(\omega)\geqslant C\}}\right). \tag{5.49}$$

Note that $\tilde{\pi}(\omega)$ and $\mathbf{1}_{\Theta}(\omega)$ are $\mathcal{F}_{s-1}$-measurable, hence if we consider an arbitrary $\mathcal{F}_{s-1}$-measurable extension of $\tilde{\pi}(\omega)$ on Θ^c and an $\mathcal{F}_s$-measurable extension of $\tilde{\varphi}_n(\omega)$ on Θ^c, (5.49) can be expressed as

$$\mathbb{E}\left(Y^{(n)}\mathbf{1}_{\{Y^{(n)}\geqslant C\}}\right) = \mathbb{E}(\tilde{\pi}(\omega)\mathbf{1}_{\Theta}(\omega)\mathbf{1}_{\{\tilde{\pi}(\omega)\tilde{\varphi}_n(\omega)\geqslant C\}}\mathbb{E}(\tilde{\varphi}_n(\omega) \mid \mathcal{F}_{s-1})). \tag{5.50}$$

Since $\tilde{\varphi}_n(\omega) \to \infty$ a.e. on Θ,

$$\mathbf{1}_{\Theta}(\omega)\mathbf{1}_{\{\tilde{\pi}(\omega)\tilde{\varphi}_n(\omega)\geqslant C\}} \to \mathbf{1}_{\Theta}(\omega), \quad \text{a.s.}, \tag{5.51}$$

and therefore

$$\lim_{n\to\infty} \mathbb{E}\left(Y^{(n)}\mathbf{1}_{\{Y^{(n)}\geqslant C\}}\right) = \mathbb{E}(\tilde{\pi}(\omega)\mathbf{1}_{\Theta}(\omega)) > 0, \tag{5.52}$$

The right-hand side in (5.52) does not depend on C, and as such it does not become infinitely small as $C \to \infty$. We conclude that the sequence $Y^{(n)}$ is not uniformly integrable, therefore the set of densities $Z_{\mathcal{P}^*}$ is not uniformly integrable, hence by the Dunford-Pettis theorem it is not weakly relatively compact. □

Remark 5.3: If a subset of $\mathbb{R}^d$ consists of isolated points only, it is countable at most. Therefore, according to Theorem 5.3, the necessary condition for relative compactness of densities in $L^1(\mathbb{P})$ is the discreteness of the joint distribution of $X = (X_{0,}X_1, \dots X_T)$. In this case, if $\mathcal{F}_t$ is generated by the family of random variables $\{X_s\}_{0\leqslant s\leqslant t}$ and $\mathcal{F} = \mathcal{F}_T$, we can choose Ω to be countable at most without loss of generality, with $\mathbb{P}(\{\omega\}) > 0$ for all $\omega \in \Omega$. Note that in a discrete case the notions of weak and strong convergence in $L^1(\mathbb{P})$ are equivalent (see e.g. Dunford and Schwartz [1988]).

Let $\mathcal{R}^*$ a set of measures $\mathbb{Q}$ such that $\mathbb{Q} \ll \mathbb{P}$ and X is a supermartingale under $\mathbb{Q}$. Denote by $Z_{\mathcal{R}^*}$ the set of densities of measures from $\mathcal{R}^*$ with respect to $\mathbb{P}$.

Proposition 5.4: *Consider a discrete-time arbitrage-free incomplete market. Let the support of the joint distribution of $X = (X_0, X_1, \dots X_T)$ be a set of isolated points with respect to $\mathbb{P}$, $\mathcal{F}_t$ is generated by the family of random variables $\{X_s\}_{0\leqslant s\leqslant t}$, and $\mathcal{F} = \mathcal{F}_T$. Then the set $Z_{\mathcal{R}^*}$ is relatively compact in $L^1(\mathbb{P})$.*

Proof. Let us note first that it is sufficient to consider the case when $\mathcal{R}^*$ is non-empty. Notice also that $\mathbb{Q} \ll \mathbb{P}$ implies $\mathrm{supp}(X \circ \mathbb{Q}) \subset \mathrm{supp}(X \circ \mathbb{P})$. As has been pointed out above in Remark 5.3, $\mathrm{supp}(X \circ \mathbb{P})$ is countable at most, so we index it as $\mathrm{supp}(X \circ \mathbb{P}) = \{x_1, \dots x_i, \dots\}$. Moreover, we can choose Ω to be countable at most and equinumerous to $\mathrm{supp}(X \circ \mathbb{P})$. In particular, we can take $\Omega = \mathrm{supp}(X \circ \mathbb{P})$ without loss of generality.

Since $\mathcal{F}_0 = \{\emptyset, \Omega\}$, the initial value of the discounted price process $X_0 = x^* \geqslant 0$ is fixed and is the same under all measures absolutely continuous with respect to $\mathbb{P}$. Choose an arbitrary $\varepsilon > 0$ and $a \geqslant 3x^*/\varepsilon$. Since X is a non-negative supermartingale under any $\mathbb{Q} \in \mathcal{R}^*$, the following inequality holds (see, for instance, Shiryaev [1995]):

$$\mathbb{Q}\left(\bigvee_{t=0}^{T} X_t \geqslant a\right) \leqslant \frac{x^*}{a} \leqslant \frac{\varepsilon}{3}. \tag{5.53}$$

Notice that (5.53) holds uniformly across all measures from $\mathcal{R}^*$. Since only a finite number of points from $\operatorname{supp}(X \circ \mathbb{P})$ is contained in the cube $[0,a]^{T+1}$, there exists such $N = N(\varepsilon)$ that $x_n \in ([0,a]^{T+1})^c$ for all $n \geqslant N(\varepsilon)$.

Consider a sequence of densities $Z^{(k)} \in Z_{\mathcal{R}^*}$ with an associated sequence of probability measures $\mathbb{Q}^{(k)} \in \mathcal{R}^*$ defined as $\mathbb{Q}^{(k)}(d\omega) = Z^{(k)}(\omega)\mathbb{P}(d\omega)$, and denote $q^{(k)}(\omega) = \mathbb{Q}^{(k)}(\{\omega\})$. By applying the standard Cantor diagonalization procedure, it is possible to construct a pointwise converging subsequence $q^{(k_j)}(\omega) \to q^*(\omega) \geqslant 0$. Then for each $\varepsilon > 0$ there exists such $M = M(\varepsilon)$ that for all $j \geqslant M$

$$\sum_{i=1}^{N} |q^{(k_j)}(x_i) - q^*(x_i)| < \frac{\varepsilon}{3}. \tag{5.54}$$

Besides,

$$\sum_{i=N+1}^{\infty} q^{(k_j)}(x_i) \leqslant \mathbb{Q}\left(X \in \left([0,a]^{T+1}\right)^c\right) \leqslant \frac{\varepsilon}{3}, \tag{5.55}$$

and, consequently,

$$\sum_{i=N+1}^{\infty} q^*(x_i) \leqslant \frac{\varepsilon}{3}. \tag{5.56}$$

Therefore, for all $j \geqslant M$

$$\begin{aligned}\sum_{i=1}^{\infty} |q^{(k_j)}(x_i) - q^*(x_i)| &\leqslant \sum_{i=1}^{N} |q^{(k_j)}(x_i) - q^*(x_i)| \\ &+ \sum_{i=N+1}^{\infty} q^{(k_j)}(x_i) + \sum_{i=N+1}^{\infty} q^*(x_i) < \varepsilon.\end{aligned} \tag{5.57}$$

We conclude that for an arbitrary sequence of densities $Z^{(k)}$ from $Z_{\mathcal{R}^*}$ we can select such subsequence that the corresponding subsequence of measures $\mathbb{Q}^{(k_j)} \in \mathcal{R}^*$ converges in total variation norm to a limiting measure $\mathbb{Q}^*$ that is absolutely continuous with respect to $\mathbb{P}$, where $\mathbb{Q}^*(\{\omega\}) = q^*(\omega)$. Therefore, the subsequence $Z^{(k_j)}$ converges in $L^1(\mathbb{P})$ to a proper density Z^*, with $\mathbb{Q}^*(d\omega) = Z^*(\omega)\mathbb{P}(d\omega)$. □

Remark 5.4: If the requirements of Proposition 5.4 are satisfied, the set $Z_{\mathcal{P}^*}$ of all equivalent measures under which X is a martingale is also relatively compact in $L^1(\mathbb{P})$.

Remark 5.5: The assumption that the filtration is generated by the process X is essential in the proof of Proposition 5.4. One could easily construct an example where the probability space also hosts another random variable Y that is independent of X, and the sequence of densities of Y can be chosen in such a way that the uniform integrability of the sequence of densities is violated.

Corollary 5.3.1: Denote by $\mathcal{A}^*$ the set of densities of all probability measures absolutely continuous with respect to $\mathbb{P}$ under which X is a martingale. Then the following statements hold:

(1) if $\mathcal{A}^*$ is compact in $L^1(\mathbb{P})$, then the support of $\mathbb{P}_s(\omega, \cdot)$ is finite, for all $s \in \mathcal{T}_+$ and $\mathbb{P}$-a.s. for all $\omega \in \Omega$;

(2) if the support of $\mathbb{P}_s(\omega, \cdot)$ is finite for all $s \in \mathcal{T}_+$ and $\mathbb{P}$-a.s. for all $\omega \in \Omega$, and in addition $\mathcal{F}_t$ is generated by $\{X_s\}_{0 \leqslant s \leqslant t}$ and $\mathcal{F} = \mathcal{F}_T$, then $\mathcal{A}^*$ is compact in $L^1(\mathbb{P})$.

Proof. **(1)** According to Theorem 5.3, relative compactness implies that the support of $\mathbb{P}_s(\omega, \cdot)$ must consist of isolated points only. Besides, in order for the martingale property to be preserved, the support set must be bounded (Theorem 5.2), thus it must be finite.

(2) Since the support of $\mathbb{P}_s(\omega, \cdot)$ is finite, then by Proposition 5.4 for an arbitrary sequence of densities of absolutely continuous martingale measures we can select a subsequence that converges in $L^1(\mathbb{P})$ to a density of a limiting measure that is absolutely continuous with respect to $\mathbb{P}$. Moreover, the finiteness of the support set implies its boundedness, hence by Theorem 5.2 X will be a martingale under the limiting measure. □

The results that we obtained for the discrete-time case led us thinking that the same sort of results regarding the weak closedness of the set of densities can be derived in continuous-time. A constructive proof of non-closedness of the set of densities of equivalent sigma-martingale measure due to non-equivalence of the limiting measure has been kindly provided to us by George Lowther[5], which we present below in full for the sake of completeness of presentation.

Consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ with a fixed time horizon $T > 0$ and an $\mathbb{R}^d$-valued semimartingale $X = (X_t)_{t\in[0,T]}$. Same as before, we denote by $\mathcal{P}^*$ the set of probability measures equivalent to $\mathbb{P}$ under which X is a sigma-martingale, and we denote the set of densities of equivalent sigma-martingale measures by $Z_{\mathcal{P}^*}$. The objective is to prove that $Z_{\mathcal{P}^*}$ is compact in $L^1(\mathbb{P})$ if and only if it is a singleton or an empty set (Theorem 5.4).

Lemma 5.4: *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, S be an integrable $\mathbb{R}^d$-valued random variable and Z be a uniformly bounded*

[5] E-mail: *george.lowther@blueyonder.co.uk*

non-negative random variable with

$$\mathbb{P}(Z \geqslant K) > 0 \tag{5.58}$$

for some $K > 0$. Then there exists a non-negative random variable X and a uniformly bounded $\mathcal{G}$-measurable random variable $Y \geqslant K$ such that

$$X \leqslant Y, \quad \mathbb{P}(X = Y) > 0 \tag{5.59}$$

and

$$\mathbb{E}(XS \mid \mathcal{G}) = \mathbb{E}(ZS \mid \mathcal{G}). \tag{5.60}$$

Denote the set of $\mathcal{G}$-measurable random variables by $\mathcal{R}$ and let

$$Y = \text{ess inf } \{\gamma : \gamma \in \mathcal{R}, \gamma \geqslant Z \vee K, \ \mathbb{P}\text{-a.s.}\}. \tag{5.61}$$

By construction, $Y \geqslant K$ and $Y \geqslant Z$. If $\mathbb{P}(Y = Z) > 0$, then we can assign $X = Z$ which will satisfy conditions (5.59)-(5.60).

Otherwise, for each $x \in (0, 1)$, define a $\mathcal{G}$-measurable random variable

$$U_x = \mathbb{E}(\mathbf{1}_{\{Z \geqslant xY\}} YS \mid \mathcal{G}), \tag{5.62}$$

which can be chosen to be jointly measurable (e.g., take it to be finite variation in x). Then

$$\int_0^1 U_x dx = \mathbb{E}\left(\int_0^1 \left(\mathbf{1}_{\{Z \geqslant xY\}} YS\right) dx \mid \mathcal{G}\right) \tag{5.63}$$

$$= \mathbb{E}\left(YS \int_0^1 \mathbf{1}_{\{Z \geqslant xY\}} dx \mid \mathcal{G}\right) \tag{5.64}$$

$$= \mathbb{E}(ZS \mid \mathcal{G}). \tag{5.65}$$

According to Lemma 5.2, for each $\omega \in \Omega$

$$\int_0^1 U_x(\omega)\ dx \in \operatorname{conv}(\{U_x(\omega) : x \in (0,1)\}), \tag{5.66}$$

therefore there exist sequences of random variables p_n and x_n such that p_n are non-negative, eventually zero with $\sum_n p_n = 1$, $x_n \in (0,1)$, and

$$\int_0^1 U_x(\omega) dx = \sum_n p_n(\omega) U_{x_n(\omega)}(\omega). \tag{5.67}$$

Due to the measurable section theorem, the convex combinations in (5.67) can be selected in a measurable way so that all of p_n and x_n are $\mathcal{G}$-measurable.

Let

$$X = \sum_n p_n \mathbf{1}_{\{Z \geqslant x_n Y\}} Y, \tag{5.68}$$

then

$$\mathbb{E}(XS \mid \mathcal{G}) = \mathbb{E}\left(\sum_n p_n \mathbb{E}\left(\mathbf{1}_{\{Z \geqslant x_n Y\}} YS \mid \mathcal{G}\right)\ \mid \mathcal{G}\right) = \mathbb{E}(ZS \mid \mathcal{G}), \tag{5.69}$$

therefore (5.60) is satisfied.

Besides,

$$X = \sum_n p_n \mathbf{1}_{\{Z \geqslant x_n Y\}} Y \leqslant \sum_n p_n Y = Y, \tag{5.70}$$

hence to finish the proof we only need to verify that $\mathbb{P}(X = Y) > 0$.

Define

$$\bar{x}(\omega) = \max_{p_n(\omega)>0} x_n(\omega), \tag{5.71}$$

then

$$P(\bar{x}Y \vee K < Y) > 0, \tag{5.72}$$

which in conjunction with the definition (5.61) of Y yields

$$\mathbb{P}(Z > \bar{x}Y) > 0. \tag{5.73}$$

However, $Z > \bar{x}Y$ implies $X = Y$, thus (5.59) is satisfied. □

Theorem 5.4: *The set $Z_{\mathcal{P}^*}$ is compact in $L^1(\mathbb{P})$ if and only if it is a singleton or an empty set.*

Assume that $\mathcal{P}^*$ contains at least two distinct elements. Choose two different measures $\mathbb{Q} \in \mathcal{P}^*$ and $\mathbb{Q}' \in \mathcal{P}^*$, and without loss of generality replace the original measure $\mathbb{P}$ with

$$\mathbb{P} := \frac{\mathbb{Q} + \mathbb{Q}'}{2}, \tag{5.74}$$

since $\mathcal{P}^*$ only depends on $\mathbb{P}$ up to equivalence.

By construction, $\mathbb{P} \in \mathcal{P}^*$ and

$$Z_{\mathbb{Q}} = \frac{d\mathbb{Q}}{d\mathbb{P}} \leqslant 2. \tag{5.75}$$

Also, since X is a sigma-martingale under $\mathbb{P}^* \in \mathcal{P}^*$, let

$$X = \varphi \circ S \equiv \int \varphi dS, \tag{5.76}$$

where φ is a strictly positive predictable process and $S = (S_t)_{t\in[0,T]}$ is an $\mathcal{H}^1$ martingale under $\mathbb{P}^* \in \mathcal{P}^*$, that is, $\sup_t ||S_t||$ is $\mathbb{P}$-integrable.

Define a $\mathbb{P}$-martingale (in what follows, we shall refer to $\mathbb{P}$-martingale as simply a martingale) $U = (U_t)_{t\in[0,T]}$ as

$$U_t = \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \mathbb{E}\left(\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t}\right). \tag{5.77}$$

This is a uniformly bounded martingale with $\mathbb{E}(U_t) = 1$. Moreover,

$$\mathbb{P}(U_t \neq 1) > 0, \tag{5.78}$$

since measures $\mathbb{P}$ and $\mathbb{Q}$ are chosen to be distinct.

Assuming that the filtration is right-continuous and complete, every martingale has a càdlàg modification. Consider a càdlàg modification of U, then it is strictly positive and non-constant martingale with S and $U \circ S$ both martingales. By integration by parts, this implies that the quadratic covariation $\langle U, S\rangle$ is a local martingale.

Suppose that we can find a non-negative martingale $M = (M_t)_{t\in[0,T]}$ which is not identically zero such that $M \circ S$ is a martingale and $\mathbb{P}(M_T = 0) > 0$. Then we can define an absolutely continuous but not-equivalent martingale measure $\mathbb{P}^*$ by

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \frac{M_T}{\mathbb{E}(M_T)}. \tag{5.79}$$

According to Lemma 5.1, this would imply that $Z_{\mathcal{P}^*}$ is not compact in $L^1(\mathbb{P})$.

Let us consider four exhaustive cases.

1. $\mathbb{P}(U_0 \neq 1) > 0$. Then we can consider a set $A = \{U_0 \geqslant 1\}$ and take M to be an almost-surely constant martingale defined as

$$M_t = \frac{\mathbf{1}_A}{\mathbb{P}(A)}. \tag{5.80}$$

2. $U_0 = 1$ a.s., and $\Delta U_t = U_t - U_{t-} \geqslant 0$ for all $t \in (0, T]$, that is, U only has upward jumps. Since U_t is assumed to be not almost-surely constant, there exists $K \in (0, 1)$ such that

$$\mathbb{P}\left(\inf_{t\in(0,T]} U_t < K\right) > 0. \tag{5.81}$$

Let τ be the first time at which $U_t \leqslant K$, and let $\tau = \infty$ if this never happens, then

$$\mathbb{P}(\tau < \infty) > 0. \tag{5.82}$$

Since U_t has only non-negative jumps, $U_\tau = K$ whenever $\tau < \infty$. Then, the martingale M can be defined as

$$M_t = U_{t\wedge\tau} - K, \tag{5.83}$$

and it hits zero whenever $\tau < \infty$, which happens with positive probability.

3. $U_0 = 1$ a.s., $\mathbb{P}(\Delta U_t < 0) > 0$ for some $t \in (0, T]$, and there exists a predictable stopping time $\tau > 0$ at which $\mathbb{P}(U_\tau \neq 0) > 0$. Then, since

U, S and $U \circ S$ are all martingales,

$$\mathbb{E}(\Delta U_\tau \Delta S_\tau \mid \mathcal{F}_{\tau-}) = \mathbb{E}(\Delta(U \circ S)_\tau - U_{\tau-}\Delta S_\tau - S_{\tau-}\Delta U_\tau \mid \mathcal{F}_{\tau-}) = 0. \tag{5.84}$$

As U_t is bounded by 2, we have $\Delta U_\tau \geqslant -2$. Also, since ΔU_τ is not identically zero and $\mathbb{E}(\Delta U_\tau \mid \mathcal{F}_{\tau-}) = 0$, there exists an $\varepsilon > 0$ such that

$$\mathbb{P}(\Delta U_\tau \geqslant \varepsilon) > 0. \tag{5.85}$$

By applying Lemma 5.4 to $\mathbb{R}^{d+1}$-valued $\mathcal{F}_\tau$-measurable random variable $(1, \Delta S_\tau)$ and non-negative random variable $1+\Delta U_\tau/2$ with $K = 1+\varepsilon/2$, there exists a uniformly bounded $\mathcal{F}_{\tau-}$-measurable random variable $Y \geqslant 1+\varepsilon/2$ and a non-negative $\mathcal{F}_{\tau-}$-measurable random variable $X \leqslant Y$ such that $\mathbb{P}(X = Y) > 0$, and

$$\mathbb{E}(X\Delta S_\tau \mid \mathcal{F}_{\tau-}) = \mathbb{E}((1 + \Delta U_\tau/2)\Delta S_\tau \mid \mathcal{F}_{\tau-}) = 0, \tag{5.86}$$

$$E(X \mid \mathcal{F}_{\tau-}) = E(1 + \Delta U_\tau/2 \mid \mathcal{F}_{\tau-}) = 1. \tag{5.87}$$

The martingale M can then be defined by

$$M_t = 1 - \mathbf{1}_{\{t \geqslant \tau\}}(X - 1)/(Y - 1). \tag{5.88}$$

Note that $M_T = 0$ whenever $X = Y$, hence $\mathbb{P}(M_T = 0) > 0$. Besides, the quadratic covariation

$$\langle M, S\rangle_t = \mathbf{1}_{\{t \geqslant \tau\}}(X - 1)\Delta S_\tau/(Y - 1) \tag{5.89}$$

is a martingale, therefore, by integration by parts, $M \circ S$ is also a martingale.

4. $U_0 = 1$ a.s., $\mathbb{P}(\Delta U_t < 0) > 0$ for some $t \in (0, T]$, and U_t is quasi-left-continuous (i.e., $\Delta U_\tau = 0$, a.s., for each predictable stopping time τ). In this case there exists an $\varepsilon > 0$ and $t \in (0, T]$ such that

$$\mathbb{P}(\Delta U_t < -\varepsilon) > 0. \tag{5.90}$$

Let τ be the first time at which $\Delta U_\tau < -\varepsilon$, and let $\tau = \infty$ if this never happens. By Lemma 5.4, there exists an $F_{\tau-}$-measurable $Y \geqslant \varepsilon$ and a non-negative $F_{\tau-}$-measurable random variable $X \leqslant Y$ such that $\mathbb{P}(X = Y) > 0$, and

$$\mathbb{E}(X \Delta S_\tau \mid \mathcal{F}_{\tau-}) = \mathbb{E}(-\Delta U_\tau \Delta S_\tau \mid \mathcal{F}_{\tau-}). \tag{5.91}$$

Let V be the step process defined by

$$V_t = \mathbf{1}_{\{t \geqslant \tau\}} (X + \Delta U_\tau) \Delta S_\tau. \tag{5.92}$$

Then the quadratic covariation

$$\langle V, S \rangle_t = \mathbf{1}_{\{t \geqslant \tau\}} (X + \Delta U_\tau) \Delta S_\tau \tag{5.93}$$

is a martingale. Since $\mathbb{P}(\eta = \tau < \infty) = 0$ for each predictable η, stopping time τ is totally inadmissible. Therefore, it has a compensator $V^\tau = (V_t^\tau)_{t \in [0,T]}$, which is a continuous finite variation adapted process starting from zero such that $V - V^\tau$ is a martingale. Consider the martingale

$$N_t = U_{t\wedge\tau} + V_t^\tau - V_t, \tag{5.94}$$

then $\langle N, S\rangle = \langle U, S\rangle_\tau - \langle V, S\rangle$ is a local martingale. Also, $\Delta N_t = -X$ and $\mathbb{P}(\Delta N_\tau \geqslant -Y) > 0$. Since Y is $\mathcal{F}_{\tau-}$-measurable, there exists a predictable process ξ with $\xi_\tau = Y$. By replacing ξ with $\xi \vee \varepsilon$, we can suppose that $\xi \geqslant \varepsilon$. Consider the martingale

$$\tilde{N} = \int \xi^{-1} dN, \tag{5.95}$$

then $\Delta\tilde{N}_t = \xi_t^{-1}\Delta N_t \geqslant -1$ and $\mathbb{P}(\Delta\tilde{N}_\tau = -1) > 0$. Let M be the solution to the following SDE:

$$M_t = 1 + \int_0^t M_{s-} d\tilde{N}_s. \tag{5.96}$$

This is the Doléans exponential of $\tilde{N}$ and is given explicitly by

$$M_t = \exp\left(\tilde{N}_t - \frac{1}{2}\langle\tilde{N}\rangle_t\right) \prod_{s\leqslant t} \exp\left(-\Delta\tilde{N}_s + \frac{1}{2}(\Delta\tilde{N}_s)^2\right)(1 + \Delta\tilde{N}_s). \tag{5.97}$$

As $\Delta\tilde{N} \geqslant -1$, we have $M_t \geqslant 0$. Also, $M_T = M_\tau = 0$ whenever $\Delta\tilde{N}_\tau = -1$, which occurs with positive probability. Since

$$dM_t = M_{t-} d\tilde{N}_t, \tag{5.98}$$

$$d\langle M, S\rangle_t = M_{t-} d\langle\tilde{N}, S\rangle_t, \tag{5.99}$$

M and $\langle M, S\rangle$ are local martingales, therefore, by integration by parts,

$M \circ S$ is also a local martingale. By localization (replacing M by M^η for a suitable stopping time η with $\mathbb{P}(\eta \geqslant \tau) > 0$), we can assume that M and $M \circ S$ are both proper martingales. Thus, M satisfies the required properties. □

5.2 Computing expectations in the telegraph market model

Consider the telegraph market model with X_t and N_t defined as in (3.67). The purpose of this section is to provide an explicit (recursive) expression for the conditional density $p_n(t,x)$ which allows to evaluate expectations of the kind $\mathbb{E}\left(g(X_t, N_t)\right)$ numerically for a given function $g(x,n)$.

Theorem 5.5: *In the telegraph market model with processes X_t and N_t defined by (3.67), for an arbitrary function $g : [0,\infty) \times \mathbb{N}_0 \mapsto \mathbb{R}$,*

$$\mathbb{E}\left(g(X_t, N_t)\right) = \sum_{\mathbb{N}_0} \int_{\mathbb{R}} g(x,n) p_n(t,x) dx, \tag{5.100}$$

where for all $t \geqslant 0$ and $x \in \mathbb{R}$,

$$p_0(t,x) = \mathrm{e}^{-\lambda_1 t} \delta(x - c_1 t), \tag{5.101}$$

and for all $k \in \mathbb{N}$,

$$p_{2k-1}(t,x) = \frac{\lambda_1 \left(\phi_1(t,x)\phi_2(t,x)\right)^{k-1}}{\mid c_2 - c_1 \mid \left((k-1)!\,\right)^2} \mathrm{e}^{-\phi_1(t,x) - \phi_2(t,x)}, \tag{5.102}$$

$$p_{2k}(t,x) = \frac{p_{2k-1}(t,x)\phi_2(t,x)}{k}, \tag{5.103}$$

with

$$\phi_1(t,x) = \lambda_1 \frac{c_2 t - x}{c_2 - c_1}, \tag{5.104}$$

$$\phi_2(t,x) = \lambda_2 \frac{x - c_1 t}{c_2 - c_1}, \tag{5.105}$$

and $x \in (c_1 t \wedge c_2 t, c_1 t \vee c_2 t)$.

Denote by T_j the time of the j-th jump of $\sigma(t)$ and denote by $\tau_j = T_j - T_{j-1}$ the length of time between two successive jumps. Since $\sigma(t)$ is a continuous-time Markov chain, random variables τ_j are exponentially distributed:

$$\tau_{2k+1} \sim \mathcal{E}(\lambda_1), \tag{5.106}$$

$$\tau_{2k} \sim \mathcal{E}(\lambda_2), \tag{5.107}$$

where $\mathcal{E}(\lambda)$ is the exponential distribution with parameter λ.

Denote

$$S_k^{\text{odd}} = \tau_1 + \tau_3 + \cdots + \tau_{2k-1}, \tag{5.108}$$

$$S_k^{\text{even}} = \tau_2 + \tau_4 + \cdots + \tau_{2k}. \tag{5.109}$$

then

$$S_k^{\text{odd}} \sim \Gamma(k, \lambda_1), \tag{5.110}$$

$$S_k^{\text{even}} \sim \Gamma(k, \lambda_2), \tag{5.111}$$

where $\Gamma(\alpha, \beta)$ is the gamma distribution with parameters α and β.

Without loss of generality, we assume that $c_2 > c_1$; the other case can be handled similarly. Note that the drift of X_t is equal to c_1 when $T_{2j} < t < T_{2j+1}$, and it is equal to c_2 when $T_{2j-1} < t < T_{2j}$.

Hence, if $N_t = 2k$, we have $T_{2k} < t < T_{2k+1}$ and $T_{2k} = S_k^{\text{odd}} + S_k^{\text{even}}$, thus

$$X_t = c_1 S_k^{\text{odd}} + c_2 S_k^{\text{even}} + c_1(t - T_{2k}) = (c_2 - c_1) S_k^{\text{even}} + c_1 t. \tag{5.112}$$

Similarly, if $N_t = 2k+1$,

$$X_t = c_1 S_{k+1}^{\text{odd}} + c_2 S_k^{\text{even}} + c_2(t - T_{2k+1}) = (c_1 - c_2) S_{k+1}^{\text{odd}} + c_2 t. \tag{5.113}$$

By using the identity

$$\{N_t = 2k\} = \{T_{2k} \leqslant t < T_{2k} + \tau_{2k+1}\}, \tag{5.114}$$

we conclude that

$$\begin{aligned} p_{2k}(t,x) = \mathbb{P}\ \Big(\{(c_2 - c_1) S_k^{\text{even}} < x - c_1 t\} \cap \\ \left\{ S_k^{\text{odd}} + S_k^{\text{even}} < t < S_k^{\text{odd}} + S_k^{\text{even}} + \tau_{2k+1} \right\}\Big). \end{aligned} \tag{5.115}$$

Finally, since S_k^{odd}, S_k^{even} and τ_{2k+1} are independent random variables with distributions $\Gamma(k, \lambda_1)$, $\Gamma(k, \lambda_2)$ and $\mathcal{E}(\lambda_1)$ respectively, we obtain

$$p_{2k}(t,x) = \frac{\lambda_1^{k+1} \lambda_2^k}{\left((k-1)!\right)^2} \int_0^{\frac{x - c_1 t}{c_2 - c_1}} dz_2 \int_0^{t - z_2} dz_1 \int_{t - z_1 - z_2}^{\infty} dz_3\ (z_1 z_2)^{k-1} \mathrm{e}^{-\lambda_1 (z_1 + z_3) - \lambda_2 z_2}, \tag{5.116}$$

By differentiating both sides of (5.116) with respect to x and computing the double integral explicitly, we arrive to (5.103). The expression (5.102) for the density at the odd indices can be derived in a similar way. □

5.3 m-th moment of the CEV distribution with $\beta = 1$

Lemma 5.5: *Let S_t be a CEV process with $\beta = 1$, as defined in (4.48). Then for any fixed $\tau > 0$*

$$\mathbb{E}(S_{t+\tau}^m \mid S_t) = \sum_{i=0}^{m-1} \alpha_i S_t^{i+1} \mathrm{e}^{(i+1)r\tau} \left(\frac{2r}{\sigma^2(\mathrm{e}^{r\tau} - 1)} \right)^{i+1-m}, \tag{5.117}$$

where $(\alpha_0, \alpha_1, \ldots \alpha_{m-1})$ is the coordinate vector of polynomial

$$p_{m-1}(n) = \prod_{j=1}^{m-1} (n + m - j + 1) \tag{5.118}$$

relative to basis $\{e_0(n), e_1(n), \ldots e_{m-1}(n)\}$, with

$$e_i(n) = \prod_{j=1}^{i} (n - j + 1), \quad i = 1, 2, \ldots m - 1. \tag{5.119}$$

In order to prove the lemma, we shall use the fact that the conditional density of a CEV stock price may be expressed in terms of power series (see e.g.. [Randal, 1998]).

Denote by $f_\tau(s)$ the conditional density of $S_{t+\tau}$ given S_t:

$$f_\tau(s \mid S_t) = \kappa \mathrm{e}^{-x-z} \sum_{n=0}^{\infty} \frac{x^{n+1} z^n}{n!\,(n+1)!}, \tag{5.120}$$

where

$$x = \kappa S_t \mathrm{e}^{r\tau}, \quad z = \kappa s, \quad \kappa = \frac{2r}{\sigma^2\left(\mathrm{e}^{r\tau} - 1\right)}. \tag{5.121}$$

Note that

$$f_\tau(s \mid S_t)ds = \mathrm{e}^{-x-z} \sum_{n=0}^{\infty} \frac{x^{n+1} z^n}{n!\,(n+1)!} dz, \tag{5.122}$$

therefore

$$\begin{aligned}
\mathbb{E}(S_{t+\tau}^m \mid S_t) &= \int_0^\infty \left(\frac{z}{\kappa}\right)^m \mathrm{e}^{-x-z} \sum_{n=0}^{\infty} \frac{x^{n+1} z^n}{n!\,(n+1)!} dz \\
&= \left(\frac{x}{\kappa}\right) \kappa^{1-m} \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-x} x^n}{n!} \int_0^\infty \frac{\mathrm{e}^{-z} z^{n+m}}{(n+1)!} dz \\
&= S_t \mathrm{e}^{r\tau} \kappa^{1-m} \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-x} x^n}{n!} \frac{(n+m)!}{(n+1)!} \\
&= S_t \mathrm{e}^{r\tau} \kappa^{1-m} \sum_{n=0}^{\infty} g_x(n) p_{m-1}(n),
\end{aligned} \tag{5.123}$$

where $g_x(n)$ is a probability mass function of a Poisson random variable with mean x.

If we choose such $\alpha_0, \alpha_1, \ldots \alpha_{m-1}$ that

$$p_{m-1}(n) = \sum_{i=0}^{m-1} \alpha_i \prod_{j=1}^{i} (n - j + 1) \tag{5.124}$$

and notice that

$$\prod_{j=1}^{i} (n - j + 1) = \begin{cases} 0 & \text{if } n = 0, 1, \ldots i - 1, \\ \dfrac{n!}{(n-i)!} & \text{if } n = i, i + 1, \ldots, \end{cases} \tag{5.125}$$

we can rewrite the expectation as

$$
\begin{aligned}
\mathbb{E}(S_{t+\tau}^m \mid S_t) &= S_t \mathrm{e}^{r\tau} \kappa^{1-m} \sum_{i=0}^{m-1} \sum_{n=i}^{\infty} \alpha_i \frac{\mathrm{e}^{-x} x^{n-i}}{(n-i)!} x^i \\
&= S_t \mathrm{e}^{r\tau} \kappa^{1-m} \sum_{i=0}^{m-1} \alpha_i x^i \sum_{n=i}^{\infty} g_x(n-i) \\
&= S_t \mathrm{e}^{r\tau} \kappa^{1-m} \sum_{i=0}^{m-1} \alpha_i \left(\kappa S_t \mathrm{e}^{r\tau}\right)^i \\
&= \sum_{i=0}^{m-1} \alpha_i S_t^{i+1} \mathrm{e}^{(i+1)r\tau} \kappa^{i+1-m}, \qquad (5.126)
\end{aligned}
$$

which concludes the proof. □

Bibliography

[Acerbi, 2002] Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7):1505–1518.

[Acerbi and Tasche, 2002a] Acerbi, C. and Tasche, D. (2002a). Expected shortfall: a natural coherent alternative to Value at Risk. *Economic Notes*, 31:379–388.

[Acerbi and Tasche, 2002b] Acerbi, C. and Tasche, D. (2002b). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503.

[Akilov and Kantorovich, 1984] Akilov, G. and Kantorovich, L. (1984). *Functional Analysis*. Nauka, Moscow, 3rd revised edition. In Russian. English translation: 2nd edition published by Pergamon Press, Oxford, UK (1984).

[Artzner et al., 1999] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

[Barlow and Perkins, 1984] Barlow, M. and Perkins, E. (1984). One dimensional stochastic differential equations involving a singular increasing process. *Stochastics*, 12:229–249.

[Basel Committee on Banking Supervision, 2012] Basel Committee on Banking Supervision (2012). Fundamental review of the trading book. Consultative document, available at http://www.bis.org/publ/bcbs219.pdf.

[Bergenthum and Ruschendorf, 2007] Bergenthum, J. and Ruschendorf, L. (2007). Comparison of Semimartingales and Lèvy Processes. *Annals of Probablity*, 35(1):228–254.

[Black and Scholes, 1973] Black, F. and Scholes, M. S. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654.

[Brennan and Schwartz, 1976] Brennan, M. and Schwartz, E. (1976). The pricing of equity-linked life insurance policies with an asset value guarantee. *Journal of Financial Economics*, 3(3):195–213.

[Cohen et al., 2010] Cohen, S., Elliott, R., and Pearce, C. (2010). A general comparison theorem for backward stochastic differential equations. *Advances in Applied Probability*, 42(3):878–898.

[Cox and Ross, 1976] Cox, J. and Ross, S. (1976). The Valuation of Options for Alternative Stochastic Processes. *Journal of Financial Economics*, 3:145–166.

[Cvitanić and Karatzas, 2001] Cvitanić, J. and Karatzas, I. (2001). Generalized Neyman-Pearson lemma via convex duality. *Bernoulli*, 7(1):79–97.

[Cvitanić and Spivak, 1999] Cvitanić, J. and Spivak, G. (1999). Maximizing the probability of a perfect hedge. *Annals of Applied Probability*, 9(4):1303–1328.

[Delbaen and Schachermayer, 2006] Delbaen, F. and Schachermayer, W. (2006). *The Mathematics of Arbitrage*. Springer, Berlin Heidelberg.

[Dellacherie and Meyer, 1975] Dellacherie, C. and Meyer, P. (1975). *Probabilités et potentiel, tome I*. Hermann, Paris.

[Ding and Wu, 1998] Ding, X. and Wu, R. (1998). A new proof for comparison theorems for stochastic differential inequalities with respect to semimartingales. *Stochastic Processes and their Applications*, 78(2):155–171.

[Dunford and Schwartz, 1988] Dunford, N. and Schwartz, J. (1988). *Linear Operators, General Theory (Part 1)*. Wiley Classics Library. Wiley-Interscience.

[El Karoui and Quenez, 1995] El Karoui, N. and Quenez, M. (1995). Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal on Control and Optimization*, 33(1):29–66.

[Föllmer and Kabanov, 1998] Föllmer, H. and Kabanov, Y. (1998). Optional decomposition and Lagrange multipliers. *Finance and Stochastics*, 2:69–81.

[Föllmer and Leukert, 1999] Föllmer, H. and Leukert, P. (1999). Quantile hedging. *Finance and Stochastics*, 3(3):251–273.

[Föllmer and Leukert, 2000] Föllmer, H. and Leukert, P. (2000). Efficient hedging: cost versus shortfall risk. *Finance and Stochastics*, 4(2):117–146.

[Föllmer and Schied, 2011] Föllmer, H. and Schied, A. (2011). *Stochastic Finance. An Introduction in Discrete Time*. Walter de Gruyter, Berlin, third revised and extended edition.

[Galtchouk, 1982] Galtchouk, L. (1982). A Comparison Theorem for Stochastic Equations with Integrals with Respect to Martingales and Random Measures. *Teoriya Veroyatnostei i ee Primeneniya*, 27(3):425–433. In Russian. English translation: Theory of Probability and Its Applications, **27**(3) (1983), 450–460.

[Hajek, 1985] Hajek, B. (1985). Mean stochastic comparison of diffusions. *Probability Theory and Related Fields*, 68:315–329.

[Jacod and Shiryaev, 1998] Jacod, J. and Shiryaev, A. (1998). Local martingale and the fundamental asset pricing theorems in the discrete-time case. *Finance and Stochastics*, 2(3):259–273.

[Jansen et al., 1986] Jansen, K., Haezendonck, J., and Goovaerts, M. (1986). Upper bounds on stop-loss premiums in case of known moments up to the fourth order. *Insurance: Mathematics and Economics*, 5:315–334.

[Karatzas and Rudloff, 2010] Karatzas, I. and Rudloff, B. (2010). Testing composite hypotheses via convex duality. *Bernoulli*, 16(4):1224–1239.

[Kloeden and Platen, 1992] Kloeden, P. and Platen, E. (1992). *Numerical solution of stochastic differential equations*. Applications of mathematics. Springer-Verlag.

[Kramkov and Schachermayer, 1999] Kramkov, D. and Schachermayer, W. (1999). The Asymptotic Elasticity of Utility Functions and Optimal Investment in Incomplete Markets. *Annals of Applied Probability*, 9(3):904–950.

[Krasin, 2010] Krasin, V. (2010). Comparison Theorem and Its Applications to Finance. PhD thesis, University of Alberta.

[Krasin and Melnikov, 2010] Krasin, V. and Melnikov, A. (2010). On comparison theorem and its applications to finance. In *Optimality and Risk – Modern Trends in Mathematical Finance*, pages 171–181. Springer Berlin Heidelberg.

[Lèvy, 1985] Lèvy, H. (1985). Upper and Lower Bounds of Put and Call Option Value: Stochastic Dominance Approach. *Journal of Finance*, 40(4):1197–1217.

[Li and Xu, 2008] Li, J. and Xu, M. (2008). Risk minimizing portfolio optimization and hedging with conditional Value-at-Risk. *Review of Futures Markets*, 16:471–506.

[MacBeth and Merville, 1980] MacBeth, J. and Merville, L. (1980). Tests of the Black-Scholes and Cox Call Option Valuation Models. *Journal of Finance*, 35(2):211–219.

[McGill et al., 2004] McGill, D., Brown, K., Haley, J., and Schieber, S. (2004). *Fundamentals of Private Pensions*. Oxford University Press.

[Melnikov, 1979a] Melnikov, A. (1979a). On the theory of stochastic equations in components of semimartingales. *Matematichiskii Sbornik*, 110(3):414–427. In Russian. English translation: Mathematics of the USSR – Sbornik, **38**(3) (1981), 381–394.

[Melnikov, 1979b] Melnikov, A. (1979b). Strong solutions of stochastic differential equations with non-smooth coefficients. *Teoriya Veroyatnostei i ee Primeneniya*, 24(1):146–149. In Russian. English translation: Theory of Probability and Its Applications, **24**(1) (1981), 147–150.

[Melnikov, 1983] Melnikov, A. (1983). On solutions of stochastic equations with driving semimartingales. *Proceedings of the Third European Young Statisticians Meeting*, pages 120–124. Catholic University, Leuven.

[Melnikov and Ratanov, 2007] Melnikov, A. and Ratanov, N. (2007). Nonhomogeneous telegraph processes and their application to financial market modeling. *Doklady Mathematics*, 75(1):115–117.

[Melnikov and Ratanov, 2008] Melnikov, A. and Ratanov, N. (2008). On financial markets based on telegraph processes. *Stochastics An International Journal of Probability and Stochastic Processes*, 80(2):247–268.

[Melnikov and Romaniuk, 2006] Melnikov, A. and Romaniuk, Y. (2006). Evaluating the performance of Gompertz,Makeham and Lee-Carter mortality models for risk management with unit-linked contracts. *Insurance: Mathematics and Economics*, 39(3):310–329.

[Melnikov and Skornyakova, 2005] Melnikov, A. and Skornyakova, V. (2005). Quantile hedging and its applications to life insurance. *Statistics and Decisions*, 23:601–615.

[Melnikov and Smirnov, 2012] Melnikov, A. and Smirnov, I. (2012). Dynamic hedging of conditional value-at-risk. *Insurance: Mathematics and Economics*, 51(1):182–190.

[Merton, 1976] Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.

[Peng and Zhu, 2006] Peng, S. and Zhu, X. (2006). Necessary and sufficient condition for comparison theorem of 1-dimensional stochastic differential equations. *Stochastic Processes and their Applications*, 116(3):370–380.

[Randal, 1998] Randal, J. (1998). The Constant Elasticity of Variance Option Pricing Model. MSc thesis, Victoria University of Wellington.

[Rockafellar, 1970] Rockafellar, R. (1970). *Convex Analysis*. Princeton Mathematical Series. Princeton University Press.

[Rockafellar and Uryasev, 2000] Rockafellar, R. T. and Uryasev, S. (2000). Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2:21–41.

[Rockafellar and Uryasev, 2002] Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471.

[Rogers, 1994] Rogers, L. (1994). Equivalent martingale measures and no-arbitrage. *Stochastics and Stochastic Reports*, 3:523–527.

[Rudloff, 2005] Rudloff, B. (2005). A Generalized Neyman-Pearson Lemma for Hedge Problems in Incomplete Markets. Proceedings of the Workshop *Stochastische Analysis*, 241–250.

[Rudloff, 2006] Rudloff, B. (2006). Hedging in Incomplete Markets and Testing Compound Hypotheses via Convex Duality. PhD thesis, Martin-Luther-Universität Halle-Wittenberg.

[Rudloff, 2009] Rudloff, B. (2009). Coherent Hedging in Incomplete Markets. *Quantitative Finance*, 9(2):197–206.

[Sarykalin et al., 2008] Sarykalin, S., Serraino, G., and Uryasev, S. (2008). Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization. *Tutorials in Operations Research*.

[Schepper and Heijnen, 2007] Schepper, A. and Heijnen, B. (2007). Distribution-free Option Pricing. *Insurance: Mathematics and Economics*, 40(2):179–199.

[Schroder, 1989] Schroder, M. (1989). Computing the Constant Elasticity of Variance Option Pricing Formula. *Journal of Finance*, 44(1):211–219.

[Schulmerich and Trautmann, 2003] Schulmerich, M. and Trautmann, S. (2003). Local Expected Shortfall-Hedging in Discrete Time. *European Finance Review*, 7:75–102.

[Shiryaev, 1995] Shiryaev, A. (1995). *Probability*. Graduate Texts in Mathematics. Springer, 2nd edition.

[Shiryaev, 1998] Shiryaev, A. (1998). *Essentials of Stochastic Finance*. Fazis, Moscow. In Russian. English translation: published by World Scientific, River Edge, NJ, USA (1999).

[Skorokhod, 1961] Skorokhod, A. (1961). *Studies in the Theory of Random Processes*. Naukova Dumka, Kyiv, USSR. In Russian. English translation: published by Addison-Wesley, Reading, MA, USA (1965).

[Steinitz, 1913] Steinitz, E. (1913). Bedingt Konvergente Reihen und Konvexe Systeme i. *Journal fr die reine und angewandte Mathematik*, 143:128–175.

[Steinitz, 1914] Steinitz, E. (1914). Bedingt Konvergente Reihen und Konvexe Systeme ii. *Journal fr die reine und angewandte Mathematik*, 144:1–40.

[Steinitz, 1916] Steinitz, E. (1916). Bedingt Konvergente Reihen und Konvexe Systeme iii. *Journal fr die reine und angewandte Mathematik*, 146:1–52.

[Tasche, 2002] Tasche, D. (2002). Expected Shortfall and Beyond. *Journal of Banking & Finance*, 26(7):1519–1533.

[Xu, 2004] Xu, M. (2004). Minimizing Shortfall Risk Using Duality Approach - An Application to Partial Hedging in Incomplete Markets. PhD thesis, Carnegie Mellon University.

[Yamada, 1973] Yamada, T. (1973). On a comparison theorem for solutions of stochastic differential equations and its applications. *Journal of Mathematics, Kyoto University*, 13(3):497–512.

[Yan, 1986] Yan, J.-A. (1986). A comparison theorem for semimartingales and its applications. *Séminaire de probabilités (Strasbourg)*, 20:349–351.