

PRICING AND HEDGING OPTIONS IN DISCRETE TIME WITH  
LIQUIDITY RISK

YEGOR SOROKIN

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE  
STUDIES  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN DEPARTMENT OF MATHEMATICS AND  
STATISTICS  
YORK UNIVERSITY  
TORONTO, ONTARIO  
JUNE 2014  
©Yegor Sorokin, 2014

## Abstract

Different derivative securities, including European options, are very popular and widely used in forms of exchange-traded instruments or over-the-counter products. For practical purposes the European options are often priced using analytic solution to the Black-Scholes formula. Hedging, according to the Black-Scholes model, is accomplished via the construction of dynamically rebalanced replicating portfolio. However, the model makes several critical assumptions. I extend the Black-Scholes model by relaxing the assumption of no trading costs and considering the market liquidity risk for the underlying asset. Liquidity risk is understood as the effect of the trade size on the price of the underlying asset. I use stochastic supply curve to model liquidity risk.

The problem is to hedge a European option in the presence of the market liquidity risk for an underlying asset. One hedges with the underlying, as the option price depends on the price of the underlying asset. The underlying asset has market liquidity risk; thus, studying the impact of market liquidity risk is important for

devising more effective and efficient option hedging algorithms.

The main contributions of the thesis arise from the investigation of mathematical techniques for hedging and pricing of European options in discrete time with liquidity risk. First, I study delta hedging in Chapter 3. I show  $L^2$  convergence of the replicating trading strategy payoff to the option payoff. In other words, the optimal strategy minimizes the mean squared replication error. I also show that for European put and call options with varying trading times the recommendation is to trade closer to expiry as the spot price of the underlying asset deviates from the strike price. Then I apply the local risk-minimizing hedge in Chapter 4. This time the optimal strategy minimizes the conditional mean squared hedging error. I prove the existence of the local risk-minimizing trading strategy and characterize its structure.

## Acknowledgements

I would like to express my very great appreciation to my supervisor Hyejin Ku for investing so much time into me.

I am also grateful to my mother Valentina Potapova for supporting me all the way through the program.

# Table of Contents

<b>Abstract</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>Table of Contents</b>	<b>v</b>
<b>List of Figures</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background and significance . . . . .	1
1.2 Basic concepts . . . . .	7
1.3 Summary . . . . .	12
<b>2 Previous results</b>	<b>17</b>
2.1 Literature review . . . . .	17
2.2 Çetin-Jarrow-Protter model . . . . .	23
2.3 Partial differential equation for option replication . . . . .	27

<b>3</b>	<b>Delta-hedging replicating strategies</b>	<b>35</b>
3.1	Convergence of the mean squared replication error . . . . .	35
3.2	Optimal strategies: minimizing the hedging error . . . . .	53
<b>4</b>	<b>Hedging via local risk-minimization</b>	<b>66</b>
4.1	Definitions . . . . .	66
4.2	Conditions on the price process and technical results . . . . .	80
4.3	Optimization problem: minimizing the conditional variance . . . . .	99
4.4	Existence and structure of an optimal strategy . . . . .	116
4.5	Algorithm for the numerical computations . . . . .	125
<b>5</b>	<b>Conclusions and future work</b>	<b>130</b>
	<b>Bibliography</b>	<b>132</b>

## List of Figures

1.1	European call option payoff. . . . .	10
1.2	European put option payoff. . . . .	11
2.1	An example of a supply curve. . . . .	26
3.1	Setting up the varying reheding times. Steeper slope of $d(t)$ corresponds to more frequent rebalancing. . . . .	40
3.2	Optimal $d(t)$ for a Euro call for decreasing strikes. . . . .	57
3.3	Optimal $d(t)$ for a Euro call for increasing strikes. . . . .	57
3.4	The numerical solution of the partial differential equation (2.6) with the final condition $V(T, S) = (S_T - K)^+$ . . . . .	58
4.1	The indices for $\delta_k$ run from 1 up to $T + 1$ . . . . .	70
4.2	The fragment of a price process approximation through a tree. . . .	127

# 1 Introduction

## 1.1 Background and significance

Pricing and hedging of derivative contracts are one of the main tasks of mathematical finance. The most widely used Black-Scholes model does not recognize market liquidity risk for the underlying asset. This thesis incorporates liquidity risk into option pricing and hedging in discrete time. We also apply two optimality criteria for the replicating trading strategy: the mean square hedging error and local risk-minimization.

An option is a contract which gives the buyer (the owner) the right, but not the obligation, to buy or sell a particular asset or instrument for an agreed amount on or before a specified time in the future. An option which brings to the owner the right to buy something at a specific price is referred to as a call; an option which conveys the right of the owner to sell something at a specific price is referred to as a put. One refers to the particular asset or an instrument on which the option is based as the underlying. The agreed amount which we can pay for the underlying



is called the exercise price or strike price. Options that can be exercised anytime during its life are called American options. Options that can only be exercised at the end of its life are called European options. We are concerned with European options in this thesis.

Options have been around for several centuries. Italian bankers were using options since the fifteenth century. Dutch were speculating in tulip bulbs using options during the Tulip mania in 1636-37. The options with regular expiration dates were traded at the Amsterdam bourse during the eighteenth century. On 26th April 1973 The Chicago Board Options Exchange first created standardized, listed options. Several months later Black and Scholes finally were able to publish their famous paper in the Journal of Political Economy. Black and Scholes [5] introduce in their paper the method for pricing and hedging European options.

The Black-Scholes model is unique from many perspectives. It brings together and connects intuitively the concepts from mathematics, finance and economics. The book written by Wilmott [26] lists twelve ways to derive the model, including “Black-Scholes for Accountants”. However the Black-Scholes model is based on a number of simplifying assumptions, which render the market complete. In a complete market it is possible to reproduce the payoff of any option using the replicating portfolio consisting of the underlying and cash. It is well-known, however, that financial markets are not complete. Asset prices depend on a multitude

of factors, and there are restrictions on composition of portfolios the investors are allowed to hold. Thus, in general, it is not possible to replicate a derivative payoff perfectly.

We review the assumptions of the Black-Scholes model for options on the stock shares. We postulate that the market has three tradeable assets: a nonrisky asset called cash, a risky asset called the stock and a derivative security. We study a special case when the derivative security is a European call or put.

1. Assumptions on the assets

- (a) The risk-free rate is constant.
- (b) The stock does not pay dividends.
- (c) The price of the stock follows a geometric Brownian motion with constant drift and volatility (a continuous time stochastic process).

2. Assumptions on the market

- (a) It is possible to buy and sell any amount, no matter how close to zero, of the stock shares (this includes short selling).
- (b) It is possible to borrow and lend any amount, no matter how close to zero, of cash.
- (c) There are no transaction costs and taxes associated with trading the assets.

- (d) There is no arbitrage opportunity (i.e., there is no way to make a riskless profit).
- (e) Hedging is done continuously.

According to the assumption (2a) market completeness implies that one is able to buy and sell any amount of the underlying asset at the current market price. However the number of stock shares involved in a particular trade (order size) may influence the quoted price. For example, selling a million of shares quickly will require accepting smaller price per share to find willing buyers, as the marketplace may not have enough buyers at a market price. Continuous trading is not possible in practice. Instead, one is forced to adjust the hedge only finite number of times. To sum up, we will relax the assumptions (2c) and (2e) by introducing liquidity risk as well as considering trading in discrete time in the thesis.

The term “liquidity risk” comes in two forms: funding liquidity risk and market liquidity risk. One talks about funding liquidity risk at the level of institutions, while market liquidity risk relates to an asset or an asset class. Funding liquidity is the ability of a financial institution to settle obligations immediately when due. Market liquidity risk for a specific security arises when one is not able to find enough buyers and sellers to trade the security. In this thesis we are dealing with the market liquidity risk in the context of option pricing and hedging.

One may perceive all financial bubbles as examples of market liquidity crises.

When everybody simultaneously tries to find buyers for a specific security or a class of securities, the market appraises the security at zero value. During the Dutch tulip mania people were preoccupied with trading great quantities of tulip bulbs and derivatives on those for about three to four months until the trading in bulbs collapsed. Seventeenth century England saw high volume of trading in the shares of South Sea Company until the share price crashes. Similarly at the outset of the Great Depression the public rushed to sell off a whole class of securities: stock shares of all sorts. As these historical examples show, one must watch the market liquidity for an asset of interest to avoid losses.

An asset-liability management function in a bank among other things is responsible for funding liquidity risk. The troubles of American International Group (AIG) during the financial crisis of 2007-2008 are examples of the funding liquidity risk. Once AIG's credit rating was downgraded, the firm suddenly needed additional funds to post collateral with its trading counterparties. Lehman Brothers provides another example of funding liquidity risk in action. Once the buyers of the securitized mortgages suddenly disappeared, Lehman could not fund itself by selling off packaged mortgages and at the same time accrued losses on the subprime mortgages it was holding on its books.

Regulators pay close attention to liquidity risk. The Third Accord of the Basel Committee on Banking Supervision [22] focuses on liquidity risk management,

while the second set of recommendations on banking laws and regulations of the Basel Committee on Banking Supervision [21] deals mostly with capital adequacy. Lehman Brothers failed because of mismanaged funding liquidity risk rather than inadequate capital.

Market liquidity risk manifests itself in wider than usual bid-ask spread for an individual security. In other words, market participants may not be able to find a counterparty for buying/selling the security at a market price. In the worst case scenario nobody in the marketplace may want to trade a particular security (liquidity “dries up” as it happened during the financial crisis of 2007-2008). One may think of market liquidity risk as a form of transaction cost. Assuming that the trade does not influence the market price of a security, the price quoted for a transaction may differ from the market price. The size of the trade for immediate execution will have an effect on a quoted price.

From the point of view of economics, one may define market liquidity risk as the absence of perfect elasticity. In our case we consider the elasticity of price of the underlying asset with respect to the buy/sell order size. Most models in mathematical finance assume perfect elasticity for the supply and demand of tradeable assets so that buy/sell orders of discretionary size do not affect asset prices. However, if there is an investor whose trades involve a substantial part of the available asset supply, the asset prices will no longer unfold separately from the trading and

hedging strategies chosen by the “big player”. Typical examples are institutional investors: insurance companies, pension funds, mutual funds, etc.

Market liquidity risk affects pricing of exchange-traded options, since option pricing theory relies on replication of the derivative payoff by trading in the underlying asset. The efficiency of the hedge using the underlying asset depends on the liquidity of the underlying security. Then there is a question of factoring market liquidity risk into pricing and hedging of over-the-counter derivatives. Over-the-counter derivatives are unique and have client-specific contracts tailored towards customer’s needs. As opposed to an exchange-traded derivative with a standard form of the contract, there is no market data as the contracts are extremely diverse and originators do not intend to provide free access to the data.

## **1.2 Basic concepts**

In this section we present the key pioneering ideas of option pricing theory and provide some intuitive explanations. We aim at giving the reader a deep understanding of these concepts.

We will now introduce some financial terms, and elaborate on the meaning of concepts like cash, stock, risky asset, etc., while trying to maintain logical coherence between concepts. However, finance, accounting and economics are social sciences, meaning that there are always several solutions to a problem and several ways of

defining basic concepts. We will take the liberty of selectively defining terms from the financial vocabulary that are used in the thesis. The definitions below are simplified and directed towards features that the models attempt to capture and quantify. The reader should consult finance and economic textbooks for the precise meaning of the notions and details of jargon.

Securities can be universally classified into three groups: debt securities (such as banknotes and bonds), equity securities (e.g. ordinary stock shares) and derivative contracts (such as forwards, futures, options and swaps). In this text we assume for simplicity and ease of presentation that the interest rate on cash is zero. The general case can be reduced to the case with zero interest rate by taking prices as being relative to some numeraire. Asset's liquidity is one of the main concepts in this thesis. An asset's market liquidity is the degree to which an asset can be bought or sold in the market without affecting the asset's price. Put differently, liquidity is an asset's capacity to be sold promptly without a substantial reduction in price. In economic terms, this means that supply meets demand and both parties agree on the price right away and are willing to transact. Cash is an example of a tradeable asset with perfect liquidity. Cash is a nonrisky asset in the sense that its value is preserved over time (recall that interest rate is zero). The state of the money market account is certain and one ends up with the same amount of cash after any period of time.

A stock share is a tradeable asset with non-constant, risky, irregular market price. In the Black Scholes model, shares are perfectly liquid. In our model we will consider the impact of liquidity risk on the price of a derivative. The market price of the share varies with time, reflecting the interplay between supply and demand, arrival of new information to the participants, etc.

Derivative is a security which derives its value from the performance of another entity such as an asset, index, or interest rate, called the “underlying”. Examples of derivatives are futures, forwards, swaps, options and variations of these such as caps, floors and credit default swaps. One may view a derivative as a legal contract between two or more parties.

In this thesis we study a particular type of derivative: European option with stock shares as the underlying asset. An option is a contract which gives the buyer (the owner) the right, but not the obligation, to buy (call option) or to sell (put option) the underlying asset, at an agreed in advance fixed price (strike price) on or before a specified date. We will usually denote the strike price as  $K \in \mathbb{R}$  and the time until expiration as  $T \in \mathbb{R}$ . An option is called European if it may only be exercised at the specified expiration date of that option.

One may visualize the option’s value at expiration via a payoff diagram. Payoff function is a relation between the option value at expiration and the market price of the underlying asset. When we say “payoff” of an option we mean the payoff



function that corresponds to that option type. Generally we will denote the payoff function of the price of the underlying asset as  $p(S)$ . The payoff function for a European call option is  $\max(S - K, 0)$  and its graph is displayed in the Figure 1.1. The payoff function for a European put option is  $\max(K - S, 0)$  and its graph is displayed in the Figure 1.2.

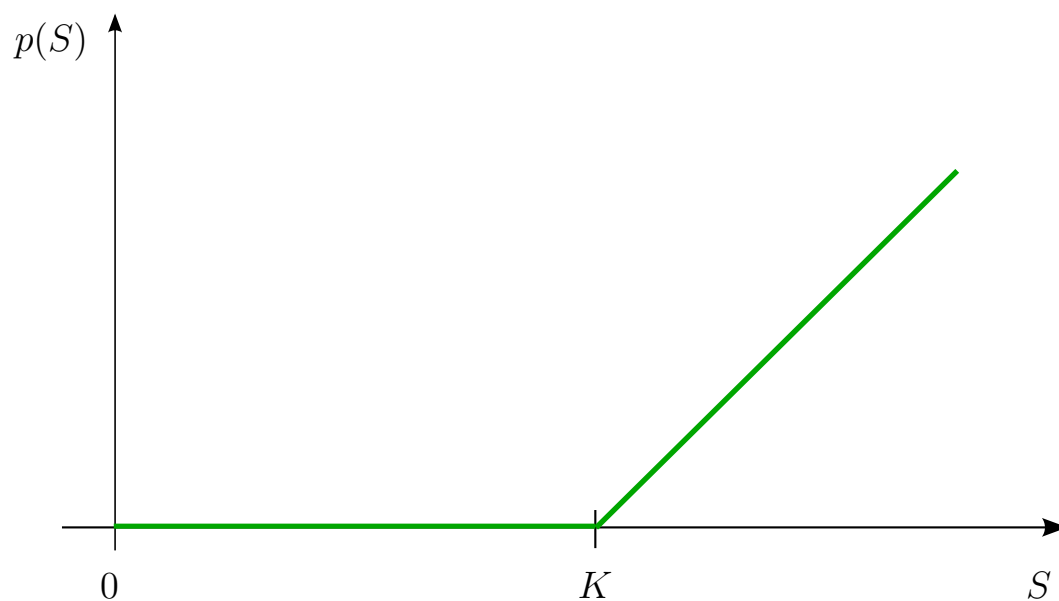


Figure 1.1: European call option payoff.

Pricing an derivative contract means finding the “fair” value of the derivatives contract, which we call a premium. Premium is the amount paid for the derivative contract initially (at time  $t = 0$ ). A portfolio is simultaneous ownership of some units of cash, stock shares and derivatives of specific type at a specific time. Mathematically, a portfolio is an ordered triple of real numbers. We mostly will look at

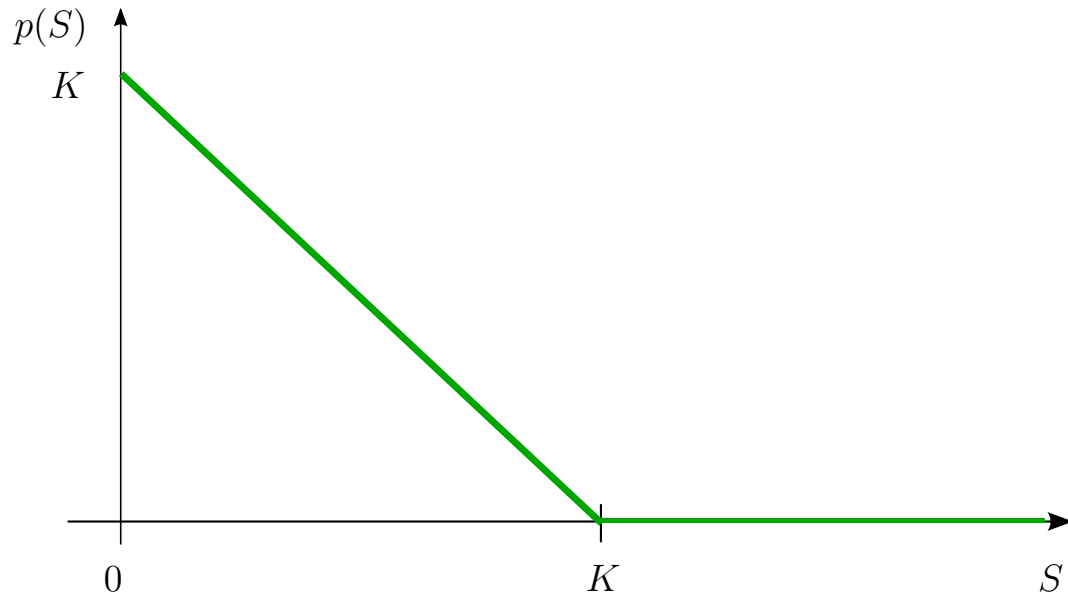


Figure 1.2: European put option payoff.

portfolios with changing amounts of cash and stock shares and fixed amounts of derivative contracts. One can calculate of the current market value of the portfolio denominated in the units of cash. To this end we multiply the number of stock shares in the portfolio by the current share price, multiply the amount of derivative contracts by its value and add the number of units of cash in the portfolio. If we denote the number of shares by  $\delta_t$ , number of units of cash  $\lambda_t$ , number of units of derivatives  $\theta_t$  and at time  $t$  the market price of one share as  $S_t$ , value of one derivative contract  $\Theta_t$ , then portfolio is an ordered pair  $(\delta_t, \lambda_t, \theta_t)$  and  $\delta_t S_t + \lambda_t + \theta_t \Theta_t$  is portfolio's market value. A holding period of a portfolio is the time period over which one owns securities in a portfolio.

Hedging is any reduction in randomness of the portfolio market value. Hedging tries to prevent both losses and gains in the portfolio market value. Cash does not need any hedging, as its market value is constant in the model at all times, which is the opposite of randomness. One may use financial instruments with linear payoff diagrams (futures for instance) to hedge an exposure to stock shares. Derivatives are usually the main tools for hedging as these instruments have nonlinear payoff diagrams. One includes some quantities of correlated assets into the portfolio to achieve the desired effect. Hedging may be static, that is the composition of assets in a portfolio is determined once and does not change over the holding period. The hedge may also be dynamic, so that the amounts of assets in a portfolio is adjusted several times over the holding period. Black-Scholes delta-hedging is an example of a dynamic hedge.

### **1.3 Summary**

This section explains the statements from the abstract and outlines the structure of this thesis. The problem of interest in this thesis is hedging and pricing a European option with market liquidity risk for the underlying asset in discrete time. The main contributions of this thesis are:

- convergence of the replicating strategy payoff to the option payoff in  $L^2$ ;

- solution to the control problem (arranging varying trading times): the suggestion for European call and put options is to re hedge more frequently closer to expiry as the strike deviates from the spot price;
- existence of the local risk-minimizing trading strategy.

We introduce the market liquidity risk for the underlying asset in Chapter 2. The market liquidity risk refers to the impact of the order size on the market price of certain asset. Accordingly, the impact of liquidity risk in the context of option hedging becomes apparent as one decides to trade in the underlying to rebalance the replicating portfolio. Building on the model advanced by Çetin et al. [10], we account for the market liquidity risk for the underlying security via the stochastic supply curve. The supply curve should satisfy five technical assumptions. We use multiplicative supply curve. To clarify, in a multiplicative supply curve some stochastic process models the current market price of the underlying security, while a certain increasing deterministic function models the effect of the order size.

It is not possible to exactly replicate the option payoff using only cash and underlying asset, due to the finite number of trading times at which we adjust the hedge, as well as presence of liquidity risk. This gives rise to the hedging error. The hedging error of a replicating trading strategy is the difference between the final payoffs of the option and the strategy. We will consider two optimality criteria:

mean square error and conditional mean square error. The first approach is called delta hedging, the second model is local risk-minimization. Both approaches are inherently designed for hedging, although the methods may also be used for pricing.

Chapter 3 contains the application of delta hedging approach. To this end we assume that the price process for the underlying asset follows geometric Brownian motion, that is specific stochastic process. We introduce a particular class of discrete-time hedging strategies which are generalizations of the discrete-time hedging strategies presented by Leland [20]. I prove that the payoff of such a strategy converges in  $L^2$  to the theoretical payoff of the option as the length of the revision interval between the rehedging times goes to zero. This is equivalent to saying that the mean square hedging error converges to zero random variable in  $L^2$ . The strategies of this type are optimal in the delta hedging framework.

I then consider an applied problem of minimizing the first order term of the mean square hedging error using a finite number of trading times. As a result, I discover that for a European option the rebalancing trades should be located closer to the expiration as the strike price deviates from the current price of the underlying asset. We introduce the idea of varying trading times and study its impact on the hedging error. Leland's class of trading strategies consists of updates adjusted at constant intervals of time. We extend this class to include strategies with rehedgings located at varying distances from one another. We fix the initial capital of the replicating

portfolio and look for the optimal hedge. We find the leading order term in the Taylor series expansion of the hedging error with respect to the distance between the trading times and then use calculus of variations to find the minimum. We compute both the location of the trading times and corresponding simulated distributions of hedging errors. We compute numerically the optimal trading strategy by solving a partial differential equation. The simulation of the geometric Brownian motion paths allows to compare the performance of the strategies with equally spaced and varying trading times. We also derive the partial differential equation in presence of stochastic volatility after making a simplifying assumption.

Chapter 4 describes the condition for optimality in the sense of local risk-minimization. Local risk-minimization, in contrast to delta hedging, is very general and much more theoretic. I still rely on stochastic supply curve to model the market liquidity risk for the underlying asset. As opposed to the delta hedging model, we do not specify the precise form of the price process to model the the underlying. Instead, we allow the price process to have any form, while satisfying several assumptions. The assumptions include the square-integrability of the price process, substantial risk and bounded mean variance tradeoff. In particular, a geometric Brownian motion satisfies these assumptions. The model can handle European options with payoffs of any form. The interval between trading times is constant, which results in equally spaced rehedging times.

Next we introduce liquidity risk for the underlying asset through the modified price process. It turns out the new process may be used to hedge an option. That is, both the modified and the original price processes yield the same class of admissible strategies.

The optimal strategy must satisfy two conditions: minimize the conditional variance of the cost process increments and turn the cost process into a martingale. The latter is accomplished with another special stochastic process, while for the former I prove the existence of the minimizer to the conditional variance. Next I show the existence of a local risk-minimizing strategy by backward induction and present strategy's structure in terms of the modified price process.

The remainder of this thesis is organized as follows. Chapter 2 advances the results of previous research on approaches to liquidity risk and introduces the Çetin-Jarrow-Protter framework that is the backbone of the model in this thesis. Chapter 3 analyzes the hedging error for delta-hedging strategies. Chapter 4 spells out the application of local risk-minimizing hedges. Chapter 5 describes the conclusions of the thesis and directions for future research.

## 2 Previous results

### 2.1 Literature review

Liquidity risk has profound implications on the derivatives pricing and hedging, since the usual techniques involve trading in the underlying asset. Classical theories of mathematical finance assume it is possible to purchase and/or sell the underlying asset at an average market price, which is not the case in practice. In fact, the prices for buy/sell orders reflected in the bid-ask spread will differ from the quoted market price. On top of the observations above, market may experience a sudden downturn and the underlying asset may not be available at economically reasonable price. A downgrade in a credit rating may trigger requests from the counterparties to post additional margin.

The model of Brunnermeier and Pedersen [6] attempts to make a connection between asset's market liquidity and speculator's funding liquidity. The authors distinguish between three types of market participants. Customers arrive sequentially to the market creating temporary market imbalances. Speculators contribute



liquidity by smoothing the price fluctuations. Financiers set margins and provide funding for the speculators through collateralized borrowing. The paper shows that, under certain conditions, market liquidity and funding liquidity are mutually upholding and may lead to liquidity crises.

There are studies investigating the degree of liquidity of exchange-traded options from the econometric point of view. Christoffersen et al. [12] found that a decrease in option liquidity increases the option price and predicts higher expected option returns. The data also show the degree of liquidity of the underlying stock influences the option return, which is consistent with the delta hedging argument. When the stock market becomes less liquid, the cost of replicating the option increases, which raises the option price and reduces its expected return. Cao and Wei [7] finds strong evidence of liquidity commonality in the option market for such liquidity measures as the bid-ask spread, volume, and price impact. In fact, smaller firms and firms with a higher volatility exhibit stronger commonalities in liquidity. Moreover, the authors show that information asymmetry has a major impact on option liquidity, and option liquidity depends on the stock market's movements.

To reduce overall market repercussion it may be desirable to split a large trade into smaller orders. One is looking to allot certain proportion for the complete order to each respective placement such that the overall price impact is minimized. Problems of this type were investigated in Alfonsi et al. [2]. Instead of specifying a

stock price model incorporating feedback effects, the authors specify the dynamics of the limit order book and provide explicit solutions of the problem.

There are suggestions in Bhaduri et al. [4] to introduce five new derivative instruments to manage liquidity risk in the hedge fund industry. Liquidity options allows investor to withdraw the investment in a tradeable asset at the market price, if the asset liquidity is low as measured by trading volume or by the width of the bid-ask spread. A withdrawal option gives the investor the right to pull out locked-up investment at the market price. A hedge fund return put option allows the buyer to sell the hedge fund investment at a strike to the option seller. Finally a hedge fund return swap and swaption exchange the return of a hedge fund into LIBOR and allow an investor to swap return into LIBOR correspondingly. The authors point out that pricing of these types of derivatives may prove to be quite challenging.

There are several approaches to tie coherent risk measures to liquidity risk framework as introduced in Artzner et al. [3]. In Jarrow and Protter [16] the computation of coherent risk measure with linear supply curve from Çetin et al. [10] results in a simple adjustment, that is multiplication by a factor with constant characterizing the slope of the supply curve. Acerbi and Scandolo [1] introduce marginal supply demand curve which gives the sorted vector, in decreasing (increasing) order, of all the bid (ask) prices available. The curve produces the best bid/ask quotes for

buying/selling certain amount of contracts and is a modified version of the supply curve. The concept of liquidity policy captures the fact that portfolio liquidity risk depends on the funding needs of the portfolio owner. The value function of a portfolio under a liquidity policy ends up being a nonlinear map. The authors suggest a new formalism of “coherent portfolio risk measures” (CPRM) that depends on the liquidity policy and show that CPRMs are convex. The liquidity policy of marking all long positions to the best bids and all short positions to the best offers available (“Uppermost Mark-to-Market Value”) reduces the formalism to exactly to the case of coherent risk measures. Ku [17] demonstrates a method to liquidate acceptable portfolios that satisfy a convex risk measure constraint. It turns out it is possible to liquidate the portfolio consisting of a long stock position and a large negative cash position at some stopping time (finite almost surely) for any initial cash holdings.

In the literature on market liquidity for a specific security there are two approaches for modeling: the effect of a large trader on the underlying (“models of feedback effect”) and the effect of liquidity costs incurred while changing position on a price-taking trader. To model the effect of the dynamic hedging strategies on the equilibrium price of the underlying asset Frey and Stremme [14] use general aggregate demand reaction function that depends on the trader’s exogenous stochastic income. The effect of a large trader is reflected in the volatility. In fact the authors derive an explicit expression for the change in the market volatility. Agents’

expectation establish the amplitude of the feedback effect, since the traders expect the next period price of underlying asset to be some random variable depending on available information.

Many researchers build on the model for market liquidity risk outlined in Çetin et al. [10]. Essentially the spot price of the underlying in the model depends also on the size of the block being traded through the stochastic supply curve. For example if one wants to purchase a large amount of shares of company X right at this moment, there may not be enough supply at the current price, so one will end up paying above the market price for such a big block of shares. The price effect is confined to the very occasion when the order is placed for execution. In Çetin et al. [11] authors use strategies with minimal super-replication cost inclusive of liquidity premium to price contingent claims in continuous time setting. The super-replication price ends up being the unique viscosity solution of the dynamic programming equation. In a discrete-time setting Çetin and Rogers [9] show that the optimization of expected utility of terminal wealth does have a solution even without the hypothesis of absence of arbitrage. Ku et al. [18] derived a partial differential equation which provides discrete-time delta-hedging strategies whose expected hedging error approaches zero almost surely as the length of the revision interval goes to zero. The equation gives the value of the call option from the seller's point of view.

Leland was the first to study the influence of the proportional transaction costs on option pricing via delta-hedging arguments in Leland [20]. Continuous trading of the classic Black-Scholes model becomes infinitely expensive no matter how modest transaction costs might be as a percentage of turnover. Leland's class of discrete time trading strategies is made up of updates adjusted at fixed constant intervals of time and relies on the modified Black-Scholes volatility dependent on the rate of transaction costs and the length of trading intervals. It is possible to approximately replicate the option's payoff as the distance between trading times becomes short. The payoff of the replicating trading strategy converges to the payoff of the option almost surely as the length of rebalance intervals approaches zero.

One may find the extension of Leland's class of discrete time strategies with rehedgings located at varying distances from one another as in Grannan and Swindle [15]. The payoff of the trading strategy with varying rehedging times converges in  $L^2$  to the payoff of the option. The authors also establish the rate of convergence of the replicating strategy to the desired payoff and provide the leading-order term of the mean squared hedging error.

Local risk-minimization in continuous time is introduced in Schweizer [23] as another optimality criterion for hedging strategies. Under five assumptions on the price process for the underlying asset and additional two assumptions yield the existence and uniqueness as well as a method for finding a locally risk-minimizing

trading strategy. The optimal trading strategy is mean-self-financing, that is its cost process is a martingale, and satisfies the stochastic optimality equation.

In Lambertson et al. [19] the authors apply local risk-minimization to a model with proportional transaction costs in discrete time. They show the existence of a locally risk-minimizing trading strategy for every square-integrable contingent claim under the assumptions of substantial risk and bounded mean-variance tradeoff on the price process of the basic asset as well as nondegeneracy requirement on the conditional variances of asset returns.

## 2.2 Çetin-Jarrow-Protter model

We use the ideas from Çetin et al. [10] to model the market liquidity risk for a stock share. Consider a filtered probability space  $[\Omega, \mathcal{F}, \mathcal{F}_{0 \leq t \leq T}, \mathbb{P}]$  satisfying the usual conditions where  $T$  is a fixed time.  $\mathbb{P}$  represents the statistical or empirical probability measure. Assume that  $\mathcal{F}_0$  is trivial, that is  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ . The stochastic process to model the underlying asset had two variables:  $\omega \in \Omega$  and  $t \in \mathbb{R}$ . Now we introduce another variable that may influence the price of the underlying asset for a particular trade: the order size  $x \in \mathbb{R}$ . The price paid for the underlying is  $S(\omega, t, x)$  - stock price, per share, at time  $t$  that the trader pays/receives for an order of size  $x$  given the state  $\omega$ . This is a stochastic supply curve. A positive order ( $x > 0$ ) represent a buy, a negative order ( $x < 0$ ) signifies a sale, and  $x = 0$  yields

the current market price without impact of the order size.

We now impose some structure on the supply curve.

**Assumption 2.1.** (Supply Curve)

1.  $S(\omega, t, x)$  is  $\mathcal{F}_t$ -measurable and non-negative.
2.  $x \mapsto S(\omega, t, x)$  is a.e.  $t$  non-decreasing in  $x$ , a.s. (i.e.  $x \leq y$  implies  $S(\omega, t, x) \leq S(\omega, t, y)$  a.s.  $\mathbb{P}$ , a.e.  $t$ ).
3.  $S$  is  $C^2$  in  $x$ ,  $\partial S(t, x)/\partial x$  and  $\partial^2 S(t, x)/\partial x^2$  are continuous in  $t$ .
4.  $S(\omega, t, 0)$  (process modeling the market price of the underlying asset) is a semi-martingale.
5.  $S(\omega, t, x)$  has continuous sample paths (including time 0) for all  $x$ .

The first assumption says the price of the underlying asset is nonnegative (which is a feature of any time series of prices) and one finds out new prices as time goes by, that is new filtrations  $\mathcal{F}_t$  become available. The second assumption states the larger the block size (or sale), the larger (the smaller) the actual price one has to pay for the order. The third assumption requires smoothness from the price process. The fourth assumption makes sure stochastic integrals with respect to the stochastic process are well-defined. In fact, semi-martingales are the largest class of stochastic processes with respect to which one may define stochastic integrals.

The fifth assumption highlights the fact that the stochastic process is continuous in time.

One way to define  $S(\omega, t, x)$  is the multiplicative supply curve:  $S(\omega, t, x) = S(\omega, t)f(x)$ , where  $S(\omega, t)$  is some stochastic process to model the price of the underlying asset multiplied by the deterministic real-valued increasing function  $f(x)$  with  $f(0) = 1$ . The requirement for  $f(x)$  to be increasing captures the following empirical fact: market participants may not be willing to sell on the spot a large number of shares at the current market price, otherwise arbitrage opportunities arise. For example, the Black Scholes model corresponds to  $f(x) = 1$ . We choose  $f(x) = e^{\alpha x}$ , where  $\alpha$  is estimated using simple regression methodology from the history of stock prices. For this choice of function clearly  $f'(0) = \alpha$ . Details of the supply curve estimation are discussed in Section 3 of Çetin et al. [8]. The value of  $\alpha$  ends up being small, usually with  $0 < \alpha < 0.0001$ . The model is focused on a small trader that cannot move the market by her transactions.

Let  $S(\omega, t, 0) = S_t$  be the marginal price of the supply curve. We denote the supply curve  $S(\omega, t, x) = S_t f(x)$ . That is, we will omit the dependence of the supply curve  $S$  on the  $\omega$  variable altogether, put the time variable  $t$  in the subscript and introduce the dependence on the order size  $x$  via the multiplication by  $f(x)$ .

Generally the supply curve is random. For the sake of illustration we treat variables  $S$  as deterministic and graph the function of the form  $S \times e^{\alpha x}$  in figure 2.1



with relatively large  $\alpha = 0.3$ . Let us assume that the current market price of the underlying asset is 20. If one wants to purchase 10 units of underlying asset, one has to pay approximately 56 per share for the block of 10 shares, that is more than the current market price of 20. Similarly if one wants to sell 10 units of underlying asset (block size of  $-10$ ), one has to accept approximately 7 units of cash per share, as opposed to the market price of 20. This is what we mean by “liquidity risk” in this thesis. Selling a lot of shares at once means one has to accept a lower price per share. Similarly buying a big number of shares at once commands a price per share higher than the market price.

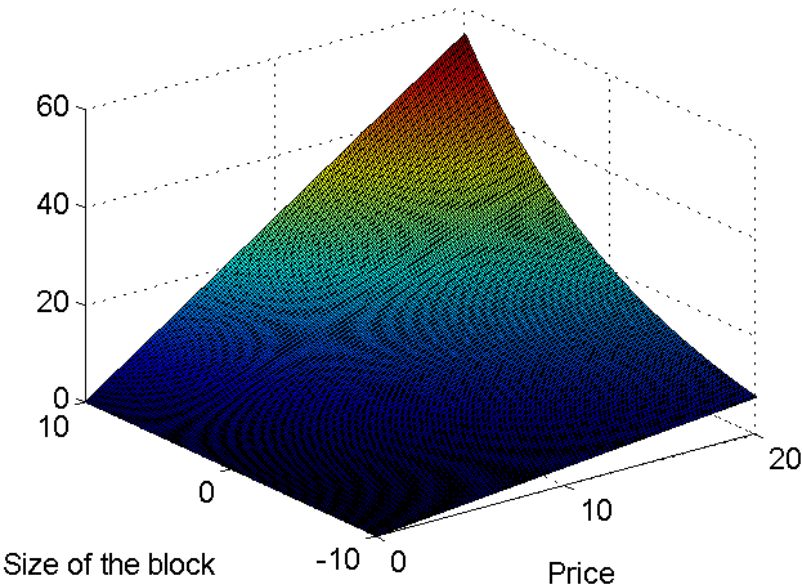


Figure 2.1: An example of a supply curve.

## 2.3 Partial differential equation for option replication

Ku et al. [18] derive the partial differential equation for delta-hedging as well as show the almost sure convergence of the hedging error to zero. Assume the price process  $S_t$  of the underlying asset follows a geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \quad 0 \leq t \leq T, \quad (2.1)$$

with constant drift  $\mu \in \mathbb{R}$ , constant volatility  $\sigma \in \mathbb{R}$  and  $B$  is a standard Brownian motion.

A general *trading strategy* (portfolio) in continuous time is a pair of stochastic processes  $(X_t, Y_t : t \in [0, T])$ , where  $X_t$  stands for trader's aggregate stock holding at time  $t$ ,  $Y_t$  is the trader's aggregate cash position at time  $t$ .  $X_t$  and  $Y_t$  are predictable, while  $X_t$  is optional process.

We introduce discrete time in the model, that is trading is possible only at  $n \in \mathbb{R}$  equally spaces times  $0 = t_0 < t_1 < \dots < t_n = T$  with  $\Delta t = t_i - t_{i-1}$  for  $i = 1, \dots, n$ . We write  $S_i$  for  $S_{t_i}$  and  $X_i$  for  $X_{t_i}$ . We will use index  $T$  for the quantities corresponding to  $T = t_n$ : denote  $S_T$  for  $S_{t_n}$  and  $Y_T$  for  $Y_{t_n}$ .

For the purpose of option replication without liquidity risk we want to start with what the hedger should charge for the option  $V(0, S_0) + X_0 S_0 [f(X_0) - 1]$  in cash ( $Y_0 = V(0, S_0) + X_0 S_0 [f(X_0) - 1]$ ,  $X_0 = 0$ ), then for any path of price process  $S_t$  we will use the cash (and maybe borrow some more at a zero interest rate)

to purchase or sell short some quantity of the underlying asset ( $X_t \neq 0$ ). Here  $V(t, S)$  is the solution to the partial differential equation (2.6). However we should return at the option expiration any money we lent during the life of the trading strategy. That means there should be no “free” money coming into the accounts of the portfolio. By the same token there should be no arbitrary withdrawals of assets from the portfolio. In others words the initial capital of the trading strategy  $Y_0 = V(0, S_0) + X_0 S_0 [f(X_0) - 1]$ , holdings of the underlying asset and the price fluctuations of  $S_t$  determine how much cash is left in the account. In accounting terms we are marking to market cash and underlying asset in the portfolio. This gives rise to the definition of the self-financing strategy.

A *discrete self-financing trading strategy* is a trading strategy  $(X_t, Y_t : t \in \{t_i\}_{i=0}^n)$  which satisfies

$$Y_t = V(0, S_0) + X_0 S_0 [f(X_0) - 1] + \sum_{t_i < t} X_i (S_{i+1} - S_i)$$

Here  $Y_t$  is the payoff of the trading strategy up to time  $t$  without impact of liquidity costs. The term  $V(0, S_0) + X_0 S_0 [f(X_0) - 1]$  corresponds to the cost of setting up initial dynamic delta hedge with impact of liquidity risk. The term  $\sum_{t_i < t} X_i (S_{i+1} - S_i)$  represents trading gains/losses of the trading strategy.

Total liquidity costs for a trading strategy up to time  $T$  are

$$L_T = \sum_{i=1}^n \Delta X_i [S_i f(\Delta X_i) - S_i f(0)] + X_0 [S_0 f(X_0) - S_0 f(0)], \quad (2.2)$$

where  $\Delta X_i = X_i - X_{i-1}$ .

The (total) hedging error of a replicating trading strategy is the difference between the payoff of the strategy  $Y_T$  and the theoretical payoff of the derivative being replicated  $p(S_T)$  less total liquidity costs:

$$\text{(total) hedging error} = V(0, S_0) + X_0 S_0 [f(X_0) - 1] + \sum_{i=0}^{n-1} X_i (S_{i+1} - S_i) - L_T - p(S_T) \quad (2.3)$$

One may also talk about a hedging error on a time interval  $\Delta t$ , that consists of profit/loss from holding  $X$  stock shares for a length of time  $\Delta t$  minus the change in the option value minus liquidity cost over time period  $\Delta t$ .

We want to hedge a European call option with strike  $K$  and expiration  $T$ . Denote the value of the option as a function of time and price of the underlying asset:  $V = V(t, S)$ . We rely on delta-hedging trading strategy to replicate the smoothed (see Theorem 2.1) payoff of the call option  $(S_T - K)^+$ :

$$X_i = V_S(t_i, S_i). \quad (2.4)$$

That is, the construction of hedge comes from the solution to the partial differential equation. We use delta-hedging to partially follow the argument of Leland [20]. Trying out other hedging schemes, possibly with a lower initial cost, may be one direction of future research. The price demanded by the hedger, that is the initial cost of the portfolio, is  $V(0, S_0) + X_0 S_0 [f(X_0) - 1]$ . There is no liquidation of the

self-financing portfolio at the expiration, instead one delivers  $1_{S_T > K}$  units of stock and  $-K1_{S_T > K}$  units of bond.

We will need the following smoothness condition on  $V(t, S)$ :

$$\|V\|_{m,n,p} = \sup_{S \geq 0, 0 \leq t \leq T} \left[ S^m \frac{\partial^{n+p} V(t, S)}{\partial S^n \partial t^p} \right] \quad (2.5)$$

is finite for all nonnegative  $m, n$  and  $p$ .

**Theorem 2.1.** *Let  $V(t, S)$  be the solution of the partial differential equation*

$$V_t(t, S) + \frac{1}{2} \sigma^2 S^2 V_{SS}(t, S) + f'(0) \sigma^2 S^3 V_{SS}^2(t, S) = 0 \quad (2.6)$$

*with the (smoothed) terminal condition*

$$V(T, S) = p(S_T),$$

*where  $p$  is the value at time  $T$  of a call with maturity  $T + \epsilon$ ,  $V(t, S)$  satisfies the smoothness condition (2.5). Then expected hedging error of a replicating trading strategy given by (2.4) over the interval  $[0, T]$  approaches zero almost surely as  $\Delta t$  goes to zero.*

*Proof.* We will represent the total hedging error as a sum of errors over equally spaced time intervals. We consider the hedging error over the small time interval  $[t_{i-1}, t_i]$  for  $i = 1, 2, \dots, n$ . The change in the call option value is

$$\Delta V = V(t + \delta t, S + \Delta S) - V(t, S) = V_S \Delta S + V_t \Delta t + \frac{1}{2} V_{SS} (\Delta S)^2 + O(\Delta t^{3/2}),$$

and the change in the current market value of the hedging strategy is according to

(2.4)

$$\Delta X = V_S(t + \delta t, S + \Delta S) - V_S(t, S) = V_{SS}\Delta S + V_{St}\Delta t + \frac{1}{2}V_{SSS}(\Delta S)^2 + O(\Delta t^{3/2}).$$

From (2.2), the liquidity cost at each interval is

$$\Delta X[Sf(\Delta X) - Sf(0)] = \Delta X(f(\Delta X) - 1)S \quad (2.7)$$

since  $f(0) = 1$ .

Taylor series expansion gives

$$f(\Delta X) - 1 = f'(0)\Delta X + \frac{1}{2}f''(0)(\Delta X)^2 + O(\Delta X^3).$$

Next we write down the discrete time version of the geometric Brownian motion

(2.1)

$$\begin{aligned} \Delta S &= \mu S \Delta t + \sigma Z S \sqrt{\Delta t} \\ (\Delta S)^2 &= S^2 \sigma^2 Z^2 \Delta t + O(\Delta t^{3/2}) \\ (\Delta S)^k &= O(\Delta t^{3/2}), \quad k = 3, 4, 5, \dots \end{aligned}$$

where  $Z$  is a standard normal random variable. Therefore the change in the market

value of the hedging strategy takes the form:

$$\begin{aligned}
\Delta X &= V_{SS}\Delta S + V_{St}\Delta t + \frac{1}{2}V_{SSS}(\Delta S)^2 + O(\Delta t^{3/2}) \\
&= V_{SS}S(\mu\Delta t + \sigma Z\sqrt{\Delta t}) + V_{St}\Delta t + \frac{1}{2}V_{SSS}S^2\sigma^2Z^2\Delta t + O(\Delta t^{3/2}), \\
(\Delta X)^2 &= V_{SS}^2(\Delta S)^2 + O(\Delta t^{3/2}) = V_{SS}^2S^2\sigma^2Z^2\Delta t + O(\Delta t^{3/2}), \\
(\Delta X)^k &= O(\Delta t^{3/2}), \quad k = 3, 4, 5, \dots
\end{aligned}$$

We rewrite the equation (2.7) for the liquidity cost per time interval using Taylor expansion for  $f(\Delta X) - 1$  as well as expressions for  $\Delta X$  and its powers:

$$\begin{aligned}
\Delta X[Sf(\Delta X) - Sf(0)] &= \Delta X(f(\Delta X) - 1)S \\
&= \Delta X \left( f'(0)\Delta X + \frac{1}{2}f''(0)(\Delta X)^2 \right) S + O(\Delta t^{3/2}) \\
&= f'(0)V_{SS}^2\sigma^2Z^2S^3\Delta t + O(\Delta t^{3/2}).
\end{aligned}$$

Therefore, the hedging error over each revision interval is

$$\begin{aligned}
\Delta H &= X\Delta S - \Delta V - \Delta X(Sf(\Delta X) - Sf(0)) \\
&= -V_t\Delta t - \frac{1}{2}V_{SS}(\Delta S)^2 - f'(0)V_{SS}^2\sigma^2Z^2S^3\Delta t + O(\Delta t^{3/2}).
\end{aligned}$$

If  $V$  satisfies the following partial differential equation

$$V_t + \frac{1}{2}\sigma^2S^2V_{SS} (1 + 2f'(0)SV_{SS}) = 0,$$

then

$$\mathbb{E}[\Delta H] = O(\Delta t^{3/2}).$$

The total hedging error over the entire interval  $[0, T]$  is the sum of individual  $\Delta H$ 's according to the definition (2.3), thus

$$\mathbb{E}[\sum \Delta H] = \sum \mathbb{E}[\Delta H] = \sum O(\Delta t^{3/2}) = O(\Delta t^{1/2}),$$

that is the expected hedging error over the period  $[0, T]$  approaches zero almost surely as  $\Delta t$  becomes small.

□

The following theorem shows the almost sure convergence of the strategy payoff to the option payoff. So the delta hedging strategy constructed with the help of the partial differential equation is optimal, in a sense that total hedging error approaches zero almost surely at the length of the interval between trading times goes to zero.

**Theorem 2.2.** *The payoff of the discrete time delta-hedging strategy  $X = V_S$  where  $V$  is the solution of the partial differential equation (2.6) converge almost surely to the smoothed payoff  $(S_T - K)^+$  of the call option including liquidity costs as  $\Delta t \rightarrow 0$ .*

Of course the argument and methodology presented in the section work well for general European contingent claims. In the next chapter we will show an analog of these two theorems for  $L^2$  convergence as opposed to almost sure convergence. The proportional transaction cost models produce linear or quasilinear partial dif-



ferential equations of the second order, while liquidity costs result in fully nonlinear partial differential equation (2.6).

### 3 Delta-hedging replicating strategies

We use the partial differential equation from Ku et al. [18] to construct replicating trading strategies. In Theorem 3.4 I show that the payoff of the replicating trading strategy converges to the payoff of the option in  $L^2$ . Then in Theorem 3.5 I find the exact first order term of the Taylor expansion with respect to the time interval length  $\Delta t$  of the mean squared hedging error. Finally, I find the optimal positioning of the rebalancing times by minimizing the first order term with respect to the deterministic “distance function”  $d(t)$ . All convergence statements between random variables are in  $L^2$  unless otherwise stated.

#### 3.1 Convergence of the mean squared replication error

Consider a filtered probability space  $[\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P}]$  satisfying the usual conditions, where  $T$  is a fixed time, and  $\mathbb{P}$  represents the statistical or empirical probability measure. We consider a market with a stock and a money market account. We assume the stock pays no dividends, and the rate of interest is zero.

Let  $S(\omega, t, x)$  represent the stock price per share at time  $[0, T]$  that a trader pays/receives for an order of size  $x \in \mathbb{R}$ . A positive order ( $x > 0$ ) represents a buy, a negative order ( $x < 0$ ) signifies a sale, and  $x = 0$  corresponds to the marginal trade. For the detailed structure of the supply curve, we refer the reader to Section 2 of Çetin et al. [10]. This model is focused on a small trader that cannot move the market by her transactions. Let  $S(\omega, t, 0) = S_t$  be the marginal price of the supply curve. We denote the multiplicative supply curve  $S(\omega, t, x) = S_t f(x)$  for some deterministic real-valued increasing function with  $f(0) = 1$ .

Assume the price process  $S_t$  follows a geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \quad 0 \leq t \leq T,$$

where the drift  $\mu$  is a constant, the volatility  $\sigma$  is a positive real number,  $B$  is a standard Brownian motion, and  $T$  is the terminal time of a European contingent claim with the payoff  $p(S_T)$ . Take  $f(x) = e^{\alpha x}$ . The slope at  $x = 0$  of supply curve,  $f'(0)$  (denoted by  $\alpha$ ) is interpreted as the parameter for liquidity risk. The parameter  $\alpha$  is estimated using simple regression methodology from the history of stock prices. Details of the supply curve estimation are discussed in Section 3 of Çetin et al. [8]. The value of  $\alpha$  turns out to be small in most cases, usually with  $0 < \alpha < 0.0001$ .

**Definition 3.1.** A (replicating) trading strategy is a pair  $(X_t, Y_t : t \in [0, T])$  where

$X_t$  represents the number of units of stock held at time  $t$  and  $Y_t$  represents the trader's cash holding at time  $t$  the trader would have, if there weren't any liquidity costs.

Here  $X_t$  and  $Y_t$  are predictable and optional (stochastic) processes respectively, with respect to the filtration  $(\mathcal{F}_t)_{0 \leq t \leq T}$ .

Let us consider equally spaced times  $0 = t_0 \leq t_1 \leq t_2 \leq \dots \leq t_n = T$ . Set  $\Delta t = t_i - t_{i-1}$  for  $i = 1, \dots, n$ . We will write  $S_{t_k} = S_k$ ,  $X_{t_k} = X_k$ . The following definition describes a self-financing condition for discrete strategies with liquidity risk.

**Definition 3.2.** *A discrete self-financing trading strategy (without liquidity costs)*

*is a trading strategy  $(X_t, Y_t : t \in \{t_i\}_{i=0}^n)$  which satisfies*

$$Y_t = V(0, S_0) + X_0 S_0 [f(X_0) - 1] + \sum_{t_i < t} X_{t_i} (S_{i+1} - S_i)$$

where  $\Delta X_i = X_i - X_{i-1}$  ( $i \geq 1$ ) and  $\Delta X_0 = X_0$ .

Considering the usual self-financing condition without liquidity costs, it is natural to define the total accumulated liquidity costs (up to time  $t < t_n = T$ ) of a discrete trading strategy  $(X_t, Y_t : t \in \{t_i\}_{i=0}^n)$  by

$$L_t = \sum_{t_i < t} \Delta X_i [S_i f(\Delta X_i) - S_i f(0)] + X_0 S_0 [f(X_0) - 1]$$

and the total liquidity costs up to time  $T$  is

$$L_T = \sum_{i=1}^n \Delta X_i [S_i f(\Delta X_i) - S_i f(0)] + X_0 S_0 [f(X_0) - 1]$$

where  $\Delta X_i = X_i - X_{i-1}$  for  $i \geq 1$  and  $\Delta X_0 = X_0$ .

In practice, continuous hedging is not possible, thus one cannot replicate the option perfectly and must accept an error. Letting  $V(0, S_0) + X_0 S_0 [f(X_0) - 1]$  denote the “true” cash position in the hedge during the interval  $(t_0, t_1)$ , the hedging error inclusive of liquidity costs is

$$HE = V(0, S_0) + X_0 S_0 [f(X_0) - 1] + \sum_{i=0}^{n-1} X_i (S_{i+1} - S_i) - L_T - p(S_T)$$

Recall that for delta-hedging approach the option premium is  $V(0, S_0) + X_0 S_0 [f(X_0) - 1]$ , that is initial capital of a replicating trading strategy.

Hedging errors are random, so we consider distributions of hedging errors. We compare efficiency of the different replicating trading strategies by looking at the corresponding distributions of the hedging errors and their parameters. We use the mean squared hedging error for a criterion to estimate the hedging error.

**Definition 3.3.** *The mean squared hedging error (MSHE) of a replicating trading strategy is*

$$\mathbb{E} \left[ V(0, S_0) + X_0 S_0 [f(X_0) - 1] + \sum_{i=0}^{n-1} X_i (S_{i+1} - S_i) - L_T - p(S_T) \right]^2$$

Not only did Black and Scholes show how to compute the fair price of the European option, but also they laid out the method for eliminating the risk of writing an option through continuous delta hedging. Implementation of continuous delta-hedging amounts to maintaining and rebalancing two positions with a money market account (in cash) and a stock. The goal of delta-hedging is to eliminate the risk of writing an option completely (continuous hedging in theory) or at least significantly reduce the level of risk (discrete hedging in practice).

Leland [20] investigated the hedging error over each revision interval in the presence of transaction costs, and modified the parameter of Black-Scholes price for delta-hedging strategies. Ku et al. [18] argue that a dynamic delta hedging according to their partial differential equation for a European contingent claim produces hedging errors over the period  $[0, T]$  whose expectation approaches 0 almost surely as the length of the revision interval goes to 0. They also show that the payoff of the discrete replicating trading strategy converges almost surely to the terminal payoff of the option  $p(S_T)$ . In these papers the trading times are equally spaced over the life of the option.

Considering alternatives to equally spaced trading times over the life of an option, the hedging error can be improved. We proceed to work with trading strategies with varying rehedging times. We parametrize varying rebalancing times with a smooth, positive, strictly increasing function  $d(t)$  via  $t_i = d^{-1}(i\Delta t)$ . We also

require that  $d(0) = 0, d(T) = T$  and  $d'(t) \neq 0$ . In other words one may recover the positioning of rehedging times through inverse of  $d(t)$ . Taking  $d(t) = t$  corresponds to the constant interval case, equally spaced trading dates. Different functions  $d(t)$  yield different locations of the rehedging times  $t_i$ . Figure 3.1 shows an example of  $d(t)$ , where the left half of the graph is concave down (which translates into more frequent rehedging at the beginning) and the right half of the graph is convex (which corresponds to more frequent rehedging toward the end of a period).

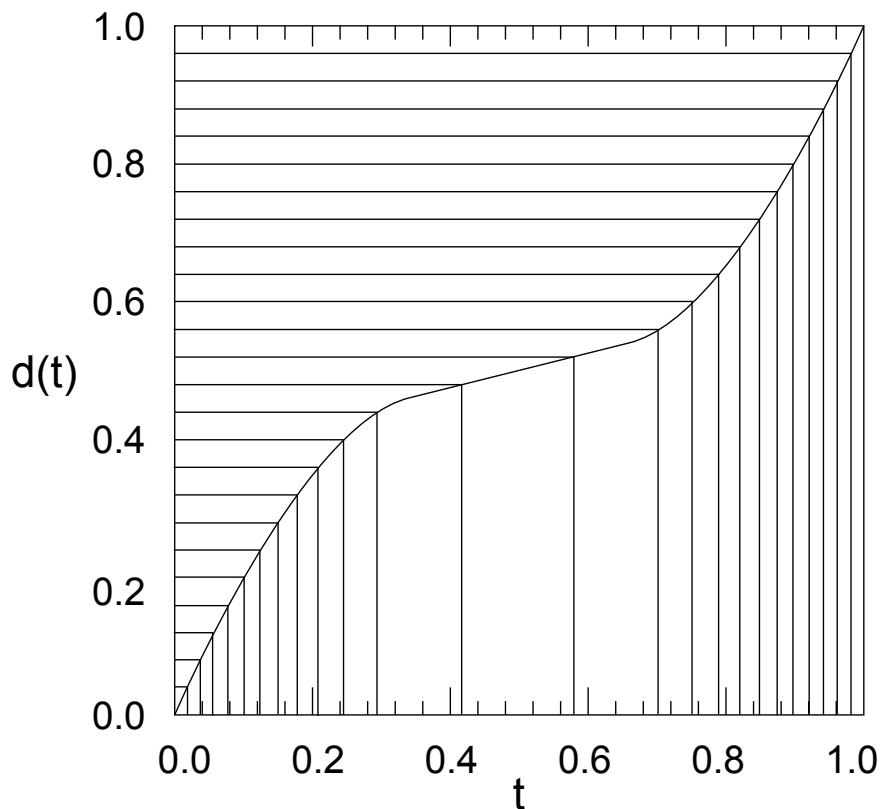


Figure 3.1: Setting up the varying rehedging times. Steeper slope of  $d(t)$  corresponds to more frequent rebalancing.

**Theorem 3.1.** *Let  $V(t, S)$  be a solution to the partial differential equation (2.6) with the final condition  $V(T, S) = p(S_T)$  and satisfies the smoothness condition (2.5), then employing delta-hedging trading strategy, that is  $X_i = V_S(t_i, S_i)$ , results in the following convergence statements:*

$$\sum_{t_i < T} V_S(t_i, S_i)(S_{i+1} - S_i) \rightarrow \int_0^T V_S(t, S_t) dS_t$$

$$L_T \rightarrow \int_0^T \sigma^2 f'(0) S_t^3 V_{SS}(t, S_t) dt + X_0 S_0 [f(X_0) - 1]$$

in  $L^2$  as  $\Delta t \rightarrow 0$ .

*Proof.* We show the convergence in  $L^2$  by explicitly calculating the expected square error and verifying that the expected square error is indeed equal to zero. Recall that a sequence of the random variables  $X_1, X_2, \dots$  converges to the random variable  $X$  in  $L^2$  if  $E[(X_n - X)^2] \rightarrow 0$  as  $n \rightarrow \infty$ . With finite number of trading times  $n$  the fact  $n \rightarrow \infty$  is equivalent to saying that the length of the time interval between the trading times converges to zero:  $\Delta t \rightarrow 0$ . The main tools are Taylor expansion and properties of the geometric Brownian motion.

**Lemma 3.2.**

$$\sum_{t_i < T} V_S(t_i, S_i)(S_{i+1} - S_i) \rightarrow \int_0^T V_S(t, S_t) dS_t$$

in  $L^2$  as  $\Delta t \rightarrow 0$ .



*Proof.* Consider

$$\begin{aligned}
& \mathbb{E} \left[ \sum_{t_i < T} V_S(t_i, S_i)(S_{i+1} - S_i) - \int_0^T V_S(t, S_t) dS_t \right]^2 \\
&= \mathbb{E} \left[ \sum_{t_i < T} \int_{t_i}^{t_{i+1}} V_S(t_i, S_i) dS_t - \sum_{t_i < T} \int_{t_i}^{t_{i+1}} V_S(t, S_t) dS_t \right]^2 \\
&= \mathbb{E} \left[ \sum_{t_i < T} \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t)) dS_t \right]^2 \\
&= \mathbb{E} \left[ \sum_i \left( \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t)) dS_t \right)^2 \right] \\
&\quad + \mathbb{E} \left[ \sum_{i \neq j} \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t)) dS_t \int_{t_j}^{t_{j+1}} (V_S(t_j, S_{t_j}) - V_S(t, S_t)) dS_t \right]
\end{aligned}$$

The cross terms (the second term of the last equation) become zero. We work with the diagonal terms first.

$$\begin{aligned}
& \mathbb{E} \left[ \sum_i \left( \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t)) dS_t \right)^2 \right] \\
&= \mathbb{E} \left[ \sum_i \left( \int_{t_i}^{t_{i+1}} [V_S(t_i, S_i) - V_S(t, S_t)] [\mu S_t dt + \sigma S_t dW_t] \right)^2 \right] \\
&= \mathbb{E} \left[ \sum_i \left( \int_{t_i}^{t_{i+1}} [V_S(t_i, S_i) - V_S(t, S_t)] \mu S_t dt + \int_{t_i}^{t_{i+1}} [V_S(t_i, S_i) - V_S(t, S_t)] \sigma S_t dW_t \right)^2 \right]
\end{aligned} \tag{3.1}$$

Now we demonstrate that the order of each sum is greater than or equal to  $O(\Delta t)$  to conclude that the whole expression converges to zero. We start by showing that

the sum of squares of the second term has expectation  $O(\Delta t)$ , that is

$$\mathbb{E} \left[ \sum_i \left( \int_{t_i}^{t_{i+1}} [V_S(t_i, S_i) - V_S(t, S_t)] \sigma S_t dW_t \right)^2 \right] = O(\Delta t) \rightarrow 0 \quad (\text{as } \Delta t \rightarrow 0)$$

To prove this we consider

$$\begin{aligned} & \mathbb{E} \left[ \sum_i \left( \int_{t_i}^{t_{i+1}} [V_S(t_i, S_i) - V_S(t, S_t)] \sigma S_t dW_t \right)^2 \right] \\ &= \mathbb{E} \left[ \sum_i \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} [V_S(t_i, S_i) - V_S(t, S_t)] \sigma S_t dW_t \right]^2 \right] \\ &= \mathbb{E} \left[ \sum_i \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t))^2 \sigma^2 S_t^2 dt \right] \right] \end{aligned}$$

by the Ito isometry (see for instance Shreve [24]). Using a Taylor series expansion,

$$\begin{aligned} & \mathbb{E} \left[ \sum_i \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} \left( V_{SS}^2(t_i, S_i) (S_t - S_i)^2 + O(\Delta t^{3/2}) \right) \sigma^2 S_t^2 dt \right] \right] \\ &= \mathbb{E} \left[ \sum_i \left( \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} V_{SS}^2(t_i, S_i) (S_t - S_i)^2 \sigma^2 S_t^2 dt \right] + O(\Delta t^{5/2}) \right) \right] \\ &= \mathbb{E} \left[ \sum_i \left( \sigma^2 V_{SS}^2(t_i, S_i) \int_{t_i}^{t_{i+1}} \mathbb{E}_{\mathcal{F}_{t_i}} [(S_t - S_i)^2 S_t^2] dt + O(\Delta t^{5/2}) \right) \right] \quad (3.2) \end{aligned}$$

We note that the fact

$$\mathbb{E}_{\mathcal{F}_{t_i}} [(S_t - S_i)^2 S_t^2] = \sigma^2 S_i^4 (t - t_i) + O(t - t_i)^2$$

which can be directly calculated from the distribution of  $S_t$ . Then equation (3.2)

equals

$$\begin{aligned}
& \mathbb{E} \left[ \sum_i \left( \sigma^2 V_{SS}^2(t_i, S_i) \int_{t_i}^{t_{i+1}} \mathbb{E}_{\mathcal{F}_{t_i}} [(S_t - S_i)^2 S_t^2] dt + O(\Delta t^{5/2}) \right) \right] \\
&= \mathbb{E} \left[ \sum_i \left( \sigma^2 V_{SS}^2(t_i, S_i) \left[ \sigma^2 S_i^4 \int_{t_i}^{t_{i+1}} (t - t_i) dt + O(\Delta t^3) \right] + O(\Delta t^{5/2}) \right) \right] \\
&= \mathbb{E} \left[ \sum_i \left( \frac{\sigma^4}{2} V_{SS}^2(t_i, S_i) S_i^4 (t_{i+1} - t_i)^2 + O(\Delta t^{5/2}) \right) \right] \\
&= \mathbb{E} \left[ \sum_i \left( O(\Delta t^2) + O(\Delta t^{5/2}) \right) \right] = O(\Delta t) \rightarrow 0 \quad (\text{as } \Delta t \rightarrow 0)
\end{aligned}$$

This is the only term with order exactly  $O(\Delta t)$ . To characterize the convergence

we find that

$$\begin{aligned}
& \mathbb{E} \left[ \sum_i \frac{\sigma^4}{2} V_{SS}^2(t_i, S_i) S_i^4 (t_{i+1} - t_i)^2 \right] \\
&= \mathbb{E} \left[ \sum_i \frac{\sigma^4}{2} V_{SS}^2(t_i, S_i) S_i^4 (t_{i+1} - t_i) (t_{i+1} - t_i) \right] \\
&= \mathbb{E} \left[ \sum_i \frac{\sigma^4}{2} V_{SS}^2(t_i, S_i) S_i^4 \left[ \frac{\Delta t}{d'(t_i)} + O(\Delta t^2) \right] (t_{i+1} - t_i) \right] \\
&= \mathbb{E} \left[ \sum_i \frac{\sigma^4}{2} V_{SS}^2(t_i, S_i) S_i^4 \frac{\Delta t}{d'(t_i)} (t_{i+1} - t_i) \right] + \sum_i O(\Delta t^3) \\
&= \Delta t \mathbb{E} \left[ \sum_i \frac{\sigma^4}{2} V_{SS}^2(t_i, S_i) S_i^4 \frac{1}{d'(t_i)} (t_{i+1} - t_i) \right] + O(\Delta t^2)
\end{aligned}$$

We also note that

$$\mathbb{E} \left[ \sum_i \frac{\sigma^4}{2} V_{SS}^2(t_i, S_i) S_i^4 \frac{1}{d'(t_i)} (t_{i+1} - t_i) \right] \rightarrow \mathbb{E} \left[ \int_0^T \frac{\sigma^4}{2} V_{SS}^2(t, S_t) S_t^4 \frac{1}{d'(t)} dt \right]$$

in  $L^1$ . We claim that the remaining terms from the equation will be higher order than  $O(\Delta t)$  and thus converge to zero as  $\Delta t$  tends to zero. We give the proof for the cross term, and the reasoning for the other terms is similar. Consider the expected value conditioned on  $\mathcal{F}_{t_i}$  of the cross term:

$$\begin{aligned} & \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t)) \mu S_t dt \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t)) \sigma S_t dW_t \right] \\ & \leq \left( \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t)) \mu S_t dt \right]^2 \right)^{1/2} \\ & \quad \times \left( \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t)) \sigma S_t dW_t \right]^2 \right)^{1/2} \end{aligned}$$

using Hölder's inequality. Then by the Ito isometry

$$\begin{aligned} & \left( \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t)) \mu S_t dt \right]^2 \right)^{1/2} \\ & \left( \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} (V_S(t_i, S_i) - V_S(t, S_t))^2 \sigma^2 S_t^2 dt \right] \right)^{1/2} \\ & = \left( \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} V_{SS}(t_i, S_i) (S_t - S_i) \mu S_t dt + O(\Delta t^2) \right]^2 \right)^{1/2} \\ & \quad \times \left( \mathbb{E}_{\mathcal{F}_{t_i}} \left[ \int_{t_i}^{t_{i+1}} V_{SS}^2(t_i, S_i) (S_t - S_i)^2 \sigma^2 S_t^2 dt + O(\Delta t^3) \right] \right)^{1/2} \\ & = O(\Delta t^{3/2}) O(\Delta t) = O(\Delta t^{5/2}) \end{aligned}$$

Then summing up the terms over all subintervals gives an estimate of

$$\mathbb{E} \left[ \sum_i O(\Delta t^{5/2}) \right] = O(\Delta t^{3/2}) > O(\Delta t).$$

□

**Lemma 3.3.**  $L_T \rightarrow \int_0^T \sigma^2 f'(0) S_t^3 V_{SS}^2(t, S_t) dt + X_0 S_0 [f(X_0) - 1]$  in  $L^2$  as  $\Delta t \rightarrow 0$ .

*Proof.* Define the following sequence of random variables

$$\Delta_i \equiv (S_i - S_{i-1})^2 - \sigma^2 S_{i-1}^2 (t_i - t_{i-1})$$

After rearranging one gets

$$(S_i - S_{i-1})^2 = \Delta_i + \sigma^2 S_{i-1}^2 (t_i - t_{i-1})$$

Let us consider

$$\begin{aligned} & \sum_{i=1}^n V_{SS}^2(t_{i-1}, S_{i-1}) (S_i - S_{i-1})^2 S_{i-1} f'(0) + X_0 S_0 [f(X_0) - 1] \\ &= \sum_{i=1}^n V_{SS}^2(t_{i-1}, S_{i-1}) (\Delta_i + \sigma^2 S_{i-1}^2 (t_i - t_{i-1})) S_{i-1} f'(0) + X_0 S_0 [f(X_0) - 1] \\ &= \sum_{i=1}^n V_{SS}^2(t_{i-1}, S_{i-1}) \Delta_i S_{i-1} f'(0) + \sum_{i=1}^n V_{SS}^2(t_{i-1}, S_{i-1}) \sigma^2 (t_i - t_{i-1}) S_{i-1}^3 f'(0) \\ &+ X_0 S_0 [f(X_0) - 1] \end{aligned}$$

It is easy to see that the second sum converges to  $\int_0^T \sigma^2 f'(0) S_t^3 V_{SS}^2(t, S_t) dt$ . We will show that for the first sum

$$\sum_{i=1}^n V_{SS}^2(t_{i-1}, S_{i-1}) \Delta_i S_{i-1} f'(0) \rightarrow 0$$

We begin by observing that  $\mathbb{E} \left[ \sum_{i=1}^n V_{SS}^2(t_{i-1}, S_{i-1}) \Delta_i S_{i-1} f'(0) \right]^2$  consists of diagonal and cross terms. Consider the diagonal terms first:

$$\begin{aligned} & \mathbb{E} \left[ \sum_{i=1}^n V_{SS}^4(t_{i-1}, S_{i-1}) \Delta_i^2 S_{i-1}^2 (f'(0))^2 \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^n V_{SS}^4(t_{i-1}, S_{i-1}) S_{i-1}^2 (f'(0))^2 \mathbb{E}_{\mathcal{F}_{t_{i-1}}} [\Delta_i^2] \right] \end{aligned} \quad (3.3)$$

Recall that  $S_i = S_{i-1} e^{Z_t - Z_{t_{i-1}}}$ , where  $Z_t = \bar{\mu}t + \sigma W_t$  is a Wiener process with drift  $\bar{\mu} = \mu - \frac{1}{2}\sigma^2$ . Since  $Z_t - Z_{t_{i-1}} \sim \mathcal{N}(\bar{\mu}(t_i - t_{i-1}), \sigma^2(t_i - t_{i-1}))$ , then the moment generating function of  $Z_t - Z_{t_{i-1}}$  is  $M_{\mathcal{N}}(s) = e^{\bar{\mu}(t_i - t_{i-1})s + \sigma^2(t_i - t_{i-1})s^2/2}$ .

Then

$$\begin{aligned} & \mathbb{E}_{\mathcal{F}_{t_{i-1}}} [\Delta_i^2] \\ &= \mathbb{E}_{\mathcal{F}_{t_{i-1}}} [(S_i - S_{i-1})^2 - \sigma^2 S_{i-1}^2 (t_i - t_{i-1})]^2 \\ &= \mathbb{E}_{\mathcal{F}_{t_{i-1}}} [(S_i - S_{i-1})^4 - 2(S_i - S_{i-1})^2 \sigma^2 S_{i-1}^2 (t_i - t_{i-1}) + \sigma^4 S_{i-1}^4 (t_i - t_{i-1})^2] \\ &= S_{i-1}^4 [3\sigma^4 (t_i - t_{i-1})^2 - 2(1 + 2\bar{\mu}(t_i - t_{i-1}) + 2\sigma^2(t_i - t_{i-1}) - 2[1 + \bar{\mu}(t_i - t_{i-1}) \\ &\quad + 1/2\sigma^2(t_i - t_{i-1})] + 1)(t_i - t_{i-1})\sigma^2 + (t_i - t_{i-1})^2\sigma^4 + O(\Delta t^3)] \\ &= 2S_{i-1}^4 \sigma^4 (t_i - t_{i-1})^2 + O(\Delta t^3) \end{aligned}$$

Use this result to rewrite (3.3) as

$$\begin{aligned}
& \mathbb{E} \left[ \sum_{i=1}^n V_{SS}^4(t_{i-1}, S_{i-1}) S_{i-1}^2 (f'(0))^2 \mathbb{E}_{\mathcal{F}_{t_{i-1}}} [\Delta_i^2] \right] \\
&= \mathbb{E} \left[ \sum_{i=1}^n V_{SS}^4(t_{i-1}, S_{i-1}) S_{i-1}^2 (f'(0))^2 (2S_{i-1}^4 \sigma^4 (t_i - t_{i-1})(t_i - t_{i-1}) + O(\Delta t^3)) \right] \\
&= 2\mathbb{E} \left[ \sum_{i=1}^n V_{SS}^4(t_{i-1}, S_{i-1}) S_{i-1}^6 (f'(0))^2 \sigma^4 \left( \frac{\Delta t}{d'(t_i)} + O(\Delta t^2) \right) (t_i - t_{i-1}) \right] + O(\Delta t^2) \\
&= 2\Delta t \mathbb{E} \left[ \sum_{i=1}^n V_{SS}^4(t_{i-1}, S_{i-1}) S_{i-1}^6 (f'(0))^2 \sigma^4 \frac{1}{d'(t_i)} (t_i - t_{i-1}) + \sum_{i=1}^n O(\Delta t^2) \right] + O(\Delta t^2) \\
&= 2\Delta t \mathbb{E} \left[ \sum_{i=1}^n \int_{t_{i-1}}^{t_i} V_{SS}^4(t_{i-1}, S_{i-1}) S_{i-1}^6 (f'(0))^2 \sigma^4 \frac{1}{d'(t_i)} dt \right] + O(\Delta t^2) \rightarrow 0 \quad (\text{as } \Delta t \rightarrow 0)
\end{aligned}$$

This is the only term with the order exactly  $O(\Delta t)$ . To characterize the convergence we note that

$$\begin{aligned}
& \mathbb{E} \left[ \sum_{i=1}^n \int_{t_{i-1}}^{t_i} V_{SS}^4(t_{i-1}, S_{i-1}) S_{i-1}^6 (f'(0))^2 \sigma^4 \frac{1}{d'(t_i)} dt \right] \\
& \rightarrow \mathbb{E} \left[ \int_0^T V_{SS}^4(t, S_t) S_t^6 (f'(0))^2 \sigma^4 \frac{1}{d'(t)} dt \right]
\end{aligned}$$

in  $L^1$ . Therefore, we conclude that

$$\begin{aligned}
& \sum_{i=1}^n V_{SS}^2(t_{i-1}, S_{i-1}) (S_i - S_{i-1})^2 S_{i-1} f'(0) + X_0 S_0 [f(X_0) - 1] \\
&= \sum_{i=1}^n V_{SS}^2(t_{i-1}, S_{i-1}) \Delta_i S_{i-1} f'(0) + \sum_{i=1}^n V_{SS}^2(t_{i-1}, S_{i-1}) \sigma^2 (t_i - t_{i-1}) S_{i-1}^3 f'(0) \\
&+ X_0 S_0 [f(X_0) - 1] \\
&\rightarrow \int_0^T V_{SS}^2(t, S_t) S_t^3 \sigma^2 f'(0) dt + X_0 S_0 [f(X_0) - 1]
\end{aligned}$$

Finally, we need to show

$$L_T - \sum_{i=1}^n V_{SS}^2(t_{i-1}, S_{i-1}) S_{i-1} f'(0) (S_i - S_{i-1})^2 - X_0 S_0 [f(X_0) - 1] \rightarrow 0$$

We omit higher-order term on the way. First we work with

$$\begin{aligned} L_T &\equiv \sum_{t_i < T} \Delta X_i [S_i f(\Delta X_i) - S_i f(0)] + X_0 S_0 [f(X_0) - 1] \\ &= \sum_{t_i < T} \Delta X_i [f(\Delta X_i) - 1] S_i + X_0 S_0 [f(X_0) - 1] \\ &= \sum_{t_i < T} \Delta X_i^2 f'(0) S_i + \sum_{t_i < T} O(\Delta X_{t_i})^3 + X_0 S_0 [f(X_0) - 1] \\ &= \sum_{i=1}^n f'(0) S_i V_{SS}^2(t_{i-1}, S_{i-1}) (S_i - S_{i-1})^2 + \sum_{t_i < T} O(\Delta X_{t_i})^3 + X_0 S_0 [f(X_0) - 1] \end{aligned}$$

Now we turn our attention to

$$\begin{aligned} &\mathbb{E} \left[ L_T - \sum_{i=1}^n f'(0) S_{i-1} V_{SS}^2(t_{i-1}, S_{i-1}) (S_i - S_{i-1})^2 - X_0 S_0 [f(X_0) - 1] \right]^2 \\ &= \mathbb{E} \left[ \sum_{i=1}^n f'(0) V_{SS}^2(t_{i-1}, S_{i-1}) (S_i - S_{i-1})^3 + \sum_{t_i < T} O(\Delta X_{t_i})^3 \right]^2 = O(\Delta t^2) \rightarrow 0 \end{aligned}$$

Therefore,

$$L_T \rightarrow \int_0^T \sigma^2 f'(0) S_t^3 V_{SS}^2(t, S_t) dt + X_0 S_0 [f(X_0) - 1]$$

The proof of the lemma is finished.  $\square$

Lemmas 3.2 and 3.3 show the convergence statements given in the Theorem 3.1.

This ends the proof of the Theorem 3.1.  $\square$



**Remark 3.1.** *The PDE in 3.1 was introduced in Ku et al. [18] and the well-posedness issues such as the existence of smooth solutions have not yet been fully resolved. However, the existence and uniqueness of solutions in the viscosity sense can be shown for options with convex payoffs, by comparing to equation (1.1) in Çetin et al. [11] with  $l = \frac{1}{4\alpha S}$  and adopting the result therein.*

**Theorem 3.4.** *Under the conditions of Theorem 3.1, we have*

$$V(0, S_0) + X_0 S_0 [f(X_0) - 1] + \sum_{t_i < T} V_S(t_i, S_i)(S_{i+1} - S_i) - L_T \rightarrow p(S_T)$$

*in  $L^2$  as  $\Delta t \rightarrow 0$ . In other words, the MSHE of this discrete delta-hedging strategy approaches zero.*

*Proof.* We prove the current theorem by relying in the result of the previous Theorem 3.1.

Rearranging the terms of the PDE one gets

$$\begin{aligned} \sigma^2 f'(0) S^3 V_{SS}^2 &= -V_t - 1/2 \sigma^2 S^2 V_{SS} \\ - \int_0^T \sigma^2 f'(0) S^3 V_{SS}^2 dt &= \int_0^T [-V_t - 1/2 \sigma^2 S^2 V_{SS}] dt \end{aligned}$$

Then, Theorem 3.1 implies

$$\begin{aligned}
& V(0, S_0) + X_0 S_0 [f(X_0) - 1] + \sum_{t_i < T} V_S(t_i, S_i)(S_{i+1} - S_{t_i}) - L_T \\
& \longrightarrow V(0, S_0) + \int_0^T V_S(t, S_t) dS_t - \int_0^T \sigma^2 f'(0) S_t^3 V_{SS}^2(t, S_t) dt \\
& = V(0, S_0) + \int_0^T V_S(t, S_t) dS_t + \int_0^T [V_t + 1/2 \sigma^2 S^2 V_{SS}] dt \\
& = V(T, S_T) = p(S_T)
\end{aligned}$$

by Ito's formula, which completes the proof of Theorem 3.4.

□

The next theorem gives the coefficient of  $\Delta t$  as a leading order term of the mean squared hedging error assuming varying rehedging times:

**Theorem 3.5.** *Under the conditions of theorem 3.1,*

$$\begin{aligned}
& \mathbb{E} \left[ V(0, S_0) + X_0 S_0 [f(X_0) - 1] + \sum_{t_i < T} V_S(t_i, S_i)(S_{i+1} - S_{t_i}) - L_T - p(S_T) \right]^2 \\
& = \Delta t \mathbb{E} \left[ \int_0^T \sigma^4 \frac{1}{d'(t)} V_{SS}^2(t, S_t) S_t^4 \left( \frac{1}{2} + 2\alpha V_{SS}(t, S_t) S_t + 2\alpha^2 V_{SS}^2(t, S_t) S_t^2 \right) dt \right] + O(\Delta t^2)
\end{aligned}$$

*Proof.* Note that we omit the higher order terms. Denote

$$\begin{aligned}\sum_1 &\equiv \sum_{t_i < T} V_S(t_i, S_i)(S_{i+1} - S_{t_i}) \\ \sum_2 &\equiv L_T = \sum_{t_i < T} \Delta X_i(S_i f(\Delta X_i) - S_i f(0)) \\ \int_1 &\equiv \int_0^T V_S(t, S_t) dS_t \\ \int_2 &\equiv \int_0^T \sigma^2 f'(0) S_t^3 V_{SS}^2(t, S_t) dt\end{aligned}$$

With this notation, the mean squared hedging error is written as:

$$\begin{aligned}&\mathbb{E} \left[ V(0, S_0) + X_0 S_0 [f(X_0) - 1] + \sum_{t_i < T} V_S(t_i, S_i)(S_{i+1} - S_i) - L_T - p(S_T) \right]^2 \\ &= \mathbb{E} \left[ \left( \sum_1 - \sum_2 - \left( \int_1 - \int_2 \right) \right)^2 \right] \\ &= \mathbb{E} \left[ \left( \sum_1 - \int_1 \right)^2 - 2 \left( \sum_1 - \int_1 \right) \left( \sum_2 - \int_2 \right) + \left( \sum_2 - \int_2 \right)^2 \right]\end{aligned}$$

The estimates for the first and third terms in the expectation are obtained from Theorem 3.1. Only the term in the middle of the expression needs to be investigated. By omitting the higher order terms and the almost identical computations to the proofs of the lemmas, the cross term of the errors is essentially

$$\begin{aligned}
& \mathbb{E} \left[ \sum_i \int_{t_{i-1}}^{t_i} (V_S(t, S_t) - V_S(t_{i-1}, S_{i-1})) dS_t V_{SS}^2(t_{i-1}, S_{i-1}) \Delta_i S_{i-1} f'(0) \right] \\
&= \mathbb{E} \left[ \sum_i V_{SS}^3(t_{i-1}, S_{i-1}) S_{i-1} f'(0) \mathbb{E}_{\mathcal{F}_{t_{i-1}}} \left[ \Delta_i \int_{t_{i-1}}^{t_i} (S_t - S_{i-1}) dS_t \right] \right] \\
&= \mathbb{E} \left[ \sum_i V_{SS}^3(t_{i-1}, S_{i-1}) S_{i-1}^5 f'(0) \sigma^4 (t_i - t_{i-1})^2 \right] \\
&= \mathbb{E} \left[ \sum_i V_{SS}^3(t_{i-1}, S_{i-1}) S_{i-1}^5 f'(0) \sigma^4 \left( \frac{\Delta t}{d'(t_{i-1})} + O(\Delta t^2) \right) (t_i - t_{i-1}) \right] \\
&\quad \longrightarrow \Delta t \mathbb{E} \left[ \int_0^T V_{SS}^3(t, S_t) S_t^5 f'(0) \sigma^4 \frac{1}{d'(t)} dt \right]
\end{aligned}$$

The result follows by collecting the estimates obtained. □

In the next section we are going to minimize the mean squared hedging error as given in the Theorem 3.5.

### 3.2 Optimal strategies: minimizing the hedging error

In this section we reduce hedging error by using varying instead of equally spaced trading times. We wish to replicate the payoff  $p(S_T)$  of a European option with a fixed initial capital prescribed by the partial differential equation (option premium) and a fixed number of available reheding times. Hedging error is nonzero due to

presence of liquidity risk as well as absence of continuous trading. We are minimizing the first order term in  $\Delta t$  of the mean squared hedging error of a trading strategy from Theorem 3.5. We investigate the optimal positioning of the rehedging times over all deterministic functions  $d(t)$ :

We employ the calculus of variations to solve the minimization problem. Essentially we find a solution  $d(t)$  giving the location of the rehedging times that will minimize the hedging error. We start by transforming the coefficient in front of  $\Delta t$  in the mean squared hedging error using Fubini theorem:

$$\begin{aligned} & \mathbb{E} \left[ \int_0^T \sigma^4 \frac{1}{d'(t)} V_{SS}^2(t, S_t) S_t^4 \left( \frac{1}{2} + 2\alpha V_{SS}(t, S_t) S_t + 2\alpha^2 V_{SS}^2(t, S_t) S_t^2 \right) dt \right] \\ &= \sigma^4 \int_0^T \frac{1}{d'(t)} \left( \frac{1}{2} \mathbb{E} [V_{SS}^2(t, S_t) S_t^4] + 2\alpha \mathbb{E} [V_{SS}^3(t, S_t) S_t^5] + 2\alpha^2 \mathbb{E} [V_{SS}^4(t, S_t) S_t^6] \right) dt \end{aligned}$$

Denote by  $A(t)$  the expression in parentheses from the previous line, that is,

$$A(t) = \frac{1}{2} \mathbb{E} [V_{SS}^2 S_t^4] + 2\alpha \mathbb{E} [V_{SS}^3 S_t^5] + 2\alpha^2 \mathbb{E} [V_{SS}^4 S_t^6]$$

Recall that in the calculus of variations the problem of finding a function  $d(t)$  which minimizes the following integral

$$I(d) = \int_{t_1}^{t_2} F(t, d(t), d'(t)) dt$$

where  $d(t_1) = d_1, d(t_2) = d_2$  and  $d'(t) \neq 0$ . We want to solve

$$\int_0^T \frac{1}{d'(t)} A(t) dt$$

with  $d(0) = 0, d(T) = T, d'(t) \neq 0$ , which means that the integrand

$$F(t, d(t), d'(t)) = \frac{1}{d'(t)} A(t)$$

in our case. The functional  $F$  depends only on  $t$  and  $d'$ , and the Euler's equation is  $\frac{d}{dt} F_{d'}(t, d') = 0$ . Next we integrate the Euler's equation with respect to  $t$  from both sides to get rid of the derivative in  $t$ . Then  $F_{d'}(t, d') = C_1$  for some constant  $C_1$ .

Differentiating  $F$  with respect to  $d'$ ,

$$F_{d'}(t, d') = -\frac{1}{(d'(t))^2} A(t) = C_1$$

Then we have  $0 \leq (d'(t))^2 = -\frac{A(t)}{C_1}$ . Next we take the square root on both sides for  $d'(t)$ :

$$d'(t) = \pm \sqrt{\frac{|A(t)|}{|-C_1|}} = \pm \frac{\sqrt{|A(t)|}}{\sqrt{|-C_1|}}$$

In order to find a value of  $\sqrt{|-C_1|}$  from the boundary condition, we integrate with respect to  $t$  from both sides. Then

$$d(t) = \pm \int_0^t \sqrt{\frac{|A(x)|}{|-C_1|}} dx + C_2$$

Recall that we require  $d(t)$  to be positive, so one is interested only in  $d(t) = + \int_0^t \sqrt{\frac{|A(x)|}{|-C_1|}} dx + C_2$ . Now we use the boundary conditions to determine the values for  $C_1$  and  $C_2$ . First we make use of  $d(0) = 0$ .

$$d(0) = 0 = \int_0^0 \sqrt{\frac{|A(x)|}{|-C_1|}} dx + C_2 = 0 + C_2$$

so  $C_2 = 0$  and  $d(t) = + \int_0^t \sqrt{\frac{|A(x)|}{|-C_1|}} dx$ . Now using  $d(T) = T$ ,

$$d(T) = T = \int_0^T \sqrt{\frac{|A(x)|}{|-C_1|}} dx = \frac{1}{\sqrt{|-C_1|}} \int_0^T \sqrt{|A(x)|} dx$$

thus

$$\frac{1}{\sqrt{|-C_1|}} = \frac{T}{\int_0^T \sqrt{|A(x)|} dx}$$

The final form of  $d(t)$  should be

$$\begin{aligned} d(t) &= \int_0^t \sqrt{\frac{|A(x)|}{|-C_1|}} dx \\ &= \frac{T}{\int_0^T \sqrt{|A(x)|} dx} \int_0^t \sqrt{|A(x)|} dx \end{aligned}$$

Now we compute  $d(t)$  numerically for a European call option with strike  $K = 100$ ,  $T = 0.5$  (expiration in half a year),  $\sigma = 0.4$ ,  $\alpha = 0.0001$ . We plot  $d(t)$  in Figure 3.2 for strikes 100, 90 and 80, while Figure 3.3 displays  $d(t)$  for strikes 100, 110 and 120. As the strike price deviates from the spot price, the recommendation is to re hedge more frequently toward the expiration (that is, the function's shape becomes steeper). The numerical simulations for put option produce same recommendations.

We use implicit finite difference scheme to solve the partial differential equation numerically. For more details, see Strikwerda [25] and Duffy [13]. Within the delta hedging framework the partial differential equation provides the number of shares to hold at each time for any trading strategy. Figure 3.4 displays the numerical

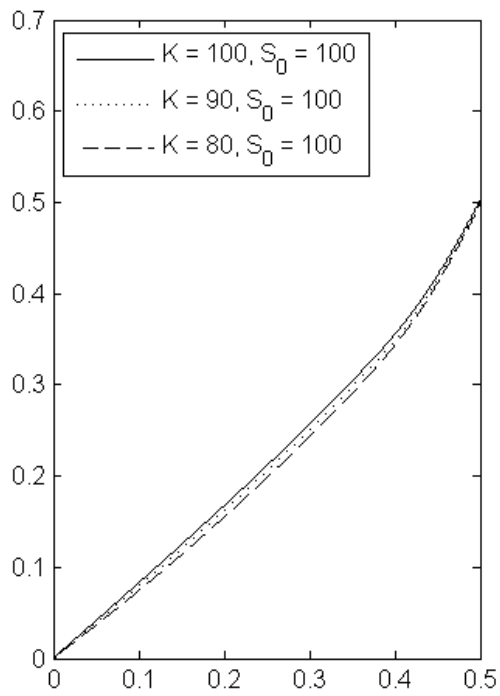


Figure 3.2: Optimal  $d(t)$  for a Euro call for decreasing strikes.

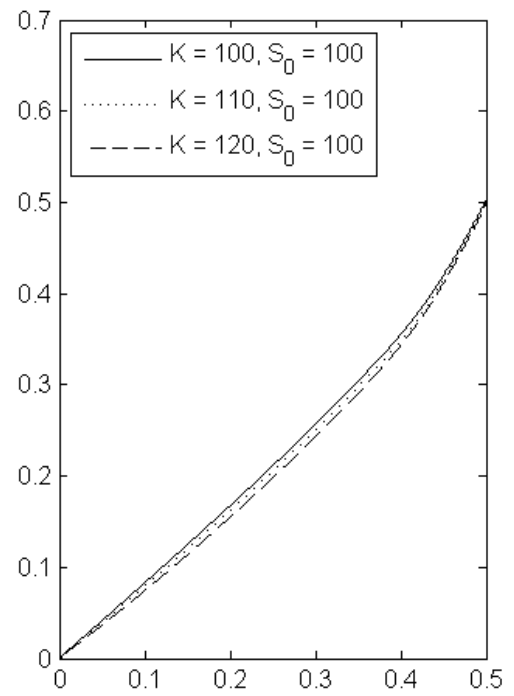


Figure 3.3: Optimal  $d(t)$  for a Euro call for increasing strikes.



solution of the partial differential equation (2.6) with the final condition  $V(T, S) = (S_T - K)^+$ , which corresponds to the call option payoff (the strike and the spot are equal to 100,  $\alpha = 0.0001$ ). As for the hedging errors, we use Monte Carlo to simulate paths of the geometric Brownian motion (actually, only the values at the expiration are utilized).

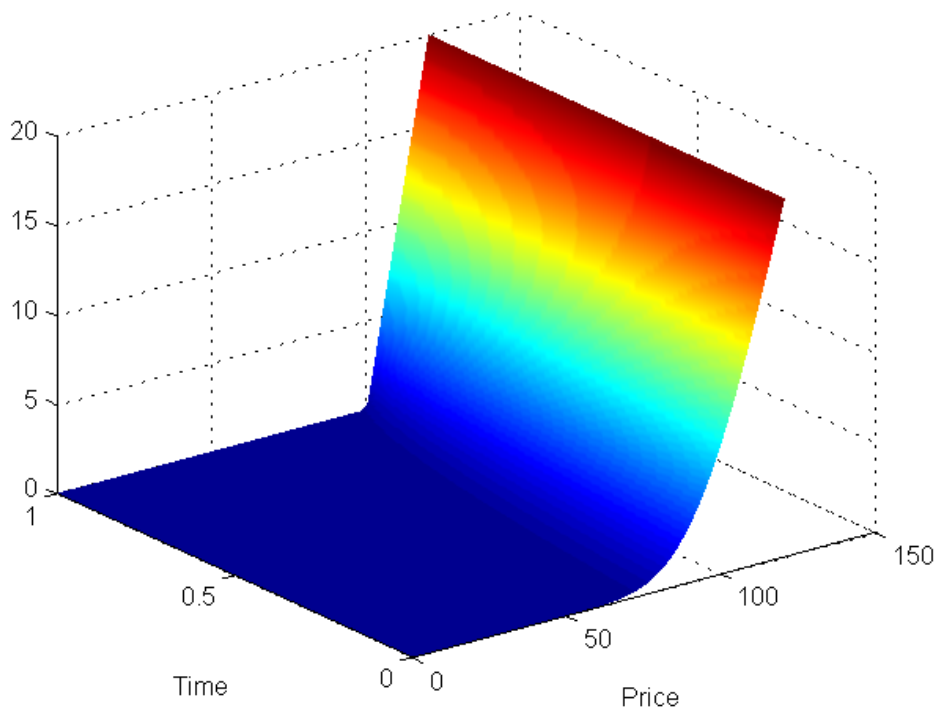


Figure 3.4: The numerical solution of the partial differential equation (2.6) with the final condition  $V(T, S) = (S_T - K)^+$ .

Table 3.1 summarizes the mean square hedging errors (*MSHE*) from using

varying rehedging times. We also list the mean hedging errors ( $MHE$ ). The subscript “e” refers to equally spaced trading times, while the subscript “v” signifies the result for the varying rebalancing times. That is,  $MHE_e$  and  $MSHE_e$  correspond to mean hedging error and mean square hedging error for equally spaced times. Similarly,  $MHE_v$  and  $MSHE_v$  correspond to mean hedging error and mean square hedging error for varying trading times. The computations are made for the European call option with  $T = 1$  (maturity date in a year from now), for varying strikes and values of  $\alpha$ . The spot price of the underlying is 100, that is  $S_0 = 100$ . Table 3.1 summarizes the hedging errors while using 250 trading times (that is, daily rebalancing) and  $\sigma = 0.2$ . Next we reduce the number of trading times from 250 to 52 (weekly rebalancing) and then down to 12 (monthly rebalancing) while keeping  $\sigma = 0.2$ . We list the results in Table 3.2 and Table 3.3 correspondingly. Table 3.4, Table 3.5 and Table 3.6 contain the results of daily, weekly and monthly rehedging with  $\sigma = 0.1$ .

We note that the hedging errors in the Black-Scholes setting (the case when  $\alpha = 0$ ) is also reduced by using varying rehedging times as well.

$\alpha$	strike	call price	mean			
			$MHE_e$	$MHE_v$	$MSHE_e$	$MSHE_v$
0.0001	95	10.5253366	0.000589	0.000537	0.013343	0.012845
	100	7.96572301	0.000671	0.000636	0.015219	0.015179
	105	5.90249744	0.003042	0.003890	0.016172	0.015938
0.0005	95	10.5083716	0.001093	0.001148	0.013608	0.013572
	100	7.92754391	0.000892	0.000731	0.015522	0.015310
	105	5.79126726	0.002353	0.001934	0.016493	0.016269
0.001	95	10.5077284	0.000911	0.000871	0.014205	0.014167
	100	7.94646865	0.001363	0.001166	0.016202	0.015746
	105	5.88314392	0.001322	0.001392	0.017216	0.016732

Table 3.1: Comparison of the hedging errors for equally spaced and varying rehedging times (subscripts  $e$  and  $v$  correspondingly) with daily rehedging for  $\sigma = 0.2$ .

$\alpha$	strike	call price	mean			
			$MHE_e$	$MHE_v$	$MSHE_e$	$MSHE_v$
0.0001	95	10.5253366	0.001250	0.001077	0.028108	0.020743
	100	7.96572301	0.001611	0.001910	0.022895	0.021902
	105	5.90249744	0.002432	0.002287	0.029164	0.028086
0.0005	95	10.5083716	0.001963	0.001294	0.027314	0.016969
	100	7.92754391	0.002153	0.001704	0.025522	0.015310
	105	5.79126726	0.001528	0.001617	0.020212	0.010256
0.001	95	10.5077284	0.001810	0.001876	0.017484	0.017427
	100	7.94646865	0.001976	0.001023	0.019862	0.019006
	105	5.88314392	0.002627	0.002587	0.020627	0.020464

Table 3.2: Comparison of the hedging errors for equally spaced and varying rehedging times (subscripts  $e$  and  $v$  correspondingly) with weekly rehedging for  $\sigma = 0.2$ .

$\alpha$	strike	call price	mean			
			$MHE_e$	$MHE_v$	$MSHE_e$	$MSHE_v$
0.0001	95	10.5253366	0.001581	0.001958	0.021522	0.020216
	100	7.96572301	0.001788	0.001329	0.028571	0.026941
	105	5.90249744	0.001362	0.001903	0.023269	0.022128
0.0005	95	10.5083716	0.002446	0.001483	0.020451	0.021208
	100	7.92754391	0.001920	0.002088	0.021412	0.016725
	105	5.79126726	0.001178	0.001757	0.021846	0.017516
0.001	95	10.5077284	0.002040	0.001379	0.024547	0.021344
	100	7.94646865	0.001836	0.002413	0.026706	0.024761
	105	5.88314392	0.002679	0.002636	0.039682	0.037551

Table 3.3: Comparison of the hedging errors for equally spaced and varying re-hedging times (subscripts  $e$  and  $v$  correspondingly) with monthly rehedging for  $\sigma = 0.2$ .

$\alpha$	strike	call price	mean			
			$MHE_e$	$MHE_v$	$MSHE_e$	$MSHE_v$
0.0001	95	6.88143273	0.000227	0.000132	0.005185	0.004467
	100	3.97951947	0.000162	0.000096	0.009780	0.008972
	105	2.05702452	0.000214	0.000198	0.001727	0.001041
0.0005	95	6.78019826	0.000247	0.000342	0.009105	0.007433
	100	3.95192241	0.000519	0.000381	0.009368	0.006045
	105	2.04212433	0.000229	0.000201	0.003033	0.001660
0.001	95	6.86598135	0.000114	0.000197	0.005159	0.004590
	100	3.96012972	0.000336	0.000224	0.008562	0.005249
	105	2.03966783	0.000032	0.000439	0.003024	0.002197

Table 3.4: Comparison of the hedging errors for equally spaced and varying rehedging times (subscripts  $e$  and  $v$  correspondingly) with daily rehedging for  $\sigma = 0.1$ .

$\alpha$	strike	call price	mean			
			$MHE_e$	$MHE_v$	$MSHE_e$	$MSHE_v$
0.0001	95	6.88143273	0.000099	0.000151	0.001951	0.001464
	100	3.97951947	0.000171	0.000138	0.005058	0.005497
	105	2.05702452	0.000222	0.000142	0.002713	0.003129
0.0005	95	6.78019826	0.000106	0.000162	0.002930	0.002802
	100	3.95192241	0.000196	0.000141	0.006242	0.005376
	105	2.04212433	0.000325	0.000576	0.006958	0.008018
0.001	95	6.86598135	0.000146	0.000078	0.002999	0.002905
	100	3.96012972	0.000206	0.000184	0.005813	0.005795
	105	2.03966783	0.000870	0.000819	0.006991	0.007391

Table 3.5: Comparison of the hedging errors for equally spaced and varying rehedging times (subscripts  $e$  and  $v$  correspondingly) with weekly rehedging for  $\sigma = 0.1$ .

$\alpha$	strike	call price	mean			
			$MHE_e$	$MHE_v$	$MSHE_e$	$MSHE_v$
0.0001	95	6.88143273	0.000156	0.000157	0.003501	0.003319
	100	3.97951947	0.000154	0.000119	0.006224	0.006175
	105	2.05702452	0.000180	0.000130	0.007158	0.007095
0.0005	95	6.78019826	0.000166	0.000150	0.003438	0.002921
	100	3.95192241	0.000173	0.000221	0.006422	0.006124
	105	2.04212433	0.000155	0.000096	0.008316	0.007607
0.001	95	6.86598135	0.000224	0.000226	0.007989	0.008012
	100	3.96012972	0.000251	0.000191	0.006487	0.006062
	105	2.03966783	0.000193	0.000201	0.008136	0.007965

Table 3.6: Comparison of the hedging errors for equally spaced and varying re-hedging times (subscripts  $e$  and  $v$  correspondingly) with monthly rehedging for  $\sigma = 0.1$ .



## 4 Hedging via local risk-minimization

In this chapter we investigate the pricing and hedging of a European option in a model with liquidity risk with *local risk-minimization* as the criterion for optimality for a replicating trading strategy. We assume a general square-integrable stochastic process to model the price of the underlying. We consider any form of European contingent claims, including convex and concave payoffs. The main result is existence of a local risk-minimizing strategy in Theorem 4.10.

### 4.1 Definitions

In this section we introduce definitions and notation for local risk-minimization in discrete time in the presence of liquidity risk. We also find the criterion for checking if a trading strategy is local risk-minimizing in Proposition 4.2.

Consider a filtered probability space  $[\Omega, \mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}, \mathbb{P}]$ , where  $T \in \mathbb{N}$  is a natural number representing the fixed time (expiration of the option).  $\mathbb{P}$  is statistical or empirical probability measure for a stock that pays no dividends.

Assume the rate of interest is zero, that is the price process is discounted and the price of non-risky asset is always equal to one.

Let  $S(\omega, t, 0) = S_t$  be the marginal price of the supply curve, that is the general price process for the underlying asset. We assume the discounted price process for the risky asset  $S_t$  is adapted to the filtration  $\mathbb{F}$ , nonnegative and square integrable:  $S_t \in L^2(\mathbb{P})$ , that is for every  $t$ ,  $\int_{\Omega} S_t^2(\omega) dP(\omega) = E[S_t^2] < \infty$ . Let the supply curve  $S(\omega, t, x)$  represent the stock price per share at time  $[0, T]$  that a trader pays/receives for an order of size  $x \in \mathbb{R}$ . A positive order ( $x > 0$ ) represents a buy, a negative order ( $x < 0$ ) signifies a sale, and  $x = 0$  corresponds to the marginal trade. The supply curve is multiplicative, and we write simply  $S(\omega, t, x) = S_t f(x)$ , where  $f(x)$  is real-valued increasing continuous function with  $f(0) = 1$ .

The *goal* is to hedge/replicate European contingent claim (strike  $K$ ) with the payoff  $p(S_T)$ . No convexity or concavity of option's payoff is required.

As before, we model the derivative payoff via random variables.

**Definition 4.1.** *A contingent claim is a contract to deliver a number of units  $\delta_{T+1}^-$  of the stock and an amount  $\bar{\lambda}_T$  in cash, where  $\delta_{T+1}^-$  and  $\bar{\lambda}_T$  are  $\mathcal{F}_T$ -measurable random variables, and  $\delta_{T+1}^-$  satisfies*

$$\delta_{T+1}^- S_T \in L^2(\mathbb{P}).$$

For example, upon exercise of the European call option with the maturity  $T$

and the strike  $K$ , the option writer has to hand over one share of stock in exchange for  $K$  units of cash so that

$$\delta_{T+1}^- = 1_{S_T > K}, \quad \bar{\lambda}_T = -K 1_{S_T > K}.$$

This is settlement with delivery.

We assume for simplicity that liquidation of a position at date  $T$  does not incur liquidity costs.

We do not specify the price process explicitly as we did in Chapter 3. Instead, we require the price process to be square-integrable for expectations and variances to be well-defined.

**Definition 4.2.** For any discrete stochastic process  $S = (S)_{k=0,1,\dots,T}$  denote by  $\Theta(S)$  the space of all stochastic processes  $\delta = (\delta)_{k=1,\dots,T+1}$  such that

- $\delta_k$  is  $\mathcal{F}_{k-1}$ -measurable (that is, predictable)
- $\delta_k \Delta S_k \in L^2(\mathbb{P})$  for  $k = 1, \dots, T$ , where  $\Delta S_k := S_k - S_{k-1}$ .

The requirement  $\delta_k \Delta S_k \in L^2(\mathbb{P})$  for  $k = 1, \dots, T$  from the following definition makes sure the objects of the form  $\sum_{j=1}^k \delta_j \Delta S_j$  are well-defined. In fact,  $\delta_j \Delta S_j$  represents trading gains/losses from holding  $\delta_j$  stock shares while the stock share price changes from date  $j - 1$  to  $j$ .

As before, a trading strategy consists of a position in the stock shares and a position in cash.

**Definition 4.3.** A (trading) strategy  $\eta$  is a pair of processes  $\delta, \lambda$  such that

- $\delta \in \Theta(S)$  (position in a risky asset),
- $\lambda = (\lambda)_{k=0,1,\dots,T}$  is adapted to  $\mathbb{F}$  (position in a non-risky asset/money market account).

**Definition 4.4.** The (adapted) value process of a (trading) strategy  $\eta = (\delta, \lambda)$  is

$$V_k(\eta) := \delta_{k+1}S_k + \lambda_k \in L^2(\mathbb{P}).$$

At each date  $k$ , one may choose the number  $\delta_{k+1}$  of stock shares and the number  $\lambda_k$  of cash to hold until the following date  $k + 1$ .

Here the change in the number of shares  $\delta_{k+1} - \delta_k$  corresponds to the order size of the supply curve introduced above. While following the trading strategy  $\eta$  we incur liquidity costs. The flow of capital at date  $k$  consists of buying or selling (depending upon the signs)  $\lambda_k - \lambda_{k-1}$  units of cash,  $\delta_{k+1} - \delta_k$  stock shares which results in an expenditure of the size  $(\delta_{k+1} - \delta_k)S_k$  and liquidity cost:

$$\begin{aligned} & \lambda_k - \lambda_{k-1} + (\delta_{k+1} - \delta_k)S_k + (\delta_{k+1} - \delta_k)(S_k f(\delta_{k+1} - \delta_k) - S_k f(0)) \\ &= \lambda_k + \delta_{k+1}S_k - \lambda_{k-1} - \cancel{\delta_k S_{k-1}} - \delta_k S_k + \cancel{\delta_k S_{k-1}} + (\delta_{k+1} - \delta_k)S_k(f(\delta_{k+1} - \delta_k) - 1) \\ &= \lambda_k + \delta_{k+1}S_k - (\lambda_{k-1} + \delta_k S_{k-1}) - \delta_k(S_k - S_{k-1}) + (\delta_{k+1} - \delta_k)S_k(f(\delta_{k+1} - \delta_k) - 1) \\ &= V_k(\eta) - V_{k-1}(\eta) - \delta_k(S_k - S_{k-1}) + (\delta_{k+1} - \delta_k)S_k(f(\delta_{k+1} - \delta_k) - 1) \end{aligned}$$

One may get the cumulative costs of the strategy  $\eta$  by summing over all dates up to  $k$ . Recall that the indices for  $\delta_k$  run from 1 to  $T + 1$ , while all other elements of the model are indexed from 0 to  $T$ , as illustrated in the Figure 4.1.

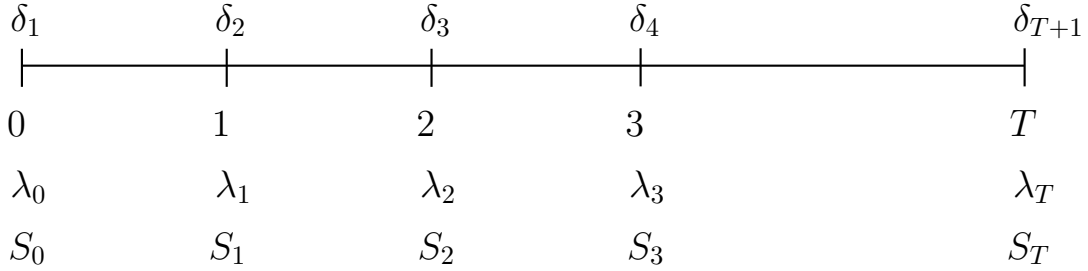


Figure 4.1: The indices for  $\delta_k$  run from 1 up to  $T + 1$ .

**Definition 4.5.** *The cost process of a strategy  $\eta = (\delta, \lambda)$  is*

$$C_k(\eta) := V_k(\eta) - \sum_{j=1}^k \delta_j \Delta S_j + \sum_{j=1}^k (\delta_{j+1} - \delta_j) (f(\delta_{j+1} - \delta_j) - 1) S_j \quad \text{for } k = 1, \dots, T$$

and  $C_0(\eta) = V_0(\eta)$ .

The cost process monitors changes in the value process that are not due to trading gains  $\delta_j(S_j - S_{j-1})$ . We note that  $V_0(\eta)$  is the minimum initial capital for the trading strategy  $\eta$ ; that is initial costs are equal to the sum invested at time 0. In the Chapter 3 the cost process is constant, since the replicating trading strategies are self-financing.

The next definition assumes the cost process of any strategy is square-integrable (please see Lemma 4.3 for the proof). If  $\Omega$  is finite or  $\alpha = 0$ , this is clearly satisfied due to the definition of a trading strategy.

**Definition 4.6.** *The risk process (conditional mean square error process) of a strategy  $\eta$  is*

$$R_k(\eta) := E[(C_T(\eta) - C_k(\eta))^2 | \mathcal{F}_k] \quad \text{for } k = 0, 1, \dots, T.$$

We define the local risk-minimizing (that is, optimal) strategy via the risk process.

We shall minimize  $R_k(\eta)$  only with respect to  $\lambda_k$  and  $\delta_{k+1}$ , as the only decision one has to take at time  $k$  is the choice of  $\lambda_k$  and  $\delta_{k+1}$ . Other variables that influence the risk process at date  $k$  are  $\delta_{k+2}, \dots, \delta_{T+1}$  and  $\lambda_{k+1}, \dots, \lambda_T$ . The following definition formalized the dependence of the risk process at date  $k$  on cash holdings  $\lambda_k$  and the number of stock shares  $\delta_{k+1}$ .

**Definition 4.7.** *Let  $\eta = (\delta, \lambda)$  be a strategy and  $k \in \{0, 1, \dots, T - 1\}$ . A local perturbation of  $\eta$  at a date  $k$  is a strategy  $\eta' = (\delta', \lambda')$  with*

- $\delta'_j = \delta_j$      for  $j \neq k + 1$
- $\lambda'_j = \lambda_j$      for  $j \neq k$ .

*A strategy  $\eta$  is called local risk-minimizing (inclusive of liquidity costs) if one has*

$$R_k(\eta) \leq R_k(\eta') \quad P\text{-a.s.}$$

*for any date  $k \in \{0, 1, \dots, T - 1\}$  and any local perturbation  $\eta'$  of  $\eta$  at the date  $k$ .*

We may restate the problem in the notation of this chapter: given a contingent claim  $(\delta_{T+1}^-, \bar{\lambda}_T)$ , find a local risk-minimizing strategy  $\eta = (\delta, \lambda)$  with  $\delta_{T+1} = \delta_{T+1}^-$  and  $\lambda_T = \bar{\lambda}_T$ .

The following lemma gives the property of optimal strategy and the form of the risk process.

**Lemma 4.1.** *If  $\eta$  is local risk-minimizing, then  $C(\eta)$  is a martingale and therefore*

$$R_k(\eta) = E[R_{k+1}(\eta)|\mathcal{F}_k] + \text{Var}[\Delta C_{k+1}(\eta)|\mathcal{F}_k] \quad P\text{-a.s. for } k = 0, 1, \dots, T-1. \quad (4.1)$$

*Proof.* We first prove that  $C(\eta)$  is a martingale, which is equivalent to showing  $E[C_T(\eta) - C_k(\eta)|\mathcal{F}_k] = 0$   $P$ -a.s.

Fix a date  $k \in \{0, 1, \dots, T-1\}$ , choose a strategy  $\eta$  and introduce a local perturbation of  $\eta$  at date  $k$  by setting  $\delta' := \delta$  (that is, same holdings of the risky asset) and  $\lambda'_j := \lambda_j$  for  $j \neq k$  where

$$\lambda'_k := E[C_T(\eta) - C_k(\eta)|\mathcal{F}_k] + \lambda_k.$$

Then  $\lambda'$  is clearly adapted to the filtration  $\mathbb{F}$  and

$$V_k(\eta') = V_k(\eta) + E[C_T(\eta) - C_k(\eta)|\mathcal{F}_k] \in L^2(\mathbb{P}) \quad (4.2)$$

Therefore  $\eta'$  is a strategy, hence a local perturbation of  $\eta$  at date  $k$ . Also by (4.2) and the definition of  $\eta'$  one gets

$$C_T(\eta') - C_k(\eta') = C_T(\eta) - C_k(\eta) - E[C_T(\eta) - C_k(\eta)|\mathcal{F}_k]$$

where the last term appears as the only difference in the trading strategies  $\eta$  and  $\eta'$ . Using the equality above we get

$$\begin{aligned}
R_k(\eta') &= \mathbb{E}[(C_T(\eta') - C_k(\eta'))^2 | \mathcal{F}_k] \\
&= \mathbb{E}[(C_T(\eta) - C_k(\eta) - \mathbb{E}[C_T(\eta) - C_k(\eta) | \mathcal{F}_k])^2 | \mathcal{F}_k] \\
&= \mathbb{E}[(C_T(\eta) - C_k(\eta))^2 | \mathcal{F}_k] - 2 \mathbb{E}[(C_T(\eta) - C_k(\eta)) \mathbb{E}[C_T(\eta) - C_k(\eta) | \mathcal{F}_k] | \mathcal{F}_k] \\
&\quad + \mathbb{E}[(\mathbb{E}[C_T(\eta) - C_k(\eta) | \mathcal{F}_k])^2 | \mathcal{F}_k] \\
&= \mathbb{E}[(C_T(\eta) - C_k(\eta))^2 | \mathcal{F}_k] - (\mathbb{E}[C_T(\eta) - C_k(\eta) | \mathcal{F}_k])^2 \\
&\leq \mathbb{E}[(C_T(\eta) - C_k(\eta))^2 | \mathcal{F}_k] = R_k(\eta)
\end{aligned}$$

Because  $\eta$  is local risk-minimizing, one must have equality  $P$ -a.s. and therefore  $\mathbb{E}[C_T(\eta) - C_k(\eta) | \mathcal{F}_k] = 0$   $P$ -a.s., showing that  $C(\eta)$  is a martingale.

The formula (4.1) for the risk process in the statement of the lemma is the immediate consequence of the martingale property of  $C(\eta)$ .  $\square$

The previous Lemma 4.1 recommends to look for a local risk-minimizing strategy by recursively minimizing  $\text{Var}[\Delta C_{k+1}(\eta) | \mathcal{F}_k]$  with respect to  $\delta_{k+1}$  and then solving for  $\lambda_k$  from the martingale property of  $C(\eta)$ .

To continue we need to use the specific formula for the supply curve for the next



proposition:

$$f(x) := \begin{cases} 1 + \alpha x & : |x| \leq N \\ 1 + \text{sign}(x)\alpha N & : |x| > N \end{cases}$$

where  $N \in \mathbb{R}$  and  $N > -\frac{1}{\alpha}$  to ensure the nonnegativity of the supply curve. Parameter  $\alpha$  is estimated using simple linear regression methodology from the history of stock prices. The value of  $\alpha$  ends up being small, usually within  $0 < \alpha < 0.001$ . The model is focused on a small trader that cannot move the market by her transactions.

Recall that in our notation  $x$  in  $f(x)$  is the order size and corresponds to the  $\delta_{k+1} - \delta_k$ , so we substitute  $\delta_{k+1} - \delta_k$  for  $x$  in the formula for  $f(x)$  above:

$$f(\delta_{k+1} - \delta_k) := \begin{cases} 1 + \alpha(\delta_{k+1} - \delta_k) & : |\delta_{k+1} - \delta_k| \leq N \\ 1 + \text{sign}(\delta_{k+1} - \delta_k)\alpha N & : |\delta_{k+1} - \delta_k| > N \end{cases}$$

Observe that two mutually exclusive events  $|\delta_{k+1} - \delta_k| \leq N$  and  $|\delta_{k+1} - \delta_k| > N$  constitute the whole sample space.

The flow of capital from time  $k - 1$  to time  $k$  is:

$$\begin{aligned} & \text{event 1 } (|\delta_{k+1} - \delta_k| \leq N) : \\ & \lambda_k - \lambda_{k-1} + (\delta_{k+1} - \delta_k)S_k + (\delta_{k+1} - \delta_k)(S_k f(\delta_{k+1} - \delta_k) - S_k f(0)) \\ & = \lambda_k - \lambda_{k-1} + (\delta_{k+1} - \delta_k)S_k + S_k(\delta_{k+1} - \delta_k)(\mathcal{I} + \alpha(\delta_{k+1} - \delta_k) - \mathcal{I}) \\ & = V_k(\eta) - V_{k-1}(\eta) - \delta_k(S_k - S_{k-1}) + \alpha S_k(\delta_{k+1} - \delta_k)^2 \end{aligned}$$

Next we use  $|\delta_{k+1} - \delta_k| = \text{sign}(\delta_{k+1} - \delta_k)(\delta_{k+1} - \delta_k)$ :

event 2 ( $|\delta_{k+1} - \delta_k| > N$ ) :

$$\begin{aligned}
& \lambda_k - \lambda_{k-1} + (\delta_{k+1} - \delta_k)S_k + (\delta_{k+1} - \delta_k)(S_k f(\delta_{k+1} - \delta_k) - S_k f(0)) \\
&= \lambda_k - \lambda_{k-1} + (\delta_{k+1} - \delta_k)S_k + S_k(\delta_{k+1} - \delta_k)(\mathcal{I} + \alpha N \text{sign}(\delta_{k+1} - \delta_k) - \mathcal{I}) \\
&= V_k(\eta) - V_{k-1}(\eta) - \delta_k(S_k - S_{k-1}) + \alpha S_k N \text{sign}(\delta_{k+1} - \delta_k)(\delta_{k+1} - \delta_k) \\
&= V_k(\eta) - V_{k-1}(\eta) - \delta_k(S_k - S_{k-1}) + \alpha S_k N |\delta_{k+1} - \delta_k|
\end{aligned}$$

One also may rewrite the flow of capital from time  $k - 1$  to time  $k$  via indicator functions to have both events in one expression:

$$\begin{aligned}
V_k(\eta) - V_{k-1}(\eta) - \delta_k(S_k - S_{k-1}) &+ 1_{\{|\Delta\delta_{k+1}| \leq N\}} \alpha S_k (\Delta\delta_{k+1})^2 \\
&+ 1_{\{|\Delta\delta_{k+1}| > N\}} \alpha N S_k |\Delta\delta_{k+1}|
\end{aligned}$$

One may get the cumulative costs of of the strategy  $\eta$  by summing over all dates up to  $k$ :

$$\begin{aligned}
C_k(\eta) &:= V_k(\eta) - \sum_{j=1}^k \delta_j \Delta S_j + \alpha \sum_{j=1}^k 1_{\{|\Delta\delta_{j+1}| \leq N\}} S_j (\Delta\delta_{j+1})^2 \\
&+ \alpha N \sum_{j=1}^k 1_{\{|\Delta\delta_{j+1}| > N\}} S_j |\Delta\delta_{j+1}| \quad \text{for } k = 0, 1, \dots, T.
\end{aligned}$$

We explicitly write out the form of  $\Delta C_{k+1}(\eta)$  using the definition of  $C_{k+1}(\eta)$ :

$$\begin{aligned}\Delta C_{k+1}(\eta) &= \Delta V_{k+1}(\eta) - \delta_{k+1} \Delta S_{k+1} \\ &\quad + 1_{\{|\Delta \delta_{k+2}| \leq N\}} \alpha S_{k+1} (\Delta \delta_{k+2})^2 + 1_{\{|\Delta \delta_{k+2}| > N\}} \alpha N S_{k+1} |\Delta \delta_{k+2}| \end{aligned} \quad (4.3)$$

The next proposition is the main tool for discovering optimal strategies. We use it once at the very end of this chapter to show that the strategy we constructed is indeed local risk-minimizing.

**Proposition 4.2.** *A strategy  $\eta = (\delta, \lambda)$  is local risk-minimizing if and only if it has the following two properties:*

1.  $C(\eta)$  is a martingale
2. For each  $k \in \{0, 1, \dots, T-1\}$ ,  $\delta_{k+1}$  minimizes

$$\begin{aligned} &\text{Var}[V_{k+1}(\eta) - \delta'_{k+1} \Delta S_{k+1} + 1_{\{|\delta_{k+2} - \delta'_{k+1}| \leq N\}} \alpha S_{k+1} (\delta_{k+2} - \delta'_{k+1})^2 \\ &\quad + 1_{\{|\delta_{k+2} - \delta'_{k+1}| > N\}} \alpha N S_{k+1} |\delta_{k+2} - \delta'_{k+1}| \mid \mathcal{F}_k] \end{aligned}$$

over all  $\mathcal{F}_k$ -measurable random variables  $\delta'_{k+1}$  such that  $\delta'_{k+1} \Delta S_{k+1} \in L^2(\mathbb{P})$ .

*Proof.* Note that  $\delta_{k+1} \Delta S_{k+1} \in L^2(\mathbb{P})$  as a part of a trading strategy.

First we obtain several helpful facts. Since  $\eta' = (\delta', \lambda')$  is a local perturbation at date  $k$ ,  $\delta'_j = \delta_j$  for  $j \neq k+1$  and  $\lambda'_j = \lambda_j$  for  $j \neq k$  by definition, so one has

$$V_{k+1}(\eta') = V_{k+1}(\eta) \quad \text{and} \quad \delta'_{k+2} = \delta_{k+2} \quad (4.4)$$

Also one has

$$C_T(\eta') - C_{k+1}(\eta') = C_T(\eta) - C_{k+1}(\eta) \quad (4.5)$$

and so  $C_i(\eta')$  is a martingale for  $i = k + 1, \dots, T$  (in particular,  $\mathbb{E}[C_T(\eta')|\mathcal{F}_{k+1}] = C_{k+1}(\eta')$ ), since the terms with index  $k$  corresponding to the local perturbation cancel out via the definition of the cost process. Next we show

$$R_k(\eta') = \mathbb{E}[R_{k+1}(\eta)|\mathcal{F}_k] + \mathbb{E}[(\Delta C_{k+1}(\eta'))^2|\mathcal{F}_k] \quad (4.6)$$

using the definition of risk process and (4.5), simple algebra and  $\mathbb{E}[C_T(\eta')|\mathcal{F}_{k+1}] = C_{k+1}(\eta')$  (condition on  $\mathcal{F}_{k+1}$  first):

$$\begin{aligned} & \mathbb{E}[R_{k+1}(\eta)|\mathcal{F}_k] + \mathbb{E}[(\Delta C_{k+1}(\eta'))^2|\mathcal{F}_k] \\ &= \mathbb{E}[R_{k+1}(\eta')|\mathcal{F}_k] + \mathbb{E}[(\Delta C_{k+1}(\eta'))^2|\mathcal{F}_k] \\ &= \mathbb{E}[\mathbb{E}[(C_T(\eta') - C_{k+1}(\eta'))^2|\mathcal{F}_{k+1}]|\mathcal{F}_k] + \mathbb{E}[(C_{k+1}(\eta') - C_k(\eta'))^2|\mathcal{F}_k] \\ &= \mathbb{E}[C_T^2(\eta')|\mathcal{F}] - 2\mathbb{E}[C_{k+1}(\eta')\mathbb{E}[C_T(\eta')|\mathcal{F}_{k+1}]|\mathcal{F}_k] + \mathbb{E}[C_{k+1}^2(\eta')|\mathcal{F}_k] \\ &+ \mathbb{E}[C_{k+1}^2(\eta')|\mathcal{F}_k] - 2C_k(\eta')\mathbb{E}[C_{k+1}(\eta')|\mathcal{F}_k] + \mathbb{E}[C_k^2(\eta')|\mathcal{F}_k] \\ &= \mathbb{E}[C_T^2(\eta')|\mathcal{F}] - 2C_k(\eta')\mathbb{E}[C_{k+1}(\eta')|\mathcal{F}_k] + \mathbb{E}[C_k^2(\eta')|\mathcal{F}_k] \\ &= R_k(\eta') \end{aligned}$$

We proceed with the proof relying on all the facts presented above.

( $\Leftarrow$ ) Suppose one has a (trading) strategy  $\eta$  and the conditions 1 and 2 from the Proposition 4.2 are met for  $\eta$ . Choose a date  $k \in \{0, 1, \dots, T - 1\}$  and let  $\eta'$

be a local perturbation of  $\eta$  at date  $k$ . We will show that  $\eta$  is local risk-minimizing strategy by the definition, that is  $R_k(\eta) \leq R_k(\eta')$ .

Using (4.4) and the form of  $\Delta C_{k+1}(\eta')$  in the formula (4.3) one gets

$$\begin{aligned} \Delta C_{k+1}(\eta') &= V_{k+1}(\eta) - V_k(\eta') - \delta'_{k+1} \Delta S_{k+1} \\ &\quad + 1_{\{|\delta_{k+2} - \delta'_{k+1}| \leq N\}} \alpha S_{k+1} (\delta_{k+2} - \delta'_{k+1})^2 + 1_{\{|\delta_{k+2} - \delta'_{k+1}| > N\}} \alpha N S_{k+1} |\delta_{k+2} - \delta'_{k+1}| \end{aligned} \quad (4.7)$$

We use (4.6) and  $E[X^2] \geq \text{Var}[X]$  to obtain the first inequality, (4.7), omit  $\mathcal{F}_k$ -measurable terms from the conditional variance, use condition 2 of the current proposition to get the second inequality, the definition of  $\Delta C_{k+1}(\eta)$  and the last

equality comes from (4.1):

$$\begin{aligned}
& R_k(\eta') \\
&= \mathbb{E}[R_{k+1}(\eta)|\mathcal{F}_k] + \mathbb{E}[(\Delta C_{k+1}(\eta'))^2|\mathcal{F}_k] \\
&\geq \mathbb{E}[R_{k+1}(\eta)|\mathcal{F}_k] + \text{Var}[\Delta C_{k+1}(\eta')|\mathcal{F}_k] \\
&= \mathbb{E}[R_{k+1}(\eta)|\mathcal{F}_k] + \text{Var}[V_{k+1}(\eta) - V_k(\eta') - \delta'_{k+1}\Delta S_{k+1} \\
&\quad + 1_{\{|\delta_{k+2}-\delta'_{k+1}|\leq N\}}\alpha S_{k+1}(\delta_{k+2} - \delta'_{k+1})^2 + 1_{\{|\delta_{k+2}-\delta'_{k+1}|>N\}}\alpha N S_{k+1}|\delta_{k+2} - \delta'_{k+1}||\mathcal{F}_k] \\
&= \mathbb{E}[R_{k+1}(\eta)|\mathcal{F}_k] + \text{Var}[V_{k+1}(\eta) - \delta'_{k+1}\Delta S_{k+1} \\
&\quad + 1_{\{|\delta_{k+2}-\delta'_{k+1}|\leq N\}}\alpha S_{k+1}(\delta_{k+2} - \delta'_{k+1})^2 + 1_{\{|\delta_{k+2}-\delta'_{k+1}|>N\}}\alpha N S_{k+1}|\delta_{k+2} - \delta'_{k+1}||\mathcal{F}_k] \\
&\geq \mathbb{E}[R_{k+1}(\eta)|\mathcal{F}_k] + \text{Var}[V_{k+1}(\eta) - \delta_{k+1}\Delta S_{k+1} \\
&\quad + 1_{\{|\delta_{k+2}-\delta_{k+1}|\leq N\}}\alpha S_{k+1}(\delta_{k+2} - \delta_{k+1})^2 + 1_{\{|\delta_{k+2}-\delta_{k+1}|>N\}}\alpha N S_{k+1}|\delta_{k+2} - \delta_{k+1}||\mathcal{F}_k] \\
&= \mathbb{E}[R_{k+1}(\eta)|\mathcal{F}_k] + \text{Var}[\Delta C_{k+1}(\eta)|\mathcal{F}_k] \\
&= R_k(\eta)
\end{aligned}$$

Then  $\eta$  is local risk-minimizing strategy by the definition.

( $\Rightarrow$ ) Conversely, suppose that  $\eta$  is local risk-minimizing. Then condition 1 ( $C(\eta)$

is a martingale) holds by Lemma 4.1.

Since  $\eta$  is local risk-minimizing, then  $R_k(\eta) - R_k(\eta') \leq 0$  from definition. Subtracting (4.6) from (4.1) yields:

$$\text{Var}[\Delta C_{k+1}(\eta)|\mathcal{F}_k] - \mathbb{E}[(\Delta C_{k+1}(\eta'))^2|\mathcal{F}_k] = R_k(\eta) - R_k(\eta') \leq 0$$

thus  $\text{Var}[\Delta C_{k+1}(\eta)|\mathcal{F}_k] \geq \text{E}[(\Delta C_{k+1}(\eta'))^2|\mathcal{F}_k]$  for any  $\mathcal{F}_k$ -measurable choice of  $\delta'_{k+1}$  and  $\lambda'_k$ . Specifically, we can fix  $\delta'_{k+1}$  and choose  $\lambda'_k$  in such a way that  $\text{E}[\Delta C_{k+1}(\eta')|\mathcal{F}_k] = 0$ . Using this fact and the definition of variance we get:

$$\begin{aligned} \text{Var}[\Delta C_{k+1}(\eta')|\mathcal{F}_k] &= \text{E}[(\Delta C_{k+1}(\eta'))^2|\mathcal{F}_k] - (\text{E}[\Delta C_{k+1}(\eta')|\mathcal{F}_k])^2 \\ &= \text{E}[(\Delta C_{k+1}(\eta'))^2|\mathcal{F}_k] \\ &\geq \text{Var}[\Delta C_{k+1}(\eta)|\mathcal{F}_k] \end{aligned}$$

so condition 2 holds. □

The proposition allows to determine whether the trading strategy is local risk-minimizing. We will use the proposition at the very end of the Theorem 4.10 to demonstrate the optimality of the trading strategy.

## 4.2 Conditions on the price process and technical results

The section introduces the new modified price process  $S_k^\xi$  and show that new process is suitable for the option hedging ( $\Theta(S_k) = \Theta(S_k^\xi)$ ).

We introduce the modified price process that incorporates liquidity costs and show that the new process may be used for hedging/replication of contingent claims. Throughout the section we assume  $S$  is a square-integrable process:  $S_k \in L^2(\mathbb{P})$  for  $k = 0, 1, \dots, T$ . Recall that a linear combination of functions in  $L^2(\mathbb{P})$  is in  $L^2(\mathbb{P})$ . Also a product of a function in  $L^2(\mathbb{P})$  and a bounded function is again in  $L^2(\mathbb{P})$ .

**Definition 4.8.** (a)  $\Xi$  is the class of all bounded adapted processes  $\xi = (\xi)_{k=0,1,\dots,T}$  with values in  $[-N, +N]$  for some fixed  $N \in \mathbb{N}$ .

(b) Given an adapted process  $\delta_{k+1}$ , we can associate with it a  $\xi = \xi^\delta \in \Xi$  as follows:

$$\xi_k(\Delta\delta_{k+1}) := \begin{cases} \Delta\delta_{k+1} & : |\Delta\delta_{k+1}| \leq N \\ N \operatorname{sign}(\Delta\delta_{k+1}) & : |\Delta\delta_{k+1}| > N \end{cases}$$

where  $\Delta\delta_{k+1}$  represents the order size.

One may think that  $\xi$  holds the structure of the supply curve.

**Definition 4.9.** For  $\xi \in \Xi$  define the modified price process  $S_k^\xi$  by

$$S_k^\xi = S_k(1 + \alpha\xi_k)$$

One may substitute the expression for  $\xi_k$  into the formula for  $S_k^\xi$  above:

$$S_k^\xi = \begin{cases} S_k(1 + \alpha\Delta\delta_{k+1}) & : |\Delta\delta_{k+1}| \leq N \\ S_k(1 + \alpha N \operatorname{sign}(\Delta\delta_{k+1})) & : |\Delta\delta_{k+1}| > N \end{cases}$$

The definition above is for the price per share inclusive of liquidity costs. It is clear from the construction that each process  $S_k^\xi$  is again adapted, nonnegative and square-integrable.

**Definition 4.10.** If  $\eta = (\delta, \lambda)$  is a strategy, the process  $V_k^\xi(\eta)$  is defined by

$$V_k^\xi(\eta) := \delta_{k+1}S_k^\xi + \lambda_k \quad \text{for } K = 0, 1, \dots, T.$$



**Definition 4.11.** *We say that  $S$  has substantial risk if there is a constant  $c < \infty$  such that*

$$\frac{S_{k-1}^2}{\mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}]} \leq c \quad P\text{-a.s. for } k = 1, \dots, T. \quad (4.8)$$

We denote  $c_{SR}$  the smallest constant satisfying (4.8). One may also rewrite the definition of substantial risk (4.8) using simple algebra to get a bound on  $S_{k-1}^2$ :

$$S_{k-1}^2 \leq c_{SR} \mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}] \quad P\text{-a.s. for } k = 1, \dots, T. \quad (4.9)$$

The following lemma provides several useful properties of the trading strategies and the modified price process  $S^\xi$ . We will use several results from the lemma in the Theorem 4.10 to prove the existence of the optimal strategy.

**Lemma 4.3.** *Assume  $S$  has substantial risk. Then:*

- (a)  $\delta_{k+1} S_k \in L^2(\mathbb{P})$  for  $k = 0, 1, \dots, T$  and for every  $\delta \in \Theta(S)$ .
- (b)  $\Theta(S) \subseteq \Theta(S^\xi)$  for every  $\xi \in \Xi$ .
- (c)  $V_k^\xi(\eta) \in L^2(\mathbb{P})$  for  $k = 0, 1, \dots, T$ , for every  $\xi \in \Xi$  and for every strategy  $\eta$ .
- (d)  $C_k(\eta) \in L^2(\mathbb{P})$  for  $k = 0, 1, \dots, T$  and for every strategy  $\eta$ .

*Proof.* First we prove (a). Since  $S$  has substantial risk and  $\delta \in \Theta(S)$ , we write:

$$\begin{aligned}
\mathbb{E}[(\delta_{k+1}S_k)^2] &= \mathbb{E} \left[ \delta_{k+1}^2 S_k^2 \frac{\Delta S_{k+1}^2}{\Delta S_{k+1}^2} \right] \\
&= \mathbb{E} \left[ \delta_{k+1}^2 S_k^2 \frac{\mathbb{E}[\Delta S_{k+1}^2 | \mathcal{F}_k]}{\mathbb{E}[\Delta S_{k+1}^2 | \mathcal{F}_k]} \right] \\
&= \mathbb{E} \left[ \mathbb{E} \left[ \delta_{k+1}^2 \frac{S_k^2 \Delta S_{k+1}^2}{\Delta S_{k+1}^2} | \mathcal{F}_k \right] \right] \\
&= \mathbb{E} \left[ (\delta_{k+1} \Delta S_{k+1})^2 \frac{S_k^2}{\mathbb{E}[\Delta S_{k+1}^2 | \mathcal{F}_k]} \right] \\
&\leq c_{SR} \mathbb{E}[(\delta_{k+1} \Delta S_{k+1})^2] < \infty
\end{aligned}$$

That is,  $\delta_{k+1}S_k \in L^2(\mathbb{P})$  by the definition.

We prove (b) next. Choose processes  $\xi \in \Xi$ ,  $\delta \in \Theta(S)$  and fix date  $k$ . Then  $\delta$  is predictable and  $\delta_k \Delta S_k \in L^2(\mathbb{P})$  since  $\delta \in \Theta(S)$ . Now we show  $\delta_k \Delta S_k^\xi$  is square-integrable.

$$\begin{aligned}
\delta_k \Delta S_k^\xi &= \delta_k S_k (1 + \alpha \xi_k) - \delta_k S_{k-1} (1 + \alpha \xi_{k-1}) \\
&= \delta_k S_k + \alpha \xi_k \delta_k S_k - \delta_k S_{k-1} - \alpha \xi_{k-1} \delta_k S_{k-1} \\
&= \delta_k \Delta S_k + \alpha \xi_k \delta_k S_k - \alpha \xi_k \delta_k S_{k-1} + \alpha \xi_k \delta_k S_{k-1} - \alpha \xi_{k-1} \delta_k S_{k-1} \\
&= \delta_k \Delta S_k + \alpha \xi_k \delta_k \Delta S_k + \alpha \delta_k S_{k-1} \Delta \xi_k \in \mathcal{L}^2(P)
\end{aligned}$$

since  $\xi_k$  is bounded and also by part (a).

To prove (c) we choose  $\xi \in \Xi$ , strategy  $\eta$  and date  $k$ . That is, we have separate

choices of  $\xi$  and  $\eta$ . Consider  $V_k^\xi(\eta)$ :

$$\begin{aligned}
V_k^\xi(\eta) &= \delta_{k+1}S_k^\xi + \lambda_k \\
&= \delta_{k+1}S_k(1 + \alpha\xi_k) + \lambda_k \\
&= \delta_{k+1}S_k + \alpha\xi_k\delta_{k+1}S_k + \lambda_k \\
&= V_k(\eta) + \alpha\xi_k\delta_{k+1}S_k \in L^2(\mathbb{P})
\end{aligned}$$

by part (a) and the definition of the value process  $V_k(\eta)$ .

Finally, we show that (d) holds. Choose date  $k$ , strategy  $\eta$  and  $\xi$  associated to the choice of  $\eta$ . Recall the form of  $C_k(\eta)$ :

$$\begin{aligned}
\Delta C_k(\eta) &= \Delta V_k(\eta) - \delta_k \Delta S_k \\
&\quad + 1_{\{|\Delta\delta_{k+1}| \leq N\}} \alpha S_k (\Delta\delta_{k+1})^2 + 1_{\{|\Delta\delta_{k+1}| > N\}} \alpha N S_k |\Delta\delta_{k+1}|
\end{aligned}$$

We consider two events to deal with the indicator functions  $1_{\{|\Delta\delta_{k+1}| \leq N\}}$  and  $1_{\{|\Delta\delta_{k+1}| > N\}}$ .

*Event 1:*  $|\Delta\delta_{k+1}| \leq N$ , then  $\xi_k = \Delta\delta_{k+1}$ ,  $S_k^\xi = S_k(1 + \alpha\Delta\delta_{k+1})$  and  $\Delta C_k(\eta)$

takes the form:

$$\Delta C_k(\eta) = \Delta V_k(\eta) - \delta_k \Delta S_k + \alpha S_k (\Delta\delta_{k+1})^2$$

Next we use simple algebra and the definition 4.9 of  $S_k^\xi$ :

$$\begin{aligned}
\Delta C_k(\eta) &= \Delta V_k(\eta) - \delta_k \Delta S_k + \alpha S_k (\Delta \delta_{k+1})^2 \\
&= V_k(\eta) - V_{k-1}(\eta) - \delta_k (S_k - S_{k-1}) + \alpha S_k (\delta_{k+1} - \delta_k) (\delta_{k+1} - \delta_k) \\
&= \delta_{k+1} S_k + \lambda_k - \cancel{\delta_k S_{k-1}} - \lambda_{k-1} - \delta_k S_k + \cancel{\delta_k S_{k-1}} + \alpha S_k (\delta_{k+1} - \delta_k) (\delta_{k+1} - \delta_k) \\
&= \delta_{k+1} S_k (1 + \alpha \Delta \delta_{k+1}) + \lambda_k - \delta_k S_k (1 + \alpha \Delta \delta_{k+1}) - \lambda_{k-1} \\
&= \delta_{k+1} S_k^\xi + \lambda_k - \cancel{\delta_k S_{k-1}^\xi} - \lambda_{k-1} + \cancel{\delta_k S_{k-1}^\xi} - \delta_k S_k^\xi \\
&= \Delta V_k^\xi(\eta) - \delta_k \Delta S_k^\xi
\end{aligned}$$

This ends the study of the event  $|\Delta \delta_{k+1}| \leq N$ .

*Event 2:*  $|\Delta \delta_{k+1}| > N$ , then  $\xi_k = N \operatorname{sign}(\Delta \delta_{k+1})$ ,  $S_k^\xi = S_k [1 + \alpha N \operatorname{sign}(\Delta \delta_{k+1})]$

and  $\Delta C_k(\eta)$  takes the form:

$$\begin{aligned}
\Delta C_k(\eta) &= \Delta V_k(\eta) - \delta_k \Delta S_k + \alpha N |\Delta \delta_{k+1}| \\
&= V_k(\eta) - \delta_k \Delta S_k + \alpha N \Delta \delta_{k+1} \operatorname{sign}(\Delta \delta_{k+1})
\end{aligned}$$

Next we use simple algebra and the definition 4.9 of  $S_k^\xi$ :

$$\begin{aligned}
\Delta C_k(\eta) &= \Delta V_k(\eta) - \delta_k \Delta S_k + \alpha N S_k \Delta \delta_{k+1} \text{sign}(\Delta \delta_{k+1}) \\
&= V_k(\eta) - V_{k-1}(\eta) - \delta_k (S_k - S_{k-1}) + \alpha N S_k (\delta_{k+1} - \delta_k) \text{sign}(\delta_{k+1} - \delta_k) \\
&= \delta_{k+1} S_k + \lambda_k - \cancel{\delta_k S_{k-1}} - \lambda_{k-1} - \delta_k S_k + \cancel{\delta_k S_{k-1}} \\
&\quad + \alpha N S_k \delta_{k+1} \text{sign}(\delta_{k+1} - \delta_k) - \alpha N S_k \delta_k \text{sign}(\delta_{k+1} - \delta_k) \\
&= \delta_{k+1} S_k [1 + \alpha N \text{sign}(\Delta \delta_{k+1})] + \lambda_k - \delta_k S_k [1 + \alpha N \text{sign}(\Delta \delta_{k+1})] - \lambda_{k-1} \\
&= \delta_{k+1} S_k^\xi + \lambda_k - \cancel{\delta_k S_{k-1}^\xi} - \lambda_{k-1} + \cancel{\delta_k S_{k-1}^\xi} - \delta_k S_k^\xi \\
&= \Delta V_k^\xi(\eta) - \delta_k \Delta S_k^\xi
\end{aligned}$$

This ends the consideration of the event  $|\Delta \delta_{k+1}| > N$ .

Note that  $\xi \in \Xi$  because  $\delta_{k+1}$  is predictable; previous equality and the definition of  $V_k^\xi$  show that (d) follows from parts (a) and (b). That is,  $C(\eta) \in L^2(\mathbb{P})$  as a finite linear combination of its increments that are functions in  $L^2(\mathbb{P})$ .  $\square$

**Definition 4.12.** For  $\xi \in \Xi$ , the mean-variance tradeoff process of  $S^\xi$  is

$$L_k^\xi := \sum_{j=1}^k \frac{(\mathbb{E}[\Delta S_j^\xi | \mathcal{F}_{j-1}])^2}{\text{Var}[\Delta S_j^\xi | \mathcal{F}_{j-1}]} \quad \text{for } k = 1, \dots, T.$$

I need the following assumption to show that the set of admissible strategies is the same both for the original and the modified price processes. The assumption is typical for the local risk-minimization literature.

**Definition 4.13.** *The mean-variance tradeoff process  $L_k^\xi$  is  $P$ -a.s. bounded by  $c_{MVT}(\xi)$  if  $c_{MVT}(\xi)$  is the smallest constant  $c < \infty$  such that*

$$\Delta L_k^\xi = \frac{(\mathbb{E}[\Delta S_k^\xi | \mathcal{F}_{k-1}])^2}{\text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}]} \leq c \quad \text{for } k = 1, \dots, T. \quad (4.10)$$

Again, one may rewrite the definition of the bounded mean-variance tradeoff process to interpret it as a bound on  $(\mathbb{E}[\Delta S_k^\xi | \mathcal{F}_{k-1}])^2$ :

$$(\mathbb{E}[\Delta S_k^\xi | \mathcal{F}_{k-1}])^2 \leq c \text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}] \quad \text{for } k = 1, \dots, T. \quad (4.11)$$

**Lemma 4.4.** *If  $S^\xi$  has a bounded mean-variance tradeoff process, then*

$$\mathbb{E}[(\Delta S_k^\xi)^2 | \mathcal{F}_{k-1}] \leq \text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}] (1 + c_{MVT}(\xi)) \quad \text{for } k = 1, \dots, T. \quad (4.12)$$

*Proof.* Using the definition of variance and the inequality (4.11) we get:

$$\begin{aligned} \text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}] &= \mathbb{E}[(\Delta S_k^\xi)^2 | \mathcal{F}_{k-1}] - (\mathbb{E}[\Delta S_k^\xi | \mathcal{F}_{k-1}])^2 \\ &\geq \mathbb{E}[(\Delta S_k^\xi)^2 | \mathcal{F}_{k-1}] - c_{MVT}(\xi) \text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}] \end{aligned}$$

Simple algebra gives the desired inequality (4.12). □

**Proposition 4.5.** *Fix  $\xi \in \Xi$ , assume that  $S$  has a bounded mean-variance tradeoff, substantial risk and there is a constant  $c > 0$  such that*

$$\text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}] \geq c \text{Var}[\Delta S_k | \mathcal{F}_{k-1}] \quad \text{for } k = 1, \dots, T. \quad (4.13)$$

*Then  $S^\xi$  has a bounded mean-variance tradeoff and  $\Theta(S^\xi) \subseteq \Theta(S)$ .*

*Proof.* First we show that (4.13) implies  $S^\xi$  has a bounded mean-variance tradeoff.

Fix  $\xi \in \Xi$  and a date  $k$ . Using simple algebra it is easy to show that

$$\begin{aligned}\Delta S_k^\xi &= S_k(1 + \alpha\xi_k) - S_{k-1}(1 + \alpha\xi_{k-1}) \\ &= \Delta S_k + \alpha\xi_k S_k - \alpha\xi_{k-1} S_{k-1} \\ &= \Delta S_k + \alpha\xi_k \Delta S_k + \alpha S_{k-1} \Delta\xi_k\end{aligned}$$

Recall that  $0 \leq \alpha < 1$ ,  $-N \leq \xi_k \leq N$ , thus  $\Delta\xi_k = \xi_k - \xi_{k-1} \leq 2N$  and  $(\Delta\xi_k)^2 \leq (2N)^2$ . Consider

$$\begin{aligned}(\mathbb{E}[\Delta S_k^\xi | \mathcal{F}_{k-1}])^2 &= (\mathbb{E}[\Delta S_k + \alpha\xi_k \Delta S_k + \alpha S_{k-1} \Delta\xi_k | \mathcal{F}_{k-1}])^2 \\ &= (\mathbb{E}[\Delta S_k(1 + \alpha\xi_k) + \alpha S_{k-1} \Delta\xi_k | \mathcal{F}_{k-1}])^2 \\ &= (\mathbb{E}[\Delta S_k(1 + \alpha\xi_k) | \mathcal{F}_{k-1}] + \mathbb{E}[\alpha S_{k-1} \Delta\xi_k | \mathcal{F}_{k-1}])^2 \\ &= (\mathbb{E}[\Delta S_k(1 + \alpha\xi_k) | \mathcal{F}_{k-1}])^2 + (\mathbb{E}[\alpha S_{k-1} \Delta\xi_k | \mathcal{F}_{k-1}])^2 \\ &\quad + 2 \mathbb{E}[\Delta S_k(1 + \alpha\xi_k) | \mathcal{F}_{k-1}] \mathbb{E}[\alpha S_{k-1} \Delta\xi_k | \mathcal{F}_{k-1}] \\ &\leq \mathbb{E}[\Delta S_k^2(1 + \alpha\xi_k)^2 | \mathcal{F}_{k-1}] + \mathbb{E}[\alpha^2 S_{k-1}^2 \Delta\xi_k^2 | \mathcal{F}_{k-1}] \\ &\quad + 2 \mathbb{E}[\Delta S_k(1 + \alpha\xi_k) | \mathcal{F}_{k-1}] \mathbb{E}[\alpha S_{k-1} \Delta\xi_k | \mathcal{F}_{k-1}] \\ &\leq (1 + \alpha N)^2 \mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}] + \alpha^2 S_{k-1}^2 \mathbb{E}[\Delta\xi_k^2 | \mathcal{F}_{k-1}] \\ &\quad + 2 \mathbb{E}[\Delta S_k(1 + \alpha\xi_k) | \mathcal{F}_{k-1}] \mathbb{E}[\alpha S_{k-1} \Delta\xi_k | \mathcal{F}_{k-1}] \\ &\leq (1 + \alpha N)^2 \mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}] + (2N)^2 \alpha^2 c_{SR} \mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}] \\ &\quad + 2 \mathbb{E}[\Delta S_k(1 + \alpha\xi_k) | \mathcal{F}_{k-1}] \mathbb{E}[\alpha S_{k-1} \Delta\xi_k | \mathcal{F}_{k-1}]\end{aligned}$$

Now we need to get an estimate for the cross term via  $E[\Delta S_k^2 | \mathcal{F}_{k-1}]$  and then use (4.12) to show that  $S^\xi$  has a bounded mean-variance tradeoff. We know that  $S$  has substantial risk, in other words  $S_{k-1}^2 \leq c_{SR} E[\Delta S_k^2 | \mathcal{F}_{k-1}]$ . Thus

$$\sqrt{S_{k-1}^2} = |S_{k-1}| \leq \sqrt{c_{SR}} \sqrt{E[\Delta S_k^2 | \mathcal{F}_{k-1}]}$$

and so

$$-\sqrt{c_{SR}} \sqrt{E[\Delta S_k^2 | \mathcal{F}_{k-1}]} \leq S_{k-1} \leq \sqrt{c_{SR}} \sqrt{E[\Delta S_k^2 | \mathcal{F}_{k-1}]}.$$

Also we observe that

$$\begin{aligned} E[\Delta S_k(1 + \alpha \xi_k) | \mathcal{F}_{k-1}] &\leq E[|\Delta S_k(1 + \alpha \xi_k)| | \mathcal{F}_{k-1}] \\ &\leq E[|\Delta S_k| | 1 + \alpha \xi_k | | \mathcal{F}_{k-1}] \\ &\leq |1 + \alpha N| E[|\Delta S_k| | \mathcal{F}_{k-1}] \\ &= |1 + \alpha N| \sqrt{(E[|\Delta S_k| | \mathcal{F}_{k-1}])^2} \\ &\leq |1 + \alpha N| \sqrt{E[\Delta S_k^2 | \mathcal{F}_{k-1}]} \end{aligned}$$



Using the inequalities above, the cross term is bounded in the following way:

$$\begin{aligned}
& 2\mathbf{E}[\Delta S_k(1 + \alpha\xi_k)|\mathcal{F}_{k-1}] \mathbf{E}[\alpha S_{k-1}\Delta\xi_k|\mathcal{F}_{k-1}] \\
& \leq 2|1 + \alpha N|\sqrt{\mathbf{E}[\Delta S_k^2|\mathcal{F}_{k-1}]} \alpha S_{k-1} \mathbf{E}[\Delta\xi_k|\mathcal{F}_{k-1}] \\
& \leq 4\alpha N|1 + \alpha N|\sqrt{\mathbf{E}[\Delta S_k^2|\mathcal{F}_{k-1}]} \sqrt{c_{SR}} \sqrt{\mathbf{E}[\Delta S_k^2|\mathcal{F}_{k-1}]} \\
& = 4\alpha N|1 + \alpha N|\sqrt{c_{SR}} \mathbf{E}[\Delta S_k^2|\mathcal{F}_{k-1}] \\
& \leq 4\alpha N|1 + \alpha N|\sqrt{c_{SR}} \mathbf{E}[\Delta S_k^2|\mathcal{F}_{k-1}]
\end{aligned}$$

Then by (4.12) and (4.13) we show that  $S^\xi$  has a bounded mean-variance tradeoff by definition. We assemble all the inequalities accumulated in the proof:

$$\begin{aligned}
& (\mathbf{E}[\Delta S_k^\xi|\mathcal{F}_{k-1}])^2 \\
& \leq (1 + \alpha N)^2 \mathbf{E}[\Delta S_k^2|\mathcal{F}_{k-1}] + (2N)^2 \alpha^2 c_{SR} \mathbf{E}[\Delta S_k^2|\mathcal{F}_{k-1}] \\
& \quad + 4\alpha N|1 + \alpha N|\sqrt{c_{SR}} \mathbf{E}[\Delta S_k^2|\mathcal{F}_{k-1}] \\
& \leq ((1 + \alpha N)^2 + \alpha^2(2N)^2 c_{SR} + 4\alpha N|1 + \alpha N|\sqrt{c_{SR}}) \mathbf{E}[\Delta S_k^2|\mathcal{F}_{k-1}] \\
& \leq \text{const} \text{Var}[\Delta S_k|\mathcal{F}_{k-1}] \\
& \leq \text{Var}[\Delta S_k^\xi|\mathcal{F}_{k-1}]
\end{aligned}$$

In the second part of the proof we show  $\Theta(S^\xi) \subseteq \Theta(S)$ . To this end we choose  $\xi \in \Xi$ , a strategy  $\eta = (\delta, \lambda)$  and a date  $k$  so that  $\delta_k \in \Theta(S^\xi)$ . By the Doob

decomposition theorem for a discrete process, we may write

$$S^\xi = S_0^\xi + M^\xi + A^\xi$$

where  $M^\xi$  is a martingale,  $A^\xi$  is an integrable predictable process (that is,  $A_k^\xi$  is  $\mathcal{F}_{k-1}$ -measurable) with  $A_0^\xi = 0$  and  $\Delta A_k^\xi = E[\Delta S_k^\xi | \mathcal{F}_{k-1}]$  are uniquely determined.

Then

$$\Delta S_k^\xi = \Delta M_k^\xi + \Delta A_k^\xi = \Delta M_k^\xi + E[\Delta S_k^\xi | \mathcal{F}_{k-1}]$$

We obtain the next equality by first omitting  $\mathcal{F}_{k-1}$ -measurable terms from the variance (that is,  $\Delta A_k^\xi$ ) and then recalling that  $M^\xi$  is a martingale (thus  $E[\Delta M_k^\xi | \mathcal{F}_{k-1}] = 0$ )

$$\text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}] = \text{Var}[\Delta M_k^\xi | \mathcal{F}_{k-1}] = E[(\Delta M_k^\xi)^2 | \mathcal{F}_{k-1}] \quad (4.14)$$

Of course equation (4.14) holds for the case  $\xi = 0$  resulting in the price process  $S_k$ . Since  $S^\xi$  has a bounded mean-variance tradeoff,  $\delta \in \Theta(S^\xi)$  if and only if  $\delta_k \Delta M_k^\xi \in L^2(\mathbb{P})$  for  $k = 1, \dots, T$  for which we shortly write  $\delta \in L^2(M^\xi)$ . If  $\delta$  is predictable and (4.14) holds, then by (4.13)

$$\begin{aligned} E[(\delta_k \Delta M_k) ^2 | \mathcal{F}_{k-1}] &= \delta_k^2 E[(\Delta M_k)^2 | \mathcal{F}_{k-1}] \\ &= \delta_k^2 \text{Var}[\Delta S_k | \mathcal{F}_{k-1}] \\ &\leq \delta_k^2 c \text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}] \\ &= c E[(\delta_k \Delta M_k^\xi)^2 | \mathcal{F}_{k-1}] \end{aligned}$$

shows that  $L^2(M^\xi) \subseteq L^2(M)$ , hence  $\Theta(S^\xi) \subseteq \Theta(S)$  since both mean-variance tradeoffs are bounded.  $\square$

Now we find the conditions on  $S$  and  $\alpha$  which ensure that (4.13) holds uniformly over all  $\xi \in \Xi$ .

**Proposition 4.6.** *If  $\text{Var}[\Delta S_k | \mathcal{F}_{k-1}] \neq 0$  and there is a constant  $b < 1$  such that*

$$2\alpha N \sqrt{\frac{\mathbb{E}[S_k^2 | \mathcal{F}_{k-1}]}{\text{Var}[\Delta S_k | \mathcal{F}_{k-1}]}} \leq b \quad P\text{-a.s. for } k = 1, \dots, T, \quad (4.15)$$

then (4.13) holds simultaneously for all  $\xi \in \Xi$ , with  $c = 1 - b$ .

*Proof.* Choose  $\xi \in \Xi$ . Recall

$$\Delta S_k^\xi = S_k(1 + \alpha\xi_k) - S_{k-1}(1 + \alpha\xi_{k-1}) = \Delta S_k + \alpha\xi_k S_k - \alpha\xi_{k-1} S_{k-1}$$

Using the Cauchy-Schwarz inequality one gets

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2 \text{Cov}(X, Y) \geq \text{Var}[X] - 2\sqrt{\text{Var}[X] \text{Var}[Y]}$$

since  $\text{Var}[X] > 0$  and  $\text{Var}[Y] > 0$ . Omitting  $\mathcal{F}_{k-1}$ -measurable terms from the conditional variance yields

$$\begin{aligned} \text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}] &= \text{Var}[\Delta S_k + \alpha\xi_k S_k - \alpha\xi_{k-1} S_{k-1} | \mathcal{F}_{k-1}] \\ &= \text{Var}[\Delta S_k + \alpha\xi_k S_k | \mathcal{F}_{k-1}] \\ &\geq \text{Var}[\Delta S_k | \mathcal{F}_{k-1}] - 2\alpha \sqrt{\text{Var}[\Delta S_k | \mathcal{F}_{k-1}] \text{Var}[\xi_k S_k | \mathcal{F}_{k-1}]} \end{aligned}$$

We rewrite (4.15) in an equivalent form  $E[S_k^2|\mathcal{F}_{k-1}] \leq \frac{b^2}{4\alpha^2 N^2} \text{Var}[\Delta S_k|\mathcal{F}_{k-1}]$ . Since  $\xi$  is bounded, one gets

$$\begin{aligned}
\text{Var}[\xi_k S_k|\mathcal{F}_{k-1}] &\leq E[(\xi_k S_k)^2|\mathcal{F}_{k-1}] \\
&\leq N^2 E[S_k^2|\mathcal{F}_{k-1}] \\
&\leq N^2 \frac{b^2}{4\alpha^2 N^2} \text{Var}[\Delta S_k|\mathcal{F}_{k-1}] \\
&\leq \frac{b^2}{4\alpha^2} \text{Var}[\Delta S_k|\mathcal{F}_{k-1}]
\end{aligned}$$

Finally putting everything together gives us the desired inequality:

$$\begin{aligned}
\text{Var}[\Delta S_k^\xi|\mathcal{F}_{k-1}] &\geq \text{Var}[\Delta S_k|\mathcal{F}_{k-1}] - 2\alpha\sqrt{\text{Var}[\Delta S_k|\mathcal{F}_{k-1}] \text{Var}[\xi_k S_k|\mathcal{F}_{k-1}]} \\
&\geq \text{Var}[\Delta S_k|\mathcal{F}_{k-1}] - 2\alpha\sqrt{\text{Var}[\Delta S_k|\mathcal{F}_{k-1}] \frac{b^2}{4\alpha^2} \text{Var}[\Delta S_k|\mathcal{F}_{k-1}]} \\
&= (1 - b) \text{Var}[\Delta S_k|\mathcal{F}_{k-1}]
\end{aligned}$$

□

**Proposition 4.7.** *If  $\text{Var}[\Delta S_k|\mathcal{F}_{k-1}] \neq 0$ ,  $S$  has a bounded mean-variance tradeoff, substantial risk and  $\alpha$  satisfies*

$$2\alpha N \sqrt{2 + c_{MVT}(0) + (c_{SR} + 2\sqrt{c_{SR}})(1 + c_{MVT}(0))} < 1, \quad (4.16)$$

then (4.15) holds with  $b < 1$  of the form

$$b := 2\alpha N \sqrt{2 + c_{MVT}(0) + (c_{SR} + 2\sqrt{c_{SR}})(1 + c_{MVT}(0))}.$$

*Proof.* Now we show how to obtain (4.15) from (4.16) and suggest a candidate for  $b < 1$ .

Use the estimate (4.9)

$$\begin{aligned} S_{k-1}^2 &\leq c_{SR} \mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}] \\ -\sqrt{c_{SR}} \sqrt{\mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}]} &\leq S_{k-1} \leq \sqrt{c_{SR}} \sqrt{\mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}]} \end{aligned}$$

also

$$\begin{aligned} S_{k-1} \mathbb{E}[\Delta S_k | \mathcal{F}_{k-1}] &\leq \sqrt{c_{SR}} \sqrt{\mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}]} \sqrt{(\mathbb{E}[\Delta S_k | \mathcal{F}_{k-1}])^2} \\ &\leq \sqrt{c_{SR}} \sqrt{\mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}]} \sqrt{\mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}]} \\ &= \sqrt{c_{SR}} \mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}] \end{aligned}$$

and the estimate (4.12)

$$\mathbb{E}[(\Delta S_k)^2 | \mathcal{F}_{k-1}] \leq \text{Var}[\Delta S_k | \mathcal{F}_{k-1}] (1 + c_{MVT}(0))$$

to find a bound for the numerator in (4.15):

$$\begin{aligned} &\mathbb{E}[S_k^2 | \mathcal{F}_{k-1}] \\ &= \text{Var}[\Delta S_k | \mathcal{F}_{k-1}] + (S_{k-1} + \mathbb{E}[\Delta S_k | \mathcal{F}_{k-1}])^2 \\ &= \text{Var}[\Delta S_k | \mathcal{F}_{k-1}] + S_{k-1}^2 + 2S_{k-1} \mathbb{E}[\Delta S_k | \mathcal{F}_{k-1}] + (\mathbb{E}[\Delta S_k | \mathcal{F}_{k-1}])^2 \\ &\leq \text{Var}[\Delta S_k | \mathcal{F}_{k-1}] + c_{SR} \mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}] + 2\sqrt{c_{SR}} \mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}] + \mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}] \\ &\leq \text{Var}[\Delta S_k | \mathcal{F}_{k-1}] (1 + c_{SR}(1 + c_{MVT}(0))) + 2\sqrt{c_{SR}}(1 + c_{MVT}(0)) + (1 + c_{MVT}(0)) \\ &= \text{Var}[\Delta S_k | \mathcal{F}_{k-1}] (2 + c_{MVT}(0) + (c_{SR} + 2\sqrt{c_{SR}})(1 + c_{MVT}(0))) \end{aligned}$$

Consider the left hand side of the inequality (4.15) and apply the bound for  $E[S_k^2|\mathcal{F}_{k-1}]$  we retrieved above. Then apply (4.16) to find the precise form for  $b$ :

$$\begin{aligned}
& 2\alpha N \sqrt{\frac{E[S_k^2|\mathcal{F}_{k-1}]}{\text{Var}[\Delta S_k|\mathcal{F}_{k-1}]}} \\
& \leq 2\alpha N \sqrt{\frac{\text{Var}[\Delta S_k|\mathcal{F}_{k-1}] (2 + c_{MVT}(0) + (c_{SR} + 2\sqrt{c_{SR}})(1 + c_{MVT}(0)))}{\text{Var}[\Delta S_k|\mathcal{F}_{k-1}]}} \\
& \leq 2\alpha N \sqrt{(2 + c_{MVT}(0) + (c_{SR} + 2\sqrt{c_{SR}})(1 + c_{MVT}(0)))} < 1
\end{aligned}$$

Therefore we may take  $b$  as follows, which satisfies the inequality (4.15) and also  $b < 1$ :

$$b := 2\alpha N \sqrt{2 + c_{MVT}(0) + (c_{SR} + 2\sqrt{c_{SR}})(1 + c_{MVT}(0))}$$

This ends the proof. □

Now we discuss the assumptions of substantial risk (4.8), bounded mean-variance tradeoff (4.10) and restrictions on  $\alpha N$  of the form (4.15) and (4.16).

First we show that assumption of substantial risk is equivalent to a lower bound on returns of  $S$ . Consider a general model where  $\mathbb{F}$  is generated by the process  $S$  and the returns  $\theta_k$  are i.i.d. random variables in  $L^2(\mathbb{P})$  independent of  $\mathcal{F}_{k-1}$  and distributed like some fixed random variable  $\theta$ . This implies

$$S_k = S_{k-1}(1 + \theta_k) \quad \text{for } k = 1, \dots, T.$$

Simple algebra shows  $\Delta S_k = S_k - S_{k-1} = S_{k-1}\theta_k$ . Then the definition of substantial risk (4.8) can equivalently be written as

$$\frac{S_{k-1}^2}{\mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}]} = \frac{S_{k-1}^2}{\mathbb{E}[S_{k-1}^2 \theta_k^2 | \mathcal{F}_{k-1}]} = \frac{\cancel{S_{k-1}^2}}{\cancel{S_{k-1}^2} \mathbb{E}[\theta_k^2 | \mathcal{F}_{k-1}]} \leq c$$

therefore

$$\mathbb{E}[\theta_k^2 | \mathcal{F}_{k-1}] \geq \frac{1}{c} > 0 \quad P - \text{a.s. for } k = 1, \dots, T.$$

In other words, this means that  $S$  has substantial risk if and only if we have some lower bound on returns of  $S$ .

Next we turn our attention to the geometric Brownian motion in discrete time and compute  $c_{SR}$  and  $c_{MVT}$ . One may discretize the analytical solution of the geometric Brownian motion

$$S_t = S_0 \exp \left( \left( \mu - \frac{\sigma^2}{2} \right) t + \sigma B_t \right)$$

by considering equally spaced trading times  $0 = t_0 \leq t_1 \leq \dots \leq t_t = T$ . Set  $\Delta t = t_k - t_{k-1}$  for  $k = 1, \dots, n$ . Note that  $t_k = k\Delta t$ . In terms of returns of  $S$  this is equivalent to  $1 + \theta$  being lognormally distributed with parameters  $(\mu - \sigma^2/2)\Delta t$  and  $\sigma^2\Delta t$

$$1 + \theta = \exp \left( \left( \mu - \frac{\sigma^2}{2} \right) \Delta t + \sigma \sqrt{\Delta t} Z \right).$$

where  $Z$  is a standard normal random variable. Denote  $g = \mu - \sigma^2/2$  for notational convenience. Then the final form of the discretization is

$$1 + \theta = \exp\left(g\Delta t + \sigma\sqrt{\Delta t}Z\right)$$

and

$$S_{t_k} = S_{k\Delta t} = S_0 \exp\left(gk\Delta t + \sigma\sqrt{k\Delta t}Z\right)$$

Thus the random variable  $\theta$  independent of  $\mathcal{F}_{k-1}$  representing the return of  $S$  takes the form

$$\theta = \exp\left(g\Delta t + \sigma\sqrt{\Delta t}Z\right) - 1$$

which allows us to compute  $c_{MVT}$  from (4.10):

$$\begin{aligned} c_{MVT}(0) &= \frac{(\mathbb{E}[\Delta S_k | \mathcal{F}_{k-1}])^2}{\text{Var}[\Delta S_k | \mathcal{F}_{k-1}]} = \frac{(\mathbb{E}[S_{k-1}\theta | \mathcal{F}_{k-1}])^2}{\text{Var}[S_{k-1}\theta | \mathcal{F}_{k-1}]} = \frac{\cancel{S_{k-1}^2}(\mathbb{E}[\theta])^2}{\cancel{S_{k-1}^2}\text{Var}[\theta]} \\ &= \frac{\left(\mathbb{E}\left[\exp\left(g\Delta t + \sigma\sqrt{\Delta t}Z\right) - 1\right]\right)^2}{\text{Var}\left[\exp\left(g\Delta t + \sigma\sqrt{\Delta t}Z\right) - 1\right]} = \frac{\left(e^{(g+\frac{1}{2}\sigma^2)\Delta t} - 1\right)^2}{e^{(2g+\sigma^2)\Delta t} (e^{\sigma^2\Delta t} - 1)} \\ &= \frac{\left(g + \frac{1}{2}\sigma^2\right)^2}{\sigma^2}\Delta t + O(\Delta t^2) \end{aligned}$$

As  $\Delta t$  approaches zero, a uniform bound on  $c_{MVT}(0)$  corresponds to the boundedness of the squared market price of risk (with interest rate  $r$  equal to zero)

$$\frac{\left(g + \frac{1}{2}\sigma^2\right)^2}{\sigma^2} = \left(\frac{\mu - r}{\sigma}\right)^2.$$



trading dates	$c_{MVT}(0)$	$c_{SR}$	$N$
12	0.0833333333	300	2618
52	0.0192307692	1300	1336
250	0.004	6250	623

Table 4.1: The values of parameters for monthly, weekly and daily rehedging.

Next we compute  $c_{SR}$  from (4.8):

$$\begin{aligned}
\frac{1}{c_{SR}} &= \frac{\mathbb{E}[\Delta S_k^2 | \mathcal{F}_{k-1}]}{S_{k-1}^2} = \frac{\mathbb{E}[S_{k-1}^2 \theta^2 | \mathcal{F}_{k-1}]}{S_{k-1}^2} = \frac{S_{k-1}^2 \mathbb{E}[\theta^2 | \mathcal{F}_{k-1}]}{S_{k-1}^2} \\
&= \mathbb{E}[\theta^2] = \mathbb{E} \left[ \left( \exp \left( g\Delta t + \sigma \sqrt{\Delta t} Z \right) - 1 \right)^2 \right] \\
&= \mathbb{E} \left[ \exp \left( 2g\Delta t + 2\sigma \sqrt{\Delta t} Z \right) - 2 \exp \left( g\Delta t + \sigma \sqrt{\Delta t} Z \right) + 1 \right] \\
&= \mathbb{E} \left[ e^{(2g+2\sigma^2)\Delta t} - 2e^{(g+\frac{1}{2}\sigma^2)\Delta t} + 1 \right] = \sigma^2 \Delta t + O(\Delta t^2)
\end{aligned}$$

In the table 4.1 we compute parameters of the model  $c_{MVT}(0)$  and  $c_{SR}$  for on option with maturity of one year ( $T = 1$ ),  $\mu = 0.2$ ,  $\sigma = 0.2$ ,  $\alpha = 10^{-5}$ ,  $r = 0$ . The number of rehedging times corresponds to the monthly ( $n = 12$ ), weekly ( $n = 52$ ) and daily ( $n = 250$ ) hedging schedule. Also we compute the value of  $N$  in the table from (4.16).

Consider the condition (4.15) for the geometric Brownian motion:

$$2\alpha N \sqrt{\frac{\mathbb{E}[S_k^2 | \mathcal{F}_{k-1}]}{\text{Var}[\Delta S_k | \mathcal{F}_{k-1}]}} = 2\alpha N \sqrt{\frac{e^{(2g+2\sigma^2)\Delta t}}{e^{(2g+\sigma^2)\Delta t}(e^{\sigma^2\Delta t} - 1)}} \leq b = O(1)$$

should be of the order  $O(1)$  for some  $b < 1$ . We solve the inequality above for  $\alpha N$  using  $e^{(2g+\sigma^2)\Delta t} = O(1)$ ,  $e^{(2g+2\sigma^2)\Delta t} = O(1)$  and  $e^{\sigma^2\Delta t} - 1 = \sigma^2\Delta t + O(\Delta t)^2$ :

$$\alpha N \leq b \frac{\sigma}{2} \sqrt{\Delta t} + O(\Delta t)$$

so that  $\alpha N$  should be of the order  $\sqrt{\Delta t}$ .

This ends the discussion of assumptions for the geometric Brownian motion.

### 4.3 Optimization problem: minimizing the conditional variance

The conditional variance of the cost process from the Proposition 4.2 will have a specific form (4.23) in the Theorem 4.10 from the next Section 4.4.

Let  $(\Omega, \mathcal{F}, P)$  be a probability space and  $\mathcal{G} \subseteq \mathcal{F}$  a sub- $\sigma$ -algebra of  $\mathcal{F}$ . Let  $U, Y, Z$  be  $\mathcal{F}$ -measurable real-valued random variables satisfying  $U \in L^2(\mathbb{P})$ ,  $Z \in L^2(\mathbb{P})$ ,  $ZY \in L^2(\mathbb{P})$ ,  $ZY^2 \in L^2(\mathbb{P})$  and let  $\alpha \in [0, 1)$  be a fixed real number. Assume there is a regular conditional distribution for  $(U, Y, Z)$  given  $\mathcal{G}$ . All conditional expectations, variances and covariances given  $\mathcal{G}$  which involve  $U, Y$  and  $Z$  will be computed with respect to this regular conditional distribution. We consider the conditional variance

$$f(\delta, \omega) := \text{Var}[U - \delta Z + 1_{\{|Y-\delta|>N\}}\alpha N Z | Y - \delta| + 1_{\{|Y-\delta|\leq N\}}\alpha Z (Y - \delta)^2 | \mathcal{G}](\omega). \quad (4.17)$$

We are interested in minimizing the conditional variance almost surely at some  $\delta^*(\omega)$  which is  $\mathcal{G}$ -measurable. One may employ first order conditions of optimality to characterize the minimizer  $\delta^*(\omega)$  more explicitly.

The first result of the section points out the existence of a  $\mathcal{G}$ -measurable minimizer  $\delta^*$ .

**Proposition 4.8.** *Assume that  $\text{Var}[Z|\mathcal{G}] > 0$  P-a.s. Then there exists a  $\mathcal{G}$ -measurable random variable  $\delta^*$  such that for P-almost every  $\omega$ ,*

$$f(\delta^*(\omega), \omega) \leq f(\delta, \omega) \quad \text{for all } \delta \in \mathbb{R}.$$

*Proof.* We first show that

$$\lim_{|\delta| \rightarrow \infty} f(\delta, \omega) = +\infty \quad \text{P-a.s.} \quad (4.18)$$

We split the conditional variance  $f(\delta, \omega)$  into four terms and study the limit of each term when  $|\delta| \rightarrow \infty$ . The first term will have positive infinity for its limit as desired, while the other terms will be bounded and of lower order in  $\delta$ . We use  $\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y] + \text{Cov}(X, Y)$ ,  $\text{Cov}(X+Z, Y) = \text{Cov}(X, Y) + \text{Cov}(Z, Y)$

and  $|Y - \delta| = (Y - \delta) \text{sign}(Y - \delta)$ :

$$\begin{aligned}
& \lim_{|\delta| \rightarrow \infty} f(\delta, \omega) \\
&= \lim_{|\delta| \rightarrow \infty} \text{Var} \left[ U - \delta Z + 1_{\{|Y-\delta|>N\}} \alpha N Z |Y - \delta| + 1_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \right] (\omega) \\
&= \lim_{|\delta| \rightarrow \infty} \text{Var} \left[ U - \delta Z + 1_{\{|Y-\delta|>N\}} \alpha N Z |Y - \delta| | \mathcal{G} \right] (\omega) \\
&+ \lim_{|\delta| \rightarrow \infty} \text{Var} \left[ 1_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \right] (\omega) \\
&+ \lim_{|\delta| \rightarrow \infty} \text{Cov} \left( U + \alpha N Z Y \text{sign}(Y - \delta), 1_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \right) (\omega) \\
&- \lim_{|\delta| \rightarrow \infty} \delta \text{Cov} \left( Z(1 + \alpha N \text{sign}(Y - \delta)), 1_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \right) (\omega)
\end{aligned}$$

We study the limit of the first variance now

$$\lim_{|\delta| \rightarrow \infty} \text{Var} \left[ U - \delta Z + 1_{\{|Y-\delta|>N\}} \alpha N Z |Y - \delta| | \mathcal{G} \right] (\omega)$$

and show it has  $+\infty$  as the limit.

Using  $|Y - \delta| = (Y - \delta) \text{sign}(Y - \delta)$ , we decompose the variance above into three

more terms through  $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + \text{Cov}(X, Y)$ :

$$\begin{aligned}
& \text{Var}[U - \delta Z + \alpha N Z | Y - \delta | \mathcal{G}](\omega) \\
&= \text{Var}[U - \delta Z + \alpha N Z (Y - \delta) \text{sign}(Y - \delta) | \mathcal{G}](\omega) \\
&= \text{Var}[U - \delta Z + \alpha N Z Y \text{sign}(Y - \delta) - \alpha N Z \delta \text{sign}(Y - \delta) | \mathcal{G}](\omega) \\
&= \text{Var}[U + \alpha N Z Y \text{sign}(Y - \delta) - \delta Z (1 + \alpha N \text{sign}(Y - \delta)) | \mathcal{G}](\omega) \\
&= \delta^2 \text{Var}[Z (1 + \alpha N \text{sign}(Y - \delta)) | \mathcal{G}](\omega) + \text{Var}[U + \alpha N Z Y \text{sign}(Y - \delta) | \mathcal{G}](\omega) \\
&\quad - 2\delta \text{Cov}(Z (1 + \alpha N \text{sign}(Y - \delta)), U + \alpha N Z Y \text{sign}(Y - \delta) | \mathcal{G})(\omega) \\
&:= \delta^2 h_1(\delta, \omega) + h_2(\delta, \omega) - 2\delta h_3(\delta, \omega)
\end{aligned}$$

Clearly  $\delta^2 h_1(\delta, \omega)$  is the highest-order term in  $\delta$ . For fixed  $\omega$ ,  $h_2(\delta, \omega)$  and  $h_3(\delta, \omega)$  are both bounded as functions of  $\delta$  due to integrability assumptions on  $U$ ,  $Y$ ,  $Z$ .

Now we show that the coefficient  $h_1(\delta, \omega)$  of the highest order term in  $\delta$  is positive in the limit. Using the assumption  $\text{Var}[Z | \mathcal{G}](\omega) > 0$   $P$ -a.s. and the dominated convergence (as we pulled out all unbounded  $\delta$ 's outside of  $h_1(\delta, \omega)$ ) we get:

$$\begin{aligned}
\lim_{|\delta| \rightarrow \infty} h_1(\delta, \omega) &= \lim_{|\delta| \rightarrow \pm\infty} \text{Var}[Z (1 + \alpha N \text{sign}(Y - \delta)) | \mathcal{G}](\omega) \\
&= \text{Var}[Z (1 \pm \alpha N) | \mathcal{G}](\omega) \\
&= (1 \pm \alpha N) \text{Var}[Z | \mathcal{G}](\omega) > 0 \quad P\text{-a.s.}
\end{aligned}$$

Thus the limit of the first variance is positive infinity:

$$\begin{aligned} & \lim_{|\delta| \rightarrow \infty} \text{Var} \left[ U - \delta Z + 1_{\{|Y-\delta| > N\}} \alpha N Z | Y - \delta | \mathcal{G} \right] (\omega) \\ &= \lim_{|\delta| \rightarrow \infty} \delta^2 h_1(\delta, \omega) + \lim_{|\delta| \rightarrow \infty} h_2(\delta, \omega) - 2 \lim_{|\delta| \rightarrow \infty} \delta h_3(\delta, \omega) = +\infty \end{aligned}$$

Now we show that the remaining variances/covariances in the decomposition of  $f(\delta, \omega)$  are either bounded or are maximum of the order one in  $\delta$ , while the first variance we just studied had order of two in  $\delta$  (the highest order).

Next we show that the following terms are bounded:

$$\begin{aligned} & \text{Var} \left[ 1_{\{|Y-\delta| \leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \right] (\omega), \\ & \text{Cov} \left( U + \alpha N Z Y \text{sign}(Y - \delta), 1_{\{|Y-\delta| \leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \right) (\omega), \\ & \text{Cov} \left( Z(1 + \alpha N \text{sign}(Y - \delta)), 1_{\{|Y-\delta| \leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \right) (\omega). \end{aligned}$$

Since  $|Y - \delta| \leq N$  from the indicator function,  $\alpha Z (Y - \delta)^2$  is bounded by  $\alpha Z N^2$ . At the same time the indicator function is bounded by 1. The terms  $Z(1 + \alpha N \text{sign}(Y - \delta))$  and  $U + \alpha N Z Y \text{sign}(Y - \delta)$  are bounded in  $\delta$  due to integrability assumptions on  $U, Y, Z$ .

Applying the dominated convergence theorem shows that (4.18) holds, that is:

$$\lim_{|\delta| \rightarrow \infty} f(\delta, \omega) = +\infty \quad \text{P-a.s.}$$

Now we show an existence of a minimizer. Due to the continuity of  $f(\delta, \omega)$  in  $\delta$ , we conclude that for  $P$ -almost all  $\omega$ ,  $\delta \mapsto f(\delta, \omega)$  admits a minimum. Let

$D_n := \{j2^{-n} | j \in \mathcal{Z}\}$  be the set of dyadic rationals of order  $n$ . We are using  $D_n$  to discretize the  $\delta$  variable. Using larger  $n$  increases the number of equally spaced points in the discretization. Define

$$\delta_n(\omega) := \inf\{\delta \in D_n | f(\delta, \omega) \leq f(\delta', \omega) \text{ for all } \delta' \in D_n\}$$

Since  $\omega \mapsto f(\delta, \omega)$  is  $\mathcal{G}$ -measurable for fixed  $\delta$ , the random variable  $\delta_n$  is clearly  $\mathcal{G}$ -measurable. (4.18) implies that  $(\delta_n(\omega))_{n \in \mathbb{N}}$  is bounded in  $n$  for  $P$ -almost every  $\omega$ , and from the continuity of  $f(\delta, \omega)$  in  $\delta$ , we conclude that  $\delta^* := \liminf_{n \rightarrow \infty} \delta_n$  has all the desired properties.  $\square$

We established the existence of the minimizer in the previous proposition. Next we find the derivative in  $\delta$  of the conditional variance  $f(\delta, \omega)$ .

**Lemma 4.9.** *For  $P$ -almost every  $\omega$ ,  $\delta \mapsto f(\delta, \omega)$  is a continuous function with the derivative*

$$f'(\delta, \omega) = -2 \text{Cov} \left( Z + 1_{\{|Y-\delta|>N\}} \alpha Z N \text{sign}(Y - \delta) + 1_{\{|Y-\delta|\leq N\}} 2\alpha Z (Y - \delta), \right. \\ \left. U - \delta Z + 1_{\{|Y-\delta|>N\}} \alpha Z N |Y - \delta| + 1_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \right) (\omega)$$

*Proof.* The continuity of  $f$  in  $\delta$  is obvious.

We are going to compute the derivative of  $f'(\delta, \omega)$  from the definition:

$$f'(\delta, \omega) = \lim_{h \rightarrow 0} \frac{f(\delta + h, \omega) - f(\delta, \omega)}{h}.$$

We take  $h > 0$ , while the same argument works for  $h < 0$ .

We rely on two simple facts below for  $h > 0$ :

$$\mathbf{1}_{\{|Y-\delta-h|\leq N\}} = \mathbf{1}_{\{|Y-\delta|\leq N\}} - \mathbf{1}_{\{\delta-N\leq Y<\delta-N+h\}} + \mathbf{1}_{\{\delta+N<Y\leq\delta+N+h\}}$$

$$\mathbf{1}_{\{|Y-\delta-h|>N\}} = \mathbf{1}_{\{|Y-\delta|>N\}} + \mathbf{1}_{\{\delta-N\leq Y<\delta-N+h\}} - \mathbf{1}_{\{\delta+N<Y\leq\delta+N+h\}}$$

Note that the intervals in the terms on the right hand side are identical.

Recalling that  $\text{Var}[X] - \text{Var}[Y] = \text{Cov}[X - Y, X + Y]$  and using previous two decompositions of the indicator functions, we write the partial definition of the derivative  $f'(\delta, \omega)$ :

$$\begin{aligned} & \frac{f(\delta + h, \omega) - f(\delta, \omega)}{h} \\ &= \frac{1}{h} \left( \text{Var} \left[ U - (\delta + h)Z + \mathbf{1}_{\{|Y-\delta-h|>N\}} \alpha NZ | Y - \delta - h | \right. \right. \\ & \quad \left. \left. + \mathbf{1}_{\{|Y-\delta-h|\leq N\}} \alpha Z (Y - \delta - h)^2 | \mathcal{G} \right] (\omega) \right. \\ & \quad \left. - \text{Var} \left[ U - \delta Z + \mathbf{1}_{\{|Y-\delta|>N\}} \alpha NZ | Y - \delta | + \mathbf{1}_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \right] (\omega) \right) \\ &= \frac{1}{h} \text{Cov} \left( -hZ + \mathbf{1}_{\{|Y-\delta-h|>N\}} \alpha NZ | Y - \delta - h | + \mathbf{1}_{\{|Y-\delta-h|\leq N\}} \alpha Z (Y - \delta - h)^2 \right. \\ & \quad \left. - \mathbf{1}_{\{|Y-\delta|>N\}} \alpha NZ | Y - \delta | - \mathbf{1}_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta)^2, \right. \\ & \quad \left. 2U - 2\delta Z - hZ + \mathbf{1}_{\{|Y-\delta-h|>N\}} \alpha NZ | Y - \delta - h | + \mathbf{1}_{\{|Y-\delta-h|\leq N\}} \alpha Z (Y - \delta - h)^2 \right. \\ & \quad \left. + \mathbf{1}_{\{|Y-\delta|>N\}} \alpha NZ | Y - \delta | + \mathbf{1}_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \right) (\omega) \end{aligned}$$



$$\begin{aligned}
&= \frac{1}{h} \text{Cov} \left( -hZ + 1_{\{|Y-\delta|>N\}} \alpha NZ | Y - \delta - h | \right. \\
&+ 1_{\{\delta-N \leq Y < \delta-N+h\}} \alpha NZ | Y - \delta - h | - 1_{\{\delta+N < Y \leq \delta+N+h\}} \alpha NZ | Y - \delta - h | \\
&+ 1_{\{|Y-\delta| \leq N\}} \alpha Z (Y - \delta - h)^2 \\
&- 1_{\{\delta-N \leq Y < \delta-N+h\}} \alpha Z (Y - \delta - h)^2 + 1_{\{\delta+N < Y \leq \delta+N+h\}} \alpha Z (Y - \delta - h)^2 \\
&- 1_{\{|Y-\delta|>N\}} \alpha NZ | Y - \delta | - 1_{\{|Y-\delta| \leq N\}} \alpha Z (Y - \delta)^2, \\
&2U - 2\delta Z - hZ + 1_{\{|Y-\delta|>N\}} \alpha NZ | Y - \delta - h | \\
&+ 1_{\{\delta-N \leq Y < \delta-N+h\}} \alpha NZ | Y - \delta - h | - 1_{\{\delta+N < Y \leq \delta+N+h\}} \alpha NZ | Y - \delta - h | \\
&+ 1_{\{|Y-\delta| \leq N\}} \alpha Z (Y - \delta - h)^2 \\
&- 1_{\{\delta-N \leq Y < \delta-N+h\}} \alpha Z (Y - \delta - h)^2 + 1_{\{\delta+N < Y \leq \delta+N+h\}} \alpha Z (Y - \delta - h)^2 \\
&+ 1_{\{|Y-\delta|>N\}} \alpha NZ | Y - \delta | + 1_{\{|Y-\delta| \leq N\}} \alpha Z (Y - \delta)^2 | \mathcal{G} \Big) (\omega)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{h} \text{Cov} \left( -hZ + 1_{\{|Y-\delta|>N\}} \alpha NZ |Y - \delta - h| - 1_{\{|Y-\delta|>N\}} \alpha NZ |Y - \delta| \right. \\
&+ 1_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta - h)^2 - 1_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta)^2 \\
&+ 1_{\{\delta-N \leq Y < \delta-N+h\}} \alpha NZ |Y - \delta - h| - 1_{\{\delta-N \leq Y < \delta-N+h\}} \alpha Z (Y - \delta - h)^2 \\
&+ 1_{\{\delta+N < Y \leq \delta+N+h\}} \alpha Z (Y - \delta - h)^2 - 1_{\{\delta+N < Y \leq \delta+N+h\}} \alpha NZ |Y - \delta - h|, \\
&2U - 2\delta Z - hZ + 1_{\{|Y-\delta|>N\}} \alpha NZ |Y - \delta - h| + 1_{\{|Y-\delta|>N\}} \alpha NZ |Y - \delta| \\
&+ 1_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta - h)^2 + 1_{\{|Y-\delta|\leq N\}} \alpha Z (Y - \delta)^2 \\
&+ 1_{\{\delta-N \leq Y < \delta-N+h\}} \alpha NZ |Y - \delta - h| - 1_{\{\delta-N \leq Y < \delta-N+h\}} \alpha Z (Y - \delta - h)^2 \\
&+ 1_{\{\delta+N < Y \leq \delta+N+h\}} \alpha Z (Y - \delta - h)^2 - 1_{\{\delta+N < Y \leq \delta+N+h\}} \alpha NZ |Y - \delta - h| \Big| \mathcal{G} \Big) (\omega) \\
&= \text{Cov} \left( -Z + 1_{\{|Y-\delta|>N\}} \alpha NZ \frac{|Y - \delta - h| - |Y - \delta|}{h} \right. \\
&+ 1_{\{|Y-\delta|\leq N\}} \alpha Z \frac{(Y - \delta - h)^2 - (Y - \delta)^2}{h} \\
&+ \frac{1}{h} 1_{\{\delta-N \leq Y < \delta-N+h\}} \alpha Z |Y - \delta - h| [N - |Y - \delta - h|] \\
&+ \frac{1}{h} 1_{\{\delta+N < Y \leq \delta+N+h\}} \alpha Z |Y - \delta - h| [|Y - \delta - h| - N], \\
&2U - 2\delta Z - hZ + 1_{\{|Y-\delta|>N\}} \alpha NZ [|Y - \delta - h| + |Y - \delta|] \\
&+ 1_{\{|Y-\delta|\leq N\}} \alpha Z [(Y - \delta - h)^2 + (Y - \delta)^2] \\
&+ \frac{1}{h} 1_{\{\delta-N \leq Y < \delta-N+h\}} \alpha Z |Y - \delta - h| [N - |Y - \delta - h|] \\
&+ \frac{1}{h} 1_{\{\delta+N < Y \leq \delta+N+h\}} \alpha Z |Y - \delta - h| [|Y - \delta - h| - N] \Big| \mathcal{G} \Big) (\omega)
\end{aligned}$$

Next we show that one may apply the dominated convergence theorem to interchange the limit and the integration in the covariance. We make sure the bounds

on the terms counteract the division by  $h$ .

For the first term with division by  $h$  clearly

$$||Y - \delta - h| - |Y - \delta|| \leq h,$$

while using the inequality  $|Y - \delta| \leq N$  from the indicator function of the second term and some simple algebra yields the following bound

$$\begin{aligned} 1_{\{|Y - \delta| \leq N\}} \alpha Z \frac{(Y - \delta - h)^2 - (Y - \delta)^2}{h} &= 1_{\{|Y - \delta| \leq N\}} \alpha Z \frac{-2(Y - \delta)h + h^2}{h} \\ &\leq 1_{\{|Y - \delta| \leq N\}} \alpha Z (-2N + h). \end{aligned}$$

Consider the third term next:

$$\frac{1}{h} 1_{\{\delta - N \leq Y < \delta - N + h\}} \alpha Z |Y - \delta - h| [N - |Y - \delta - h|]$$

The inequality from the indicator function may be rewritten as  $-N - h \leq Y - \delta - h < -N$ , so  $|Y - \delta - h|$  is bounded, and  $|N - |Y - \delta - h|| \leq h$ , while the indicator function itself is bounded by 1. Similar reasoning shows that the fourth term

$$\frac{1}{h} 1_{\{\delta + N < Y \leq \delta + N + h\}} \alpha Z |Y - \delta - h| [|Y - \delta - h| - N]$$

is bounded as well.

Therefore we may use the dominated convergence theorem and interchange the

limit with covariance:

$$\begin{aligned}
f'(\delta, \omega) &= \lim_{h \rightarrow 0} \frac{f(\delta + h, \omega) - f(\delta, \omega)}{h} \\
&= \text{Cov} \left( -Z - 1_{\{|Y-\delta|>N\}} \alpha N Z \text{sign}(Y - \delta) - 1_{\{|Y-\delta|\leq N\}} 2\alpha Z(Y - \delta), \right. \\
&\quad \left. 2U - 2\delta Z + 1_{\{|Y-\delta|>N\}} 2\alpha N Z |Y - \delta| + 1_{\{|Y-\delta|\leq N\}} 2\alpha Z(Y - \delta)^2 | \mathcal{G} \right) (\omega) \\
&= -2 \text{Cov} \left( Z + 1_{\{|Y-\delta|>N\}} \alpha N Z \text{sign}(Y - \delta) + 1_{\{|Y-\delta|\leq N\}} 2\alpha Z(Y - \delta), \right. \\
&\quad \left. U - \delta Z + 1_{\{|Y-\delta|>N\}} \alpha N Z |Y - \delta| + 1_{\{|Y-\delta|\leq N\}} \alpha Z(Y - \delta)^2 | \mathcal{G} \right) (\omega)
\end{aligned}$$

since

$$\begin{aligned}
\lim_{h \rightarrow 0} \frac{(Y - \delta - h)^2 - (Y - \delta)^2}{h} &= -2(Y - \delta) \\
\lim_{h \rightarrow 0} \frac{|Y - \delta - h| - |Y - \delta|}{h} &= -\text{sign}(Y - \delta)
\end{aligned}$$

This ends the proof.  $\square$

The optimality condition  $f'(\delta, \omega) = 0$  will render the candidates for the extrema of a continuous function  $f$ . Recall the form of the derivative:

$$\begin{aligned}
f'(\delta, \omega) &= -2 \text{Cov} \left( Z + 1_{\{|Y-\delta|>N\}} \alpha N Z \text{sign}(Y - \delta) + 1_{\{|Y-\delta|\leq N\}} 2\alpha Z(Y - \delta), \right. \\
&\quad \left. U - \delta Z + 1_{\{|Y-\delta|>N\}} \alpha N Z |Y - \delta| + 1_{\{|Y-\delta|\leq N\}} \alpha Z(Y - \delta)^2 | \mathcal{G} \right) (\omega) \\
&= -2 \text{Cov} \left( 1_{\{|Y-\delta|>N\}} (Z + \alpha N Z \text{sign}(Y - \delta)) + 1_{\{|Y-\delta|\leq N\}} (Z + 2\alpha Z(Y - \delta)), \right. \\
&\quad \left. 1_{\{|Y-\delta|>N\}} (U - \delta Z + \alpha N Z |Y - \delta|) + 1_{\{|Y-\delta|\leq N\}} (U - \delta Z + \alpha Z(Y - \delta)^2) | \mathcal{G} \right) (\omega)
\end{aligned}$$

We do mostly algebraic manipulations and use the property of covariance:

$$\begin{aligned}\text{Cov}(a_1X_1 + a_2X_2, b_1Y_1 + b_2Y_2) &= a_1b_1 \text{Cov}(X_1, Y_1) + a_1b_2 \text{Cov}(X_2, Y_2) \\ &\quad + a_2b_1 \text{Cov}(X_2, Y_1) + a_2b_2 \text{Cov}(X_2, Y_2)\end{aligned}$$

to represent  $f'(\delta, \omega)$  as a sum of covariances:

$$\begin{aligned}& -\frac{1}{2}f'(\delta, \omega) \\ &= \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}}(Z + \alpha NZ \text{sign}(Y - \delta)) \right. \\ &\quad \left. + \mathbf{1}_{\{|Y-\delta|\leq N\}}(Z + 2\alpha Z(Y - \delta)), \right. \\ &\quad \left. \mathbf{1}_{\{|Y-\delta|>N\}}(U - \delta Z + \alpha NZ|Y - \delta|) \right. \\ &\quad \left. + \mathbf{1}_{\{|Y-\delta|\leq N\}}(U - \delta Z + \alpha Z(Y - \delta)^2) | \mathcal{G} \right) (\omega) \\ &= \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}}Z(1 + \alpha N \text{sign}(Y - \delta)) \right. \\ &\quad \left. + \mathbf{1}_{\{|Y-\delta|\leq N\}}(Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\ &\quad \left. \mathbf{1}_{\{|Y-\delta|>N\}}(U - \delta Z + \alpha NZ(Y - \delta) \text{sign}(Y - \delta)) \right. \\ &\quad \left. + \mathbf{1}_{\{|Y-\delta|\leq N\}}(U - \delta Z + \alpha ZY(Y - \delta) - \alpha Z\delta(Y - \delta)) | \mathcal{G} \right) (\omega) \\ &= \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}}Z(1 + \alpha N \text{sign}(Y - \delta)) \right. \\ &\quad \left. + \mathbf{1}_{\{|Y-\delta|\leq N\}}(Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\ &\quad \left. \mathbf{1}_{\{|Y-\delta|>N\}}(U + \alpha NZY \text{sign}(Y - \delta) - \delta Z[1 + \alpha N \text{sign}(Y - \delta)]) \right. \\ &\quad \left. + \mathbf{1}_{\{|Y-\delta|\leq N\}}(U + \alpha ZY(Y - \delta) - \delta Z[1 + \alpha(Y - \delta)]) | \mathcal{G} \right) (\omega)\end{aligned}$$

$$\begin{aligned}
&= \text{Cov} \left( 1_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \right. \\
&\quad \left. 1_{\{|Y-\delta|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta) - \delta Z [1 + \alpha N \text{sign}(Y - \delta)]) | \mathcal{G} \right) (\omega) \\
&+ \text{Cov} \left( 1_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \right. \\
&\quad \left. 1_{\{|Y-\delta|\leq N\}} (U + \alpha Z Y (Y - \delta) - \delta Z [1 + \alpha (Y - \delta)]) | \mathcal{G} \right) (\omega) \\
&+ \text{Cov} \left( 1_{\{|Y-\delta|\leq N\}} (Z [1 + \alpha (Y - \delta)] + \alpha Z (Y - \delta)), \right. \\
&\quad \left. 1_{\{|Y-\delta|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta) - \delta Z [1 + \alpha N \text{sign}(Y - \delta)]) | \mathcal{G} \right) (\omega) \\
&+ \text{Cov} \left( 1_{\{|Y-\delta|\leq N\}} (Z [1 + \alpha (Y - \delta)] + \alpha Z (Y - \delta)), \right. \\
&\quad \left. 1_{\{|Y-\delta|\leq N\}} (U + \alpha Z Y (Y - \delta) - \delta Z [1 + \alpha (Y - \delta)]) | \mathcal{G} \right) (\omega)
\end{aligned}$$

$$\begin{aligned}
&= \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \mathbf{1}_{\{|Y-\delta|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta)) | \mathcal{G} \right) (\omega) \\
&- \delta \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)) | \mathcal{G} \right) (\omega) \\
&+ \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \mathbf{1}_{\{|Y-\delta|\leq N\}} [U + \alpha Z Y (Y - \delta)] | \mathcal{G} \right) (\omega) \\
&- \delta \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \mathbf{1}_{\{|Y-\delta|\leq N\}} Z[1 + \alpha(Y - \delta)] | \mathcal{G} \right) (\omega) \\
&+ \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
&\quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta)) | \mathcal{G} \right) (\omega) \\
&- \delta \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
&\quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} Z[1 + \alpha N \text{sign}(Y - \delta)] | \mathcal{G} \right) (\omega) \\
&+ \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \mathbf{1}_{\{|Y-\delta|\leq N\}} (U + \alpha Z Y (Y - \delta)) | \mathcal{G} \right) (\omega) \\
&- \delta \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \mathbf{1}_{\{|Y-\delta|\leq N\}} Z[1 + \alpha(Y - \delta)] | \mathcal{G} \right) (\omega)
\end{aligned}$$

Recall that

$$\xi_k := \begin{cases} \Delta \delta_{k+1}^* & : |\Delta \delta_{k+1}^*| \leq N \\ N \text{sign}(\Delta \delta_{k+1}^*) & : |\Delta \delta_{k+1}^*| > N \end{cases}$$

Also

$$U := \mathbb{E}[I_k^* | \mathcal{F}_k]$$

$$Z := S_k$$

$$Y := \delta_{k+1}^*$$

$$\delta := \delta_k^*$$

$$Y - \delta := \Delta \delta_{k+1}^*$$

$$\mathcal{G} := \mathcal{F}_{k-1}.$$

We substitute the variables for  $U, Z, Y$  to get the  $\delta$  in the notation of the discrete model, that is equation (4.23) from the next section.



Set  $f'(\delta, \omega)$  equal to zero and solve for  $\delta$ . Numerator is:

$$\begin{aligned}
& \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \mathbf{1}_{\{|Y-\delta|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta)) | \mathcal{G} \right) (\omega) \\
& + \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \mathbf{1}_{\{|Y-\delta|\leq N\}} [U + \alpha Z Y (Y - \delta)] | \mathcal{G} \right) (\omega) \\
& + \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta)) | \mathcal{G} \right) (\omega) \\
& + \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \mathbf{1}_{\{|Y-\delta|\leq N\}} (U + \alpha Z Y (Y - \delta)) | \mathcal{G} \right) (\omega) \\
& = \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} [U + \alpha Z Y (Y - \delta)] | \mathcal{G} \right) (\omega) \\
& + \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} (U + \alpha Z Y (Y - \delta)) | \mathcal{G} \right) (\omega) \\
& = \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} [U + \alpha Z Y (Y - \delta)] \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} (U + \alpha Z Y (Y - \delta)) | \mathcal{G} \right) (\omega) \\
& = \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} 2(U + \alpha N Z Y \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} 2[U + \alpha Z Y (Y - \delta)] | \mathcal{G} \right) (\omega) \\
& = 2 \text{Cov} \left( \mathbf{1}_{\{|\Delta\delta_{k+1}^*|>N\}} S_k^\xi + \mathbf{1}_{\{|\Delta\delta_{k+1}^*|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|\Delta\delta_{k+1}^*|>N\}} (U + \alpha N Z Y \text{sign}(Y - \delta)) + \mathbf{1}_{\{|\Delta\delta_{k+1}^*|\leq N\}} [U + \alpha Z Y (Y - \delta)] | \mathcal{G} \right) (\omega)
\end{aligned}$$

Denominator is:

$$\begin{aligned}
& \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)) | \mathcal{G} \right) (\omega) \\
& + \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \mathbf{1}_{\{|Y-\delta|\leq N\}} Z[1 + \alpha(Y - \delta)] | \mathcal{G} \right) (\omega) \\
& + \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} Z[1 + \alpha N \text{sign}(Y - \delta)] | \mathcal{G} \right) (\omega) \\
& + \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \mathbf{1}_{\{|Y-\delta|\leq N\}} Z[1 + \alpha(Y - \delta)] | \mathcal{G} \right) (\omega) \\
& = \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} Z[1 + \alpha(Y - \delta)] | \mathcal{G} \right) (\omega) \\
& \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} Z[1 + \alpha N \text{sign}(Y - \delta)] + \mathbf{1}_{\{|Y-\delta|\leq N\}} Z[1 + \alpha(Y - \delta)] | \mathcal{G} \right) (\omega) \\
& = \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} Z[1 + \alpha(Y - \delta)] \right. \\
& \quad \left. + \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} Z[1 + \alpha(Y - \delta)] | \mathcal{G} \right) (\omega) \\
& = 2 \text{Cov} \left( \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} (Z[1 + \alpha(Y - \delta)] + \alpha Z(Y - \delta)), \right. \\
& \quad \left. \mathbf{1}_{\{|Y-\delta|>N\}} Z(1 + \alpha N \text{sign}(Y - \delta)) + \mathbf{1}_{\{|Y-\delta|\leq N\}} Z[1 + \alpha(Y - \delta)] | \mathcal{G} \right) (\omega)
\end{aligned}$$

So the final condition for  $\delta_k^*$  is:

$$\delta_k^* = \frac{\text{Cov} \left( S_k^\xi + \mathbf{1}_{\{|\Delta\delta_{k+1}^*|\leq N\}} \alpha S_k \xi_k, \mathbb{E}[I_k^* | \mathcal{F}_k] + \alpha S_k \delta_{k+1}^* \xi_k | \mathcal{F}_{k-1} \right) (\omega)}{\text{Cov} \left( S_k^\xi + \mathbf{1}_{\{|\Delta\delta_{k+1}^*|\leq N\}} \alpha S_k \xi_k, S_k^\xi | \mathcal{F}_{k-1} \right) (\omega)}.$$

## 4.4 Existence and structure of an optimal strategy

In this section we prove the existence of a local risk-minimizing strategy under liquidity risk and describe its structure via backward induction.

We assume that  $S$  is a square-integrable process throughout the section.

**Theorem 4.10.** *Assume that  $S$  has a bounded mean-variance tradeoff, substantial risk, satisfies (4.15) and*

$$\text{Var}[\Delta S_k | \mathcal{F}_{k-1}] > 0 \quad P\text{-a.s. for } k = 1, \dots, T.$$

Then for any contingent claim  $(\delta_{T+1}^-, \bar{\lambda}_T)$ , there exists a local risk-minimizing strategy  $\eta^* = (\delta^*, \lambda^*)$  with  $\delta_{T+1}^- = \delta_{T+1}^*$ ,  $\bar{\lambda}_T = \lambda_T^*$ . Its first component  $\delta^*$  can be described as follows: define process  $\xi \in \Xi$  by setting  $\xi_0 := 0$  and

$$\xi_k := \begin{cases} \Delta \delta_{k+1}^* & : |\Delta \delta_{k+1}^*| \leq N \\ N \text{ sign}(\Delta \delta_{k+1}^*) & : |\Delta \delta_{k+1}^*| > N \end{cases} \quad P\text{-a.s. for } k = 1, \dots, T, \quad (4.19)$$

then local risk-minimizing strategy has the following structure

$$\delta_k^* = \frac{\text{Cov} \left( \Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \Delta V_k^\xi(\eta^*) | \mathcal{F}_{k-1} \right) (\omega)}{\text{Cov} \left( \Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \Delta S_k^\xi | \mathcal{F}_{k-1} \right) (\omega)} \quad (4.20)$$

$$P\text{-a.s. for } k = 1, \dots, T.$$

*Proof.* We employ the backward induction argument to demonstrate the existence of the optimal trading strategy  $\eta^* = (\delta^*, \lambda^*)$  with  $\delta_{T+1}^- = \delta_{T+1}^*$  and  $\bar{\lambda}_T = \lambda_T^*$ . The trading strategy is optimal, if it satisfies the assertions below for  $k = 0, 1, \dots, T$ :

a)  $\delta_{k+1}^* S_k \in L^2(\mathbb{P})$

b)  $I_k^* \in L^2(\mathbb{P})$ , where

$$I_k^* := \delta_{T+1}^- S_T + \bar{\lambda}_T - \sum_{j=k+1}^T \delta_j^* \Delta S_j \quad (4.21)$$

$$+ \sum_{j=k+1}^T 1_{\{|\Delta \delta_{j+1}| \leq N\}} \alpha S_j (\Delta \delta_{j+1})^2 + \sum_{j=k+1}^T 1_{\{|\Delta \delta_{j+1}| > N\}} \alpha N S_j |\Delta \delta_{j+1}|$$

c) Define the second component  $\lambda^*$  of the trading strategy  $\eta^*$  by

$$\lambda_k^* := \mathbb{E}[I_k^* | \mathcal{F}_k] - \delta_{k+1}^* S_k \in L^2(\mathbb{P}) \quad (4.22)$$

d) There exists an  $\mathcal{F}_{k-1}$ -measurable random variable  $\delta_k^*$  such that (equivalent form of (4.20))

$$\delta_k^* = \frac{\text{Cov} \left( \Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \mathbb{E}[I_k^* | \mathcal{F}_k] + \alpha S_k \delta_{k+1}^* \xi_k | \mathcal{F}_{k-1} \right) (\omega)}{\text{Cov} \left( \Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \Delta S_k^\xi | \mathcal{F}_{k-1} \right) (\omega)}. \quad (4.23)$$

e)  $\delta_k^* \Delta S_k \in L^2(\mathbb{P})$ .

f)  $\delta_k^*$  minimizes

$$\text{Var}[\mathbb{E}[I_k^* | \mathcal{F}_k] - \delta_k \Delta S_k + 1_{\{|\delta_{k+1}^* - \delta_k| \leq N\}} \alpha S_k (\delta_{k+1}^* - \delta_k)^2$$

$$+ 1_{\{|\delta_{k+1}^* - \delta_k| > N\}} \alpha N S_k |\delta_{k+1}^* - \delta_k| | \mathcal{F}_{k-1}] \quad (4.24)$$

over all  $\mathcal{F}_{k-1}$ -measurable random variables  $\delta_k$  satisfying  $\delta_k \Delta S_k \in L^2(\mathbb{P})$ .

We are going to show that if a) and b) hold for  $k$ , then there exists a  $\mathcal{F}_{k-1}$ -measurable random variable  $\delta_k^*$  satisfying c) - f) for  $k$ , and that in turn implies the validity of a) and b) for  $k - 1$ .

To start the induction we work out the base case  $k = T$ . Define  $\delta_{T+1}^* := \delta_{T+1}^-$  and  $\lambda_T^* := \lambda_T^-$ . Part a) of the theorem holds for  $k = T$  by part (a) of Lemma 4.3:  $\delta_{T+1}^* S_T \in L^2(\mathbb{P})$ . Part b) of the current theorem has the following form for  $k = T$ :  $I_T^* = H = \delta_{T+1}^* S_T + \lambda_T^* \in L^2(\mathbb{P})$ . Part b) holds by the definition of contingent claim:  $H = \delta_{T+1}^* S_T + \lambda_T^* \in L^2(\mathbb{P})$ . The process  $I_k^*$  is vital to showing that the cost process of the strategy  $\eta^* = (\delta^*, \lambda^*)$  we are about to construct,  $C(\eta^*)$  is a martingale. The process  $I_k^*$  collects and saves the previous choices for stock shares holdings  $\delta_k^*$  ( $I_k^*$  is a “backward value process” inclusive of liquidity costs), since liquidity costs are path-dependent.

Next we proceed to the inductive step. Assume a) and b) hold for  $k$ .

The item c) is just the definition of  $\lambda_k^*$ . Define the second component  $\lambda^*$  of the trading strategy  $\eta^*$  by

$$\lambda_k^* := E[I_k^* | \mathcal{F}_k] - \delta_{k+1}^* S_k$$

Note that  $\lambda_k^* \in L^2(\mathbb{P})$  by parts a) and b) for  $k$ . The construction of  $\lambda_k^*$  requires the introduction of the additional process  $I_k^*$ . Moreover, the definition of  $\eta^*$  implies

that

$$\begin{aligned}
\mathbb{E}[I_k^* | \mathcal{F}_k] + \alpha \xi_k S_k \delta_{k+1}^* &= \cancel{\delta_{k+1}^* S_k} + \alpha \xi_k S_k \delta_{k+1}^* + \mathbb{E}[I_k^* | \mathcal{F}_k] - \cancel{\delta_{k+1}^* S_k} \\
&= \delta_{k+1}^* S_k (1 + \alpha \xi_k) + \mathbb{E}[I_k^* | \mathcal{F}_k] - \delta_{k+1}^* S_k \\
&= \delta_{k+1}^* S_k^\xi + \lambda_k^* = V_k^\xi(\eta^*)
\end{aligned}$$

Using the form of  $V_k^\xi(\eta^*)$  above we are able to show that (4.23) is just a restatement of (4.20) (here we omit  $\mathcal{F}_{k-1}$ -measurable terms from the covariances):

$$\begin{aligned}
\delta_k^* &= \frac{\text{Cov}(\Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \Delta V_k^\xi(\eta^*) | \mathcal{F}_{k-1})}{\text{Cov}(\Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \Delta S_k^\xi | \mathcal{F}_{k-1})} \\
&= \frac{\text{Cov}(S_k^\xi - S_{k-1}^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, V_k^\xi(\eta^*) - V_{k-1}^\xi(\eta^*) | \mathcal{F}_{k-1})}{\text{Cov}(S_k^\xi - S_{k-1}^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, S_k^\xi - S_{k-1}^\xi | \mathcal{F}_{k-1})} \\
&= \frac{\text{Cov}(S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \mathbb{E}[I_k^* | \mathcal{F}_k] + \alpha S_k \delta_{k+1}^* \xi_k | \mathcal{F}_{k-1})}{\text{Cov}(S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, S_k^\xi | \mathcal{F}_{k-1})}.
\end{aligned}$$

Next we show d), that is, there exists a random variable  $\delta_k^*$  given by the previous formula. The existence and the form of  $\delta_k^*$  will follow from the material in Section 4.3. Define the function

$$\begin{aligned}
f_k(\delta, \omega) &:= \text{Var}[\mathbb{E}[I_k^* | \mathcal{F}_k] - \delta \Delta S_k + 1_{\{|\delta_{k+1}^* - \delta| \leq N\}} \alpha S_k (\delta_{k+1} - \delta)^2 \\
&\quad + 1_{\{|\delta_{k+1} - \delta| > N\}} \alpha N S_k |\delta_{k+1} - \delta| | \mathcal{F}_{k-1}](\omega)
\end{aligned}$$

where the conditional variances and covariances are all computed with respect to a regular conditional distribution of  $(\mathbb{E}[I_k^* | \mathcal{F}_k], S_k, \delta_{k+1}^*)$  given  $\mathcal{F}_{k-1}$ . From Proposition 4.8 in Section 4.3, we get the existence of an  $\mathcal{F}_{k-1}$ -measurable random

variable  $\delta_k^*$  such that

$$f(\delta^*(\omega), \omega) \leq f(\delta, \omega) \quad P\text{-a.s. for all } \delta. \quad (4.25)$$

and

$$f(\delta^*(\omega), \omega) = 0 \quad P\text{-a.s.}$$

Using the representation of  $\delta_k^*$  from the optimality conditions gets us precisely the form (4.23) for the  $\delta_k^*$  so that d) holds for  $k$ . Please see the end of Section 4.3 for the detailed exposition.

Now we are ready to show e) holds for  $k$ , that is  $\delta_k^* \Delta S_k \in L^2(\mathbb{P})$ . Define

$$I_k^\xi := \mathbb{E}[I_k^* | \mathcal{F}_k] + \alpha \xi_k S_k \delta_{k+1}^*$$

By item b), we get  $\mathbb{E}[I_k^* | \mathcal{F}_k] \in L^2(\mathbb{P})$ . Using the boundedness of  $\xi_k$  and the part a) for  $k$ ,  $\alpha \xi_k S_k \delta_{k+1}^* \in L^2(\mathbb{P})$  as a product of bounded and square-integrable functions. Therefore  $I_k^\xi$  is square-integrable as well:  $I_k^\xi \in L^2(\mathbb{P})$ . We use the definition of  $I_k^\xi$  to shorten the notation in the formula for  $\delta_k^*$  (4.23). Assume the correlation coefficient of the denominator in (4.23) is nonzero, that is  $\rho_d \neq 0$ . That is, the covariance in the denominator in (4.23) is nonzero:  $\text{Cov}(\Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \Delta S_k^\xi | \mathcal{F}_{k-1}) \neq 0$ . To show e) for  $k$ , namely  $\mathbb{E}[(\delta_k^* \Delta S_k)^2] \in L^2(\mathbb{P})$ , we rewrite the covariance via the correlation coefficient and variances:  $\text{Cov}(X, Y) = \rho_{XY} \sqrt{\text{Var}[X] \text{Var}[Y]}$ , use the Cauchy-Schwarz inequality in the form of  $(\text{Cov}(X, Y))^2 \leq \text{Var}[X] \text{Var}[Y]$ , estimate

(4.12) for the process  $X$  with  $\xi = 0$  then result in the following estimates:

$$\begin{aligned}
\mathbb{E}[(\delta_k^* \Delta S_k)^2] &= \mathbb{E}[(\Delta S_k)^2 (\delta_k^*)^2] \\
&= \mathbb{E} \left[ (\Delta S_k)^2 \frac{(\text{Cov}(\Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, I_k^\xi | \mathcal{F}_{k-1}))^2}{(\text{Cov}(\Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \Delta S_k^\xi | \mathcal{F}_{k-1}))^2} \right] \\
&\leq \mathbb{E} \left[ (\Delta S_k)^2 \frac{\text{Var}[\Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k | \mathcal{F}_{k-1}] \text{Var}[I_k^\xi | \mathcal{F}_{k-1}]}{\rho_d^2 \text{Var}[\Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k | \mathcal{F}_{k-1}] \text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}]} \right] \\
&= \frac{1}{\rho_d^2} \mathbb{E} \left[ \mathbb{E} \left[ (\Delta S_k)^2 \frac{\text{Var}[I_k^\xi | \mathcal{F}_{k-1}]}{\text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}]} \middle| \mathcal{F}_{k-1} \right] \right] \\
&= \frac{1}{\rho_d^2} \mathbb{E} \left[ \frac{\text{Var}[I_k^\xi | \mathcal{F}_{k-1}]}{\text{Var}[\Delta S_k^\xi | \mathcal{F}_{k-1}]} \mathbb{E} [(\Delta S_k)^2 | \mathcal{F}_{k-1}] \right] \\
&\leq \frac{1}{\rho_d^2 c} \mathbb{E} \left[ \mathbb{E}[(I_k^\xi)^2 | \mathcal{F}_{k-1}] \frac{\mathbb{E}[(\Delta S_k)^2 | \mathcal{F}_{k-1}]}{\text{Var}[\Delta S_k | \mathcal{F}_{k-1}]} \right] \\
&\leq \frac{1}{\rho_d^2 c} (1 + c_{MVT}(0)) \mathbb{E}[(I_k^\xi)^2] < \infty
\end{aligned}$$

as  $\rho_d^2$  is nonzero and  $c > 0$  according to the statement of the Proposition 4.5. Then

e) holds for  $k$ .

To show f) for  $k$ , we observe that  $\mathbb{E}[I_k^* | \mathcal{F}_k] \in \mathcal{L}^2(P)$  by part b) for  $k$ ,  $\delta_k^* \Delta S_k \in \mathcal{L}^2(P)$  by part e) for  $k$ , then  $\delta_k^*$  minimizes the variance below by the result obtained from the inequality (4.25):

$$\begin{aligned}
&\text{Var}[\mathbb{E}[I_k^* | \mathcal{F}_k] - \delta_k \Delta S_k + 1_{\{|\delta_{k+1}^* - \delta_k| \leq N\}} \alpha S_k (\delta_{k+1}^* - \delta_k)^2 \\
&\quad + 1_{\{|\delta_{k+1}^* - \delta_k| > N\}} \alpha N S_k |\delta_{k+1}^* - \delta_k| | \mathcal{F}_{k-1}] (\omega) \\
&= f_k(\delta_k(\omega), \omega) \quad P\text{-a.s.}
\end{aligned}$$



This gives f) for  $k$ .

Since  $S$  has substantial risk, one may mimic the proof of part (a) of the Lemma 4.3 to get  $\delta_k^* S_{k-1} \in \mathcal{L}^2(P)$ , which establishes a) for  $k - 1$ .

Also we obtain

$$\begin{aligned} I_{k-1}^* &= I_k^* - \delta_k^* \Delta S_k + 1_{\{|\delta_{k+1}^* - \delta_k| \leq N\}} \alpha S_k (\delta_{k+1}^* - \delta_k)^2 \\ &\quad + 1_{\{|\delta_{k+1}^* - \delta_k| > N\}} \alpha N S_k |\delta_{k+1}^* - \delta_k| \in L^2(\mathbb{P}) \end{aligned}$$

due to b) for  $k$ , e) for  $k$  and boundedness of the last two terms with the indicator function. Thus b) holds for  $k - 1$ , and this completes the induction.

Finally we show that the strategy we constructed is indeed local risk-minimizing by verifying that it satisfies Proposition 4.2. First we show that  $C(\eta^*)$  is a martingale, that is  $E[C_k(\eta^*) | \mathcal{F}_{k-1}] = C_{k-1}(\eta^*)$ . From the part c) the component  $\lambda^*$  of the trading strategy  $\eta^*$  is adapted: all the terms of  $\lambda_k^*$  are  $\mathcal{F}_k$ -measurable. Item d) shows that  $\delta_{k+1}^*$  is predictable ( $\mathcal{F}_k$ -measurable) by construction. Thus the value process of the strategy  $\eta^*$  is adapted:  $V_k(\eta^*) = \delta_{k+1}^* S_k + \lambda_k^*$  is  $\mathcal{F}_k$ -measurable. Moreover, the value process is also square-integrable  $\delta_{k+1}^* S_k + \lambda_k^* \in L^2(\mathbb{P})$  by a) and c). In part e) above we established  $\delta_k^* \Delta S_k \in \mathcal{L}^2(P)$ , that is  $\delta^* \in \Theta(S)$ . Thus  $\eta^* = (\delta^*, \lambda^*)$

is a trading strategy by the definition. Using the definitions of  $\lambda^*$  we get

$$\begin{aligned}
V_k(\eta^*) &= \delta_{k+1}^* S_k + \lambda_k^* \\
&= \delta_{k+1}^* S_k + \mathbb{E}[I_k^* | \mathcal{F}_k] - \delta_{k+1}^* S_k \\
&= \mathbb{E}[I_k^* | \mathcal{F}_k]
\end{aligned}$$

for all  $k$ . We use the definition of  $I_k^*$  in the part b) to rewrite  $C_k(\eta^*)$  using the conditional expectation:

$$\begin{aligned}
C_k(\eta^*) &= V_k(\eta^*) - \sum_{i=1}^k \delta_i^* \Delta S_i + \sum_{i=1}^k 1_{\{|\Delta \delta_{i+1}| \leq N\}} \alpha S_i (\Delta \delta_{i+1})^2 \\
&\quad + \sum_{i=1}^k 1_{\{|\Delta \delta_{i+1}| > N\}} \alpha N S_i |\Delta \delta_{i+1}| \\
&= \mathbb{E}[I_k^* | \mathcal{F}_k] - \sum_{i=1}^k \delta_i^* \Delta S_i + \sum_{i=1}^k 1_{\{|\Delta \delta_{i+1}| \leq N\}} \alpha S_i (\Delta \delta_{i+1})^2 \\
&\quad + \sum_{i=1}^k 1_{\{|\Delta \delta_{i+1}| > N\}} \alpha N S_i |\Delta \delta_{i+1}| \\
&= \mathbb{E}[\delta_{T+1}^- S_T + \bar{\lambda}_T - \sum_{j=k+1}^T \delta_j^* \Delta S_j + \sum_{j=k+1}^T 1_{\{|\Delta \delta_{j+1}| \leq N\}} \alpha S_j (\Delta \delta_{j+1})^2 \\
&\quad + \sum_{j=k+1}^T 1_{\{|\Delta \delta_{j+1}| > N\}} \alpha N S_j |\Delta \delta_{j+1}| | \mathcal{F}_k] \\
&\quad - \sum_{i=1}^k \delta_i^* \Delta S_i + \sum_{i=1}^k 1_{\{|\Delta \delta_{i+1}| \leq N\}} \alpha S_i (\Delta \delta_{i+1})^2 \\
&\quad + \sum_{i=1}^k 1_{\{|\Delta \delta_{i+1}| > N\}} \alpha N S_i |\Delta \delta_{i+1}|
\end{aligned}$$

We exploit the properties of conditional expectation below to show that  $C(\eta^*)$  is a

martingale:

$$\begin{aligned}
& \mathbb{E}[C_k(\eta^*)|\mathcal{F}_{k-1}] \\
&= \mathbb{E} \left[ \mathbb{E} \left[ \delta_{T+1}^- S_T + \bar{\lambda}_T - \sum_{j=k+1}^T \delta_j^* \Delta S_j + \sum_{j=k+1}^T 1_{\{|\Delta\delta_{j+1}| \leq N\}} \alpha S_j (\Delta\delta_{j+1})^2 \right. \right. \\
&\quad \left. \left. + \sum_{j=k+1}^T 1_{\{|\Delta\delta_{j+1}| > N\}} \alpha N S_j |\Delta\delta_{j+1}| \middle| \mathcal{F}_k \right] \middle| \mathcal{F}_{k-1} \right] \\
&\quad + \mathbb{E} \left[ - \sum_{i=1}^k \delta_i^* \Delta S_i + \sum_{i=1}^k 1_{\{|\Delta\delta_{i+1}| \leq N\}} \alpha S_i (\Delta\delta_{i+1})^2 \right. \\
&\quad \left. + \sum_{i=1}^k 1_{\{|\Delta\delta_{i+1}| > N\}} \alpha N S_i |\Delta\delta_{i+1}| \middle| \mathcal{F}_{k-1} \right] \\
&= \mathbb{E} \left[ \delta_{T+1}^- S_T + \bar{\lambda}_T - \sum_{j=k}^T \delta_j^* \Delta S_j + \sum_{j=k}^T 1_{\{|\Delta\delta_{j+1}| \leq N\}} \alpha S_j (\Delta\delta_{j+1})^2 \right. \\
&\quad \left. + \sum_{j=k}^T 1_{\{|\Delta\delta_{j+1}| > N\}} \alpha N S_j |\Delta\delta_{j+1}| \middle| \mathcal{F}_{k-1} \right] \\
&\quad - \sum_{i=1}^{k-1} \delta_i^* \Delta S_i + \sum_{i=1}^{k-1} 1_{\{|\Delta\delta_{i+1}| \leq N\}} \alpha S_i (\Delta\delta_{i+1})^2 \\
&\quad + \sum_{i=1}^{k-1} 1_{\{|\Delta\delta_{i+1}| > N\}} \alpha N S_i |\Delta\delta_{i+1}| \\
&= C_{k-1}(\eta^*)
\end{aligned}$$

Thus  $C(\eta^*)$  is a martingale. Item f) ensures that  $\delta_k^*$  complies with the second condition in the Proposition 4.2 (minimizing the conditional variance). Then by Proposition 4.2 the strategy  $\eta^*$  is local risk-minimizing. This ends the proof of the main theorem.  $\square$

The optimal local risk-minimizing trading strategy  $\eta^* = (\delta^*, \lambda^*)$  for a contingent

claim  $(\delta_{T+1}^-, \bar{\lambda}_T)$  is summarized as follows:

$$\delta_{T+1}^* := \delta_{T+1}^-$$

$$\lambda_T^* := \bar{\lambda}_T$$

$$\delta_k^* := \frac{\text{Cov}\left(\Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \Delta V_k^\xi(\eta^*) | \mathcal{F}_{k-1}\right)(\omega)}{\text{Cov}\left(\Delta S_k^\xi + 1_{\{|\Delta \delta_{k+1}^*| \leq N\}} \alpha S_k \xi_k, \Delta S_k^\xi | \mathcal{F}_{k-1}\right)(\omega)}$$

$P$ -a.s. for  $k = 1, \dots, T$

$$\lambda_k^* := \mathbb{E}[I_k^* | \mathcal{F}_k] - \delta_{k+1}^* S_k \quad P\text{-a.s. for } k = 0, 1, \dots, T-1$$

## 4.5 Algorithm for the numerical computations

Assume we are a writer of a European call option with the strike price  $K$  and expiration date  $T$ . The goal is to compute the local risk-minimization hedging strategy  $\eta = (\delta, \lambda)$ , that is  $\delta$  stock shares and  $\lambda$  units of cash. We remove the asterisk superscript in the notation for the optimal strategy to allow for space index as in  $\delta^j$ . Assume the price process  $S_t$  follows the geometric Brownian motion (2.1) with the drift  $\mu$  and volatility  $\sigma$ :

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \quad 0 \leq t \leq T,$$

where  $B$  is a standard Brownian motion. We approximate the stochastic process  $S_t = S(\omega, t)$  above via a binomial tree  $S_k^j = S(j, k)$  with a large number  $N$  of time periods. Here the discretization of  $\omega \in \Omega$  refers to the possible values of the stock

price. The choice of the parameters for the tree ensures that if the number of time periods in the tree is increased, in the limiting case, the discrete binomial process converges to the continuous process (2.1).

The numerical computations consist of the two main parts: first one computes the prescription of how many shares to hold for the hedge, and then one generates the synthetic paths of the geometric Brownian motion using the drift  $\mu$  and computes the hedging errors to check how the strategy performs. For the points of each stock price path we determine the share and cash holdings. The algorithm goes back in time and finds the  $\delta_k^j$  and  $\lambda_k^j$  via backward induction. The output of the computation are two trees stored in computer memory via two-dimensional arrays, each with  $M < N$  time periods, that approximate the stochastic processes  $\delta$  and  $\lambda$  of the local risk-minimizing strategy  $\eta = (\delta, \lambda)$ . Here  $M$  is the number of rebalancing times. Specifically, the space  $j$  indices from the tree for the price process for the underlying produces the index for the corresponding trees for share  $\delta_k^j$  and cash  $\lambda_k^j$  holdings at some trading time  $t_k$ .

Next we approximate the distribution of the price process through discrete random variables at the trading dates. There are only  $M < N$  equally spaced trading dates  $0 = t_0 < t_1 < \dots < t_{M-1} < t_M := N$  available for the hedge rebalancing. For all  $0 \leq k \leq M$ , at time  $t_k$  there are  $n_k = t_k + 1$  possible states for the stock price and given state  $j$  at time  $t_k$  in the binomial tree. The stock price can only move

to  $\beta_k = t_{k+1} - t_k + 1$  possible states at time  $t_{k+1}$  as illustrated in the Figure 4.2, where  $p_i = \binom{\beta_k-1}{i} p^{\beta_k-1-i} (1-p)^i$  for all  $0 \leq i \leq \beta_k - 1$ . The parameters  $u, d, p$  are defined as follows:

$$u = e^{\sigma\sqrt{\frac{T}{N}}}, \quad d = \frac{1}{u}, \quad p = \frac{e^{\mu\frac{T}{N}} - d}{u - d}$$

where  $\sigma$  and  $\mu$  are parameters of the geometric Brownian motion. Recall that the indices for  $\delta$  run from 1 to  $M + 1$ , while the indices for  $\lambda, S$ , etc. run from 0 to  $M$ . For example, for a call option for all  $0 \leq j \leq n_M - 1$ ,  $\delta_{M+1}^j = 1_{S_M^j > K}$  and  $\lambda_M = -K 1_{S_M^j > K}$ .

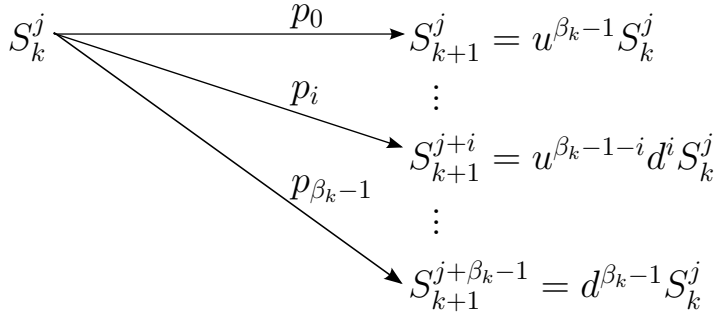


Figure 4.2: The fragment of a price process approximation through a tree.

Next we show that minimization of the conditional variance (4.24) with respect to a random variable  $\delta_k$  boils down to the unconstrained minimization of a function of one variable for each node of the tree  $\delta_{k+1}^j$ . Consider the filtration  $(\mathcal{F})_{k=0,1,\dots,M}$ , given by  $\mathcal{F}_k = \sigma(S_j | j \leq k)$ , the  $\sigma$ -field generated by the variables  $S_0, \dots, S_k$ . Here

we look for the  $\delta_k$  that minimizes the conditional variance (4.24):

$$\begin{aligned}
& \text{Var}[\mathbb{E}[I_k^*|\mathcal{F}_k] - \delta_k \Delta S_k + 1_{\{|\delta_{k+1} - \delta_k| \leq N\}} \alpha S_k (\delta_{k+1} - \delta_k)^2 \\
& \quad + 1_{\{|\delta_{k+1} - \delta_k| > N\}} \alpha N S_k |\delta_{k+1} - \delta_k| | \mathcal{F}_{k-1}] \\
& = \text{Var}[\mathbb{E}[I_k^*|\mathcal{F}_k] - \delta_k S_k + 1_{\{|\delta_{k+1} - \delta_k| \leq N\}} \alpha S_k (\delta_{k+1} - \delta_k)^2 \\
& \quad + 1_{\{|\delta_{k+1} - \delta_k| > N\}} \alpha N S_k |\delta_{k+1} - \delta_k| | \mathcal{F}_{k-1}] \\
& = \text{Var}[\mathbb{E}[I_k^*|S_k = S_k^j] - \delta_k S_k + 1_{\{|\delta_{k+1} - \delta_k| \leq N\}} \alpha S_k (\delta_{k+1} - \delta_k)^2 \\
& \quad + 1_{\{|\delta_{k+1} - \delta_k| > N\}} \alpha N S_k |\delta_{k+1} - \delta_k| | S_{k-1} = S_{k-1}^j] \\
& = \sum_{l=0}^{\beta_{k-1}-1} p_l \left( [\mathbb{E}[I_k^*|S_k = S_k^{j+l}] - \delta_k S_k^{j+l} + 1_{\{|\delta_{k+1}^{j+l} - \delta_k| \leq N\}} \alpha S_k^{j+l} (\delta_{k+1}^{j+l} - \delta_k)^2 \right. \\
& \quad \left. + 1_{\{|\delta_{k+1}^{j+l} - \delta_k| > N\}} \alpha N S_k |\delta_{k+1}^{j+l} - \delta_k| \right)^2 \\
& - \left[ \sum_{l=0}^{\beta_{k-1}-1} p_l \left( [\mathbb{E}[I_k^*|S_k = S_k^{j+l}] - \delta_k S_k^{j+l} + 1_{\{|\delta_{k+1}^{j+l} - \delta_k| \leq N\}} \alpha S_k^{j+l} (\delta_{k+1}^{j+l} - \delta_k)^2 \right. \right. \\
& \quad \left. \left. + 1_{\{|\delta_{k+1}^{j+l} - \delta_k| > N\}} \alpha N S_k |\delta_{k+1}^{j+l} - \delta_k| \right) \right]^2
\end{aligned}$$

over all  $\mathcal{F}_{k-1}$ -measurable random variables  $\delta_k$  satisfying  $\delta_k \Delta S_k \in L^2(\mathbb{P})$ . Basically this is a problem of the form “minimize  $f(x)$  over all  $x \in \mathbb{R}$ ”, since  $\delta_k$  is  $\mathcal{F}_{k-1}$ -measurable by its definition. Conditioning on the filtration  $\mathcal{F}_{k-1}$  in the context of the binomial tree basically means computing the expectations and variances separately for all the nodes  $S_{k-1}^j$  with the states  $j$  and fixed time index  $k - 1$ . Moreover,  $\delta_k$  is constant, since it is  $\mathcal{F}_{k-1}$ -measurable. At the same time random variables  $\mathbb{E}[I_k^*|S_k = S_k^j]$ ,  $\delta_{k+1}^*$ ,  $S_k$  have the same number of states, since those are

$\mathcal{F}_k$ -measurable. Thus, for each node  $j$  for a fixed time index  $k - 1$  the conditional variance for a discrete random variable is a just a function with respect to a real number  $\delta_k$ , recall the formula for variance for a discrete random variable with states  $X_i$  and corresponding probabilities  $p_i$ :  $\text{Var}[X] = \sum_i p_i [X_i^2] - (\sum_i p_i [X_i])^2$ . The random variable  $E[I_k^* | \mathcal{F}_k]$  for the “backward value process”  $I_k^*$  depends on known random variables from the previous steps according to the formula (4.21) for  $I_k^*$ , so its computation within binomial model is straightforward although tedious.

In Section 4.3 we obtained the implicit relation (4.20) that optimal  $\delta_k^j$  should satisfy for each  $k = M, \dots, 1$  and for each  $0 \leq j \leq n_k - 1$ . We employed standard analysis of finding the derivative and setting it equal to zero. Here we apply the brute force approach to find the minimum: evaluate the conditional variance for a large set of possible values for  $\delta_k$  for each node, take the minimal value and make sure it satisfies the implicit relation (4.20). That gives the values of  $\delta_k^j$ , that is the approximation of  $\delta_k$  by a discrete random variable. Finally, we find the distribution of  $\lambda_k^j$  according to the formula (4.22) for each  $k = M, \dots, 1$  and for each  $0 \leq j \leq n_k - 1$ :

$$\lambda_k^j := E[I_k^* | S_k = S_k^j] - \delta_{k+1}^j S_k^j.$$



## 5 Conclusions and future work

It is possible to successfully include the treatment of the market liquidity risk for the underlying asset in a discrete time model for hedging options. Moreover, the models lend themselves to numerical computations, which yield specific algorithms for hedging option payoffs using the underlying asset. However, there are multiple ways to assess the performance of a hedging strategy in an incomplete market. For example, while  $L^2$  norms of the hedging errors were minimal as demonstrated in the proofs, mean hedging errors showed less encouraging results, as I did not provide any theoretical background for the latter. Overall, choosing the optimality criterion depends on multiple factors: the form of the derivative payoff, maturity date, etc. The selection of the optimality criterion is nontrivial in general and beyond the scope of this thesis.

In this thesis I investigated discrete time hedging of a contingent claim under market liquidity risk. We modeled liquidity costs through a stochastic supply curve with an underlying asset price depending on order flow. That is, the purchases of

the large blocks are executed at higher prices, while sales are executed at lower prices. I used a partial differential equation to define a delta-hedging strategy and showed that the payoff of this discrete replicating strategy converges in  $L^2$  to the payoff of the option as the length of revision interval goes to zero. I introduced the class of discrete delta-hedging trading strategies with varying rebalancing times and showed the mean squared hedging error improves with the transition from equally spaced to varying rehedging times. I found an optimal hedging strategy given an initial portfolio value which minimizes the expected square error of the strategy payoff versus the option payoff. At the same time I showed the existence of the local risk-minimizing trading strategy.

This work sets out several issues for the future research and expansion. Delta hedging approach may permit generalization to delta-gamma hedging. One may also want to seek some inequalities, including bounds on option prices, to characterize the model more fully. Using coherent risk measures as a norm for hedging errors to characterize optimality of trading strategies is another possible extension of the model. One also may consider studying the path-dependent derivative payoffs within the framework of liquidity risk.

## Bibliography

- [1] Carlo Acerbi and Giacomo Scandolo. Liquidity risk theory and coherent measures of risk. *Quant. Finance*, 8(7):681–692, 2008. ISSN 1469-7688.
- [2] Aurélien Alfonsi, Antje Fruth, and Alexander Schied. Optimal execution strategies in limit order books with general shape functions. *Quant. Finance*, 10(2):143–157, 2010. ISSN 1469-7688.
- [3] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999. ISSN 1467-9965.
- [4] Ranjan Bhaduri, Gunter Meissner, and James Youn. Hedging liquidity risk: Potential solutions for hedge funds. *Journal of Alternative Investments*, pages 80–90, 2007.
- [5] Fischer Black and Myron S Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–54, May-June 1973.
- [6] Markus K. Brunnermeier and Lasse H. Pedersen. Market Liquidity and Funding Liquidity. *Review of Financial Studies*, 22:2201–2238, 2009.
- [7] Melanie Cao and Jason Wei. Option market liquidity: Commonality and other characteristics. *Journal of Financial Markets*, 13(1):20–48, February 2010.
- [8] U. Çetin, R. Jarrow, P. Protter, and M. Warachka. Pricing options in an extended black scholes economy with illiquidity: Theory and empirical evidence. *Review of Financial Studies*, 19(2):493–529, 2006.
- [9] Umut Çetin and L. C. G. Rogers. Modeling liquidity effects in discrete time. *Mathematical Finance*, 17(1):15–29, 2007. ISSN 1467-9965.
- [10] Umut Çetin, Robert A. Jarrow, and Philip Protter. Liquidity risk and arbitrage pricing theory. *Finance Stoch.*, 8(3):311–341, 2004. ISSN 0949-2984.

- [11] Umut Çetin, H. Soner, and Nizar Touzi. Option hedging for small investors under liquidity costs. *Finance and Stochastics*, 14:317–341, 2010. ISSN 0949-2984.
- [12] Peter Christoffersen, Ruslan Goyenko, Kris Jacobs, and Mehdi Karoui. Illiquidity Premia in the Equity Options Market. Technical Report 2011-43, School of Economics and Management, University of Aarhus, April 2011.
- [13] Daniel J. Duffy. *Finite difference methods in financial engineering*. Wiley Finance Series. John Wiley & Sons, Ltd., Chichester, 2006. ISBN 978-0-470-85882-0; 0-470-85882-6. A partial differential equation approach, With 1 CD-ROM (Windows, Macintosh and UNIX).
- [14] Rüdiger Frey and Alexander Stremme. Market volatility and feedback effects from dynamic hedging. *Mathematical Finance*, 7:351–374, 1996.
- [15] E. R. Grannan and G. H. Swindle. Minimizing transaction costs of option hedging strategies. *Mathematical Finance*, 6(4):341–364, 1996.
- [16] Robert A. Jarrow and Philip Protter. Liquidity risk and risk measure computation. *Review of Futures Markets*, pages 27–39, 2005.
- [17] Hyejin Ku. Liquidity risk with coherent risk measures. *Applied Mathematical Finance*, 13(2):131–141, 2006.
- [18] Hyejin Ku, Kiseop Lee, and Huaiping Zhu. Discrete time hedging with liquidity risk. *Finance Research Letters*, 9(3):135 – 143, 2012. ISSN 1544-6123.
- [19] Damien Lamberton, Huyên Pham, and Martin Schweizer. Local risk-minimization under transaction costs. *Math. Oper. Res.*, 23(3):585–612, 1998. ISSN 0364-765X.
- [20] Hayne E. Leland. Option pricing and replication with transactions costs. *The Journal of Finance*, 40(5):1283–1301, 1985. ISSN 1540-6261.
- [21] Basel Committee on Banking Supervision. Basel II capital accord: International convergence of capital measurements and capital standards: A revised framework (comprehensive version). Technical report, Basel Committee on Banking Supervision, June 2006. URL <http://www.bis.org/publ/bcbs128.pdf>.

- [22] Basel Committee on Banking Supervision. Basel III capital accord: The liquidity coverage ratio and liquidity risk monitoring tools. Technical report, Basel Committee on Banking Supervision, January 2013. URL <http://www.bis.org/publ/bcbs238.pdf>.
- [23] Martin Schweizer. Option hedging for semimartingales. *Stochastic Processes and their Applications*, 37(2):339 – 363, 1991. ISSN 0304-4149.
- [24] Steven E. Shreve. *Stochastic Calculus Models for Finance: Continuous Time Models (Springer Finance)*. Springer, jun 2004. ISBN 0387401016.
- [25] John C. Strikwerda. *Finite difference schemes and partial differential equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2004. ISBN 0-89871-567-9.
- [26] Paul Wilmott. *Frequently asked questions in quantitative finance including key models, important formulæ, common contracts, a history of quantitative finance, sundry lists, brainteasers and more*. John Wiley, Chichester, England; Hoboken, NJ, 2007. ISBN 0470058269.