Vom Fachbereich Mathematik der Technischen Universität Kaiserslautern
zur Verleihung des akademischen Grades Doktor der Naturwissenschaften
(Doctor rerum naturalium, Dr. rer. nat.) genehmigte Dissertation

Continuous-Time Portfolio Optimization under Partial Information and Convex Constraints: Deriving Explicit Results

Christian Vonwirth

Disputation: July 7th, 2017

1st Referee, Supervisor: Prof. Dr. Jörn Saß
2nd Referee: Prof. Dr. Luitgard A. M. Veraart

Technical University of Kaiserslautern

D386

TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

Abstract

In this thesis we explicitly solve several portfolio optimization problems in a very realistic setting. The fundamental assumptions on the market setting are motivated by practical experience and the resulting optimal strategies are challenged in numerical simulations.

We consider an investor who wants to maximize expected utility of terminal wealth by trading in a high-dimensional financial market with one riskless asset and several stocks. The stock returns are driven by a Brownian motion and their drift is modelled by a Gaussian random variable. We consider a partial information setting, where the drift is unknown to the investor and has to be estimated from the observable stock prices in addition to some analyst's opinion as proposed in [CLMZ06]. The best estimate given these observations is the well known Kalman-Bucy-Filter. We then consider an innovations process to transform the partial information setting into a market with complete information and an observable Gaussian drift process.

The investor is restricted to portfolio strategies satisfying several convex constraints. These constraints can be due to legal restrictions, due to fund design or due to client's specifications. We cover in particular no-short-selling and no-borrowing constraints. One popular approach to constrained portfolio optimization is the convex duality approach of Cvitanic and Karatzas. In [CK92] they introduce auxiliary stock markets with shifted market parameters and obtain a dual problem to the original portfolio optimization problem that can be better solvable than the primal problem.

Hence we consider this duality approach and using stochastic control methods we first solve the dual problems in the cases of logarithmic and power utility. Here we apply a reverse separation approach in order to obtain areas where the corresponding Hamilton-Jacobi-Bellman differential equation can be solved. It turns out that these areas have a straightforward interpretation in terms of the resulting portfolio strategy. The areas differ between active and passive stocks, where active stocks are invested in, while passive stocks are not.

Afterwards we solve the auxiliary market given the optimal dual processes in a more general setting, allowing for various market settings and various dual processes. We obtain explicit analytical formulas for the optimal portfolio policies and provide an algorithm that determines the correct formula for the optimal strategy in any case. We also show optimality of our resulting portfolio strategies in different verification theorems.

Subsequently we challenge our theoretical results in a historical and an artificial simulation that are even closer to the real world market than the setting we used to derive our theoretical results. However, we still obtain compelling results indicating that our optimal strategies can outperform any benchmark in a real market in general.

Zusammenfassung

In dieser Arbeit lösen wir explizit mehrere Portfoliooptimierungsprobleme in einer sehr realistischen Umgebung. Die grundlegenden Annahmen an den Aufbau des Marktes sind durch praktische Erfahrungen motiviert und die resultierenden optimalen Portfoliostrategien werden in numerischen Simulationen überprüft.

Wir betrachten eine Investorin, die den erwarteten Nutzen ihres Endvermögens maximieren möchte, indem sie in einem hoch-dimensionalen Finanzmarkt mit einer risikolosen Anlage und vielen Aktien handelt. Die Steigerungsraten der Aktien werden durch eine Brownsche Bewegung getrieben und ihr Drift wird durch eine normalverteilte Zufallsvariable modelliert. Da der Investorin diese Drift nicht bekannt ist, sind wir in einem Modell unter partiellen Informationen, wo die Drift durch die beobachtbaren Aktienpreise, sowie die Meinungen von Analysten geschätzt werden muss, wie es in [CLMZ06] vorgeschlagen wird. Gegeben diese Beobachtungen ist der beste Schätzer der gut bekannte Kalman-Bucy-Filter. Anschließend benutzen wir einen Innovationsprozess, um das Modell unter partiellen Informationen in einen Markt mit vollständigen Informationen und einem beobachtbaren normal-verteilten Driftprozess zu transformieren.

Die Investorin darf nur Portfoliostrategien implementieren, die gewisse konvexe Nebenbedingungen einhalten. Diese Nebenbedingungen können von gesetzlichen Vorgaben, vom Fondsdesign oder von Kundenvorgaben können. Wir behandeln insbesondere die Verbote von Leerverkäufen und dem Leihen vom Bargeld. Ein bekannter Ansatz in der Portfoliooptimierung unter Nebenbedingungen ist der konvexe Dualitätsansatz von Cvitanic und Karatzas. In [CK92] führen sie Hilfsaktienmärkte mit verschobenen Marktparametern ein und erhalten ein duales Problem zum ursprünglichen Portfoliooptimierungsproblem, das besser lösbar als das ursprüngliche Problem sein kann.

Daher betrachten wir diesen Dualitätsansatz und lösen die dualen Probleme mittels stochastischer Kontrollmethoden für logarithmischen und potenzierten Nutzen. Dabei nutzen wir einen inversen Separationsansatz, um Bereiche zu erhalten, auf denen die zugehörige Hamilton-Jacobi-Bellman Differentialgleichung gelöst werden kann. Wir stellen fest, dass diese Bereiche mit aktiven und passiven Aktien in der resultierenden Portfoliostrategie zusammenhängen, wobei die Investorin in aktive Aktien investiert und in passive nicht.

Anschließend lösen wir den Hilfsmarkt mit dem optimal dualen Prozess unter sehr allgemeinen Bedingungen, die auf viele Marktmodelle und duale Prozesse zutreffen. Wir leiten explizite analytische Formeln für die optimalen Portfoliostrategien her und erhalten einen Algorithmus, der die richtigen Formeln für die optimalen Portfoliostrategien in jedem Fall bestimmt. Außerdem zeigen wir die Optimalität unserer resultierenden Portfoliostrategien in unterschiedlichen Verifizierungssätzen.

Schließlich testen wir unsere theoretischen Ergebnisse in einer historischen und einer künstlichen Simulation, die beide sogar näher am echten Markt sind als unser theoretisches Marktmodell. Dennoch erhalten wir überzeugende Ergebnisse, die zeigen, dass unsere optimalen Strategien eine beliebige Benchmark in einem echten Markt übertreffen können.

Acknowledgements

First and foremost, I thank my supervisor Prof. Dr. Jörn Saß for his tremendous support during my time at the University of Kaiserslautern. His friendly mentoring was always determined by his generous availability and backup, accompanied by fruitful discussions on and beyond my thesis. I also thank Prof. Dr. Luitgard Veraart very much for acting as referee for my thesis.

Further gratitude goes to all of my former and current colleagues at the financial mathematics group at the University of Kaiserslautern. Together we always provided a pleasant atmosphere to work in and to discuss mathematical issues of any kind.

I gratefully acknowledge the financial and social assistance I experienced from the Studienstiftung des deutschen Volkes and from Konrad-Adenauer-Stiftung during my time in Kaiserslautern. Particular thanks goes to the CIO-Office team at (the former) DWS Investments where I experienced practical fund management during my internship that led to the idea of this thesis.

Last but not least I personally thank my family and friends and in particular my wife Tabea for the continuous support and encouragement that led to the completion of this thesis.

I couldn't have done it without you!

"One might say that to attempt to estimate the expected return on the market is to embark on a fool's errand."
[Merton, 1980]

Contents

Abstract i
Acknowledgements iii
Contents v

1 Introduction **1**
1.1 Motivation 1
1.2 Outline of this thesis 3
1.3 The market setting 4

2 Basic Results **9**
2.1 The optimal deterministic portfolio 10
2.2 Partial Information 12
2.2.1 From partial to completely observable information 14
2.3 Constrained Optimization 17

3 Constrained optimization under logarithmic utility **23**
3.1 The convex dual problem 26
3.2 Determining the correct case and active dimensions 33

4 Constrained optimization under power utility **37**
4.1 The unconstrained optimization problem 39
4.1.1 Martinagale approaches 39
4.1.2 Stochastic control approach 41
4.2 The convex dual problem: a reverse separation approach 49
4.2.1 The dual problem in Case 3 56
4.2.2 The dual problem in Case 4 64
4.3 The auxiliary market 79
4.3.1 The auxiliary market in the Cases 3 and 4 80
4.4 Determining the correct case and active dimensions 92
4.4.1 The domain of the value function 98
4.5 The complete portfolio optimization problem 106

5 Extensions **109**
5.1 Further Constraints 112
5.1.1 L^p-constraints 113

6 Simulations **121**
6.1 The historical market 122

6.1.1 Analysts and Filtering . . . 123
6.1.2 Deriving the optimal portfolio strategies . . . 125
6.2 Evaluation of the optimal portfolio strategies . . . 129
6.3 Path-wise simulation . . . 134

7 Outlook on further research **141**

A Appendix **145**
A.1 Proofs from Chapter 1 . . . 145
A.2 Proofs from Chapter 2 . . . 145
A.2.1 Convex Duality of [CK92] . . . 145
A.3 Proofs from Chapter 3 . . . 152
A.3.1 Choice of I in Case 3 . . . 152
A.3.2 Choice of I in Case 4 . . . 153
A.3.3 Dimensions on the boundary . . . 157
A.4 Proofs from Chapter 4 . . . 163
A.4.1 Martingale approach from [BDL10]: . . . 163
A.4.2 Martingale approach from [Lak98] . . . 166
A.4.3 Another stochastic control approach . . . 169
A.4.4 Choice of I in Case 3 . . . 171
A.4.5 Dimensions on the boundary in Case 3 . . . 172
A.4.6 The optimal solutions in Case 4 . . . 174
A.4.7 Shifted unconstrained solution . . . 177
A.4.8 Exact shifting in Case 4 . . . 180
A.5 Notations . . . 184

B Bibliography **187**

C Scientific Career **191**

D Declaration of Academic Honesty **193**

1 Introduction

1.1 Motivation

In financial mathematics the main objective is to develop models that describe the financial market and its actors. This way the theory of financial mathematics aims to reproduce and explain observations made in the real world financial market. In some areas like portfolio optimization an additional goal is to derive instructions for an agent who wants to optimizes his actions.
In general there are two important assumptions to meet in the set-up of a new model:

1. The model has to be realistic.
   The more properties and details of the real world market are met in the model, the better its results can be applied to a real world problem. The more simplifications and assumptions are introduced to the model, the less reliable its outcome becomes.

2. The results have to be applicable.
   In the best case scenario the resulting instructions to the agent are given analytically or explicitly so she can act immediately following precise instructions. If the results are just given up to a numerical simulation, then on the one hand the agent needs to wait before being able to apply the results and on the other hand the results are inaccurate which might lead to an inaccurate action of the agent.

Of course these assumption oppose each other since the most realistic model might be too complicated to be sufficiently solvable, while many explicitly solvable models are simplified too much to be able to adequately model the real financial market.
From a theoretical point of view the first assumption is most important because it leads to scientific advances in research. Therefore most literature on portfolio optimization focuses on developing more complex models including more details of the real market. However from a practical point of view the first assumption is only of interest if the second assumption is met after all, since otherwise there is no application possible. Subsequently the practical approach of course also wants its model to be as realistic as possible.

In this thesis the objective will be to derive a compromise such that the resulting model is as realistic as possible given the instructions to the agent can be computed explicitly in a reasonable amount of time. To our knowledge there is no literature yet that provides explicit solutions in a setting as detailed as ours.

The literature on financial markets goes back to Markowitz' Mean-Variance-Approach [Mar52] being the first result to explain the effect of diversification and deriving an optimal asset allocation in a one-period model.

Modern portfolio theory goes back in particular to Merton in [Mer71] who was among the first to set-up and solve a continuous-time portfolio optimization problem. At around the same time the infamous Black-Scholes-Formula to estimate implied volatility got derived in [BS73]. A broad introduction to the early concepts in probability theory is given in [Shi80]. Karatzas et al were among the first to derive very general solutions and methods to continuous-time portfolio optimization in [KLS87] and [KLSX91].

The Master's thesis [Von14] preceding this work was motivated by [GV13] who examine the effect of estimating unknown parameters on the utility of portfolio strategies. This thesis is motivated by the challenge of finding optimal strategies under various restrictions like partial information, convex constraints, high dimensionality and explicit computability.

The problem of estimating the unknown growth rate of the stocks was first discussed in [Mer80]. A wide summary on filtering results including the Kalman or Kalman-Bucy filter can be found in the books [EAM95], [LS01] or [BC09]. A Bayesian approach to this filter is presented in [BW96].

One of the first to formulate the continuous-time portfolio optimization problem under partial information is Lakner in [Lak95] and [Lak98]. Our setting of partial information is similar to those in [CLMZ06] and [BUV12] who use the same filter for the unknown growth rate of the stocks as we do.
There are various other papers discussing partial information and general parameter uncertainty. [Rog01] show that parameter uncertainty is way more serious than transaction costs and [KZ07] show that plug-in strategies in general won't work. Partial information in a Hidden-Markow-Model are considered for instance in [SH04] and [HPS07], while [BDL10] and [PS08] consider the growth rate to follow various stochastic differential equations. [Bre06] additionally derived the structure of the value function in a partial information setting and derived ordinary differential equations that are very similar to the differential equations that we will solve. [GKSW14] additionally introduce expert opinions to their partial information setting.

Cvitanic and Karatzas developed in [CK92] the very important duality approach to continuous-time portfolio optimization under convex constraints. A summary of methods to dealing with constrained portfolio optimization is given in [Cvi97].

Constrained portfolio optimization under partial information is dealt with in [Sas07] using a continuous-time Markov chain model for the drift. [DV15] constrain their portfolio strategies with the L^1-norm and examine the loss due to estimation.

[PS11] consider portfolio optimization under partial information where they derive dynamic constraints depending on a risk measure.

1.2 Outline of this thesis

In the following section we start by introducing the general setting of this thesis. This includes defining the market model and stating any assumptions or restrictions on the market or the acting agents. In particular we emphasise the necessity for the fund manager to be able to calculate any resulting portfolio strategy. Therefore the setting is as realistic as possible and any restricting assumptions are used to make the results analytically computable.

In the second chapter we present some basic results. We solve our portfolio optimization problem for deterministic portfolios that turns out to be the optimal stochastic portfolio at $t = 0$. We introduce our notion of partial information and how to use filtering results to get from partial to completely observable information using the innovations process. We also introduce convex constraints in the notion of [CK92] and recall their basic results on convex duality that we will apply in the following.

The third chapter is about solving the portfolio optimization problem under logarithmic utility but with a no short-selling and a no borrowing constraint. The corresponding dual problem can be solved using basic techniques from linear algebra. We present an algorithm that determines the optimal portfolio strategy explicitly.

The fourth chapter is about solving the same constrained portfolio optimization problem under power utility. We start in Section 4.1 by solving the unconstrained portfolio optimization problem under partial information using a stochastic control approach. This problem was first solved in [CLMZ06] using a martingale approach, leading to the same optimal strategy in a slightly more complicated form. In order to solve the constrained problem in Section 4.2, we consider a stochastic control approach for the corresponding dual problem. The resulting HJB-equation contains non-differentiable terms that we get rid of by a localization argument, the reverse separation approach. This leads to various HJB-equations on subsets of the parameter-space that locally solve the original HJB-equation. The corresponding primal problems are solved in Section 4.3.

The main effort to solving our optimization problem lies in finding the structure of the optimal dual process. It turns out that the investor only needs to determine which stocks are active and passive, referring to those stocks the optimal strategy does invest in respectively doesn't. This structure is mostly determined by the convex support function given the constraints. Section 4.4 is about describing this structure of the dual problem and the partitioning of the domain of the value function with respect to the structure of the optimal dual process. This is used in Section 4.5 to sketch the solution of the complete portfolio optimization problem.

In Chapter 5 we prove a generalized approach to solving the auxiliary market for any utility function. Also we consider further convex constraints for which portfolio optimization problems can be solved explicitly.

In Chapter 6 we simulate our derived optimal portfolio strategies to monitor their performance under realistic conditions. First we consider a historical market and observe an impressive outperformance of our strategies in Section 6.2. We challenge this result in Section 6.3 with simulated markets where we can still observe outperformance.

In the last chapter we present an outlook on further research that might follow this thesis and open questions that still have to be considered. The appendix begins with a short list of the most important notations used throughout this thesis and continues with a list of technical proofs of several results.

1.3 The market setting

Most of the results in the following chapters aim at solving a specific class of portfolio optimization problems. Therefore we start by introducing the basic settings of the market that we want to work with and state and reason most of the assumptions needed throughout the following chapters.

The basic idea is very conventional. We consider an investor whose objective is to maximize her initial wealth by investing in some given stock market. We want this market and the investor to behave as realistically as possible, where 'as possible' presumes that the resulting market still has to be solvable and the resulting optimal strategy has to be computable. We introduce the following properties to our market to make it more realistic:

- A high-dimensional stock market:
  Most investors are only limited to some specific sector, region or investment grade and hence remain with up to 100 or more stocks to choose from. Therefore we need to allow for a high-dimensional stock market.
  Unfortunately lots of important one-dimensional results in portfolio optimization and in particular under convex constraints cannot be transformed trivially to a high-dimensional setting. Additionally lots of results are only given implicitly and therefore need numerics to get solved. This often becomes too time-consuming when the number of stocks increases to some reasonable level.

- Only observable information:
  Another important assumption that is often neglected is to restrict the knowledge of the investor. It is fairly realistic for the investor to be able to observe the market and read out the stock prices. On the other hand it is rather optimistic to assume the investor to know the underlying market factors.
  We will specify the investor's knowledge in detail below when introducing the market. In particular this means that any resulting portfolio strategy has to be adapted to some observation filtration $\mathcal{F}^S$.

- Expert's opinions:
  The investor is allowed to use one external input other than her observations of the stock prices: there will be experts providing estimates on the future growth rates of the available stocks. In practice these estimates form the foundation of most investment decisions anyway. Also they are necessary since the investor cannot base her future investments just on historical observations since historical estimates often are too unreliable.

- Convex constraints:
  It is well known that most optimal unconstrained portfolio strategies invest huge amounts into single stocks and therefore tend to go bankrupt very easily. Also the investor might have to face several restrictions from his client, his supervisor and legal regulations.
  Hence it is reasonable to limit the admissibility set for the portfolio strategies a priori to avoid these kinds of problems anyway.

- Computable results:
  In order to generate results that can actually be applied, it is not enough to show uniqueness and existence. Often it is not even enough to determine the solution up to some implicit formula, since the subsequent numerics might be too time-consuming when the number of stocks gets larger. This is in particular important for portfolio strategies that have to be implemented almost instantly where the investor cannot wait several hours for the solution to be derived only approximately.
  Hence it is necessary to either generate analytical results or at least explicit formulas such that an algorithm does not need numerics to determine the solution.

Of course the last point contradicts the idea of modelling completely realistic, but the goal will always be to model as realistically as possible such that the computability is still given.

Our model consists of a Black-Scholes-type market with one risk-less bond $B_t \in \mathbb{R}$ and $d > 1$ risky assets $S_t \in \mathbb{R}^d$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ with dynamics

$$\begin{aligned} dB_t &= B_t r_t dt \\ dS_t &= \operatorname{diag}(S_t)\left(\mu_t dt + \sigma_t dW_t\right). \end{aligned}$$

Here $r_t \in \mathbb{R}$ is the interest rate of the bond, $\mu_t \in \mathbb{R}^d$ are the growth rates of the stocks, $\Sigma_t = \sigma_t \sigma_t^\top$ is the variance-covariance matrix of the stocks and $(W_t)_{t\in[0,T]}$ is a standard d-dimensional $\mathcal{F}$-Brownian motion.
The investor is allowed to buy and sell stocks while her remaining wealth is invested into the bond. Her portfolio strategy is given by a progressively measurable process $(\pi_t)_{t\in[0,T]}$ that satisfies the usual conditions on admissibility. $\pi_t \in \mathbb{R}^d$ are the percentages of wealth invested in the d stocks at time t. The remaining wealth $1 - \pi_t^\top \mathbb{1}$ is automatically invested in the bond.
The wealth process $(X_t)_{t\in[0,T]}$ is a real-valued non-negative process given by its dynamics with initial wealth $X_0 = x_0$:

$$dX_t = X_t\left(\left(r_t + \pi_t^\top(\mu_t - r_t\mathbb{1})\right)dt + \pi_t^\top \sigma_t dW_t\right)$$

The investor's aim will always be to maximize her expected utility of terminal wealth.

$$\pi^* = \underset{\pi\in\mathcal{A}}{\arg\max}\ \mathbb{E}\left[U(X_T^\pi)\right]$$

Here $\mathcal{A}$ is the set of admissible strategies, $U : \mathbb{R}^+ \mapsto \mathbb{R} \cup [-\infty]$ is a utility function, $[0,T]$ is the investment horizon and X_t^π is the wealth process when investing according to portfolio strategy $\pi \in \mathcal{A}$. By restricting the set $\mathcal{A}$ we can introduce any constraint on the admissible strategies.

In order to keep our market setting solvable and the resulting optimal strategies computable, we have to specify some of the previous parameters in more detail and introduce some simplifications:

- The risk-aversion α:
  The investor needs to know her (client's) risk-aversion, described by a utility function U that satisfies the usual conditions. We will mostly focus on the cases of logarithmic utility $U(x) = \log(x)$ and power utility $U(x) = \frac{1}{\alpha}x^\alpha$ for $\alpha < 0$.
  The case of logarithmic utility is almost always the limiting case of power utility for $\alpha \to 0$. Hence the risk-aversion parameter $\alpha \in (-\infty, 0]$ entirely describes the risk-aversion of the portfolio.
  Although there are good reasons for using logarithmic utility like the easy solvability of most portfolio optimization problems, there are also major drawbacks. Logarithmic utility tends to lead to very risky strategies and in our cases it mostly ignores several sources of uncertainty.
  We use negative power utility since these utility functions are more risk averse. We do not additionally calculate the respective results for positive power utility since the calculations and results are essentially the same. Also these utility functions are even less risk averse than logarithmic utility and therefore sometimes lead to degenerated solutions. However the calculations of Chapter 4 can be repeated for positive power utility as well.

- Deterministic and known market parameters: r_t and σ_t
  Obviously all market parameters evolve stochastically in the future. However we cannot model them entirely arbitrary, if we want to achieve an applicable and computable strategy. Fortunately the short term risk-free interest rate is usually known to the portfolio manager quite well and the current variance-covariance matrix of the market can usually be estimated from observing the market quite adequately. Also both parameters don't change significantly in a short period of time, when compared to the possible changes of the growth rates of the stocks.
  Therefore we model both parameters deterministic and known.

- Time-independent market parameters: $r_t = r$, $\sigma_t = \sigma$ and $\mu_t = \mu$
  Obviously the market parameters change over time. However, neither the interest rate r nor the covariance matrix $\sigma\sigma^\top$ change significantly in a short period of time, when compared to the possible changes of the growth rates of the stocks. Hence a constant approximation of these parameters is often sufficient.
  On the other hand the growth rates of the stocks are very volatile and very hard to estimate such that a time-independent average most likely won't be worse than a bad fitting time-dependent approach.
  Also we will not be able to trade continuously in practice anyway (for instance due

to transaction costs). Therefore we only need to know the average future behaviour of the stocks until our next trading opportunity. These "expected average market parameters" can hence be modelled time-independently.

- Normally distributed average growth rate μ:
  We already motivated to use a time-independent average future growth rate. However we should not consider it being deterministic as this is the most crucial parameter to our portfolio strategy.
  In practice analysts give estimates about the future behaviour of the stocks in the market. Assuming they provide their expectations μ_0 of the future growth rates μ together with their uncertainty (covariances Σ_0) of their estimates, we may model the future growth rates as a multidimensional normal distributed random variable:
  $$\mu \sim \mathcal{N}(\mu_0, \Sigma_0)$$
  where μ_0 and Σ_0 are provided by the analyst.
  On the one hand the assumption of a normal distribution is the natural choice for a parameter like the growth rate. One can hardly expect an analyst to specify a whole distribution of his expectations but rather at most two parameters. On the other hand the assumption of a normal distribution enables us to calculate explicit filtering results and therefore lots of explicit portfolio strategies.

- Time-continuous trading:
  While the continuous-time portfolio optimization problem is very well understood in general, any discrete-time factor makes handling and solving this problem significantly harder.
  Therefore we need to allow the investor to trade continuously in time, while knowing that this is not possible in practice. When actually applying a continuous portfolio strategy we need to define points of time, when the investor may update his stock positions to the current optimal amount.

- No transaction costs:
  We will also pass on introducing transaction costs into our market setting, although these may influence the structure of any time-continuous portfolio strategy considerably. However by restricting the admissibility set significantly and hence reducing the amount of wealth that can be shifted with each trade (and by not trading very often in practice) we reduce the impact of transaction costs anyway.

The market setting as described here will be used in any derivation or result in the following unless explicitly stated otherwise.

2 Basic Results

As describe in the above Section 1.3, our model consists of a multidimensional Black-Scholes-type market with one risk-less bond B and d risky assets S on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$.

$$\begin{aligned} dB_t &= B_t r dt & B(0) &= 1 \\ dS_t &= \operatorname{diag}(S_t)\, dR_t & S(0) &= S_0 \\ &= \operatorname{diag}(S_t)\,(\mu dt + \sigma dW_t) \end{aligned} \tag{2.1}$$

The initial values of the stocks are observable to the investor, the market parameters r and σ are time-independent, deterministic and observable, while μ is time-independent, but stochastic and not observable. W_t is a standard d-dimensional $\mathcal{F}$-Brownian motion that is independent of μ.
Any resulting optimal strategy has to be adapted to the observation filtration $(\mathcal{F}_t^S)$, that is generated by the observable information from the stock prices augmented by null sets.

The average future growth rate μ is modelled normally distributed

$$\mu \sim \mathcal{N}(\mu_0, \Sigma_0) \tag{2.2}$$

where μ_0 and Σ_0 are provided by the analyst being his expectation and uncertainty about the future average growth rate μ.

The investor's aim will be to maximize her expected utility of terminal wealth:

$$\pi^* = \arg\max_{\pi \in \mathcal{A}} \mathbb{E}\left[U(X_T^\pi)\right] \tag{2.3}$$

where $\mathcal{A}$ is the set of admissible strategies, $U : \mathbb{R}^+ \mapsto \mathbb{R} \cup [-\infty]$ is a utility function, $[0, T]$ is the investment horizon and X_t^π is the wealth process when investing according to portfolio strategy $\pi \in \mathcal{A}$.

2.1 The optimal deterministic portfolio

As a motivation we will first calculate the optimal deterministic portfolio strategy in our market setting with stochastic growth rate μ. Later on this strategy can be used as a benchmark to the stochastic strategy. We consider the wealth process X_t for a self-financing trading strategy π_t:

$$dX_t = X_t\left(\left(r + \pi_t^\top(\mu - r\mathbb{1})\right)dt + \pi_t^\top \sigma dW_t\right), \qquad X_0 = x_0,$$
$$X_t = x_0 \exp\left(rt + \int_0^t \left(\pi_s^\top(\mu - r\mathbb{1}) - \frac{1}{2}\pi_s^\top \Sigma \pi_s\right) ds + \int_0^t \pi_s^\top \sigma dW_s\right)$$

For logarithmic utility we maximize the following expected utility of terminal wealth:

$$\begin{aligned}\mathbb{E}\left[\log X_T\right] &= \log x_0 + rT + \mathbb{E}\left[\int_0^T \left(\pi_t^\top(\mu - r\mathbb{1}) - \frac{1}{2}\pi_t^\top \Sigma \pi_t\right) dt\right] + \mathbb{E}\left[\int_0^T \pi_t^\top \sigma dW_t\right] \\ &= \log x_0 + rT + \int_0^T \left(\pi_t^\top(\mu_0 - r\mathbb{1}) - \frac{1}{2}\pi_t^\top \Sigma \pi_t\right) dt\end{aligned}$$

Hence we want to maximize the integral over $\left(\pi_t^\top(\mu_0 - r\mathbb{1}) - \frac{1}{2}\pi_t^\top \Sigma \pi_t\right)$ for all admissible processes π. Point-wise maximization of this term already leads to an admissible solution, hence we consider the first order condition: $0 \overset{!}{=} (\mu_0 - r\mathbb{1})T - \Sigma\pi_t T$. Since Σ is symmetric and positive definite the second order condition ensures the optimizer to be a maximizer:

$$\pi^*_{\mathrm{log,det}} = \Sigma^{-1}(\mu_0 - r\mathbb{1})$$

Hence the optimal deterministic strategy is the Merton Plug-In strategy. This is shown in more Detail in [BUV12, (3.6)]. However, note that the uncertainty of the analyst (the covariance matrix Σ_0) doesn't even enter the formula supporting the idea that logarithmic utility is not very risk averse.

For power utility we need to maximize the following expected utility of terminal wealth:

$$\mathbb{E}\left[\frac{1}{\alpha}X_T^\alpha\right] = \frac{1}{\alpha}x_0^\alpha \mathbb{E}\left[\exp\left(\alpha rT + \int_0^T \alpha\pi_t^\top(\mu - r\mathbb{1})dt - \int_0^T \frac{\alpha}{2}\pi_t^\top \Sigma \pi_t dt + \int_0^T \alpha \pi_t^\top \sigma W_t\right)\right]$$

For simplicity we will only consider the optimal time-independent deterministic strategy such that the expectation value becomes

$$\frac{1}{\alpha}x_0^\alpha \exp\left(\alpha rT + \alpha\pi^\top(\mu_0 - r\mathbb{1})T + \frac{1}{2}\alpha^2T^2\pi^\top\Sigma_0\pi - \frac{1}{2}\alpha T\pi^\top\Sigma\pi + \frac{1}{2}\alpha^2 T\pi^\top\Sigma\pi\right).$$

Now for both, positive and negative α, maximizing $\mathbb{E}\left[\frac{1}{\alpha}X_T^\alpha\right]$ is equivalent to maximizing the inner term $\pi^\top(\mu_0 - r\mathbb{1})T + \frac{1}{2}\alpha T^2\pi^\top\Sigma_0\pi - \frac{1}{2}T\pi^\top\Sigma\pi + \frac{1}{2}\alpha T\pi^\top\Sigma\pi$. Again we consider the first order condition: $0 \overset{!}{=} (\mu_0 - r\mathbb{1}) + \alpha T\Sigma_0\pi - \Sigma\pi + \alpha\Sigma\pi = (\mu_0 - r\mathbb{1}) + (\alpha T\Sigma_0 - (1-\alpha)\Sigma)\,\pi$.

Here we need to assume that α is small enough or negative such that $(1-\alpha)\Sigma - \alpha T\Sigma_0$ is positive definite and the second order condition ensures the optimizer to be a maximizer. The optimal portfolio strategy then becomes:

$$\begin{aligned}
\pi^*_{\text{det,pow}} &= \left(\alpha T\Sigma_0 - (1-\alpha)\Sigma\right)^{-1}(\mu_0 - r\mathbb{1}) \\
&= \frac{1}{1-\alpha}\left(\Sigma - \frac{\alpha}{1-\alpha}T\Sigma_0\right)^{-1}(\mu_0 - r\mathbb{1}) \\
&= \frac{1}{1-\alpha}\Sigma^{-1}\left(\Sigma_0^{-1} - \frac{\alpha}{1-\alpha}T\Sigma^{-1}\right)^{-1}\Sigma_0^{-1}(\mu_0 - r\mathbb{1})
\end{aligned}$$

Now for power utility the optimal portfolio strategy does account for the uncertainty Σ_0 of the analyst, shifting the variance Σ by some fraction of Σ_0. In case of negative power utility the optimal strategy gets reduced compared to the logarithmic case above. Positive power utility vice versa leads to increasing the optimal investment, hence increasing the involved risk.

Obviously, a deterministic portfolio strategy cannot account for any observations of the market evolution, therefore this approach does not lead to satisfactory results. However, these optimal strategies both start at $t = 0$ with the same values than their corresponding optimal stochastic strategies that we will see later on.

2.2 Partial Information

Our market setting is in particular motivated by [CLMZ06] and [BUV12] where the investor doesn't know all the values of the underlying processes and random variables which is interpreted as 'partial information'. This section will give a short introduction to the theoretical background on partial information and we prove some results that are used within this thesis.
Since the resulting portfolio strategies have to be adapted to the observation filtration $(\mathcal{F}_t^S)_{t\in[0,T]}$, the best estimator for μ at time t is the conditional expectation given the information up to time t. This conditional expectation is called the 'filter'.

$$\hat{\mu}_t := \mathbb{E}\left[\mu | \mathcal{F}_t^S\right] \tag{2.4}$$

As pointed out by [CLMZ06] this filter can be motivated via a Bayesian approach with normal prior and posterior distribution:

Proposition 2.1
Let $R_t = \mu t + \sigma W_t$ be the observable return on the stocks and let the prior distribution of μ be $\mathcal{N}(\mu_0, \Sigma_0)$. Then the filter $\hat{\mu}_t$ is the posterior distribution of μ given the observation until time t that is normally distributed as follows:

$$\mu | R_t \sim \mathcal{N}\Big(\gamma_t \left(\Sigma_0^{-1}\mu_0 + \Sigma^{-1}R_t\right), \gamma_t\Big) \tag{2.5}$$
$$\text{where} \quad \gamma_t = \left(\Sigma_0^{-1} + t\Sigma^{-1}\right)^{-1}$$

Proof. The Theorem of Bayes claims for conditional densities:

$$f_{\mu|R_t}(x) = \frac{f_{R_t|\mu}(R_t) f_\mu(x)}{f_{R_t}(R_t)} \propto f_{R_t|\mu}(R_t) f_\mu(x)$$

where $\propto$ denotes 'proportional to', hence equal up to some constant.
With $\mu \sim \mathcal{N}(\mu_0, \Sigma_0)$ and $R_t|\mu \sim \mathcal{N}(\mu t, \Sigma t)$ we get:

$$\begin{aligned}
f_{\mu|R_t}(x) &\propto f_{R_t|\mu}(R_t) f_\mu(x) \\
&\propto \exp\left(-\frac{1}{2t}(R_t - xt)^\top \Sigma^{-1}(R_t - xt) - \frac{1}{2}(x-\mu_0)^\top \Sigma_0^{-1}(x-\mu_0)\right) \\
&\propto \exp\left(-\frac{1}{2}\left(x^\top \left(t\Sigma^{-1} + \Sigma_0^{-1}\right) x - 2x^\top \left(\Sigma^{-1}R_t - \Sigma_0^{-1}\mu_0\right)\right)\right) \\
&\propto \exp\Bigg(-\frac{1}{2}\bigg(\left(x - \left(\Sigma_0^{-1} + t\Sigma^{-1}\right)^{-1}\left(\Sigma_0^{-1}\mu_0 + \Sigma^{-1}R_t\right)\right)^\top \left(t\Sigma^{-1} + \Sigma_0^{-1}\right) \\
&\qquad \cdot \left(x - \left(\Sigma_0^{-1} + t\Sigma^{-1}\right)^{-1}\left(\Sigma_0^{-1}\mu_0 + \Sigma^{-1}R_t\right)\right)\bigg)\Bigg) \\
&\propto \exp\left(-\frac{1}{2}\left((x-\hat{\mu}_t)^\top \gamma_t^{-1}(x-\hat{\mu}_t)\right)\right)
\end{aligned}$$

As $f_{\mu|R_t}$ has to be a density and the resulting formula is proportional to the density of a normal distribution, the assumption follows immediately. □

Alternatively it is well known that this filter can also be calculated as a degenerated version of the Kalman-Bucy filter (cf. [BUV12]):

Lemma 2.2
Let $R_t := \mu t + \sigma W_t$ be the observable returns on the stocks and $\mu \sim \mathcal{N}(\mu_0, \Sigma_0)$. Then

$$\begin{aligned} \hat{\mu}_t &:= \mathbb{E}\left[\mu | \mathcal{F}_t^S\right] = \gamma_t \left(\Sigma_0^{-1}\mu_0 + \Sigma^{-1}R_t\right) \\ \gamma_t &:= \mathrm{Var}\big(\mu|\mathcal{F}_t^S\big) = \left(\Sigma_0^{-1} + t\Sigma^{-1}\right)^{-1} \end{aligned} \tag{2.6}$$

where γ_t is the conditional variance given the information in $\mathcal{F}_t^S$.

Proof. We prove this by one step of the (degenerated) time-discrete Kalman-Filter as described in chapter 4.5 in [EAM95] with signal $\mu \sim \mathcal{N}(\mu_0, \Sigma_0)$ and observation R_t.
We set their parameters to $x_0 = \mu, A = Id, Q = 0$ for the signal and $C = t \cdot Id, R = t\Sigma$ for the observation. Following formulas (5.12) and (5.13) in [EAM95] we start with the first 'time updates' $\hat{\mu}_{1|0} = \mu_0$ and $\hat{\Sigma}_{1|0} = \Sigma_0$ and we follow with

$$\begin{aligned} \mathbb{E}\left[\mu | \mathcal{F}_t^S\right] = \hat{\mu}_t &= \mu_0 + \Sigma_0 t \left(t\Sigma_0 t + \Sigma t\right)^{-1}\left(R_t - t\mu_0\right) \\ &= \mu_0 + \left(tI + \Sigma\Sigma_0^{-1}\right)^{-1}\left(R_t - t\mu_0\right) \\ &= \left(tI + \Sigma\Sigma_0^{-1}\right)^{-1}\Big(\left(tI + \Sigma\Sigma_0^{-1}\right)\mu_0 - t\mu_0 + R_t\Big) \\ &= \left(tI + \Sigma\Sigma_0^{-1}\right)^{-1}\left(\Sigma\Sigma_0^{-1}\mu_0 + R_t\right) \\ &= \left(t\Sigma^{-1} + \Sigma_0^{-1}\right)^{-1}\left(\Sigma_0^{-1}\mu_0 + \Sigma^{-1}R_t\right) \\ \mathrm{Var}\big(\mu|\mathcal{F}_t^S\big) = \gamma_t &= \Sigma_0 - \Sigma_0 t \left(t\Sigma_0 t + \Sigma t\right)^{-1} t\Sigma_0 \\ &= \Sigma_0 - \Sigma_0\left(\Sigma_0 t + \Sigma\right)^{-1} t\Sigma_0 \\ &= \Sigma_0 - \left(tI + \Sigma\Sigma_0^{-1}\right)^{-1} t\Sigma_0 \\ &= \Sigma_0 - \left(t\Sigma^{-1} + \Sigma_0^{-1}\right)^{-1} t\Sigma^{-1}\Sigma_0 \\ &= \left(t\Sigma^{-1} + \Sigma_0^{-1}\right)^{-1}\left(\left(t\Sigma^{-1} + \Sigma_0^{-1}\right)\Sigma_0 - t\Sigma^{-1}\Sigma_0\right) \\ &= \left(t\Sigma^{-1} + \Sigma_0^{-1}\right)^{-1} \end{aligned}$$

□

Remark 2.3 (Interpretation)
Note that we can actually observe the returns $R(t) = \mu t + \sigma W_t$ from the stock prices via

$$R_i(t) = \log\frac{S_i(t)}{S_i(0)} + \frac{1}{2}\Sigma_{ii}t \tag{2.7}$$

where the stock prices $S_i(s)$ are observable for $s \leq t$ and the variance Σ is known to the investor. Therefore the value of the filter is known at any time.

The analysts provide their estimates (μ_0, Σ_0) at $t = 0$ and the investor observes the stock returns $R_t = R(t)$ until time t. We observe for the resulting filter:

$$\hat{\mu}_0 = \mu_0 \quad \text{and} \quad \lim_{t\to\infty}\hat{\mu}_t = \mu \tag{2.8}$$

Hence at time $t = 0$ the best guess for μ is the estimate of the analyst itself, while for $t \to \infty$ the filter converges to the true (and unknown) value of μ. Note that the second equality is rather theoretical as it depends crucially on the unrealistic assumption that μ is time-independent, which is only reasonable for a small investment horizon.

Remark 2.4
Alternatively one could also model the risk premium $\theta := \sigma^{-1}(\mu - r\mathbb{1})$ instead of the growth rate μ as it is done in [CLMZ06]. Then the risk premium gets modelled normally distributed $\theta \sim \mathcal{N}(\theta_0, \Delta)$ with $\theta_0 = \sigma^{-1}(\mu_0 - r\mathbb{1})$ and $\Delta = \sigma^{-1}\Sigma_0(\sigma^{-1})^\top$. The corresponding filter $\hat{\theta}_t := \mathbb{E}\left[\theta | \mathcal{F}_t^S\right]$ hence is given by the following formulas by [LS01, Thm.10.3]:

$$\begin{aligned}
\hat{\theta}_t &= \gamma_t^\theta \left(W_t^Q + \gamma_0^{-1}\theta_0\right) = P^\top \hat{D}_t \left(PW_t^Q + D^{-1}P\theta_0\right) \\
\gamma_t^\theta &= \mathrm{Var}\big(\theta|\mathcal{F}_t^S\big) = P^\top \hat{D}_t P
\end{aligned}$$

Here $W_t^Q = W_t + \theta t = \sigma^{-1}(R_t - rt)$ is a Brownian motion with respect to the risk-neutral measure Q. P is an orthogonal matrix that diagonalizes Δ such that $\Delta = P^\top D P$ with $D = \mathrm{diag}\,(...d_i...)$ and $\hat{D}_t := \mathrm{diag}\,(...\delta_i...)$ is diagonal with $\delta_i := \frac{d_i}{1+d_i t}$.

One can easily show that both approaches are equivalent:

$$\hat{\theta}_t = \sigma^{-1}(\hat{\mu}_t - r\mathbb{1}).$$

2.2.1 From partial to completely observable information

Our market model (2.1) includes the unknown and unobservable parameters μ and W_t.

$$\begin{aligned}
dB_t &= B_t r dt \\
dS_t &= \mathrm{diag}\,(S_t)\, dR_t \\
\text{where} \quad dR_t &= \mu dt + \sigma dW_t \\
&= r dt + \sigma dW_t^Q
\end{aligned}$$

where $W_t^Q = W_t + \theta t = \sigma^{-1}(R_t - rt)$ is a Brownian motion with respect to the risk-neutral measure Q. Therefore it is equivalent whether the investor can observe the stock prices S_t, the returns R_t or the risk-neutral Brownian motion W_t^Q. We observe for the observation filtration $\mathcal{F}^S = \mathcal{F}^R = \mathcal{F}^Q \subset \mathcal{F}$.
We won't use the Brownian motion W_t^Q with respect to the risk-neutral measure Q, but instead we define the innovation process V_t via its dynamics:

$$\begin{aligned}
dV_t &:= dW_t + \sigma^{-1}(\mu - \hat{\mu}_t)\, dt & (2.9) \\
&= \sigma^{-1}(dR_t - \hat{\mu}_t dt) \\
V_t &= \sigma^{-1}\left(R_t - \int_0^t \hat{\mu}_s ds\right) & (2.10)
\end{aligned}$$

From (2.10) we observe that the innovations process is completely observable, since it only consists of σ, R_t and $\hat{\mu}_t$.
Hence we may update our market setting from the original partial information setting to a setting with complete information using the innovation process V_t:

$$\begin{aligned} dS_t &= \operatorname{diag}(S_t)\left(\mu dt + \sigma dW_t\right) \\ &= \operatorname{diag}(S_t)\left(\hat{\mu}_t dt + \sigma dV_t\right) \end{aligned} \tag{2.11}$$

While we cannot observe μ and W_t, all parameters in (2.11) are known at time t.
Additionally we get the following nice results:

Theorem 2.5 (Fujisaki, Kallianpur, Kunita)
The innovation process V_t with $dV_t = dW_t + \sigma^{-1}(\mu - \hat{\mu}_t)\, dt$ is a standard $\mathcal{F}^S$-Brownian motion under the physical measure $\mathbb{P}$.

Proof. The proof is done via Lévy's characterization of a Brownian motion.
Obviously $V_0 = 0$ since W_t is a Brownian motion.
First we show that V_t is a martingale with respect to $\mathcal{F}^R = \mathcal{F}^S$. Let $s \leq t$.

$$\begin{aligned} \mathbb{E}\left[R_t \middle| \mathcal{F}_s^R\right] &= R_s + \mathbb{E}\left[R_t - R_s\right] \\ &= R_s + \mathbb{E}\left[\mu(t-s) + \sigma(W_t - W_s)\right] \\ &= R_s + \mathbb{E}\left[\mu\right](t-s) \\ \mathbb{E}\left[\int_0^t \hat{\mu}_u du \middle| \mathcal{F}_s^R\right] &= \int_0^s \hat{\mu}_u du + \mathbb{E}\left[\int_s^t \hat{\mu}_u du\right] \\ &= \int_0^s \hat{\mu}_u du + \int_s^t \mathbb{E}\left[\mathbb{E}\left[\mu \middle| \mathcal{F}_u^R\right]\right] du \\ &= \int_0^s \hat{\mu}_u du + \mathbb{E}\left[\mu\right](t-s) \end{aligned}$$

hence

$$\begin{aligned} \mathbb{E}\left[V_t \middle| \mathcal{F}_s^R\right] &= \sigma^{-1}\left(\mathbb{E}\left[R_t \middle| \mathcal{F}_s^R\right] - \mathbb{E}\left[\int_0^t \hat{\mu}_u du \middle| \mathcal{F}_s^R\right]\right) \\ &= \sigma^{-1}\left(R_s - \int_0^s \hat{\mu}_u du\right) \\ &= V_s \end{aligned}$$

Secondly we consider the quadratic covariation of V_t where we use formula (2.9).

$$V_t = W_t + \int_0^t \sigma^{-1}\left(\mu - \mathbb{E}\left[\mu \middle| \mathcal{F}_s^R\right]\right) ds$$

The quadratic covariation of a standard Lebegues integral $\int ...dt$ is zero and the quadratic covariation of a multidimensional Brownian motion is $t \cdot Id$, hence

$$\langle V \rangle_t = \langle W \rangle_t + 0 = t \cdot Id.$$

□

Lemma 2.6
The filter $\hat{\mu}_t$ is a martingale with respect to the filtration $\mathcal{F}^S$ under the innovations process V_t and the physical measure $\mathbb{P}$. Its dynamics are explicitly given with respect to the innovations process V_t by

$$d\hat{\mu}_t = \gamma_t(\sigma^{-1})^\top dV_t. \tag{2.12}$$

Proof. We apply the Itô-formula to the d+1-dimensional function $f(t, w)$

$$f(t,w) = \gamma_t \left(\Sigma_0^{-1}\mu_0 + \Sigma^{-1}\mu t + \Sigma^{-1}\sigma w\right)$$

with derivatives (note that $\partial_t\gamma_t = -\gamma_t\Sigma^{-1}\gamma_t$)

$$\begin{aligned}
\partial_t f(t,w) &= \gamma_t\Sigma^{-1}\mu - \gamma_t\Sigma^{-1}\gamma_t\left(\Sigma_0^{-1}\mu_0 + \Sigma^{-1}\mu t + \Sigma^{-1}\sigma w\right)\\
&= \gamma_t\Sigma^{-1}\mu - \gamma_t\Sigma^{-1}f(t,w)\\
\partial_w f(t,w) &= \gamma_t\Sigma^{-1}\sigma\\
\partial^2_{ww} f(t,w) &= 0
\end{aligned}$$

Hence we observe for the filter $\hat{\mu}_t = f(t, W_t)$

$$\begin{aligned}
d\hat{\mu}_t = df(t,W_t) &= \partial_t f(t,W_t)dt + \partial_w f(t,W_t)dW_t + 0\\
&= \gamma_t\Sigma^{-1}\big((\mu - f(t,W_t))dt + \sigma dW_t\big)\\
&= \gamma_t\Sigma^{-1}\big((\mu - \hat{\mu}_t)dt + \sigma dW_t\big)\\
&= \gamma_t(\sigma^{-1})^\top dV_t
\end{aligned}$$

Therefore $\hat{\mu}_t$ is a local martingale with

$$\begin{aligned}
\sup_{t\in[0,T]} \mathbb{E}\left[\langle\hat{\mu}\rangle_t\right] &= \sup_{t\in[0,T]} \mathbb{E}\left[\int_0^t \mathrm{tr}(\gamma_s\Sigma^{-1}\gamma_s)ds\right]\\
&= \sup_{t\in[0,T]} \int_0^t \partial_s\mathrm{tr}(\gamma_s)ds\\
&= \sup_{t\in[0,T]} \mathrm{tr}(\gamma_t) - \mathrm{tr}(\Sigma_0) < \infty
\end{aligned}$$

hence $\hat{\mu}_t$ is an L^2-martingale. □

Remark 2.7
For the filter for the risk premium θ we observe the same results with dynamics

$$d\hat{\theta}_t = \gamma^\theta_t dV_t.$$

2.3 Constrained Optimization

Our portfolio optimization setting relies in particular on the ability to include convex constraints on the admissible portfolio strategies. This section will give a short introduction to the theoretical background of handling convex constraints as it was introduced by [CK92] including their main results needed for this thesis.

In our market setting we consider the constrained portfolio optimization problem of [CK92]. We want to maximize expected utility of terminal wealth (2.3) where the portfolio strategy $(\pi_t)_{t\in[0,T]} \in \mathcal{A}$ additionally has to remain in some closed convex set $K \subseteq \mathbb{R}^d$ for all $t \in [0,T]$. This problem is dealt with in the following.

We start by defining the support function of $-K$ as it is known from convex analysis,

$$\delta(x) := \delta(x|K) := \sup_{\pi\in K}(-\pi^\top x). \tag{2.13}$$

The support function δ is a convex function on its effective domain $\tilde{K}$,

$$\tilde{K} := \left\{ x \in \mathbb{R}^d \middle| \, \delta(x|K) < \infty \right\}. \tag{2.14}$$

We observe immediately

$$\delta(x) + \pi^\top x \geq 0 \qquad \text{for all } \pi \in K \text{ and } x \in \tilde{K}. \tag{2.15}$$

Now we introduce the auxiliary markets $\mathcal{M}^\nu$ of [CK92].

$$\begin{aligned} dB_t^\nu &= B_t^\nu\big(r + \delta\left(\nu_t\right)\big)dt \\ dS_t^\nu &= \operatorname{diag}\left(S_t^\nu\right)\big(\left(\mu + \nu_t + \delta(\nu_t)\mathbb{1}\right)dt + \sigma dW_t\big) \\ &= \operatorname{diag}\left(S_t^\nu\right)\big(\left(\hat{\mu}_t + \nu_t + \delta(\nu_t)\mathbb{1}\right)dt + \sigma dV_t\big) \end{aligned} \tag{2.16}$$

The process $\nu \in \mathcal{D}$ is defined to be any square-integrable and $\mathcal{F}_t^S$-progressively measurable processes with $\mathbb{E}\left[\int_0^T \delta\left(\nu_t\right)dt\right] < \infty$. Hence for $\nu \in \mathcal{D}$ we observe in particular $\nu_t \in \tilde{K}$ for all $t \in [0,T]$.

For $\nu = 0$ we observe our original market. Hence these auxiliary markets are considered to be 'versions' of the original market, where the market parameters μ and r get updated by functionals of ν_t. We will see below that these updates always result in a 'better' market with respect to maximizing utility of wealth.

In each of these auxiliary markets $\mathcal{M}^\nu$ we define the processes (2.17):

the risk-premium: $\quad \hat{\theta}_t^\nu := \sigma^{-1}\left(\hat{\mu}_t + \nu_t - r\mathbb{1}\right) = \hat{\theta}_t + \sigma^{-1}\nu_t \quad (2.17)$

the deflator: $\quad \beta_t^\nu := \exp\left(-\int_0^t r + \delta\left(\nu_s\right)ds\right)$

the 'change of measure'-process: $Z_t^\nu := \exp\left(-\int_0^t \hat{\theta}_s^{\nu\top} dV_s - \frac{1}{2}\int_0^t \hat{\theta}_s^{\nu\top}\hat{\theta}_s^\nu ds\right)$

$dZ_t^\nu = -Z_t^\nu \hat{\theta}_s^{\nu\top} dV_t$

the 'state of the world'-process: $H_t^\nu := \beta_t^\nu Z_t^\nu$

the risk-neutral measure: $\mathbb{Q}^\nu$ defined via $\left.\frac{d\mathbb{Q}^\nu}{d\mathbb{P}}\right|_{\mathcal{F}_t^S} = Z_t^\nu$

the $\mathbb{Q}^\nu$-Brownian motion: $V_t^{\mathbb{Q},\nu} := V_t + \int_0^t \hat{\theta}_s^\nu ds$

Note that the $\mathbb{Q}^\nu$-Brownian motion $V_t^{\mathbb{Q},\nu}$ that is defined via the innovations process equals the $\mathbb{Q}^\nu$-Brownian motion $W_t^{\mathbb{Q},\nu}$ that gets defined via the original Brownian motion:

$$\begin{aligned} W_t^{\mathbb{Q},\nu} &:= W_t + \int_0^t \theta_s^\nu ds = W_t + \int_0^t \theta_s + \sigma^{-1}\nu_s ds \\ &= V_t + \int_0^t \hat{\theta}_s + \sigma^{-1}\nu_s ds = V_t + \int_0^t \hat{\theta}_s^\nu ds = V_t^{\mathbb{Q},\nu} \end{aligned}$$

The auxiliary market with respect to the risk-neutral measure $\mathbb{Q}^\nu$ reads as follows:

$$\begin{aligned} dB_t^\nu &= B_t^\nu \big(r + \delta(\nu_t)\big) dt \\ dS_t^\nu &= \operatorname{diag}(S_t^\nu)\,\big(\,(r + \delta(\nu_t))\,\mathbb{1}dt + \sigma dV_t^{\mathbb{Q},\nu}\big) \end{aligned}$$

For any self-financing portfolio process with $\pi_t \in K$ for all $t \in [0,T]$, we observe the dynamics of the wealth process X_t^ν in the auxiliary market $\mathcal{M}^\nu$.

$$\begin{aligned} dX_t^\nu &= X_t^\nu \left(\left(r + \delta(\nu_t) + \pi_t^\top (\hat{\mu}_t + \nu_t - r\mathbb{1})\right) dt + \pi_t^\top \sigma dV_t\right) \\ &= X_t^\nu \left(\big(r + \delta(\nu_t)\big) dt + \pi_t^\top \sigma dV_t^{\mathbb{Q},\nu}\right) \\ &= X_t^\nu \left(\left(r + \delta(\nu_t) + \pi_t^\top \nu_t\right) dt + \pi_t^\top \sigma dV_t^{\mathbb{Q},0}\right) \end{aligned}$$

In each of these auxiliary markets we can solve the standard (unconstrained) portfolio optimization problem of maximizing utility of terminal wealth. The resulting optimal portfolio strategy will be denoted by π^ν.

Remark 2.8
We have seen in equation (2.15) that $\delta(\nu_t) + \pi_t^\top \nu_t \geq 0$ for all t. However, since the wealth process in the original market X_t equals the wealth process in the auxiliary market $\mathcal{M}^0$, we observe:

$$\frac{dX_t^\nu}{X_t^\nu} = \left(r + \delta(\nu_t) + \pi_t^\top \nu_t\right) dt + \pi_t^\top \sigma dV_t^{\mathbb{Q},0} \geq r dt + \pi_t^\top \sigma dV_t^{\mathbb{Q},0} = \frac{dX_t^0}{X_t^0} = \frac{dX_t}{X_t}$$

Hence the wealth process for an admissible portfolio strategy π in any of the auxiliary markets is at least as good (large) as the wealth process for this strategy in the original market. Hence if we find a process $\nu \in \mathcal{D}$ such that the optimal portfolio strategy π^ν in the auxiliary market $\mathcal{M}^\nu$ satisfies $\delta(\nu_t) + \nu_t^\top \pi_t^\lambda = 0$ then we also found an optimal portfolio strategy for our original market. This is the conclusion of the following theorem. The optimal choice for the process ν will be called λ.

Theorem 2.9 ([CK92] and [Sas07] for partial information)
Suppose there is a process $\lambda \in \mathcal{D}$, such that the optimal unconstrained portfolio strategy π^λ of the auxiliary market $\mathcal{M}^\lambda$ satisfies for all $t \in [0, T]$

$$\pi_t^\lambda \in K \qquad \text{and} \qquad \delta(\lambda_t) + \lambda_t^\top \pi_t^\lambda = 0$$

Then π_t^λ is the optimal portfolio strategy for the original constrained portfolio optimization problem (2.3).

Proof. The proof is given in detail in [CK92, Prop.8.3], respectively in [Sas07] and is based upon the idea described above in Remark 2.8:

$$X_t^\nu = X_t^0 \quad \Leftrightarrow \quad dX_t^\nu = X_t^\nu \left(rdt + \pi_t^\top \sigma dV_t^{\mathbb{Q},0} \right) \quad \Leftrightarrow \quad \delta\left(\nu_t\right) + \nu_t^\top \pi_t^\nu = 0$$

□

Remark 2.10 (On the support function)
Although Theorem 2.9 provides an explicit equation to solve for λ, in most cases we cannot find λ explicitly this way.
In case δ is differentiable at λ we observe from $\delta(\lambda_t) + \lambda_t^\top \pi_t^\lambda = 0$:

$$\pi_t^\lambda = -\delta'(\lambda_t)$$

Unfortunately we will see in the following examples that if δ is differentiable at some point, then the derivative is often 0 or $-e_i$. The resulting optimal strategies then become $\pi^* = 0$ or $\pi^* = e_i$, hence not investing at all or investing everything in one stock. Both results don't seem to be reasonable results in general, so we have to consider that the optimal λ can be on non-differentiability points of the support function δ.

However, additionally to Theorem 2.9 we have the following duality theorem:

Theorem 2.11 (Dual optimality, [CK92])
If there exists an optimal portfolio strategy π^* for the constrained portfolio optimization problem (2.3), then there exists an optimal dual process $\lambda \in \mathcal{D}$ that satisfies Theorem 2.9 with $\pi^* = \pi^\lambda$.
Vice versa, if there exists an optimal dual process $\lambda \in \mathcal{D}$, satisfying Theorem 2.9, then $\pi^\lambda \in K$ is the optimal portfolio strategy for the original constrained optimization problem. In both cases the process λ solves the following dual problem:

$$\mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^\lambda(x_0) H_T^\lambda\big)\right] \leq \mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^\lambda(x_0) H_T^\nu\big)\right] \qquad \text{for all } \nu \in \mathcal{D} \tag{2.18}$$

Here $\tilde{U}(y) = U(I(y)) - yI(y)$ with $I(y) = U'^{-1}(y)$ is the convex dual function of U and $\mathcal{Y}^\nu(x) = (\mathcal{X}^\nu)^{-1}(x)$ with $\mathcal{X}^\nu(y) = \mathbb{E}\left[H_T^\nu I\big(yH_T^\nu\big)\right]$.

Proof. The proof is given in Appendix A.2.1, conditions (A) and (D). □

Example 2.12 (choices of K)
First we state some examples for possible sets of constraints K and their corresponding support functions.

1. $K = \mathbb{R}^d$ i.e. the unconstrained case

$$\delta(x) = \begin{cases} 0, & \text{if } x = 0 \\ \infty, & \text{else} \end{cases} \qquad \text{i.e. } \tilde{K} = \{0\}$$

2. $K = [0, \infty)^d$ i.e. no short-selling

$$\delta(x) = \begin{cases} 0, & \text{if } x \in \tilde{K} = [0, \infty)^d \\ \infty, & \text{if } x \notin \tilde{K} \end{cases}$$

   In particular $\delta(x)$ is differentiable on the interior of $\tilde{K}$ with $\partial_x \delta(x) = 0$.

3. $K = \left\{\pi \in \mathbb{R}^d \middle| \pi^\top \mathbb{1} \le 1\right\}$ i.e. no borrowing

$$\delta(x) = \begin{cases} -x_1, & \text{if } x \in \tilde{K} = \{x \in \mathbb{R}^d | x_i = x_1 \le 0 \text{ for all } i\} \\ \infty, & \text{if } x \notin \tilde{K} \end{cases}$$

4. $K = \left\{\pi \in \mathbb{R}^d \middle| \, \|\pi\|_2 \le 1\right\}$ i.e. bounded L^2-norm

$$\delta(x) = \|x\|_2 \qquad \text{for all } x \in \tilde{K} = \mathbb{R}^d$$

   In particular $\delta(x)$ is differentiable on $\mathbb{R}^d$ with $\partial_x \delta(x) = \frac{x}{\|x\|_2}$.

5. $K = \left\{x \in [0, \infty)^d \middle| x^\top \mathbb{1} \le 1\right\}$ i.e. no short-selling and no borrowing

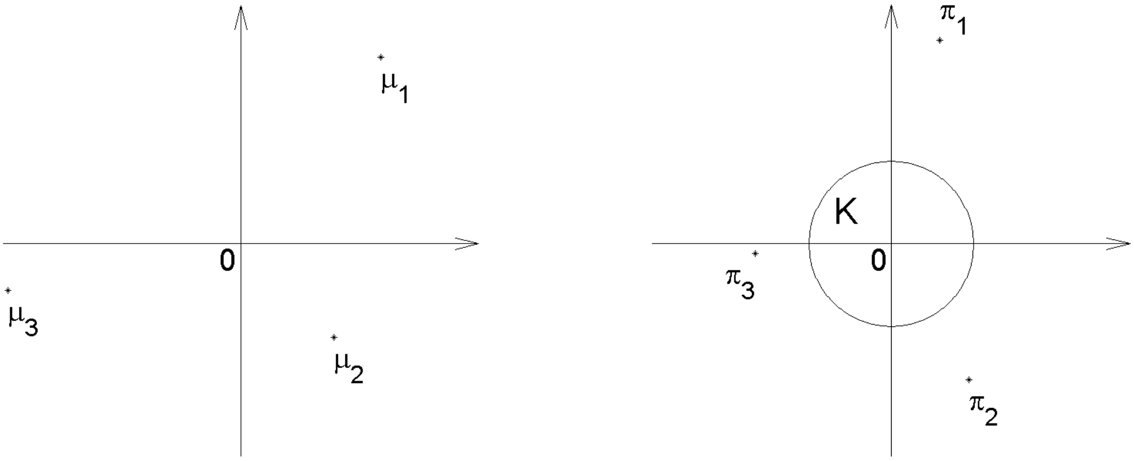
$$\delta(x) = \max_j \, (x_j)^- = \begin{cases} 0, & \text{if } x \in [0, \infty)^d \\ -x_i, & \text{if } x_i = \min\limits_j x_j < 0 \end{cases} \qquad \text{i.e. } \tilde{K} = \mathbb{R}^d$$

   In particular $\delta(x)$ is differentiable on the relative interiors of each part of $\tilde{K}$ with

$$\partial_x \delta(x) = \begin{cases} 0, & \text{if } x \in (0, \infty)^d \\ -e_i, & \text{if } x_i = \min\limits_j x_j < 0 \text{ and } x_i < x_j \text{ for all } j \end{cases}$$

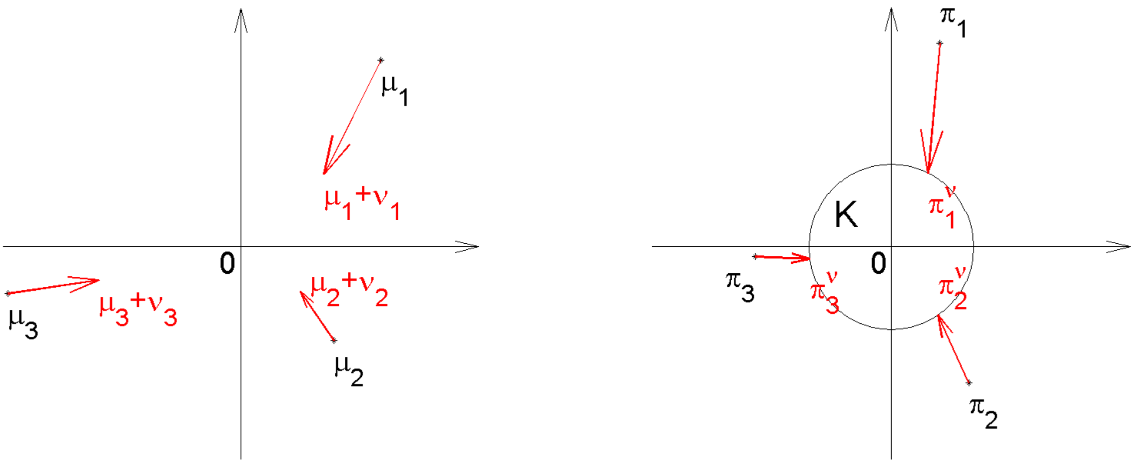
In Section 5.1 we present further constraints and their corresponding support functions.

Example 2.13 (the idea of [CK92])
In order to understand the idea behind the auxiliary markets approach of [CK92] we consider the following easy example in 2 dimensions, where the set of admissible strategies K is given by a bounded L^2-norm: $K = \left\{\pi \in \mathbb{R}^d \middle| \, \|\pi\|_2 \leq 1\right\}$.
Let the market parameters be given by $r = 0$ and $\Sigma = \begin{pmatrix} 0.3 & 0.05 \\ 0.05 & 0.15 \end{pmatrix}$.
Then under logarithmic utility we consider three different examples μ_1, μ_2, μ_3 for the growth rate μ with their corresponding optimal Merton strategies $\pi_i = \Sigma^{-1}\mu_i$.



$$\mu_1 = \begin{pmatrix} 0.3 \\ 0.4 \end{pmatrix}, \ \mu_2 = \begin{pmatrix} 0.2 \\ -0.2 \end{pmatrix}, \ \mu_3 = \begin{pmatrix} -0.5 \\ -0.1 \end{pmatrix}; \quad \pi_1 = \begin{pmatrix} 0.59 \\ 2.47 \end{pmatrix}, \ \pi_2 = \begin{pmatrix} 0.94 \\ -1.65 \end{pmatrix}, \ \pi_3 = \begin{pmatrix} -1.65 \\ -0.12 \end{pmatrix}$$

All three optimal unconstrained strategies π_i are not admissible and we need to introduce auxiliary markets $\mathcal{M}^\nu$ as described above. In Proposition 5.5 below we will describe how to get the optimal choices for the dual processes ν in this setting. The optimal strategies in the auxiliary markets are given by $\pi^\nu = \Sigma^{-1}(\mu + \nu + \delta(\nu)\mathbb{1} - r\mathbb{1} - \delta(\nu)\mathbb{1}) = \Sigma^{-1}(\mu + \nu)$ and hence the optimal constrained strategies in the original market are also given by $\pi_i^\nu = \Sigma^{-1}(\mu_i + \nu_i)$.



We observe that the additional term ν shifts the original μ non-linearly closer to zero such that the strategy $\pi^\nu = \Sigma^{-1}(\mu + \nu)$ just touches the border of the admissibility region K. The direction of the arrows drawn above goes approximately to zero, slightly rotated by the correlation in Σ.

3 Constrained optimization under logarithmic utility

In this chapter we consider the constrained portfolio optimization problem (2.3) under partial information as introduced in Section 2.2 for the case of logarithmic utility. It is well known that the logarithmic utility function has favorable properties like the easy derivation of optimal results, but unfavorable properties like the low risk-aversion. However, due to the easy handling we are able to derive the optimal portfolio strategies in our market setting without too much effort. In Chapter 4 we will also consider power utility functions to be able to vary the investor's risk-aversion.

We restrict the set of admissible strategies $\pi \in \mathcal{A}$ to strategies with $\pi_t \in K$ for all $t \in [0,T]$ for some closed convex set K. In equation (3.7) we specify the constraints K further to forbid short-selling and borrowing.

According to Section 2.3 and [CK92], in order to solve the constrained portfolio optimization problem, we can solve the unconstrained portfolio optimization problem in the auxiliary market $\mathcal{M}^\lambda$ where λ is chosen according to Theorem 2.11:

$$\begin{aligned} dB_t^\lambda &= B_t^\lambda\big(r + \delta\left(\lambda_t\right)\big)dt \qquad (3.1)\\ dS_t^\lambda &= \operatorname{diag}\left(S_t^\lambda\right)\big(\left(\hat{\mu}_t + \lambda_t + \delta(\lambda_t)\mathbb{1}\right)dt + \sigma dV_t\big) \\ \text{where} \quad \lambda &= \underset{\nu\in\mathcal{D}}{\arg\min}\ \mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^\lambda(x_0)H_T^\nu\big)\right] \qquad (3.2)\end{aligned}$$

It is well known that the auxiliary market under logarithmic utility can be solved quite easily. [BUV12] and [CK92] showed that the optimal portfolio strategy in our setting is just given by the Merton Plug-In strategy.

$$\pi_t^\lambda = \Sigma^{-1}\left(\hat{\mu}_t + \lambda_t - r\mathbb{1}\right) \qquad (3.3)$$

Hence the only problem left to solve is the dual optimization problem (3.2). For logarithmic utility we get:

$$\begin{aligned} U(x) &= \log(x), \\ I(y) &= (U')^{-1}(y) = 1/y, \\ \tilde{U}(y) &= U(I(y)) - yI(y) = \log(1/y) - 1, \\ \mathcal{X}^\lambda(y) &= \mathbb{E}\left[H_T^\lambda I\big(yH_T^\lambda\big)\right] = 1/y, \\ \mathcal{Y}^\lambda(x) &= \left(\mathcal{X}^\lambda\right)^{-1}(x) = 1/x. \end{aligned}$$

and H_t^ν is given in (2.17). Hence

$$\begin{aligned}
\lambda &= \underset{\nu\in\mathcal{D}}{\arg\min}\ \mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^\lambda(x_0)H_T^\nu\big)\right] \\
&= \underset{\nu\in\mathcal{D}}{\arg\min}\ \mathbb{E}\left[\log\left(\frac{x_0}{H_T^\nu}\right)-1\right] \\
&= \underset{\nu\in\mathcal{D}}{\arg\min}\ \mathbb{E}\left[\log\Big((H_T^\nu)^{-1}\Big)\right] \\
&= \underset{\nu\in\mathcal{D}}{\arg\min}\ \mathbb{E}\left[\int_0^T r+\delta(\nu_t)+\frac{1}{2}\left\|\hat{\theta}_t^\nu\right\|_2^2 dt+\int_0^T \hat{\theta}_t^{\nu\top}dV_t\right] \\
&= \underset{\nu\in\mathcal{D}}{\arg\min}\ \mathbb{E}\left[\int_0^T \delta(\nu_t)+\frac{1}{2}\left\|\hat{\theta}_t^\nu\right\|_2^2 dt\right]. \qquad (3.4)
\end{aligned}$$

Here the last equality assumes $\mathbb{E}\left[\int_0^T \hat{\theta}_t^{\nu\top}\hat{\theta}_t^\nu dt\right]<\infty$ which follows from $\hat{\mu}_t$ and $\nu_t\in\mathcal{D}$ being squared integrable. Point-wise minimization now leads to

$$\lambda_t = \underset{\nu\in\tilde{K}}{\arg\min}\ \left(\delta(\nu)+\frac{1}{2}(\hat{\mu}_t+\nu-r\mathbb{1})^\top\Sigma^{-1}(\hat{\mu}_t+\nu-r\mathbb{1})\right) \quad \text{for all } t \qquad (3.5)$$

Since [CK92, Remark 11.1] showed that the resulting dual process λ defined by (3.5) is already $\mathcal{F}^S$-progressively measurable, this indeed solves the dual problem (3.4). Hence the dual optimization problem (3.2) results in the point-wise minimization of the convex function $\delta(\nu_t)+\frac{1}{2}\left\|\hat{\theta}_t^\nu\right\|_2^2$ over all admissible values $\nu_t\in\tilde{K}$.

In the following we fix some $t\in[0,T]$ and abbreviate $\mu=\hat{\mu}_t$ and $\lambda=\lambda_t$ such that we have to solve

$$\lambda = \underset{\nu\in\tilde{K}}{\arg\min}\ \left(\delta(\nu)+\frac{1}{2}(\mu+\nu-r\mathbb{1})^\top\Sigma^{-1}(\mu+\nu-r\mathbb{1})\right). \qquad (3.6)$$

Remark 3.1
The point-wise minimization in equation (3.5) could also be derived by applying a stochastic control approach to (3.4). The controlled process would be defined as follows.

$$\begin{aligned}
X_t^\nu &:= \int_0^t \delta(\nu_s)+\frac{1}{2}\hat{\theta}_s^{\nu\top}\hat{\theta}_s^\nu ds+\int_0^t \hat{\theta}_s^{\nu\top}dV_s \\
\text{i.e.}\quad dX_t^\nu &= \left(\delta(\nu_t)+\frac{1}{2}\hat{\theta}_t^{\nu\top}\hat{\theta}_t^\nu\right)dt+\hat{\theta}_t^{\nu\top}dV_t \\
\text{such that}\quad \lambda &= \underset{\nu\in\tilde{K}}{\arg\min}\ \mathbb{E}\left[X_T^\nu\right].
\end{aligned}$$

Remark 3.2
If we assume δ to be differentiable at λ we can derive the solution of (3.6) by differentiating the argument of the 'argmin', leading to $0=\delta'(\lambda)+\Sigma^{-1}(\mu+\lambda-r\mathbb{1})$. Hence

$$\lambda = -\Sigma\delta'(\lambda)-(\mu-r\mathbb{1})$$

$$\text{and} \quad \pi^\lambda = \Sigma^{-1}(\mu + \lambda - r\mathbb{1}) = -\delta'(\lambda)$$

We have already seen this result in Remark 2.10 and motivated why the assumption of differentiability is unrealistic.

From a theoretical point of view the minimization of a convex, quadratic function like in problem (3.6) is rather trivial and can be solved numerically. Unfortunately we have already motivated in Section 1.3 why numerical solutions are not satisfactory and for a large number of dimensions this won't work properly enough anyway as the following lemma indicates:

Lemma 3.3 (accuracy of λ)
Let's consider the p-norm to be the underlying metric.
The necessary condition: To gain an accuracy of ε for π^λ, λ needs an accuracy of $\varepsilon\|\Sigma\|_p$.
The sufficient condition: If λ has an accuracy of $\varepsilon/\|\Sigma^{-1}\|_p$, then π^λ has an accuary of ε.

Proof. Consider an arbitrary solution: $\pi^{\lambda'} := \Sigma^{-1}(\mu - r + \lambda')$. Then $\pi^\lambda - \pi^{\lambda'} = \Sigma^{-1}(\lambda - \lambda')$.
The necessary condition:
If $\|\pi^\lambda - \pi^{\lambda'}\|_p < \varepsilon$ then $\|\lambda - \lambda'\|_p \leq \|\Sigma\|_p \|\pi^\lambda - \pi^{\lambda'}\|_p < \varepsilon\|\Sigma\|_p$.
The sufficient condition:
If $\|\lambda - \lambda'\|_p < \varepsilon/\|\Sigma^{-1}\|_p$ then $\|\pi^\lambda - \pi^{\lambda'}\|_p \leq \|\Sigma^{-1}\|_p \|\lambda - \lambda'\|_p < \varepsilon$. □

Remark 3.4
If we are looking for an accuracy per digit we need to use $p = \infty$ in Lemma 3.3. The induced matrix-norm is the row-sum-norm.
In the DAX-30-Simulation (in [Von14]) we observe $\|\Sigma\|_\infty \approx 0.6$ and $\|\Sigma^{-1}\|_\infty \approx 1000$. Hence in order to gain an accuracy of 0.1% for π^λ, we need $\varepsilon = 0.05\% = 5 \cdot 10^{-4}$ and therefore a λ-accuracy in $[5 \cdot 10^{-7}, 3 \cdot 10^{-4}]$ in every dimension.

3.1 The convex dual problem

Following the previous lemma and remarks it seems necessary to solve the dual problem (3.6) as explicitly as possible. Therefore we fix some time $t \in [0,T]$ and the set of convex constraints K to **no short-selling and no borrowing**:

$$K = \left\{x \in [0,\infty)^d \middle| x^\top \mathbb{1} \leq 1\right\} \tag{3.7}$$

$$\text{with} \quad \delta(x) = \max_j \,(x_j)^- = \begin{cases} 0, & \text{if } x \in [0,\infty)^d \\ -x_i, & \text{if } x_i = \min_j x_j < 0 \end{cases} \quad , \quad \tilde{K} = \mathbb{R}^d$$

Additionally we state the following 'matrix inversion formulas' that we will use frequently in the following sections:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix} \tag{3.8}$$

Note in particular the matrix notations described in Appendix A.5.

We start with some minor lemmas that lead to the solution of the dual problem (3.6):

$$\lambda = \arg\min_{\nu \in \tilde{K}} \left(\delta(\nu) + \frac{1}{2}(\mu + \nu - r\mathbb{1})^\top \Sigma^{-1}(\mu + \nu - r\mathbb{1})\right)$$

Lemma 3.5 (reduction of dimension)
Let the optimal portfolio strategy π^λ (3.3) of the d-dimensional portfolio optimization problem fulfil $\pi_i^\lambda = 0$. Then the optimal solution $\bar{\pi}^{\bar{\lambda}}$ of the d-1-dimensional problem generated by the dimensions $I = \{1,...,d\} \setminus \{i\}$ equals $(\pi^\lambda)_I$. Also $\bar{\lambda} = \lambda_I$.

Proof. To simplify notation we assume without loss of generality $r = 0$.
The original portfolio optimization problem in d dimensions leads to the optimal solutions:

$$\pi^\lambda = \Sigma^{-1}(\mu + \lambda) \qquad \text{where} \quad \lambda = \arg\min_{\nu \in \mathbb{R}^d} \left(\delta(\nu) + \frac{1}{2}(\mu + \nu)^\top \Sigma^{-1}(\mu + \nu)\right)$$

The reduced portfolio optimization problem in d-1 dimensions has the following solutions:

$$\bar{\pi}^{\bar{\lambda}} = \bar{\Sigma}^{-1}(\bar{\mu} + \bar{\lambda}) \qquad \text{where} \quad \bar{\lambda} = \arg\min_{\bar{\nu} \in \mathbb{R}^{d-1}} \left(\delta(\bar{\nu}) + \frac{1}{2}(\bar{\mu} + \bar{\nu})^\top \bar{\Sigma}^{-1}(\bar{\mu} + \bar{\nu})\right)$$

where $\bar{\Sigma} = \Sigma_{I,I}$ and $\bar{\mu} = \mu_I$. Obviously if $\bar{\lambda} = \lambda_I$, then

$$\Sigma_{I,I}\bar{\pi}^{\bar{\lambda}} = \bar{\Sigma}\bar{\pi}^{\bar{\lambda}} = \bar{\mu} + \bar{\lambda} = \mu_I + \lambda_I = \Sigma_{I,\cdot}\pi^\lambda = \Sigma_{I,I}(\pi^\lambda)_I + \Sigma_{I,i}(\pi^\lambda)_i = \Sigma_{I,I}(\pi^\lambda)_I$$

and since $\Sigma_{I,I}$ is a covariance matrix and hence invertible, the proof is completed.

Now we show that $\bar{\lambda} = \lambda_I$. Therefore note

$$
\begin{aligned}
0 = \pi_i^\lambda &= (\Sigma^{-1})_{i,\cdot}(\mu + \lambda) && (3.9)\\
\Rightarrow \quad (\Sigma^{-1})_{i,i}(\mu_i + \lambda_i) &= -(\Sigma^{-1})_{i,I}(\mu_I + \lambda_I)\\
\Rightarrow \quad (\mu_i + \lambda_i) &= -((\Sigma^{-1})_{i,i})^{-1}(\Sigma^{-1})_{i,I}(\mu_I + \lambda_I)\\
&= \Sigma_{i,I}(\Sigma_{I,I})^{-1}(\mu_I + \lambda_I) && (3.10)
\end{aligned}
$$

and from the matrix inversion formula (3.8):

$$
(\Sigma^{-1})_{I,I} = (\Sigma_{I,I})^{-1} + (\Sigma_{I,I})^{-1}\Sigma_{I,i}\left(\Sigma_{i,i} - \Sigma_{i,I}(\Sigma_{I,I})^{-1}\Sigma_{I,i}\right)^{-1}\Sigma_{i,I}(\Sigma_{I,I})^{-1}
$$

and hence $\nu = \lambda$ minimizes the following term:

$$
\begin{aligned}
&\delta(\nu) + \frac{1}{2}(\mu + \nu)^\top \Sigma^{-1}(\mu + \nu)\\
&= \max_j\{0, -\nu_j\} + \frac{1}{2}(\mu_I + \nu_I)^\top(\Sigma^{-1})_{I,\cdot}(\mu + \nu) + \frac{1}{2}(\mu_i + \nu_i)^\top(\Sigma^{-1})_{i,\cdot}(\mu + \nu)\\
&\overset{(3.9)}{=} \max_j\{0, -\nu_j\} + \frac{1}{2}(\mu_I + \nu_I)^\top(\Sigma^{-1})_{I,\cdot}(\mu + \nu)\\
&\overset{(3.9)}{=} \max_j\{0, -\nu_j\} + \frac{1}{2}(\mu_I + \nu_I)^\top(\Sigma^{-1})_{I,\cdot}(\mu + \nu) - \frac{1}{2}(\mu_i + \nu_i)^\top(\Sigma^{-1})_{i,\cdot}(\mu + \nu)\\
&= \max_j\{0, -\nu_j\} + \frac{1}{2}(\mu_I + \nu_I)^\top(\Sigma^{-1})_{I,I}(\mu_I + \nu_I) - \frac{1}{2}(\mu_i + \nu_i)^\top(\Sigma^{-1})_{i,i}(\mu_i + \nu_i)\\
&\overset{(3.8)}{=} \max_j\{0, -\nu_j\} + \frac{1}{2}(\mu_I + \nu_I)^\top(\Sigma_{I,I})^{-1}(\mu_I + \nu_I)\\
&\quad + \frac{1}{2}(\mu_I + \nu_I)^\top(\Sigma_{I,I})^{-1}\Sigma_{I,i}\left(\Sigma_{i,i} - \Sigma_{i,I}(\Sigma_{I,I})^{-1}\Sigma_{I,i}\right)^{-1}\Sigma_{i,I}(\Sigma_{I,I})^{-1}(\mu_I + \nu_I)\\
&\quad - \frac{1}{2}(\mu_i + \nu_i)^\top\left(\Sigma_{i,i} - \Sigma_{i,I}(\Sigma_{I,I})^{-1}\Sigma_{I,i}\right)^{-1}(\mu_i + \nu_i)\\
&\overset{(3.10)}{=} \max_j\{0, -\nu_j\} + \frac{1}{2}(\mu_I + \nu_I)^\top(\Sigma_{I,I})^{-1}(\mu_I + \nu_I) && (3.11)
\end{aligned}
$$

Now let's assume for the optimal dual process: $\lambda_i < \min_{j\in I}\{0, \lambda_j\}$.
But then δ is differentiable in λ with $\delta'(\lambda) = -e_i$. Hence $\pi^\lambda = e_i$ with $\pi_i^\lambda = 1 \neq 0$.
This is a contradiction to the assumptions of the lemma. Therefore

$$
\begin{aligned}
\lambda_i &\geq \min_{j\in I}\{0, \lambda_j\}\\
\Rightarrow \quad -\lambda_i &\leq \max_{j\in I}\{0, -\lambda_j\}\\
\Rightarrow \quad \max_j\{0, -\nu_j\} &= \max_{j\in I}\{0, -\nu_j\} \quad \text{for the optimal } \nu = \lambda
\end{aligned}
$$

But then (3.11) is just the equation that we have to solve to determine $\bar{\lambda}$ and it gets minimized by λ_I. □

Hence under logarithmic utility those stocks we don't invest in don't affect the optimal strategy. It is rather undesirable that not even their correlation gets taken into account for the optimal solution. In the case of power utility in Chapter 4 this results does not hold true any more.

Now let's start to determine the optimal dual solution.

Definition 3.6 (active and passive)
Let π^λ be the optimal constrained portfolio strategy with optimal dual process λ.
Then we define the 'active dimensions' I and the 'passive dimensions' J:

$$I := \{i \in \{1, ..., d\} \big| \pi_i^\lambda > 0\}$$
$$J := \{j \in \{1, ..., d\} \big| \pi_j^\lambda = 0\}$$

Sometimes we will also refer to I and J as the 'active stocks' and 'passive stocks', identifying a dimension with the corresponding stock.

Lemma 3.5 allows to delete all passive stocks from the market setting before solving the portfolio optimization problem under logarithmic utility.

Lemma 3.7 (areas of λ)
For each active stock i, i.e. $\pi_i^\lambda > 0$, we get $\lambda_i \leq 0$.
Vice versa if $\lambda_j > 0$ for some j, then j is a passive stock, hence $\pi_j^\lambda = 0$.

Proof. Let $\pi_i^\lambda > 0$ and assume $\lambda_i > 0$.
Then consider the strategy $\pi^{\lambda_\varepsilon}$ with $\lambda_\varepsilon = \lambda - \varepsilon e_i$ for some $0 < \varepsilon < \lambda_i$.
Note that $\delta(\lambda) = \delta(\lambda_\varepsilon)$ since $\lambda_i > 0$, $(\lambda_\varepsilon)_i > 0$ and $\lambda_j = (\lambda_\varepsilon)_j$ for all $j \neq i$.
Then for all $\varepsilon > 0$

$$\begin{aligned}
& \lambda = \arg\min_{\nu\in\mathbb{R}^d} 2\delta(\nu) + (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1}(\mu - r\mathbb{1} + \nu) \\
\Rightarrow\quad & 2\delta(\lambda_\varepsilon) + (\mu - r\mathbb{1} + \lambda_\varepsilon)^\top \Sigma^{-1}(\mu - r\mathbb{1} + \lambda_\varepsilon) \geq 2\delta(\lambda) + (\mu - r\mathbb{1} + \lambda)^\top \Sigma^{-1}(\mu - r\mathbb{1} + \lambda) \\
\Rightarrow\quad & -2(\mu - r\mathbb{1} + \lambda)^\top \Sigma^{-1}(\varepsilon e_i) + (\varepsilon e_i)^\top \Sigma^{-1}(\varepsilon e_i) \geq 0 \\
\Rightarrow\quad & -2\varepsilon(\pi^\lambda)^\top e_i + \varepsilon^2(\Sigma^{-1})_{ii} \geq 0 \\
\Rightarrow\quad & (\Sigma^{-1})_{ii}\varepsilon \geq 2\pi_i^\lambda \\
\Rightarrow\quad & \pi_i^\lambda \leq \frac{1}{2}\varepsilon(\Sigma^{-1})_{ii} \qquad \text{for all } \varepsilon > 0 \\
\Rightarrow\quad & 0 < \pi_i^\lambda \leq \lim_{\varepsilon\to 0} \frac{1}{2}\varepsilon(\Sigma^{-1})_{ii} = 0
\end{aligned}$$

But the last line is a contradiction, hence $\lambda_i \leq 0$.
Vice versa if $\lambda_i > 0$ we get by the same argumentation $\pi_i^\lambda = 0$. □

Corollary 3.8
If all stocks are positively correlated, i.e. $\Sigma_{ij} \geq 0$ for all i, j, then $\lambda_i \geq -(\mu_i - r)$ for all i.
Hence if $\pi_i^\lambda > 0$, then $\lambda_i \in \big(-(\mu_i - r), 0\,\big]$ and if $\lambda_i > 0$ then $\pi_i^\lambda = 0$.

Proof. $\pi_j^\lambda \geq 0$ for all j, hence $(\mu - r\mathbb{1} + \lambda)_i = \Sigma_{i,\cdot}\pi^\lambda \geq \Sigma_{i,i}\pi_i^\lambda \geq 0$,
hence $\lambda_i \geq r - \mu_i$ for all i.
If additionally $\pi_i^\lambda > 0$, we get $(\mu - r\mathbb{1} + \lambda)_i = \Sigma_{i,\cdot}\pi^\lambda \geq \Sigma_{i,i}\pi_i^\lambda > 0$,
hence $\lambda_i > r - \mu_i$ for all i.
The rest follows by Lemma 3.7. □

Remark 3.9
The condition in Corollary 3.8 is strict in the sense that we really need all stocks to be positively correlated. Otherwise we can construct a counterexample to Corollary 3.8 as in the following setting:

$$\Sigma = \begin{pmatrix} 0.10 & 0.05 \\ -0.05 & 0.10 \end{pmatrix} \quad , \quad \mu = \begin{pmatrix} 0.25 \\ -0.25 \end{pmatrix} \quad , \quad r = 0 \quad \Rightarrow \quad \pi^M = \Sigma^{-1}(\mu - r\mathbb{1}) = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

Due to the extreme values in μ and the resulting extreme values in the Merton-strategy, the optimal constrained strategy is trivially observed. (Also compare Algorithm 3.20 to derive the optimal solutions analytically.)

$$\pi^\lambda = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \lambda = \begin{pmatrix} -0.15 \\ 0.20 \end{pmatrix} \ngeq \begin{pmatrix} -0.25 \\ 0.25 \end{pmatrix} = -\mu$$

Then the corresponding optimal dual process λ does not fulfil $\lambda_i \geq -(\mu_i - r)$ for $i = 2$.

Corollary 3.10
Let all stocks be positively correlated.
Then all dimensions i with $\mu_i - r \leq 0$ can be deleted.

Proof. Let $\mu_i - r \leq 0$ for some i and assume $\pi_i^\lambda > 0$.
Then by Corollary 3.8 we get $-(\mu_i - r) < \lambda_i \leq 0$, hence $\mu_i - r > 0$.
This is a contradiction, hence $\pi_i^\lambda = 0$ and by Lemma 3.5 dimension i can be deleted. □

At this point we may state the first important result describing the structure of the optimal dual process λ.

Theorem 3.11 (structure of λ)
Let I and J be the active and passive dimensions of the optimal constrained portfolio strategy. Define $\bar{\lambda} := \min_k \lambda_k$.
Then $\lambda_i = \bar{\lambda}$ for all $i \in I$ (and $\lambda_j \geq \bar{\lambda}$ for all $j \in J$).

Proof. Assume there is some $i \in I$ with $\lambda_i > \bar{\lambda}$.
Consider $\lambda_\varepsilon = \lambda - \varepsilon e_i$ for $0 < \varepsilon < \lambda_i - \bar{\lambda}$. Then $\delta(\lambda_\varepsilon) = \delta(\lambda)$ and for all $\varepsilon > 0$

$$\lambda = \underset{\nu \in \mathbb{R}^d}{\arg\min}\ 2\delta(\nu) + (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1}(\mu - r\mathbb{1} + \nu)$$

$$\Rightarrow \quad 2\delta(\lambda_\varepsilon) + (\mu - r\mathbb{1} + \lambda_\varepsilon)^\top \Sigma^{-1}(\mu - r\mathbb{1} + \lambda_\varepsilon) \geq \ 2\delta(\lambda) + (\mu - r\mathbb{1} + \lambda)^\top \Sigma^{-1}(\mu - r\mathbb{1} + \lambda)$$

$$\Rightarrow \quad -2(\mu - r\mathbb{1} + \lambda)^\top \Sigma^{-1}(\varepsilon e_i) + (\varepsilon e_i)^\top \Sigma^{-1}(\varepsilon e_i) \geq 0$$

$$\Rightarrow \quad -2\varepsilon(\pi^\lambda)^\top e_i + \varepsilon^2(\Sigma^{-1})_{ii} \geq 0$$

$$
\begin{aligned}
\Rightarrow \qquad & (\Sigma^{-1})_{ii}\varepsilon \geq 2\pi_i^\lambda \\
\Rightarrow \qquad & 0 < \pi_i^\lambda \leq \frac{1}{2}\varepsilon(\Sigma^{-1})_{ii} \qquad \text{for all } \varepsilon \\
\Rightarrow \qquad & 0 < \pi_i^\lambda \leq \lim_{\varepsilon\to 0} \frac{1}{2}\varepsilon(\Sigma^{-1})_{ii} = 0
\end{aligned}
$$

This is a contradiction, hence $\lambda_i = \bar{\lambda}$ for all $i \in I$.
On the other hand $\lambda_j \geq \min_k \lambda_k = \bar{\lambda}$ is obvious. □

Remark 3.12 (Interpretation)
Theorem 3.11 is the most important result in deriving the optimal dual process λ for logarithmic utility in this chapter since it describes the structure of λ given the active and passive dimensions. This result is quite surprising, since it reveals that each active dimensions i (or stock that will be invested in) gets treated equally by reducing the corresponding expected growth rate μ_i by the same amount (adding $\bar{\lambda} < 0$). Additionally the growth rates of each passive dimensions (or stock that will not be invested in) get reduced less.
Most surprisingly we will see throughout this thesis that the structure of the optimal dual process in general only gets effected by the structure behind the support function δ and not (as one might expect) by other assumptions like the choice of a utility function.

Remark 3.13 (the boundary between active and passive)
There can be situations with $j \in J$ where $\pi_j^\lambda = 0$ and $\lambda_j = \bar{\lambda}$.
This happens if dimension j is just not good enough to be invested in.
Consider the following setting:

$$
\Sigma = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.2 \end{pmatrix} \quad , \quad \mu = \begin{pmatrix} 0.15 \\ 0.05 \end{pmatrix} \quad , \quad r = 0 \quad \Rightarrow \quad \pi^M = \Sigma^{-1}(\mu - r\mathbb{1}) = \begin{pmatrix} 1.50 \\ 0.25 \end{pmatrix}
$$

Algorithm 3.20 leads to the following optimal solution.

$$
\lambda = \begin{pmatrix} -0.05 \\ -0.05 \end{pmatrix} \quad \Rightarrow \quad \pi^\lambda = \begin{pmatrix} 1 \\ 0 \end{pmatrix}
$$

Hence $\bar{\lambda} = -0.05 = \lambda_2$ although $\pi_2^\lambda = 0$.

Theorem 3.14 (explicit calculation of λ)
Let I and J be the active and passive dimensions of the optimal constrained portfolio strategy. Define $\bar{\lambda} := \min_k \lambda_k$. Then one of the following cases holds true:

1. If $\pi^M \in K$ then $\lambda = 0$. Else:
2. If $\bar{\lambda} > 0 \;\Leftrightarrow\; \mu - r\mathbb{1} < 0$ then $\lambda = -(\mu - r\mathbb{1})$.
3. If $\bar{\lambda} = 0$ then

$$
\begin{aligned}
\lambda_J &= -(\mu - r\mathbb{1})_J + \Sigma_{JI}\pi_I^\lambda \\
&= -(\mu - r\mathbb{1})_J + \Sigma_{JI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I
\end{aligned}
$$

4. If $\bar\lambda < 0$ then
$$\begin{aligned}\bar\lambda &= \frac{1-(\mu_I - r\mathbb{1}_I)^\top(\Sigma_{II})^{-1}\mathbb{1}_I}{\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I}\\ \lambda_J &= -(\mu - r\mathbb{1})_J + \Sigma_{JI}\pi_I^\lambda\\ &= -(\mu - r\mathbb{1})_J + \Sigma_{JI}(\Sigma_{II})^{-1}(\mu_I - r\mathbb{1}_I + \bar\lambda\mathbb{1}_I)\end{aligned}$$

Proof. The optimal dual solution λ minimizes the positive function $F(\nu)$ where
$$F(\nu) := 2\delta(\nu) + (\mu - r\mathbb{1} + \nu)^\top\Sigma^{-1}(\mu - r\mathbb{1} + \nu)$$

Case 1 is trivial. The unconstrained strategy is already admissible, therefore we don't need to consider auxiliary markets.

In Case 2 let $\mu - r\mathbb{1} < 0$.
Then for $\lambda = -(\mu - r\mathbb{1}) > 0$ we get $F(\lambda) = 2\delta(\lambda) = 0$ since $\lambda > 0$.
But F is a positive function, hence $\lambda = -(\mu - r\mathbb{1})$ minimizes F.
Vice versa let $\bar\lambda > 0$ and hence $\lambda > 0$.
Then by Lemma 3.7 $\pi^\lambda = 0$ and hence $\lambda = -(\mu - r\mathbb{1})$ and hence $\mu - r < 0$.

In the Cases 3 and 4 we observe
$$\begin{aligned} & \begin{pmatrix}\pi_I^\lambda\\ 0_J\end{pmatrix} = \pi^\lambda = \Sigma^{-1}\begin{pmatrix}\mu_I - r\mathbb{1}_I + \bar\lambda\mathbb{1}_I\\ \mu_J - r\mathbb{1}_J + \lambda_J\end{pmatrix}\\ \Rightarrow\quad & \begin{pmatrix}\Sigma_{II}\pi_I^\lambda\\ \Sigma_{JI}\pi_I^\lambda\end{pmatrix} = \Sigma\pi^\lambda = \begin{pmatrix}\mu_I - r\mathbb{1}_I + \bar\lambda\mathbb{1}_I\\ \mu_J - r\mathbb{1}_J + \lambda_J\end{pmatrix}\\ \Rightarrow\quad & \lambda_J = -(\mu - r\mathbb{1})_J + \Sigma_{JI}\pi_I^\lambda\\ \text{and}\quad & \pi_I^\lambda = (\Sigma_{II})^{-1}\left(\mu_I - r\mathbb{1}_I + \bar\lambda\mathbb{1}_I\right)\end{aligned} \tag{3.12}$$

In Case 4 we have additionally $\bar\lambda < 0$ and hence $\delta(\bar\lambda)$ is differentiable in $\mathbb{R}$.
Hence
$$\begin{aligned}F(\lambda) &= 2\delta(\lambda) + (\mu - r\mathbb{1} + \lambda)^\top\Sigma^{-1}(\mu - r\mathbb{1} + \lambda)\\ &= 2\delta(\bar\lambda) + (\mu - r\mathbb{1} + \lambda)^\top\pi^\lambda\\ &= -2\bar\lambda + (\mu - r\mathbb{1} + \lambda)_I^\top\pi_I^\lambda\\ &= -2\bar\lambda + \left(\mu_I - r\mathbb{1}_I + \bar\lambda\mathbb{1}_I\right)^\top(\Sigma_{II})^{-1}\left(\mu_I - r\mathbb{1}_I + \bar\lambda\mathbb{1}_I\right)\\ &= \bar\lambda^2\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I + 2\bar\lambda\big(-1 + (\mu_I - r\mathbb{1}_I)^\top(\Sigma_{II})^{-1}\mathbb{1}_I\big) + (\mu_I - r\mathbb{1}_I)^\top(\Sigma_{II})^{-1}(\mu_I - r\mathbb{1}_I)\\ &=: F(\bar\lambda)\end{aligned}$$

Differentiating leads to the optimal result.
$$\begin{aligned} & 0 = \partial_{\bar\lambda}F = 2\bar\lambda\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I - 2 + 2(\mu_I - r\mathbb{1}_I)^\top(\Sigma_{II})^{-1}\mathbb{1}_I\\ \text{with}\quad & \partial_{\bar\lambda}^2F = 2\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I > 0\\ \Rightarrow\quad & \bar\lambda = \frac{1-(\mu_I - r\mathbb{1}_I)^\top(\Sigma_{II})^{-1}\mathbb{1}_I}{\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I}\end{aligned}$$

□

Remark 3.15
Note that the Cases 1-3 in Theorem 3.14 are equivalent to only using the 'no-short-selling' constraint. On the other hand Case 4 corresponds to the restriction to always invest 100% of your wealth:

$$\begin{aligned}\mathbb{1}^\top \pi_I^\lambda &\overset{(3.12)}{=} \mathbb{1}^\top (\Sigma_{II})^{-1} \left(\mu_I - r\mathbb{1}_I + \mathbb{1} \frac{1 - (\mu_I - r\mathbb{1}_I)^\top (\Sigma_{II})^{-1}\mathbb{1}}{\mathbb{1}^\top (\Sigma_{II})^{-1}\mathbb{1}} \right) \\ &= \mathbb{1}^\top (\Sigma_{II})^{-1} (\mu_I - r\mathbb{1}_I) + \left(1 - (\mu_I - r\mathbb{1}_I)^\top (\Sigma_{II})^{-1}\mathbb{1} \right) = 1\end{aligned}$$

3.2 Determining the correct case and active dimensions

Theorem 3.14 provides four different cases for the optimal dual solution λ, but it does not determine the correct case for some specific time and filter value. Additionally it requires knowledge about the active and passive dimensions I and J that we don't know yet. In this section we will provide several results about choosing the correct case and the correct active dimensions and close with an algorithm that provides the respective correct choices.

Remark 3.16
In each active dimension i the given growth rate μ_i gets reduced by the same amount $\lambda_i = \bar{\lambda}$. However we cannot conclude to invest in those stocks with largest μ_i, with largest $(\sigma^{-1}(\mu - r\mathbb{1}))_i$ or with largest $\pi_i^M = (\Sigma^{-1}(\mu - r\mathbb{1}))_i$ as the following easy counterexamples show.

Consider the following setting with large volatility for stock 1:

$$\sigma = \begin{pmatrix} 0.6 & 0.1 \\ 0.1 & 0.1 \end{pmatrix} \quad , \quad \mu = \begin{pmatrix} 0.10 \\ 0.06 \end{pmatrix} \quad , \quad r = 0 \quad \Rightarrow \quad \pi^M = \Sigma^{-1}(\mu - r\mathbb{1}) = \begin{pmatrix} -0.88 \\ 6.08 \end{pmatrix}$$

The optimal solution is given by Algorithm 3.20.

$$I = \{2\} \quad , \quad \lambda = \begin{pmatrix} -0.03 \\ -0.04 \end{pmatrix} \quad \Rightarrow \quad \pi^\lambda = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Although $\mu_1 > \mu_2$ only stock 2 enters the optimal constrained strategy.

Now consider the following setting with large correlation between the stocks:

$$\sigma = \begin{pmatrix} 0.25 & 0.16 \\ 0.1 & 0.1 \end{pmatrix} \quad , \quad \mu = \begin{pmatrix} 0.10 \\ 0.05 \end{pmatrix} \quad , \quad r = 0 \quad \Rightarrow \quad \pi^M = \Sigma^{-1}(\mu - r\mathbb{1}) \approx \begin{pmatrix} -0.62 \\ 3.77 \end{pmatrix}$$

The optimal solution is given by Algorithm 3.20.

$$I = \{1\} \quad , \quad \lambda \approx \begin{pmatrix} -0.012 \\ -0.009 \end{pmatrix} \quad \Rightarrow \quad \pi^\lambda = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Although $(\sigma^{-1}\mu)_1 \approx 0.22 < 0.28 \approx (\sigma^{-1}\mu)_2$ and $\pi_1^M \approx -0.62 < 3,77 \approx \pi_2^M$ only stock 1 enters the optimal constrained strategy.

However, we can show in the following lemmas that several intuitively false choices of active and passive dimensions in fact lead to non-admissible solutions, revealing that the choice was wrong:

Lemma 3.17 (Choice of I in Case 3)
Let I and J be the active and passive dimensions such that (λ, π) is the optimal solution in Case 3 and let (λ', π') be the solution if we chose I' and J' instead.

1) Let $i \in I$ with $\pi_i > 0$ be an active dimension.
   If we choose $I' = I \setminus \{i\}$ and $J' = J \cup \{i\}$ then we get $\lambda'_i < 0$.
2) Let $j \in J$ with $\lambda_j > 0$ be a passive dimension (by Lemma 3.7).
   If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ then we get $\pi'_j < 0$.

Proof. The proof is given in Appendix A.3.1. □

Lemma 3.18 (Choice of I in Case 4)
Let I and J be the active and passive dimensions such that (λ, π) is the optimal solution in Case 4 and let (λ', π') be the solution if we chose I' and J' instead.
1) Let $i \in I$ with $\pi_i > 0$ be an active dimension.
   If we choose $I' = I \setminus \{i\}$ and $J' = J \cup \{i\}$ then we get $\lambda'_i < \bar{\lambda}'$.
2) Let $j \in J$ with $\lambda_j > \bar{\lambda}$ be a passive dimension (by Theorem 3.11).
   If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ then we get $\pi'_j < 0$.

Proof. The proof is given in Appendix A.3.2. □

Corollary 3.19 (Dimensions on the boundary)
Let I and J be the active and passive dimensions such that (λ, π) is the optimal solution in Case 4 and let (λ', π') be the solution if we chose I' and J' instead.

If $j \in J$ is a passive dimension that is almost invested in, hence $\lambda_j = \bar{\lambda}$ according to Remark 3.13, then j can also be considered an active dimension.
In particular:

Let $j \in J$ with $\pi_j = 0$ and $\lambda_j = \bar{\lambda}$ be a passive dimension that is almost invested in.
Then: If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ we get $\lambda' = \lambda$ and $\pi' = \pi$.

Proof. The proof is given in Appendix A.3.3. □

Obviously the same holds true in Case 3 for $\bar{\lambda} = 0$.

Algorithm 3.20 (solving the portfolio optimization problem explicitly)
The constrained portfolio optimization problem with dual problem (3.6) gets solved via going through the following four cases until hitting an admissible solution.
When hitting an admissible solution this provides the optimal dual process and the optimal constrained portfolio strategy.
Recall that μ is short for the filter value $\hat{\mu}_t$ at our current time t.

1. Compute the optimal unconstrained strategy $\pi^M = \Sigma^{-1}(\mu - r\mathbb{1})$.
   If $\pi^M \in K$ then $\lambda = 0$ and $\pi^\lambda = \pi^M$.
   Else continue.

2. If $\mu - r\mathbb{1} < 0$ then $\lambda = -(\mu - r\mathbb{1})$ and $\pi^\lambda = 0$.
   Else continue.

3. For each $\emptyset \neq I \subset \{1, ..., d\}$ (with $J = \{1, ..., d\} \setminus I$) compute

$$\begin{aligned} \pi_I^\lambda &= (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I && \text{with} \quad \pi_J^\lambda = 0_J \\ \lambda_J &= -(\mu - r\mathbb{1})_J + \Sigma_{JI}\pi_I^\lambda && \text{with} \quad \lambda_I = 0_I \end{aligned}$$

   If $\pi^\lambda \notin K$ or $\lambda_j < 0$ for some $j \in J$, then I was the wrong choice.
   Else continue.

4. For each $\emptyset \neq I \subseteq \{1, ..., d\}$ (with $J = \{1, ..., d\} \setminus I$) compute

$$\bar{\lambda}^{(I)} = \frac{1 - (\mu_I - r\mathbb{1}_I)^\top (\Sigma_{II})^{-1} \mathbb{1}_I}{\mathbb{1}_I^\top (\Sigma_{II})^{-1} \mathbb{1}_I}$$
$$\pi_I^\lambda = (\Sigma_{II})^{-1} \big(\mu_I - r\mathbb{1}_I + \bar{\lambda}^{(I)} \mathbb{1}_I\big) \qquad \text{with} \quad \pi_J^\lambda = 0_J$$
$$\lambda_J = -(\mu - r\mathbb{1})_J + \Sigma_{JI} \pi_I^\lambda$$

If $\bar{\lambda}^{(I)} \geq 0$ or $\pi^\lambda \notin K$ or $\lambda_j < \bar{\lambda}^{(I)}$ for some $j \in J$, then I was the wrong choice. Stop whenever a solution is admissible.

Recall that any relational operator applied to vectors is meant component-wisely.

Theorem 3.21
The above algorithm always results in the unique strategy.

Proof. Case 1:
If the optimal unconstrained strategy is admissible, there is nothing left to solve. Otherwise Theorem 3.14 ensures that the optimal λ is in one of the Cases 2 to 4. Therefore by checking all possibilities we find in particular the optimal choice.

Case 2: $(\bar{\lambda} > 0)$
Theorem 3.14 shows that Case 2 is equivalent to checking the condition $\mu - r\mathbb{1} < 0$ and leads to $\pi^\lambda = 0$.

Cases 3 $(\bar{\lambda} = 0)$ and 4 $(\bar{\lambda} < 0)$:
By Theorem 2.9 we know that if $\lambda \in \tilde{K}$, $\pi^\lambda \in K$ and $\lambda^\top \pi^\lambda + \delta(\lambda) = 0$ then π^λ is already the optimal portfolio strategy for the constrained portfolio optimization problem. Obviously $\lambda \in \tilde{K} = \mathbb{R}^d$ is always true.

By the structure of the optimal solution in Case 3 we observe if $\pi^\lambda \in K$:

$$\lambda = \begin{pmatrix} 0_I \\ \lambda_J \end{pmatrix} \quad \text{and} \quad \pi^\lambda = \begin{pmatrix} \pi_I^\lambda \\ 0_J \end{pmatrix} \quad \Rightarrow \quad \lambda^\top \pi^\lambda = 0$$

Hence if $\lambda_j \geq 0$ for all $j \in J$ then $\delta(\lambda) = 0$ and hence $\lambda^\top \pi^\lambda + \delta(\lambda) = 0$. Therefore (λ, π^λ) is the optimal solution.

On the other hand in Case 4 we observe if $\pi^\lambda \in K$:

$$\lambda = \begin{pmatrix} \bar{\lambda} \mathbb{1}_I \\ \lambda_J \end{pmatrix} \quad \text{and} \quad \pi^\lambda = \begin{pmatrix} \pi_I^\lambda \\ 0_J \end{pmatrix} \text{ with } \mathbb{1}_I^\top \pi_I^\lambda = 1 \quad \Rightarrow \quad \lambda^\top \pi^\lambda = \bar{\lambda}$$

Hence if $\lambda_j \geq \bar{\lambda}$ for all $j \in J$ and $\bar{\lambda} < 0$ then $\delta(\lambda) = -\bar{\lambda}$ and hence $\lambda^\top \pi^\lambda + \delta(\lambda) = 0$. Therefore (λ, π^λ) is the optimal solution. □

Remark 3.22 (Non-unique active and passive dimensions)
The proof of Theorem 3.21 shows that every choice of I that does not harm the structural assumptions made in Algorithm 3.20, satisfies the assumptions of Theorem 2.9 and hence the resulting admissible strategy is the optimal solution.

As the optimal portfolio strategy in our market setting is unique, two different but admissible choices of I can only lead to the same resulting optimal strategy. This happens in view of Corollary 3.19 whenever some stock is on the boundary between being an active or passive dimension. For instance in the situation as described in remark (3.13) both choices $I = \{1\}$ and $I = \{1, 2\}$ lead the same (optimal) strategy.

Remark 3.23
The Algorithm 3.20 is a very fast and explicit way to determine the optimal solution.
If we wanted to solve the dual optimization problem (3.6) with a standard non-linear optimization approach, we would need to consider $\mathbb{R}^d$ possible choices for λ with some numerical approach. Algorithm 3.20 on the other hand only needs to check four cases and several choices of active dimensions and hence at most $1+1+(2^d-2)+(2^d-1) = 2^{d+1}-1$ possible choices for the optimal dual process λ. Additionally the algorithm can stop as soon as it detects an admissible solution.

Therefore when implementing the algorithm in Chapter 6 we add an additional step to sort the different choices for I approximately such that the algorithm checks those choices first that are more plausible to appear. This decreases the running time of the algorithm substantial as we will see in Chapter 6.

Remark 3.24
When comparing our structured form of the optimal solution to the quite involved formulas provided in [CK92, Ex. 14.9] for the two-dimensional case, one realizes the immense gain of using Algorithm 3.20.

4 Constrained optimization under power utility

This chapter contains the main part of this thesis. We consider the constrained portfolio optimization problem (2.3) under partial information and convex constraints. We restrict the set of admissible strategies $\pi \in \mathcal{A}$ to strategies in the closed convex set K that consists of the no short-selling and the no borrowing constraints as introduced in (3.7):

$$\begin{aligned} \pi \in \mathcal{A} &= \left\{ \pi \text{ admissible strategy } \middle| \pi_t \in K \text{ for all } t \in [0,T] \right\} \\ K &= \left\{ x \in [0,\infty)^d \middle| x^\top \mathbb{1} \leq 1 \right\} \end{aligned}$$

This results in the following support function δ.

$$\delta(x) = \max_j \, (x_j)^- = \begin{cases} 0, & \text{if } x \in [0,\infty)^d \\ -x_i, & \text{if } x_i = \min_j x_j < 0 \end{cases} \quad , \quad \tilde{K} = \mathbb{R}^d$$

Remark 4.1 (Choice of utility function)
In the previous chapter we solved this problem under logarithmic utility. We observed that the resulting optimal portfolio strategy only uses the filter value $\hat{\mu}_t$, but not its variance γ_t. This is a quite common feature of results under logarithmic utility and gives reason to consider logarithmic utility being too less risk averse.

Power utility (4.1) on the other hand generalizes logarithmic utility in the class of constant relative risk aversion utility functions. Also it provides the opportunity to adjust any personal risk-aversion level via the risk-aversion parameter α, where the risk aversion increases when α decreases towards $-\infty$.

$$U(x) = \frac{1}{\alpha} x^\alpha \qquad \begin{cases} \text{for } \alpha \in (-\infty, 0) & \text{negative power utility} \\ \text{for } \alpha \in (0,1) & \text{positive power utility} \end{cases} \tag{4.1}$$

Note that we also see in the literature that most results under power utility converge to the respective results for logarithmic utility when α goes to 0.

Since considering both, positive and negative power utility, might significantly complicate the notation in our derivations we will only focus on one. *Positive power utility* is known to be even less risk averse than logarithmic utility and hence it is considered economically less important. Therefore we will focus on *negative power utility* in the following derivations of this chapter.
However note that the whole approach of this chapter can also be conducted for positive power utility.

In Section 4.1 we start by solving the unconstrained portfolio optimization problem under partial information and observe a different optimal portfolio strategy than the Merton strategy.
In Sections 4.2 and 4.3 we solve the constrained portfolio optimization problem by a reverse separation approach. We separate the dual admissibility region into several cases that can each be solved analytically. In Section 4.4 we examine these cases in more detail and in particular the regions of the domain of the dual value function that correspond to these cases.
Finally in Section 4.5 we combine the separated cases and sketch the proof to derive optimality for the whole portfolio optimization problem.

4.1 The unconstrained optimization problem

Under logarithmic utility the optimal unconstrained solution is given by the Merton strategy. Since this strategy is known to be too simple, this fortunately does not hold true for power utility. In the following subsections we present different approaches leading to the optimal unconstrained strategy (4.2).

$$\pi_t^{\text{unc}} = \frac{1}{1-\alpha}\Sigma^{-1}C(t)\gamma_t^{-1}(\hat{\mu}_t - r\mathbb{1}) \tag{4.2}$$
$$\text{where} \quad C(t) = \left(\gamma_T^{-1} - \frac{1}{1-\alpha}(T-t)\Sigma^{-1}\right)^{-1}$$

4.1.1 Martinagale approaches

There are several martingale approaches that deal with unconstrained portfolio optimization problems under partial information. One of the first to derive a quite explicit solution are Cvitanic et al:

Theorem 4.2 ([CLMZ06])
The optimal unconstrained portfolio strategy in our market setting is given by

$$\pi_t^{\text{unc}} = (\sigma^{-1})^\top P^\top A_t^{-1} P\hat{\theta}_t \tag{4.3}$$

where $\hat{\theta}_t = \sigma^{-1}(\hat{\mu}_t - r\mathbb{1})$ is the filter for the risk-premium $\theta = \sigma^{-1}(\mu - r\mathbb{1})$ with initial distribution $\theta \sim \mathcal{N}(\theta_0, \Delta)$. This filter is described in Section 2.2 as

$$\hat{\theta}_t = \mathbb{E}\left[\theta \middle| \mathcal{F}_t^S\right] = P^\top \bar{D}_t \left(PW_t^{\mathbb{Q}} + \bar{D}_0^{-1}P\theta_0\right).$$

Here P is an orthogonal matrix that diagonalizes Δ such that

$$D = \operatorname{diag}(..., d_i, ...) \quad \text{, where} \quad D = P\Delta P^\top$$
$$\bar{D}_t := \operatorname{diag}(..., \delta_i(t), ...) \quad \text{, where} \quad \delta_i(t) = \frac{d_i}{1+d_i t}$$
$$A_t := \operatorname{diag}(..., A_i(t), ...) \quad \text{, where} \quad A_i(t) = (1-\alpha) - \alpha\delta_i(t)(T-t)$$

Proof. The proof is given in [CLMZ06, Appendix A.1].
Additionally we show in Appendix A.4.3 that this form of the optimal unconstrained solution can also be derived via another approach, verifying that (4.2) and (4.3) are the same formula. □

The main ingredient of the proof in [CLMZ06] is the ability to explicitly calculate the following conditional expectation:

Lemma 4.3
Define

$$Z_t := \exp\left(-\int_0^t \hat{\theta}_s^\top dW_s^{\mathbb{Q}} - \frac{1}{2}\int_0^t \left\|\hat{\theta}_s\right\|^2 ds\right)$$

$$\bar{\theta}_t := P\hat{\theta}_t \quad , \quad \hat{\theta}_t = P^\top \bar{D}_t \left(PW_t^{\mathbb{Q}} + D_0^{-1} P\theta_0 \right)$$

then

$$\mathbb{E}^{\mathbb{Q}} \left[\left(\frac{Z_T}{Z_t} \right)^\alpha \middle| \mathcal{F}_t^S \right] = \prod_{i=1}^n g_\alpha \left(T - t, \bar{\theta}_i(t), \delta_i(t) \right)$$

$$\text{where} \quad g_\alpha(\tau, x, y) = \sqrt{\frac{(1 + y\tau)^{\alpha+1}}{1 + y\tau(\alpha + 1)}} \exp \frac{\alpha(1 + \alpha)x^2\tau}{2(1 + y\tau(\alpha + 1))}$$

Proof. Note that $\bar{W}_t^{\mathbb{Q}} := PW_t^{\mathbb{Q}}$ is also a Brownian motion and hence the components of $\bar{\theta}_t$ are independent of each other.

$$\bar{\theta}_t = P\hat{\theta}_t = \bar{D}_t \left(\bar{W}_t^{\mathbb{Q}} + D_0^{-1}\bar{\theta}_0 \right)$$

$$\Rightarrow \qquad \bar{\theta}_i(t) = \delta_i(t)\bar{W}_i^{\mathbb{Q}}(t) + \frac{\delta_i(t)}{d_i}\bar{\theta}_i(0)$$

Therefore we get

$$\begin{aligned} Z_t &= \exp \left(- \int_0^t \hat{\theta}_s^\top dW_s^{\mathbb{Q}} - \frac{1}{2} \int_0^t \left\| \hat{\theta}_s \right\|^2 ds \right) \\ &= \exp \left(- \int_0^t \bar{\theta}_s^\top d\bar{W}_s^{\mathbb{Q}} - \frac{1}{2} \int_0^t \left\| \bar{\theta}_s \right\|^2 ds \right) \\ &= \prod_{i=1}^n \underbrace{\exp \left(- \int_0^t \bar{\theta}_i(s) d\bar{W}_i^{\mathbb{Q}}(s) - \frac{1}{2} \int_0^t \bar{\theta}_i(s)^2 ds \right)}_{=:Z_i(t)} \end{aligned}$$

$$\Rightarrow \quad \mathbb{E}^{\mathbb{Q}} \left[\left(\frac{Z_T}{Z_t} \right)^\alpha \middle| \mathcal{F}_t^S \right] = \prod_{i=1}^n \mathbb{E}^{\mathbb{Q}} \left[\left(\frac{Z_i(T)}{Z_i(t)} \right)^\alpha \middle| \mathcal{F}_t^S \right]$$

Now the calculation continues like in [CLMZ06]. □

In particular we get for $t = 0$

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[\left(\frac{Z_T}{Z_0} \right)^\alpha \middle| \mathcal{F}_0^S \right] &= \mathbb{E}^{\mathbb{Q}} \left[\exp \left(-\alpha \int_0^T \hat{\theta}_t^\top dW_t^{\mathbb{Q}} - \frac{\alpha}{2} \int_0^T \left\| \hat{\theta}_t \right\|^2 dt \right) \right] \\ &= \prod_{i=1}^n \sqrt{\frac{(1 + d_iT)^{\alpha+1}}{1 + d_iT(\alpha + 1)}} \exp \frac{\alpha(1 + \alpha)\bar{\theta}_i(0)^2T}{2(1 + d_iT(\alpha + 1))} \end{aligned}$$

whenever Z_t can be expressed component-wisely

Remark 4.4

There are lots of martingale approaches trying to solve similar problems. In particular [BDL10] and [Lak98] can be used to calculate this form of the optimal solution, which is shown in Appendix A.4.1 and A.4.2.

However each of these approaches requires values that are not trivially observed like the orthogonal diagonalization of Δ or the square root σ of the covariance matrix $\Sigma = \sigma\sigma^\top$. Hence this form of the optimal unconstrained portfolio strategy is not yet in perfect shape.

4.1.2 Stochastic control approach

In this section we want to present a stochastic control approach to deal with the unconstrained portfolio optimization problem under partial information. We will come up with the explicit solution (4.2) for the optimal portfolio strategy without the drawbacks of the above martingale approaches and provide a verification theorem.

Our goal is to maximize expected power utility $U(x) = \frac{1}{\alpha}x^\alpha$ for $\alpha < 0$ under the observation filtration $\mathcal{F}^S$. The resulting stochastic control problem is to determine $V(0, x_0, \mu_0)$ where the value function is given by

$$V(t,x,\mu) := \sup_{\pi\in\mathcal{A}(t,x,\mu)} \mathbb{E}\left[\frac{1}{\alpha}X_T^\alpha \middle| (X_t,\hat{\mu}_t) = (x,\mu)\right] \tag{4.4}$$

Here $\mathcal{A}(t,x,\mu)$ is the set of admissible strategies in $[t,T]$ such that $(X_t,\hat{\mu}_t) = (x,\mu)$. Note that $\mathcal{A} = \mathcal{A}(0,x_0,\mu_0)$.

In order to make the controlled process Markov, we also need to consider the dynamics of the filter given in (2.12). Then the controlled process evolves like

$$d\begin{pmatrix} X_t \\ \hat{\mu}_t \end{pmatrix} = \begin{pmatrix} X_t\left(r + \pi_t^\top(\hat{\mu}_t - r\mathbb{1})\right) \\ 0 \end{pmatrix} dt + \begin{pmatrix} X_t\pi_t^\top\sigma \\ \gamma_t(\sigma^{-1})^\top \end{pmatrix} dV_t$$

If the Bellmann principle holds for $t_1 > t$ we get

$$V(t,x,\mu) = \sup_{\pi\in\mathcal{A}(t,x,\mu)} \mathbb{E}\left[V\big(t_1, X_{t_1}, \hat{\mu}_{t_1}\big)\big| (X_t,\hat{\mu}_t) = (x,\mu)\right]$$

With Itô's formula it follows

$$\begin{aligned} V(t,x,\mu) = \sup_{\pi\in\mathcal{A}(t,x,\mu)} \mathbb{E}\Bigg[V(t) &+ \int_t^{t_1} V_t(s) + V_x(s)X_s\left(r + \pi_s^\top(\hat{\mu}_s - r\mathbb{1})\right)ds \\ &+ \int_t^{t_1} V_x(s)X_s\pi_s^\top\sigma + V_\mu(s)^\top\gamma_s(\sigma^{-1})^\top dV_s \\ &+ \frac{1}{2}\int_t^{t_1} \operatorname{tr}\big(HV(s)\cdot a(s)\big)ds \,\Bigg|(X_t,\hat{\mu}_t) = (x,\mu)\Bigg] \end{aligned}$$

where we abbreviate $V(s) := V(s, X_s, \hat{\mu}_s)$ with derivatives V_t, V_x, V_μ and Hessian matrix HV. $a(s)$ is called diffusion matrix.

$$\begin{aligned} a(s) &:= \begin{pmatrix} X_s\pi_s^\top\sigma \\ \gamma_s(\sigma^{-1})^\top \end{pmatrix}\begin{pmatrix} X_s\pi_s^\top\sigma \\ \gamma_s(\sigma^{-1})^\top \end{pmatrix}^\top = \begin{pmatrix} X_s\pi_s^\top\Sigma\pi_sX_s & X_s\pi_s^\top\gamma_s \\ \gamma_s\pi_sX_s & \gamma_s\Sigma^{-1}\gamma_s \end{pmatrix} \\ \operatorname{tr}\big(HV(s)\cdot a(s)\big) &= V_{xx}a_{11} + V_{x\mu}^\top a_{21} + \operatorname{tr}\left(V_{\mu x}a_{12} + V_{\mu\mu}a_{22}\right) \\ &= V_{xx}a_{11} + 2V_{x\mu}^\top a_{21} + \operatorname{tr}\left(V_{\mu\mu}a_{22}\right) \end{aligned}$$

Now if $\int_t^{t_1} V_x(s)X_s\pi_s^\top\sigma + V_\mu(s)^\top\gamma_s(\sigma^{-1})^\top dV_s$ is a martingale and we may switch the supremum with the integration, we observe for $t_1 \to t$ the following HJB-equation.

$$0 = \sup_{\pi\in\mathcal{A}(t,x,\mu)} \mathbb{E}\left[\int_t^{t_1} V_t(s) + V_x(s)X_s\left(r + \pi_s^\top(\hat{\mu}_s - r\mathbb{1})\right)ds\right.$$

$$+\frac{1}{2}\int_t^{t_1}\operatorname{tr}\big(HV(s)\cdot a(s)\big)ds\;\bigg|(X_t,\hat{\mu}_t)=(x,\mu)\bigg]$$

$$\Rightarrow\quad 0=\sup_{\pi}\left(V_t+V_x x\left(r+\pi^\top(\mu-r\mathbb{1})\right)+\frac{1}{2}\operatorname{tr}\left(HV\cdot a\right)\right) \tag{4.5}$$

where a is a quadratic function in x. The HJB-equation (4.5) has to be solved with respect to the boundary condition

$$V(T,x,\mu)=\sup_{\pi\in\mathcal{A}(T,x,\mu)}\mathbb{E}\left[\frac{1}{\alpha}X_T^\alpha\bigg|(X_T,\hat{\mu}_T)=(x,\mu)\right]=\frac{1}{\alpha}x^\alpha\qquad\text{for all }x,\mu$$

Now we can find the optimal control π^* that maximizes the HJB-equation (4.5) by considering the derivative of its inner part with respect to π:

$$0\overset{!}{=}V_x x(\mu-r\mathbb{1})+V_{xx}x^2\Sigma\pi^*+x\gamma_t V_{x\mu}$$

$$\Rightarrow\quad \pi^*=-\frac{V_x}{xV_{xx}}\Sigma^{-1}(\mu-r\mathbb{1})-\frac{1}{xV_{xx}}\Sigma^{-1}\gamma_t V_{x\mu} \tag{4.6}$$

Solving the HJB-equation

We solve the HJB-equation by first plugging in π^*:

$$\begin{aligned}
0&=V_t+rxV_x-(\mu-r\mathbb{1})^\top\Sigma^{-1}(\mu-r\mathbb{1})\frac{V_x^2}{V_{xx}}-(\mu-r\mathbb{1})^\top\Sigma^{-1}\gamma_t\frac{V_x}{V_{xx}}V_{x\mu}\\
&\quad+\frac{1}{2}x^2\pi^\top\Sigma\pi V_{xx}+x\pi^\top\gamma_t V_{x\mu}+\frac{1}{2}\operatorname{tr}\left(V_{\mu\mu}\gamma_t\Sigma^{-1}\gamma_t\right)\\
&=V_t+rxV_x-(\mu-r\mathbb{1})^\top\Sigma^{-1}(\mu-r\mathbb{1})\frac{V_x^2}{V_{xx}}-(\mu-r\mathbb{1})^\top\Sigma^{-1}\gamma_t\frac{V_x}{V_{xx}}V_{x\mu}\\
&\quad+\frac{1}{2}\frac{1}{V_{xx}}\left(V_x(\mu-r\mathbb{1})^\top+V_{x\mu}^\top\gamma_t\right)\Sigma^{-1}\big((\mu-r\mathbb{1})V_x+\gamma_t V_{x\mu}\big)\\
&\quad-\frac{V_x}{V_{xx}}(\mu-r\mathbb{1})^\top\Sigma^{-1}\gamma_t V_{x\mu}-\frac{1}{V_{xx}}V_{x\mu}^\top\gamma_t\Sigma^{-1}\gamma_t V_{x\mu}+\frac{1}{2}\operatorname{tr}\left(V_{\mu\mu}\gamma_t\Sigma^{-1}\gamma_t\right)\\
&=V_t+rxV_x-\frac{1}{2}(\mu-r\mathbb{1})^\top\Sigma^{-1}(\mu-r\mathbb{1})\frac{V_x^2}{V_{xx}}\\
&\quad-\frac{1}{2}\frac{1}{V_{xx}}V_{x\mu}^\top\gamma_t\Sigma^{-1}\gamma_t V_{x\mu}-\frac{V_x}{V_{xx}}(\mu-r\mathbb{1})^\top\Sigma^{-1}\gamma_t V_{x\mu}+\frac{1}{2}\operatorname{tr}\left(V_{\mu\mu}\gamma_t\Sigma^{-1}\gamma_t\right)
\end{aligned}$$

Now we use the multiplicative ansatz

$$V(t,x,\mu)=U(x)e^{f(t,\mu)}=\frac{x^\alpha}{\alpha}e^{f(t,\mu)}$$

with corresponding boundary condition $f(T,\mu)=0$ for all μ.
This leads to the following HJB-equation:

$$0=f_t+\alpha r-\frac{1}{2}\frac{\alpha}{\alpha-1}(\mu-r\mathbb{1})^\top\Sigma^{-1}\left(\mu-r\mathbb{1}\right)-\frac{\alpha}{\alpha-1}(\mu-r\mathbb{1})^\top\Sigma^{-1}\gamma_t f_\mu$$

$$
\begin{aligned}
&- \frac{1}{2}\frac{\alpha}{\alpha-1} f_\mu^\top \gamma_t \Sigma^{-1} \gamma_t f_\mu + \frac{1}{2} f_\mu^\top \gamma_t \Sigma^{-1} \gamma_t f_\mu + \frac{1}{2}\mathrm{tr}\left(f_{\mu\mu} \gamma_t \Sigma^{-1} \gamma_t\right) \\
&= f_t + \alpha r - \frac{1}{2}\frac{\alpha}{\alpha-1} \left((\mu - r\mathbb{1}) + \gamma_t f_\mu\right)^\top \Sigma^{-1} \left((\mu - r\mathbb{1}) + \gamma_t f_\mu\right) \\
&+ \frac{1}{2} f_\mu^\top \gamma_t \Sigma^{-1} \gamma_t f_\mu + \frac{1}{2}\mathrm{tr}\left(f_{\mu\mu} \gamma_t \Sigma^{-1} \gamma_t\right)
\end{aligned}
$$

We improve the ansatz further:

$$
\begin{aligned}
f(t,\mu) &= \frac{1}{2}(\mu - r\mathbb{1})^\top A(t)(\mu - r\mathbb{1}) + k(t) \qquad \text{(with symmatric } A\text{)} \\
\text{with} \quad f_t(t,\mu) &= \frac{1}{2}(\mu - r\mathbb{1})^\top A'(t)(\mu - r\mathbb{1}) + k'(t) \\
f_\mu(t,\mu) &= A(t)(\mu - r\mathbb{1}) \\
f_{\mu\mu}(t,\mu) &= A(t)
\end{aligned}
$$

with corresponding boundary conditions $A(T) = 0$ and $k(T) = 0$.
The HJB-equation then becomes a quadratic differential equation in $(\mu - r\mathbb{1})$.

$$
\begin{aligned}
&\frac{1}{2}(\mu - r\mathbb{1})^\top A'(t)(\mu - r\mathbb{1}) + k'(t) \\
&= f_t = -\alpha r - \frac{1}{2}\frac{\alpha}{1-\alpha}(\mu - r\mathbb{1})^\top \big(Id + \gamma_t A(t)\big)^\top \Sigma^{-1} \big(Id + \gamma_t A(t)\big)(\mu - r\mathbb{1}) \\
&- \frac{1}{2}(\mu - r\mathbb{1})^\top A(t)\gamma_t \Sigma^{-1} \gamma_t A(t)(\mu - r\mathbb{1}) - \frac{1}{2}\mathrm{tr}\left(A(t)\gamma_t \Sigma^{-1} \gamma_t\right)
\end{aligned}
$$

Hence solving the HJB-equation (4.5) boils down to solving two ordinary differential equations in t.

$$
A'(t) = -\frac{\alpha}{1-\alpha}\big(Id + \gamma_t A(t)\big)^\top \Sigma^{-1} \big(Id + \gamma_t A(t)\big) - A(t)\gamma_t \Sigma^{-1} \gamma_t A(t) \tag{4.7}
$$

$$
k'(t) = -\alpha r - \frac{1}{2}\mathrm{tr}\left(A(t)\gamma_t \Sigma^{-1} \gamma_t\right) \tag{4.8}
$$

Solving the ordinary differential equations

The first ODE is a matrix Ricatti equation that can only be solved numerically in general. However in our case, we can solve it explicitly by defining $B(t) := Id + \gamma_t A(t)$ and vice versa $A(t) = \gamma_t^{-1}(B(t) - Id)$ with corresponding boundary condition $B(T) = Id$. Hence $A'(t) = \gamma_t^{-1} B'(t) + \Sigma^{-1} B(t)$ and the ODE (4.7) then becomes

$$
\gamma_t^{-1} B'(t) = -\frac{\alpha}{1-\alpha} B(t)^\top \Sigma^{-1} B(t) - B(t)^\top \Sigma^{-1} (B(t) - I)
$$

Now we define $C(t) := B(t)\gamma_t$ hence $B(t) = C(t)\gamma_t^{-1}$ with corresponding boundary condition $C(T) = \gamma_T$ and derivative $B'(t) = C'(t)\gamma_t^{-1} + C(t)\Sigma^{-1}$.
If additionally C is symmetric then ODE (4.7) becomes

$$
\gamma_t^{-1} C'(t) \gamma_t^{-1} = -\frac{\alpha}{1-\alpha}\gamma_t^{-1} C(t)\Sigma^{-1} C(t)\gamma_t^{-1} - \gamma_t^{-1} C(t)\Sigma^{-1}(C(t)\gamma_t^{-1} - Id) - \gamma_t^{-1} C(t)\Sigma^{-1}
$$

$$\Rightarrow \quad C'(t) = -\frac{1}{1-\alpha} C(t)\Sigma^{-1}C(t)$$

To solve this remaining differential equation we use $D(t) := C^{-1}(t)$ with corresponding boundary condition $D(T) = \gamma_T^{-1} = \Sigma_0^{-1} + T\Sigma^{-1}$. Then $C'(t) = -D^{-1}(t)D'(t)D^{-1}(t)$ and ODE (4.7) becomes

$$-D^{-1}(t)D'(t)D^{-1}(t) = C'(t) = -\frac{1}{1-\alpha} D^{-1}(t)\Sigma^{-1}D^{-1}(t)$$

Hence $D'(t) = \frac{1}{1-\alpha}\Sigma^{-1}$ and the boundary condition leads to the resulting

$$\begin{aligned} D(t) &= \gamma_T^{-1} - \frac{1}{1-\alpha}(T-t)\Sigma^{-1} \\ \text{and} \quad C(t) &= \left(\gamma_T^{-1} - \frac{1}{1-\alpha}(T-t)\Sigma^{-1}\right)^{-1} \\ &= \left(\gamma_t^{-1} - \frac{\alpha}{1-\alpha}(T-t)\Sigma^{-1}\right)^{-1}. \end{aligned}$$

Indeed $C(t)$ is symmetric.

Now consider the second ordinary differential equation (4.8).

$$\begin{aligned} k'(t) &= -\alpha r - \frac{1}{2}\mathrm{tr}\left(A(t)\gamma_t\Sigma^{-1}\gamma_t\right) \quad , \quad k(T) = 0 \\ \Rightarrow \quad k(t) &= \alpha r(T-t) + \frac{1}{2}\int_t^T \mathrm{tr}\left((C(s)-\gamma_s)\Sigma^{-1}\right) ds \end{aligned}$$

To solve the remaining integral, note that the matrix logarithm only fulfils the following chain rule: $\partial_t \log\det M(t) = \mathrm{tr}\left(M(t)^{-1}\partial_t M(t)\right)$.
Applying this chain rule twice, with $M(s) := \gamma_s^{-1}$ respectively $M(s) := C(s)^{-1}$ we observe

$$\begin{aligned} \int_t^T \mathrm{tr}\left(\gamma_s\Sigma^{-1}\right) ds &= \int_t^T \mathrm{tr}\left(M(s)^{-1}\partial_s M(s)\right) ds \\ &= \int_t^T \partial_s \log\det M(s) ds \\ &= \log\det\gamma_T^{-1} - \log\det\gamma_t^{-1} \\ \int_t^T \mathrm{tr}\left(C(s)\Sigma^{-1}\right) ds &= (1-\alpha)\int_t^T \mathrm{tr}\left(C(s)\frac{1}{1-\alpha}\Sigma^{-1}\right) ds \\ &= (1-\alpha)\int_t^T \mathrm{tr}\left(M(s)^{-1}\partial_s M(s)\right) ds \\ &= (1-\alpha)\int_t^T \partial_s \log\det M(s) ds \\ &= (1-\alpha)\log\det C(T)^{-1} - (1-\alpha)\log\det C(t)^{-1} \end{aligned}$$

$$= (1-\alpha)\log\det\gamma_T^{-1} - (1-\alpha)\log\det C(t)^{-1}$$

Note that $\det\gamma_t^{-1}$ and $\det C(t)^{-1}$ are positive as γ_t and $C(t)$ are symmetric and positive definite. Hence the formulas are well-defined and result in the solution for $k(t)$.

$$\begin{aligned}
k(t) &= \alpha r(T-t) + \frac{1}{2}\int_t^T \operatorname{tr}\left((C(s)-\gamma_s)\Sigma^{-1}\right) ds \\
&= \alpha r(T-t) - \frac{1}{2}\log\det\gamma_T^{-1} + \frac{1}{2}\log\det\gamma_t^{-1} + \frac{1-\alpha}{2}\log\det\gamma_T^{-1} - \frac{1-\alpha}{2}\log\det C(t)^{-1} \\
&= \alpha r(T-t) - \frac{\alpha}{2}\log\det\gamma_T^{-1} + \frac{1}{2}\log\det\gamma_t^{-1} - \frac{1-\alpha}{2}\log\det\left(\gamma_t^{-1} - \frac{\alpha}{1-\alpha}(T-t)\Sigma^{-1}\right)
\end{aligned}$$

The candidate solutions of the stochastic control approach

Finally we are able to plug the derived value function $V = V(t,x,\mu)$ into the equation (4.6) for the optimal strategy at time t:

$$\begin{aligned}
\pi^*(t,x,\mu) &= -\frac{V_x}{xV_{xx}}\Sigma^{-1}(\mu - r\mathbb{1}) - \frac{1}{xV_{xx}}\Sigma^{-1}\gamma_t V_{x\mu} \\
&= \frac{1}{1-\alpha}\Sigma^{-1}(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}\Sigma^{-1}\gamma_t f_\mu \\
&= \frac{1}{1-\alpha}\Sigma^{-1}(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}\Sigma^{-1}\gamma_t A(t)(\mu - r\mathbb{1}) \\
&= \frac{1}{1-\alpha}\Sigma^{-1}B(t)(\mu - r\mathbb{1}) \\
&= \frac{1}{1-\alpha}\Sigma^{-1}C(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \\
&= \frac{1}{1-\alpha}\Sigma^{-1}\left(\gamma_T^{-1} - \frac{1}{1-\alpha}(T-t)\Sigma^{-1}\right)^{-1}\gamma_t^{-1}(\mu - r\mathbb{1})
\end{aligned}$$

Hence we get the optimal unconstrained portfolio strategy as presented in (4.2):

$$\begin{aligned}
\pi_t^{\text{unc}} &= \pi^*(t, X_t, \hat{\mu}_t) \\
&= \frac{1}{1-\alpha}\Sigma^{-1}C(t)\gamma_t^{-1}(\hat{\mu}_t - r\mathbb{1}) \\
&= \frac{1}{1-\alpha}\Sigma^{-1}\left(\gamma_T^{-1} - \frac{1}{1-\alpha}(T-t)\Sigma^{-1}\right)^{-1}\gamma_t^{-1}(\hat{\mu}_t - r\mathbb{1})
\end{aligned}$$

Remark 4.5
Note the similarity of the optimal unconstrained portfolio strategy to the optimal deterministic portfolio strategy derived in Section 2.1: $\pi_t^{\text{det}} = \pi_0^{\text{unc}}$ for all t.

It is also possible to use this stochastic control approach to derive the optimal portfolio strategy in the form of Section 4.1.1. This is shown in Appendix A.4.3.

Theorem 4.6 (verification)
Define the candidates $\bar{V}$ and $\bar{\pi}$ for value function and optimal control as derived in the previous stochastic control approach.

$$\bar{V}(t,x,\mu) := \frac{x^\alpha}{\alpha}\exp\left(\frac{1}{2}(\mu - r\mathbb{1})^\top A(t)(\mu - r\mathbb{1}) + k(t)\right)$$
$$\bar{\pi}(t,x,\mu) := \frac{1}{1-\alpha}\Sigma^{-1}C(t)\gamma_t^{-1}(\mu - r\mathbb{1})$$

where $A(t)$, $C(t)$ and $k(t)$ are given as above.

Then they solve the stochastic control approach (4.4) of our constrained portfolio optimization problem, hence $\bar{V} = V$ and $\bar{\pi} = \pi^*$ where

$$V(t,x,\mu) := \sup_{\pi\in\mathcal{A}(t,x,\mu)} \mathbb{E}\left[\frac{1}{\alpha}X_T^\alpha \,\middle|\, (X_t,\hat{\mu}_t) = (x,\mu)\right]$$
$$\text{and} \quad \pi^* := \arg\max_{\pi\in\mathcal{A}} \mathbb{E}\left[\frac{1}{\alpha}X_T^\alpha\right]$$

Proof. Let $(t,x,\mu) \in [0,T) \times \mathbb{R}^+ \times \mathbb{R}^d$, $\pi \in \mathcal{A}$ and $s \in [t,T)$ be arbitrary and let τ be an arbitrary stopping time with values in $[t,T]$.
As $\bar{V} \in C^{1,2,2}$ we get with the Itô formula

$$\begin{aligned}\bar{V}(s\wedge\tau, X_{s\wedge\tau}, \hat{\mu}_{s\wedge\tau}) &= \bar{V}(t,x,\mu) + \int_t^{s\wedge\tau} \bar{V}_t(u) + \mathcal{L}^\pi\bar{V}(u)du \\ &\quad + \int_t^{s\wedge\tau} \bar{V}_x(u)X_u\pi_u^\top\sigma + \bar{V}_\mu(u)^\top\gamma_u(\sigma^{-1})^\top dV_u \qquad (4.9)\end{aligned}$$
$$\begin{aligned}\text{where} \quad \mathcal{L}^\pi\bar{V}(u) &= \bar{V}_x(u)X_u\left(r + \pi_u^\top(\hat{\mu}_u - r\mathbb{1})\right) + \frac{1}{2}\bar{V}_{xx}(u)X_u\pi_u^\top\Sigma\pi_uX_u \\ &\quad + \bar{V}_{x\mu}^\top(u)\gamma_u\pi_uX_u + \frac{1}{2}\mathrm{tr}\big(\bar{V}_{\mu\mu}(u)\gamma_u\Sigma^{-1}\gamma_u\big)\end{aligned}$$

where we use the notation $\bar{V}(u) := \bar{V}(u,X_u,\hat{\mu}_u)$. Here the sub-indices u indicate time-dependency of the corresponding variables, the other sub-indices of $\bar{V}$ are the respective derivatives of $\bar{V}(t,x,\mu)$.
Now for each $n \in \mathbb{N}$ define the stopping time

$$\tau_n := \inf\left\{s \geq t \,\middle|\, \int_t^s \left\|\sigma^{-1}\gamma_u\bar{V}_\mu(u)\right\|_2^2 du \geq n \quad \text{or} \quad \int_t^s \left\|\bar{V}_x(u)X_u\sigma^\top\pi_u\right\|_2^2 du \geq n\right\} \wedge T$$

Hence $\tau_n \to T$ as $n \to \infty$ and therefore for any $n \in \mathbb{N}$ the stopped process

$$\int_t^{s\wedge\tau_n} \bar{V}_x(u)X_u\pi_u^\top\sigma + \bar{V}_\mu(u)^\top\gamma_u(\sigma^{-1})^\top dV_u$$

is a martingale with zero expectation.
Now we use τ_n in (4.9) and take conditional expectations on both sides.

$$\mathbb{E}\left[\bar{V}(s\wedge\tau_n, X_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\,\middle|\, (X_t,\hat{\mu}_t) = (x,\mu)\right]$$

$$= \bar{V}(t,x,\mu) + \mathbb{E}\left[\int_t^{s\wedge\tau_n} \bar{V}_t(u,X_u,\hat{\mu}_u) + \mathcal{L}^\pi \bar{V}(u,X_u,\hat{\mu}_u) du \middle| (X_t,\hat{\mu}_t) = (x,\mu) \right]$$

$\bar{V}(t,x,\mu)$ satisfies the HJB-equation (4.5), hence we observe for the right hand side:

$$\begin{aligned} \bar{V}_t(u,X_u,\hat{\mu}_u) + \mathcal{L}^\pi \bar{V}(u,X_u,\hat{\mu}_u) &\leq 0 \qquad \text{for all } \pi \in \mathcal{A} \\ \bar{V}_t(u,X_u,\hat{\mu}_u) + \mathcal{L}^\pi \bar{V}(u,X_u,\hat{\mu}_u) &= 0 \qquad \text{if } \pi = \pi^* \end{aligned}$$

This leads for each $u \in [t,T]$ and $\pi \in \mathcal{A}(t,x,\mu)$ to the following equations:

$$\begin{aligned} \mathbb{E}\left[\bar{V}(s\wedge\tau_n, X_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n}) \middle| (X_t,\hat{\mu}_t) = (x,\mu) \right] &\leq \bar{V}(t,x,\mu) \qquad \text{for all } \pi \in \mathcal{A}(t,x,\mu) \\ \mathbb{E}\left[\bar{V}(s\wedge\tau_n, X_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n}) \middle| (X_t,\hat{\mu}_t) = (x,\mu) \right] &= \bar{V}(t,x,\mu) \qquad \text{if } \pi = \pi^* \end{aligned}$$

By Lemma 4.7 we can apply dominated convergence and observe for $n \to \infty$

$$\begin{aligned} \bar{V}(t,x,\mu) &\geq \mathbb{E}\left[\bar{V}(s,X_s,\hat{\mu}_s) \middle| (X_t,\hat{\mu}_t) = (x,\mu) \right] \qquad \text{for all } \pi \in \mathcal{A}(t,x,\mu) \\ \bar{V}(t,x,\mu) &= \mathbb{E}\left[\bar{V}(s,X_s,\hat{\mu}_s) \middle| (X_t,\hat{\mu}_t) = (x,\mu) \right] \qquad \text{if } \pi = \pi^* \end{aligned}$$

Since $\bar{V}(s,X_s,\hat{\mu}_s) \to \bar{V}(T,X_T,\hat{\mu}_T) = \frac{1}{\alpha}(X_T)^\alpha$ for $s \to T$ we get by dominated convergence

$$\begin{aligned} & \bar{V}(t,x,\mu) \geq \mathbb{E}\left[\frac{1}{\alpha}(X_T)^\alpha \middle| (X_t,\hat{\mu}_t) = (x,\mu) \right] \qquad \text{for all } \pi \in \mathcal{A}(t,x,\mu) \\ \Rightarrow \quad & \bar{V}(t,x,\mu) \geq \sup_{\pi\in\mathcal{A}(t,x,\mu)} \mathbb{E}\left[\frac{1}{\alpha}(X_T)^\alpha \middle| (X_t,\hat{\mu}_t) = (x,\mu) \right] = V(t,x,\mu) \end{aligned}$$

Since we get equality for $\pi = \bar{\pi}$ we get $V(t,x,\mu) = \bar{V}(t,x,\mu)$ with optimizer $\pi^* = \bar{\pi}$. □

Lemma 4.7
There is an integrable random variable Y with $\left|\bar{V}(s\wedge\tau_n, X_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\right| \leq Y$.

Proof. The function f is a second order polynomial in μ, hence

$$f(t,\mu) \leq C_t(1 + \|\mu\|_2^2)$$

where the leading constant C_t may depend on time t. However $t \in [0,T]$ is in a compact interval, hence $f(\cdot,\mu)$ is uniformly continuous in t and attains its supremum as a maximum.

$$f(t,\mu) \leq C(1 + \|\mu\|_2^2)$$

for some upper bound $C \geq \max_{t\in[0,T]} C_t < \infty$. Therefore

$$\begin{aligned} & \left|\bar{V}(t,x,\mu)\right| = -\frac{1}{\alpha}x^\alpha e^{f(t,\mu)} \leq -\frac{1}{\alpha}x^\alpha e^{C(1+\|\mu\|_2^2)} \\ \Rightarrow \quad & \left|\bar{V}(s\wedge\tau_n, X_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\right| \leq -\frac{1}{\alpha} X^\alpha_{s\wedge\tau_n} \exp\left(C(1 + \|\hat{\mu}_{s\wedge\tau_n}\|_2^2) \right) \\ & \qquad \leq -\frac{1}{\alpha} \sup_{s\in[t,T]} X_s^\alpha \exp\left(C(1 + \|\hat{\mu}_s\|_2^2) \right) \end{aligned}$$

$$\le -\frac{1}{\alpha} \sup_{s\in[t,T]} X_s^\alpha \cdot \sup_{s\in[t,T]} \exp\Big(C(1+\|\hat{\mu}_s\|_2^2)\Big) =: Y$$

Now on the one hand $X_s > 0$ is uniformly continuous in $[t,T]$, hence $M := \inf_{s\in[t,T]} X_s > 0$ and $\sup_{s\in[t,T]} X_s^\alpha \le M^\alpha < \infty$.
And on the other hand $\hat{\mu}_t$ is a martingale, hence $\|\hat{\mu}_s\|_2^2$ is a submartingale since $\|\cdot\|_2^2$ is a convex function. Hence $\exp\Big(C(1+\|\hat{\mu}_s\|_2^2)\Big)$ is also a submartingale and by Doob's maximal inequality we get

$$\mathbb{E}\left[\sup_{s\in[t,T]} \exp\Big(C(1+\|\hat{\mu}_s\|_2^2)\Big)\right] \le \left(\frac{C}{C-1}\right)^C \mathbb{E}\left[\exp\Big(C(1+\|\hat{\mu}_T\|_2^2)\Big)\right] < \infty$$

Hence $\mathbb{E}\left[Y\right] < \infty$. □

4.2 The convex dual problem: a reverse separation approach

The main point of Chapter 4 is to solve our constrained portfolio optimization problem (2.3) under partial information as introduced in Section 2.2 and under the convex constraints (3.7).
According to Section 2.3 we may equivalently solve the unconstrained portfolio optimization problem in the auxiliary market $\mathcal{M}^\lambda$ (2.16) where the dual process λ is chosen according to Theorem 2.11:

$$\begin{aligned} dB_t^\lambda &= B_t^\lambda\big(r + \delta\left(\lambda_t\right)\big)dt \\ dS_t^\lambda &= \operatorname{diag}\left(S_t^\lambda\right)\big(\left(\hat{\mu}_t + \lambda_t + \delta(\lambda_t)\mathbb{1}\right)dt + \sigma dV_t\big) \\ \text{and} \quad \lambda &= \operatorname*{arg\,min}_{\nu\in\mathcal{D}} \mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^\lambda(x_0)H_T^\nu\big)\right] \end{aligned} \tag{4.10}$$

In our setting we get the following components:

$$\begin{aligned} I(y) = U'^{-1}(y) \quad &= y^{\frac{1}{\alpha-1}} \\ \tilde{U}(y) = U(I(y)) - yI(y) &= \frac{1-\alpha}{\alpha} y^{\frac{\alpha}{\alpha-1}} \\ \mathcal{X}^\lambda(y) = \mathbb{E}\left[H_T^\lambda I\big(yH_T^\lambda\big)\right] &= y^{\frac{1}{\alpha-1}}\mathbb{E}\left[\big(H_T^\lambda\big)^{\frac{\alpha}{\alpha-1}}\right] \\ \mathcal{Y}^\lambda(x) = \big(\mathcal{X}^\lambda\big)^{-1}(x) \quad &= x^{\alpha-1}\mathbb{E}\left[\big(H_T^\lambda\big)^{\frac{\alpha}{\alpha-1}}\right]^{1-\alpha} \end{aligned}$$

and $dH_t^\nu = -H_t^\nu\Big((r+\delta(\nu_t))dt + \hat{\theta}_t^{\nu\top}dV_t\Big)$ with $\hat{\theta}_t^\nu = \sigma^{-1}(\hat{\mu}_t - r\mathbb{1} + \nu_t)$ is given in (2.17).
In this section we solve the convex dual problem (4.10) for the optimal dual process λ:

$$\begin{aligned} \lambda &= \operatorname*{arg\,min}_{\nu\in\mathcal{D}} \mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^\lambda(x_0)H_T^\nu\big)\right] \\ &= \operatorname*{arg\,min}_{\nu\in\mathcal{D}} \frac{1-\alpha}{\alpha}\mathbb{E}\left[\big(\mathcal{Y}^\lambda(x_0)H_T^\nu\big)^{\frac{\alpha}{\alpha-1}}\right] \\ &= \operatorname*{arg\,min}_{\nu\in\mathcal{D}} \frac{1-\alpha}{\alpha}x_0^\alpha\mathbb{E}\left[\mathbb{E}\left[(H_T^\lambda)^{\frac{\alpha}{\alpha-1}}\right]^{-\alpha}(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\right] \\ &= \operatorname*{arg\,min}_{\nu\in\mathcal{D}} \frac{1-\alpha}{\alpha}x_0^\alpha\mathbb{E}\left[(H_T^\lambda)^{\frac{\alpha}{\alpha-1}}\right]^{-\alpha}\mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\right] \\ &= \operatorname*{arg\,min}_{\nu\in\mathcal{D}} \frac{1-\alpha}{\alpha}\mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\right] \qquad \text{since } x_0>0 \text{ and } H_T^\lambda>0 \\ &= \operatorname*{arg\,max}_{\nu\in\mathcal{D}} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\right] \qquad \text{since } \alpha<0 \end{aligned} \tag{4.11}$$

Unfortunately the dual problem does not simplify as much as under logarithmic utility. The remaining dual problem (4.11) can be solved using a stochastic control approach with value function

$$V(t,h,\mu) := \sup_{\nu\in\mathcal{D}(t,h,\mu)} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\,\middle|\,(H_t^\nu,\hat{\mu}_t)=(h,\mu)\right] \tag{4.12}$$

Here $\mathcal{D}(t,h,\mu)$ is the set of admissible control processes $\nu \in \mathcal{D}$ such that $(H_t^\nu, \hat{\mu}_t) = (h,\mu)$. The stochastic control problem is to determine $V(0,1,\mu_0)$ and the optimal dual process λ such that

$$V(0,1,\mu_0) = \sup_{\nu \in \mathcal{D}(0,1,\mu_0)} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}} \middle| (H_0^\nu, \hat{\mu}_0) = (1,\mu_0) \right] = \mathbb{E}\left[(H_T^\lambda)^{\frac{\alpha}{\alpha-1}} \right]$$

In order to make the controlled process Markov, we also need the dynamics of the filter.

$$d\begin{pmatrix} H_t^\nu \\ \hat{\mu}_t \end{pmatrix} = \begin{pmatrix} -H_t^\nu(r+\delta(\nu_t)) \\ 0 \end{pmatrix} dt + \begin{pmatrix} -H_t^\nu(\hat{\mu}_t - r\mathbb{1} + \nu_t)^\top (\sigma^{-1})^\top \\ \gamma_t (\sigma^{-1})^\top \end{pmatrix} dV_t$$

If the Bellmann principle holds for $t_1 > t$ we get

$$V(t,h,\mu) = \sup_{\nu \in \mathcal{D}(t,h,\mu)} \mathbb{E}\left[V\big(t_1, H_{t_1}^\nu, \hat{\mu}_{t_1}\big) \middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu) \right]$$

With Itô's formula we get

$$\begin{aligned} V(t,h,\mu) = \sup_{\nu \in \mathcal{D}(t,h,\mu)} \mathbb{E}\Bigg[& V(t) + \int_t^{t_1} V_t(s) - V_h(s) H_s^\nu (r+\delta(\nu_s)) ds \\ & + \int_t^{t_1} V_\mu(s)^\top \gamma_s (\sigma^{-1})^\top - V_h(s) H_s^\nu (\hat{\mu}_s - r\mathbb{1} + \nu_s)^\top (\sigma^{-1})^\top dV_s \\ & + \int_t^{t_1} \frac{1}{2} \mathrm{tr}(HV(s)\cdot a(s)) ds \Bigg| (H_t^\nu, \hat{\mu}_t) = (h,\mu) \Bigg] \end{aligned}$$

where we use the notation $V(s) := V(s, H_s^\nu, \hat{\mu}_s)$ with derivatives V_t, V_h, V_μ and Hessian matrix HV. $a(s)$ is called the diffusion matrix.

$$\begin{aligned} a(s) &:= \begin{pmatrix} -H_s^\nu(\hat{\mu}_s - r\mathbb{1} + \nu_s)^\top (\sigma^{-1})^\top \\ \gamma_s(\sigma^{-1})^\top \end{pmatrix} \begin{pmatrix} -H_s^\nu(\hat{\mu}_s - r\mathbb{1} + \nu_s)^\top (\sigma^{-1})^\top \\ \gamma_s(\sigma^{-1})^\top \end{pmatrix}^\top \\ &= \begin{pmatrix} (H_s^\nu)^2(\hat{\mu}_s - r\mathbb{1} + \nu_s)^\top \Sigma^{-1} (\hat{\mu}_s - r\mathbb{1} + \nu_s) & -H_s^\nu (\hat{\mu}_s - r\mathbb{1} + \nu_s)^\top \Sigma^{-1} \gamma_s \\ -H_s^\nu \gamma_s \Sigma^{-1} (\hat{\mu}_s - r\mathbb{1} + \nu_s) & \gamma_s \Sigma^{-1} \gamma_s \end{pmatrix} \\ \mathrm{tr}(HV(s)\cdot a(s)) &= V_{hh}(s)\cdot a_{11}(s) + V_{h\mu}(s)^\top \cdot a_{21}(s) + \mathrm{tr}(V_{\mu h}(s)\cdot a_{12}(s) + V_{\mu\mu}(s)\cdot a_{22}(s)) \\ &= V_{hh}(s)\cdot a_{11}(s) + 2a_{12}(s) V_{h\mu}(s) + \mathrm{tr}(V_{\mu\mu}(s)\cdot a_{22}(s)) \\ &= (H_s^\nu)^2 (\hat{\mu}_s - r\mathbb{1} + \nu_s)^\top \Sigma^{-1} (\hat{\mu}_s - r\mathbb{1} + \nu_s) V_{hh}(s) \\ &\quad - 2H_s^\nu (\hat{\mu}_s - r\mathbb{1} + \nu_s)^\top \Sigma^{-1} \gamma_s V_{h\mu}(s) + \mathrm{tr}(V_{\mu\mu}(s)\cdot \gamma_s \Sigma^{-1} \gamma_s) \end{aligned}$$

If the usual suitable conditions hold, we observe the HJB-equation (4.13) for any fixed t. Of course these conditions have to be verified in the verification later on.

$$\begin{aligned} 0 = \sup_{\nu \in \mathcal{D}(t,h,\mu)} \mathbb{E}\Bigg[& \int_t^{t_1} V_t(s) - V_h(s) H_s^\nu (r+\delta(\nu_s)) ds \\ & + \int_t^{t_1} \frac{1}{2} (H_s^\nu)^2 (\hat{\mu}_s - r\mathbb{1} + \nu_s)^\top \Sigma^{-1} (\hat{\mu}_s - r\mathbb{1} + \nu_s) V_{hh}(s) ds \end{aligned}$$

$$- \int_t^{t_1} H_s^\nu (\hat{\mu}_s - r\mathbb{1} + \nu_s)^\top \Sigma^{-1} \gamma_s V_{h\mu}(s) ds$$
$$\left. + \int_t^{t_1} \frac{1}{2} \mathrm{tr}(V_{\mu\mu}(s) \cdot \gamma_s \Sigma^{-1} \gamma_s) ds \right| (H_t^\nu, \hat{\mu}_t) = (h, \mu) \Bigg]$$
$$\Rightarrow \quad 0 = \sup_{\nu \in \tilde{K}} \Bigg(V_t - V_h h(r + \delta(\nu)) + \frac{1}{2} h^2 (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} (\mu - r\mathbb{1} + \nu) V_{hh}$$
$$- h(\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} \gamma_t V_{h\mu} + \frac{1}{2} \mathrm{tr}(V_{\mu\mu} \cdot \gamma_t \Sigma^{-1} \gamma_t) \Bigg) \tag{4.13}$$

The HJB-equation (4.13) has to be solved with respect to the boundary condition

$$V(T, h, \mu) = \sup_{\nu \in \mathcal{D}(T,h,\mu)} \mathbb{E} \left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}} \middle| (H_T^\nu, \hat{\mu}_T) = (h, \mu) \right] = h^{\frac{\alpha}{\alpha-1}}$$

Then the optimal control λ at time t gets chosen by:

$$\lambda = \arg\max_{\nu \in \tilde{K} = \mathbb{R}^d} \Bigg(- V_h \delta(\nu) + \frac{1}{2} h (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} (\mu - r\mathbb{1} + \nu) V_{hh} - (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} \gamma_t V_{h\mu} \Bigg) \tag{4.14}$$

The reverse separation approach

Unfortunately we cannot expect δ to be differentiable at the optimal λ as we have motivated in Remark 2.10. Therefore we need to consider the following reverse separation approach:

> We split the domain $\tilde{K}$ of the dual optimization parameter ν into several disjoint regions, such that in each region $\delta(\nu)$ is differentiable *('separation')*. These regions are defined with respect to the yet unknown optimal control, the resulting dual process λ *('reverse')*.

We will call these regions 'Cases', indicating that in each case λ is in one of these regions. Consider Example 4.44 and in particular the left hand side of Figure 4.1 to get an idea of how this separation can look like for $\tilde{K} = \mathbb{R}^2$. In order to be able to use this reverse separation approach we need the following assumption for the moment. We will stick to this assumption for the following two sections.

Assumption 4.8
Assume that the optimal dual control process λ stays in the same case for all t.

Remark 4.9 (Interpretation)
Note that this assumption has to be understood as a simplification with two purposes. On the one hand it simplifies notation and on the other hand it allows us to solve the different cases separately.

Obviously the optimal dual process will almost surely not stay in whatever current case it starts in. However the main idea behind the stochastic control approach is to derive an HJB-equation that solves the optimization problem locally. But locally we will always stay in whatever current case we start in, hence locally Assumption 4.8 actually holds. And therefore given this assumption we are able to solve our optimization problem. Additionally note that we will see in section 4.4 that we are almost surely always in the interior of some region such that there actually is some local neighbourhood that is completely contained in the region of the current case and hence the HJB-equation is well defined. Again consider Figure 4.1 to get an idea of how this separation can look like.

Remark 4.10 (Interpretation of the value function)
Another important point to be interpreted are the resulting (primal or dual) value functions. By definition, the (primal) value function is the expected utility of terminal wealth given some current time and values of the state processes X and $\hat{\mu}$. However the value functions resulting from any separated case of the reverse separation approach are *purely theoretical tools* to derive the locally optimal solutions of the respective case.

The value of one of these separated value functions on the other hand would be the expected utility of terminal wealth if the investor was to invest according to the rules of the starting case but for the whole investment period even when the parameters leave this current case. This could be made explicit after determining the correct cases in Section 4.4 by stopping at the boundaries of the corresponding region, say with stopping time τ, and using a boundary condition $V(\tau, \cdot, \cdot)$ computed from continuing with the same strategy and the same δ as before up to terminal time T.

If one wanted to derive the complete value function of the whole optimization problem, one would most likely have to consider another modelling approach than ours. However we think that this value function has to be some weighted mean of our separated value functions, weighted by the probabilities of how much the optimal solution is in the different cases.

Under the no short-selling and no borrowing constraints with $\delta(\nu) = \max_j\{0, -\nu_j\}$, we use the following reverse separation approach on the optimal dual process λ:

Case 1) $\lambda = 0$.
This case is equivalent to $\pi^\lambda = \pi_{\text{unc}} \in K$. Hence the optimal unconstrained portfolio strategy is already admissible.

Case 2) $\lambda \in \mathbb{R}^d_{>0}$
Then $\delta(\lambda) = 0$ and we will show below that $\pi^\lambda = 0$ as with logarithmic utility.

Case 3) $\lambda \in \mathbb{R}^d_{\geq 0} \setminus \left(\mathbb{R}^d_{>0} \cup \{0\}\right)$.
Then $\delta(\lambda) = 0$ and we will see later on that this would be the relevant case to consider if we were to only consider the no short-selling constraint.

Case 4) $\lambda \in \mathbb{R}^d \setminus \mathbb{R}^d_{\geq 0}$ (equivalent to $\min_i \lambda_i < 0$)
Then $\delta(\lambda) = -\min_i \lambda_i$ is still not differentiable, but we will show that δ can be made differentiable by an easy transformation. Also we will see later on that this will be the relevant case to consider for most real-life situations.

Note that these cases are the same cases that we already derived under logarithmic utility in Theorem 3.14. This is due to the observation that the underlying structure of λ mainly depends on the admissibility set K and not on the utility function U.

Remark 4.11
In Case 1 there obviously is nothing left to solve. In particular this is a case that is detectable a priori as we only need to calculate the optimal unconstrained portfolio strategy. Case 2 on the other hand is detectable a priori, too, as the following theorem shows.

Proposition 4.12 (Case 2)
The optimal dual λ satisfies $\lambda > 0$ if and only if $\mu - r\mathbb{1} < 0$.
In that case $\lambda = -(\mu - r\mathbb{1})$ and $\pi^\lambda = 0$.

Proof. "$\Rightarrow$"
As $\delta(\nu) = 0$ for $\nu \in (0, \infty)^d$ the HJB-equation (4.13) becomes:

$$0 = \sup_{\nu \in \mathbb{R}^d} \Big(V_t - V_h hr + \frac{1}{2} h^2 (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} (\mu - r\mathbb{1} + \nu) V_{hh} - h(\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} \gamma_t V_{h\mu} + \frac{1}{2} \mathrm{tr}(V_{\mu\mu} \cdot \gamma_t \Sigma^{-1} \gamma_t) \Big)$$

$$\lambda = \arg\max_{\nu \in \mathbb{R}^d} \Big(\frac{1}{2} h^2 (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} (\mu - r\mathbb{1} + \nu) V_{hh} - h(\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} \gamma_t V_{h\mu} \Big)$$

with boundary condition $V(T, h, \mu) = h^{\frac{\alpha}{\alpha-1}}$. Differentiation leads to the optimal dual λ:

$$0 \overset{!}{=} h\Sigma^{-1}(\mu - r\mathbb{1} + \lambda) V_{hh} - \Sigma^{-1} \gamma_t V_{h\mu}$$
$$\Rightarrow \quad \lambda = -(\mu - r\mathbb{1}) + \gamma_t \frac{V_{h\mu}}{h V_{hh}}$$

Now the HJB-equation becomes

$$0 = V_t - V_h hr - \frac{1}{2} V_{h\mu}^\top \gamma_t \Sigma^{-1} \gamma_t \frac{V_{h\mu}}{V_{hh}} + \frac{1}{2} \mathrm{tr}(V_{\mu\mu} \cdot \gamma_t \Sigma^{-1} \gamma_t) \quad , \quad V(T, h, \mu) = h^{\frac{\alpha}{\alpha-1}}$$

This differential equation gets solved easily as follows:

$$V(t, h, \mu) = h^{\frac{\alpha}{\alpha-1}} \exp\left(\frac{\alpha}{1-\alpha} r(T-t) \right)$$
$$\Rightarrow \quad \lambda = -(\mu - r\mathbb{1}) + \gamma_t \frac{V_{h\mu}}{h V_{hh}} = -(\mu - r\mathbb{1}) \tag{4.15}$$

However solving the auxiliary market $\mathcal{M}^\lambda$ (2.16) for $\lambda = -(\mu - r\mathbb{1})$ is trivial and leads to $\pi^\lambda = 0$. In particular as $\lambda > 0$ we conclude $\mu - r\mathbb{1} = -\lambda < 0$.
"$\Leftarrow$"
If vice versa $\mu - r\mathbb{1} < 0$ then define $\lambda := -(\mu - r\mathbb{1}) > 0$. Hence $\delta(\lambda) = 0$ and the function $V(t, h, \mu) = h^{\frac{\alpha}{\alpha-1}} \exp\left(\frac{\alpha}{1-\alpha} r(T-t)\right)$ solves the optimization problem given above with optimal control λ. In particular we get $\pi^\lambda = 0$.
Note that the verification in both cases is trivially given by Theorem 4.13. □

Theorem 4.13 (Verification in Case 2)
Define the candidates for value function and optimal control as derived in Proposition 4.12:

$$\bar{V}(t,h,\mu) := h^{\frac{\alpha}{\alpha-1}} \exp\left(\frac{\alpha}{1-\alpha} r(T-t)\right)$$
$$\bar{\lambda}(t,h,\mu) := -(\mu - r\mathbb{1})$$

Then they solve the stochastic control approach (4.12) in Case 2, hence $\bar{V} = V$ and $\bar{\lambda} = \nu^*$, where

$$V(t,h,\mu) := \sup_{\nu \in \mathcal{D}(t,h,\mu)} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}} \middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu) \right]$$
$$\text{and} \quad \nu^* := \arg\max_{\nu \in \mathcal{D}} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}} \right]$$

Proof. Let $(t,h,\mu) \in [0,T) \times \mathbb{R}^+ \times \mathbb{R}^d$, $\nu \in \mathcal{D}$ and $s \in [t,T)$ be arbitrary and let τ be an arbitrary stopping time with values in $[t,T]$.
As $\bar{V} \in C^{1,2,2}$ we get with Itô's formula

$$\bar{V}(s\wedge\tau, H^\nu_{s\wedge\tau}, \hat{\mu}_{s\wedge\tau}) = \bar{V}(t,h,\mu) + \int_t^{s\wedge\tau} \bar{V}_t(u) + \mathcal{L}^\nu \bar{V}(u) du \qquad (4.16)$$
$$- \int_t^{s\wedge\tau} \bar{V}_h(u) H_u^\nu (\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top (\sigma^{-1})^\top dV_u$$
$$\text{where} \quad \mathcal{L}^\nu \bar{V}(u) = -\bar{V}_h(u) H_u^\nu \left(r + \delta(\nu_u)\right) + \frac{1}{2}(H_u^\nu)^2 (\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top \Sigma^{-1} (\hat{\mu}_u - r\mathbb{1} + \nu_u) V_{hh}(u)$$

Here we use the notation $\bar{V}(u) := \bar{V}(u, H_u^\nu, \hat{\mu}_u)$. While most sub-indices indicate time-dependency of the corresponding variables, sub-indices of $\bar{V}$ are the respective derivatives of $\bar{V}(t,h,\mu)$.
Now for each $n \in \mathbb{N}$ define the stopping time

$$\tau_n := \inf\left\{ s \geq t \,\middle|\, \int_t^s \left\| \bar{V}_h(u) H_u^\nu (\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top (\sigma^{-1})^\top \right\|_2^2 du \geq n \right\} \wedge T$$

Hence $\tau_n \to T$ as $n \to \infty$ and therefore for any $n \in \mathbb{N}$ the stopped process

$$\int_t^{s\wedge\tau} \bar{V}_h(u) H_u^\nu (\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top (\sigma^{-1})^\top dV_u$$

is a martingale with zero expectation.
Now we use τ_n in (4.16) and take conditional expectations on both sides.

$$\mathbb{E}\left[\bar{V}(s\wedge\tau_n, H^\nu_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n}) \middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu) \right]$$
$$= \bar{V}(t,h,\mu) + \mathbb{E}\left[\int_t^{s\wedge\tau_n} \bar{V}_t(u, H_u^\nu, \hat{\mu}_u) + \mathcal{L}^\nu \bar{V}(u, H_u^\nu, \hat{\mu}_u) du \middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu) \right]$$

$\bar{V}(t,h,\mu)$ satisfies the HJB-equation (4.13), hence we observe for the right hand side:

$$\begin{aligned} \bar{V}_t(u,H_u^\nu,\hat{\mu}_u) + \mathcal{L}^\nu \bar{V}(u,H_u^\nu,\hat{\mu}_u) &\leq 0 \qquad \text{for all } \nu \in \mathcal{D} \\ \bar{V}_t(u,H_u^\nu,\hat{\mu}_u) + \mathcal{L}^\nu \bar{V}(u,H_u^\nu,\hat{\mu}_u) &= 0 \qquad \text{if } \nu = \bar{\lambda} \end{aligned}$$

This leads for each $u \in [t,T]$ and $\nu \in \mathcal{D}(t,h,\mu)$ to the following equations:

$$\begin{aligned} \mathbb{E}\left[\bar{V}(s\wedge\tau_n, H^\nu_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\middle| (H_t^\nu,\hat{\mu}_t) = (h,\mu)\right] &\leq \bar{V}(t,h,\mu) \qquad \text{for all } \nu \in \mathcal{D}(t,h,\mu) \\ \mathbb{E}\left[\bar{V}(s\wedge\tau_n, H^\nu_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\middle| (H_t^\nu,\hat{\mu}_t) = (h,\mu)\right] &= \bar{V}(t,h,\mu) \qquad \text{if } \nu = \bar{\lambda} \end{aligned}$$

By Lemma 4.14 we can apply dominated convergence and observe for $n \to \infty$

$$\begin{aligned} \bar{V}(t,h,\mu) &\geq \mathbb{E}\left[\bar{V}(s,H_s^\nu,\hat{\mu}_s)\middle| (H_t^\nu,\hat{\mu}_t) = (h,\mu)\right] \qquad \text{for all } \nu \in \mathcal{D}(t,h,\mu) \\ \bar{V}(t,h,\mu) &= \mathbb{E}\left[\bar{V}(s,H_s^\nu,\hat{\mu}_s)\middle| (H_t^\nu,\hat{\mu}_t) = (h,\mu)\right] \qquad \text{if } \nu = \bar{\lambda} \end{aligned}$$

Since $\bar{V}(s,H_s^\nu,\hat{\mu}_s) \to \bar{V}(T,H_T^\nu,\hat{\mu}_T) = (H_T^\nu)^{\frac{\alpha}{\alpha-1}}$ for $s \to T$ we conclude by dominated convergence

$$\begin{aligned} & \bar{V}(t,h,\mu) \geq \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\middle| (H_t^\nu,\hat{\mu}_t) = (h,\mu)\right] \qquad \text{for all } \nu \in \mathcal{D}(t,h,\mu) \\ \Rightarrow \quad & \bar{V}(t,h,\mu) \geq \sup_{\nu\in\mathcal{D}(t,h,\mu)} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\middle| (H_t^\nu,\hat{\mu}_t) = (h,\mu)\right] = V(t,h,\mu) \end{aligned}$$

Since we get equality for $\nu = \bar{\lambda}$ we get $V(t,h,\mu) = \bar{V}(t,h,\mu)$ with optimizer $\nu^* = \bar{\lambda}$. □

Lemma 4.14
There is an integrable random variable Y with $\left|\bar{V}(s\wedge\tau_n, H^\nu_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\right| \leq Y$.

Proof. Since $\frac{\alpha}{\alpha-1} \in (0,1)$ and $\frac{\alpha}{1-\alpha} < 0$ we observe

$$\begin{aligned} & h^{\frac{\alpha}{\alpha-1}} \leq \begin{cases} 1 & \text{if } h \leq 1 \\ h & \text{if } h > 1 \end{cases} \\ \Rightarrow \quad & 0 < h^{\frac{\alpha}{\alpha-1}} \leq (1+h) \\ \text{and} \quad & \exp\left(\frac{\alpha}{1-\alpha} r(T-t)\right) \leq 1 \end{aligned}$$

Hence $\left|\bar{V}(t,h,\mu)\right| = h^{\frac{\alpha}{\alpha-1}} \exp\left(\frac{\alpha}{1-\alpha} r(T-t)\right) \leq 1+h$ and

$$\begin{aligned} \left|\bar{V}(s\wedge\tau_n, H^\nu_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\right| &\leq 1 + H^\nu_{s\wedge\tau_n} \\ &\leq 1 + M =: Y \end{aligned}$$

where the last line holds with $\mathbb{E}[Y] < \infty$ because $H_u^\nu > 0$ is uniformly continuous in $[t,T]$ and hence attains its supremum $M := \sup_{s\in[t,T]} H_u^\nu < \infty$. □

4.2.1 The dual problem in Case 3

In the next step we want to solve the dual problem in Case 3. Note that together with the trivial Cases 1 and 2 this case would be the correct reverse separation approach when only considering the no-short-selling constraint. This follows in particular from the observations on the domain of the value function in Section 4.4. Recall Case 3:

$$\begin{aligned} & \lambda \in \mathbb{R}^d_{\geq 0} \setminus \left(\mathbb{R}^d_{>0} \cup \{0\}\right) \\ \text{hence} \quad & \lambda_i = 0 \quad \text{for some dimensions } i \\ & \lambda_j > 0 \quad \text{for some dimensions } j \end{aligned}$$

During this section we will allow without loss of generality $\lambda \in \mathbb{R}^d_{\geq 0}$. This simplification works because we know a priori if we were in the Cases 1 or 2 and then we would not consider Case 3 anyway. On the other hand if we are in Case 3 the optimal result cannot be in the regions of Case 1 or 2 as shown in Proposition 4.12 and Remark 4.11. Also the optimal result that we will derive in the following is still in $\mathbb{R}^d_{\geq 0} \setminus \left(\mathbb{R}^d_{>0} \cup \{0\}\right)$.

In Case 3 we observe in particular $\delta(\lambda) = 0$ on $\mathbb{R}^d_{\geq 0}$. However we still cannot differentiate to find the optimal solution for λ, since differentiability would only be given on the interior $\mathbb{R}^d_{>0}$, which is the region of Case 2, that we don't want to consider here. Hence we need a different approach, that widens the dual domain to the whole of $\mathbb{R}^d$.

First we reconsider the HJB-equation (4.13):

$$\begin{aligned} 0 = \sup_{\nu \in \mathbb{R}^d_{\geq 0}} \Bigg(& V_t - V_h h r + \frac{1}{2} h^2 (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} (\mu - r\mathbb{1} + \nu) V_{hh} \\ & - h(\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} \gamma_t V_{h\mu} + \frac{1}{2} \operatorname{tr}(V_{\mu\mu} \cdot \gamma_t \Sigma^{-1} \gamma_t) \Bigg) \end{aligned}$$

with boundary condition $V(T, h, \mu) = h^{\frac{\alpha}{\alpha - 1}}$. The optimal control λ gets chosen by (4.14):

$$\lambda = \operatorname*{arg\,max}_{\nu \in \mathbb{R}^d_{\geq 0}} \left(\frac{1}{2} h^2 (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} (\mu - r\mathbb{1} + \nu) V_{hh} - h(\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} \gamma_t V_{h\mu} \right)$$

We use the multiplicative ansatz $V(t, h, \mu) = h^{\frac{\alpha}{\alpha-1}} e^{f(t,\mu)}$ and get since $h > 0$ and $\alpha < 0$:

$$\begin{aligned} \lambda &= \operatorname*{arg\,max}_{\nu \in \mathbb{R}^d_{\geq 0}} \left(\frac{1}{2} \frac{\alpha}{(1-\alpha)^2} (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} (\mu - r\mathbb{1} + \nu) + \frac{\alpha}{1-\alpha} (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} \gamma_t f_\mu \right) \\ &= \operatorname*{arg\,min}_{\nu \in \mathbb{R}^d_{\geq 0}} \left(\frac{1}{2} \frac{1}{(1-\alpha)^2} (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} (\mu - r\mathbb{1} + \nu) + \frac{1}{1-\alpha} (\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1} \gamma_t f_\mu \right) \end{aligned}$$

The corresponding HJB-equation with this λ becomes

$$f_t = -\frac{\alpha}{1-\alpha} r - \frac{1}{2} \frac{\alpha}{(1-\alpha)^2} (\mu - r\mathbb{1} + \lambda)^\top \Sigma^{-1} (\mu - r\mathbb{1} + \lambda)$$

$$-\frac{\alpha}{1-\alpha}(\mu - r\mathbb{1} + \lambda)^\top \Sigma^{-1}\gamma_t f_\mu - \frac{1}{2} f_\mu^\top \gamma_t \Sigma^{-1}\gamma_t f_\mu - \frac{1}{2}\mathrm{tr}(f_{\mu\mu}\gamma_t\Sigma^{-1}\gamma_t) \tag{4.17}$$

with boundary condition $f(T,\mu) = 0$.
Standard non-linear optimization provides the 'penalty method' to handle this kind of constrained optimization problems. The penalty method allows the optimization parameter ν to be in the whole space $\mathbb{R}^d$ while some penalty function P penalizes it if being in the former constrained region. Therefore we introduce the following auxiliary unconstrained problems for $k \in \mathbb{N}$:

$$\lambda^{(k)} = \arg\min_{\nu\in\mathbb{R}^d}\Bigg(\frac{1}{2}\frac{1}{(1-\alpha)^2}(\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1}(\mu - r\mathbb{1} + \nu) + \frac{1}{1-\alpha}(\mu - r\mathbb{1} + \nu)^\top \Sigma^{-1}\gamma_t f_\mu + c^{(k)}P(\nu)\Bigg) \tag{4.18}$$

where $(c^{(k)})_{k\in\mathbb{N}} > 0$ is an increasing sequence of constants with $c^{(k)} \to \infty$ and P is a penalty function with $P(\nu) = 0$ on $\mathbb{R}^d_{\geq 0}$ and $P(\nu) > 0$ otherwise. Then the optimal solutions of the auxiliary problems converge to the optimal solution of the original problem $\lambda^{(k)} \to \lambda$ as shown in [Kru12, Sec.8.1]. We choose the following penalty function $P \in C^1$:

$$P(\nu) := \frac{1}{2}\frac{1}{(1-\alpha)^2}\sum_{i=1}^d (\nu_i^-)^2 = \frac{1}{2}\frac{1}{(1-\alpha)^2}\left\|\nu^-\right\|_2^2 \quad \left(= 0 \Leftrightarrow \nu \in \mathbb{R}^d_{\geq 0}\right)$$

$$\partial_\nu P(\nu) = \frac{1}{(1-\alpha)^2}\sum_{i=1}^d (-\nu_i^-)e_i = -\frac{1}{(1-\alpha)^2}\cdot \nu^-$$

Since the penalty function is differentiable on the whole of $\mathbb{R}^d$, the unconstrained auxiliary problems (4.18) can be solved explicitly by differentiating:

$$0 \stackrel{!}{=} \frac{1}{(1-\alpha)^2}\Sigma^{-1}(\mu - r\mathbb{1} + \lambda^{(k)}) + \frac{1}{1-\alpha}\Sigma^{-1}\gamma_t f_\mu - \frac{c_k}{(1-\alpha)^2}(\lambda^{(k)})^-$$

$$\Rightarrow \quad 0 = (\mu - r\mathbb{1} + \lambda^{(k)}) + (1-\alpha)\gamma_t f_\mu - c_k\Sigma(\lambda^{(k)})^-$$

$$\Rightarrow \quad \lambda^{(k)} - c_k\Sigma(\lambda^{(k)})^- = -(\mu - r\mathbb{1}) - (1-\alpha)\gamma_t f_\mu$$

Now we need to differ whether the $\lambda_i^{(k)}$ are positive or negative. Therefore we collect their indices in $I^{(k)}$ and $J^{(k)}$ where

$$I^{(k)} := \left\{i \,\middle|\, \lambda_i^{(k)} \leq 0\right\} \qquad J^{(k)} := \left\{j \,\middle|\, \lambda_j^{(k)} > 0\right\} \qquad \text{with } I^{(k)}\dot\cup J^{(k)} = \{1,...,d\}$$

However we know from non-linear optimization that $\lambda^{(k)} \to \lambda$ and $\lambda \in \mathbb{R}^d_{\geq 0} \setminus \left(\mathbb{R}^d_{>0} \cup \{0\}\right)$ hence $\lambda_i = 0$ for some $i \in I$ and $\lambda_j > 0$ for some $j \in J$.
Therefore there is some k_0 such that $I^{(k)} = I^{(k_0)}$ and $J^{(k)} = J^{(k_0)}$ for all $k \geq k_0$. We will

call $I = I^{(k_0)}$ the *active dimensions* and $J = J^{(k_0)}$ the *passive dimensions* in Case 3 and we will see later on that this is well-defined in the sense of Definition 3.6.

Now we solve (4.18) for $k \geq k_0$.

$$\begin{aligned}
&\begin{pmatrix} \lambda_I^{(k)} \\ \lambda_J^{(k)} \end{pmatrix} - c_k \begin{pmatrix} \Sigma_{II} & \Sigma_{IJ} \\ \Sigma_{JI} & \Sigma_{JJ} \end{pmatrix} \begin{pmatrix} -\lambda_I^{(k)} \\ 0_J \end{pmatrix} = \begin{pmatrix} -(\mu - r\mathbb{1})_I \\ -(\mu - r\mathbb{1})_J \end{pmatrix} - (1-\alpha) \begin{pmatrix} (\gamma_t f_\mu)_I \\ (\gamma_t f_\mu)_J \end{pmatrix} \\
\Rightarrow \quad & \begin{cases} \lambda_I^{(k)} + c_k \Sigma_{II} \lambda_I^{(k)} = -(\mu - r\mathbb{1})_I - (1-\alpha)(\gamma_t f_\mu)_I \\ \lambda_J^{(k)} + c_k \Sigma_{JI} \lambda_I^{(k)} = -(\mu - r\mathbb{1})_J - (1-\alpha)(\gamma_t f_\mu)_J \end{cases} \\
\Rightarrow \quad & \begin{cases} \lambda_I^{(k)} = -\big(Id_I + c_k \Sigma_{II}\big)^{-1} \big((\mu - r\mathbb{1})_I + (1-\alpha)(\gamma_t f_\mu)_I\big) \\ \lambda_J^{(k)} = -c_k \Sigma_{JI} \lambda_I^{(k)} - (\mu - r\mathbb{1})_J - (1-\alpha)(\gamma_t f_\mu)_J \end{cases}
\end{aligned} \tag{4.19}$$

Now for $k \to \infty$ we get $c_k \to \infty$ and $\lambda^{(k)} \to \lambda$. Hence the first line of (4.19) leads to:

$$\begin{aligned}
\big(Id_I + c_k \Sigma_{II}\big)^{-1} &= \frac{1}{c_k} \left(\frac{1}{c_k} Id_I + \Sigma_{II} \right)^{-1} \underset{k\to\infty}{\longrightarrow} 0_{II} \\
\Rightarrow \quad \lambda_I &= \lim_{k\to\infty} \lambda_I^{(k)} = 0_{II} \big((\mu - r\mathbb{1})_I + (1-\alpha)(\gamma_t f_\mu)_I\big) = 0_I
\end{aligned}$$

The second line of (4.19) leads to the solution for λ_J:

$$\begin{aligned}
-c_k \Sigma_{JI} \lambda_I^{(k)} &= c_k \Sigma_{JI} \left(Id_I + c_k \Sigma_{II}\right)^{-1} \left((\mu - r\mathbb{1})_I + (1-\alpha)(\gamma_t f_\mu)_I\right) \\
&= \Sigma_{JI} \left(\frac{1}{c_k} Id_I + \Sigma_{II} \right)^{-1} \left((\mu - r\mathbb{1})_I + (1-\alpha)(\gamma_t f_\mu)_I\right) \\
&\underset{k\to\infty}{\longrightarrow} \Sigma_{JI} (\Sigma_{II})^{-1} \big((\mu - r\mathbb{1})_I + (1-\alpha)(\gamma_t f_\mu)_I\big) \\
\Rightarrow \quad \lambda_J &= \lim_{k\to\infty} \lambda_J^{(k)} \\
&= \Sigma_{JI} (\Sigma_{II})^{-1} \big((\mu - r\mathbb{1})_I + (1-\alpha)(\gamma_t f_\mu)_I\big) - \big((\mu - r\mathbb{1})_J + (1-\alpha)(\gamma_t f_\mu)_J\big)
\end{aligned}$$

Combining these solutions we obtain the optimal λ.

$$\lambda = \mathbb{J}^I \cdot \big((\mu - r\mathbb{1}) + (1-\alpha)\gamma_t f_\mu\big) \tag{4.20}$$

$$\text{where} \quad \mathbb{J}^I := \begin{pmatrix} 0_{II} & 0_{IJ} \\ \Sigma_{JI} (\Sigma_{II})^{-1} & -Id_J \end{pmatrix} \tag{4.21}$$

At this point we observe that Case 3 has to be divided further into 2^d subcases, depending on which dimensions are active or passive. For $J = \emptyset$ we immediately get $\lambda = 0$ hence this is Case 1 and we will see in the following that $I = \emptyset$ just leads to Case 2.
Finally we need to consider $2^d - 2$ subcases that we will call Cases $(3, I)$ for each possible choice of active dimensions I.

Solving the HJB-equation

In the following we will solve the HJB-equation (4.17) in subcase $(3, I)$ with the following ansatz.

$$f(t,\mu) = \frac{1}{2}\frac{1}{1-\alpha}(\mu - r\mathbb{1})^\top A(t)(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}k(t) \qquad \text{with symmetric } A$$

This ansatz needs the boundary conditions $A(T) = 0$, $k(T) = 0$ and the following derivatives:

$$\begin{aligned} f_t(t,\mu) &= \frac{1}{2}\frac{1}{1-\alpha}(\mu - r\mathbb{1})^\top A'(t)(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}k'(t) \\ f_\mu(t,\mu) &= \frac{1}{1-\alpha}A(t)(\mu - r\mathbb{1}) \\ f_{\mu\mu}(t,\mu) &= \frac{1}{1-\alpha}A(t) \end{aligned}$$

The optimal λ from (4.20) becomes using this ansatz

$$\lambda = \mathbb{J}^I\big(Id + \gamma_t A(t)\big)(\mu - r\mathbb{1})$$

and the HJB-equation (4.17) becomes:

$$\begin{aligned} &\frac{1}{2}\frac{1}{1-\alpha}(\mu - r\mathbb{1})^\top A'(t)(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}k'(t) \\ = f_t &= -\frac{\alpha}{1-\alpha}r - \frac{1}{2}\frac{\alpha}{(1-\alpha)^2}(\mu - r\mathbb{1} + \lambda)^\top\Sigma^{-1}(\mu - r\mathbb{1} + \lambda) \\ &\quad - \frac{\alpha}{1-\alpha}(\mu - r\mathbb{1} + \lambda)^\top\Sigma^{-1}\gamma_t f_\mu - \frac{1}{2}f_\mu^\top\gamma_t\Sigma^{-1}\gamma_t f_\mu - \frac{1}{2}\mathrm{tr}(f_{\mu\mu}\gamma_t\Sigma^{-1}\gamma_t) \\ &= -\frac{\alpha}{1-\alpha}r \\ &\quad - \frac{1}{2}\frac{\alpha}{(1-\alpha)^2}(\mu - r\mathbb{1})^\top\big(Id + \mathbb{J}^I + \mathbb{J}^I\gamma_t A(t)\big)^\top\Sigma^{-1}\big(Id + \mathbb{J}^I + \mathbb{J}^I\gamma_t A(t)\big)(\mu - r\mathbb{1}) \\ &\quad - \frac{\alpha}{(1-\alpha)^2}(\mu - r\mathbb{1})^\top\big(Id + \mathbb{J}^I + \mathbb{J}^I\gamma_t A(t)\big)^\top\Sigma^{-1}\gamma_t A(t)(\mu - r\mathbb{1}) \\ &\quad - \frac{1}{2}\frac{1}{(1-\alpha)^2}(\mu - r\mathbb{1})^\top A(t)\gamma_t\Sigma^{-1}\gamma_t A(t)(\mu - r\mathbb{1}) - \frac{1}{2}\frac{1}{1-\alpha}\mathrm{tr}(A(t)\gamma_t\Sigma^{-1}\gamma_t) \\ &= -\frac{\alpha}{1-\alpha}r \\ &\quad - \frac{1}{2}\frac{\alpha}{(1-\alpha)^2}(\mu - r\mathbb{1})^\top(Id + \gamma_t A(t))^\top(Id + \mathbb{J}^I)^\top\Sigma^{-1}(Id + \mathbb{J}^I)(Id + \gamma_t A(t))(\mu - r\mathbb{1}) \\ &\quad - \frac{1}{2}\frac{1}{1-\alpha}(\mu - r\mathbb{1})^\top A(t)\gamma_t\Sigma^{-1}\gamma_t A(t)(\mu - r\mathbb{1}) - \frac{1}{2}\frac{1}{1-\alpha}\mathrm{tr}(A(t)\gamma_t\Sigma^{-1}\gamma_t) \end{aligned} \tag{4.22}$$

To simplify notation we introduce the following matrix in Case $(3, I)$

$$\Sigma_{\mathbb{J}}^{(3,I)} := \left(Id + \mathbb{J}^I\right)^\top \Sigma^{-1}\left(Id + \mathbb{J}^I\right) \tag{4.23}$$

$$
\begin{aligned}
&= \begin{pmatrix} Id_I & (\Sigma_{II})^{-1}\Sigma_{IJ} \\ 0_{JI} & 0_{JJ} \end{pmatrix} \cdot \begin{pmatrix} (\Sigma^{-1})_{II} & (\Sigma^{-1})_{IJ} \\ (\Sigma^{-1})_{JI} & (\Sigma^{-1})_{JJ} \end{pmatrix} \cdot \begin{pmatrix} Id_I & 0_{IJ} \\ \Sigma_{JI}(\Sigma_{II})^{-1} & 0_{JJ} \end{pmatrix} \\
&= \begin{pmatrix} Id_I & (\Sigma_{II})^{-1}\Sigma_{IJ} \\ 0_{JI} & 0_{JJ} \end{pmatrix} \cdot \begin{pmatrix} (\Sigma^{-1})_{II} + (\Sigma^{-1})_{IJ}\Sigma_{JI}(\Sigma_{II})^{-1} & 0_{IJ} \\ (\Sigma^{-1})_{JI} + (\Sigma^{-1})_{JJ}\Sigma_{JI}(\Sigma_{II})^{-1} & 0_{JJ} \end{pmatrix} \\
&= \begin{pmatrix} Id_I & (\Sigma_{II})^{-1}\Sigma_{IJ} \\ 0_{JI} & 0_{JJ} \end{pmatrix} \cdot \begin{pmatrix} (\Sigma_{II})^{-1} & 0_{IJ} \\ 0_{JI} & 0_{JJ} \end{pmatrix} \\
&= \begin{pmatrix} (\Sigma_{II})^{-1} & 0_{IJ} \\ 0_{JI} & 0_{JJ} \end{pmatrix}
\end{aligned}
$$

Note that the HJB-equation (4.22) is quadratic in $(\mu - r\mathbb{1})$ without linear term and therefore can be split up. We observe the following two ODEs for the quadratic and the constant part of (4.22):

$$A'(t) = -\frac{\alpha}{1-\alpha}(Id + \gamma_t A(t))^\top \Sigma_{\mathbb{J}}^{(3,I)}(Id + \gamma_t A(t)) - A(t)\gamma_t\Sigma^{-1}\gamma_t A(t) \tag{4.24}$$

$$k'(t) = -\alpha r - \frac{1}{2}\operatorname{tr}\left(A(t)\gamma_t\Sigma^{-1}\gamma_t\right) \tag{4.25}$$

where the boundary conditions $A(T) = 0$ and $k(T) = 0$ stay unchanged.

Solving the ordinary differential equations

Fortunately in this case the matrix Ricatti equation (4.24) can also be solved explicitly. We define $B(t) := Id + \gamma_t A(t)$ and vice versa $A(t) = \gamma_t^{-1}(B(t) - Id)$ with corresponding boundary condition $B(T) = Id$. Hence $A'(t) = \Sigma^{-1}(B(t) - Id) + \gamma_t^{-1}B'(t)$ and ODE (4.24) becomes

$$\gamma_t^{-1}B'(t) = -\frac{\alpha}{1-\alpha}B(t)^\top\Sigma_{\mathbb{J}}^{(3,I)}B(t) - B(t)^\top\Sigma^{-1}(B(t) - Id)$$

Now we define $C(t) := B(t)\gamma_t$ hence $B(t) = C(t)\gamma_t^{-1}$ with corresponding boundary condition $C(T) = \gamma_T$ and derivative $B'(t) = C'(t)\gamma_t^{-1} + C(t)\Sigma^{-1}$. If additionally C is symmetric then ODE (4.24) becomes

$$
\begin{aligned}
\gamma_t^{-1}C'(t)\gamma_t^{-1} &= -\frac{\alpha}{1-\alpha}\gamma_t^{-1}C(t)\Sigma_{\mathbb{J}}^{(3,I)}C(t)\gamma_t^{-1} - \gamma_t^{-1}C(t)\Sigma^{-1}(C(t)\gamma_t^{-1} - Id) - \gamma_t^{-1}C(t)\Sigma^{-1} \\
\Rightarrow \quad C'(t) &= -\frac{\alpha}{1-\alpha}C(t)\Sigma_{\mathbb{J}}^{(3,I)}C(t) - C(t)\Sigma^{-1}C(t) \\
&= -C(t)\left(\frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(3,I)} + \Sigma^{-1}\right)C(t)
\end{aligned}
$$

To solve this remaining ODE we use $D(t) = C(t)^{-1}$ with corresponding boundary condition $D(T) = \gamma_T^{-1}$. Then $C'(t) = -D^{-1}(t)D'(t)D^{-1}(t)$ and ODE (4.24) becomes

$$-D^{-1}(t)D'(t)D^{-1}(t) = C'(t) = -D^{-1}(t)\left(\frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(3,I)} + \Sigma^{-1}\right)D^{-1}(t)$$

Hence $D'(t) = \frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(3,I)} + \Sigma^{-1}$ and the boundary condition leads to the resulting

$$\begin{aligned} D(t) &= \gamma_T^{-1} - \Sigma^{-1}(T-t) - \frac{\alpha}{1-\alpha}(T-t)\Sigma_{\mathbb{J}}^{(3,I)} \\ \text{and} \quad C(t) &= \left(\gamma_t^{-1} - \frac{\alpha}{1-\alpha}(T-t)\Sigma_{\mathbb{J}}^{(3,I)}\right)^{-1}. \end{aligned}$$

Indeed $C(t)$ is symmetric.

Now we can solve the second ODE (4.25):

$$\begin{aligned} k'(t) &= -\alpha r - \frac{1}{2}\mathrm{tr}\left(A(t)\gamma_t\Sigma^{-1}\gamma_t\right) \\ &= -\alpha r - \frac{1}{2}\mathrm{tr}\left((C(t)-\gamma_t)\Sigma^{-1}\right) \\ \Rightarrow \quad k(t) &= \alpha(T-t)r + \frac{1}{2}\int_t^T \mathrm{tr}\left(C(s)\Sigma^{-1}\right)ds - \frac{1}{2}\int_t^T \mathrm{tr}\left(\gamma_s\Sigma^{-1}\right)ds \\ &= \alpha(T-t)r + \frac{1}{2}\int_t^T \mathrm{tr}\left(C(s)\Sigma^{-1}\right)ds - \frac{1}{2}\log\det\gamma_T^{-1} + \frac{1}{2}\log\det\gamma_t^{-1} \end{aligned}$$

Here the last line holds du to the chain rule $\partial_t \log\det M(t) = \mathrm{tr}\left(M(t)^{-1}\partial_t M(t)\right)$ applied to $M(s) = \gamma_s^{-1}$ with $\partial_t\gamma_t^{-1} = \Sigma^{-1}$.
Since γ_t^{-1} is symmetric and positive definite, $\det\gamma_t^{-1}$ is positive and hence the log is well-defined. Unfortunately the other integral cannot be calculated explicitly, it needs to be approximated numerically.

The candidate solutions of the stochastic control approach

In order to point out the underlying case we will refer in the following to the optimal dual solution λ in Case $(3, I)$ as follows:

$$\begin{aligned} \lambda^{(3,I)}(t,\mu) &= \mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \qquad (4.26) \\ \text{where} \qquad \mathbb{J}^I &= \begin{pmatrix} 0_{II} & 0_{IJ} \\ \Sigma_{JI}\Sigma_{II}^{-1} & -Id_J \end{pmatrix} \\ \text{and} \quad C^{(3,I)}(t) &:= \left(\gamma_t^{-1} - \frac{\alpha}{1-\alpha}(T-t)\Sigma_{\mathbb{J}}^{(3,I)}\right)^{-1} \end{aligned}$$

Now we only need to verify this stochastic control approach.

Theorem 4.15 (Verification in Case $(3, I)$)
Define the candidates for value function and optimal control as derived above.

$$\begin{aligned} \bar{V}(t,h,\mu) &:= h^{\frac{\alpha}{\alpha-1}}e^{f(t,\mu)} \\ \bar{\lambda}(t,h,\mu) &:= \lambda^{(3,I)}(t,\mu) \end{aligned}$$

Then they solve the stochastic control approach (4.12) in Case $(3, I)$, hence $\bar{V} = V$ and $\bar{\lambda} = \nu^*$, where

$$V(t,h,\mu) := \sup_{\nu \in \mathcal{D}(t,h,\mu)} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}} \middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu) \right]$$
$$\text{and} \quad \nu^* := \arg\max_{\nu \in \mathcal{D}} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}} \right]$$

Proof. Let $(t,h,\mu) \in [0,T) \times \mathbb{R}^+ \times \mathbb{R}^d$, $\nu \in \mathcal{D}$ and $s \in [t,T)$ be arbitrary and let τ be an arbitrary stopping time with values in $[t,T]$.
As $\bar{V} \in C^{1,2,2}$ we get with Itô's formula

$$\begin{aligned} \bar{V}(s\wedge\tau, H^\nu_{s\wedge\tau}, \hat{\mu}_{s\wedge\tau}) &= \bar{V}(t,h,\mu) + \int_t^{s\wedge\tau} \bar{V}_t(u) + \mathcal{L}^\nu \bar{V}(u) du \\ &\quad + \int_t^{s\wedge\tau} \bar{V}_\mu(u)^\top \gamma_u (\sigma^{-1})^\top - \bar{V}_h(u) H_u^\nu (\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top (\sigma^{-1})^\top dV_u \end{aligned} \tag{4.27}$$

$$\begin{aligned} \text{where} \quad \mathcal{L}^\nu \bar{V}(u) &= -\bar{V}_h(u) H_u^\nu \left(r + \delta(\nu_u)\right) \\ &\quad + \frac{1}{2}(H_u^\nu)^2 (\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top \Sigma^{-1} (\hat{\mu}_u - r\mathbb{1} + \nu_u) V_{hh}(u) \\ &\quad + H_u^\nu (\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top \Sigma^{-1} \gamma_u V_{h\mu}(u) + \frac{1}{2} \mathrm{tr}\big(V_{\mu\mu}(u) \cdot \gamma_u \Sigma^{-1} \gamma_u\big) \end{aligned}$$

Here we use the notation $\bar{V}(u) := \bar{V}(u, H_u^\nu, \hat{\mu}_u)$.
While most sub-indices indicate time-dependency of the corresponding variables, sub-indices of $\bar{V}$ are the respective derivatives of $\bar{V}(t,h,\mu)$.
Now for each $n \in \mathbb{N}$ define the stopping time

$$\begin{aligned} \tau_n := \inf \Bigg\{ s \geq t \;\Bigg|\; & \int_t^s \left\| \bar{V}_\mu(u)^\top \gamma_u (\sigma^{-1})^\top \right\|_2^2 du \geq n \\ & \text{or} \int_t^s \left\| \bar{V}_h(u) H_u^\nu (\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top (\sigma^{-1})^\top \right\|_2^2 du \geq n \Bigg\} \wedge T \end{aligned}$$

Hence $\tau_n \to T$ as $n \to \infty$ and therefore for any $n \in \mathbb{N}$ the stopped process

$$\int_t^{s\wedge\tau} \bar{V}_\mu(u)^\top \gamma_u (\sigma^{-1})^\top - \bar{V}_h(u) H_u^\nu (\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top (\sigma^{-1})^\top dV_u$$

is a martingale with zero expectation.
Now we use τ_n in (4.27) and take conditional expectations on both sides.

$$\begin{aligned} &\mathbb{E}\left[\bar{V}(s\wedge\tau_n, H^\nu_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n}) \middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu) \right] \\ &= \bar{V}(t,h,\mu) + \mathbb{E}\left[\int_t^{s\wedge\tau_n} \bar{V}_t(u, H_u^\nu, \hat{\mu}_u) + \mathcal{L}^\nu \bar{V}(u, H_u^\nu, \hat{\mu}_u) du \middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu) \right] \end{aligned}$$

Since $\bar{V}(t,h,\mu)$ satisfies the HJB-equation (4.13) in Case $(3, I)$, we observe for the above right hand side:

$$\bar{V}_t(u, H_u^\nu, \hat{\mu}_u) + \mathcal{L}^\nu \bar{V}(u, H_u^\nu, \hat{\mu}_u) \leq 0 \qquad \text{for all } \nu \in \mathcal{D}$$

$$\bar{V}_t(u, H_u^\nu, \hat{\mu}_u) + \mathcal{L}^\nu \bar{V}(u, H_u^\nu, \hat{\mu}_u) = 0 \qquad \text{if } \nu = \bar{\lambda}$$

This leads for each $u \in [t, T]$ and $\nu \in \mathcal{D}(t, h, \mu)$ to the following equations:

$$\begin{aligned} \mathbb{E}\left[\bar{V}(s \wedge \tau_n, H^\nu_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\middle| (H_t^\nu, \hat{\mu}_t) = (h, \mu)\right] &\leq \bar{V}(t, h, \mu) \qquad \text{for all } \nu \in \mathcal{D}(t, h, \mu) \\ \mathbb{E}\left[\bar{V}(s \wedge \tau_n, H^\nu_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\middle| (H_t^\nu, \hat{\mu}_t) = (h, \mu)\right] &= \bar{V}(t, h, \mu) \qquad \text{if } \nu = \bar{\lambda} \end{aligned}$$

By Lemma 4.16 we can apply dominated convergence and observe for $n \to \infty$

$$\begin{aligned} \bar{V}(t, h, \mu) &\geq \mathbb{E}\left[\bar{V}(s, H_s^\nu, \hat{\mu}_s)\middle| (H_t^\nu, \hat{\mu}_t) = (h, \mu)\right] \qquad \text{for all } \nu \in \mathcal{D}(t, h, \mu) \\ \bar{V}(t, h, \mu) &= \mathbb{E}\left[\bar{V}(s, H_s^\nu, \hat{\mu}_s)\middle| (H_t^\nu, \hat{\mu}_t) = (h, \mu)\right] \qquad \text{if } \nu = \bar{\lambda} \end{aligned}$$

Since $\bar{V}(s, H_s^\nu, \hat{\mu}_s) \to \bar{V}(T, H_T^\nu, \hat{\mu}_T) = (H_T^\nu)^{\frac{\alpha}{\alpha-1}}$ for $s \to T$ we conclude by dominated convergence

$$\begin{aligned} &\bar{V}(t, h, \mu) \geq \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\middle| (H_t^\nu, \hat{\mu}_t) = (h, \mu)\right] \qquad \text{for all } \nu \in \mathcal{D}(t, h, \mu) \\ \Rightarrow \quad &\bar{V}(t, h, \mu) \geq \sup_{\nu \in \mathcal{D}(t,h,\mu)} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\middle| (H_t^\nu, \hat{\mu}_t) = (h, \mu)\right] = V(t, h, \mu) \end{aligned}$$

Since we get equality for $\nu = \bar{\lambda}$ we get $V(t, h, \mu) = \bar{V}(t, h, \mu)$ with optimizer $\nu^* = \bar{\lambda}$. □

Lemma 4.16
There is an integrable random variable Y with $\left|\bar{V}(s \wedge \tau_n, H^\nu_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\right| \leq Y$.

Proof. Since $\frac{\alpha}{\alpha-1} \in (0, 1)$ we observe

$$h^{\frac{\alpha}{\alpha-1}} \leq \begin{cases} 1 & \text{if } h \leq 1 \\ h & \text{if } h > 1 \end{cases} \qquad \text{hence} \qquad 0 < h^{\frac{\alpha}{\alpha-1}} \leq 1 + h$$

The continuous function f is a second order polynomial in μ, hence there are constants C_t such that

$$f(t, \mu) \leq C_t(1 + \|\mu\|_2^2)$$

The leading constants C_t may depend on time t, however $t \in [0, T]$ is in a closed interval, hence $f(\cdot, \mu)$ is uniformly continuous in t and attains its supremum as a maximum.

$$f(t, \mu) \leq C(1 + \|\mu\|_2^2)$$

for some upper bound $C \geq \max_{t\in[0,T]} C_t < \infty$.
Hence

$$\begin{aligned} &\left|\bar{V}(t, h, \mu)\right| = h^{\frac{\alpha}{\alpha-1}} e^{f(t,\mu)} \leq (1 + h)e^{C(1+\|\mu\|_2^2)} \\ \Rightarrow \quad &\left|\bar{V}(s \wedge \tau_n, H^\nu_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\right| \leq (1 + H^\nu_{s\wedge\tau_n}) \cdot e^{C(1+\|\hat{\mu}_{s\wedge\tau_n}\|_2^2)} \\ &\qquad \leq (1 + M) \cdot e^{C(1+\|\hat{\mu}_{s\wedge\tau_n}\|_2^2)} =: Y \end{aligned}$$

$$\text{where} \qquad M := \sup_{s \in [t,T]} H_u^\nu$$

Now on the one hand $H_s^\nu > 0$ is uniformly continuous in $[t, T]$, and hence attains its supremum $M = \sup_{s\in[t,T]} H_u^\nu < \infty$.
And on the other hand $\hat{\mu}_t$ is a martingale, hence $\|\hat{\mu}_s\|_2^2$ is a submartingale since $\|\cdot\|_2^2$ is a convex function. Hence $\exp\Big(C(1 + \|\hat{\mu}_s\|_2^2)\Big)$ is also a submartingale and by Doob's maximal inequality we get

$$\mathbb{E}\left[\sup_{s\in[t,T]} \exp\Big(C(1 + \|\hat{\mu}_s\|_2^2)\Big)\right] \leq \left(\frac{C}{C-1}\right)^C \mathbb{E}\left[\exp\Big(C(1 + \|\hat{\mu}_T\|_2^2)\Big)\right] < \infty$$

Hence $\mathbb{E}\left[Y\right] < \infty$. □

4.2.2 The dual problem in Case 4

Case 4 is the most important case to consider, since it contains in particular most of those cases where the optimal unconstrained solution invests more money than admissible. This can be seen in more detail from the observations on the domain of the value function in Section 4.4. In Case 4 we are in the following situation

$$\begin{aligned} & \lambda \in \mathbb{R}^d \setminus \mathbb{R}^d_{\geq 0} \\ \text{hence} \quad & \lambda_i = \bar{\lambda} \ \text{ for the dimensions } i \in I \\ & \lambda_j > \bar{\lambda} \ \text{ for the dimensions } j \in J \\ \text{where} \quad & \bar{\lambda} := \min_{k\in\{1,\ldots,d\}} \lambda_k \ < 0 \end{aligned}$$

Again we will call I the *active* and J the *passive dimensions*. Since the corresponding optimal portfolio strategy will satisfy $\pi_J = 0_J$ this is compatible with Definition 3.6.
Additionally we will see in the following that there can be several choices of I and J that lead to the unique optimal solution. This happens for the same reasons than in Corollary 3.19 for logarithmic utility: If a dimension k is on the boundary between being active and being passive, than it satisfies $\pi_k = 0$ ('passive') but also $\lambda_i = \bar{\lambda}$ ('active') and for both choices $k \in I$ or $k \in J$ we get the same optimal solution.

To avoid notational difficulties we will therefore assume in the following that there are no parameters (t, h, μ) in our current stochastic control problem (4.12) such that the optimal solution contains any boundary dimensions. This can be done without loss of generality since this is a null-set anyway and hence not relevant for maximizing a standard expectation value. However the observation from logarithmic utility will also hold true for power utility: It does not change the optimal solution whether a boundary dimension is considered active or passive.

At this point we observe that Case 4 also has to be divided further into 2^d sub-cases, similar to Case 3. In the following we will therefore assume to be in Case $(4, I)$, being Case

4 with active dimensions I. It is obvious that the active dimensions cannot be empty, since they collect those dimensions i with minimal λ_i out of finitely many dimensions. Hence we get $2^d - 1$ subcases to consider.

We want to solve the stochastic control approach (4.12) in Case $(4, I)$. In order to solve the corresponding HJB-equation (4.13) we need to consider an auxiliary problem such that $\delta(\lambda) = -\bar{\lambda}$ gets differentiable in λ. Note that the distance between the $\lambda_i = \bar{\lambda}$ and the $\lambda_j > \bar{\lambda}$ is larger than zero, hence there is some ε-ball around the parameters (t, h, μ) such that the resulting active dimensions I don't switch within this ε-ball.

In the following we consider an auxiliary problem that maps the domain of the support function δ in Case $(4, I)$ bijectively into a new domain such that $\delta(\lambda) = -\bar{\lambda}$ gets differentiable.

Without loss of generality we simplify notation by assuming $1 \in I$. First we define the new domain G^I and the linear bijection φ^I, both depending on the active dimensions I.

$$G^I := \left\{ x \in \mathbb{R} \times \mathbb{R}^{|J|} \;\middle|\; x_1 < 0,\ x_j > x_1 \text{ for all } j \in J \right\}$$

$$\varphi^I : G^I \to H^I : \begin{pmatrix} x_1 \\ x_J \end{pmatrix} \mapsto \begin{pmatrix} x_1 \mathbb{1}_I \\ x_J \end{pmatrix}$$

$$\Rightarrow \quad (\varphi^I)^{-1} : H^I \to G^I : \begin{pmatrix} x_I \\ x_J \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_J \end{pmatrix}$$

Obviously the image $H^I := \mathrm{Im}(\varphi^I) \subset \mathbb{R}^d$ collects just all admissible solutions for the optimal λ in Case $(4, I)$ and satisfies

$$\dot{\bigcup_{\emptyset \neq I \subseteq \{1,\dots,d\}}} H^I = \mathbb{R}^d \setminus \mathbb{R}^d_{\geq 0}$$

The mapping φ^I is bijective and linear in the sense that it can be written using matrices.

$$\varphi^I(x) = \Phi x \quad \text{where} \quad \Phi := \begin{pmatrix} \mathbb{1}_I & 0_{IJ} \\ 0_J & Id_J \end{pmatrix} \in \mathbb{R}^{d \times 1+|J|}$$

$$\text{and} \quad (\varphi^I)^{-1}(x) = \Psi x \quad \text{where} \quad \Psi := \begin{pmatrix} e_1^\top & 0_J^\top \\ 0_{JI} & Id_J \end{pmatrix} \in \mathbb{R}^{1+|J| \times d}$$

Since $\lambda \in \mathrm{Im}(\varphi^I) = H^I$ we may define the auxiliary parameter $\tilde{\lambda}$.

$$\tilde{\lambda} := (\varphi^I)^{-1}(\lambda) = \begin{pmatrix} \lambda_1 \\ \lambda_J \end{pmatrix} \in G^I$$

And since φ^I is a bijection we can reduce our problem from finding the optimal $\lambda \in H^I$ to finding the optimal $\tilde{\lambda} \in G^I$.

Additionally we get the following nice property in G^I:
Although the support function δ is not differentiable with respect to λ we observe that $\delta \circ \varphi^I$ is differentiable with respect to $\tilde{\lambda}$ on G^I.

$$(\delta \circ \varphi^I)(\tilde{\lambda}) = \delta\left(\begin{pmatrix} \lambda_1 \mathbb{1}_I \\ \lambda_J \end{pmatrix}\right) = \max\{-\lambda_1, 0\} = -\lambda_1$$

$$\Rightarrow \quad (\delta \circ \varphi^I)'(\tilde{\lambda}) = -e_1$$

Deriving the optimizer

Now we can go back to solving our original stochastic control problem. With $\lambda = \varphi^I(\tilde{\lambda})$ respectively $\nu = \varphi^I(\tilde{\nu})$ its HJB-equation (4.13) becomes:

$$\begin{aligned} 0 = \sup_{\tilde{\nu} \in G^I} \Bigg(& V_t - V_h h(r + (\delta \circ \varphi^I)(\tilde{\nu})) + \frac{1}{2} h^2 \left(\mu - r\mathbb{1} + \varphi^I(\tilde{\nu})\right)^\top \Sigma^{-1} \left(\mu - r\mathbb{1} + \varphi^I(\tilde{\nu})\right) V_{hh} \\ & - h \left(\mu - r\mathbb{1} + \varphi^I(\tilde{\nu})\right)^\top \Sigma^{-1} \gamma_t V_{h\mu} + \frac{1}{2} \mathrm{tr}(V_{\mu\mu} \cdot \gamma_t \Sigma^{-1} \gamma_t) \Bigg) \end{aligned} \tag{4.28}$$

with boundary condition $V(T, h, \mu) = h^{\frac{\alpha}{\alpha-1}}$. The optimal $\tilde{\lambda}$ gets chosen by (4.14):

$$\begin{aligned} \tilde{\lambda} &= \arg\max_{\tilde{\nu} \in G^I} \Bigg(- h\delta(\varphi^I(\tilde{\nu})) V_h + \frac{1}{2} h^2 \left(\mu - r\mathbb{1} + \varphi^I(\tilde{\nu})\right)^\top \Sigma^{-1} \left(\mu - r\mathbb{1} + \varphi^I(\tilde{\nu})\right) V_{hh} \\ &\qquad - h \left(\mu - r\mathbb{1} + \varphi^I(\tilde{\nu})\right)^\top \Sigma^{-1} \gamma_t V_{h\mu} \Bigg) \\ &= \arg\max_{\tilde{\nu} \in G^I} \Bigg(- h(\delta \circ \varphi^I)(\tilde{\nu}) V_h + \frac{1}{2} h^2 (\mu - r\mathbb{1} + \Phi\tilde{\nu})^\top \Sigma^{-1} (\mu - r\mathbb{1} + \Phi\tilde{\nu}) V_{hh} \\ &\qquad - h(\mu - r\mathbb{1} + \Phi\tilde{\nu})^\top \Sigma^{-1} \gamma_t V_{h\mu} \Bigg) \end{aligned}$$

Now we get the optimal $\tilde{\lambda}$ by differentiating:

$$\begin{aligned} 0 &\overset{!}{=} h V_h e_1 + h^2 \Phi^\top \Sigma^{-1} (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) V_{hh} - h \Phi^\top \Sigma^{-1} \gamma_t V_{h\mu} \\ \Rightarrow \quad \Phi^\top \Sigma^{-1} (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) &= -\frac{V_h}{h V_{hh}} e_1 + \Phi^\top \Sigma^{-1} \gamma_t \frac{V_{h\mu}}{h V_{hh}} \end{aligned} \tag{4.29}$$

In (4.29) there are $1 + |J|$ equations for $1 + |J|$ unknown variables. In the first line we get:

$$\begin{aligned} -\frac{V_h}{h V_{hh}} + \left(\mathbb{1}_I^\top \ 0_J^\top\right) \Sigma^{-1} \gamma_t \frac{V_{h\mu}}{h V_{hh}} &= \left(\mathbb{1}_I^\top \ 0_J\right) \Sigma^{-1} (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) \\ &= \left(\mathbb{1}_I^\top (\Sigma^{-1})_{II} \quad \mathbb{1}_I^\top (\Sigma^{-1})_{IJ}\right) (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) \\ &= \mathbb{1}_I^\top (\Sigma^{-1})_{II} (\mu_I - r\mathbb{1}_I + \lambda_1 \mathbb{1}_I) + \mathbb{1}_I^\top (\Sigma^{-1})_{IJ} (\mu_J - r\mathbb{1}_J + \lambda_J) \end{aligned} \tag{4.30}$$

and in the other $|J|$ lines of (4.29) we get:

$$\begin{aligned} \left(0_{JI} \ Id_J\right) \Sigma^{-1} \gamma_t \frac{V_{h\mu}}{h V_{hh}} &= \left(0_{JI} \ Id_J\right) \Sigma^{-1} (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) \\ &= \left((\Sigma^{-1})_{JI} \quad (\Sigma^{-1})_{JJ}\right) (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) \\ &= (\Sigma^{-1})_{JI} (\mu_I - r\mathbb{1}_I + \lambda_1 \mathbb{1}_I) + (\Sigma^{-1})_{JJ} (\mu_J - r\mathbb{1}_J + \lambda_J) \end{aligned}$$

$$\Rightarrow \quad (\Sigma^{-1})_{JJ}(\mu_J - r\mathbb{1}_J + \lambda_J) = -(\Sigma^{-1})_{JI}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I) + \big(0_{JI} \quad Id_J\big)\Sigma^{-1}\gamma_t \frac{V_{h\mu}}{hV_{hh}}$$

Now the matrix inversion formula (3.8) leads to

$$\begin{aligned}(\mu_J - r\mathbb{1}_J + \lambda_J) = &-\left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)(\Sigma^{-1})_{JI}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I) \\ &+ \left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)\big(0_{JI} \quad Id_J\big)\Sigma^{-1}\gamma_t \frac{V_{h\mu}}{hV_{hh}}\end{aligned} \tag{4.31}$$

When plugging (4.31) into equation (4.30) and frequently applying the matrix inversion formula (3.8), we get the optimal solution for $\bar{\lambda} = \lambda_1 = \tilde{\lambda}_1$.

$$\begin{aligned}
&\mathbb{1}_I^\top(\Sigma^{-1})_{II}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I) \\
&\qquad = -\mathbb{1}_I^\top(\Sigma^{-1})_{IJ}(\mu_J - r\mathbb{1}_J + \lambda_J) - \frac{V_h}{hV_{hh}} + \big(\mathbb{1}_I^\top \quad 0_J^\top\big)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\qquad = \mathbb{1}_I^\top(\Sigma^{-1})_{IJ}\left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)(\Sigma^{-1})_{JI}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I) \\
&\qquad\quad - \mathbb{1}_I^\top(\Sigma^{-1})_{IJ}\left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)\big(0_{JI} \quad Id_J\big)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\qquad\quad - \frac{V_h}{hV_{hh}} + \big(\mathbb{1}_I^\top \quad 0_J^\top\big)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\qquad = -\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{IJ}(\Sigma^{-1})_{JI}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I) \\
&\qquad\quad + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{IJ}\big(0_{JI} \quad Id_J\big)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\qquad\quad - \frac{V_h}{hV_{hh}} + \big(\mathbb{1}_I^\top \quad 0_J^\top\big)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
\Rightarrow\quad &\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I) \\
&\qquad = \mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{IJ}\big(0_{JI} \quad Id_J\big)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\qquad\quad - \frac{V_h}{hV_{hh}} + \big(\mathbb{1}_I^\top \quad 0_J^\top\big)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\qquad = -\frac{V_h}{hV_{hh}} + \big(\mathbb{1}_I^\top \quad \mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{IJ}\big)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\qquad = -\frac{V_h}{hV_{hh}} + \big(\mathbb{1}_I^\top(\Sigma_{II})^{-1} \quad 0_J^\top\big)\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
\Rightarrow\quad &\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I\lambda_1 \\
&\qquad = -\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu_I - r\mathbb{1}_I) - \frac{V_h}{hV_{hh}} + \big(\mathbb{1}_I^\top(\Sigma_{II})^{-1} \quad 0_J^\top\big)\gamma_t\frac{V_{h\mu}}{hV_{hh}}
\end{aligned}$$

When introducing the notation $g^I := \left(\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I\right)^{-1}$ the optimal $\bar{\lambda} = \lambda_1 = \tilde{\lambda}_1$ becomes

$$\bar{\lambda} = -g^I\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I - g^I\frac{V_h}{hV_{hh}} + g^I\big(\mathbb{1}_I^\top(\Sigma_{II})^{-1} \quad 0_J^\top\big)\gamma_t\frac{V_{h\mu}}{hV_{hh}} \tag{4.32}$$

Finally from equation (4.31) we get the optimal λ_J:

$$\begin{aligned}
\lambda_J &= -(\mu - r\mathbb{1})_J - \left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)(\Sigma^{-1})_{JI}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I) \\
&\qquad + \left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)\begin{pmatrix}0_{JI} & Id_J\end{pmatrix}\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&= -(\mu - r\mathbb{1})_J + \Sigma_{JI}(\Sigma_{II})^{-1}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I) \\
&\qquad + \left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)\begin{pmatrix}(\Sigma^{-1})_{JI} & (\Sigma^{-1})_{JJ}\end{pmatrix}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&= -(\mu - r\mathbb{1})_J + \Sigma_{JI}(\Sigma_{II})^{-1}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I) \\
&\qquad + \begin{pmatrix}-\Sigma_{JI}(\Sigma_{II})^{-1} & Id_J\end{pmatrix}\gamma_t\frac{V_{h\mu}}{hV_{hh}}
\end{aligned} \tag{4.33}$$

Solving the HJB-equation

The remaining step is to solve the HJB-equation (4.28). Therefore we consider the multiplicative ansatz $V(t,h,\mu) = h^{\frac{\alpha}{\alpha-1}}e^{f(t,\mu)}$ leading to

$$\frac{V_h}{hV_{hh}} = (\alpha - 1) \quad \text{and} \quad \frac{V_{h\mu}}{hV_{hh}} = (\alpha - 1)f_\mu$$

and update the recently derived optimal λ:

$$\begin{aligned}
\bar{\lambda} &= -g^I\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I - g^I(\alpha - 1) + g^I\begin{pmatrix}\mathbb{1}_I^\top(\Sigma_{II})^{-1} & 0_J^\top\end{pmatrix}\gamma_t(\alpha - 1)f_\mu \\
&= (1-\alpha)g^I - g^I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\left((\mu - r\mathbb{1})_I + (1-\alpha)(\gamma_t f_\mu)_I\right) \qquad (4.34) \\
\lambda_J &= -(\mu - r\mathbb{1})_J - (1-\alpha)(\gamma_t f_\mu)_J \\
&\quad + \Sigma_{JI}(\Sigma_{II})^{-1}\left((\mu - r\mathbb{1})_I + \bar{\lambda}\mathbb{1}_I + (1-\alpha)(\gamma_t f_\mu)_I\right)
\end{aligned}$$

The former HJB-equation (4.28) becomes with this ansatz and $\lambda = \varphi^I(\tilde{\lambda}) = \begin{pmatrix}\bar{\lambda}\mathbb{1}_I \\ \lambda_J\end{pmatrix}$:

$$\begin{aligned}
f_t &= -\frac{\alpha}{1-\alpha}(r - \bar{\lambda}) - \frac{1}{2}\frac{\alpha}{(1-\alpha)^2}(\mu - r\mathbb{1} + \lambda)^\top\Sigma^{-1}(\mu - r\mathbb{1} + \lambda) \\
&\quad - \frac{\alpha}{1-\alpha}(\mu - r\mathbb{1} + \lambda)^\top\Sigma^{-1}\gamma_t f_\mu - \frac{1}{2}f_\mu^\top\gamma_t\Sigma^{-1}\gamma_t f_\mu - \frac{1}{2}\mathrm{tr}(f_{\mu\mu}\cdot\gamma_t\Sigma^{-1}\gamma_t)
\end{aligned} \tag{4.35}$$

We solve (4.35) with a quadratic ansatz for f, where the matrix $A(t)$ is assumed to be symmetric.

$$f(t,\mu) = \frac{1}{2}\frac{1}{1-\alpha}(\mu - r\mathbb{1})^\top A(t)(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}b(t)^\top(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}k(t)$$

This ansatz needs the boundary conditions $A(T) = 0$, $b(T) = 0$ and $k(T) = 0$ and results in the following derivatives of f:

$$f_t(t,\mu) = \frac{1}{2}\frac{1}{1-\alpha}(\mu - r\mathbb{1})^\top A'(t)(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}b'(t)^\top(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}k'(t)$$

$$
\begin{aligned}
f_\mu(t,\mu) &= \frac{1}{1-\alpha}A(t)(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}b(t)\\
f_{\mu\mu}(t,\mu) &= \frac{1}{1-\alpha}A(t)
\end{aligned}
$$

The optimal λ from (4.34) becomes using this ansatz:

$$
\begin{aligned}
\bar\lambda &= (1-\alpha)g^I - g^I \begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top \big((Id+\gamma_t A(t))(\mu - r\mathbb{1}) + \gamma_t b(t)\big)\\
\lambda_J &= -\big((Id+\gamma_t A(t))(\mu - r\mathbb{1}) + \gamma_t b(t)\big)_J\\
&\quad + \Sigma_{JI}(\Sigma_{II})^{-1}\left(\bar\lambda\mathbb{1}_I + \big((Id+\gamma_t A(t))(\mu - r\mathbb{1}) + \gamma_t b(t)\big)_I\right)\\
\Rightarrow \quad \lambda &= \begin{pmatrix}\bar\lambda\mathbb{1}_I\\ \lambda_J\end{pmatrix} = (\mathbb{J}^I + Id)\bar\lambda\mathbb{1} + \mathbb{J}^I\big((Id+\gamma_t A(t))(\mu - r\mathbb{1}) + \gamma_t b(t)\big) \qquad (4.36)
\end{aligned}
$$

where J^I is given in (4.21).
Now the HJB-equation (4.35) becomes:

$$
\begin{aligned}
&\frac{1}{2}\frac{1}{1-\alpha}(\mu - r\mathbb{1})^\top A'(t)(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}b'(t)^\top(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}k'(t)\\
= f_t &= -\frac{\alpha}{1-\alpha}(r-\bar\lambda) - \frac{1}{2}\frac{\alpha}{(1-\alpha)^2}(\mu - r\mathbb{1} + \lambda)^\top\Sigma^{-1}(\mu - r\mathbb{1}+\lambda)\\
&\quad - \frac{\alpha}{1-\alpha}(\mu - r\mathbb{1}+\lambda)^\top\Sigma^{-1}\gamma_t f_\mu - \frac{1}{2}f_\mu^\top\gamma_t\Sigma^{-1}\gamma_t f_\mu - \frac{1}{2}\mathrm{tr}(f_{\mu\mu}\cdot\gamma_t\Sigma^{-1}\gamma_t)\\
&= -\frac{\alpha}{1-\alpha}r + \alpha g^I\\
&\quad - \frac{\alpha}{1-\alpha}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top (Id+\gamma_t A(t))(\mu - r\mathbb{1})\\
&\quad - \frac{\alpha}{1-\alpha}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top \gamma_t b(t)\\
&\quad - \frac{1}{2}\frac{\alpha}{(1-\alpha)^2}(\mu - r\mathbb{1})^\top\left(Id + \mathbb{J}^I(Id+\gamma_t A(t)) - g^I(\mathbb{J}^I+Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top (Id+\gamma_t A(t))\right)^\top\\
&\qquad \cdot\Sigma^{-1}\left(Id + \mathbb{J}^I(Id+\gamma_t A(t)) - g^I(\mathbb{J}^I+Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top (Id+\gamma_t A(t))\right)(\mu - r\mathbb{1})\\
&\quad - \frac{\alpha}{(1-\alpha)^2}(\mu - r\mathbb{1})^\top\left(Id + \mathbb{J}^I(Id+\gamma_t A(t)) - g^I(\mathbb{J}^I+Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top (Id+\gamma_t A(t))\right)^\top\\
&\qquad \cdot\Sigma^{-1}\left((1-\alpha)g^I(\mathbb{J}^I+Id)\mathbb{1} - g^I(\mathbb{J}^I+Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top \gamma_t b(t) + \mathbb{J}^I\gamma_t b(t)\right)\\
&\quad - \frac{1}{2}\frac{\alpha}{(1-\alpha)^2}\left((1-\alpha)g^I(\mathbb{J}^I+Id)\mathbb{1} - g^I(\mathbb{J}^I+Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top \gamma_t b(t) + \mathbb{J}^I\gamma_t b(t)\right)^\top
\end{aligned}
$$

$$
\begin{aligned}
&\quad \cdot \Sigma^{-1}\left((1-\alpha)g^I(\mathbb{J}^I+Id)\mathbb{1} - g^I(\mathbb{J}^I+Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top \gamma_t b(t) + \mathbb{J}^I\gamma_t b(t)\right)\\
&-\frac{\alpha}{(1-\alpha)^2}(\mu-r\mathbb{1})^\top\left(Id-(\mathbb{J}^I+Id)\mathbb{1}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top(Id+\gamma_t A(t))+\mathbb{J}^I(Id+\gamma_t A(t))\right)^\top\\
&\quad \cdot\Sigma^{-1}\gamma_t A(t)(\mu-r\mathbb{1})\\
&-\frac{\alpha}{(1-\alpha)^2}(\mu-r\mathbb{1})^\top\left(Id-(\mathbb{J}^I+Id)\mathbb{1}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top(Id+\gamma_t A(t))+\mathbb{J}^I(Id+\gamma_t A(t))\right)^\top\\
&\quad \cdot\Sigma^{-1}\gamma_t b(t)\\
&-\frac{\alpha}{(1-\alpha)^2}\left((1-\alpha)g^I(\mathbb{J}^I+Id)\mathbb{1}-(\mathbb{J}^I+Id)\mathbb{1}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top\gamma_t b(t)+\mathbb{J}^I\gamma_t b(t)\right)^\top\\
&\quad \cdot\Sigma^{-1}\gamma_t A(t)(\mu-r\mathbb{1})\\
&-\frac{\alpha}{(1-\alpha)^2}\left((1-\alpha)g^I(\mathbb{J}^I+Id)\mathbb{1}-(\mathbb{J}^I+Id)\mathbb{1}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top\gamma_t b(t)+\mathbb{J}^I\gamma_t b(t)\right)^\top\\
&\quad \cdot\Sigma^{-1}\gamma_t b(t)\\
&-\frac{1}{2}\frac{1}{(1-\alpha)^2}(A(t)(\mu-r\mathbb{1})+b(t))^\top\gamma_t\Sigma^{-1}\gamma_t(A(t)(\mu-r\mathbb{1})+b(t))\\
&-\frac{1}{2}\frac{1}{1-\alpha}\operatorname{tr}(A(t)\gamma_t\Sigma^{-1}\gamma_t)
\end{aligned}
$$

We continue solving this HJB-equation as in Case 3 with $B(t) := Id+\gamma_t A(t)$, respectively $A(t) = \gamma_t^{-1}(B(t) - Id)$ and with corresponding boundary condition $B(T) = Id$. Additionally we define $c(t) := \gamma_t b(t)$, respectively $b(t) = \gamma_t^{-1}c(t)$ with resulting boundary condition $c(T) = 0$.
Since $A'(t) = \Sigma^{-1}(B(t)-Id)+\gamma_t^{-1}B'(t)$ and $b'(t) = \Sigma^{-1}c(t)+\gamma_t^{-1}c'(t)$ the HJB-equation becomes:

$$
\begin{aligned}
&\frac{1}{2}(\mu-r\mathbb{1})^\top\left(\Sigma^{-1}(B(t)-Id)+\gamma_t^{-1}B'(t)\right)(\mu-r\mathbb{1})+(\mu-r\mathbb{1})^\top\left(\Sigma^{-1}c(t)+\gamma_t^{-1}c'(t)\right)+k'(t)\\
&=\frac{1}{2}(\mu-r\mathbb{1})^\top A'(t)(\mu-r\mathbb{1})+b'(t)^\top(\mu-r\mathbb{1})+k'(t)\\
&=-\alpha r+\alpha(1-\alpha)g^I\\
&\quad-\alpha g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top B(t)(\mu-r\mathbb{1})\\
&\quad-\alpha g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top c(t)\\
&\quad-\frac{1}{2}\frac{\alpha}{1-\alpha}(\mu-r\mathbb{1})^\top\left(Id+\mathbb{J}^I B(t)-g^I(\mathbb{J}^I+Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top B(t)\right)^\top
\end{aligned}
$$

$$
\begin{aligned}
&\cdot \Sigma^{-1}\left(Id + \mathbb{J}^I B(t) - g^I(\mathbb{J}^I + Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top B(t)\right)(\mu - r\mathbb{1})\\
&- \frac{\alpha}{1-\alpha}(\mu - r\mathbb{1})^\top \left(Id + \mathbb{J}^I B(t) - g^I(\mathbb{J}^I + Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top B(t)\right)^\top\\
&\cdot \Sigma^{-1}\left((1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} - g^I(\mathbb{J}^I + Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top c(t) + \mathbb{J}^I c(t)\right)\\
&- \frac{1}{2}\frac{\alpha}{1-\alpha}\left((1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} - g^I(\mathbb{J}^I + Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top c(t) + \mathbb{J}^I c(t)\right)^\top\\
&\cdot \Sigma^{-1}\left((1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} - g^I(\mathbb{J}^I + Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top c(t) + \mathbb{J}^I c(t)\right)\\
&- \frac{\alpha}{1-\alpha}(\mu - r\mathbb{1})^\top \left(Id - (\mathbb{J}^I + Id)\mathbb{1}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top B(t) + \mathbb{J}^I B(t)\right)^\top\\
&\cdot \Sigma^{-1}(B(t) - Id)(\mu - r\mathbb{1})\\
&- \frac{\alpha}{1-\alpha}(\mu - r\mathbb{1})^\top \left(Id - (\mathbb{J}^I + Id)\mathbb{1}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top B(t) + \mathbb{J}^I B(t)\right)^\top \Sigma^{-1}c(t)\\
&- \frac{\alpha}{1-\alpha}(\mu - r\mathbb{1})^\top (B(t) - Id)^\top\\
&\cdot \Sigma^{-1}\left((1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} - (\mathbb{J}^I + Id)\mathbb{1}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top c(t) + \mathbb{J}^I c(t)\right)\\
&- \frac{\alpha}{1-\alpha}\left((1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} - (\mathbb{J}^I + Id)\mathbb{1}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top c(t) + \mathbb{J}^I c(t)\right)^\top \Sigma^{-1}c(t)\\
&- \frac{1}{2}\frac{1}{1-\alpha}((B(t) - Id)(\mu - r\mathbb{1}) + c(t))^\top \Sigma^{-1}((B(t) - Id)(\mu - r\mathbb{1}) + c(t))\\
&- \frac{1}{2}\mathrm{tr}\left((B(t) - Id)\gamma_t\Sigma^{-1}\right)
\end{aligned}
$$

Now there are only terms that are either quadratic, linear or constant in $(\mu - r\mathbb{1})$, hence the HJB-equation can be split up into the ordinary differential equations (4.37), (4.40) and (4.42), corresponding to its quadratic, linear and constant part.

Solving the ordinary differential equations

First we consider the quadratic terms leading to ODE (4.37):

$$
\begin{aligned}
&\Sigma^{-1}(B(t) - Id) + \gamma_t^{-1}B'(t)\\
&\qquad = -\frac{\alpha}{1-\alpha}\left(Id + \mathbb{J}^I B(t) - g^I(\mathbb{J}^I + Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top B(t)\right)^\top
\end{aligned}
$$

$$
\begin{aligned}
&\cdot \Sigma^{-1}\left(Id + \mathbb{J}^I B(t) - g^I(\mathbb{J}^I + Id)\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top B(t)\right)\\
&- 2\frac{\alpha}{1-\alpha}\left(Id - (\mathbb{J}^I + Id)\mathbb{1}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top B(t) + \mathbb{J}^I B(t)\right)^\top\\
&\cdot \Sigma^{-1}(B(t) - Id)\\
&- \frac{1}{1-\alpha}(B(t) - Id)^\top \Sigma^{-1}(B(t) - Id)\\
\Rightarrow \quad \gamma_t^{-1}B'(t) = &- \frac{\alpha}{1-\alpha}B(t)^\top\left(Id - g^I\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top\right)^\top\\
&\cdot (\mathbb{J}^I + Id)^\top\Sigma^{-1}(\mathbb{J}^I + Id)\left(Id - g^I\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top\right)B(t)\\
&- B(t)^\top\Sigma^{-1}(B(t) - Id)
\end{aligned} \tag{4.37}
$$

Similarly to (4.23) we define (4.38) in Case $(4, I)$:

$$
\begin{aligned}
\Sigma_{\mathbb{J}}^{(4,I)} &:= \left(Id - g^I\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top\right)^\top(\mathbb{J}^I + Id)^\top\Sigma^{-1}(\mathbb{J}^I + Id)\left(Id - g^I\mathbb{1}\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\ 0_J\end{pmatrix}^\top\right)\\
&= \begin{pmatrix}\bar{\Sigma}^I & 0_{IJ}\\ 0_{JI} & 0_{JJ}\end{pmatrix}
\end{aligned} \tag{4.38}
$$

$$
\text{where} \quad \bar{\Sigma}^I = (\Sigma_{II})^{-1}\left(Id_I - g^I\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\right) \quad \text{is symmetric.}
$$

Hence (4.37) simplifies to

$$
\gamma_t^{-1}B'(t) = -\frac{\alpha}{1-\alpha}B(t)^\top\Sigma_{\mathbb{J}}^{(4,I)}B(t) - B(t)^\top\Sigma^{-1}(B(t) - Id)
$$

Now we define $C(t) := B(t)\gamma_t$, hence $B(t) = C(t)\gamma_t^{-1}$ with corresponding boundary condition $C(T) = \gamma_T$. With derivative $B'(t) = C(t)\Sigma^{-1} + C'(t)\gamma_t^{-1}$ we observe

$$
C'(t) = -C(t)^\top\left(\Sigma^{-1} + \frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)}\right)C(t) \tag{4.39}
$$

To solve this remaining IDE we consider the ansatz $D(t) = C^{-1}(t)$ with corresponding boundary condition $D(T) = \gamma_T^{-1}$. Then $C'(t) = -D^{-1}(t)D'(t)D^{-1}(t)$ and the ODE becomes

$$
-D^{-1}(t)D'(t)D^{-1}(t) = C'(t) = -D^{-1}(t)\left(\frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)} + \Sigma^{-1}\right)D^{-1}(t)
$$

Hence $D'(t) = \frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)} + \Sigma^{-1}$ and the boundary condition leads to the resulting

$$
D(t) = \gamma_T^{-1} - \Sigma^{-1}(T - t) - \frac{\alpha}{1-\alpha}(T - t)\Sigma_{\mathbb{J}}^{(4,I)}
$$

$$\text{and} \quad C(t) = \left(\gamma_T^{-1} - (T-t)\left(\Sigma^{-1} + \frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)}\right)\right)^{-1}$$
$$= \left(\gamma_t^{-1} - \frac{\alpha}{1-\alpha}(T-t)\Sigma_{\mathbb{J}}^{(4,I)}\right)^{-1}$$

Indeed $C(t)$ is symmetric.

Secondly we consider the terms linear in $(\mu - r\mathbb{1})$.

$$\begin{aligned}
&\Sigma^{-1}c(t) + \gamma_t^{-1}c'(t) \\
&\quad = -\alpha g^I B(t)^\top \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} \\
&\quad - \frac{\alpha}{1-\alpha}\left(Id + \mathbb{J}^I B(t) - g^I(\mathbb{J}^I + Id)\mathbb{1}\begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top B(t)\right)^\top \\
&\qquad \cdot \Sigma^{-1}\left((1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} - g^I(\mathbb{J}^I + Id)\mathbb{1}\begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) + \mathbb{J}^I c(t)\right) \\
&\quad - \frac{\alpha}{1-\alpha}\left(Id - (\mathbb{J}^I + Id)\mathbb{1}g^I\begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top B(t) + \mathbb{J}^I B(t)\right)^\top \Sigma^{-1}c(t) \\
&\quad - \frac{\alpha}{1-\alpha}(B(t) - Id)^\top \\
&\qquad \cdot \Sigma^{-1}\left((1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} - (\mathbb{J}^I + Id)\mathbb{1}g^I\begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) + \mathbb{J}^I c(t)\right) \\
&\quad - \frac{1}{1-\alpha}(B(t) - Id)^\top \Sigma^{-1}c(t)
\end{aligned}$$

This leads to ODE (4.40).

$$\gamma_t^{-1}c'(t) = -\alpha g^I B(t)^\top \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} - B(t)^\top\left(\frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)} + \Sigma^{-1}\right)c(t) \tag{4.40}$$

When using $C(t) = B(t)\gamma_t$ from above we observe a linear ODE for $c(t)$.

$$c'(t) = -\alpha g^I C(t)\begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} - C(t)\left(\frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)} + \Sigma^{-1}\right)c(t)$$

Since $C'(t) = -C(t)^\top\left(\Sigma^{-1} + \frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)}\right)C(t)$ by (4.39) we get the solution for $c(t)$ immediately with boundary condition $c(T) = 0$.

$$c(t) = \alpha(T-t)g^I C(t)\begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} \tag{4.41}$$

Finally we consider the terms that are constant in $(\mu - r\mathbb{1})$ and lead to ODE (4.42).

$$\begin{aligned}
k'(t) &= -\alpha r + \alpha(1-\alpha)g^I - \alpha g^I \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) \\
&\quad - \frac{1}{2}\frac{\alpha}{1-\alpha}\left((1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} - g^I(\mathbb{J}^I + Id)\mathbb{1} \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) + \mathbb{J}^I c(t) \right)^\top \\
&\qquad \cdot \Sigma^{-1} \left((1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} - g^I(\mathbb{J}^I + Id)\mathbb{1} \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) + \mathbb{J}^I c(t) \right) \\
&\quad - \frac{\alpha}{1-\alpha}\left((1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} - (\mathbb{J}^I + Id)\mathbb{1} g^I \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) + \mathbb{J}^I c(t) \right)^\top \\
&\qquad \cdot \Sigma^{-1} c(t) \\
&\quad - \frac{1}{2}\frac{1}{1-\alpha} c(t)^\top \Sigma^{-1} c(t) \\
&\quad - \frac{1}{2}\operatorname{tr}\left((B(t) - Id)\gamma_t \Sigma^{-1} \right) \\
&= -\alpha r + \frac{1}{2}\alpha(1-\alpha)g^I - \alpha g^I \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) \\
&\quad - \frac{1}{2} c(t)^\top \left(\Sigma^{-1} + \frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)} \right) c(t) - \frac{1}{2}\operatorname{tr}\left((C(t) - \gamma_t)\Sigma^{-1} \right)
\end{aligned} \tag{4.42}$$

With the boundary condition $k(T) = 0$ we get

$$\begin{aligned}
k(t) &= \alpha(T-t)r - \frac{1}{2}\alpha(1-\alpha)(T-t)g^I + \int_t^T \frac{1}{2}\operatorname{tr}\left((C(s) - \gamma_s)\Sigma^{-1} \right) ds \\
&\quad + \int_t^T \alpha g^I \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(s)\ ds + \int_t^T \frac{1}{2} c(s)^\top \left(\Sigma^{-1} + \frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)} \right) c(s)\ ds
\end{aligned}$$

Here the last line can be simplified further.

$$\begin{aligned}
&\int_t^T \alpha g^I \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(s)\ ds + \int_t^T \frac{1}{2} c(s)^\top \left(\Sigma^{-1} + \frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)} \right) c(s)\ ds \\
&= \int_t^T \alpha^2 (g^I)^2 (T-s) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top C(s) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} ds \\
&\quad + \int_t^T \frac{1}{2}\alpha^2 (T-s)^2 (g^I)^2 \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top C(s) \left(\Sigma^{-1} + \frac{\alpha}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)} \right) C(s) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} ds \\
&= -\int_t^T \alpha g^I (T-s) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c'(s)\ ds \\
&\quad + \int_t^T \frac{1}{2}\alpha^2 (T-s)^2 (g^I)^2 \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top C'(s) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} ds
\end{aligned}$$

$$
\begin{aligned}
&= \alpha g^I (T-t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) - \int_t^T \alpha g^I \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(s)\, ds \\
&\quad - \frac{1}{2}\alpha^2 (T-t)^2 (g^I)^2 \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top C(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} \\
&\quad + \int_t^T \alpha^2 (T-s) (g^I)^2 \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top C(s) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} ds \\
&= \alpha g^I (T-t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) \\
&\quad - \frac{1}{2}\alpha^2 (T-t)^2 (g^I)^2 \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top C(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} \\
&= \frac{1}{2}\alpha^2 (T-t)^2 (g^I)^2 \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top C(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}
\end{aligned}
$$

Hence the solution for $k(t)$ is given by (4.43).

$$
\begin{aligned}
k(t) &= \alpha(T-t)r - \frac{1}{2}\alpha(1-\alpha)(T-t)g^I \\
&\quad + \frac{1}{2}\alpha^2 (T-t)^2 (g^I)^2 \mathbb{1}_I^\top (\Sigma_{II})^{-1} C(t)_{II} (\Sigma_{II})^{-1}\mathbb{1}_I \\
&\quad + \int_t^T \frac{1}{2}\mathrm{tr}\left((C(s) - \gamma_s)\Sigma^{-1}\right) ds \\
&= \alpha(T-t)r - \frac{1}{2}\alpha(1-\alpha)(T-t)g^I \\
&\quad + \frac{1}{2}\alpha^2 (T-t)^2 (g^I)^2 \mathbb{1}_I^\top (\Sigma_{II})^{-1} C(t)_{II} (\Sigma_{II})^{-1}\mathbb{1}_I \\
&\quad + \frac{1}{2}\int_t^T \mathrm{tr}\left(C(s)\Sigma^{-1}\right) ds - \frac{1}{2}\log\det\gamma_T^{-1} + \frac{1}{2}\log\det\gamma_t^{-1}
\end{aligned}
\tag{4.43}
$$

Here the last line follows from the chain rule as applied in the previous Section 4.2.1. Unfortunately the remaining integral cannot be calculated explicitly, it needs to be approximated numerically.

The candidate solutions of the stochastic control approach

In order to point out the underlying case we will refer to the optimal dual solution λ in Case $(4, I)$ as $\lambda^{(4,I)}(t,\mu)$.

$$
\lambda^{(4,I)}(t,\mu) = \begin{pmatrix} \bar{\lambda}^I \mathbb{1}_I \\ \lambda_J \end{pmatrix} = (\mathbb{J}^I + Id)\mathbb{1}\bar{\lambda}^I + \mathbb{J}^I C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) + \mathbb{J}^I c(t)
$$

where

$$
\bar{\lambda}^I = (1-\alpha)g^I - g^I \mathbb{1}_I^\top (\Sigma_{II})^{-1} (C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}))_I - g^I \mathbb{1}_I^\top (\Sigma_{II})^{-1} c(t)_I
$$

$$
\begin{aligned}
C^{(4,I)}(t) &= \left(\gamma_t^{-1} - \frac{\alpha}{1-\alpha}(T-t)\Sigma_{\mathbb{J}}^{(4,I)}\right)^{-1} \\
c(t) &= \alpha(T-t)g^I C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} \\
\mathbb{J}^I &= \begin{pmatrix} 0_{II} & 0_{IJ} \\ \Sigma_{JI}(\Sigma_{II})^{-1} & -Id_J \end{pmatrix}
\end{aligned}
$$

The formula for $\lambda^{(4,I)}(t,\mu)$ can though be simplified further.

$$
\lambda^{(4,I)}(t,\mu) = \mathbb{J}^{(4,I)}\left(C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) + c(t)\right) + (1-\alpha)g^I\left(\mathbb{J}^I + Id\right)\mathbb{1} \tag{4.44}
$$

where

$$
\begin{aligned}
\mathbb{J}^{(4,I)} &:= \mathbb{J}^I - (\mathbb{J}^I + Id)\mathbb{1}g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}^\top \\
&= \begin{pmatrix} -g^I\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1} & 0_{IJ} \\ \Sigma_{JI}(\Sigma_{II})^{-1}\left(Id_I - g^I\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\right) & -Id_J \end{pmatrix}
\end{aligned}
$$

Note in particular the relation ship between $\mathbb{J}^{(4,I)}$ and $\Sigma_{\mathbb{J}}^{(4,I)}$.

$$
\left(Id + \mathbb{J}^{(4,I)}\right)^\top \Sigma^{-1}\left(Id + \mathbb{J}^{(4,I)}\right) = \Sigma_{\mathbb{J}}^{(4,I)} \tag{4.45}
$$

Now we only need to verify the stochastic control approach in Case $(4,I)$.

Theorem 4.17 (Verification in Case $(4,I)$)
Define the candidates for value function and optimal control as derived in this section.

$$
\begin{aligned}
\bar{V}(t,h,\mu) &:= h^{\frac{\alpha}{\alpha-1}}e^{f(t,\mu)} \\
\bar{\lambda}(t,h,\mu) &:= \lambda^{(4,I)}(t,\mu)
\end{aligned}
$$

Then they solve the stochastic control approach (4.12) in Case $(3,I)$, hence $\bar{V} = V$ and $\bar{\lambda} = \nu^*$, where

$$
\begin{aligned}
V(t,h,\mu) &:= \sup_{\nu\in\mathcal{D}(t,h,\mu)} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\,\middle|\,(H_t^\nu,\hat{\mu}_t) = (h,\mu)\right] \\
\text{and}\quad \nu^* &:= \arg\max_{\nu\in\mathcal{D}} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\right]
\end{aligned}
$$

Proof. Let $(t,h,\mu) \in [0,T) \times \mathbb{R}^+ \times \mathbb{R}^d$, $\nu \in \mathcal{D}$ and $s \in [t,T)$ be arbitrary and let τ be an arbitrary stopping time with values in $[t,T]$.
As $\bar{V} \in C^{1,2,2}$ we get with Itô's formula

$$
\bar{V}(s\wedge\tau, H_{s\wedge\tau}^\nu, \hat{\mu}_{s\wedge\tau}) = \bar{V}(t,h,\mu) + \int_t^{s\wedge\tau} \bar{V}_t(u) + \mathcal{L}^\nu\bar{V}(u)du \tag{4.46}
$$

$$+\int_t^{s\wedge\tau} \bar{V}_\mu(u)^\top \gamma_u(\sigma^{-1})^\top - \bar{V}_h(u)H_u^\nu(\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top(\sigma^{-1})^\top dV_u$$

$$\text{where}\quad \mathcal{L}^\nu \bar{V}(u) = -\bar{V}_h(u)H_u^\nu\left(r + \delta(\nu_u)\right) + \frac{1}{2}(H_u^\nu)^2(\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top\Sigma^{-1}(\hat{\mu}_u - r\mathbb{1} + \nu_u)V_{hh}(u) + H_u^\nu(\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top\Sigma^{-1}\gamma_u V_{h\mu}(u) + \frac{1}{2}\mathrm{tr}\big(V_{\mu\mu}(u)\cdot\gamma_u\Sigma^{-1}\gamma_u\big)$$

Here we use the notation $\bar{V}(u) := \bar{V}(u, H_u^\nu, \hat{\mu}_u)$.
While most sub-indices indicate time-dependency of the corresponding variables, sub-indices of $\bar{V}$ are the respective derivatives of $\bar{V}(t, h, \mu)$.
Now for each $n \in \mathbb{N}$ define the stopping time

$$\tau_n := \inf\left\{ s \geq t \;\middle|\; \int_t^s \left\|\bar{V}_\mu(u)^\top\gamma_u(\sigma^{-1})^\top\right\|_2^2 du \geq n \right.$$
$$\left. \text{or} \int_t^s \left\|\bar{V}_h(u)H_u^\nu(\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top(\sigma^{-1})^\top\right\|_2^2 du \geq n \right\} \wedge T$$

Hence $\tau_n \to T$ as $n \to \infty$ and therefore for any $n \in \mathbb{N}$ the stopped process

$$\int_t^{s\wedge\tau} \bar{V}_\mu(u)^\top\gamma_u(\sigma^{-1})^\top - \bar{V}_h(u)H_u^\nu(\hat{\mu}_u - r\mathbb{1} + \nu_u)^\top(\sigma^{-1})^\top dV_u$$

is a martingale with zero expectation.
Now we use τ_n in (4.46) and take conditional expectations on both sides.

$$\mathbb{E}\left[\bar{V}(s\wedge\tau_n, H_{s\wedge\tau_n}^\nu, \hat{\mu}_{s\wedge\tau_n})\middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu)\right]$$
$$= \bar{V}(t,h,\mu) + \mathbb{E}\left[\int_t^{s\wedge\tau_n} \bar{V}_t(u, H_u^\nu, \hat{\mu}_u) + \mathcal{L}^\nu\bar{V}(u, H_u^\nu, \hat{\mu}_u)du \middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu)\right]$$

Since $\bar{V}(t,h,\mu)$ satisfies the HJB-equation (4.13) in Case $(3, I)$, we observe for the above right hand side:

$$\bar{V}_t(u, H_u^\nu, \hat{\mu}_u) + \mathcal{L}^\nu\bar{V}(u, H_u^\nu, \hat{\mu}_u) \leq 0 \qquad \text{for all } \nu \in \mathcal{D}$$
$$\bar{V}_t(u, H_u^\nu, \hat{\mu}_u) + \mathcal{L}^\nu\bar{V}(u, H_u^\nu, \hat{\mu}_u) = 0 \qquad \text{if } \nu = \bar{\lambda}$$

This leads for each $u \in [t, T]$ and $\nu \in \mathcal{D}(t, h, \mu)$ to the following equations:

$$\mathbb{E}\left[\bar{V}(s\wedge\tau_n, H_{s\wedge\tau_n}^\nu, \hat{\mu}_{s\wedge\tau_n})\middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu)\right] \leq \bar{V}(t,h,\mu) \qquad \text{for all } \nu \in \mathcal{D}(t,h,\mu)$$
$$\mathbb{E}\left[\bar{V}(s\wedge\tau_n, H_{s\wedge\tau_n}^\nu, \hat{\mu}_{s\wedge\tau_n})\middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu)\right] = \bar{V}(t,h,\mu) \qquad \text{if } \nu = \bar{\lambda}$$

By Lemma 4.18 we can apply dominated convergence and observe for $n \to \infty$

$$\bar{V}(t,h,\mu) \geq \mathbb{E}\left[\bar{V}(s, H_s^\nu, \hat{\mu}_s)\middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu)\right] \qquad \text{for all } \nu \in \mathcal{D}(t,h,\mu)$$
$$\bar{V}(t,h,\mu) = \mathbb{E}\left[\bar{V}(s, H_s^\nu, \hat{\mu}_s)\middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu)\right] \qquad \text{if } \nu = \bar{\lambda}$$

Since $\bar{V}(s, H_s^\nu, \hat{\mu}_s) \to \bar{V}(T, H_T^\nu, \hat{\mu}_T) = (H_T^\nu)^{\frac{\alpha}{\alpha-1}}$ for $s \to T$ we conclude by dominated convergence

$$\begin{aligned} & \bar{V}(t,h,\mu) \geq \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}} \middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu) \right] && \text{for all } \nu \in \mathcal{D}(t,h,\mu) \\ \Rightarrow \quad & \bar{V}(t,h,\mu) \geq \sup_{\nu \in \mathcal{D}(t,h,\mu)} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}} \middle| (H_t^\nu, \hat{\mu}_t) = (h,\mu) \right] = V(t,h,\mu) \end{aligned}$$

Since we get equality for $\nu = \bar{\lambda}$ we get $V(t,h,\mu) = \bar{V}(t,h,\mu)$ with optimizer $\nu^* = \bar{\lambda}$. □

Lemma 4.18
There is an integrable random variable Y with $\left|\bar{V}(s \wedge \tau_n, H_{s\wedge\tau_n}^\nu, \hat{\mu}_{s\wedge\tau_n})\right| \leq Y$.

Proof. Since $\frac{\alpha}{\alpha-1} \in (0,1)$ we observe

$$h^{\frac{\alpha}{\alpha-1}} \leq \begin{cases} 1 & \text{if } h \leq 1 \\ h & \text{if } h > 1 \end{cases} \qquad \text{hence} \qquad 0 < h^{\frac{\alpha}{\alpha-1}} \leq 1 + h$$

The continuous function f is a second order polynomial in μ, hence there are constants C_t such that

$$f(t,\mu) \leq C_t(1 + \|\mu\|_2^2)$$

The leading constants C_t may depend on time t, however $t \in [0,T]$ is in a closed interval, hence $f(\cdot,\mu)$ is uniformly continuous in t and attains its supremum as a maximum.

$$f(t,\mu) \leq C(1 + \|\mu\|_2^2)$$

for some upper bound $C \geq \max_{t\in[0,T]} C_t < \infty$.
Hence

$$\begin{aligned} & \left|\bar{V}(t,h,\mu)\right| = h^{\frac{\alpha}{\alpha-1}} e^{f(t,\mu)} \leq (1+h) e^{C(1+\|\mu\|_2^2)} \\ \Rightarrow \quad & \left|\bar{V}(s\wedge\tau_n, H_{s\wedge\tau_n}^\nu, \hat{\mu}_{s\wedge\tau_n})\right| \leq (1 + H_{s\wedge\tau_n}^\nu) \cdot e^{C(1+\|\hat{\mu}_{s\wedge\tau_n}\|_2^2)} \\ & \qquad \leq (1+M) \cdot e^{C(1+\|\hat{\mu}_{s\wedge\tau_n}\|_2^2)} =: Y \\ & \text{where} \quad M := \sup_{s\in[t,T]} H_u^\nu \end{aligned}$$

Now on the one hand $H_s^\nu > 0$ is uniformly continuous in $[t,T]$, and hence attains its supremum $M = \sup_{s\in[t,T]} H_u^\nu < \infty$.
And on the other hand $\hat{\mu}_t$ is a martingale, hence $\|\hat{\mu}_s\|_2^2$ is a submartingale since $\|\cdot\|_2^2$ is a convex function. Hence $\exp\left(C(1+\|\hat{\mu}_s\|_2^2)\right)$ is also a submartingale and by Doob's maximal inequality we get

$$\mathbb{E}\left[\sup_{s\in[t,T]} \exp\left(C(1+\|\hat{\mu}_s\|_2^2)\right)\right] \leq \left(\frac{C}{C-1}\right)^C \mathbb{E}\left[\exp\left(C(1+\|\hat{\mu}_T\|_2^2)\right)\right] < \infty$$

Hence $\mathbb{E}[Y] < \infty$. □

4.3 The auxiliary market

The main point of Chapter 4 is to solve our constrained portfolio optimization problem (2.3) under partial information as introduced in Section 2.2 and under the convex constraints (3.7).
According to Section 2.3 we can equivalently solve the unconstrained portfolio optimization problem in the auxiliary market $\mathcal{M}^\lambda$ (2.16) where the dual process λ is chosen according to Theorem 2.11:

$$\begin{aligned} dB_t^\lambda &= B_t^\lambda\big(r + \delta\left(\lambda_t\right)\big)dt \\ dS_t^\lambda &= \operatorname{diag}\left(S_t^\lambda\right)\big(\left(\hat{\mu}_t + \lambda_t + \delta(\lambda_t)\mathbb{1}\right)dt + \sigma dV_t\big) \\ \text{where}\quad \lambda &= \underset{\nu\in\mathcal{D}}{\arg\min}\ \mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^\lambda(x_0)H_T^\nu\big)\right] \end{aligned}$$

In the previous Section 4.2 we derived the optimal dual process λ. Via a reverse separation approach we observed explicit formulas for λ for each possible case.
We still stick to Assumption 4.8 and assume that λ stays in the same case during the whole investment period. We observed the following cases:

Case 1) $\lambda = 0$.
There is nothing left to solve and $\lambda^{(1,\cdot)} = 0$.

Case 2) $\lambda \in \mathbb{R}^d_{>0}$.
Then by Proposition 4.12 we get $\lambda^{(2,\cdot)}(t,\mu) = -(\mu - r\mathbb{1}) > 0$.

Case 3) $\lambda \in \mathbb{R}^d_{\geq 0} \setminus \big(\mathbb{R}^d_{>0} \cup \{0\}\big)$ with active dimensions I.
Then by Section 4.2.1 and (4.26) we get

$$\lambda^{(3,I)}(t,\mu) = \mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1})$$

Case 4) $\lambda \in \mathbb{R}^d \setminus \mathbb{R}^d_{\geq 0}$ with active dimensions I.
Then by Section 4.2.2 and (4.44) we get

$$\begin{aligned} \lambda^{(4,I)}(t,\mu) = \mathbb{J}^{(4,I)}C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) + (1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} \\ + \alpha(T-t)g^I\mathbb{J}^{(4,I)}C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} \end{aligned}$$

In Cases 1 and 2 the auxiliary market can be solved very easily. In Case 1 with $\lambda = 0$ and $\delta(\lambda) = 0$ we need to solve $\mathcal{M}^0$ which is equal to the original unconstrained market. However the optimal unconstrained solution has already been derived in Section 4.1:

$$\begin{aligned} \pi^{(1,\cdot)}(t,\mu) &= \frac{1}{1-\alpha}\Sigma^{-1}C(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \\ \text{where}\quad C(t) &= \left(\gamma_T^{-1} - \frac{1}{1-\alpha}(T-t)\Sigma^{-1}\right)^{-1} \end{aligned}$$

In Case 2 with $\lambda = -(\mu - r\mathbb{1}) > 0$ and $\delta(\lambda) = 0$ the auxiliary market simplifies to all stocks having the same drift r as the bond. Since there is no risk in the bond, the optimal strategy results to be trivial:

$$\pi^{(2,\cdot)}(t,\mu) = 0$$

4.3.1 The auxiliary market in the Cases 3 and 4

Solving the auxiliary market for the Cases $(3, I)$ and $(4, I)$ with active dimensions I is slightly more complicated. In the following we solve the auxiliary market $\mathcal{M}^\lambda$ for the generalized linear λ (4.47) that involves all remaining cases.

$$\begin{aligned}\lambda(t,\mu) &= \mathbb{J}_\lambda(t) \cdot (\mu - r\mathbb{1}) + \lambda_c(t) \\ \delta(\lambda(t,\mu)) &= \mathbb{J}_\delta(t) \cdot (\mu - r\mathbb{1}) + \delta_c(t)\end{aligned} \tag{4.47}$$

The unconstrained portfolio optimization problem in the auxiliary market $\mathcal{M}^\lambda$ consists of maximizing expected utility of terminal wealth, where the wealth process X^π depends on the portfolio strategy π.

$$\pi^* = \arg\max_{\pi\in\mathcal{A}} \mathbb{E}\left[\frac{1}{\alpha}\Big(X_T^\pi\Big)^\alpha\right]$$

This problem can be solved using a stochastic control approach with value function $\bar{V}$.

$$\bar{V}(t,x,\mu) := \sup_{\pi\in\mathcal{A}(t,x,\mu)} \mathbb{E}\left[\frac{1}{\alpha}\Big(X_T^\pi\Big)^\alpha \,\middle|\, (X_t^\pi, \hat{\mu}_t) = (x,\mu)\right] \tag{4.48}$$

Here $\mathcal{A}(t,x,\mu)$ is the set of admissible strategies $\pi \in \mathcal{A}$ such that $(X_t^\pi, \hat{\mu}_t) = (x,\mu)$.
The stochastic control problem is to determine $\bar{V}(0,1,\mu_0)$ and the optimal portfolio strategy π^* such that

$$\bar{V}(0,1,\mu_0) = \sup_{\pi\in\mathcal{A}(0,1,\mu_0)} \mathbb{E}\left[\frac{1}{\alpha}\Big(X_T^\pi\Big)^\alpha \,\middle|\, (X_t^\pi, \hat{\mu}_t) = (x,\mu)\right] = \mathbb{E}\left[\frac{1}{\alpha}\Big(X_T^{\pi^*}\Big)^\alpha\right]$$

The controlled process in this stochastic control approach consists of the wealth process X^π together with the filter $\hat{\mu}$ such that the controlled process is Markov.

$$d\begin{pmatrix} X_t^\pi \\ \hat{\mu}_t \end{pmatrix} = \begin{pmatrix} X_t^\pi \left(r + \delta(\lambda_t) + \pi_t^\top (\hat{\mu}_t - r\mathbb{1} + \lambda_t)\right) \\ 0 \end{pmatrix} dt + \begin{pmatrix} X_t^\pi \pi_t^\top \sigma \\ \gamma_t (\sigma^{-1})^\top \end{pmatrix} dV_t$$

If the Bellmann principle holds for $t_1 > t$ we get

$$\bar{V}(t,x,\mu) = \sup_{\pi\in\mathcal{A}(t,x,\mu)} \mathbb{E}\left[\bar{V}(t_1, X_{t_1}^\pi, \hat{\mu}_{t_1}) \,\middle|\, (X_t^\pi, \hat{\mu}_t) = (x,\mu)\right]$$

With Itô's formula we get

$$\bar{V}(t,x,\mu) = \sup_{\pi\in\mathcal{A}(t,x,\mu)} \mathbb{E}\Bigg[\bar{V}(t) + \int_t^{t_1} \bar{V}_t(s) + \bar{V}_x(s) X_s^\pi \Big(r + \delta(\lambda_s) + \pi_s^\top (\hat{\mu}_s - r\mathbb{1} + \lambda_s)\Big)\, ds$$

$$+\int_t^{t_1} \bar{V}_x(s) X_s^\pi \pi_s^\top \sigma + \bar{V}_\mu(s)^\top \gamma_s (\sigma^{-1})^\top dV_s$$
$$+\frac{1}{2}\int_t^{t_1} \operatorname{tr}\big(H\bar{V}(s)\cdot a(s)\big)\, ds \bigg| (X_t^\pi, \hat{\mu}_t) = (x,\mu)\bigg]$$

where we use the notation $\bar{V}(s) := \bar{V}(s, X_s^\pi, \hat{\mu}_s)$ with derivatives $\bar{V}_t, \bar{V}_x, \bar{V}_\mu$ and Hessian matrix $H\bar{V}$ of the value function. $a(s)$ is called the diffusion matrix.

$$\begin{aligned}
a(s) &:= \begin{pmatrix} X_s^\pi \pi_s^\top \sigma \\ \gamma_s(\sigma^{-1})^\top \end{pmatrix} \begin{pmatrix} X_s^\pi \pi_s^\top \sigma \\ \gamma_s(\sigma^{-1})^\top \end{pmatrix}^\top = \begin{pmatrix} X_s^\pi \pi_s^\top \Sigma \pi_s X_s^\pi & X_s^\pi \pi_s^\top \gamma_s \\ \gamma_s \pi_s X_s^\pi & \gamma_s \Sigma^{-1} \gamma_s \end{pmatrix} \\
\Rightarrow \quad \operatorname{tr}\big(H\bar{V}(s)\cdot a(s)\big) &= \bar{V}_{xx} a_{11} + \bar{V}_{x\mu}^\top a_{21} + \operatorname{tr}\big(\bar{V}_{\mu x} a_{12} + \bar{V}_{\mu\mu} a_{22}\big) \\
&= \bar{V}_{xx} a_{11} + 2 a_{12} \bar{V}_{\mu x} + \operatorname{tr}\big(\bar{V}_{\mu\mu} a_{22}\big) \\
&= (X_s^\pi)^2 \pi_s^\top \Sigma \pi_s \bar{V}_{xx} + 2 X_s^\pi \pi_s^\top \gamma_s \bar{V}_{\mu x} + \operatorname{tr}\big(\bar{V}_{\mu\mu} \gamma_s \Sigma^{-1} \gamma_s\big)
\end{aligned}$$

If $\int_t^{t_1} \bar{V}_x(s) X_s^\pi \pi_s^\top \sigma + \bar{V}_\mu(s)^\top \gamma_s (\sigma^{-1})^\top dV_s$ is a martingale and the usual suitable conditions hold, we observe the heuristic HJB-equation (4.49) for any fixed time t. Of course these conditions have to be verified in a verification theorem later on.

$$\begin{aligned}
0 &= \sup_{\pi \in \mathcal{A}(t,x,\mu)} \mathbb{E}\bigg[\int_t^{t_1} \bar{V}_t(s) + \bar{V}_x(s) X_s^\pi \Big(r + \delta(\lambda_s) + \pi_s^\top \left(\hat{\mu}_s - r\mathbb{1} + \lambda_s\right)\Big) ds \\
&\quad + \int_t^{t_1} \frac{1}{2}(X_s^\pi)^2 \pi_s^\top \Sigma \pi_s \bar{V}_{xx} + X_s^\pi \pi_s^\top \gamma_s \bar{V}_{\mu x} ds \\
&\quad + \int_t^{t_1} \frac{1}{2}\operatorname{tr}\left(\bar{V}_{\mu\mu}\gamma_s \Sigma^{-1}\gamma_s\right) ds \bigg| (X_t^\pi, \hat{\mu}_t) = (x,\mu)\bigg] \\
\Rightarrow \quad 0 &= \sup_{\pi \in \mathbb{R}^d} \bigg(\bar{V}_t + \bar{V}_x x\Big(r + \delta(\lambda(t,\mu)) + \pi^\top \left(\mu - r\mathbb{1} + \lambda(t,\mu)\right)\Big) \\
&\quad + \frac{1}{2}\pi^\top \Sigma \pi x^2 \bar{V}_{xx} + \pi^\top \gamma_t x \bar{V}_{x\mu} + \frac{1}{2}\operatorname{tr}\left(\bar{V}_{\mu\mu}\gamma_t \Sigma^{-1}\gamma_t\right)\bigg)
\end{aligned} \tag{4.49}$$

The HJB-equation (4.49) has to be solved with respect to the boundary condition

$$\bar{V}(T,x,\mu) = \sup_{\pi \in \mathcal{A}(T,x,\mu)} \mathbb{E}\left[\frac{1}{\alpha}\Big(X_T^\pi\Big)^\alpha \bigg| (X_T^\pi, \hat{\mu}_t) = (x,\mu)\right] = \frac{1}{\alpha}x^\alpha$$

Now we can derive the optimal control π^* that maximizes the HJB-equation (4.49) via differentiating the inner part of (4.49) with respect to π.

$$\begin{aligned}
0 &\overset{!}{=} \bar{V}_x x \left(\mu - r\mathbb{1} + \lambda(t,\mu)\right) + \bar{V}_{xx} x^2 \Sigma \pi + x \gamma_t \bar{V}_{x\mu} \\
\Rightarrow \quad \pi^* &= -\frac{\bar{V}_x}{x\bar{V}_{xx}} \Sigma^{-1}\left(\mu - r\mathbb{1} + \lambda(t,\mu)\right) - \frac{1}{x\bar{V}_{xx}}\Sigma^{-1}\gamma_t \bar{V}_{x\mu}
\end{aligned} \tag{4.50}$$

Solving the HJB-equation

We solve the HJB-equation (4.49) by plugging in π^*:

$$
\begin{aligned}
0 &= \bar{V}_t + (r + \delta(\lambda(t,\mu)))x\bar{V}_x \\
&\quad - (\mu - r\mathbb{1} + \lambda(t,\mu))^\top \Sigma^{-1} (\mu - r\mathbb{1} + \lambda(t,\mu)) \frac{\bar{V}_x^2}{\bar{V}_{xx}} \\
&\quad - (\mu - r\mathbb{1} + \lambda(t,\mu))^\top \Sigma^{-1} \gamma_t \frac{\bar{V}_x}{\bar{V}_{xx}} \bar{V}_{x\mu} \\
&\quad + \frac{1}{2}\frac{1}{\bar{V}_{xx}} \left(\bar{V}_x (\mu - r\mathbb{1} + \lambda(t,\mu))^\top + \bar{V}_{x\mu}^\top \gamma_t\right) \Sigma^{-1} \big((\mu - r\mathbb{1} + \lambda(t,\mu)) \bar{V}_x + \gamma_t \bar{V}_{x\mu}\big) \\
&\quad - \frac{\bar{V}_x}{\bar{V}_{xx}} (\mu - r\mathbb{1} + \lambda(t,\mu))^\top \Sigma^{-1} \gamma_t \bar{V}_{x\mu} \\
&\quad - \frac{1}{\bar{V}_{xx}} \bar{V}_{x\mu}^\top \gamma_t \Sigma^{-1} \gamma_t \bar{V}_{x\mu} + \frac{1}{2} \operatorname{tr}\left(\bar{V}_{\mu\mu} \gamma_t \Sigma^{-1} \gamma_t\right) \\
&= \bar{V}_t + (r + \delta(\lambda(t,\mu)))x\bar{V}_x \\
&\quad - \frac{1}{2} (\mu - r\mathbb{1} + \lambda(t,\mu))^\top \Sigma^{-1} (\mu - r\mathbb{1} + \lambda(t,\mu)) \frac{\bar{V}_x^2}{\bar{V}_{xx}} \\
&\quad - \frac{1}{2}\frac{1}{\bar{V}_{xx}} \bar{V}_{x\mu}^\top \gamma_t \Sigma^{-1} \gamma_t \bar{V}_{x\mu} \\
&\quad - \frac{\bar{V}_x}{\bar{V}_{xx}} (\mu - r\mathbb{1} + \lambda(t,\mu))^\top \Sigma^{-1} \gamma_t \bar{V}_{x\mu} \\
&\quad + \frac{1}{2} \operatorname{tr}\left(\bar{V}_{\mu\mu} \gamma_t \Sigma^{-1} \gamma_t\right)
\end{aligned}
$$

Now we use the multiplicative ansatz

$$\bar{V}(t,x,\mu) = U(x) e^{\bar{f}(t,\mu)} = \frac{x^\alpha}{\alpha} e^{\bar{f}(t,\mu)}$$

with corresponding boundary condition $\bar{f}(T,\mu) = 0$.
This simplifies the HJB-equation (4.49) as follows:

$$
\begin{aligned}
0 &= \bar{f}_t + \alpha(r + \delta(\lambda(t,\mu))) \\
&\quad - \frac{1}{2}\frac{\alpha}{(\alpha - 1)} (\mu - r\mathbb{1} + \lambda(t,\mu))^\top \Sigma^{-1} (\mu - r\mathbb{1} + \lambda(t,\mu)) \\
&\quad - \frac{1}{2}\frac{1}{\alpha - 1} \bar{f}_\mu^\top \gamma_t \Sigma^{-1} \gamma_t \bar{f}_\mu \\
&\quad - \frac{\alpha}{\alpha - 1} (\mu - r\mathbb{1} + \lambda(t,\mu))^\top \Sigma^{-1} \gamma_t \bar{f}_\mu \\
&\quad + \frac{1}{2} \operatorname{tr}\left(\bar{f}_{\mu\mu} \gamma_t \Sigma^{-1} \gamma_t\right)
\end{aligned}
$$

We improve the above ansatz further, where the matrix $\bar{A}(t)$ is assumed to be symmetric.

$$\bar{f}(t,\mu) = \frac{1}{2}(\mu - r\mathbb{1})^\top \bar{A}(t)(\mu - r\mathbb{1}) + \bar{b}(t)^\top (\mu - r\mathbb{1}) + \bar{k}(t)$$

This needs the boundary conditions $\bar{A}(T) = 0$, $\bar{b}(T) = 0$ and $\bar{k}(T) = 0$.
We also observe the following derivatives of $\bar{f}$:

$$\begin{aligned}
\bar{f}_t(t,\mu) &= \frac{1}{2}(\mu - r\mathbb{1})^\top \bar{A}'(t)(\mu - r\mathbb{1}) + \bar{b}'(t)^\top(\mu - r\mathbb{1}) + \bar{k}'(t)\\
\bar{f}_\mu(t,\mu) &= \bar{A}(t)(\mu - r\mathbb{1}) + b(t)\\
\bar{f}_{\mu\mu}(t,\mu) &= \bar{A}(t)
\end{aligned}$$

Using this ansatz and the formulas for λ from (4.47) the HJB-equation (4.49) becomes:

$$\begin{aligned}
&\frac{1}{2}(\mu - r\mathbb{1})^\top \bar{A}'(t)(\mu - r\mathbb{1}) + \bar{b}'(t)^\top(\mu - r\mathbb{1}) + \bar{k}'(t)\\
= \bar{f}_t &= -\alpha r - \alpha \mathbb{J}_\delta(t)\cdot(\mu - r\mathbb{1}) - \alpha\delta_c(t)\\
&\quad + \frac{1}{2}\frac{\alpha}{\alpha-1}\left((Id + \mathbb{J}_\lambda(t))(\mu - r\mathbb{1}) + \lambda_c(t)\right)^\top \Sigma^{-1}\left((Id + \mathbb{J}_\lambda(t))(\mu - r\mathbb{1}) + \lambda_c(t)\right)\\
&\quad + \frac{1}{2}\frac{1}{\alpha-1}\big(\bar{A}(t)(\mu - r\mathbb{1}) + \bar{b}(t)\big)^\top \gamma_t\Sigma^{-1}\gamma_t\big(\bar{A}(t)(\mu - r\mathbb{1}) + \bar{b}(t)\big)\\
&\quad + \frac{\alpha}{\alpha-1}\left((Id + \mathbb{J}_\lambda(t))(\mu - r\mathbb{1}) + \lambda_c(t)\right)^\top \Sigma^{-1}\gamma_t\big(\bar{A}(t)(\mu - r\mathbb{1}) + \bar{b}(t)\big)\\
&\quad - \frac{1}{2}\mathrm{tr}\left(\bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\right)\\
&= -\alpha r - \alpha \mathbb{J}_\delta(t)\cdot(\mu - r\mathbb{1}) - \alpha\delta_c(t)\\
&\quad + \frac{1}{2}\frac{\alpha}{\alpha-1}(\mu - r\mathbb{1})^\top(Id + \mathbb{J}_\lambda(t))^\top\Sigma^{-1}(Id + \mathbb{J}_\lambda(t))(\mu - r\mathbb{1})\\
&\quad + \frac{\alpha}{\alpha-1}(\mu - r\mathbb{1})^\top(Id + \mathbb{J}_\lambda(t))^\top\Sigma^{-1}\lambda_c(t)\\
&\quad + \frac{1}{2}\frac{\alpha}{\alpha-1}\lambda_c(t)^\top\Sigma^{-1}\lambda_c(t)\\
&\quad + \frac{1}{2}\frac{1}{\alpha-1}(\mu - r\mathbb{1})^\top\bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\bar{A}(t)(\mu - r\mathbb{1})\\
&\quad + \frac{1}{\alpha-1}(\mu - r\mathbb{1})^\top\bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\bar{b}(t)\\
&\quad + \frac{1}{2}\frac{1}{\alpha-1}\bar{b}(t)^\top\gamma_t\Sigma^{-1}\gamma_t\bar{b}(t)\\
&\quad + \frac{\alpha}{\alpha-1}(\mu - r\mathbb{1})^\top(Id + \mathbb{J}_\lambda(t))^\top\Sigma^{-1}\gamma_t\bar{A}(t)(\mu - r\mathbb{1})\\
&\quad + \frac{\alpha}{\alpha-1}(\mu - r\mathbb{1})^\top(Id + \mathbb{J}_\lambda(t))^\top\Sigma^{-1}\gamma_t\bar{b}(t)\\
&\quad + \frac{\alpha}{\alpha-1}(\mu - r\mathbb{1})^\top\bar{A}(t)\gamma_t\Sigma^{-1}\lambda_c(t)\\
&\quad + \frac{\alpha}{\alpha-1}\lambda_c(t)^\top\Sigma^{-1}\gamma_t\bar{b}(t)\\
&\quad - \frac{1}{2}\mathrm{tr}\left(\bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\right)
\end{aligned}$$

Now the resulting differential equation only consists of terms that are either quadratic, linear or constant in $(\mu - r\mathbb{1})$, hence it can be split up into the following three separate ordinary differential equations:

$$\begin{aligned}\bar{A}'(t) &= \frac{\alpha}{\alpha-1}(Id+\mathbb{J}_\lambda(t))^\top\Sigma^{-1}(Id+\mathbb{J}_\lambda(t)) + \frac{1}{\alpha-1}\bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\bar{A}(t)\\ &\quad + 2\frac{\alpha}{\alpha-1}(Id+\mathbb{J}_\lambda(t))^\top\Sigma^{-1}\gamma_t\bar{A}(t)\\ &= -\frac{\alpha}{1-\alpha}\Big(Id+\mathbb{J}_\lambda(t)+\gamma_t\bar{A}(t)\Big)^\top\Sigma^{-1}\Big(Id+\mathbb{J}_\lambda(t)+\gamma_t\bar{A}(t)\Big) - \bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\bar{A}(t)\end{aligned} \tag{4.51}$$

$$\begin{aligned}\bar{b}'(t) &= -\alpha\mathbb{J}_\delta(t)^\top + \frac{\alpha}{\alpha-1}(Id+\mathbb{J}_\lambda(t))^\top\Sigma^{-1}\lambda_c(t) + \frac{1}{\alpha-1}\bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\bar{b}(t)\\ &\quad + \frac{\alpha}{\alpha-1}(Id+\mathbb{J}_\lambda(t))^\top\Sigma^{-1}\gamma_t\bar{b}(t) + \frac{\alpha}{\alpha-1}\bar{A}(t)\gamma_t\Sigma^{-1}\lambda_c(t)\\ &= -\alpha\mathbb{J}_\delta(t)^\top - \frac{\alpha}{1-\alpha}\Big(Id+\mathbb{J}_\lambda(t)+\gamma_t\bar{A}(t)\Big)^\top\Sigma^{-1}\Big(\gamma_t\bar{b}(t)+\lambda_c(t)\Big) - \bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\bar{b}(t)\end{aligned} \tag{4.52}$$

$$\begin{aligned}\bar{k}'(t) &= -\alpha r - \alpha\delta_c(t) + \frac{1}{2}\frac{\alpha}{\alpha-1}\lambda_c(t)^\top\Sigma^{-1}\lambda_c(t)\\ &\quad + \frac{1}{2}\frac{1}{\alpha-1}\bar{b}(t)^\top\gamma_t\Sigma^{-1}\gamma_t\bar{b}(t) + \frac{\alpha}{\alpha-1}\lambda_c(t)^\top\Sigma^{-1}\gamma_t\bar{b}(t) - \frac{1}{2}\mathrm{tr}\left(\bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\right)\\ &= -\alpha r - \alpha\delta_c(t) - \frac{1}{2}\frac{\alpha}{1-\alpha}\Big(\gamma_t\bar{b}(t)+\lambda_c(t)\Big)^\top\Sigma^{-1}\Big(\gamma_t\bar{b}(t)+\lambda_c(t)\Big)\\ &\quad - \frac{1}{2}\bar{b}(t)^\top\gamma_t\Sigma^{-1}\gamma_t\bar{b}(t) - \frac{1}{2}\mathrm{tr}\left(\bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\right)\end{aligned} \tag{4.53}$$

These differential equations need to be solved with the correct values for λ (4.47) in each case. Then the optimal portfolio strategy as derived in (4.50) becomes:

$$\begin{aligned}\pi^*(t,x,\mu) &= -\frac{\bar{V}_x}{x\bar{V}_{xx}}\Sigma^{-1}(\mu - r\mathbb{1} + \lambda(t,\mu)) - \frac{1}{x\bar{V}_{xx}}\Sigma^{-1}\gamma_t\bar{V}_{x\mu}\\ &= \frac{1}{1-\alpha}\Sigma^{-1}(\mu - r\mathbb{1} + \lambda(t,\mu)) + \frac{1}{1-\alpha}\Sigma^{-1}\gamma_t\bar{f}_\mu\\ &= \frac{1}{1-\alpha}\Sigma^{-1}(Id+\mathbb{J}_\lambda(t)+\gamma_t\bar{A}(t))(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}\Sigma^{-1}(\gamma_t\bar{b}(t)+\lambda_c(t))\end{aligned} \tag{4.54}$$

Theorem 4.19
In Case $(3,I)$ the optimal portfolio strategy is given by

$$\pi^{(3,I)}(t,x,\mu) = \frac{1}{1-\alpha}\Sigma_{\mathbb{J}}^{(3,I)}C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1})$$

where

$$C^{(3,I)}(t) = \left(\gamma_t^{-1} - \frac{\alpha}{1-\alpha}(T-t)\Sigma_{\mathbb{J}}^{(3,I)}\right)^{-1}$$

$$\text{and}\quad \Sigma_{\mathbb{J}}^{(3,I)} = \begin{pmatrix}(\Sigma_{II})^{-1} & 0_{IJ}\\ 0_{JI} & 0_{JJ}\end{pmatrix}$$

Theorem 4.20
In Case $(4, I)$ the optimal portfolio strategy is given by

$$\begin{aligned}\pi^{(4,I)}(t,x,\mu) = &\frac{1}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)}C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \\ &+ \frac{\alpha}{1-\alpha}(T-t)g^I\Sigma_{\mathbb{J}}^{(4,I)}C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} + g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}\end{aligned}$$

where

$$\begin{aligned}C^{(4,I)}(t) &= \left(\gamma_t^{-1} - \frac{\alpha}{1-\alpha}(T-t)\Sigma_{\mathbb{J}}^{(4,I)}\right)^{-1} \\ \text{and} \quad \Sigma_{\mathbb{J}}^{(4,I)} &= \begin{pmatrix}(\Sigma_{II})^{-1}\left(Id_I - g^I\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\right) & 0_{IJ} \\ 0_{JI} & 0_{JJ}\end{pmatrix}\end{aligned}$$

Proof. (of Theorem 4.19)
In Case $(3, I)$ with λ from (4.26) we get in (4.47):

$$\begin{aligned}&\lambda^{(3,I)}(t,\mu) = \mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) &&\text{hence} \quad \mathbb{J}_\lambda(t) = \mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1} &&\text{and} \quad \lambda_c = 0 \\ &\delta(\lambda(t,\mu)) = 0 &&\text{hence} \quad \mathbb{J}_\delta = 0 &&\text{and} \quad \delta_c = 0\end{aligned}$$

Using these parameters the ODEs (4.51),(4.52) and (4.53) simplify significantly.

$$\begin{aligned}\bar{A}'(t) = &-\frac{\alpha}{1-\alpha}\left(\mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1} + Id + \gamma_t\bar{A}(t)\right)^\top \Sigma^{-1}\left(\mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1} + Id + \gamma_t\bar{A}(t)\right) \\ &- \bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\bar{A}(t) \\ \bar{b}'(t) = &-\frac{\alpha}{1-\alpha}\left(\mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1} + Id + \gamma_t\bar{A}(t)\right)^\top \Sigma^{-1}\gamma_t\bar{b}(t) - \bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\bar{b}(t) \\ \bar{k}'(t) = &-\alpha r - \frac{1}{2}\frac{\alpha}{1-\alpha}\bar{b}(t)^\top\gamma_t\Sigma^{-1}\gamma_t\bar{b}(t) \\ &- \frac{1}{2}\bar{b}(t)^\top\gamma_t\Sigma^{-1}\gamma_t\bar{b}(t) - \frac{1}{2}\operatorname{tr}\left(\bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\right)\end{aligned}$$

To solve the first ODE for $\bar{A}(t)$, note the solution $A(t)$ of ODE (4.24):

$$\begin{aligned}A(t) &= \gamma_t^{-1}(C^{(3,I)}(t)\gamma_t^{-1} - Id) \\ \Rightarrow \quad \mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1} &= \mathbb{J}^I(Id + \gamma_t A(t))\end{aligned}$$

For $\bar{A}(t) = A(t)$ we get immediately that ODEs (4.51) and (4.24) are the same, hence $\bar{A}(t) = A(t)$ is the correct solution for ODE (4.51).
The second ODE above gets solved trivially by $\bar{b}(t) = 0$ and the third ODE for $\bar{k}(t)$ results to be the same ODE as (4.25) for $k(t)$.

$$\begin{aligned}&\bar{k}'(t) = -\alpha r - \frac{1}{2}\operatorname{tr}\left(A(t)\gamma_t\Sigma^{-1}\gamma_t\right) = k'(t) \quad \text{from (4.25)} \\ \Rightarrow \quad &\bar{k}(t) = \alpha(T-t)r + \frac{1}{2}\int_t^T \operatorname{tr}\left((C(s) - \gamma_s)\Sigma^{-1}\right)ds\end{aligned}$$

$$= \alpha(T-t)r + \frac{1}{2}\int_t^T \operatorname{tr}\left(C(s)\Sigma^{-1}\right) ds - \frac{1}{2}\log\det\gamma_T^{-1} + \frac{1}{2}\log\det\gamma_t^{-1}$$

In particular we get for the optimal trading strategy (4.54)

$$\begin{aligned}
\pi^{(3,I)}(t,x,\mu) &= -\frac{\bar{V}_x}{x\bar{V}_{xx}}\Sigma^{-1}\left(\mu - r\mathbb{1} + \lambda^{(3,I)}(t,\mu)\right) - \frac{1}{x\bar{V}_{xx}}\Sigma^{-1}\gamma_t\bar{V}_{x\mu} \\
&= \frac{1}{1-\alpha}\Sigma^{-1}\left(\mu - r\mathbb{1} + \lambda^{(3,I)}(t,\mu)\right) + \frac{1}{1-\alpha}\Sigma^{-1}\gamma_t\bar{f}_\mu \\
&= \frac{1}{1-\alpha}\Sigma^{-1}(Id + \mathbb{J}_\lambda(t) + \gamma_t\bar{A}(t))(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}\Sigma^{-1}(\gamma_t\bar{b}(t) + \lambda_c(t)) \\
&= \frac{1}{1-\alpha}\Sigma^{-1}\left(Id + \mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1}\right)(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}\Sigma^{-1}\gamma_t\bar{A}(t)(\mu - r\mathbb{1}) \\
&= \frac{1}{1-\alpha}\Sigma^{-1}(Id + \mathbb{J}^I)C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \\
&= \frac{1}{1-\alpha}\Sigma_{\mathbb{J}}^{(3,I)}C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1})
\end{aligned}$$

Finally the verification in Case $(3, I)$ follows from the following Theorem 4.21. □

Proof. (of Theorem 4.20)
In Case $(4, I)$ with λ from (4.44) we get in (4.47):

$$\begin{aligned}
\lambda^{(4,I)}(t,\mu) &= \mathbb{J}^{(4,I)}C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) + (1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} \\
&\quad + \alpha(T-t)g^I\mathbb{J}^{(4,I)}C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} \\
\delta\big(\lambda^{(4,I)}(t,\mu)\big) &= g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}^\top C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \\
&\quad + \alpha(T-t)(g^I)^2\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}^\top C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} - (1-\alpha)g^I \\
\Rightarrow\quad \mathbb{J}_\lambda(t) &= \mathbb{J}^{(4,I)}C^{(4,I)}(t)\gamma_t^{-1} \\
\lambda_c &= \alpha(T-t)g^I\mathbb{J}^{(4,I)}C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} + (1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} \\
\mathbb{J}_\delta &= g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}^\top C^{(4,I)}(t)\gamma_t^{-1} \\
\delta_c &= \alpha(T-t)(g^I)^2\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}^\top C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} - (1-\alpha)g^I
\end{aligned}$$

Using these parameters the ODEs (4.51),(4.52) and (4.53) simplify significantly.

$$\begin{aligned}
\bar{A}'(t) = &-\frac{\alpha}{1-\alpha}\left(Id + \gamma_t\bar{A}(t) + \mathbb{J}^{(4,I)}C^{(4,I)}(t)\gamma_t^{-1}\right)^\top\Sigma^{-1}\left(Id + \gamma_t\bar{A}(t) + \mathbb{J}^{(4,I)}C^{(4,I)}(t)\gamma_t^{-1}\right) \\
&- \bar{A}(t)\gamma_t\Sigma^{-1}\gamma_t\bar{A}(t)
\end{aligned}$$

$$\begin{aligned}
\bar{b}'(t) &= -\alpha g^I \gamma_t^{-1} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} \\
&\quad - \frac{\alpha}{1-\alpha} \left(Id + \gamma_t \bar{A}(t) + \mathbb{J}^{(4,I)} C^{(4,I)}(t)\gamma_t^{-1} \right)^\top \Sigma^{-1}\gamma_t \bar{b}(t) - \bar{A}(t)\gamma_t \Sigma^{-1}\gamma_t \bar{b}(t) \\
&\quad - \frac{\alpha}{1-\alpha} \left(Id + \gamma_t \bar{A}(t) + \mathbb{J}^{(4,I)} C^{(4,I)}(t)\gamma_t^{-1} \right)^\top \Sigma^{-1}\alpha(T-t) g^I \mathbb{J}^{(4,I)} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} \\
&\quad - \frac{\alpha}{1-\alpha} \left(Id + \gamma_t \bar{A}(t) + \mathbb{J}^{(4,I)} C^{(4,I)}(t)\gamma_t^{-1} \right)^\top \Sigma^{-1}(1-\alpha) g^I (\mathbb{J}^I + Id)\mathbb{1} \\
\bar{k}'(t) &= -\alpha r + \alpha(1-\alpha) g^I - \alpha\alpha(T-t)(g^I)^2 \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}^\top C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} \\
&\quad - \frac{1}{2}\frac{\alpha}{1-\alpha} \left(\alpha(T-t) g^I \mathbb{J}^{(4,I)} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} + (1-\alpha) g^I (\mathbb{J}^I + Id)\mathbb{1} \right)^\top \\
&\quad \cdot \Sigma^{-1} \left(\alpha(T-t) g^I \mathbb{J}^{(4,I)} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} + (1-\alpha) g^I (\mathbb{J}^I + Id)\mathbb{1} \right) \\
&\quad - \frac{\alpha}{1-\alpha} \left(\alpha(T-t) g^I \mathbb{J}^{(4,I)} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} + (1-\alpha) g^I (\mathbb{J}^I + Id)\mathbb{1} \right)^\top \Sigma^{-1}\gamma_t \bar{b}(t) \\
&\quad - \frac{1}{2}\frac{\alpha}{1-\alpha} \bar{b}(t)^\top \gamma_t \Sigma^{-1}\gamma_t \bar{b}(t) \\
&\quad - \frac{1}{2}\bar{b}(t)^\top \gamma_t \Sigma^{-1}\gamma_t \bar{b}(t) - \frac{1}{2}\operatorname{tr}\left(\bar{A}(t)\gamma_t \Sigma^{-1}\gamma_t \right)
\end{aligned}$$

Fortunately when comparing the first ODE for $\bar{A}(t)$ with the ODEs (4.37, 4.39) for $A(t)$ in Section 4.2.2, we observe that both equations get solved for the same parameters $\bar{A}(t) = A(t)$, where $A(t) = \gamma_t^{-1}(B(t) - Id) = \gamma_t^{-1}(C^{(4,I)}(t)\gamma_t^{-1} - Id)$.
Now the second ODE for $\bar{b}(t)$ becomes:

$$\begin{aligned}
\bar{b}'(t) &= -\alpha g^I \gamma_t^{-1} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} \\
&\quad - \frac{\alpha}{1-\alpha}\gamma_t^{-1} C^{(4,I)}(t) \left(Id + \mathbb{J}^{(4,I)} \right)^\top \Sigma^{-1}\gamma_t \bar{b}(t) - \left(\gamma_t^{-1} C^{(4,I)}(t) - Id \right) \Sigma^{-1}\gamma_t \bar{b}(t) \\
&\quad - \frac{\alpha}{1-\alpha}\gamma_t^{-1} C^{(4,I)}(t) \left(Id + \mathbb{J}^{(4,I)} \right)^\top \Sigma^{-1}\alpha(T-t) g^I \mathbb{J}^{(4,I)} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} \\
&\quad - \frac{\alpha}{1-\alpha}\gamma_t^{-1} C^{(4,I)}(t) \left(Id + \mathbb{J}^{(4,I)} \right)^\top \Sigma^{-1}(1-\alpha) g^I (\mathbb{J}^I + Id)\mathbb{1} \\
&= -\alpha g^I \gamma_t^{-1} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix} \\
&\quad - \frac{\alpha}{1-\alpha}\gamma_t^{-1} C^{(4,I)}(t) \left(Id + \mathbb{J}^{(4,I)} \right)^\top \Sigma^{-1}\Big(Id \Big)\gamma_t \bar{b}(t) - \left(\gamma_t^{-1} C^{(4,I)}(t) - Id \right) \Sigma^{-1}\gamma_t \bar{b}(t) \\
&\quad - \frac{\alpha}{1-\alpha}\gamma_t^{-1} C^{(4,I)}(t) \left(Id + \mathbb{J}^{(4,I)} \right)^\top \Sigma^{-1}\Big(\mathbb{J}^{(4,I)} \Big)\alpha(T-t) g^I C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J \end{pmatrix}
\end{aligned}$$

Again when comparing this ODE to ODE (4.40) for $b(t)$ in Section 4.2.2, we observe that both equations get solved for the same parameters $\bar{b}(t) = b(t) = \gamma_t^{-1} c(t)$ with

$$c(t) = \alpha(T-t) g^I C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1} \mathbb{1}_I \\ 0_J \end{pmatrix}$$

Finally the third ODE above remains with:

$$\begin{aligned}
\bar{k}'(t) = & -\alpha r + \alpha(1-\alpha) g^I - \alpha g^I \begin{pmatrix} (\Sigma_{II})^{-1} \mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) \\
& - \frac{1}{2} \frac{\alpha}{1-\alpha} \left(\mathbb{J}^{(4,I)} c(t) + (1-\alpha) g^I (\mathbb{J}^I + Id) \mathbb{1} \right)^\top \\
& \quad \cdot \Sigma^{-1} \left(\mathbb{J}^{(4,I)} c(t) + (1-\alpha) g^I (\mathbb{J}^I + Id) \mathbb{1} \right) \\
& - \frac{\alpha}{1-\alpha} \left(\mathbb{J}^{(4,I)} c(t) + (1-\alpha) g^I (\mathbb{J}^I + Id) \mathbb{1} \right)^\top \Sigma^{-1} c(t) \\
& - \frac{1}{2} \frac{\alpha}{1-\alpha} c(t)^\top \Sigma^{-1} c(t) \\
& - \frac{1}{2} c(t)^\top \Sigma^{-1} c(t) - \frac{1}{2} \operatorname{tr} \left((C^{(4,I)}(t) - \gamma_t) \Sigma^{-1} \right) \\
= & -\alpha r + \frac{1}{2} \alpha(1-\alpha) g^I - \alpha g^I \begin{pmatrix} (\Sigma_{II})^{-1} \mathbb{1}_I \\ 0_J \end{pmatrix}^\top c(t) \\
& - \frac{1}{2} c(t)^\top \left(\Sigma^{-1} + \frac{\alpha}{1-\alpha} \Sigma_{\mathbb{J}}^{(4,I)} \right) c(t) - \frac{1}{2} \operatorname{tr} \left((C^{(4,I)}(t) - \gamma_t) \Sigma^{-1} \right) \\
= & \, k'(t)
\end{aligned}$$

Hence this is the same ODE as (4.42), such that we get $\bar{k}(t) = k(t)$ as given in (4.43). In particular we get for the optimal trading strategy (4.54)

$$\begin{aligned}
\pi^{(4,I)}(t, x, \mu) = & -\frac{\bar{V}_x}{x \bar{V}_{xx}} \Sigma^{-1} \left(\mu - r\mathbb{1} + \lambda^{(4,I)}(t, \mu) \right) - \frac{1}{x \bar{V}_{xx}} \Sigma^{-1} \gamma_t \bar{V}_{x\mu} \\
= & \frac{1}{1-\alpha} \Sigma^{-1} \left(\mu - r\mathbb{1} + \lambda^{(4,I)}(t, \mu) \right) + \frac{1}{1-\alpha} \Sigma^{-1} \gamma_t \bar{f}_\mu \\
= & \frac{1}{1-\alpha} \Sigma^{-1} (Id + \mathbb{J}_\lambda(t) + \gamma_t A(t))(\mu - r\mathbb{1}) + \frac{1}{1-\alpha} \Sigma^{-1} (\gamma_t b(t) + \lambda_c(t)) \\
= & \frac{1}{1-\alpha} \Sigma^{-1} \left(\mathbb{J}^{(4,I)} C^{(4,I)}(t) \gamma_t^{-1} + B(t) \right) (\mu - r\mathbb{1}) + \Sigma^{-1} g^I (\mathbb{J}^I + Id) \mathbb{1} \\
& + \frac{1}{1-\alpha} \Sigma^{-1} \left(c(t) + \alpha(T-t) g^I \mathbb{J}^{(4,I)} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1} \mathbb{1}_I \\ 0_J \end{pmatrix} \right) \\
= & \frac{1}{1-\alpha} \Sigma^{-1} \left(Id + \mathbb{J}^{(4,I)} \right) C^{(4,I)}(t) \gamma_t^{-1} (\mu - r\mathbb{1}) + \Sigma^{-1} g^I (\mathbb{J}^I + Id) \mathbb{1} \\
& + \frac{\alpha}{1-\alpha} (T-t) g^I \Sigma^{-1} \left(Id + \mathbb{J}^{(4,I)} \right) C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1} \mathbb{1}_I \\ 0_J \end{pmatrix} \\
= & \frac{1}{1-\alpha} \Sigma_{\mathbb{J}}^{4,I} C^{(4,I)}(t) \gamma_t^{-1} (\mu - r\mathbb{1})
\end{aligned}$$

$$+\frac{\alpha}{1-\alpha}(T-t)g^I\Sigma_{\mathbb{J}}^{4,I}C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\0_J\end{pmatrix}+g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I\\0_J\end{pmatrix}$$

Finally the verification in Case $(4,I)$ follows from Theorem 4.21. □

Theorem 4.21 (Verification of the auxiliary market in the Cases 3 and 4)
Define the candidates $\bar{V}$ and $\bar{\pi}$ for value function and optimal control as derived in the previous theorems.

$$\begin{aligned}\bar{V}(t,x,\mu)&:=\frac{x^\alpha}{\alpha}e^{\bar{f}(t,\mu)}\\\bar{\pi}(t,x,\mu)&:=\pi^{(3,I)}(t,\mu)\quad\text{resp.}\quad\pi^{(4,I)}(t,\mu)\end{aligned}$$

where $\bar{f}$ is the quadratic function given in the respective ansatz above. Then they solve the stochastic control approach (4.48) of our constrained portfolio optimization problem in the Cases 3 and 4, hence $\bar{V}=V$ and $\bar{\pi}=\pi^*$ where

$$\begin{aligned}V(t,x,\mu)&:=\sup_{\pi\in\mathcal{A}(t,x,\mu)}\mathbb{E}\left[\frac{1}{\alpha}(X_T^\pi)^\alpha\,\middle|\,(X_t^\pi,\hat{\mu}_t)=(x,\mu)\right]\\\text{and}\quad\pi^*&:=\operatorname*{arg\,max}_{\pi\in\mathcal{A}}\mathbb{E}\left[\frac{1}{\alpha}(X_T^\pi)^\alpha\right]\end{aligned}$$

Proof. Let $(t,x,\mu)\in[0,T)\times\mathbb{R}^+\times\mathbb{R}^d$, $\pi\in\mathcal{A}$ and $s\in[t,T)$ be arbitrary and let τ be an arbitrary stopping time with values in $[t,T]$.
As $\bar{V}\in C^{1,2,2}$ we get with the Itô formula

$$\begin{aligned}\bar{V}(s\wedge\tau,X_{s\wedge\tau}^\pi,\hat{\mu}_{s\wedge\tau})&=\bar{V}(t,x,\mu)+\int_t^{s\wedge\tau}\bar{V}_t(u)+\mathcal{L}^\pi\bar{V}(u)du\\&\quad+\int_t^{s\wedge\tau}\bar{V}_x(u)X_u^\pi\pi_u^\top\sigma+\bar{V}_\mu(u)^\top\gamma_u(\sigma^{-1})^\top dV_u\qquad(4.55)\\\text{where}\quad\mathcal{L}^\pi\bar{V}(u)&=\bar{V}_x(u)X_u^\pi\Big(r+\delta(\lambda_u)+\pi_u^\top(\hat{\mu}_u-r\mathbb{1}+\lambda_u)\Big)\\&\quad+\frac{1}{2}\bar{V}_{xx}(u)(X_u^\pi)^2\pi_u^\top\Sigma\pi_u+\bar{V}_{x\mu}^\top(u)\gamma_u\pi_uX_u^\pi+\frac{1}{2}\mathrm{tr}\big(\bar{V}_{\mu\mu}(u)\gamma_u\Sigma^{-1}\gamma_u\big)\end{aligned}$$

where we use the notation $\bar{V}(u):=\bar{V}(u,X_u^\pi,\hat{\mu}_u)$. While most sub-indices indicate time-dependency of the corresponding variables, sub-indices of $\bar{V}$ are the respective derivatives of $\bar{V}(t,x,\mu)$.
Now for each $n\in\mathbb{N}$ define the stopping time

$$\tau_n:=\inf\left\{s\geq t\,\middle|\,\int_t^s\left\|\sigma^{-1}\gamma_u\bar{V}_\mu(u)\right\|_2^2du\geq n\quad\text{or}\quad\int_t^s\left\|\bar{V}_x(u)X_u^\pi\sigma^\top\pi_u\right\|_2^2du\geq n\right\}\wedge T$$

Hence $\tau_n\to T$ as $n\to\infty$ and therefore for any $n\in\mathbb{N}$ the stopped process

$$\int_t^{s\wedge\tau_n}\bar{V}_x(u)X_u^\pi\pi_u^\top\sigma+\bar{V}_\mu(u)^\top\gamma_u(\sigma^{-1})^\top\,dV_u$$

is a martingale with zero expectation.
Now we use τ_n in (4.55) and take conditional expectations on both sides.

$$\begin{aligned}&\mathbb{E}\left[\bar{V}(s\wedge\tau_n, X^\pi_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\middle|\,(X^\pi_t,\hat{\mu}_t)=(x,\mu)\right]\\ &=\bar{V}(t,x,\mu)+\mathbb{E}\left[\int_t^{s\wedge\tau_n}\bar{V}_t(u,X^\pi_u,\hat{\mu}_u)+\mathcal{L}^\pi\bar{V}(u,X^\pi_u,\hat{\mu}_u)du\middle|\,(X^\pi_t,\hat{\mu}_t)=(x,\mu)\right]\end{aligned}$$

Since $\bar{V}(t,x,\mu)$ satisfies the HJB-equation (4.49), we observe for the right hand side:

$$\begin{aligned}\bar{V}_t(u,X^\pi_u,\hat{\mu}_u)+\mathcal{L}^\pi\bar{V}(u,X^\pi_u,\hat{\mu}_u)&\le 0 \qquad \text{for all } \pi\in\mathcal{A}\\ \bar{V}_t(u,X^\pi_u,\hat{\mu}_u)+\mathcal{L}^\pi\bar{V}(u,X^\pi_u,\hat{\mu}_u)&= 0 \qquad \text{if } \pi=\pi^*\end{aligned}$$

This leads for each $u\in[t,T]$ and $\pi\in\mathcal{A}(t,x,\mu)$ to the following equations:

$$\begin{aligned}\mathbb{E}\left[\bar{V}(s\wedge\tau_n, X^\pi_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\middle|\,(X^\pi_t,\hat{\mu}_t)=(x,\mu)\right]&\le\bar{V}(t,x,\mu) \qquad \text{for all } \pi\in\mathcal{A}(t,x,\mu)\\ \mathbb{E}\left[\bar{V}(s\wedge\tau_n, X^\pi_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\middle|\,(X^\pi_t,\hat{\mu}_t)=(x,\mu)\right]&=\bar{V}(t,x,\mu) \qquad \text{if } \pi=\pi^*\end{aligned}$$

By Lemma 4.22 we can apply dominated convergence and observe for $n\to\infty$

$$\begin{aligned}\bar{V}(t,x,\mu)&\ge\mathbb{E}\left[\bar{V}(s,X^\pi_s,\hat{\mu}_s)\middle|\,(X^\pi_t,\hat{\mu}_t)=(x,\mu)\right] \qquad \text{for all } \pi\in\mathcal{A}(t,x,\mu)\\ \bar{V}(t,x,\mu)&=\mathbb{E}\left[\bar{V}(s,X^\pi_s,\hat{\mu}_s)\middle|\,(X^\pi_t,\hat{\mu}_t)=(x,\mu)\right] \qquad \text{if } \pi=\pi^*\end{aligned}$$

Since $\bar{V}(s,X^\pi_s,\hat{\mu}_s)\to\bar{V}(T,X^\pi_T,\hat{\mu}_T)=\frac{1}{\alpha}(X^\pi_T)^\alpha$ for $s\to T$ we get by dominated convergence

$$\begin{aligned}&\bar{V}(t,x,\mu)\ge\mathbb{E}\left[\frac{1}{\alpha}(X^\pi_T)^\alpha\middle|\,(X^\pi_t,\hat{\mu}_t)=(x,\mu)\right] \qquad \text{for all } \pi\in\mathcal{A}(t,x,\mu)\\ \Rightarrow\quad &\bar{V}(t,x,\mu)\ge\sup_{\pi\in\mathcal{A}(t,x,\mu)}\mathbb{E}\left[\frac{1}{\alpha}(X^\pi_T)^\alpha\middle|\,(X^\pi_t,\hat{\mu}_t)=(x,\mu)\right]=V(t,x,\mu)\end{aligned}$$

Since we get equality for $\pi=\bar{\pi}$ we get $V(t,x,\mu)=\bar{V}(t,x,\mu)$ with optimizer $\pi^*=\bar{\pi}$. □

Lemma 4.22
There is an integrable random variable Y with $\left|\bar{V}(s\wedge\tau_n, X^\pi_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\right|\le Y$.

Proof. The continuous function f is a second order polynomial in μ, hence

$$f(t,\mu)\le C_t(1+\|\mu\|_2^2)$$

where the leading constant C_t may depend on time t. However $t\in[0,T]$ is in a closed interval, hence $f(\cdot,\mu)$ is uniformly continuous in t and attains its supremum as a maximum.

$$f(t,\mu)\le C(1+\|\mu\|_2^2)$$

for some upper bound $C\ge\max_{t\in[0,T]}C_t<\infty$.
Therefore

$$\left|\bar{V}(t,x,\mu)\right|=-\frac{1}{\alpha}x^\alpha e^{f(t,\mu)}\le-\frac{1}{\alpha}x^\alpha e^{C(1+\|\mu\|_2^2)}$$

$$\begin{aligned}
\Rightarrow \quad \left|\bar{V}(s \wedge \tau_n, X^\pi_{s\wedge\tau_n}, \hat{\mu}_{s\wedge\tau_n})\right| &\leq -\frac{1}{\alpha}(X^\pi_{s\wedge\tau_n})^\alpha \exp\Big(C(1 + \|\hat{\mu}_{s\wedge\tau_n}\|_2^2)\Big) \\
&\leq -\frac{1}{\alpha} \sup_{s\in[t,T]} (X^\pi_s)^\alpha \exp\Big(C(1 + \|\hat{\mu}_s\|_2^2)\Big) \\
&\leq -\frac{1}{\alpha} \sup_{s\in[t,T]} (X^\pi_s)^\alpha \cdot \sup_{s\in[t,T]} \exp\Big(C(1 + \|\hat{\mu}_s\|_2^2)\Big) =: Y
\end{aligned}$$

Now on the one hand $X^\pi_s > 0$ is uniformly continuous in $[t, T]$, hence $M := \inf_{s\in[t,T]} X^\pi_s > 0$ and therefore $\sup_{s\in[t,T]}(X^\pi)^\alpha_s \leq M^\alpha < \infty$.
And on the other hand $\hat{\mu}_t$ is a martingale, hence $\|\hat{\mu}_s\|_2^2$ is a submartingale since $\|\cdot\|_2^2$ is a convex function. Hence $\exp\Big(C(1 + \|\hat{\mu}_s\|_2^2)\Big)$ is also a submartingale and by Doob's maximal inequality we get

$$\mathbb{E}\left[\sup_{s\in[t,T]} \exp\Big(C(1 + \|\hat{\mu}_s\|_2^2)\Big)\right] \leq \left(\frac{C}{C-1}\right)^C \mathbb{E}\left[\exp\Big(C(1 + \|\hat{\mu}_T\|_2^2)\Big)\right] < \infty$$

Hence $\mathbb{E}[Y] < \infty$. $\square$

4.4 Determining the correct case and active dimensions

Following the reverse separation approach of Section 4.2, we have observed four different cases with several subcases that the optimal dual process λ and the corresponding optimal portfolio strategy π^λ could be in. This section is devoted to deciding which case is the correct choice for some given set of market parameters. Additionally we will provide an algorithm that leads to the correct optimal case.

First we note that the dual processes ν were created depending on $(t, h, \mu) = (t, H_t^\nu, \hat{\mu}_t)$ and the portfolio strategies π were created depending on $(t, x, \mu) = (t, X_t^\pi, \hat{\mu}_t)$. However both, the optimal dual solution λ and the optimal portfolio strategy π^λ only depend on the time and the current filter value $(t, \mu) = (t, \hat{\mu}_t)$. Note that during this whole section we will continue using (t, μ) for $(t, \hat{\mu}_t)$.

We have already seen in Remark 4.11 and Proposition 4.12 that Cases 1 and 2 are previsible given the current market parameters. Case 1 happens if and only if the optimal unconstrained solution π_{unc} is admissible and Case 2 happens if and only if $\mu - r\mathbb{1} < 0$. In these cases we get the following optimal solutions at time t:

$$
\begin{aligned}
\lambda_t &= \lambda^{(1,\cdot)}(t,\mu) = 0 && \Rightarrow \quad \pi_t^\lambda = \pi^{(1,\cdot)}(t,\mu) = \frac{1}{1-\alpha}\Sigma^{-1}C(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \\
\lambda_t &= \lambda^{(2,\cdot)}(t,\mu) = -(\mu - r\mathbb{1}) && \Rightarrow \quad \pi_t^\lambda = \pi^{(2,\cdot)}(t,\mu) = 0
\end{aligned}
$$

Unfortunately we have also seen in the previous sections that if neither of the trivial cases holds true, we cannot directly observe whether Case 3 or Case 4 is the right choice and in particular which dimensions have to be active. This can be seen similarly to Remark 3.16. In Case 3 with active dimensions I, we get the following optimal solution at time t:

$$
\begin{aligned}
\lambda_t &= \lambda^{(3,I)}(t,\mu) = \mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \\
\Rightarrow \quad \pi_t^\lambda &= \pi^{(3,I)}(t,\mu) = \frac{1}{1-\alpha}\Sigma_{\mathbb{J}}^{(3,I)}C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1})
\end{aligned}
$$

And in Case 4 with active dimensions I the optimal solution becomes:

$$
\begin{aligned}
\lambda_t = \lambda^{(4,I)}(t,\mu) &= \mathbb{J}^{(4,I)}C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \\
&\quad + \alpha(T-t)g^I\mathbb{J}^{(4,I)}C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} \\
&\quad + (1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} \\
\Rightarrow \quad \pi_t^\lambda = \pi^{(4,I)}(t,\mu) &= \frac{1}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)}C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \\
&\quad + \frac{\alpha}{1-\alpha}(T-t)g^I\Sigma_{\mathbb{J}}^{(4,I)}C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} \\
&\quad + g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}
\end{aligned}
$$

In the following we will collect some properties of these optimal solutions.

Example 4.23
In Lemma 3.5 we have seen that under logarithmic utility we can eliminate passive dimensions from the market model without changing the optimal solution. In case of power utility the elimination of passive dimensions does not preserve the correct optimal solution in general.

We consider the following easy counterexample in two dimensions:

$$\alpha = -1 \quad , \quad t = 1 \quad , \quad T = 10$$
$$\mu - r\mathbb{1} = \begin{pmatrix} 0.04 \\ -0.04 \end{pmatrix} , \ \Sigma = \begin{pmatrix} 0.01 & 0.005 \\ 0.005 & 0.04 \end{pmatrix} , \ \Sigma_0 = \begin{pmatrix} 0.01 & 0.002 \\ 0.002 & 0.01 \end{pmatrix}$$
$$\Rightarrow \quad \pi_{\text{unc}} \approx \begin{pmatrix} 0.74 \\ -0.37 \end{pmatrix} \notin K$$

Later on we will see that the optimal solution in this setting is in Case $(3, \{1\})$. Hence

$$\pi^{(3,I)} = \frac{1}{2} \begin{pmatrix} 100 & 0 \\ 0 & 0 \end{pmatrix} C^{(3,I)}(1)\gamma_1^{-1}(\mu - r\mathbb{1}) \approx \begin{pmatrix} 0.6194 \\ 0 \end{pmatrix}$$
$$\text{where} \quad C^{(3,I)}(1) = \left(\gamma_1^{-1} + \frac{9}{2}\begin{pmatrix} 100 & 0 \\ 0 & 0 \end{pmatrix}\right)^{-1} \approx \begin{pmatrix} 0.0015 & 0.0004 \\ 0.0004 & 0.0077 \end{pmatrix}$$

On the other hand, we consider the setting without passive dimension 2:

$$\mu' - r = 0.04 \quad , \quad \Sigma' = 0.01 \quad , \quad \Sigma_0' = 0.01$$
$$\Rightarrow \quad \pi'_{\text{unc}} = 0.6154 \in K$$

However we observe that due to missing correlation effects $\pi'_{\text{unc}} \neq (\pi_{\text{unc}})_{\{1\}}$.

Remark 4.24
The above example points out some very important characteristics of optimal strategies in general.

On first sight it seems unnatural that the optimal constrained strategy should also depend on those stocks that we don't even invest in. However just because some stock has a very poor performance and does not make it into the optimal strategy, it does not loose its correlation to the other stocks. Deleting this stock from the market model changes the covariance-matrices Σ and Σ_0 and therefore the resulting optimal strategies.

Also note that this has extensive consequences to any practical application since it means that in order to get perfectly fitting results the investor would have to include any existing stock in the world into her model. Theoretically this also expands to having to include any investment opportunity that correlates to her model.

For logarithmic utility we showed in Lemma 3.5 that we can eliminate some of those stocks without changing the resulting strategy. While this is a computational advantage, this is also a major drawback of logarithmic utility and its low risk-aversion.

Remark 4.25 (notation)
In the following we will frequently use the matrix $D(t)$:

$$D(t) := \left(\Sigma^{-1}C(t)\gamma_t^{-1}\right)^{-1}$$

$$= \gamma_t \left(\gamma_t^{-1} - \frac{\alpha}{1-\alpha}(T-t)\Sigma^{-1} \right) \Sigma$$
$$= \Sigma - \frac{\alpha}{1-\alpha}(T-t)\gamma_t$$

Note $\pi_{\text{unc}} = \frac{1}{1-\alpha} D(t)^{-1}(\mu - r\mathbb{1})$.

Proposition 4.26 (The optimal solutions in Case 3)
Let I and J be the (correct) choices for active and passive dimensions.
Then for $\lambda = \lambda^{(3,I)}(t,\mu)$ and $\pi = \pi^{(3,I)}(t,\mu)$ we get:

$$\begin{aligned} \pi_J = 0 \quad &\text{and} \quad 0_I \le \pi_I = \frac{1}{1-\alpha}\left(D(t)_{II}\right)^{-1}(\mu - r\mathbb{1})_I \\ \lambda_I = 0 \quad &\text{and} \quad 0_J \le \lambda_J = -(\mu - r\mathbb{1})_J + D(t)_{JI}(1-\alpha)\pi_I \end{aligned}$$

Proof. By construction the optimal solutions satisfy $\pi \geq 0$ and $\lambda \geq 0$. Define

$$B(t) := C^{(3,I)}(t)\gamma_t^{-1} = \left(Id - \frac{\alpha}{1-\alpha}(T-t)\gamma_t \Sigma_{\mathbb{J}}^{(3,I)} \right)^{-1}$$
$$b(t) := B(t)^{-1} = \begin{pmatrix} Id_I - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(\Sigma_{II})^{-1} & 0_{IJ} \\ -\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{JI}(\Sigma_{II})^{-1} & Id_{JJ} \end{pmatrix}$$

By the matrix inversion formula (3.8) we get (omitting the (t))

$$\begin{aligned} B_{II} = (b^{-1})_{II} &= \left(b_{II} - b_{IJ}(b_{JJ})^{-1}b_{JI}\right)^{-1} \\ &= (b_{II})^{-1} \\ B_{JJ} = (b^{-1})_{JJ} &= (b_{JJ})^{-1} + (b_{JJ})^{-1}b_{JI}(b^{-1})_{II}b_{IJ}(b_{JJ})^{-1} \\ &= Id_J \\ B_{JI} = (b^{-1})_{JI} &= -(b_{JJ})^{-1}b_{JI}(b^{-1})_{II} \\ &= -b_{JI}(b_{II})^{-1} \\ B_{IJ} = (b^{-1})_{IJ} &= -(b^{-1})_{II}b_{IJ}(b_{JJ})^{-1} \\ &= 0_{IJ} \end{aligned}$$

Hence

$$\begin{aligned} \pi &= \frac{1}{1-\alpha}\Sigma_{\mathbb{J}}^{(3,I)} C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) \\ &= \frac{1}{1-\alpha}\begin{pmatrix} (\Sigma_{II})^{-1} & 0_{IJ} \\ 0_{JI} & 0_{JJ} \end{pmatrix} B(t)(\mu - r\mathbb{1}) \\ \Rightarrow \quad \pi_J &= 0_J \\ \pi_I &= \frac{1}{1-\alpha}(\Sigma_{II})^{-1}B(t)_{II}(\mu - r\mathbb{1})_I \\ &= \frac{1}{1-\alpha}(b(t)_{II}\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \end{aligned}$$

$$= \frac{1}{1-\alpha} \left(D(t)_{II}\right)^{-1} (\mu - r\mathbb{1})_I$$

and

$$\lambda = \mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) = \begin{pmatrix} 0_{II} & 0_{IJ} \\ \Sigma_{JI}(\Sigma_{II})^{-1} & -Id_J \end{pmatrix} B(t)(\mu - r\mathbb{1})$$

$$\begin{aligned}
\Rightarrow \quad \lambda_I &= 0_I \\
\lambda_J &= \Sigma_{JI}(\Sigma_{II})^{-1}B(t)_{II}(\mu - r\mathbb{1})_I - B(t)_{JI}(\mu - r\mathbb{1})_I - B(t)_{JJ}(\mu - r\mathbb{1})_J \\
&= -(\mu - r\mathbb{1})_J + \Sigma_{JI}(\Sigma_{II})^{-1}(b(t)_{II})^{-1}(\mu - r\mathbb{1})_I + b(t)_{JI}(b(t)_{II})^{-1}(\mu - r\mathbb{1})_I \\
&= -(\mu - r\mathbb{1})_J + \left(\Sigma_{JI} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{JI}\right)(\Sigma_{II})^{-1}(b(t)_{II})^{-1}(\mu - r\mathbb{1})_I \\
&= -(\mu - r\mathbb{1})_J + D(t)_{JI}(1-\alpha)\pi_I
\end{aligned}$$

□

Now we show in Lemmas 4.27 and 4.31 that several intuitively false choices of active and passive dimensions in fact lead to non-admissible solutions, revealing that the choice was wrong:

Lemma 4.27 (Choice of I in Case 3)
Let I and J be the active and passive dimensions such that (λ, π) is the optimal solution in Case 3 and let (λ', π') be the solution if we chose I' and J' instead.
1) Let $i \in I$ with $\pi_i > 0$ be an active dimension.
   If we choose $I' = I \setminus \{i\}$ and $J' = J \cup \{i\}$ then we get $\lambda'_i < 0$.
2) Let $j \in J$ with $\lambda_j > 0$ be a passive dimension (by Proposition 4.26).
   If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ then we get $\pi'_j < 0$.

Proof. The proof is given in Appendix A.4.4 □

Remark 4.28
So far we know $\pi_I \geq 0$, $\pi_J = 0$ and $\lambda_I = 0$, $\lambda_J \geq 0$.
However similarly to Remark 3.13 under logarithmic utility there can be situations with $\pi_k = 0$ and $\lambda_k = 0$ for some dimension k. We call these dimensions 'boundary dimensions' since they are on the boundary between being active and being passive. The corresponding stocks are almost invested in, meaning that if the market parameters change slightly into the right direction, then $\pi_k > 0$.
These dimensions can be considered both, active and passive as Corollary 4.29 shows.

Corollary 4.29 (Dimensions on the boundary in Case 3)
If $j \in J$ is a passive dimensions that is almost invested in, than j can also be considered an active dimension.

In particular:
Let I and J be the active and passive dimensions such that (λ, π) is the optimal solution in Case 3 and let (λ', π') be the solution if we chose I' and J' instead.
Let $j \in J$ with $\pi_j = 0$ and $\lambda_j = 0$ be a passive dimension, that is almost invested in.
Then: If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ we still get $\lambda' = \lambda$ and $\pi' = \pi$.

Proof. The proof is given in Appendix A.4.5 □

In Case 4 the formulas look significantly more complicated:

Proposition 4.30 (The optimal solutions in Case 4)
Let I and J be the (correct) choices for active and passive dimensions. To simplify notation we use following abbreviations:

$$\begin{aligned}\mathbb{I}d^I &:= Id_I - g^I(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top \\ \mathbb{D}^I(t) &:= \Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}\mathbb{I}d^I \\ (\mu - r\mathbb{1})_I^{+\mathbb{I}} &:= (\mu - r\mathbb{1})_I + \alpha(T-t)g^I(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I\end{aligned}$$

Then for $\lambda = \lambda^{(4,I)}(t,\mu)$ and $\pi = \pi^{(4,I)}(t,\mu)$ we get:

$$\begin{aligned}\pi_I &= \frac{1}{1-\alpha}\mathbb{I}d^I\ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}}\ + g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\ \pi_J &= 0\end{aligned}$$

with $\mathbb{1}_I^\top\pi_I = 1$ and

$$\begin{aligned}\bar\lambda &= -g^I\ \mathbb{1}_I^\top\ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}}\ + g^I(1-\alpha) \\ \bar\lambda\mathbb{1}_I &= (1-\alpha)\Sigma_{II}\pi_I - \Sigma_{II}\ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\ \lambda_J &= -(\mu - r\mathbb{1})_J + D(t)_{JI}(1-\alpha)\pi_I\end{aligned}$$

Proof. The proof is given in A.4.6. □

We get the analogue result to Lemma 4.27 for Case 4, showing that several intuitively false choices of active and passive dimensions in fact lead to non-admissible solutions, revealing that the choice was wrong:

Lemma 4.31 (Choice of I in Case 4)
Let I and J be the active and passive dimensions such that (λ, π) is the optimal solution in Case 4 and let (λ', π') be the solution if we chose I' and J' instead.
1) Let $i \in I$ with $\pi_i > 0$ be an active dimension.
If we choose $I' = I \setminus \{i\}$ and $J' = J \cup \{i\}$ then we get $\lambda'_i < \bar\lambda'$.
2) Let $j \in J$ with $\lambda_j > \bar\lambda$ be a passive dimension (by Proposition 4.30).
If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ then we get $\pi'_j < 0$.
Hence several false choices of active and passive dimensions lead to a non-admissible solutions, revealing that the choice was wrong.

Proof. Proving this directly is way too circumstantial and lengthy, hence we conclude this from Lemma 4.39 and Theorem 4.36. □

Corollary 4.32 (Dimensions on the boundary in Case 4)
If $j \in J$ is a passive dimensions that is almost invested in, than j can also be considered an active dimension.

In particular:
Let I and J be the active and passive dimensions such that (λ, π) is the optimal solution in Case 4 and let (λ', π') be the solution if we chose I' and J' instead.
Let $j \in J$ with $\pi_j = 0$ and $\lambda_j = \bar{\lambda}$ be a passive dimension that is almost invested in.
Then: If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ we still get $\lambda' = \lambda$ and $\pi' = \pi$.

Proof. Proving this directly is way too circumstantial and lengthy, hence we conclude this from Lemma 4.39 and Theorem 4.36. □

Remark 4.33
Note that each of the Results 4.27, 4.29, 4.31 and 4.32 follow immediately from Lemma 4.39 and Theorem 4.36. The proofs given above only show the underlying technical connections.

Now we are able to formulate an algorithm to determine the correct case:

Algorithm 4.34 (Solving the portfolio optimization problem explicitly)
The constrained portfolio optimization problem under power utility gets solved via going through the following four cases until hitting an admissible solution. When hitting an admissible solution this provides the optimal dual process and the optimal constrained portfolio strategy at the current time t with the current filter value $\hat{\mu}_t$.

1. Compute the optimal unconstrained strategy $\pi_{\text{unc}} = \frac{1}{1-\alpha}\Sigma^{-1}C(t)\gamma_t^{-1}(\mu - r\mathbb{1})$.
   If $\pi_{\text{unc}} \in K$ then $\lambda = 0$ and $\pi^\lambda = \pi_{\text{unc}}$.
2. If $\mu - r\mathbb{1} < 0$ then $\lambda = -(\mu - r\mathbb{1})$ and $\pi^\lambda = 0$.
3. For each $\emptyset \neq I \subset \{1, ..., d\}$ compute $\lambda^{(3,I)}$ and $\pi^{(3,I)}$.
   If $\pi^{(3,I)} \in K$ and $\lambda^{(3,I)} \geq 0$, then $\lambda = \lambda^{(3,I)}$ and $\pi^\lambda = \pi^{(3,I)}$.
4. For each $\emptyset \neq I \subseteq \{1, ..., d\}$ compute $\lambda^{(4,I)}$ and $\pi^{(4,I)}$.
   If $\pi^{(4,I)} \in K$ and $\lambda^{(4,I)} \geq \bar{\lambda}^I \mathbb{1} < 0$, then $\lambda = \lambda^{(4,I)}$ and $\pi^\lambda = \pi^{(4,I)}$.

Remark 4.35
There are several additional conditions that can be introduced to Algorithm 4.34 to accelerate his speed.
One example would be to check whether $\pi_{\text{unc}} > 0$. In that case the no short-selling constraint is not harmed and it is not necessary to consider Case 3 at all. Also sorting the possible choices for the active dimensions can increase the algorithm's speed as described in Remark 3.23.
Several more possibilities can be concluded from the derivations in Section 4.4.1.

Theorem 4.36
The above algorithm always results in the unique solution.

Proof. Case 1:
If the optimal unconstrained strategy is admissible, there is nothing left to solve. Otherwise Theorem 2.11 ensures that there is an optimal λ whose corresponding auxiliary market leads to the optimal constrained portfolio strategy. Depending on the sign of $\bar{\lambda} = \min_i \lambda_i$ we need to find λ in Case 2 ($\bar{\lambda} > 0$), Case 3 ($\bar{\lambda} = 0$) or Case 4 ($\bar{\lambda} < 0$).

Case 2:
Proposition 4.12 shows that Case 2 is equivalent to checking the condition $\mu - r\mathbb{1} < 0$ and leads to $\lambda = -(\mu - r\mathbb{1})$ and $\pi^\lambda = 0$.

Cases 3 and 4:
By Theorem 2.9 we know that if $\lambda \in \tilde{K}$, $\pi^\lambda \in K$ and $\lambda^\top \pi^\lambda + \delta(\lambda) = 0$ then (λ, π^λ) is already the optimal solution. Obviously $\lambda \in \tilde{K} = \mathbb{R}^d$ is always true.
By the structure of the optimal solution in Case 3 we observe if $\pi^\lambda \in K$:

$$\lambda = \begin{pmatrix} 0_I \\ \lambda_J \end{pmatrix} \quad \text{and} \quad \pi^\lambda = \begin{pmatrix} \pi_I^\lambda \\ 0_J \end{pmatrix} \quad \Rightarrow \quad \lambda^\top \pi^\lambda = 0$$

Hence if $\lambda \geq 0$ then $\delta(\lambda) = 0$ and hence $\lambda^\top \pi^\lambda + \delta(\lambda) = 0$.

On the other hand in Case 4 we observe if $\pi^\lambda \in K$:

$$\lambda = \begin{pmatrix} \bar{\lambda}\mathbb{1}_I \\ \lambda_J \end{pmatrix} \quad \text{and} \quad \pi^\lambda = \begin{pmatrix} \pi_I^\lambda \\ 0_J \end{pmatrix} \text{ with } \mathbb{1}_I^\top \pi_I^\lambda = 1 \quad \Rightarrow \quad \lambda^\top \pi^\lambda = \bar{\lambda}$$

Hence if $\lambda \geq \bar{\lambda}\mathbb{1} < 0$ then $\delta(\lambda) = -\bar{\lambda}$ and hence $\lambda^\top \pi^\lambda + \delta(\lambda) = \bar{\lambda} - \bar{\lambda} = 0$.
In both cases we can follow that (λ, π^λ) is already the optimal solution. □

Remark 4.37
Of course in the case of Algorithm 4.34 under power utility the same remarks hold true than with Algorithm 3.20 under logarithmic utility:
Every choice of I that does not harm the structural assumptions made in Algorithm 4.34, satisfies the assumptions of Theorem 2.9 and hence the resulting admissible strategy is the optimal solution. As the optimal portfolio strategy in our market setting is unique, two different but admissible choices of I can only lead to the same resulting optimal strategy.

Contrary to Algorithm 3.20, Algorithm 4.34 solves the constrained portfolio optimization problem not only faster than a standard non-linear optimization approach. There is no numerical approach at all that could solve the HJB-equation (4.13) since it depends significantly on the structure of its optimizer (4.14) that depends on the solution of the HJB-equation itself.

4.4.1 The domain of the value function

In this section we will analyse the domains of the dual and primal value functions. In particular we will analyse the effect made by the input parameters $(t, \mu) = (t, \hat{\mu}_t)$ on the choice of the correct case and active dimensions.

The correct case is chosen in the reverse separation approach during stochastic control. Hence we need to consider the parameters $(t,\xi,\mu) \in [0,T] \times \mathbb{R}_+ \times \mathbb{R}^d$, where $\xi = x = X_t$ for the primal problem and $\xi = h = H_t$ for the dual problem. However we have already seen that the parameters $x = X_t$, respectively $h = H_t$ influence neither the choice of case nor the optimal solutions. Therefore it is enough to only consider the parameters $(t,\mu) \in \mathbb{D}_+ := [0,T] \times \mathbb{R}^d$ and to consider $\mathbb{D}_+$ as the interesting part of the domain of either of the value functions.
Note the following disjoint partitioning of the domain $\mathbb{D}_+ = [0,T] \times \mathbb{R}^d$ with respect to the four cases:

$$\begin{aligned}
\mathbb{D}_+^1 &:= \left\{(t,\mu) \in \mathbb{D}_+ \middle| \lambda(t,\mu) = 0\right\} \\
\mathbb{D}_+^2 &:= \left\{(t,\mu) \in \mathbb{D}_+ \middle| \lambda(t,\mu) > 0\right\} \\
\mathbb{D}_+^3 &:= \left\{(t,\mu) \in \mathbb{D}_+ \middle| \lambda(t,\mu) \geq 0 \text{ and } 0 \neq \lambda(t,\mu) \not> 0\right\} \\
\mathbb{D}_+^4 &:= \left\{(t,\mu) \in \mathbb{D}_+ \middle| \lambda(t,\mu) \not\geq 0\right\}
\end{aligned}$$

Obviously it is enough to only consider either the optimal dual process $\lambda(t,\mu)$ or the optimal portfolio strategy $\pi(t,\mu)$ to define these sub-domains.

Now at any time t the investor (respectively the algorithm) needs to decide which case is the correct choice. The only information that is not given to the investor ahead of time t will be the information from observing the market, hence the filter value $\hat{\mu}_t$. Therefore it could be of great importance to the investor to be able to compute regions of possible filter values $\hat{\mu}_t \in \mathbb{D} := \mathbb{R}^d$ that belong to the different Case (i,I). This way whenever a new filter value is generated, he could immediately decide which region and case it belongs to and hence he would immediately know the correct optimal strategy.
These regions are defined for any fixed time t and form a disjoint partitioning of $\mathbb{D} = \mathbb{R}^d$:

$$\begin{aligned}
\mathbb{D}^{(1,\cdot)} &:= \mathbb{D}_t^{(1,\cdot)} := \left\{\mu \in \mathbb{D} \middle| \lambda(t,\mu) = 0\right\} \\
\mathbb{D}^{(2,\cdot)} &:= \mathbb{D}_t^{(2,\cdot)} := \left\{\mu \in \mathbb{D} \middle| \lambda(t,\mu) > 0\right\} \\
\mathbb{D}^{(3,\cdot)} &:= \mathbb{D}_t^{(3,\cdot)} := \left\{\mu \in \mathbb{D} \middle| \lambda(t,\mu) \geq 0 \text{ and } 0 \neq \lambda \not> 0\right\} \\
\mathbb{D}^{(4,\cdot)} &:= \mathbb{D}_t^{(4,\cdot)} := \left\{\mu \in \mathbb{D} \middle| \lambda(t,\mu) \not\geq 0\right\}
\end{aligned}$$

Note that while the optimal solutions λ and π depend affine linear on $\mu = \hat{\mu}_t$ their dependence on t is highly non-linear, hence one cannot expect to derive nice regions when trying to describe these subregions also with respect to time t.

We have already seen that Cases 1 and 2 are previsible but Cases 3 and 4 have to be split up further to account for the various choices of active dimensions.
Define $I(t,\mu)$ to be the set of all possible correct choices of active dimensions I for the given parameters (t,μ). Note by Corollaries 4.29 and 4.32 that the choice of active dimensions does not have to be unique if the active dimensions are on the boundary between two subcases. For each $\emptyset \neq I \subseteq \{1,...,d\}$ and fixed time t define:

$$\mathbb{D}^{(3,I)} := \left\{\mu \in \mathbb{D}^{(3,\cdot)} \middle| \; I \in I(t,\mu)\right\}$$

$$
\begin{aligned}
&= \Big\{ \mu \in \mathbb{D}^{(3,\cdot)} \Big| \ \pi^{(3,I)} \in K,\ \lambda^{(3,I)} \in \tilde{K} \text{ and } \lambda^{(3,I)}{}^\top \pi^{(3,I)} + \delta\big(\lambda^{(3,I)}\big) = 0 \Big\} \\
\mathbb{D}^{(4,I)} &:= \Big\{ \mu \in \mathbb{D}^{(4,\cdot)} \Big| \ I \in I(t,\mu) \Big\} \\
&= \Big\{ \mu \in \mathbb{D}^{(4,\cdot)} \Big| \ \pi^{(4,I)} \in K,\ \lambda^{(4,I)} \in \tilde{K} \text{ and } \lambda^{(4,I)}{}^\top \pi^{(4,I)} + \delta\big(\lambda^{(4,I)}\big) = 0 \Big\}
\end{aligned}
$$

Obviously the union of all $\mathbb{D}^{(3,I)}$ becomes $\mathbb{D}^{(3,\cdot)}$ and the union of the $\mathbb{D}^{(4,I)}$ becomes $\mathbb{D}^{(4,\cdot)}$ but the subregions $\mathbb{D}^{(\cdot,I)}$ are not disjoint any more.
Technically we have only derived the optimal solutions in the relative interiors of the subregions $\mathbb{D}^{(\cdot,I)}$. However in Section 4.5 we will show that these solutions also hold true on the boundaries of the respective regions. Also we have already seen that the limits of the respective solutions in adjacent cases coincide when converging to their common boundary. This follows in particular from Lemma 4.39.

Lemma 4.38
The sets $\mathbb{D}^{(1,\cdot)}, \mathbb{D}^{(2,\cdot)}, \mathbb{D}^{(3,I)}$ and $\mathbb{D}^{(4,I)}$ are multi-dimensional convex polyhedrons.

Proof. We use the notation $M \cdot (K - c) := \big\{ m = M(k-c) \big| k \in K \big\} \subset \mathbb{R}^d$ for some matrix $M \in \mathbb{R}^{d\times d}$ and some set $K \subset \mathbb{R}^d$. Hence

$$
\begin{aligned}
\mathbb{D}^{(1,\cdot)} &= \big\{ \mu \in \mathbb{D} \ \big| \lambda(t,\mu) = 0 \big\} \\
&= \big\{ \mu \in \mathbb{D} \ \big| \pi_{\text{unc}} \in K \big\} \\
&= \Big\{ \mu \in \mathbb{D} \ \Big| \frac{1}{1-\alpha} D(t)^{-1}(\mu - r\mathbb{1}) \in K \Big\} \\
&= \Big\{ \mu \in \mathbb{D} \ \Big| \mu \in r\mathbb{1} + (1-\alpha) D(t) \cdot K \Big\} \\
\mathbb{D}^{(2,\cdot)} &= \big\{ \mu \in \mathbb{D} \ \big| \lambda(t,\mu) > 0 \big\} \\
&= \Big\{ \mu \in \mathbb{D} \ \Big| \mu < r\mathbb{1} \Big\}
\end{aligned}
$$

Hence $\mathbb{D}^{(1,\cdot)}$ is an affine linear shift of the convex polyhedron K and $\mathbb{D}^{(2,\cdot)}$ is the solution to a multi-dimensional linear inequality. Hence both are convex polyhedrons that can easily be observed. For $\mathbb{D}^{(3,I)}$ and $\mathbb{D}^{(4,I)}$ note that the condition $\lambda^\top \pi^\lambda + \delta(\lambda) = 0$ is redundant by the proof of Theorem 4.36 if the condition $\lambda \in \tilde{K}$ is replaced by the structure of λ in the respective case. In the following we use the notation $K^I := \big\{ k \in \mathbb{R}^I \big| \mathbb{1}_I^\top k \le 1 \big\}$ for the admissibility set in the dimensions I. Hence

$$
\begin{aligned}
\mathbb{D}^{(3,I)} &= \Big\{ \mu \in \mathbb{D}^{(3,\cdot)} \Big| \pi^{(3,I)} \in K \ ,\ \lambda^{(3,I)} \ge 0 \Big\} \\
&= \Big\{ \mu \in \mathbb{D}^{(3,\cdot)} \Big| \frac{1}{1-\alpha} \Sigma_{\mathbb{J}}^{(3,I)} C(t)^{(3,I)} \gamma_t^{-1} (\mu - r\mathbb{1}) \in K \\
&\qquad \text{and } \ \mathbb{J}^I C^{(3,I)}(t) \gamma_t^{-1} (\mu - r\mathbb{1}) \ge 0 \Big\} \\
&\overset{(4.26)}{=} \Big\{ \mu \in \mathbb{D}^{(3,\cdot)} \Big| \frac{1}{1-\alpha} \left(D(t)_{II}\right)^{-1} (\mu - r\mathbb{1})_I \in K^I \\
&\qquad \text{and } \ -(\mu - r\mathbb{1})_J + D(t)_{JI} \left(D(t)_{II}\right)^{-1} (\mu - r\mathbb{1})_I \ge 0_J \Big\}
\end{aligned}
$$

$$= \Big\{\mu \in \mathbb{D}^{(3,\cdot)} \Big| (\mu - r\mathbb{1})_I \in (1-\alpha) D(t)_{II} \cdot K^I \\ \text{and } \ (\mu - r\mathbb{1})_J \leq D(t)_{JI} \left(D(t)_{II}\right)^{-1} (\mu - r\mathbb{1})_I \Big\}$$

and

$$\mathbb{D}^{(4,I)} = \Big\{\mu \in \mathbb{D}^{(4,\cdot)} \Big| \pi^{(4,I)} \in K \ , \ \lambda^{(4,I)} \geq \bar{\lambda}^I \mathbb{1} < 0 \Big\}$$

$$\overset{(4.30)}{=} \Big\{\mu \in \mathbb{D}^{(4,\cdot)} \Big| \pi_I = \frac{1}{1-\alpha} \mathbb{I}d^I \ \mathbb{D}^I(t)^{-1} (\mu - r\mathbb{1})_I^{+\mathbb{I}} + g^I (\Sigma_{II})^{-1} \mathbb{1}_I \in K \\ \text{and } \ \bar{\lambda} = -g^I \ \mathbb{1}_I^\top \ \mathbb{D}^I(t)^{-1} (\mu - r\mathbb{1})_I^{+\mathbb{I}} + g^I (1-\alpha) < 0 \\ \text{and } \ \lambda_J = -(\mu - r\mathbb{1})_J + D(t)_{JI} (1-\alpha) \pi_I \geq \bar{\lambda} \mathbb{1}_J \Big\}$$

$$= \Big\{\mu \in \mathbb{D}^{(4,\cdot)} \Big| (\mu - r\mathbb{1})_I^{+\mathbb{I}} \in (1-\alpha) \mathbb{D}^I(t) (\mathbb{I}d^I)^{-1} \cdot \left(K - g^I (\Sigma_{II})^{-1} \mathbb{1}_I\right) \\ \text{and } \ \mathbb{1}_I^\top \ \mathbb{D}^I(t)^{-1} (\mu - r\mathbb{1})_I^{+\mathbb{I}} > (1-\alpha) \\ \text{and } \ (\mu - r\mathbb{1})_J \leq \bar{\lambda} \mathbb{1}_J + D(t)_{JI} (1-\alpha) \pi_I \Big\}$$

$$\text{where } \ (\mu - r\mathbb{1})_I^{+\mathbb{I}} := (\mu - r\mathbb{1})_I + \alpha (T-t) g^I (\gamma_t)_{II} (\Sigma_{II})^{-1} \mathbb{1}_I$$

Hence all sets $\mathbb{D}^{(3,I)}$ and $\mathbb{D}^{(4,I)}$ are solutions to linear inequalities in μ and hence they are multi-dimensional convex polyhedrons. □

The following lemma shows in particular that all solutions can be extended continuously to the boundaries of all subregions $\mathbb{D}^{(\cdot,\cdot)}$ and coincide on each point on their boundaries.

Lemma 4.39 (shifted unconstrained solution)
The optimal constrained solution π^λ in all cases is equal to the optimal unconstrained solution with $\mu + \lambda$ plugged in instead of μ, i.e.

$$\pi^\lambda = \pi_{\text{unc}}^{+\lambda} := \frac{1}{1-\alpha} D(t)^{-1} (\mu - r\mathbb{1} + \lambda) \\ = \frac{1}{1-\alpha} \Sigma^{-1} C(t) \gamma_t^{-1} (\mu - r\mathbb{1} + \lambda)$$

In particular all solutions in the domains $\mathbb{D}^{(\cdot,\cdot)}$ can be extended continuously to the boundary of their respective domain and coincide at the boundary points, such that π^λ is continuous in μ.

Proof. This is a special case of Proposition 5.2.
However it can also be proved directly as shown in Appendix A.4.7. □

Remark 4.40
Lemma 4.39 shows that the optimal dual λ is always chosen such that $\mu+\lambda \in \mathbb{D}^{(1,\cdot)}$. Hence any incoming non-admissible filter value $\mu \notin \mathbb{D}^{(1,\cdot)}$ respectively $\pi_{\text{unc}} \notin K$ basically just gets shifted by λ such that the 'updated expected growth rate' $\mu + \lambda$ becomes admissible, hence $\mu + \lambda \in \mathbb{D}^{(1,\cdot)}$.

More precisely we will see in Proposition 4.41 that $\mu + \lambda \in \partial\mathbb{D}^{(1,\cdot)}$, hence the observed filter value gets shifted to the boundary of $\mathbb{D}^{(1,\cdot)}$. This can be motivated as in Example 2.13, where a further shift into the interior of $\mathbb{D}^{(1,\cdot)}$ is not necessary to become admissible, but it would decrease the terminal wealth.
Proposition 4.41 answers the question, which input parameters $\mu \notin \mathbb{D}^{(1,\cdot)}$ lead to a given $\mu' \in \partial\mathbb{D}^{(1,\cdot)}$.

Proposition 4.41 (Exact shifting in Case 3)
Let $\rho \in \partial K$ be arbitrary with $\mathbb{1}_I^\top \rho_I < 1$ and

$$\begin{aligned} \rho_I &> 0_I \quad \text{where } \emptyset \neq I \subset \{1, ..., d\} \\ \rho_J &= 0_J \quad \text{where } \emptyset \neq J = \{1, ..., d\} \setminus I \end{aligned}$$

and consider the parameter μ' on the boundary of $\mathbb{D}^{(1,\cdot)}$

$$\mu' := r\mathbb{1} + (1-\alpha)D(t) \cdot \rho \quad \in \partial\mathbb{D}^{(1,\cdot)}$$

Let $\mu := \mu' + \begin{pmatrix} 0_I \\ c_J \end{pmatrix}$ be an arbitrary input parameter for any $c_J \leq 0$.
Then the optimal dual solution for the constrained portfolio optimization problem with input parameter μ is in Case $(3, I)$ with $\lambda = -\begin{pmatrix} 0_I \\ c_J \end{pmatrix}$ and hence $\mu + \lambda = \mu' \in \partial\mathbb{D}^{(1,\cdot)}$.

Proof. We get the optimal dual solution with Algorithm 4.34:
If $c_J = 0$ then $\mu = \mu' \in \partial\mathbb{D}^{(1,\cdot)} \subset \mathbb{D}^{(1,\cdot)}$ hence $\lambda = 0$ and we are done. Else $c_J \neq 0$.
Then Case 1 does not fit:

$$\begin{aligned} (1-\alpha)\pi_{\text{unc}} &= D(t)^{-1}(\mu - r\mathbb{1}) \\ &= D(t)^{-1}\left(\mu' - r\mathbb{1} + \begin{pmatrix} 0_I \\ c_J \end{pmatrix}\right) \\ &= (1-\alpha)\rho + D(t)^{-1}\begin{pmatrix} 0_I \\ c_J \end{pmatrix} \\ \Rightarrow \quad (\pi_{\text{unc}})_J &= \frac{1}{1-\alpha}(D(t)^{-1})_{JJ}c_J \end{aligned}$$

Since $D(t)$ and $D(t)^{-1}$ are positive definite and $c_J \neq 0$, we get

$$(c_J)^\top(\pi_{\text{unc}})_J = \frac{1}{1-\alpha}c_J^\top(D(t)^{-1})_{JJ}c_J > 0$$

But with $c_J \leq 0_J$ and $(\pi_{\text{unc}})_J \geq 0_J$ this cannot work, hence $\pi_{\text{unc}} \ngeq 0$ and Case 1 does not fit. Also Case 2 does not fit:

$$\begin{aligned} \mu - r\mathbb{1} &= \mu' - r\mathbb{1} + \begin{pmatrix} 0_I \\ c_J \end{pmatrix} = (1-\alpha)D(t) \cdot \rho + \begin{pmatrix} 0_I \\ c_J \end{pmatrix} \\ \Rightarrow \quad (\mu - r\mathbb{1})_I &= (1-\alpha)D(t)_{II}\rho_I \end{aligned}$$

Again $D(t)_{II}$ is positive definite and $\rho_I \neq 0$, hence

$$\rho_I^\top(\mu - r\mathbb{1})_I = \rho_I^\top D(t)_{II}\rho_I > 0$$

But with $\rho_I > 0_I$ and $(\mu - r\mathbb{1})_I < 0_I$ this cannot work, hence $(\mu - r\mathbb{1})_I \nless 0_I$ and Case 2 does not fit.
Now we consider Case $(3, I)$ and get by Proposition 4.26:

$$\begin{aligned}
\lambda_I &= 0_I \\
\lambda_J &= -(\mu - r\mathbb{1})_J + D(t)_{JI}(D(t)_{II})^{-1}(\mu - r\mathbb{1})_I \\
&= -(\mu' - r\mathbb{1})_J - c_J + D(t)_{JI}(D(t)_{II})^{-1}(\mu' - r\mathbb{1})_I \\
&= -(1-\alpha)D(t)_{JI}\rho_I - c_J + (1-\alpha)D(t)_{JI}(D(t)_{II})^{-1}D(t)_{II}\rho_I \\
&= -c_J \geq 0_J \\
\pi_J^\lambda &= 0_J \\
\pi_I^\lambda &= \frac{1}{1-\alpha}(D(t)_{II})^{-1}(\mu - r\mathbb{1})_I \\
&= \frac{1}{1-\alpha}(D(t)_{II})^{-1}(\mu' - r\mathbb{1})_I \\
&= \rho_I > 0_I \\
\mathbb{1}^\top\pi^\lambda &= \mathbb{1}_I^\top\pi_I^\lambda = \mathbb{1}_I^\top\rho_I < 1
\end{aligned}$$

Since this solution is admissible it is already the optimal solution by the proof of Theorem 4.36. □

Proposition 4.42 (exact shifting in Case 4)
Let $\rho \in \partial K$ be arbitrary with $\mathbb{1}_I^\top\rho_I = 1$ and

$$\begin{aligned}
\rho_I &> 0_I \quad \text{where } \emptyset \neq I \subset \{1, ..., d\} \\
\rho_J &= 0_J \quad \text{where } J = \{1, ..., d\} \setminus I
\end{aligned}$$

and consider the parameter μ' on the boundary of $\mathbb{D}^{(1,\cdot)}$

$$\mu' := r\mathbb{1} + (1-\alpha)D(t)\cdot\rho \quad \in \partial\mathbb{D}^{(1,\cdot)}$$

Let $\mu := \mu' + \begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix}$ be an arbitrary input parameter for any $c_J \leq c\mathbb{1}_J$ and $c > 0$.
Then the optimal dual solution for the constrained portfolio optimization problem with input parameter μ is in Case $(4, I)$ with $\lambda = -\begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix}$ and hence $\mu + \lambda = \mu' \in \partial\mathbb{D}^{(1,\cdot)}$.

Proof. The proof is given in Appendix A.4.8. □

Corollary 4.43
The Propositions 4.41 and 4.42 describe exactly those μ leading to Cases $(3, I)$ and $(4, I)$. In particular we observe that all solutions in the domains $\mathbb{D}^{(\cdot,\cdot)}$ can be extended continuously to their boundaries where they coincide with any adjacent case such that λ and π^λ are continuous in μ.

Proof. Proposition 4.41 and 4.42 describe disjoint subsets of $\mathbb{R}^d \setminus \left(\mathbb{D}^{(1,\cdot)} \cup \mathbb{D}^{(2,\cdot)}\right)$, hence we get equivalence in both lemmas.
Continuity of the optimal solution has already been observed in Lemma 4.39. $\square$

Example 4.44
We consider the following easy example in two dimensions:

$$\alpha = -1 \quad , \quad t = 0.5 \quad , \quad T = 1 \quad , \quad r = 0.03$$
$$\Sigma = \begin{pmatrix} 0.2 & -0.05 \\ -0.05 & 0.15 \end{pmatrix} \quad , \quad \Sigma_0 = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.2 \end{pmatrix}$$

Then we can easily visualize the three edges $\eta^{(i)}$ and hence by Propositions 4.41 and 4.42 the complete partitioning of $\mathbb{D} = \mathbb{R}^2$ into the respective subregions indicating the respective cases. This is shown on the left hand side in Figure 4.1. On its right hand side we can see how any incoming filter value $\mu = \hat{\mu}_t$ can easily be classified to its correct case and how it gets shifted by some λ on the boundary $\partial\mathbb{D}^1$.

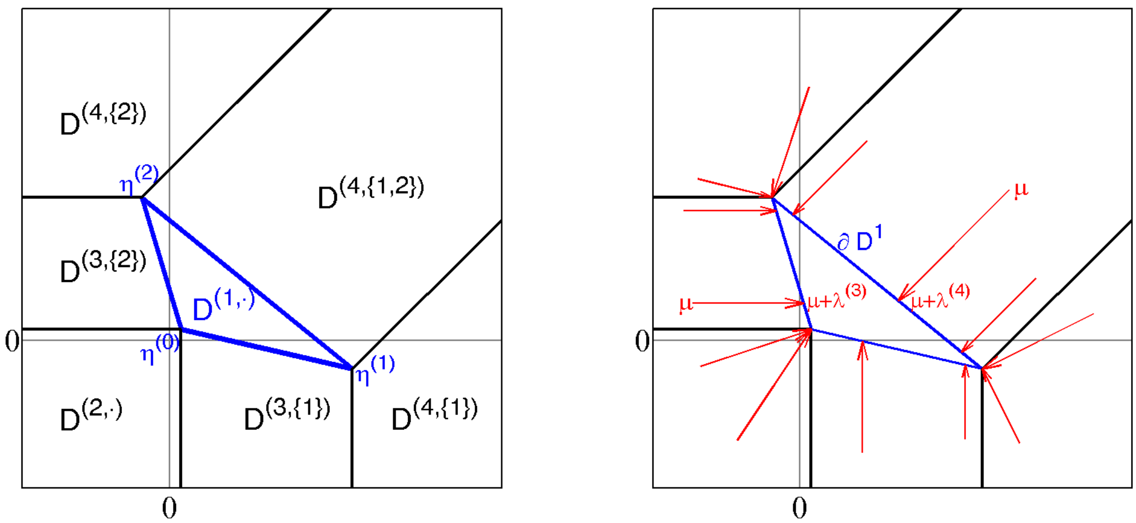


Figure 4.1: (lhs) The partitioning of $\mathbb{D} = \mathbb{R}^2$ into the cases
(rhs) The exact shifting of μ into the admissibility region $\mathbb{D}^1$

Remark 4.45 (Dependence of the optimal solutions on the time)
Obviously the filter $\hat{\mu}_t$ depends significantly on time t. However this is not important for determining the correct case.

First the domain $\mathbb{D}^{(2,\cdot)}$ does not depend on t anyway. Secondly the Propositions 4.41 and 4.42 show that the partitioning of $\mathbb{D}^{(3,\cdot)}$ and $\mathbb{D}^{(4,\cdot)}$ and the positions of their subregions $\mathbb{D}^{(i,I)}$ only depend on the position of the boundary $\partial\mathbb{D}^{(1,\cdot)}$ and not on time t. This can also be seen in Figure 4.1.
Finally the region $\mathbb{D}^{(1,\cdot)}$ and hence its boundary is described by its edges $\eta^{(0)} := r\mathbb{1}$ and

$\eta^{(i)}(t) := r\mathbb{1} + (1-\alpha)D(t)\cdot e_i \in \partial\mathbb{D}^{(1,\cdot)}$. Therefore $\mathbb{D}^{(1,\cdot)}$ actually depends on time t. However the changes of the edges $\eta^{(i)}$ are previsible, hence any resulting shift of the subregions $\mathbb{D}^{(3,I)}$ or $\mathbb{D}^{(4,I)}$ is also previsible.
In particular the optimal solution (λ, π^λ) is continuous in t, hence continuous in (t, μ).

Now we can consider the following alternative algorithm to determine the optimal solution.

Algorithm 4.46 (Alternative approach)
A priori calculation before observing the next filter value:

- Compute the edges $\eta^{(0)} := r\mathbb{1}$ and $\eta^{(i)}(t) := r\mathbb{1} + (1-\alpha)D(t)\cdot e_i$.
- Determine the regions $\mathbb{D}^{(i,I)}$ according to Propositions 4.41 and 4.42 and Example 4.44.

A posteriori calculation after observing the next filter value μ.

- Find the Case (i, I) with $\mu \in \mathbb{D}^{(i,I)}$.
- Calculate the optimal solutions $\lambda^{(i,I)}$ and $\pi^{(i,I)}$.

Theorem 4.47
The above algorithm yields the optimal solution.

Proof. This is a trivial consequence from the proof of the original Algorithm 4.34. □

Remark 4.48
Although Algorithm 4.46 seems explicit and straight forward, its computational effort will usually be much larger than with the original Algorithm 4.34. The main problem is to determine the correct case (i, I) with $\mu \in \mathbb{D}^{(i,I)}$. This can be arbitrarily complicated if the number of dimensions is larger than two.

4.5 The complete portfolio optimization problem

In the previous sections we solved our constrained portfolio optimization problem 2.3 under partial information using a localization argument that we called a reverse separation approach. We separated the region $\mathbb{D}$ of admissible values for the dual process λ into several regions according to the different cases. In these regions we successfully applied a stochastic control argument and solved it for the locally optimal dual processes.

In this section we sketch how to derive the globally optimal solution. Therefore note that if the globally optimal solution is in the interior of some specific region then it should locally behave like the locally optimal solution of this specific region since both maximize the same objective.

Additionally we need to specify what happens if we are not in the interior of some region but on the boundary between some of them. However since there are only finitely many regions that are all polyhedrons, their boundaries form a null set in the admissibility region. But our objective is to optimize an expectation value hence null sets do not matter.

Fortunately the locally optimal solutions of the different regions coincide in their limits when approaching any point on the boundaries of their admissibility regions. Therefore we may combine the respective locally optimal solution to get a continuous version of the globally optimal solution.

Remark 4.49 (the global value function)
Of course the value functions corresponding to the localized problems do not contain much information for the global problem since they only give the expected utility of terminal wealth if the optimal strategy would never leave the current case until terminal time, or if one continues sub-optimally with the strategy of the current case after leaving the region, cf. Remarks 4.9 and 4.10.

Also one cannot expect to be able to somehow combine these value functions to get the global value function. From an intuitive point of view the global value function has to be some weighted mean of the local value functions, weighted by the possible paths that the optimal solution can follow through the different cases.

In fact we expect it to be very hard to derive any analytical form of this value function. Anyway we expect the value function to be continuously differentiable since the underlying processes don't have any irregularities and we have continuity of the strategies at the boundaries.

To simplify notation we first consider the following definition.

Definition 4.50
Define the set of all possible cases $\mathcal{C}$ of the optimal solution:

$$\mathcal{C} := \big\{(1,\cdot),(2,\cdot)\big\} \cup \big\{(3,I)\big|\emptyset \neq I \subset \{1,...,d\}\big\} \cup \big\{(4,I)\big|\emptyset \neq I \subseteq \{1,...,d\}\big\}$$

And define $C(t,\mu) \in \mathcal{C}$ to be the current case given the current parameters (t,μ). Then we observe

$$(t,\mu) \in \mathbb{D}_+^{C(t,\mu)} \subset \bigcup_{C\in\mathcal{C}} \mathbb{D}_+^C = [0,T] \times \mathbb{R}^d$$

$$\text{and} \qquad \mu \in \mathbb{D}_t^{C(t,\mu)} \subset \bigcup_{C \in \mathcal{C}} \mathbb{D}_t^C = \mathbb{R}^d.$$

where the sets $\mathbb{D}_+^C$ and $\mathbb{D}_t^C$ are defined in Section 4.4.1.

Steps to verifying the global solution

Consider the composition π_{loc} of the locally optimal portfolio strategies π^C:

$$\pi_{\text{loc}}(t,\mu) = \sum_{C \in \mathcal{C}} \mathbb{1}_{C(t,\mu)=C} \cdot \pi^C(t,\mu)$$

Then the globally optimal portfolio strategy π^* is given by a continuous modification of π_{loc}. However the complete rigorous proof to verify the optimality of the global strategy will be very circumstantial and challenging since several details are very hard to prove accurately. Therefore we will only give the framework and provide the relevant steps.

Step 1: the dual problem.
By Lemma 4.39 the locally optimal portfolio strategy is linear in the locally optimal dual process. By its generalization, Proposition 5.2, this also holds true for the globally optimal solutions.
Therefore we only need to show that the composition λ_{loc} of the locally optimal dual processes as given below is equal to the globally optimal dual process λ. The locally optimal solutions λ^C are derived point-wisely in Section 4.2 and the globally optimal process is defined by the duality Theorem 2.11:

$$\lambda_{\text{loc}}(t,\mu) = \sum_{C \in \mathcal{C}} \mathbb{1}_{C(t,\mu)=C} \cdot \lambda^C(t,\mu)$$
$$\lambda = \arg\max_{\nu \in \mathcal{D}} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}}\right].$$

Step 2: the stochastic control approach.
The stochastic control approach on the dual optimization problem uses the value function V to determine the optimal dual process λ for some given parameters (t,h,μ).

$$V(t,h,\mu) = \sup_{\nu \in \mathcal{D}(t,h,\mu)} \mathbb{E}\left[(H_T^\nu)^{\frac{\alpha}{\alpha-1}} \,\middle|\, (H_t^\nu, \hat{\mu}_t) = (h,\mu)\right]$$

We have already seen in Chapter 4 that the value of h is irrelevant to any further optimization and in Section 4.4.1 that the set of possible values for (t,μ) can be partitioned into subregions corresponding to the respective cases $C \in \mathcal{C}$. This can also be shown for the HJB-equation resulting from this global value function.

Step 3: the Bellman principle.
Note that the current parameters $(t,\mu) = (t,\hat{\mu}_t)$ are almost surely in the interior of the region of some case $(\mathbb{D}^{C(t,\mu)})^\circ$ since all of the finitely many regions $\mathbb{D}^C$ for $C \in \mathcal{C}$ are

polyhedrons and hence the union of their boundaries form a null set in the admissibility region.

Now consider the stopping time τ:

$$\tau := \inf\left\{s > t \,\middle|\, (s, \hat{\mu}_s) \notin (\mathbb{D}^{C(t,\mu)})^\circ\right\} \wedge T$$

The stopping time τ stops whenever the parameters (t, μ) get on the boundary between some cases.

According to the Bellman principle we get to solve the optimization problem by acting optimally until the stopping time τ and continuing optimally afterwards. Since we are in a Markovian setting there are no dependencies or correlations of the market parameters over time. Only the filter itself gets updated using past information, but the current dynamics (and hence the optimal control) are not affected by the history of the filter, but just by its dynamics. Hence we get for the value function

$$V(t, h, \mu) = \sup_{\nu \in \mathcal{D}(t,h,\mu)} \mathbb{E}\left[V\big(\tau, H_\tau^\nu, \hat{\mu}_\tau\big)\,\big|\,(H_t^\nu, \hat{\mu}_t) = (h, \mu)\right]$$

where the optimal control ν^* in $[t, \tau)$ is given by solving the corresponding HJB-equation before leaving the region of the current case $\mathbb{D}^{C(t,\mu)}$.

Step 4: deriving the optimal solution.
But now this is just one of the separated problems dealt with in Section 4.2. Since the parameters $(t, \mu) = (t, \hat{\mu}_t)$ do not leave the current case in $[t, \tau)$ we can use the optimal control already derived in the respective case in Section 4.2.

Therefore this local optimizer is also the global optimizer.

Step 5: a continuous modification.
So far we derived the optimal control for all current parameters that are in the interior of the region of some case $(t, \mu) \in (\mathbb{D}^{C(t,\mu)})^\circ$. However since all of the finitely many regions $\mathbb{D}^C$ for $C \in \mathcal{C}$ are polyhedrons the union of their boundaries form a null set in the admissibility region. Therefore the current parameters are almost surely not on some boundary and the exact choice of the optimal control inside this null set is irrelevant for optimizing an expectation value. Also note that the filter $\hat{\mu}_t$ essentially behaves like a Brownian motion and it is well known that if some Brownian motion hits some null set, then it will hit this null set uncountably often but the set of hitting times is still a Lebesgue-null set and hence irrelevant for maximizing an expectation value.

Since we derived in the previous sections that the optimal local solutions coincide with each other on the boundaries of their admissibility regions we can choose to use the continuous modification of the composition of the local solutions as the global solution.

Nevertheless, to make the above arguments rigorous a lot of technical details need to be proved.

5 Extensions

In this chapter we show some additional interesting results and we consider in particular several further types of constraints on the optimal portfolio strategy.

Remark 5.1
The optimal constrained portfolio strategy in our setting is given by the optimal unconstrained strategy in the auxiliary market $\mathcal{M}^\lambda$, given the optimal dual process λ. For logarithmic utility we have seen that the optimal strategy in this auxiliary market is the Merton plug-in strategy $\Sigma^{-1}(\hat{\mu}_t + \lambda_t - r\mathbb{1})$. For power utility Lemma 4.39 shows that the optimal solution in the auxiliary market is equal to the optimal unconstrained solution in the original market with $\hat{\mu} + \lambda$ plugged in instead of $\hat{\mu}$. Now we want to generalize this result to arbitrary utility functions.

Similarly to Section 4.1.2 we consider the maximization of expected utility of terminal wealth in the original unconstrained market $\mathcal{M}^0$ where the value function of the corresponding stochastic control problem is

$$V(t,x,\mu) = \sup_{\pi \in \mathcal{A}(t,x,\mu)} \mathbb{E}\left[U(X_T) \middle| (X_t, \hat{\mu}_t) = (x,\mu) \right]$$

Under suitable conditions we derived the following HJB-equation:

$$0 = \sup_{\pi} \left(V_t + V_x x \left(r + \pi^\top (\mu - r\mathbb{1}) \right) + \frac{1}{2} x^2 \pi^\top \Sigma \pi V_{xx} + x \pi^\top \gamma_t V_{x\mu} + \frac{1}{2} \mathrm{tr}\left(V_{\mu\mu} \gamma_t \Sigma^{-1} \gamma_t \right) \right)$$

with boundary condition $V(T,x,\mu) = U(x)$ and with optimal control π_{unc}

$$\pi_{\text{unc}}(t,\mu,r,x) = -\frac{V_x}{x V_{xx}} \Sigma^{-1}(\mu - r\mathbb{1}) - \frac{1}{x V_{xx}} \Sigma^{-1} \gamma_t V_{x\mu} \tag{5.1}$$

After plugging in π_{unc} into the above HJB-equation we get:

$$\begin{aligned} 0 = {} & V_t + r x V_x - \frac{1}{2} (\mu - r\mathbb{1})^\top \Sigma^{-1} (\mu - r\mathbb{1}) \frac{V_x^2}{V_{xx}} \\ & - \frac{1}{2} \frac{1}{V_{xx}} V_{x\mu}^\top \gamma_t \Sigma^{-1} \gamma_t V_{x\mu} - \frac{V_x}{V_{xx}} (\mu - r\mathbb{1})^\top \Sigma^{-1} \gamma_t V_{x\mu} + \frac{1}{2} \mathrm{tr}\left(V_{\mu\mu} \gamma_t \Sigma^{-1} \gamma_t \right) \end{aligned} \tag{5.2}$$

This is a deterministic partial differential equation for the value function V, depending in particular on the parameters t, μ, r and x. After solving this PDE for V we derive an explicit solution for π_{unc}.

Proposition 5.2 (shifted unconstrained solution)
Consider the constrained portfolio optimization problem (2.3) under convex constraints K with a utility function $U(x)$ such that the optimal unconstrained solution (5.1) can be derived and verified via solving the HJB-equation (5.2). Let the corresponding optimal dual process λ of Theorem 2.11 exist and be well-defined as $\lambda_t = \lambda(t, \hat{\mu}_t)$.
Then the optimal K-constrained solution π^λ is equal to the optimal unconstrained solution augmented by λ, i.e.

$$\pi^\lambda(t, \hat{\mu}_t, r) = \pi^{+\lambda}_{\text{unc}}(t, \hat{\mu}_t, r) := \pi_{\text{unc}}\big(t, \hat{\mu}_t + \lambda_t + \delta(\lambda_t)\mathbb{1}, r + \delta(\lambda_t)\big)$$

If $\hat{\mu}_t$ and r only appear together in the form $\hat{\mu}_t - r\mathbb{1}$, the above formula reduces to

$$\pi^\lambda(t, \hat{\mu}_t) = \pi^{+\lambda}_{\text{unc}}(t, \hat{\mu}_t) = \pi_{\text{unc}}\big(t, \hat{\mu}_t + \lambda_t\big)$$

Proof. π^λ is the optimal constrained strategy of the original market $\mathcal{M}^0$ or by dual optimality of [CK92] the optimal unconstrained strategy in the auxiliary market $\mathcal{M}^\lambda$, given by:

$$\begin{aligned}
dB_t^\lambda &= B_t^\lambda\big(r + \delta\left(\lambda(t, \hat{\mu}_t)\right)\big)dt \\
dS_t^\lambda &= \operatorname{diag}\left(S_t^\lambda\right)\big(\left(\hat{\mu}_t + \lambda(t, \hat{\mu}_t) + \delta(\lambda(t, \hat{\mu}_t))\mathbb{1}\right)dt + \sigma dV_t\big) \\
\Rightarrow \quad dX_t^\lambda &= X_t^\lambda\left(\left(r + \delta\left(\lambda(t, \hat{\mu}_t)\right) + \pi_t^\top\left(\hat{\mu}_t + \lambda(t, \hat{\mu}_t) - r\mathbb{1}\right)\right)dt + \pi_t^\top \sigma dV_t\right) \\
\Rightarrow \quad \pi^\lambda &= \operatorname*{arg\,max}_{\pi \in \mathcal{A}} \mathbb{E}\left[U(X_T^\lambda)\right]
\end{aligned}$$

In the corresponding stochastic control problem we need to determine the value function

$$\bar{V}(t, x, \mu) = \sup_{\pi \in \mathcal{A}(t,x,\mu)} \mathbb{E}\left[U(X_T^\lambda)\,\Big|\,(X_t, \hat{\mu}_t) = (x, \mu)\right]$$

Under suitable conditions we derive the following HJB-equation:

$$\begin{aligned}
0 = \sup_\pi \Bigg(&\bar{V}_t + \bar{V}_x x\Big(r + \delta(\lambda(t, \mu)) + \pi^\top\left(\mu - r\mathbb{1} + \lambda(t, \mu)\right)\Big) \\
&+ \frac{1}{2}\pi^\top \Sigma \pi x^2 \bar{V}_{xx} + \pi^\top \gamma_t x \bar{V}_{x\mu} + \frac{1}{2}\operatorname{tr}\left(\bar{V}_{\mu\mu}\gamma_t \Sigma^{-1}\gamma_t\right)\Bigg)
\end{aligned}$$

with boundary condition $\bar{V}(T, x, \mu) = U(x)$ and with optimal control π^λ

$$\pi^\lambda(t, \mu, r) = -\frac{\bar{V}_x}{x\bar{V}_{xx}}\Sigma^{-1}\left(\mu - r\mathbb{1} + \lambda(t, \mu)\right) - \frac{1}{x\bar{V}_{xx}}\Sigma^{-1}\gamma_t \bar{V}_{x\mu} \tag{5.3}$$

After plugging in π^λ into the above HJB-equation we get to solve:

$$0 = \bar{V}_t + (r + \delta(\lambda(t, \mu)))x\bar{V}_x - \frac{1}{2}\left(\mu - r\mathbb{1} + \lambda(t, \mu)\right)^\top \Sigma^{-1}\left(\mu - r\mathbb{1} + \lambda(t, \mu)\right)\frac{\bar{V}_x^2}{\bar{V}_{xx}} \tag{5.4}$$

$$-\frac{1}{2}\frac{1}{\bar{V}_{xx}}\bar{V}_{x\mu}^{\top}\gamma_t\Sigma^{-1}\gamma_t\bar{V}_{x\mu}-\frac{\bar{V}_x}{\bar{V}_{xx}}\left(\mu-r\mathbb{1}+\lambda(t,\mu)\right)^{\top}\Sigma^{-1}\gamma_t\bar{V}_{x\mu}+\frac{1}{2}\mathrm{tr}\left(\bar{V}_{\mu\mu}\gamma_t\Sigma^{-1}\gamma_t\right)$$

The HJB-equation (5.4) is just the same as equation (5.2) if we update their parameters r by $r+\delta(\lambda(t,\mu))$ and μ by $\mu+\lambda(t,\mu)+\delta(\lambda(t,\mu))$. (This is obviously possible as at the fixed time t of the HJB-equation these are deterministic.) Hence $\bar{V}$ is just the solution V of (5.2) with updated parameters.
Additionally then (5.3) is the same formula as (5.1) with updated parameters, such that $\pi^{\lambda}(t,\mu,r)=\pi_{\mathrm{unc}}(t,\mu+\lambda(t,\mu)+\delta(\lambda(t,\mu)),r+\delta(\lambda(t,\mu)))$. □

5.1 Further Constraints

In the previous Chapters 3 and 4 we used the no short-selling and the no borrowing constraints together. Note that using only one of these constraints would just result in a special case of our derived results.

When only using the no short-selling constraint, the optimal constrained solution is determined by just using the cases 1-3. The former domains $\mathbb{D}^{(4,I)}$ of the value function resulting from Case 4 become part of the domain $\mathbb{D}^{(1,\cdot)}$. This can be seen in Section 4.2.1 and in particular in the interpretations in Section 4.4.

In case of only using the no borrowing constraint, our setting even reduces further to only using Cases 1 and $(4, I)$ with full $I = \{1, ..., d\}$. This follows in particular from the linear form of the dual process $\lambda = -c\mathbb{1}$ as derived below. Hence this is a particularly simple case that can be generalized further:

We consider the constraint where investment is allowed up to a fraction $k \geq 0$ of the wealth. Hence $K := \{\pi | \mathbb{1}^\top \pi \leq k\}$ and the no borrowing constraint follows for $k = 1$. The support function is then given by

$$\begin{aligned}\delta(\nu) &= \sup_{\pi \in K}(-\pi^\top \nu) \\ &= \begin{cases} ck & \text{if } \nu = -c\mathbb{1} \text{ for } c \geq 0 \\ \infty & \text{else} \end{cases} \\ &< \infty \quad \text{for } \nu \in \tilde{K} = \{-c\mathbb{1} | c \geq 0\}\end{aligned}$$

Under logarithmic utility we have to solve (if $\pi^M \notin K$)

$$\begin{aligned} & \lambda_t = \arg\min_{\nu \in \tilde{K}} \left(\delta(\nu) + \frac{1}{2}(\hat{\mu}_t + \nu - r)^\top \Sigma^{-1} (\hat{\mu}_t + \nu - r) \right) \\ \Rightarrow \quad & \lambda_t = -c_t \mathbb{1} \\ \text{where} \quad & c_t = \arg\min_{c \geq 0} \left(ck + \frac{1}{2}(\hat{\mu}_t - c\mathbb{1} - r\mathbb{1})^\top \Sigma^{-1} (\hat{\mu}_t - c\mathbb{1} - r\mathbb{1}) \right) \\ & \quad = g\mathbb{1}^\top \Sigma^{-1}(\hat{\mu}_t - r\mathbb{1}) - gk \qquad \left(> 0 \ \Leftrightarrow \ \mathbb{1}^\top \pi^M > k \right) \\ \Rightarrow \quad & \lambda_t = gk\mathbb{1} - g\mathbb{1}^\top \Sigma^{-1}(\hat{\mu}_t - r\mathbb{1})\mathbb{1} \quad , \quad g := (\mathbb{1}^\top \Sigma^{-1} \mathbb{1})^{-1} \end{aligned}$$

Under power utility we get from the previous chapter in Case 4 with full I

$$\begin{aligned} & \lambda_t = -c_t \mathbb{1} \\ \text{where} \quad & c_t = -(1-\alpha)gk + g\mathbb{1}^\top \Sigma^{-1} C'(t) \gamma_t^{-1} (\mu - r\mathbb{1}) + \alpha k (T-t) g^2 \mathbb{1}^\top \Sigma^{-1} C'(t) \Sigma^{-1} \mathbb{1} \\ \text{with} \quad & C'(t) = \left(\gamma_t^{-1} - \frac{\alpha}{1-\alpha}(T-t)\Sigma^{-1}\big(Id - g\mathbb{1}\mathbb{1}^\top \Sigma^{-1}\big) \right)^{-1} \end{aligned}$$

Note that for $\alpha = 0$ we get the logarithmic solution.

In both cases we observe an explicit optimal dual process λ and hence an explicit optimal strategy π^λ.

5.1.1 L^p-constraints

Proposition 5.3
Under general L^p-constraints with $K := \{\pi \mid \|\pi\|_p \leq k\}$ for $p \in [1,\infty]$ we get

$$\delta(\nu) = k\,\|\nu\|_q$$

for $\nu \in \tilde{K} = \mathbb{R}^d$ where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. Let $p = 1$. Then $K := \left\{\pi \;\middle|\; \sum |\pi_i| \leq k\right\}$. Hence

$$\begin{aligned}\delta(\nu) &= \sup_{\pi\in K}(-\pi^\top\nu) = \sup_{\pi\in K}(\pi^\top\nu) \qquad \text{since } K \text{ is symmetric}\\ &\leq \sup_{\pi\in K}(\pi^\top\mathbb{1}\,\|\nu\|_\infty)\\ &\leq k\,\|\nu\|_\infty \quad < \infty \quad \text{for } \nu \in \tilde{K} = \mathbb{R}^d\end{aligned}$$

On the other hand choose $\pi^* = \operatorname{sgn}(\nu_i)ke_i$ where $i = \arg\max_j |\nu_j|$. Then

$$\delta(\nu) = \sup_{\pi\in K}(\pi^\top\nu) \geq (\pi^*)^\top\nu = k\,|\nu_i| = k\,\|\nu\|_\infty$$

Hence $\delta(\nu) = k\,\|\nu\|_\infty$.

Now let $p = \infty$. Then $K := \left\{\pi \;\middle|\; \max_i |\pi_i| \leq k\right\}$. Hence

$$\begin{aligned}\delta(\nu) &= \sup_{\pi\in K}(-\pi^\top\nu) = \sup_{\pi\in K}(\pi^\top\nu) \qquad \text{since } K \text{ is symmetric}\\ &\leq \sup_{\pi\in K}\left(\sum_i |\pi_i\nu_i|\right)\\ &\leq \sup_{\pi\in K}\left(\|\pi\|_\infty \sum_i |\nu_i|\right)\\ &= k\,\|\nu\|_1 \quad < \infty \quad \text{for } \nu \in \tilde{K} = \mathbb{R}^d\end{aligned}$$

On the other hand choose π^* with $\pi_i^* = \operatorname{sgn}(\nu_i)k$. Then

$$\delta(\nu) = \sup_{\pi\in K}(\pi^\top\nu) \geq (\pi^*)^\top\nu = \sum_i k\,|\nu_i| = k\,\|\nu\|_1$$

Hence $\delta(\nu) = k\,\|\nu\|_1$.

Now let $p \in (1,\infty)$.

$$\begin{aligned}\delta(\nu) &= \sup_{\pi\in K}(-\pi^\top\nu) = \sup_{\pi\in K}(\pi^\top\nu) \qquad \text{by symmetry of } K\\ &= \sup_{\pi\in K}\left(\sum_i \pi_i\nu_i\right) \leq \sup_{\pi\in K}\left(\sum_i |\pi_i\nu_i|\right)\end{aligned}$$

$$\begin{aligned}
&\leq \sup_{\pi\in K}\left(\|\pi\|_p \|\nu\|_q\right) \qquad \text{by Hölder with } \frac{1}{p}+\frac{1}{q}=1\\
&= k\|\nu\|_q \quad < \infty \quad \text{for } \nu \in \mathbb{R}^d = \tilde{K}
\end{aligned}$$

On the other hand choose π^* with $\pi_i^* = \text{sgn}(\nu_i)k\left|\frac{\nu_i}{\|\nu\|_q}\right|^{q/p}$ such that

$$\|\pi^*\|_p^p = \sum_i |\pi_i^*|^p = \sum_i \left|k\left|\frac{\nu_i}{\|\nu\|_q}\right|^{q/p}\right|^p = \sum_i k^p \frac{|\nu_i|^q}{\|\nu\|_q^q} = k^p$$

Hence $\pi \in K$ and

$$\begin{aligned}
\delta(\nu) &= \sup_{\pi\in K}\left(\sum_i \pi_i\nu_i\right)\\
&\geq \sum_i \pi_i^*\nu_i\\
&= \sum_i \text{sgn}(\nu_i)k\left|\frac{\nu_i}{\|\nu\|_q}\right|^{q/p} \nu_i\\
&= k\sum_i \left|\frac{\nu_i}{\|\nu\|_q}\right|^{q/p} |\nu_i|\\
&= k\sum_i \frac{|\nu_i|^q}{\|\nu\|_q^{q-1}} \quad \text{by } \frac{q}{p} = q-1\\
&= k\|\nu\|_q
\end{aligned}$$

Hence $\delta(\nu) = k\|\nu\|_q$. □

Remark 5.4
The δ-function is differentiable for $p \in (1,\infty)$ with partial derivative:

$$\begin{aligned}
\partial_{x_i}\|x\|_p &= \frac{1}{p}\left(\sum_j |x_j|^p\right)^{(1-p)/p} p\,|x_i|^{p-1}\,\text{sgn}(x_i)\\
&= \left(\frac{|x_i|}{\|x\|_p}\right)^{p-1} \text{sgn}(x_i)\cdot \mathbb{1}_{\{x\neq 0\}}
\end{aligned}$$

where $\text{sgn}(x) = \begin{cases} 1, & x>0\\ 0, & x=0\,.\\ -1, & x<0\end{cases}$

Hence by Theorem 2.9 we get the optimal constrained solution of the original problem as the optimal unconstrained solution in the auxiliary market $\mathcal{M}^\lambda$ with (the yet unknown) optimal $\lambda \neq 0$:

$$\begin{aligned} \pi^\lambda &= -\delta'(\lambda) \\ \Rightarrow \quad \pi_i^\lambda &= -k\partial_i \|\lambda\|_q = -k\left(\frac{|\lambda_i|}{\|\lambda\|_q}\right)^{q-1} \operatorname{sgn}(\lambda_i) \\ \Rightarrow \quad \pi^\lambda &= -k\frac{1}{\|\lambda\|_q^{q-1}} \cdot L \quad \text{where } L_i := |\lambda_i|^{q-1}\operatorname{sgn}(\lambda_i) \end{aligned}$$

hence $\pi^\lambda \propto L$.

Proposition 5.5
For $p = 2$ we can solve the constrained portfolio optimization problem explicitly up to one trivial numerical step.

Proof. Since $\lambda = 0$ if and only if $\pi_{\text{unc}} \in K$, we only need to consider $\lambda \neq 0$ with $\pi_{\text{unc}} \notin K$.
From Proposition 5.3 we get $K := \{\pi | \, \|\pi\|_2 \leq k\}$, $\tilde{K} = \mathbb{R}^d$ and $\delta(\nu) = k\|\nu\|_2$.
In particular we get $\delta'(\nu) = \frac{k\nu}{\|\nu\|_2} \cdot \mathbb{1}_{\nu\neq 0}$.
By Remark 5.4 we get the optimal constrained solution as the optimal unconstrained solution in the auxiliary market $\mathcal{M}^\lambda$:

$$\pi^\lambda = -\delta'(\lambda) = -\frac{k}{\|\lambda\|_2}\lambda \quad \Rightarrow \quad \pi^\lambda \propto \lambda.$$

1. Under logarithmic utility we know:

$$\begin{aligned} \pi^\lambda &= \Sigma^{-1}(\mu - r\mathbb{1} + \lambda) \\ \Rightarrow \quad -\frac{k}{\|\lambda\|_2}\lambda &= \Sigma^{-1}(\mu - r\mathbb{1} + \lambda) \\ \Rightarrow \quad -\left(\frac{k}{\|\lambda\|_2}Id + \Sigma^{-1}\right)\lambda &= \Sigma^{-1}(\mu - r\mathbb{1}) \\ \Rightarrow \quad \lambda &= -\left(\frac{k}{\|\lambda\|_2}Id + \Sigma^{-1}\right)^{-1}\Sigma^{-1}(\mu - r\mathbb{1}) \\ &= -\left(\frac{k}{\|\lambda\|_2}\Sigma + Id\right)^{-1}(\mu - r\mathbb{1}) \end{aligned}$$

If we can determine $x := \|\lambda\|_2 > 0$ then

$$\lambda = -\left(\frac{k}{x}\Sigma + Id\right)^{-1}(\mu - r\mathbb{1}) \tag{5.5}$$

$$\Rightarrow \quad \pi^\lambda = -\frac{k}{x}\lambda = \left(\Sigma + \frac{x}{k}Id\right)^{-1}(\mu - r\mathbb{1}) \tag{5.6}$$

Now we observe

$$x = \|\lambda\|_2 = \left\|\left(\frac{k}{x}\Sigma + Id\right)^{-1}(\mu - r\mathbb{1})\right\|_2$$

$$\Leftrightarrow \quad k = \left\| \left(\Sigma + \frac{x}{k} Id \right)^{-1} (\mu - r\mathbb{1}) \right\|_2 = \left\| \pi^\lambda \right\|_2 =: \varphi(x)$$

This cannot be solved explicitly, even for $d = 2$.
However this is trivially solved numerically as $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is one-dimensional, continuous and monotone decreasing with $\lim_{x\to\infty} \varphi(x) = 0 < k$ and $\varphi(0) = \left\|\pi^M\right\|_2 > k$ since $\pi^M \notin K$. The optimal x^* with $\varphi(x^*) = k$ is unique.

2. Under power utility we observe similarly, using Proposition 5.2:

$$-\frac{k}{\|\lambda\|_2}\lambda = \pi^\lambda = \frac{1}{1-\alpha}\Sigma^{-1}C(t)\gamma_t^{-1}(\mu + \lambda - r\mathbb{1})$$

$$\Rightarrow \qquad \lambda = -\left((1-\alpha)\frac{k}{\|\lambda\|_2} Id + \Sigma^{-1}C(t)\gamma_t^{-1}\right)^{-1} \Sigma^{-1}C(t)\gamma_t^{-1}(\mu - r\mathbb{1})$$

$$= -\left(\frac{k}{\|\lambda\|_2}\Big((1-\alpha)\Sigma - \alpha(T-t)\gamma_t\Big) + Id\right)^{-1} (\mu - r\mathbb{1})$$

If we can determine $x := \|\lambda\|_2 > 0$ then

$$\lambda = -\left(\frac{k}{x}\Big((1-\alpha)\Sigma - \alpha(T-t)\gamma_t\Big) + Id\right)^{-1} (\mu - r\mathbb{1}) \tag{5.7}$$

$$\Rightarrow \quad \pi^\lambda = -\frac{k}{x}\lambda = \left(\Big((1-\alpha)\Sigma - \alpha(T-t)\gamma_t\Big) + \frac{x}{k} Id\right)^{-1} (\mu - r\mathbb{1}) \tag{5.8}$$

Now we observe

$$x = \|\lambda\|_2 = \left\| \left(\frac{k}{x}\Big((1-\alpha)\Sigma - \alpha(T-t)\gamma_t\Big) + Id\right)^{-1} (\mu - r\mathbb{1}) \right\|_2$$

$$\Leftrightarrow \quad k = \left\| \left(\Big((1-\alpha)\Sigma - \alpha(T-t)\gamma_t\Big) + \frac{x}{k} Id\right)^{-1} (\mu - r\mathbb{1}) \right\|_2 = \left\|\pi^\lambda\right\|_2 =: \varphi(x)$$

Again, this can only be solved numerically. $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is one-dimensional, continuous and monotone decreasing with $\lim_{x\to\infty} \varphi(x) = 0 < k$ and $\varphi(0) = \|\pi_{\text{unc}}\|_2 > k$ since $\pi_{\text{unc}} \notin K$. The optimal x^* with $\varphi(x^*) = k$ is unique.

Having numerically determined the optimal $x^* = \|\lambda\|_2$ we get the optimal dual process via the formulas (5.5) and (5.7), respectively the optimal portfolio strategies via the formulas (5.6) and (5.8). □

Remark 5.6
For $p = 1$ we get an L^1-constrained portfolio strategy that generalizes the no-borrowing case. [DV15] also consider L^1-constraints in their setting and refer to several more papers dealing with these constraints.

We want to consider the case of L^1-constraints as a generalization of Case 4 in our setting. Therefore we repeat the stochastic control approach as in Section 4.2 until we observe the HJB-equation (4.13). Under the L^1-constraints there are only two cases to consider:

Case 1) $\delta(\lambda) = 0$ if and only if $\lambda = 0$ if and only if $\pi_{\text{unc}} \in K$

Case 2) $\delta(\lambda) > 0$ if and only if $\bar{\lambda} := \max_k |\lambda_k| > 0$

Let's abbreviate the calculations of the second case with $\lambda \in \mathbb{R}^d \setminus \{0\}$. There is

- $I \subseteq \{1, ..., d\}$ such that $|\lambda_i| = \bar{\lambda}$ for all $i \in I$ where
  - $I^+ \subseteq I$ such that $\lambda_i = -\bar{\lambda} < 0$ for all $i \in I^+$
  - $I^- \subseteq I$ such that $\lambda_i = \bar{\lambda} > 0$ for all $i \in I^-$
- $J = \{1, ..., d\} \setminus I$ such that $|\lambda_j| < \bar{\lambda}$ for all $j \in J$.

We will call I and J the *active and passive dimensions* and I^+ and I^- the *positive and negative active dimensions*. In our finite dimensional setting the distance between $\bar{\lambda}$ and the $\lambda_j, j \in J$ is larger than zero, hence there is some ε-ball around (t, h, μ) such that the active dimensions don't switch in $\lambda = \lambda(t, h, \mu)$. We observe that this case has to be divided further into several sub-cases, depending on which dimensions are active or passive. (By definition the active dimensions I cannot be empty.)

We solve the HJB-equation (4.13) for some fixed choice of active dimensions just like in Section 4.2.2, hence we consider an auxiliary problem on a subspace of $\mathbb{R}^d$, such that $\delta(\lambda) = k\bar{\lambda}$ is differentiable in λ. Without loss of generality we simplify notation by assuming $1 \in I^-$. (If I^- was empty, we'd be in the well known case of only having a no-borrowing constraint.)

First define $G := \left\{ x \in \mathbb{R} \times \mathbb{R}^{|J|} \middle| x_1 > 0, |x_j| < x_1 \text{ for all } j \in J \right\}$ and the linear bijection φ via:

$$\varphi : G \to H : \begin{pmatrix} x_1 \\ x_J \end{pmatrix} \mapsto \begin{pmatrix} x_1 \mathbb{1}_I^{\pm} \\ x_J \end{pmatrix} \qquad \text{where } \mathbb{1}_I^{\pm} := \begin{pmatrix} \mathbb{1}_{I^-} \\ -\mathbb{1}_{I^+} \end{pmatrix}$$

where $H := \text{Im}(\varphi) \subset \mathbb{R}^d$. Obviously φ is bijective and linear, where 'linear' means that both φ and $(\varphi)^{-1}$ can be written using matrices:

$$\varphi(x) = \Phi x \quad \text{where } \Phi := \begin{pmatrix} \mathbb{1}_I^{\pm} & 0_{IJ} \\ 0_J & Id_J \end{pmatrix} \in \mathbb{R}^{d \times 1+|J|}$$

$$\text{and } \ (\varphi)^{-1}(x) = \Psi x \quad \text{where } \Psi := \begin{pmatrix} e_1^{\top} & 0_J^{\top} \\ 0_{JI} & Id_J \end{pmatrix} \in \mathbb{R}^{1+|J| \times d}$$

Since $\lambda \in \text{Im}(\varphi)$ we may define $\tilde{\lambda} := (\varphi)^{-1}(\lambda) = \begin{pmatrix} \lambda_1 \\ \lambda_J \end{pmatrix} \in G$. Also $\lambda(t, h, \mu) \in \text{Im}(\varphi)$ for any (t, h, μ) in the ε-ball given above, hence we may reduce our problem from finding the optimal $\lambda \in H$ to finding the optimal $\tilde{\lambda} \in G$.

Additionally we get the following nice property in G: Although δ is not differentiable with respect to λ we observe that $\delta \circ \varphi$ is differentiable with respect to $\tilde{\lambda}$ on G with

$$(\delta \circ \varphi)(\tilde{\lambda}) = \delta \left(\begin{pmatrix} \lambda_1 \mathbb{1}_I^{\pm} \\ \lambda_J \end{pmatrix} \right) = k\, |\lambda_1| = k\lambda_1$$

$$\Rightarrow \quad (\delta \circ \varphi)'(\tilde{\lambda}) = ke_1$$

Hence we can rewrite equation (4.13) as follows:

$$0 = \sup_{\tilde{\nu} \in G} \Bigg(V_t - V_h h(r + (\delta \circ \varphi)(\tilde{\nu})) + \frac{1}{2} h^2 \left(\mu - r\mathbb{1} + \varphi(\tilde{\nu})\right)^\top \Sigma^{-1} \left(\mu - r\mathbb{1} + \varphi(\tilde{\nu})\right) V_{hh} \\ - h \left(\mu - r\mathbb{1} + \varphi(\tilde{\nu})\right)^\top \Sigma^{-1} \gamma_t V_{h\mu} + \frac{1}{2} \mathrm{tr}(V_{\mu\mu} \cdot \gamma_t \Sigma^{-1} \gamma_t) \Bigg)$$

with boundary condition $V(T, h, \mu) = h^{\frac{\alpha}{\alpha-1}}$ and optimal $\tilde{\lambda}$ chosen by (4.14):

$$\tilde{\lambda} = \arg\max_{\tilde{\nu} \in G} \Bigg(- h\delta(\varphi(\tilde{\nu}))V_h + \frac{1}{2} h^2 \left(\mu - r\mathbb{1} + \varphi(\tilde{\nu})\right)^\top \Sigma^{-1} \left(\mu - r\mathbb{1} + \varphi(\tilde{\nu})\right) V_{hh} \\ - h \left(\mu - r\mathbb{1} + \varphi(\tilde{\nu})\right)^\top \Sigma^{-1} \gamma_t V_{h\mu} \Bigg) \\ = \arg\max_{\tilde{\nu} \in G} \Bigg(- h(\delta \circ \varphi)(\tilde{\nu})V_h + \frac{1}{2} h^2 (\mu - r\mathbb{1} + \Phi\tilde{\nu})^\top \Sigma^{-1} (\mu - r\mathbb{1} + \Phi\tilde{\nu}) V_{hh} \\ - h(\mu - r\mathbb{1} + \Phi\tilde{\nu})^\top \Sigma^{-1} \gamma_t V_{h\mu} \Bigg)$$

We get the optimal $\tilde{\lambda}$ by differentiating:

$$0 \overset{!}{=} -hV_h ke_1 + h^2 \Phi^\top \Sigma^{-1} (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) V_{hh} - h \Phi^\top \Sigma^{-1} \gamma_t V_{h\mu}$$

$$\Rightarrow \quad \Phi^\top \Sigma^{-1} (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) = \frac{V_h}{hV_{hh}} ke_1 + \Phi^\top \Sigma^{-1} \gamma_t \frac{V_{h\mu}}{hV_{hh}} \tag{5.9}$$

In (5.9) there are $1 + |J|$ equations for $1 + |J|$ unknown variables. In the first line we get:

$$k\frac{V_h}{hV_{hh}} + \left(\mathbb{1}_I^{\pm\top} \; 0_J^\top\right) \Sigma^{-1} \gamma_t \frac{V_{h\mu}}{hV_{hh}} = \left(\mathbb{1}_I^{\pm\top} \; 0_J\right) \Sigma^{-1} (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) \\ = \left(\mathbb{1}_I^{\pm\top} (\Sigma^{-1})_{II} \quad \mathbb{1}_I^{\pm\top} (\Sigma^{-1})_{IJ}\right) (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) \tag{5.10} \\ = \mathbb{1}_I^{\pm\top} (\Sigma^{-1})_{II} (\mu_I - r\mathbb{1}_I + \lambda_1 \mathbb{1}_I^\pm) + \mathbb{1}_I^{\pm\top} (\Sigma^{-1})_{IJ} (\mu_J - r\mathbb{1}_J + \lambda_J)$$

and in the other $|J|$ lines of (5.9) we get:

$$\left(0_{JI} \; Id_J\right) \Sigma^{-1} \gamma_t \frac{V_{h\mu}}{hV_{hh}} = \left(0_{JI} \; Id_J\right) \Sigma^{-1} (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) \\ = \left((\Sigma^{-1})_{JI} \quad (\Sigma^{-1})_{JJ}\right) (\mu - r\mathbb{1} + \Phi\tilde{\lambda}) \\ = (\Sigma^{-1})_{JI} (\mu_I - r\mathbb{1}_I + \lambda_1 \mathbb{1}_I^\pm) + (\Sigma^{-1})_{JJ} (\mu_J - r\mathbb{1}_J + \lambda_J)$$

Here the matrix inversion formula (3.8) on Σ leads to

$$(\Sigma^{-1})_{JJ} (\mu_J - r\mathbb{1}_J + \lambda_J) = -(\Sigma^{-1})_{JI} (\mu_I - r\mathbb{1}_I + \lambda_1 \mathbb{1}_I^\pm) + \left(0_{JI} \quad Id_J\right) \Sigma^{-1} \gamma_t \frac{V_{h\mu}}{hV_{hh}}$$

$$\Rightarrow \qquad (\mu_J - r\mathbb{1}_J + \lambda_J) = -\left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)(\Sigma^{-1})_{JI}(\mu_I - r\mathbb{1}_I + \lambda_1 \mathbb{1}_I^{\pm}) + \left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)\left(0_{JI} \quad Id_J\right)\Sigma^{-1}\gamma_t \frac{V_{h\mu}}{hV_{hh}} \tag{5.11}$$

When plugging this into equation (5.10) we result in:

$$\begin{aligned}
&\mathbb{1}_I^{\pm\top}(\Sigma^{-1})_{II}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I^{\pm}) \\
&\quad = -\mathbb{1}_I^{\pm\top}(\Sigma^{-1})_{IJ}(\mu_J - r\mathbb{1}_J + \lambda_J) + k\frac{V_h}{hV_{hh}} + \left(\mathbb{1}_I^{\pm\top} \quad 0_J^\top\right)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\quad = \mathbb{1}_I^{\pm\top}(\Sigma^{-1})_{IJ}\left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)(\Sigma^{-1})_{JI}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I^{\pm}) \\
&\qquad - \mathbb{1}_I^{\pm\top}(\Sigma^{-1})_{IJ}\left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)\left(0_{JI} \quad Id_J\right)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\qquad + k\frac{V_h}{hV_{hh}} + \left(\mathbb{1}_I^{\pm\top} \quad 0_J^\top\right)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\quad = -\mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1}\Sigma_{IJ}(\Sigma^{-1})_{JI}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I^{\pm}) \\
&\qquad + \mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1}\Sigma_{IJ}\left(0_{JI} \quad Id_J\right)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} + k\frac{V_h}{hV_{hh}} + \left(\mathbb{1}_I^{\pm\top} \quad 0_J^\top\right)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
\Rightarrow \quad &\mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I^{\pm}) \\
&\quad = k\frac{V_h}{hV_{hh}} + \left(\mathbb{1}_I^{\pm\top} \quad \mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&\quad = k\frac{V_h}{hV_{hh}} + \left(\mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1} \quad 0_J^\top\right)\gamma_t\frac{V_{h\mu}}{hV_{hh}}
\end{aligned}$$

again due to the matrix inversion formula (3.8). In particular we get

$$\bar{\lambda} = \lambda_1 = -g\mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I + gk\frac{V_h}{hV_{hh}} + g\left(\mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1} \quad 0_J^\top\right)\gamma_t\frac{V_{h\mu}}{hV_{hh}} \tag{5.12}$$
$$\text{where} \quad g := \left(\mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1}\mathbb{1}_I^{\pm}\right)^{-1}$$

and from (5.11) we get λ_J:

$$\begin{aligned}
\lambda_J &= -(\mu - r\mathbb{1})_J - \left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)(\Sigma^{-1})_{JI}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I^{\pm}) \\
&\quad + \left(\Sigma_{JJ} - \Sigma_{JI}(\Sigma_{II})^{-1}\Sigma_{IJ}\right)\left(0_{JI} \quad Id_J\right)\Sigma^{-1}\gamma_t\frac{V_{h\mu}}{hV_{hh}} \\
&= -(\mu - r\mathbb{1})_J + \Sigma_{JI}(\Sigma_{II})^{-1}(\mu_I - r\mathbb{1}_I + \lambda_1\mathbb{1}_I^{\pm}) \\
&\quad + \left(-\Sigma_{JI}(\Sigma_{II})^{-1} \quad Id_J\right)\gamma_t\frac{V_{h\mu}}{hV_{hh}}
\end{aligned} \tag{5.13}$$

where λ and V together solve the problems (4.13) and (4.14).
Therefore we make the multiplicative ansatz $V(t,h,\mu) = h^{\frac{\alpha}{\alpha-1}}e^{f(t,\mu)}$. Then

$$\frac{V_h}{hV_{hh}} = (\alpha - 1) \quad \text{and} \quad \frac{V_{h\mu}}{hV_{hh}} = (\alpha - 1)f_\mu$$

$$
\begin{aligned}
\Rightarrow \quad \bar{\lambda} &= -g\mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I + gk(\alpha-1) + g\big(\mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1} \quad 0_J^\top\big)\gamma_t(\alpha-1)f_\mu \\
&= -(1-\alpha)gk - g\mathbb{1}_I^{\pm\top}(\Sigma_{II})^{-1}\left((\mu - r\mathbb{1})_I + (1-\alpha)(\gamma_t f_\mu)_I\right) && (5.14) \\
\lambda_J &= -(\mu - r\mathbb{1})_J - (1-\alpha)(\gamma_t f_\mu)_J \\
&\quad + \Sigma_{JI}(\Sigma_{II})^{-1}\big((\mu - r\mathbb{1})_I + \bar{\lambda}\mathbb{1}_I^{\pm} + (1-\alpha)(\gamma_t f_\mu)_I\big) && (5.15)
\end{aligned}
$$

where $f = f(t,\mu)$ solves the former HJB-equation (4.13) with $\lambda = \varphi(\tilde{\lambda}) = \begin{pmatrix} \bar{\lambda}\mathbb{1}_I^{\pm} \\ \lambda_J \end{pmatrix}$:

$$
\begin{aligned}
f_t = &-\frac{\alpha}{1-\alpha}(r+\bar{\lambda}) - \frac{1}{2}\frac{\alpha}{(1-\alpha)^2}(\mu - r\mathbb{1} + \lambda)^\top\Sigma^{-1}(\mu - r\mathbb{1} + \lambda) \\
&- \frac{\alpha}{1-\alpha}(\mu - r\mathbb{1} + \lambda)^\top\Sigma^{-1}\gamma_t f_\mu - \frac{1}{2}f_\mu^\top\gamma_t\Sigma^{-1}\gamma_t f_\mu - \frac{1}{2}\mathrm{tr}(f_{\mu\mu}\cdot\gamma_t\Sigma^{-1}\gamma_t) \qquad (5.16)
\end{aligned}
$$

This partial differential equation is solved with a quadratic ansatz for f as it is done with equation (4.35) in Section 4.2.2. This extensive calculation is left to the interested reader.

6 Simulations

In the previous chapters we considered a portfolio optimization problem under convex constraints and partial information. We derived explicit analytical formulas for the optimal portfolio strategies in various cases and an algorithm to determine the respective correct case for the optimal solutions. In this chapter we will apply these theoretical results in different practical simulations. We examine the performance of the optimal portfolio strategies under conditions as realistic as possible.

In the first two sections we consider a historical simulation on the DAX.
In Section 6.1 we describe the historical market and how it fits into our market model. We introduce several types of analysts who provide the important estimates for the filter of the growth rate and we examine our implementation of the algorithm that derives the optimal strategy. Then we derive the resulting wealth processes in Section 6.2 under various risk-aversions, with different analysts opinions and under non-zero transaction costs. In most cases we observe that our derived optimal strategies strongly out-perform their benchmark, the DAX.

In Section 6.3 we essentially repeat the same simulation but with simulated paths of a generated market. We observe the stochastic behaviour of the wealth processes and examine in particular the distribution of the terminal wealth in various settings. When compared to the benchmark $1/d$-strategy we come to the conclusion that our derived optimal strategies are worth its additional effort.

6.1 The historical market

In our first simulation we use historical data from the DAX-30-Performance-Index such that we can challenge our derived portfolio strategies in a very realistic scenario. We use data from the past 16 years as given in Figure 6.1. This period includes twice dropping and rising markets with a drop of two-thirds from 2000 to 2002, a rise of more than 200% from 2003 to 2007, another drop of 50% in 2008 and another rise of almost 200% from 2009 to 2015.

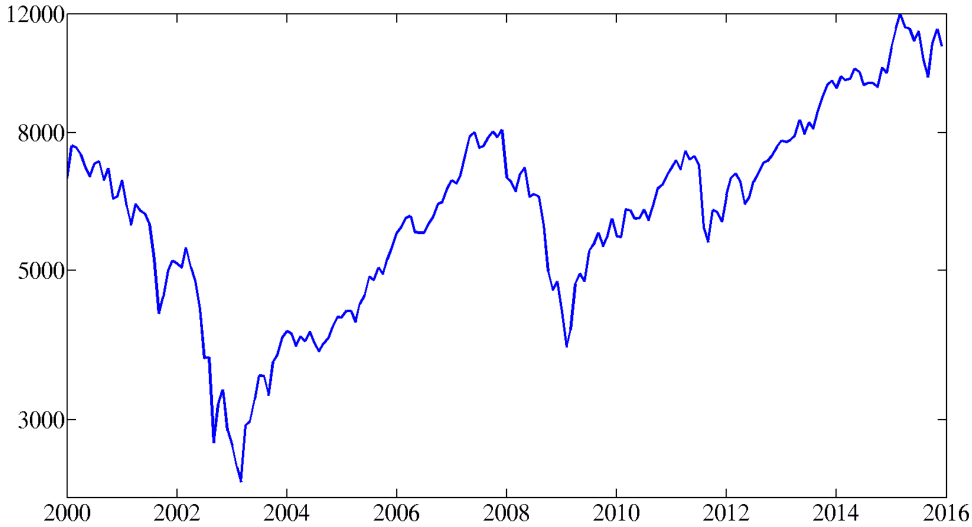


Figure 6.1: the DAX-30-Performance-Index

Since the components of the DAX may change each year, we choose those 30 stocks that have been part of the DAX mostly during the given period: Adidas, Allianz, BASF, Bayer, Beiersdorf, BMW, Commerzbank, Continental, Daimler, Deutsche Bank, Deutsche Telekom, E.ON, Fresenius, Fresenius Medical Care, HeidelbergCement, Henkel, Infineon, K+S, Linde, Lufthansa, MAN, Merck, Metro, MünchnerRück, RWE, Salzgitter, SAP, Siemens, ThyssenKrupp and Volkswagen. The historical prices of these stocks have been corrected for splits and dividend payments where the latter got reinvested in order to observe the actual performance of the respective stocks. The data is provided by 'yahoo finance'.

The investor is forced to only change her portfolio strategy once a month to avoid immediate bankruptcies by transaction costs, respectively to account for real life behaviour. She starts trading in January 2001 and stops in December 2014. During this investment period of 168 months, the DAX rose on average 3% per annum.

We also fix the risk-free interest rate at $r = 0$ to simplify notational efforts. The investor is able to observe the stock prices S_t, respectively their returns R_t and she estimates the covariance-matrix $\Sigma = \sigma\sigma^\top$ from the stock prices as follows.

$$\hat{\Sigma}_{ij} := \frac{1}{\Delta t}\frac{1}{N-1}\sum_{k=1}^{N}\left(\hat{R}^{(i)}_{t_i} - \bar{R}^{(i)}\right)\left(\hat{R}^{(j)}_{t_i} - \bar{R}^{(j)}\right)$$

$$\text{where } \hat{R}^{(i)}_{t_k} := \log\frac{S^{(i)}_{t_k}}{S^{(i)}_{t_k-\Delta t}} \quad \text{and} \quad \bar{R}^{(i)} := \frac{1}{N}\sum_{k=1}^{N}\hat{R}^{(i)}_{t_k}$$

Derivation (in one dimension). On the one hand we get from $R_t = \mu t + \sigma W_t$ that

$$(R_t - R_{t-\Delta t}) = \mu\Delta t + \sigma(W_t - W_{t-\Delta t}). \tag{6.1}$$

Therefore we get $\mathrm{Var}\left(R_t - R_{t-\Delta t}\right) = \sigma^2\Delta t$ and hence

$$\Sigma := \sigma^2 = \frac{1}{\Delta t}\mathrm{Var}\left(R_t - R_{t-\Delta t}\right).$$

On the other hand we get from $S_t = \exp\left(\mu t - \frac{1}{2}\sigma^2 t + \sigma W_t\right)$ and (6.1) that

$$\begin{aligned}\log\frac{S_t}{S_{t-\Delta t}} &= \mu\Delta t - \frac{1}{2}\sigma^2\Delta t + \sigma(W_t - W_{t-\Delta t})\\ &= (R_t - R_{t-\Delta t}) - \frac{1}{2}\sigma^2\Delta t \end{aligned}\tag{6.2}$$

and therefore we observe $\mathrm{Var}\left(R_t - R_{t-\Delta t}\right) = \mathrm{Var}\left(\log\frac{S_t}{S_{t-\Delta t}}\right)$ and hence

$$\Sigma = \frac{1}{\Delta t}\mathrm{Var}\left(\log\frac{S_t}{S_{t-\Delta t}}\right)$$

□

6.1.1 Analysts and Filtering

The only external source of information other than the stock prices are the so-called analysts or experts who provide estimates μ_0 for the future growth rate $\mu \sim \mathcal{N}(\mu_0, \Sigma_0)$. We assume these estimates to be uncorrelated with variance $\Sigma_0 = \mathrm{diag}(0.2, ..., 0.2)$.

Obviously it is very hard to simulate results whose main ingredient should be non-observable expertise. Therefore we consider four different types of analysts whose estimates cover a wide range of possible opinions. Any real-life analyst should be somewhere in between these theoretical analysts:

- The optimal analyst: *'opt'*
  This analysts looks one year into the future to determine his estimate μ_0 and hence simulates an analyst with perfect expertise. The resulting strategies should provide reasonable upper bounds to any real-life analyst.

- The bad analyst: *'bad'*
  This analysts just reverses the estimate of the optimal analyst and hence simulates an analyst whose estimates are always wrong. The resulting strategies might even lead to bankruptcy.

- The random analyst: *'rand'*
  This analysts uses the average growth rate of the whole market to any stock in his portfolio and hence simulates an analyst without any knowledge or opinion. He is used as a benchmark-analyst that is better fitting to our market model than the usual $1/d$-strategies. The resulting strategies should provide reasonable lower bounds to any real-life analyst.

- The historical analyst: *'hist'*
  This analysts looks one year into the past to determine his estimate μ_0 and hence simulates an analyst who doesn't have a personal opinion and can't detect market changes before they happen. Since the real-life analysts have at least these (historical) information, the resulting strategies should again provide reasonable lower bounds to any real-life analyst.

Additionally to these deterministic analysts we consider several stochastic analysts in each of the above cases, where 'stochastic' means that the analyst estimates his μ_0 not deterministically, but $\mathcal{N}(\mu_0, 0.2)$-distributed around his original deterministic value. This way we may increase the sample size significantly and reduce the influence of some accidental irregularities in the data. This is important in particular for the evaluation of the running time of the algorithm.

Our portfolio optimization model is set up with analysts that provide at $t = 0$ their estimate μ_0 of the future growth rate μ of the stocks. The filter $\hat{\mu}_t$ for this growth rate is initialized with μ_0 and gets continuously updated using the observations from the stock prices.

We want in particular to be able to observe the effect that originates from these estimates. Therefore we allow the analysts in our simulation study to provide their estimates more often than just at the beginning. Whenever a new estimates arises we will just reset the time to $t = 0$ and restart the whole investment procedure.

We consider the following three cases:

- New estimates every year: *'12 months'*
  This shall be the standard case, where new estimates are provided once every year and in between we use filtering to determine the current $\hat{\mu}_t$.

- New estimates every month: *'1 months'*
  In this setting the analysts provide new estimates every time we may trade, such that we won't use filtering at all.

- New estimates only once in the beginning: *'192 months'*
  This is the original setting of our model. The analysts' estimate is only given once in the beginning and afterwards we need to rely entirely on filtering.

Note that the terminal time T of the portfolio optimization has to be at least as long as the period between two points of estimation. We use $T = 1$ [year] for the first two cases and $T = 16$ [years] for the last case.

In the following sections we will simulate and evaluate the optimal strategies and performances of these analysts. For most graphical results we will use the 'hist12-analyst' (i.e. the historical analyst, that estimates every 12 months), because this is the most reasonable analyst providing exactly the lower bound that every real-life analyst should aim to beat.

Now the filter for the growth rate gets calculated using the result from (2.6):

$$\hat{\mu}_{t_k} = \hat{\gamma}_t \left(\Sigma_0^{-1} \mu_0 + \hat{\Sigma}^{-1} R_{t_k} \right) \qquad \text{where} \quad \hat{\gamma}_t = \left(\Sigma_0^{-1} + t_k \hat{\Sigma}^{-1} \right)^{-1}$$

Here the time t_k is set to zero whenever the analysts provide new estimates. The only yet unspecified parameter R_{t_k} are the (observable log-)returns of the stock prices since the last estimate, i.e. the sum of the previous monthly returns as given in (6.2):

$$R_{t_k} = R_{t_k} - R_{t_0} = \sum_{j=1}^{k} \left(R_{t_j} - R_{t_{j-1}} \right)$$

$$\text{where} \quad \left(R_{t_j}^{(i)} - R_{t_{j-1}}^{(i)} \right) = \log \frac{S_{t_j}^{(i)}}{S_{t_{j-1}}^{(i)}} + \frac{1}{2} \Delta t \cdot \sigma_i^2$$

6.1.2 Deriving the optimal portfolio strategies

After initializing the market we can calculate the optimal portfolio strategies.
Therefore we need to fix the investor's utility function via the risk-aversion level α, where $\alpha = 0$ corresponds to logarithmic utility and $\alpha < 0$ to negative power utility. Note that any result under logarithmic utility appearing in this work is the continuous limit of the same results under power utility for $\alpha \searrow 0$.

As convex constraints we consider throughout this simulation no-short-selling and no-borrowing.

The formulas for the Merton Plug-In strategy and the optimal unconstrained strategy (using stochastic filtering) are straightforward given by:

$$\pi_t^{\mathrm{M}} = \frac{1}{1-\alpha} \hat{\Sigma}^{-1} \hat{\mu}_t$$

$$\pi_t^{\mathrm{unc}} = \frac{1}{1-\alpha} \hat{\Sigma}^{-1} \left(\hat{\gamma}_T^{-1} - \frac{1}{1-\alpha} (T-t) \hat{\Sigma}^{-1} \right)^{-1} \hat{\gamma}_t^{-1} \hat{\mu}_t$$

However, Algorithm 4.34 to determine the optimal constrained strategy usually requires computations for every subset I of the set of dimensions $\{1, ..., d\}$ until it hits the correct subset of active dimensions I. Unfortunately there are 2^d possibilities for the correct active dimensions. If each attempt under d=30 stocks only needs 1ms (which is a realistic value, given there are several high-dimensional computations to conduct) we still need on average 150 hours to find the correct choice. This is far to much computation time for any application. Fortunately we can avoid having to check all 2^d possibilities by taking good guesses. For instance, there is a high chance that a dimension (a stock) i with a large expected growth rate $\hat{\mu}_t^{(i)}$ will enter the optimal portfolio strategy, while a small (or even negative) growth rate indicates a high chance that this stock won't be invested in.

Following this approach we can reorder the set of subsets of $\{1, ..., d\}$ such that we first check those combinations most likely to succeed. Of course, finding the optimal reordering would be too time-consuming, too. It is hence important to find a good balance between adding computational effort and gaining speed improvements by the chosen reordering procedure.

In our simulation we just reorder the set of dimensions with respect to their expected growth rates $\hat{\mu}_t^{(i)}$ and then start by checking combinations of those stocks with highest expected growth rates. Our results show that this approach reduces the average computation time for lots of cases to less than 1 second, which corresponds to only 2^{10} attempts. Note that there still occur rare combinations of growth rates and covariances leading to computation times of more than 1000 seconds (corresponding to 2^{20} attempts out of 2^{30}). Also note that the computation time obviously rises with an increasing diversification within the optimal strategy.

The figures on the following pages show results for the historical 12-months analyst. The other types of analysts presented above show similar behaviour.

The table in Figure 6.2 shows the average computation time for the historical 12-months analyst for various combinations of risk-aversions α and the number of stocks d in our market. It is very obvious that the computation time increases approximately exponentially in the number of stocks if the reordering corresponding to the respective setting stays unchanged.

d	$\alpha = 0$	$\alpha = -2$	$\alpha = -4$	$\alpha = -6$	$\alpha = -8$	$\alpha = -10$
5	0.0006	0.0013	0.0015	0.0016	0.0016	0.0015
10	0.0005	0.0015	0.0039	0.0058	0.0063	0.0063
15	0.0007	0.0044	0.0148	0.0422	0.0533	0.0545
20	0.0005	0.0072	0.0542	0.1752	0.3817	0.4721
24	0.0006	0.0161	0.0627	0.5227	1.5472	2.1064
26	0.0007	0.0155	0.0720	0.8738	3.0975	3.9837
28	0.0008	0.0154	0.1172	1.0687	10.9296	12.8380
30	0.0010	0.0177	0.1420	2.0968	20.7158	27.3027

Figure 6.2: average computation times for the hist12-analyst

On the other hand the computation time with respect to the risk-aversion does not increase arbitrarily. It stabilizes for a high risk-aversion $-\alpha$ as shown on the left hand side in Figure 6.3.

This effect happens because the choice of active dimensions entering the optimal strategy becomes constant for increasing $-\alpha$. On the right hand side in Figure 6.3 this is shown for the mean value of active stocks in the optimal strategies.

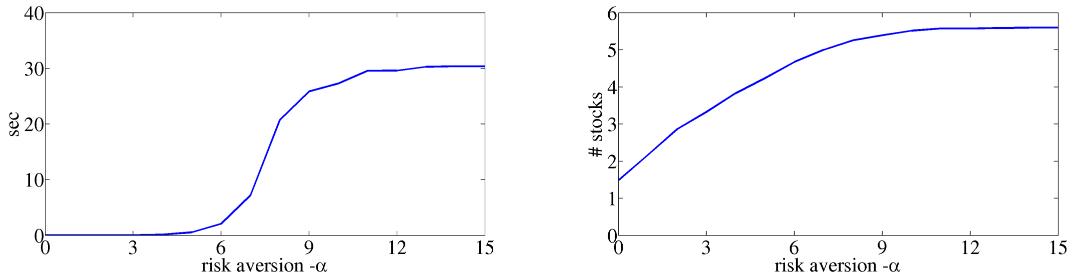


Figure 6.3: lhs: computation times for different risk aversions
rhs: mean value of active stocks in the optimal portfolio strategies

In the histogram in Figure 6.4, showing the distribution of active dimensions for various risk-aversions, the same effect is shown in more detail. For low risk-aversions ($-\alpha \in [0,3]$) most optimal strategies just invest in up to 3 different stocks, for medium risk-aversions ($-\alpha \in [4,8]$) the optimal strategies mainly invest in 3 to 6 different stocks, while for high risk-aversion ($-\alpha \geq 9$) most optimal strategies invest in at least 6 different stocks.

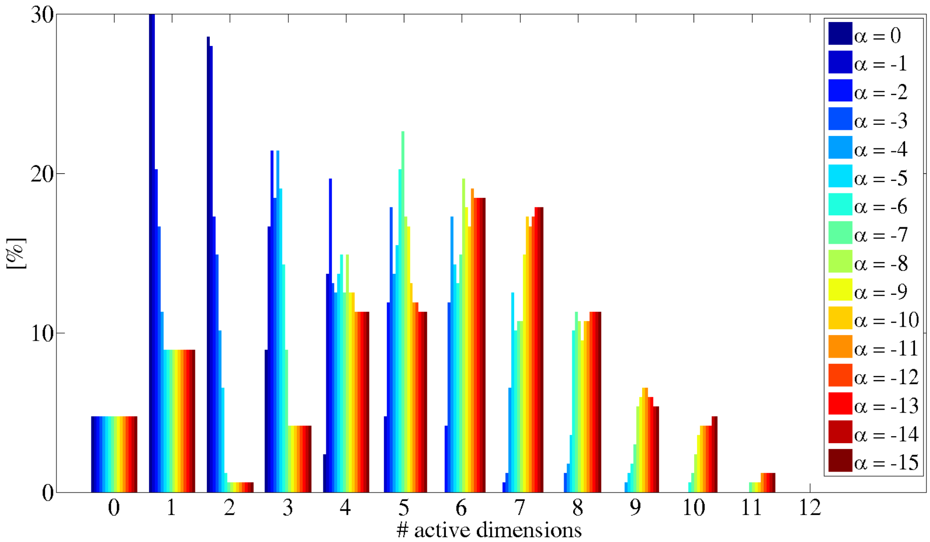


Figure 6.4: histograms of active stocks

Note that no strategy invests in more than 11 stocks. This is quite reasonable since we are in a setting with 30 stocks of whom some may have negative expected future growth rates while most are rather positively correlated.

Of course we cannot prove any effect for $1-\alpha \to \infty$ since the optimal unconstrained strategy then tends towards 0 and eventually enters the admissibility set anyway.

6.2 Evaluation of the optimal portfolio strategies

In the next step we want to evaluate the performance of an investor who would have invested according to the optimal constrained strategies in our market with 30 stocks from January 2001 until December 2014. First we compare the performance of constrained versus unconstrained strategies and then we examine the effect of transaction costs. Afterwards we will compare the several types of analysts.

In Figure 6.5 we see the standard case of a historical analyst for several risk-aversions between $\alpha = 0$ (logarithmic, very risky) and $\alpha = -15$ (very risk averse).

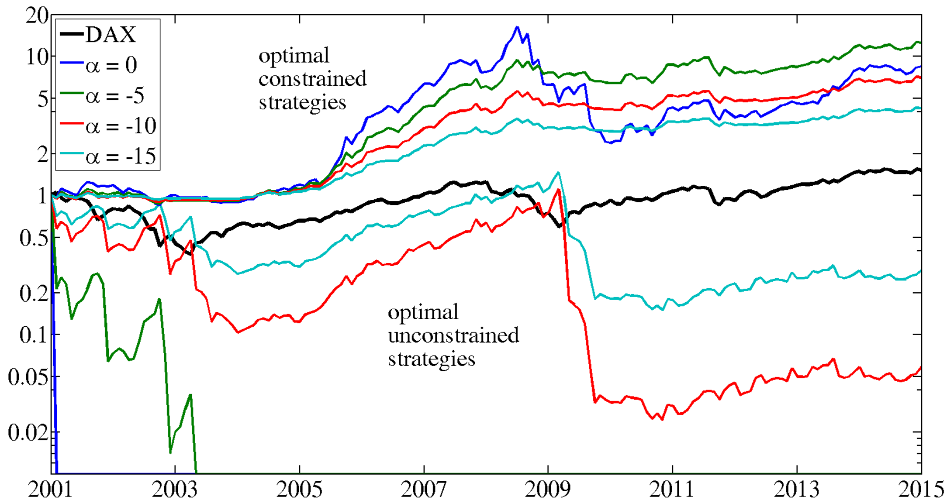


Figure 6.5: Wealth performance of the *hist-12*-analyst

We observe that the unconstrained strategies eventually lead to bankruptcy or at least to an immense loss of wealth no matter how risk-averse the investor may be. Note that in a theoretical continuous market bankruptcy is impossible by definition of the admissible strategies. However in our setting of trading only once a month bankruptcy is possible whenever the unconstrained strategy uses short-selling or money lending while the analyst provides a bad estimate. Although it is not shown in the above figure, the Merton-Plug-In-Strategies perform even worse then the unconstrained strategies.

The optimal constrained strategies on the other hand outperform both, the market and the unconstrained strategies by a lot. This holds true for any choice of risk-aversion. Figure 6.5 also demonstrates two effects arising from the various risk-aversions. For risky investors like $\alpha \in \{0, -1, -2\}$ we observe a rather volatile behaviour of the wealth process that doesn't necessarily lead to a higher utility of wealth in the end. The wealth process resulting from the logarithmic investor ($\alpha = 0$) for instance outperforms the other strategies until 2008, but then drops by more than 70%. This might be a highly undesirable behaviour that one can avoid by choosing a less risky utility function. On the other hand we observe that a rising risk-aversion of $\alpha \in \{-5, -6, ...\}$ doesn't really change the general behaviour of the wealth performance at all. The wealth process basically gets shrunk without a significant gain in terms of volatility.

However although the investor's choice of risk-aversion is essential for his resulting wealth performance, it is not too crucial since most of the optimal constrained strategies remain within some corridor that does not stretch too much.

Our theoretical market model does not yet account for transaction costs because this would complicate the calculations way further. Unfortunately many optimal strategy that get developed in a continuous-time setting without transactions costs get bankrupt really fast when applied in a setting with transaction costs, in particular when applied in discrete time as in our simulation.

Figure 6.6 shows with the blue line the optimal constrained and unconstrained wealth processes resulting from the same historical analyst as above under a risk-aversion of $\alpha = -10$. Note that the other risk-aversions don't change the statement of the figure significantly. The other colours illustrate the a posteriori effect of proportional transaction costs of $0.2\%, 1\%$ and 2% on the wealth performance of the optimal constrained and unconstrained strategies.

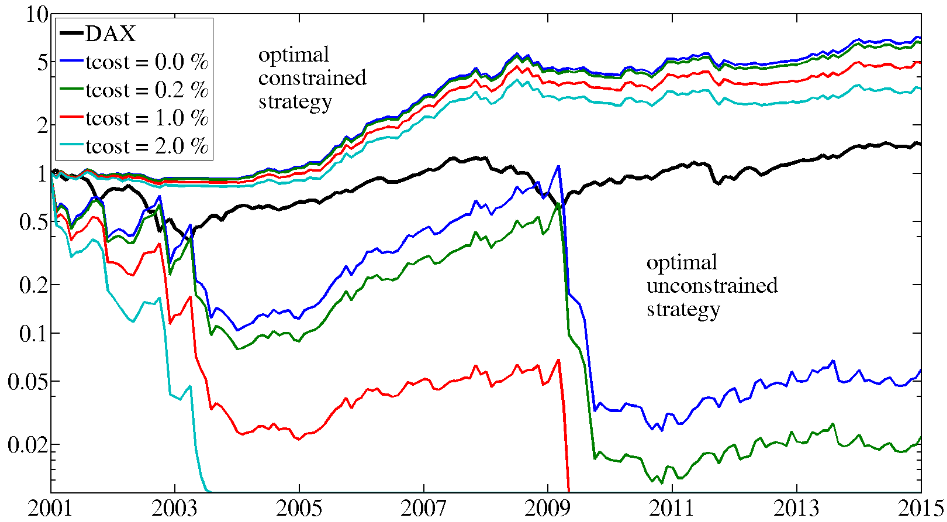
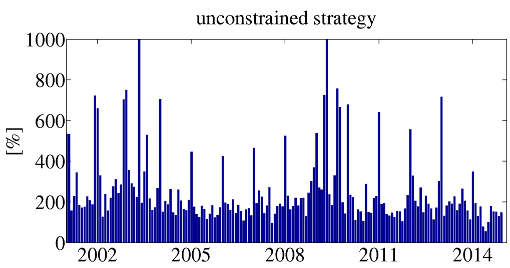


Figure 6.6: Wealth performance of the optimal strategies for $\alpha = -10$

Obviously we observe that higher transaction costs lead to higher losses of utility. However while the loss due to transaction costs in case of the constrained strategies is acceptable, the situation is very different for the optimal unconstrained strategies. Here transaction costs of 0.2% already lead to a loss of more than 60% of terminal wealth while higher transaction costs lead to bankruptcies before reaching the terminal time in 2015. We observe that transaction costs of 0.2% lead to a loss of 6.1% (p.a.) of wealth for the optimal unconstrained strategy while the respective constrained strategy loooses only 0.5% (p.a.) of its wealth. This difference gets even more impressive when considering transaction costs of 1%. While the optimal constrained strategy looses 2.6% (p.a.) of its wealth, the respective unconstrained strategy looses 27.1% (p.a.) of its wealth.

These numbers point out that the optimal unconstrained strategies are far from being applicable in the real world. Their high amount of transaction fees comes from their possibly extreme positions that any unconstrained strategy may invest in.
As we see in Figure 6.7 for the historical analyst with $\alpha = -10$ as given above, the unconstrained strategy needs to shift approximately 260% of its wealth on average each month. This means the investor needs to sell and rebuy around 130% of her wealth each month. The constrained strategy on the other hand only shifts around 22% each months, hence sells and buys around 11% each month. It is very obvious that the latter is a far more realistic behaviour. Note in particular the different scaling of the plots in Figure 6.7.



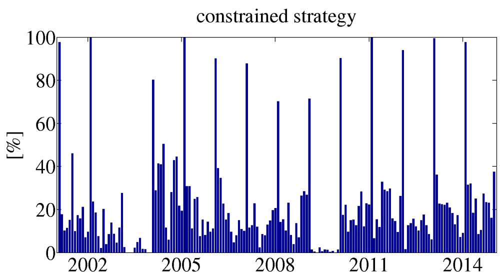


Figure 6.7: Amount of wealth to be shifted

In Figure 6.7 we also observe another interesting effect that arises from using an analyst that provides his new estimate every 12 months. During the months where the analyst is inactive, the investor updates her filter $\hat{\mu}_t$ just by filtering (hence by observing the market). These updates do not change her filter significantly, so the average amount of wealth to be shifted stays rather low. However in the particular months, where the analyst provides a new estimate, the filter might change a lot, possibly leading to a significant different optimal strategy. This results in a considerably elevated amount of wealth to be shifted, that is visible via the peaks on most Januarys in Figure 6.7. The average amount of wealth to be shifted in January for the optimal constrained strategy is 91% (versus 22% on average for the whole period) and the average amount of wealth to be shifted in January for the optimal unconstrained strategy is even 543% (versus 260% for the whole period).

Last but not least we want to consider the different results that the various analysts might produce.

In Figure 6.8 we compare the wealth performances of the optimal constrained strategies using those analysts that provide their estimates every 12 months. It is very obvious that the *optimal* analyst outperforms any other strategy by achieving immense gains while the *bad* analyst looses most his wealth. These analysts have been introduced as sort of an upper and lower bound to what might happen with any analyst's estimates, since no real analyst will be always right or always wrong. The *historical* analyst that is examined in detail above should be the realistic lower benchmark to any wealth performance, since any real-life analyst has at least the historical information.

An interesting observation is the poor performance of the *random* analyst. The random analyst should be the analogue in our setting to the usual $1/d$-strategy as explained above in Section 6.1.1. However while the $1/d$-strategy achieves approximately 7.5% (p.a.) during our investment period, the optimal strategy using the random analyst doesn't even outperform the market that achieves around 2.6% (p.a., the black line).

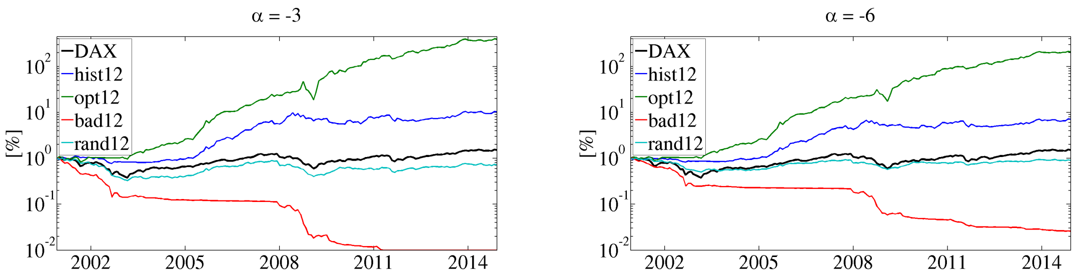


Figure 6.8: Wealth performance of various *12-months*-analysts

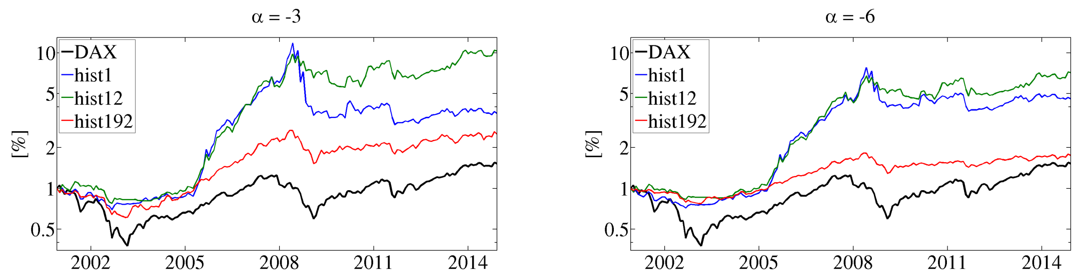


Figure 6.9: Wealth performance of various *historical* analysts

In Figure 6.9 we compare the wealth performances of the optimal constrained strategies using the different historical analysts.

The *12-months* historical analyst that is examined in detail above seems to produce the best wealth performance. This is no surprise since he combines the analyst's expertise with filtering the market observations. This way he is able to identify market changes faster than without the filtering. The *1-month* historical analyst that provides new estimates every month cannot use any filtering and hence the resulting strategy is way more prone to market changes. The resulting wealth process is way more volatile, in particular in a falling market. The *192-months* historical analyst that only provides one estimate in the beginning leads to a rather poor performance of the resulting optimal strategy when compared to the other two analysts. This is due to the fact that long-term filtering can only lead to useful results if its assumptions are satisfied. However one of the most crucial assumptions - for the drift $\mu_t \equiv \mu$ to be constant - is barely realistic for a long horizon. Therefore the filter $\hat{\mu}_t$ does not converge to a constant μ but to the long term average $\bar{\mu}_T = \frac{1}{T}\int_0^T \mu_t dt$.

Summing up the results for the various types of analysts it seems still reasonable to consider the historical 12-months analyst to be a very applicable benchmark to compare with any real-life analyst.

6.3 Path-wise simulation

Regarding the outstanding results of the historical simulation described in the previous sections the natural question arising is whether this might be coincidental due to luckily suitable data. Therefore in this section we repeat the whole simulation using random data to observe that the previous outstanding results on average still hold true.

In order to be able to conduct lots of simulations, we construct a market with only 10 stocks. This way the algorithm to figure out the correct case and active dimensions needs at most $2^{10} = 1024$ attempts and hence at most one second per step. In fact this speeds up the simulation a lot, such that each step needs less than 0.1 second on average.

We use the following values for the market's drift and volatility (each given in %) such that the simulated market is close to some possible real world market:

$$\mu^\top = \begin{pmatrix} 15 & 12 & 10 & 8 & 6 & 4 & 2 & 0 & -3 & -8 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 5.44 & 3.36 & 3.37 & 0.78 & 2.61 & 2.27 & 1.23 & 2.38 & -0.41 & 0.32 \\ 3.36 & 8.77 & 1.83 & 5.50 & 2.02 & 1.28 & 2.50 & 3.06 & -0.12 & -0.58 \\ 3.37 & 1.83 & 5.20 & 3.39 & 2.84 & 2.31 & 0.93 & 2.28 & 1.74 & 2.91 \\ 0.78 & 5.50 & 3.39 & 8.11 & 2.18 & 1.30 & 3.09 & 2.32 & 2.92 & 2.52 \\ 2.61 & 2.02 & 2.84 & 2.18 & 2.62 & 2.05 & 1.41 & 1.94 & 1.90 & 1.72 \\ 2.27 & 1.28 & 2.31 & 1.30 & 2.05 & 5.81 & 2.42 & 1.42 & 4.95 & 2.09 \\ 1.23 & 2.50 & 0.93 & 3.09 & 1.41 & 2.42 & 3.23 & 1.07 & 2.12 & 1.00 \\ 2.38 & 3.06 & 2.28 & 2.32 & 1.94 & 1.42 & 1.07 & 4.85 & 0.90 & 4.39 \\ -0.41 & -0.12 & 1.74 & 2.92 & 1.90 & 4.95 & 2.12 & 0.90 & 9.05 & 3.49 \\ 0.32 & -0.58 & 2.91 & 2.52 & 1.72 & 2.09 & 1.00 & 4.39 & 3.49 & 7.11 \end{pmatrix}$$

We want to be able to compare this simulation to the historical simulation described above, therefore we use the same investment period of 14 years. The investor is allowed to trade once every month, hence 168 times from January 2001 until December 2014. In order to be able to observe the distribution of any resulting values, we generate 1000 simulations of this 10-dimensional market consisting of 168 months.

In the historical simulation in Section 6.1.1 we created 4 different analysts to cover various real-life cases. Without being able to look into the future we can only use the *historical* and the *random* analyst. However both should provide reasonable lower bounds to any result from real-life analysts, since they should have at least this information. The analysts provide new estimates once every year and we set the rolling terminal time to $T = 1$ [year].

After preparing this market setting we compute as described in Section 6.1.2

- the optimal constrained and unconstrained strategies
- for the historical and the random analyst,
- for various choices of the risk-aversion α from 0 to -20

- for each of the 1000 simulations.

Additionally we implement the standard $1/d$-strategy as a benchmark that just reallocates $1/d$ of its wealth to each of the d stocks in each of the simulations.

At this point we are able to evaluate the performance of all the optimal strategies in our simulated markets. Therefore we compute for each optimal strategy the monthly portfolio allocations and the resulting wealth process. We also allow for proportional transaction costs that are calculated each month as a fraction of the wealth that has to be shifted between different stocks. If not stated otherwise we always imply 0.2% proportional transaction costs.

A first observation is as clearly as obvious: all of the unconstrained strategies go bankrupt and most of them within the first few months. Those strategies buy and (short) sell multiples of their wealth into single stocks, leading to immediate bankruptcies the moment some of these stocks move slightly too much into the wrong direction. With transaction costs this process obviously runs even faster. Increasing the risk-aversion of course leads to more unconstrained strategies not going bankrupt, but this needs really unrealistically high risk-aversions like $\alpha = -50$.

In Figure 6.10 we see several realisations of the resulting wealth process of the optimal constrained strategy resulting from the historical analyst. Additionally we see the monthly medians and 5% and 95%-quantiles of the wealth processes.

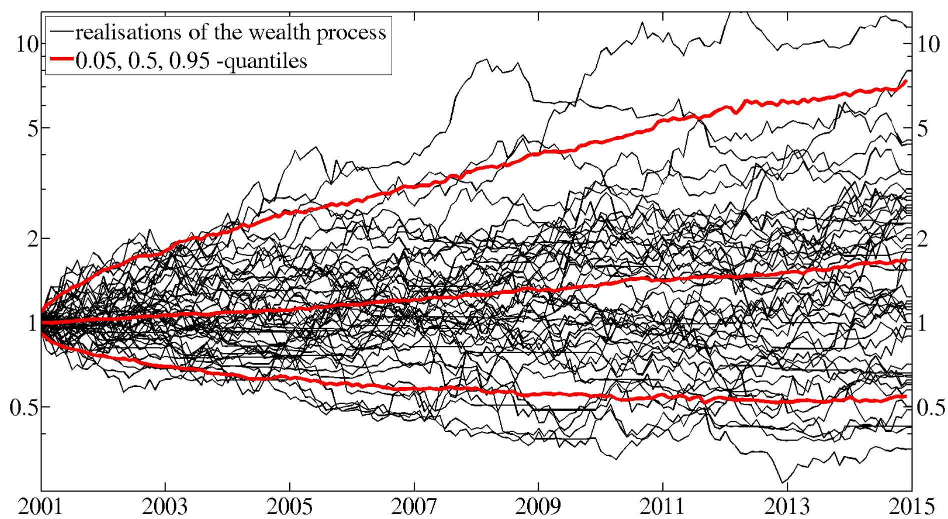


Figure 6.10: Wealth processes of the historical analyst for $\alpha = -3$

With Figure 6.11 we can compare these results of the historical analyst to the wealth processes of the random analyst as well as to the wealth processes of the $1/d$-strategy.

We observe that the random analyst performs less volatile but also way worse than the historical analyst. The benchmark-$1/d$-strategy on the other hand also performs less volatile but just slightly worse than the historical analyst. Since both strategies are only little volatile they reduce the downturn risk significantly, however at the same time they reduce the chance of high gains as significantly.

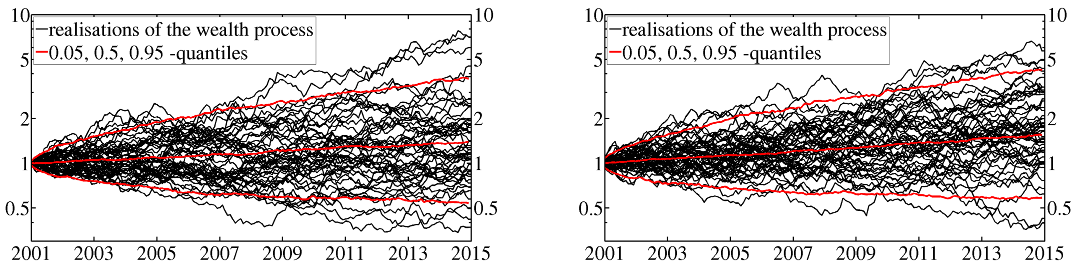


Figure 6.11: The random analyst for $\alpha = -3$ (left) and the $1/d$-strategy (right)

Of course the volatility of the wealth process can be controlled via the risk-aversion α of the investor. In Figure 6.12 we compare the various realisations of the wealth processes for the historical analyst for the risky logarithmic utility ($\alpha = 0$) to the more risk-averse power utility with $\alpha = -10$. Obviously we will always find risk-aversions that lead to both, riskier and more risk-averse strategies than any benchmark.

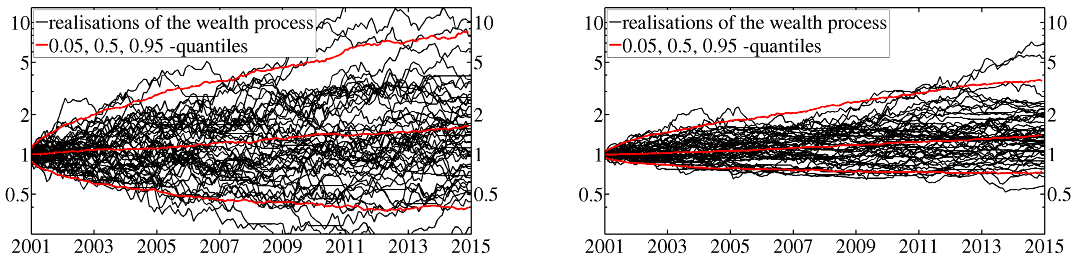


Figure 6.12: The historical analyst for $\alpha = 0$ (left) and for $\alpha = -10$ (right)

In all the above figures we observe behaviours of the realisations of the wealth processes that look very much like the behaviour of paths of a Brownian motion. Additionally the distributions of the respective wealth processes are quite skew. This can be seen by comparing their means and medians or the distances of their symmetric quantiles to their medians. In Figure 6.10 (for $\alpha = -3$) the mean of the terminal wealth is 2.43 while the median is only 1.68.

However if we compare the logarithm of the terminal wealth instead, this distance vanishes mostly. In Figure 6.10, the mean (of the log terminal wealth) is 0.55 and the median is 0.52. Also the given quantiles have a similar distance to the median. This is highly suggesting a log-normal distribution for the terminal wealth. The wealth process hence might look like $X_t = \exp\left(\int_0^t a_t dt + \int_0^t b_t dW_t\right)$ for some functions a_t and b_t.

In Figure 6.13 we see histograms of the logarithmic terminal wealth for the various strategies. In order to be able to better interpret these numbers they are scaled down to one year, hence a value of 0.05 corresponds to a gain of $e^{0.05} - 1 = 5.13\%$ p.a.. Additionally we see the curve of the density of a normal distribution with mean and standard deviation of the corresponding data to see the fit of a log normal distribution.

We compare the distribution of the logarithmic terminal wealth of the benchmark $1/d$-strategy to the distributions of the logarithmic terminal wealth of the historical analyst for various risk-aversions.

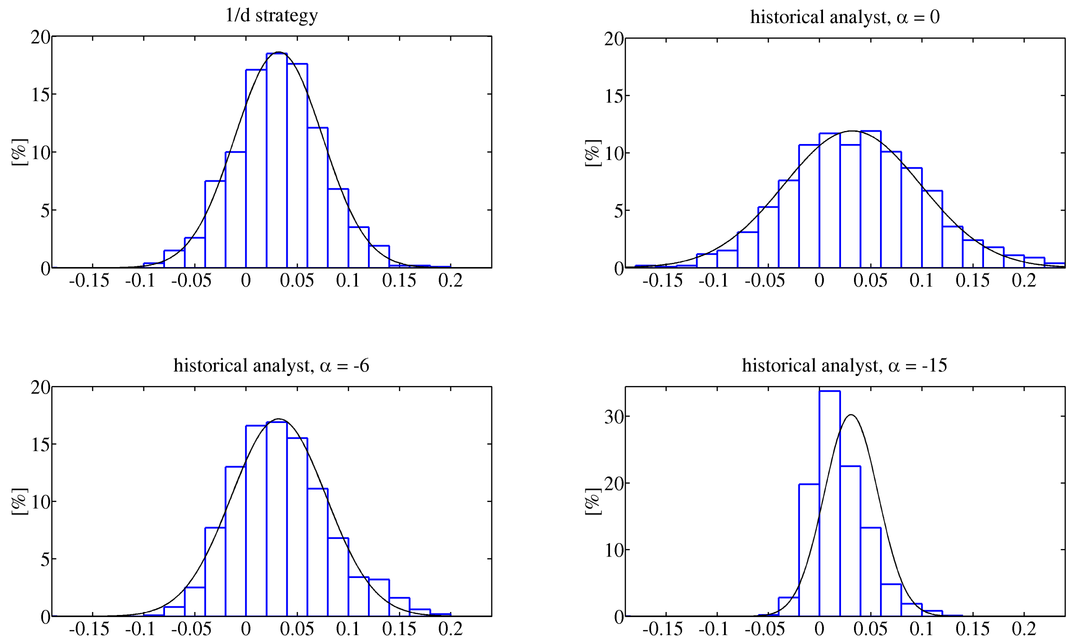


Figure 6.13: Histograms of the logarithmic terminal wealth p.a.

We observe that the distribution of the terminal wealth of the $1/d$-strategy is very close to a log normal distribution, as it is the case for the historical analyst for low risk-aversion. For higher risk-aversions we observe the distribution getting positively very skew, essentially reducing the shortfall risk while still allowing for a large gain with a very small probability.

For the historical analyst with $\alpha = -6$ we observe a quite similar mean than for the $1/d$-strategy while the historical analyst allows for way higher gains although reducing the risk of larger losses.

This can also be seen in Figure 6.14 showing the medians and quantiles of both, the $1/d$-strategy and the historical analyst for various risk-aversions. We observe that the median for the historical analyst is always quite close ($\pm 1\%$) to the median of the $1/d$-strategy while the quantiles vary a lot. For low risk-aversions we observe a slightly higher median for the logarithmic terminal wealth and quantiles of the historical analyst that are far away from the median spreading its distribution a lot to possibly high gains and losses.

However for medium-high risk-aversions like $\alpha \in [-4, -7]$ we observe similar medians and 5%-quantiles for the benchmark and the historical strategy, while the 95%-quantiles of the historical strategy is way higher then the 95%-quantile of the $1/d$-strategy. This could be exactly the results wished for in practical applications: reduce the shortfall-risk while maintaining the chance for large gains.

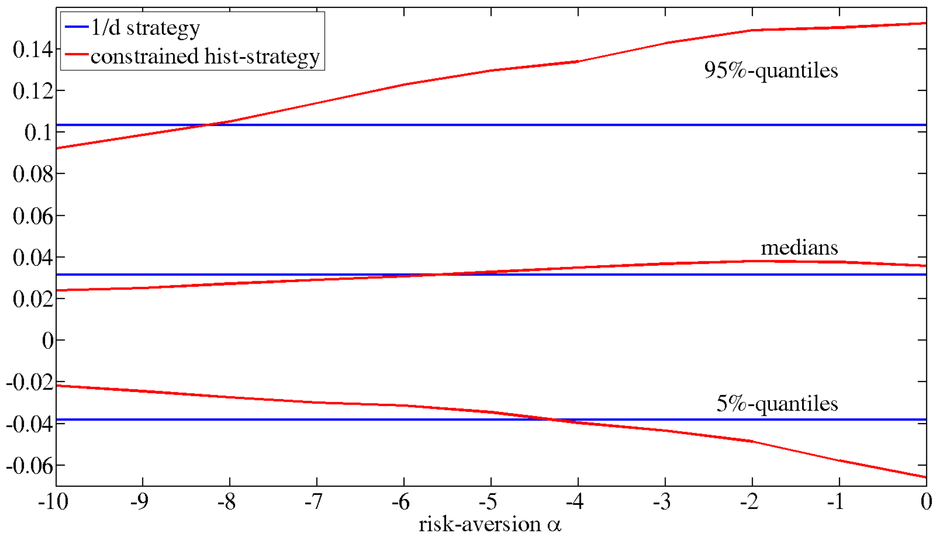


Figure 6.14: Quantiles and medians of the logarithmic terminal wealth p.a.

Finally we want to compare directly the terminal wealth of the respective strategies within each simulation. Within each simulated market we calculate the wealth processes of all three strategies (constrained historical, constrained random, $1/d$-strategy) and compare their terminal wealth.

In Figure 6.15 we see the percentage of simulations with one strategy outperforming another one, i.e. resulting in a higher terminal wealth than another strategy. For starters there seems to be no strategy strongly dominating the others. The maximal value in Figure 6.15 can be found for $\alpha = -10$ and t-cost $= 0.2\%$ and it says that 73% of these $1/d$-strategies outperformed the random analyst. However 73% is far from being able to consider one strategy significantly better. After all we need to notice that these are only probabilities whether some strategy performs better or worse on average.

Nevertheless we do observe some interesting tendencies. Considering the red line we observe that no matter the risk-aversion and no matter the transaction costs, the random analyst can only beat the historical analyst in about 30% to 40% of cases. Only in the case without transaction costs and with very low risk-aversion the random analyst manages to

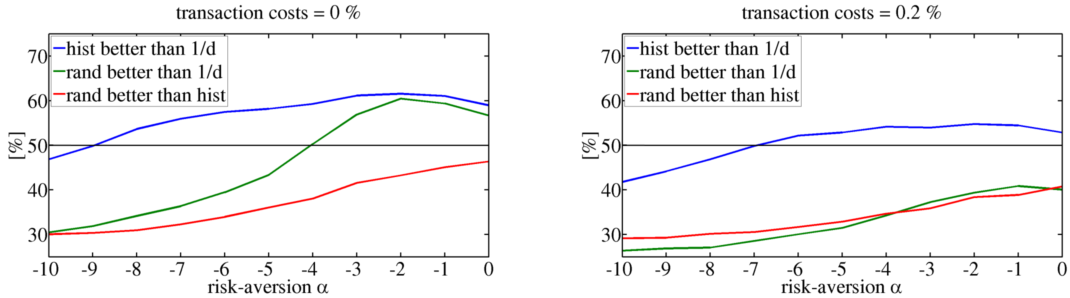


Figure 6.15: Percentage of better running strategies

dominate the benchmark-$1/d$-strategy in more than 50% of the cases (the green line). This is not very surprising since the random analyst does not do any analysing work at all.

However although we have seen that the historical analyst outperforms the $1/d$-strategy significantly on average (with respect to the mean) we observe that he only beats the benchmark in about 40% to 60% of the simulations (the blue line). Also for higher risk-aversions this advantages vanishes.

Last but not least note the main impact that comes from transaction costs. When comparing both plots in Figure 6.15 we observe a major shift of most data points towards those strategies that trade less. Hence the impact of transaction costs on highly active strategies is still high. Also note that this must of course be the case, since transaction costs are not part of the model solved in the previous chapters, hence the resulting optimal strategies need not be optimal in a setting with transaction costs any more.

7 Outlook on further research

In the previous chapters we defined and solved several portfolio optimization problems explicitly. First we introduced our notion of partial information and convex constraints and how to deal with it using standard filtering techniques and the convex dual methods of [CK92]. Then we considered the unconstrained and constrained portfolio optimization problems under logarithmic and power utility with a no short-selling and a no borrowing constraint. In order to solve the constrained problem under power utility we considered a reverse separation approach leading to several cases for the optimal solution that we treated separately. The main effort to solving this problem lies in finding the structure of the optimal dual process. Afterwards we transferred our approach to further convex constraints whose corresponding problems can also be solved quite explicitly and we provided some general results for arbitrary utility functions. Finally we simulated the performance of our optimal portfolio strategies under realistic conditions.

One can think of several extensions of our results. Therefore in the following we will present some further ideas and generalizations that can and should be considered in future research. The most challenging part however might be to keep the explicit solvability of the resulting solutions.

Remark 7.1 (Transaction costs and non-continuous trading)
The most unrealistic part of our market setting is the absence of transaction costs. Our resulting portfolio strategies require time-continuous trading. Therefore they are not feasible in practice since they lead to immediate bankruptcy when introducing transaction costs a posteriori.

In our simulation we have already seen in Figure 6.7 that the amount of transactions costs gets reduced significantly when applying constrained strategies due to less trading in general. An interesting investigation could therefore be to figure out lower bounds to what extent we can reduce transaction costs with different strategies.

There are various approaches in the literature that deal with non-continuous trading like the 'relaxed h-investor' as it is described in [Rog01] or in [BUV12] who use the same setting of partial information than we do. However limiting the investor to only be allowed to trade at fixed points in time seems undesirable when the investor is able to observe the market continuously allowing him to trade as needed.

Anyway the problem of continuous trading vanishes automatically when introducing transaction costs. As soon as they contain a constant part, any optimal portfolio strategy cannot trade continuously without getting bankrupt. Hence introducing transaction costs does not only lead to a way more realistic market model, but it also leads to time-discrete trading.

The challenging part of transaction costs however remains. Finding analytical expressions for the optimal portfolio strategy might be arbitrarily hard, as results for the unconstrained problems under transaction costs already indicate.

Remark 7.2 (Infinite horizon)
When implementing our market model and the corresponding algorithm there occurs another problem. We assume a terminal time T to optimize the utility of terminal wealth. However most actual funds do not have a terminal time, they continue to exist as long as they have good performance or their clients remain satisfied.

Therefore it would be interesting to consider also a portfolio optimization problem in the same market model but with infinite time horizon. This could be done by introducing a constant consumption process whose discounted future payouts add up to the argument that gets optimized. Typically this kind of problem leads to the same optimal portfolio strategy as the maximization of terminal wealth, at least for logarithmic and power utility.

One advantage of this approach might be that the resulting HJB-equation could be independent of t and therefore the problem would be easier to solve. The only remaining time-dependency would be the values of the filter and its variance themselves, who only enter the strategy as given constants. On the other hand this approach might not be realistic if the investor's incentive was to rather optimize her current 1-year-performance than to include her future performance, since she might have to please her client or employer in the present time.

Remark 7.3 (recurring estimates of the analysts)
Our portfolio optimization model is set up with analysts that provide their estimate of the future growth rate of the stocks only at $t = 0$. Afterwards the filter for this growth rate gets only updated using the following observations from the stock prices.

However in practice we would like to be able to get entirely new estimates of these analysts whenever the market might change. In the simulation study in Chapter 6 we therefore restarted the whole investment procedure whenever we wanted to introduce new analysts' estimates.

After all, recurring expert opinions can be a totally different problem, like in [SWW17].

Remark 7.4 (estimates of single stocks at different points of time)
Another problem with recurring analysts' estimates than in Remark 7.3 could be incoming estimates of single stocks. We assume in our market setting that all the analysts provide their estimates of the different stocks simultaneously. This is important to our market setting since the filter $\hat{\mu}_t$ of the growth rates starts with this estimate and evolves simultaneously in all stocks. In particular the conditional variance γ_t of the filter depends on the correlation between these simultaneous estimates and the time that has passed since the last estimate.

To include the possibility for estimates of different stocks at different points of time one needs to set up a whole new problem. However we want to present two ad hoc ideas of how to deal with this problem when given our setting.

1) In the beginning we start with the parameters (μ_0, Σ_0). Now assume there occurs a new estimate μ_i to stock i at time $t_1 > 0$. Then the idea of this first approach is to just treat the current filter values $\hat{\mu}_j$ for $j \neq i$ as if they were new estimates to the other stocks, too. This way our whole theory stays intact and applicable. However treating the conditional variance γ_t of the filter for $t > t_1$ might be challenging. Consider the following three possibilities to deal with the conditional variance:

- At time t_1 γ_t stays unchanged, hence $\gamma_t = \left(\Sigma_0^{-1} + t\Sigma^{-1}\right)^{-1}$ where t is the time that past since the very beginning.
  The problem is that then $\gamma_t \underset{t\to\infty}{\longrightarrow} 0$ indicating that the observed filter value is the real growth rate μ while in reality we do not get closer over time.

- At time t_1 γ_t gets set back to its original value, hence $\gamma_t = \left(\Sigma_0^{-1} + (t - t_1)\Sigma^{-1}\right)^{-1}$ where $t - t_1$ is the time since the last new estimate.
  The problem with this approach would be that we go back to $\gamma_{t_1} = \Sigma_0$ every time a new estimate occurs although the other d-1 estimates have a longer history.

- At time t_1 γ_t gets updated using the previous information $\gamma_t = \left(\gamma_{t_1-}^{-1} + (t - t_1)\Sigma^{-1}\right)^{-1}$ where γ_{t_1-} is the value of the conditional variance just before the new estimate occurred. This approach is the most consistent with the above idea of updating the filter value.
  However the problem with this approach is that it is the same as the first approach: $\gamma_t = \left(\gamma_{t_1-}^{-1} + (t - t_1)\Sigma^{-1}\right)^{-1} = \left(\Sigma_0^{-1} + t_1\Sigma^{-1} + (t - t_1)\Sigma^{-1}\right)^{-1} = \left(\Sigma_0^{-1} + t\Sigma^{-1}\right)^{-1}$

After all the second approach of resetting the conditional variance every time a new estimate occurs seems to be reasonable. Alternatively one might think of resetting the time more slowly such that $\gamma_t = \left(\Sigma_0^{-1} + (t - t^*)\Sigma^{-1}\right)^{-1}$ where t^* is some mean of the previous times when new estimates occurred.

2) Another idea could be to use a direct approach that does not need computing the whole successive path of the filter values $\hat{\mu}_t$. To compute the filter $\hat{\mu}_t = \gamma_t \left(\Sigma_0^{-1}\mu_0 + \Sigma^{-1}R_t\right)$ we only need the conditional variance γ_t and the current observable returns R_t.

Therefore at time t assume that the analysts provided their estimate for μ_i at time $t - t_i$, hence t_i is the age of the estimate for stock i. Then the observable return for stock i given the age of the last estimate becomes

$$R_i(t_i) = \log \frac{S_i(t)}{S_i(t - t_i)} + \frac{1}{2}\Sigma_{ii}t_i$$

Hence we can define the return vector $R_\tau := \left(R_1(t_1), ..., R_d(t_d)\right)$, where $\tau = \operatorname{diag}(t_1, ..., t_d)$. Also we need to modify the conditional variance $\gamma_t = \left(\Sigma_0^{-1} + t\Sigma^{-1}\right)^{-1}$ to γ_τ using the approach

$$\gamma_\tau := \left(\Sigma_0^{-1} + \sigma^{-\top}\tau\sigma^{-1}\right)^{-1}$$

This way we observe a symmetric generalization of γ_t that accounts for the different estimation times via $(\sigma^{-\top}\tau\sigma^{-1})_{ij} = \sum_k t_k(\sigma^{-1})_{ki}(\sigma^{-1})_{kj}$. Note that neither $\Sigma\tau$ nor $\tau\Sigma$ would be symmetric or contain a reasonable interpretation.

The disadvantage of using this second approach is obvious. We redefined some parameters hence it becomes necessary to reprove every subsequent result like the derivation of the optimal strategies.

A Appendix

A.1 Proofs from Chapter 1

There are no proofs in Chapter 1.

A.2 Proofs from Chapter 2

A.2.1 Convex Duality of [CK92]

Cvitanic and Karatzas (1992) proofed dual optimality for utility functions that satisfy lots of assumptions, in particular

$$x \mapsto xU'(x) = x^{\alpha} \text{ should be nondecreasing} \qquad [\text{CK92}, (5.8)]$$

Unfortunately negative power utility doesn't satisfy this assumption, that is used to show that existence of an optimal constrained strategy implies existence of an optimal dual λ.

However in the case without consumption one can easily rewrite the whole proof by using negative power utility directly to obtain the exact same results. Additionally we will rewrite the whole proof below and allow for stochastic parameters and an arbitrary physical measure.
In the following any undefined notation is taken from [CK92]!

Theorem A.1 (Dual optimality)
The following conditions are equivalent in the sense that optimality of π_t^* (A) implies existence of λ_t satisfying (B) - (E), while conditions (B) - (E) are equivalent and existence of their λ_t implies existence of the optimal $\pi_t^* = \pi_t^{\lambda}$.

(A) Optimality of $\pi_t^* \in K$:

$$\begin{aligned} & \mathbb{E}\left[U(X_T^*)\right] \geq \mathbb{E}\left[U(X_T^{\pi})\right] \ \forall \pi \in K \\ \text{and} \quad & \mathbb{E}\left[X_T^* U'(X_T^*)\right] < \infty \end{aligned}$$

(B) Financibility of ξ^{λ} $(= X_T^{\lambda,*})$:

$$\exists \pi_t^{\lambda} \in K, \quad \delta(\lambda_t) + \lambda_t^{\top} \pi_t^{\lambda} = 0 \quad \text{and} \quad X_t^{\pi^{\lambda}} = X_t^{\lambda,*}$$

(C) Minimality of λ_t:

$$\mathbb{E}\left[U\big(\xi^{\lambda}\big)\right] \leq \mathbb{E}\left[U\big(\xi^{\nu}\big)\right] \forall \nu \in \mathcal{D}$$

(D) Dual optimality of λ_t:

$$\mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^{\lambda}(x_0)H_T^{\lambda}\big)\right] \leq \mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^{\lambda}(x_0)H_T^{\nu}\big)\right] \forall \nu_t \in \mathcal{D}$$

(E) Parsimony of λ_t:

$$\mathbb{E}\left[H_T^{\nu}\xi^{\lambda}\right] \leq x_0 \ \forall \nu \in \mathcal{D}$$

Notation:
X_t^{π} is the wealth process in the original market when investing π_t,
X_t^{*} is the optimal wealth process in the original market,
$X_t^{\nu,\pi}$ is the wealth process in the ν-market when investing π_t
$X_t^{\nu,*}$ is the optimal wealth process in the ν-market,
$\xi^{\nu} := I\big(\mathcal{Y}^{\nu}(x_0)H_T^{\nu}\big)$ is the optimal terminal wealth in the ν-market, $\mathcal{Y}^{\nu}(x) = (\mathcal{X}^{\nu})^{-1}(x)$ and $\mathcal{X}^{\nu}(y) = \mathbb{E}\left[H_T^{\nu}I\big(yH_T^{\nu}\big)\right]$.

Proof. (similar to [CK92, Thm.10.1])
(B) $\Rightarrow$ (A) is trivial.
(A) $\Rightarrow$ (B)
The wealth process X_t^* according to the optimal portfolio process π_t^* satisfies

$$H_t^0 X_t^* = x_0 + \int_0^t H_s^0 X_s^* \left(\sigma^{\top}\pi_s^* - \theta_s\right)^{\top} dW_s$$

We have $\mathbb{P}\left(X_t^* > 0, \forall t\right) = 1$. (proof: [CK92])
We want to show existence of $\lambda \in \mathcal{D}$ such that

$$H_t^{\lambda} X_t^* = \mathbb{E}\left[H_T^{\lambda}\xi^{\lambda}\middle|\,\mathcal{F}_t\right]$$
$$\text{and} \quad \delta(\lambda_t) + \lambda_t^{\top}\pi_t^* = 0$$

(because then $\pi_t^* = \pi_t^{\lambda}$ and $X_t^* = X_t^{\pi^{\lambda}} = X_t^{\lambda,*}$)
Therefore consider the continuous martingale (with martingale representation)

$$M_t := \mathbb{E}\left[X_T^* U'(X_T^*)\middle|\,\mathcal{F}_t\right] = \mathbb{E}\left[(X_T^*)^{\alpha}\middle|\,\mathcal{F}_t\right] = m_0 + \int_0^t \psi_s^{\top} dW_s$$

then the process $\quad \lambda_t := -(\mu - r\mathbb{1}) + \Sigma\pi_t^* - \sigma\frac{1}{M_t}\psi_t$

satisfies the conditions of Financibility (B).
This is shown in [CK92, Appendix A.1] where we want to stress out that the following conclusions have to be altered for negative power utility:

Proof. (the original version of [CK92])
Let $\hat{X}(t)e^{-3\varepsilon n} \leq X_{\varepsilon}(t) \leq \hat{X}(t)e^{3\varepsilon n}$ $(*)$. Then

$$\frac{1}{\varepsilon}\left|U\big(X_{\varepsilon}(T)\big) - U\big(\hat{X}(T)\big)\right| \overset{MVT}{\leq} \frac{1}{\varepsilon}\left|X_{\varepsilon}(T) - \hat{X}(T)\right| \cdot U'\left(\hat{X}(T)e^{-3\varepsilon n}\right)$$

$$\overset{(*)}{\leq} U'\left(\hat{X}(T)e^{-3\varepsilon n}\right)\hat{X}(T)e^{-3\varepsilon n}\frac{e^{3\varepsilon n}-1}{\varepsilon}e^{3\varepsilon n}$$
$$\overset{(5.8)}{\leq} K_n U'\big(\hat{X}(T)\big)\hat{X}(T)$$

where $K_n := e^{3n}\sup_{0<\varepsilon<1}\frac{e^{3\varepsilon n}-1}{\varepsilon}$ is finite. □

Proof. (the new version for negative power utility)
Let $\hat{X}(t)e^{-3\varepsilon n} \leq X_\varepsilon(t) \leq \hat{X}(t)e^{3\varepsilon n}$ $(*)$. Then

$$\begin{aligned}\frac{1}{\varepsilon}\Big|U\big(X_\varepsilon(T)\big)-U\big(\hat{X}(T)\big)\Big| &\overset{MVT}{\leq} \frac{1}{\varepsilon}\Big|X_\varepsilon(T)-\hat{X}(T)\Big|\cdot U'\left(\hat{X}(T)e^{-3\varepsilon n}\right)\\ &\overset{(*)}{\leq} U'\left(\hat{X}(T)e^{-3\varepsilon n}\right)\hat{X}(T)\frac{e^{3\varepsilon n}-1}{\varepsilon}\\ &= \hat{X}(T)^\alpha e^{-3\varepsilon n(\alpha-1)}\frac{e^{3\varepsilon n}-1}{\varepsilon}\\ &\leq \hat{X}(T)^\alpha\cdot K_n = K_n U'\big(\hat{X}(T)\big)\hat{X}(T)\end{aligned}$$

where $K_n := \sup_{0<\varepsilon<1}\left(e^{-3\varepsilon n(\alpha-1)}\frac{e^{3\varepsilon n}-1}{\varepsilon}\right)$ is finite as $x \mapsto e^{-3\varepsilon n(\alpha-1)}\frac{e^{3\varepsilon n}-1}{\varepsilon}$ is uniformly continuous on $[0,1]$. □

Now we can continue the proof of the above theorem:
(B) ⇒ (E)
We know from the dynamics of H_t^ν in the ν-market and X_t^π in the original market:

$$\begin{aligned} dH_t^\nu X_t^\pi &= H_t^\nu X_t^\pi\left(\left(-\delta(\nu_t)-\pi_t^\top\nu_t\right)dt+\left(\pi_t^\top\sigma-\theta_t^\top\right)dW_t\right)\\ \Rightarrow\quad & \mathbb{E}\left[H_T^\nu X_T^\pi\right]\leq x_0\end{aligned}$$

As $X_t^{\lambda,*} = X_t^{\pi^\lambda}$ (B) can be financed by π_t^λ in the original market, we get for $\xi^\lambda = X_T^{\lambda,*}$

$$\mathbb{E}\left[H_T^\nu\xi^\lambda\right]\leq x_0$$

(E) ⇒ (D)
By using $\tilde{U}\left(U'(x)\right)+x\left(U'(x)-y\right)\leq\tilde{U}(y)$ with $x=\xi^\lambda=I\big(\mathcal{Y}^\lambda(x_0)H_T^\lambda\big)$
and $y=\mathcal{Y}^\lambda(x_0)H_T^\nu$ we get

$$\begin{aligned}\tilde{U}\big(\mathcal{Y}^\lambda(x_0)H_T^\nu\big) &\geq \tilde{U}\big(\mathcal{Y}^\lambda(x_0)H_T^\lambda\big)+\mathcal{Y}^\lambda(x_0)\big(H_T^\lambda\xi^\lambda-H_T^\nu\xi^\lambda\big)\\ \Rightarrow\mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^\lambda(x_0)H_T^\nu\big)\right] &\geq \mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^\lambda(x_0)H_T^\lambda\big)\right]+\mathcal{Y}^\lambda(x_0)\big(x_0-\mathbb{E}\left[H_T^\nu\xi^\lambda\right]\big)\\ &\geq \mathbb{E}\left[\tilde{U}\big(\mathcal{Y}^\lambda(x_0)H_T^\lambda\big)\right]\end{aligned}$$

(B) ⇒ (C)
If we invest π_t^λ in a ν-market and in the original market, we get

$$dX_t^{\nu,\pi^\lambda} = X_t^{\nu,\pi^\lambda}\left(\left(r+\delta\left(\nu_t\right)+{\pi_t^\lambda}^\top\nu_t\right)dt+{\pi_t^\lambda}^\top\sigma dW_t^{\mathbb{Q}}\right)$$

$$dX_t^{\pi^\lambda} = X_t^{\pi^\lambda}\left(rdt + \pi_t^{\lambda^\top}\sigma dW_t^{\mathbb{Q}}\right)$$

and observe $X_t^{\nu,\pi^\lambda} \geq X_t^{\pi^\lambda} > 0$, hence $\xi^\nu = X_T^{\nu,*} \geq X_T^{\nu,\pi^\lambda} \geq X_T^{\pi^\lambda} = \xi^\lambda$, hence (C).

(C) $\Rightarrow$ (D)

We know from the dynamics of H_t^ν in the ν-market and X_t^π in the original market:

$$dH_t^\nu X_t^\pi = H_t^\nu X_t^\pi\left(\left(-\delta(\nu_t) - \pi_t^\top\nu_t\right)dt + \left(\pi_t^\top\sigma - \theta_t^\top\right)dW_t\right)$$

$$\Rightarrow \quad \mathbb{E}\left[H_T^\nu X_T^\pi\right] \leq x_0$$

Using $\tilde{U}(y) \geq U(x) - xy \ \forall x, y$ we get

$$\begin{aligned} U(X_T^\pi) &\leq \tilde{U}(yH_T^\nu) + yH_T^\nu X_T^\pi \\ \Rightarrow \quad \mathbb{E}\left[U(X_T^\pi)\right] &\leq \mathbb{E}\left[\tilde{U}(yH_T^\nu)\right] + y\mathbb{E}\left[H_T^\nu X_T^\pi\right] \\ &\leq \mathbb{E}\left[\tilde{U}(yH_T^\nu)\right] + yx_0 \end{aligned}$$

as this holds for any π_t admissible for the ν-market, it also holds for the optimum

$$\begin{aligned} \Rightarrow \quad & \mathbb{E}\left[U(\xi^\nu)\right] \leq \mathbb{E}\left[\tilde{U}(yH_T^\nu)\right] + yx_0 \ \forall y \\ \Rightarrow \quad & \mathbb{E}\left[U(\xi^\nu)\right] \leq \mathbb{E}\left[\tilde{U}(\mathcal{Y}^\lambda(x_0)H_T^\nu)\right] + x_0\mathcal{Y}^\lambda(x_0) \end{aligned}$$

However on the other hand by (C) and $\tilde{U}(y) = U(I(y)) - yI(y)$ we get

$$\begin{aligned} \mathbb{E}\left[U(\xi^\nu)\right] &\geq \mathbb{E}\left[U(\xi^\lambda)\right] \\ &= \mathbb{E}\left[\tilde{U}(\mathcal{Y}^\lambda(x_0)H_T^\lambda)\right] + \mathcal{Y}^\lambda(x_0)\mathbb{E}\left[H_T^\lambda\xi^\lambda\right] \\ &= \mathbb{E}\left[\tilde{U}(\mathcal{Y}^\lambda(x_0)H_T^\lambda)\right] + x_0\mathcal{Y}^\lambda(x_0) \\ \Rightarrow \quad \mathbb{E}\left[\tilde{U}(\mathcal{Y}^\lambda(x_0)H_T^\nu)\right] &\geq \mathbb{E}\left[\tilde{U}(\mathcal{Y}^\lambda(x_0)H_T^\lambda)\right] \quad \text{(D)} \end{aligned}$$

(D) $\Rightarrow$ (B) ([CK92, Thm.10.1 + Thm.9.1]:)

By martingale representation of $\mathbb{E}\left[H_T^\lambda\xi^\lambda \middle| \mathcal{F}_t\right]$ there exists a portfolio process π_t in the λ-market such that $X_t^{\lambda,\pi}$ satisfies

$$H_t^\lambda X_t^{\lambda,\pi} = x_0 + \int_0^t H_s^\lambda X_s^{\lambda,\pi}\left(\pi_s^\top\sigma - \theta_s^{\lambda^\top}\right)dW_s = \text{mart.rep.} = \mathbb{E}\left[H_T^\lambda\xi^\lambda \middle| \mathcal{F}_t\right]$$

$$\text{where} \quad dX_t^{\lambda,\pi} = X_t^{\lambda,\pi}\left(\left(r + \delta(\lambda_t) + \pi_t^\top\lambda_t\right)dt + \pi_t^\top\sigma_t dW_t^0\right)$$

i.e. π is the optimal strategy in the λ-market.

If now $\pi_t \in K$ and $\delta(\lambda_t) + \pi_t^\top\lambda_t = 0$ then $X_t^\pi = X_t^{\lambda,\pi}$ is also a wealth process in the original market and we are done.

Step I: Fix some ν_t, some $\varepsilon \in (0,1)$ and the stopping times $\tau_n \nearrow T$ below. Then define

$$\lambda_n^{\nu,\epsilon}(t) := \lambda_t + \varepsilon(\nu_t - \lambda_t)\mathbb{1}_{\{t \leq \tau_n\}}$$

$$L_t^\nu := \int_0^t \tilde{\delta}^\nu(\lambda_s)ds \quad , \text{ where } \tilde{\delta}^\nu(\lambda_t) := \begin{cases} -\delta(\lambda_t), & \text{if } \nu_t = 0 \\ \delta(\nu_t - \lambda_t), & \text{else} \end{cases}$$

$$N_t^\nu := \int_0^t \left(\sigma^{-1}(\nu_s - \lambda_s)\right)^\top dW_s^\lambda$$

$$\begin{aligned} \tau_n := T \wedge \inf \Bigg\{ & t \in [0,T] : |L_t^\nu| \geq n \text{ or } |N_t^\nu| \geq n \\ & \text{or } \int_0^t \left\|\sigma^{-1}(\nu_s - \lambda_s)\right\|^2 ds \geq n \text{ or } \int_0^t \left\|\theta_s^\lambda\right\|^2 ds \geq n \\ & \text{or } \int_0^t (\beta_s^\lambda X_s^{\lambda,\pi})^2 \left\|\sigma^{-1}(\nu_s - \lambda_s) + (L_s^\nu + N_s^\nu)\sigma^\top \pi_s\right\|^2 ds \geq n \Bigg\} \end{aligned}$$

Step II: define (for any $y > 0$, $\varepsilon \in (0,1)$)

$$Y_n^\varepsilon := \frac{1}{\varepsilon}\left(\tilde{U}\left(yH_T^{\lambda_n^{\nu,\epsilon}}\right) - \tilde{U}\left(yH_T^\lambda\right)\right)$$

then

$$\limsup_{\varepsilon \searrow 0} \mathbb{E}\left[Y_n^\varepsilon\right] \leq y\mathbb{E}\left[\xi^\lambda H_T^\lambda \left(L_{\tau_n}^\nu + N_{\tau_n}^\nu\right)\right] \tag{A.1}$$

$$= y\mathbb{E}\left[\int_0^{\tau_n} H_t^\lambda X_t^{\lambda,\pi}\pi_t^\top(\nu_t - \lambda_t)dt + \int_0^{\tau_n} H_t^\lambda X_t^{\lambda,\pi} dL_t^\nu\right] \tag{A.2}$$

$$= \begin{cases} y\mathbb{E}\left[\int_0^{\tau_n} H_t^\lambda X_t^{\lambda,\pi}\left(-\pi_t^\top \lambda_t - \delta(\lambda_t)\right)dt\right], & \text{if } \nu = 0 \\ y\mathbb{E}\left[\int_0^{\tau_n} H_t^\lambda X_t^{\lambda,\pi}\left(\pi_t^\top \rho_t + \delta(\rho_t)\right)dt\right], & \text{if } \nu \neq 0, \rho := \nu - \lambda \end{cases}$$

Proof of inequality (A.1):

Note that the random variables Y_n^ε are bounded from above:
By the mean value theorem, the increasing $\tilde{U}'(y) = -I(y)$ and (D) we get

$$\exists y^* : \frac{\tilde{U}(yH_T^{\lambda_n^{\nu,\epsilon}}) - \tilde{U}(yH_T^\lambda)}{yH_T^{\lambda_n^{\nu,\epsilon}} - yH_T^\lambda} = \tilde{U}'(y^*)$$

CASE I : $H_T^{\lambda_n^{\nu,\epsilon}} \geq H_T^\lambda \quad \Rightarrow \tilde{U}'(y^*) \leq -I\left(yH_T^{\lambda_n^{\nu,\epsilon}}\right)$
CASE II: $H_T^{\lambda_n^{\nu,\epsilon}} \leq H_T^\lambda \quad \Rightarrow \tilde{U}'(y^*) \geq -I\left(yH_T^{\lambda_n^{\nu,\epsilon}}\right)$
In both cases holds

$$Y_n^\varepsilon \leq \frac{1}{\varepsilon}\left(yH_T^\lambda - yH_T^{\lambda_n^{\nu,\epsilon}}\right) I\left(yH_T^{\lambda_n^{\nu,\epsilon}}\right)$$

Also we know by definition of τ_n

$$\frac{H_t^{\lambda_n^{\nu,\epsilon}}}{H_t^\lambda} = \exp\left(-\int_0^{t\wedge\tau_n} \delta\left(\lambda_s + \varepsilon(\nu_s - \lambda_s)\right) - \delta(\lambda_s)ds - \varepsilon N_{t\wedge\tau_n} - \frac{\varepsilon^2}{2}\int_0^{t\wedge\tau_n} \left\|\sigma^{-1}(\nu_s - \lambda_s\right\|^2 ds\right)$$

$$\geq \exp\left(-\varepsilon(L^{\nu}_{t\wedge\tau_n} + N^{\nu}_{t\wedge\tau_n}) - \frac{\varepsilon^2}{2}\int_0^{t\wedge\tau_n} \left\|\sigma^{-1}(\nu_s - \lambda_s\right\|^2 ds\right) =: \Delta_n^{\varepsilon}$$
$$\geq e^{-3\varepsilon n}$$

hence

$$\begin{aligned} Y_n^{\varepsilon} &\leq \frac{1}{\varepsilon}\left(yH_T^{\lambda} - yH_T^{\lambda_n^{\nu,\epsilon}}\right) I\left(yH_T^{\lambda_n^{\nu,\epsilon}}\right) \\ &= \frac{1}{\varepsilon}\left(1 - \frac{H_T^{\lambda_n^{\nu,\epsilon}}}{H_T^{\lambda}}\right) yH_T^{\lambda} I\left(yH_T^{\lambda_n^{\nu,\epsilon}}\right) \\ &\leq \frac{1}{\varepsilon}\left(1 - e^{-3\varepsilon n}\right) yH_T^{\lambda} I\left(ye^{-3\varepsilon n}H_T^{\lambda}\right) \\ \leq Y_n &:= K_n y H_T^{\lambda} I(e^{-3n} yH_T^{\lambda}) \\ \text{where } K_n &:= \sup_{\varepsilon\in(0,1)} \frac{1 - e^{-3\varepsilon n}}{\varepsilon} \end{aligned}$$

note that [CK92] get a slightly different formula: $K_n = \sup_{\varepsilon\in(0,1)} \frac{e^{3\varepsilon n} - 1}{\varepsilon}$
However, as $\mathbb{E}[Y_n] = K_n y \mathcal{X}^{\lambda}(ye^{-3n}) < \infty$ we can apply Fatou's Lemma:

$$\limsup_{\varepsilon\searrow 0} \mathbb{E}[Y_n^{\varepsilon}] \leq \mathbb{E}\left[\limsup_{\varepsilon\searrow 0} Y_n^{\varepsilon}\right]$$

On the other hand the random variables Y_n^{ε} are also bounded from above by:

$$Y_n^{\varepsilon} \leq \frac{1}{\varepsilon}(1 - \Delta_n^{\varepsilon})\, yH_T^{\lambda} I\left(yH_T^{\lambda}\right)$$
$$\text{where } \frac{1}{\varepsilon}(1 - \Delta_n^{\varepsilon}) \underset{\varepsilon\to 0}{\overset{l'H}{\to}} L^{\nu}_{t\wedge\tau_n} + N^{\nu}_{t\wedge\tau_n}$$

Therefore (A.1) follows immediately for $y = \mathcal{Y}^{\lambda}(x_0)$, i.e. $\xi^{\lambda} = I\left(yH_T^{\lambda}\right)$.

Proof of equality (A.2)
We know in the λ-market

$$\begin{aligned} d\big(\beta_t^{\lambda} X_t^{\lambda,\pi}\big) &= \beta_t^{\lambda} X_t^{\lambda,\pi} \pi_t^{\top} \sigma dW_t^{\lambda} \\ \Rightarrow d\big(\beta_t^{\lambda} X_t^{\lambda,\pi}(L_t^{\nu} + N_t^{\nu})\big) &= \beta_t^{\lambda} X_t^{\lambda,\pi}(dL_t^{\nu} + dN_t^{\nu}) + (L_t^{\nu} + N_t^{\nu})\beta_t^{\lambda} X_t^{\lambda,\pi}\pi_t^{\top}\sigma dW_t^{\lambda} + \beta_t^{\lambda} X_t^{\lambda,\pi}\pi_t^{\top}(\nu_t - \lambda_t)dt \\ &= \beta_t^{\lambda} X_t^{\lambda,\pi}\left(\left(\sigma^{-1}(\nu_t - \lambda_t)\right)^{\top} dW_t^{\lambda} + (L_t^{\nu} + N_t^{\nu})\pi_t^{\top}\sigma dW_t^{\lambda} + \pi_t^{\top}(\nu_t - \lambda_t)dt + dL_t^{\nu}\right) \end{aligned}$$

Now, $W^{\lambda}_{t\wedge\tau_n}$ is a Brownian motion under R_n, where $\frac{dR_n}{d\mathbb{P}} = Z^{\lambda}_{\tau_n}$. Hence taking the integral from 0 to τ_n and then the expectation w.r.t R_n on both sides leads to:

$$\mathbb{E}^R\left[\beta^{\lambda}_{\tau_n} X^{\lambda,\pi}_{\tau_n}(L^{\nu}_{\tau_n} + N^{\nu}_{\tau_n})\right] = \mathbb{E}^R\left[\int_0^{\tau_n} \beta_t^{\lambda} X_t^{\lambda,\pi}\left(\pi_t^{\top}(\nu_t - \lambda_t)dt + dL_t^{\nu}\right) dt\right]$$

$$\Rightarrow \quad \mathbb{E}\left[H^{\lambda}_{\tau_n} X^{\lambda,\pi}_{\tau_n} (L^{\nu}_{\tau_n} + N^{\nu}_{\tau_n})\right] = \mathbb{E}\left[\int_0^{\tau_n} H^{\lambda}_t X^{\lambda,\pi}_t \left(\pi_t^{\top}(\nu_t - \lambda_t)dt + dL^{\nu}_t\right) dt\right]$$

Note that $H^{\lambda}_{\tau_n} X^{\lambda,\pi}_{\tau_n} = \mathbb{E}\left[H^{\lambda}_T \xi^{\lambda} \middle| \mathcal{F}_{\tau_n}\right]$ and hence

$$\mathbb{E}\left[H^{\lambda}_{\tau_n} X^{\lambda,\pi}_{\tau_n} (L^{\nu}_{\tau_n} + N^{\nu}_{\tau_n})\right] = \mathbb{E}\left[\mathbb{E}\left[H^{\lambda}_T \xi^{\lambda} \middle| \mathcal{F}_{\tau_n}\right] (L^{\nu}_{\tau_n} + N^{\nu}_{\tau_n})\right] = \mathbb{E}\left[H^{\lambda}_T \xi^{\lambda} (L^{\nu}_{\tau_n} + N^{\nu}_{\tau_n})\right]$$

which finishes the proof of equality (A.2)

Step III:
By assumption (D) we get in (A.1) LHS ≥ 0 for $y = \mathcal{Y}^{\lambda}(x_0)$, hence (A.2) ≥ 0. Hence

$$\begin{aligned} -\pi_t^{\top}\lambda_t - \delta(\lambda_t) &\geq 0 \\ \pi_t^{\top}\rho_t + \delta(\rho_t) &\geq 0 \ \forall \rho_t \neq -\lambda_t \end{aligned}$$

By considering $\rho_t = \lambda_t$ we get $\pi_t^{\top}\lambda_t + \delta(\lambda_t) = 0$.
For the case $\rho_t = -\lambda_t$ we get

$$\begin{aligned} & -\delta(\lambda) = -\sup_{\pi \in K} -\pi^{\top}\lambda = \inf_{\pi \in K} \pi^{\top}\lambda \leq \sup_{\pi \in K} \pi^{\top}\lambda = \delta(-\lambda) \\ \Rightarrow \quad & -\pi_t^{\top}\rho_t = \pi_t^{\top}\lambda_t = -\delta(\lambda_t) \leq \delta(-\lambda_t) = \delta(\rho_t) \\ \Rightarrow \quad & -\pi_t^{\top}\rho_t \leq \delta(\rho_t) \qquad \forall \rho_t \\ \Rightarrow \quad & \pi_t \in K \quad \text{by [Roc70, Thm.13.1]} \end{aligned}$$

□

A.3 Proofs from Chapter 3

A.3.1 Choice of I in Case 3

Let I and J be the active and passive dimensions such that (λ, π) are the optimal solution in Case 3 and let (λ', π') be the solution if we chose I' and J' instead.
1) Let $i \in I$ with $\pi_i > 0$ be an active dimension.
If we choose $I' = I \setminus \{i\}$ and $J' = J \cup \{i\}$ then we get $\lambda'_i < 0$.
2) Let $j \in J$ with $\lambda_j > 0$ be a passive dimension (by Lemma 3.7).
If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ then we get $\pi'_j < 0$.
Hence several false choices of active and passive dimensions lead to non-admissible solutions, reveal that the choice was wrong.

Proof. Note the (correct) optimal solution:

$$\begin{aligned}\lambda_J &= -(\mu - r\mathbb{1})_J + \Sigma_{JI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I = -(\mu - r\mathbb{1})_J + \Sigma_{JI}\pi_I \\ \pi_I &= (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I\end{aligned}$$

and the (wrong) optimal solution:

$$\begin{aligned}\lambda'_{J'} &= -(\mu - r\mathbb{1})_{J'} + \Sigma_{J'I'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1}+)_{I'} = -(\mu - r\mathbb{1})_{J'} + \Sigma_{J'I'}\pi'_{I'} \\ \pi'_{I'} &= (\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'}\end{aligned}$$

First assume $I' = I \setminus \{i\}$ and $J' = J \cup \{i\}$. To show: $\lambda'_i < 0$.
We know for π_i:

$$\begin{aligned}0 < \pi_i &= ((\Sigma_{II})^{-1})_{iI}(\mu - r\mathbb{1})_I \\ &= ((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i + ((\Sigma_{II})^{-1})_{iI'}(\mu - r\mathbb{1})_{I'} \\ &= ((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i - ((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\ \Leftrightarrow \quad 0 &< (\mu - r\mathbb{1})_i - \Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'}\end{aligned}$$

Now consider λ'_i (where $i \in J'$)

$$\lambda'_i = -(\mu - r\mathbb{1})_i + \Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} < 0$$

and we have shown the first part.
For the second part assume $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$. To show: $\pi'_j < 0$.
We know for λ_j:

$$0 < \lambda_j = -(\mu - r\mathbb{1})_j + \Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I$$

Now consider π'_j (where $j \in I'$).

$$\begin{aligned}\pi'_j &= ((\Sigma_{I'I'})^{-1})_{jI'}(\mu - r\mathbb{1})_{I'} \\ &= ((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j + ((\Sigma_{I'I'})^{-1})_{jI}(\mu - r\mathbb{1})_I\end{aligned}$$

$$
\begin{aligned}
&= ((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j - ((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&= -((\Sigma_{I'I'})^{-1})_{jj}\lambda_j \\
&< 0
\end{aligned}
$$

hence we have shown the second part. □

A.3.2 Choice of I in Case 4

Let I and J be the active and passive dimensions such that (λ, π) are the optimal solution in Case 4 and let (λ', π') be the solution if we chose I' and J' instead.
1) Let $i \in I$ with $\pi_i > 0$ be an active dimension.
If we choose $I' = I \setminus \{i\}$ and $J' = J \cup \{i\}$ then we get $\lambda'_i < \bar{\lambda}'$.
2) Let $j \in J$ with $\lambda_j > \bar{\lambda}$ be a passive dimension (by Theorem 3.11).
If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ then we get $\pi'_j < 0$.
Hence several false choices of active and passive dimensions lead to non-admissible solutions, reveal that the choice was wrong.

Proof. Note the (correct) optimal solution:

$$
\begin{aligned}
\bar{\lambda} &= \frac{1 - \mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I}{\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I} \\
\lambda_J &= -(\mu - r\mathbb{1})_J + \Sigma_{JI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}\mathbb{1})_I = -(\mu - r\mathbb{1})_J + \Sigma_{JI}\pi_I \\
\pi_I &= (\Sigma_{II})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}\mathbb{1})_I
\end{aligned}
$$

and the (wrong) optimal solution:

$$
\begin{aligned}
\bar{\lambda}' &= \frac{1 - \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'}}{\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}} \\
\lambda'_{J'} &= -(\mu - r\mathbb{1})_{J'} + \Sigma_{J'I'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}'\mathbb{1})_{I'} = -(\mu - r\mathbb{1})_{J'} + \Sigma_{J'I'}\pi'_{I'} \\
\pi'_{I'} &= (\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}'\mathbb{1})_{I'}
\end{aligned}
$$

First assume $I' = I \setminus \{i\}$ and $J' = J \cup \{i\}$. To show: $\lambda'_i < \bar{\lambda}'$.
We know for π_i:

$$
\begin{aligned}
0 < \pi_i &= ((\Sigma_{II})^{-1})_{iI}(\mu - r\mathbb{1} + \bar{\lambda}\mathbb{1})_I \\
&= ((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1} + \bar{\lambda}\mathbb{1})_i + ((\Sigma_{II})^{-1})_{iI'}(\mu - r\mathbb{1} + \bar{\lambda}\mathbb{1})_{I'} \\
&= ((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i + ((\Sigma_{II})^{-1})_{ii}\bar{\lambda} \\
&\quad - ((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} - ((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\bar{\lambda} \\
\Leftrightarrow \quad 0 &< (\mu - r\mathbb{1})_i + \bar{\lambda} \\
&\quad - \Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&\quad - \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\bar{\lambda} \\
\Leftrightarrow \quad 0 &< \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_i + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I\bar{\lambda}
\end{aligned}
$$

$$
\begin{aligned}
&\quad - \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&\quad - \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\bar{\lambda} \\
&= 1 + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_i \\
&\quad - \mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&\quad - \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'} \\
&\quad + \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
\Leftrightarrow \quad 1 &> -\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_i \\
&\quad + \mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&\quad + \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'} \\
&\quad - \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&= \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'} \\
&\quad - ((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i \\
&\quad - \mathbb{1}_{I'}^\top((\Sigma_{II})^{-1})_{I'i}(\mu - r\mathbb{1})_i \\
&\quad - ((\Sigma_{II})^{-1})_{iI'}\mathbb{1}_{I'}(\mu - r\mathbb{1})_i \\
&\quad - \mathbb{1}_{I'}^\top((\Sigma_{II})^{-1})_{I'I'}\mathbb{1}_{I'}(\mu - r\mathbb{1})_i \\
&\quad + ((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i \\
&\quad + \mathbb{1}_{I'}^\top((\Sigma_{II})^{-1})_{I'i}(\mu - r\mathbb{1})_i \\
&\quad + ((\Sigma_{II})^{-1})_{iI'}(\mu - r\mathbb{1})_{I'} \\
&\quad + \mathbb{1}_{I'}^\top((\Sigma_{II})^{-1})_{I'I'}(\mu - r\mathbb{1})_{I'} \\
&\quad + ((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&\quad + \mathbb{1}_{I'}^\top((\Sigma_{II})^{-1})_{I'i}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&\quad + ((\Sigma_{II})^{-1})_{iI'}\mathbb{1}_{I'}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&\quad + \mathbb{1}_{I'}^\top((\Sigma_{II})^{-1})_{I'I'}\mathbb{1}_{I'}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&\quad - \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i \\
&\quad - \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\mathbb{1}_{I'}^\top((\Sigma_{II})^{-1})_{I'i}(\mu - r\mathbb{1})_i \\
&\quad - \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}((\Sigma_{II})^{-1})_{iI'}(\mu - r\mathbb{1})_{I'} \\
&\quad - \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\mathbb{1}_{I'}^\top((\Sigma_{II})^{-1})_{I'I'}(\mu - r\mathbb{1})_{I'} \\
&= \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'} \\
&\quad - ((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i \\
&\quad + \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\Sigma_{I'i}((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i \\
&\quad + ((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}(\mu - r\mathbb{1})_i
\end{aligned}
$$

$$
\begin{aligned}
&- \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}(\mu - r\mathbb{1})_i \\
&- \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\Sigma_{I'i}((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}(\mu - r\mathbb{1})_i \\
&+ ((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i \\
&- \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\Sigma_{I'i}((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i \\
&- ((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&+ \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&+ \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\Sigma_{I'i}((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&+ ((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&- \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\Sigma_{I'i}((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&- ((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&+ \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&+ \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\Sigma_{I'i}((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&- \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i \\
&+ \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\Sigma_{I'i}((\Sigma_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i \\
&+ \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&- \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&- \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\Sigma_{I'i}((\Sigma_{II})^{-1})_{ii}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
= {} & \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'} \\
&- \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}(\mu - r\mathbb{1})_i \\
&+ \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&+ \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&- \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'}
\end{aligned}
$$

because all other terms cancel out.
Now consider λ'_i (where $i \in J'$)

$$
\begin{aligned}
\lambda'_i &= -(\mu - r\mathbb{1})_i + \Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}'\mathbb{1})_{I'} \\
\Rightarrow \quad \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\lambda'_i &= -\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}(\mu - r\mathbb{1})_i \\
&\quad + \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\Sigma_{iI'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&\quad + \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'} - \Sigma_{iI'}(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'}
\end{aligned}
$$

hence

$$
\begin{aligned}
\bar{\lambda}' &> \lambda'_i \\
\Leftrightarrow \quad 1 - \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} &> \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\lambda'_i \\
&= -\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}(\mu - r\mathbb{1})_i
\end{aligned}
$$

$$
\begin{aligned}
&& &+ \mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1} \mathbb{1}_{I'} \Sigma_{iI'} (\Sigma_{I'I'})^{-1} (\mu - r\mathbb{1})_{I'} \\
&& &+ \Sigma_{iI'} (\Sigma_{I'I'})^{-1} \mathbb{1}_{I'} \\
&& &- \Sigma_{iI'} (\Sigma_{I'I'})^{-1} \mathbb{1}_{I'} \mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1} (\mu - r\mathbb{1})_{I'} \\
\Leftrightarrow && 1 &> \Sigma_{iI'} (\Sigma_{I'I'})^{-1} \mathbb{1}_{I'} \\
&& &- \mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1} \mathbb{1}_{I'} (\mu - r\mathbb{1})_i \\
&& &+ \mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1} (\mu - r\mathbb{1})_{I'} \\
&& &+ \mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1} \mathbb{1}_{I'} \Sigma_{iI'} (\Sigma_{I'I'})^{-1} (\mu - r\mathbb{1})_{I'} \\
&& &- \Sigma_{iI'} (\Sigma_{I'I'})^{-1} \mathbb{1}_{I'} \mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1} (\mu - r\mathbb{1})_{I'} \\
\Leftrightarrow && \pi_i &> 0
\end{aligned}
$$

hence we have shown the first part.
For the second part assume $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$. To show: $\pi_j' < 0$.
We know for λ_j:

$$
\begin{aligned}
&& \bar\lambda &< \lambda_j = -(\mu - r\mathbb{1})_j + \Sigma_{jI} (\Sigma_{II})^{-1} (\mu - r\mathbb{1} + \bar\lambda \mathbb{1})_I \\
\Leftrightarrow && 1 - \mathbb{1}_I^\top (\Sigma_{II})^{-1} (\mu - r\mathbb{1})_I &< \mathbb{1}_I^\top (\Sigma_{II})^{-1} \mathbb{1}_I \lambda_j \\
&& &= -\mathbb{1}_I^\top (\Sigma_{II})^{-1} \mathbb{1}_I (\mu - r\mathbb{1})_j \\
&& &\quad + \mathbb{1}_I^\top (\Sigma_{II})^{-1} \mathbb{1}_I \Sigma_{jI} (\Sigma_{II})^{-1} (\mu - r\mathbb{1})_I \\
&& &\quad + \Sigma_{jI} (\Sigma_{II})^{-1} \mathbb{1}_I \\
&& &\quad - \Sigma_{jI} (\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top (\Sigma_{II})^{-1} (\mu - r\mathbb{1})_I \\
\Leftrightarrow && 1 &< \Sigma_{jI} (\Sigma_{II})^{-1} \mathbb{1}_I \\
&& &\quad - \mathbb{1}_I^\top (\Sigma_{II})^{-1} \mathbb{1}_I (\mu - r\mathbb{1})_j \\
&& &\quad + \mathbb{1}_I^\top (\Sigma_{II})^{-1} (\mu - r\mathbb{1})_I \\
&& &\quad + \mathbb{1}_I^\top (\Sigma_{II})^{-1} \mathbb{1}_I \Sigma_{jI} (\Sigma_{II})^{-1} (\mu - r\mathbb{1})_I \\
&& &\quad - \Sigma_{jI} (\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top (\Sigma_{II})^{-1} (\mu - r\mathbb{1})_I
\end{aligned}
$$

Now consider π_j' (where $j \in I'$). To show:

$$
\begin{aligned}
0 > \pi_j' &= ((\Sigma_{I'I'})^{-1})_{jI'} (\mu - r\mathbb{1} + \bar\lambda' \mathbb{1})_{I'} \\
&= ((\Sigma_{I'I'})^{-1})_{jj} (\mu - r\mathbb{1})_j \\
&\quad + ((\Sigma_{I'I'})^{-1})_{jj} \bar\lambda' \\
&\quad + ((\Sigma_{I'I'})^{-1})_{jI} (\mu - r\mathbb{1})_I \\
&\quad + ((\Sigma_{I'I'})^{-1})_{jI} \mathbb{1}_I \bar\lambda' \\
&= ((\Sigma_{I'I'})^{-1})_{jj} (\mu - r\mathbb{1})_j \\
&\quad + ((\Sigma_{I'I'})^{-1})_{jj} \bar\lambda' \\
&\quad - ((\Sigma_{I'I'})^{-1})_{jj} \Sigma_{jI} (\Sigma_{II})^{-1} (\mu - r\mathbb{1})_I
\end{aligned}
$$

$$
\begin{aligned}
& && - ((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\bar{\lambda}' \\
\Leftrightarrow & && 0 > \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}(\mu - r\mathbb{1})_j \\
& && + 1 - \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
& && - \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
& && - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I \\
& && + \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
\Leftrightarrow & && 1 < \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I \\
& && - \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}(\mu - r\mathbb{1})_j \\
& && + \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
& && + \mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
& && - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
& && = \ldots \text{ (as above, using the matrix inversion from below)} \\
& && = \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I \\
& && - \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_j \\
& && + \mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
& && + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
& && - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
\Leftrightarrow & && \bar{\lambda} < \lambda_j
\end{aligned}
$$

hence we have shown the second part. □

A.3.3 Dimensions on the boundary

Let I and J be the active and passive dimensions such that (λ, π) are the optimal solution in Case 4 and let (λ', π') be the solution if we chose I' and J' instead.

If $j \in J$ is a passive dimension that is almost invested in, hence $\lambda_j = \bar{\lambda}$ according to Remark 3.13, then j can also be considered an active dimension.
In particular:

Let $j \in J$ with $\pi_j = 0$ and $\lambda_j = \bar{\lambda}$ be a passive dimension that is almost invested in.
Then: If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ we get $\lambda' = \lambda$ and $\pi' = \pi$.

Proof. (Only for case 4 as case 3 is a trivial consequence for $\bar{\lambda}00$.)
Note the (correct) optimal solution:

$$
\begin{aligned}
\bar{\lambda} &= \frac{1 - \mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I}{\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I} \\
\lambda_J &= -(\mu - r\mathbb{1})_J + \Sigma_{JI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}\mathbb{1})_I = -(\mu - r\mathbb{1})_J + \Sigma_{JI}\pi_I \\
\pi_I &= (\Sigma_{II})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}\mathbb{1})_I
\end{aligned}
$$

and the other solution:

$$\begin{aligned}
\bar{\lambda}' &= \frac{1 - \mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'}}{\mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}} \\
\lambda'_{J'} &= -(\mu - r\mathbb{1})_{J'} + \Sigma_{J'I'}(\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}'\mathbb{1})_{I'} = -(\mu - r\mathbb{1})_{J'} + \Sigma_{J'I'}\pi'_{I'} \\
\pi'_{I'} &= (\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}'\mathbb{1})_{I'}
\end{aligned}$$

Let $j \in J$ with $\lambda_j = \bar{\lambda}$ and $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$.
From the proof of Lemma 3.18 we get for λ_j:

$$\begin{aligned}
& \lambda_j = \bar{\lambda} \\
\Leftrightarrow \quad & 1 = \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I \\
& \quad - \mathbb{1}_I^\top (\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_j \\
& \quad + \mathbb{1}_I^\top (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
& \quad + \mathbb{1}_I^\top (\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
& \quad - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \qquad (*) \\
\Leftrightarrow \quad & \pi'_j = 0
\end{aligned}$$

note for $\bar{\lambda} = \frac{1 - \mathbb{1}_I^\top (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I}{\mathbb{1}_I^\top (\Sigma_{II})^{-1}\mathbb{1}_I}$:

$$\begin{aligned}
\bar{\lambda} = \lambda_j &= -(\mu - r\mathbb{1})_j + \Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}\mathbb{1})_I \\
&= -(\mu - r\mathbb{1})_j + \Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I + \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\bar{\lambda} \\
\Rightarrow \quad \bar{\lambda} &= \frac{-(\mu - r\mathbb{1})_j + \Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I}{1 - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I}
\end{aligned}$$

also

$$\begin{aligned}
\bar{\lambda}' &= \frac{1 - \mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'}}{\mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}} \\
\Rightarrow \quad \bar{\lambda}'(\mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1}\mathbb{1}_{I'}) &= 1 - \mathbb{1}_{I'}^\top (\Sigma_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
&= 1 - ((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j \\
&\quad - ((\Sigma_{I'I'})^{-1})_{jI}(\mu - r\mathbb{1})_I \\
&\quad - \mathbb{1}_I^\top ((\Sigma_{I'I'})^{-1})_{Ij}(\mu - r\mathbb{1})_j \\
&\quad - \mathbb{1}_I^\top ((\Sigma_{I'I'})^{-1})_{II}(\mu - r\mathbb{1})_I \\
&= 1 - ((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j \\
&\quad + ((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad + \mathbb{1}_I^\top (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j \\
&\quad - \mathbb{1}_I^\top (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - \mathbb{1}_I^\top (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I
\end{aligned}$$

$$
\begin{aligned}
\Rightarrow \quad & \bar{\lambda}'(\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'})(1-\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I) \\
&= 1-\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I \\
&\quad -(1-\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I)((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad +(1-\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I)((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad +(1-\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I)\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad -(1-\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I)\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad -(1-\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I)\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&= 1-\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I \\
&\quad -((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad +\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad +((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad -\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad +\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad -\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad -\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad +\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad -\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad +\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&= 1-\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I \\
&\quad -((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad +2\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad -\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad -\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad +((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad +\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad -2\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad +\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\overset{(*)}{=} -\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I(\mu-r\mathbb{1})_j \\
&\quad -((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad +2\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad -\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu-r\mathbb{1})_j \\
&\quad +\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I \\
&\quad +((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu-r\mathbb{1})_I
\end{aligned}
$$

$$
\begin{aligned}
&- 2\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&+ \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I
\end{aligned}
$$

from

$$
\begin{aligned}
\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'} &= ((\Sigma_{I'I'})^{-1})_{jj} + ((\Sigma_{I'I'})^{-1})_{jI}\mathbb{1}_I + \mathbb{1}_I^\top((\Sigma_{I'I'})^{-1})_{Ij} + \mathbb{1}_I^\top((\Sigma_{I'I'})^{-1})_{II}\mathbb{1}_I \\
&= ((\Sigma_{I'I'})^{-1})_{jj} \\
&\quad - ((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I \\
&\quad - \mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj} \\
&\quad + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I \\
&\quad + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I
\end{aligned}
$$

we get

$$
\begin{aligned}
&\bar{\lambda}(\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'})(1 - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I) \\
&\qquad = (\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'})(-(\mu - r\mathbb{1})_j + \Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I) \\
&\qquad = -((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j \\
&\qquad\quad + ((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_j \\
&\qquad\quad + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j \\
&\qquad\quad - \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_j \\
&\qquad\quad - \mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_j \\
&\qquad\quad + ((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\qquad\quad - ((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\qquad\quad - \mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\qquad\quad + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\qquad\quad + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&= \bar{\lambda}'(\mathbb{1}_{I'}^\top(\Sigma_{I'I'})^{-1}\mathbb{1}_{I'})(1 - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I)
\end{aligned}
$$

Hence $\bar{\lambda}' = \bar{\lambda}$
Now consider π_I and π_I':

$$
\begin{aligned}
\pi_I &= (\Sigma_{II})^{-1}(\mu - r\mathbb{1} + \bar{\lambda}\mathbb{1})_I \\
\Rightarrow \quad \pi_I(1 - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I) &= (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I(1 - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I) \\
&\quad + (\Sigma_{II})^{-1}\mathbb{1}_I(-(\mu - r\mathbb{1})_j + \Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I) \\
&= (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - (\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_j \\
&\quad + (\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I
\end{aligned}
$$

and using $\bar{\lambda}' = \bar{\lambda}$

$$\begin{aligned}
\pi'_I &= ((\Sigma_{I'I'})^{-1})_{II'}(\mu - r\mathbb{1} + \bar{\lambda}'\mathbb{1})_{I'} \\
&= ((\Sigma_{I'I'})^{-1})_{II}(\mu - r\mathbb{1})_I \\
&\quad + ((\Sigma_{I'I'})^{-1})_{II}\mathbb{1}_I\bar{\lambda} \\
&\quad + ((\Sigma_{I'I'})^{-1})_{Ij}(\mu - r\mathbb{1})_j \\
&\quad + ((\Sigma_{I'I'})^{-1})_{Ij}\bar{\lambda} \\
&= (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad + (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad + (\Sigma_{II})^{-1}\mathbb{1}_I\bar{\lambda} \\
&\quad + (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\bar{\lambda} \\
&\quad - (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j \\
&\quad - (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\bar{\lambda} \\
\Rightarrow \quad \pi'_I(1 - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I) &= (1 - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I)(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad + (1 - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I)(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad + (\Sigma_{II})^{-1}\mathbb{1}_I(-(\mu - r\mathbb{1})_j + \Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I) \\
&\quad + (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I(-(\mu - r\mathbb{1})_j + \Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I) \\
&\quad - (1 - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I)(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j \\
&\quad - (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(-(\mu - r\mathbb{1})_j + \Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I) \\
&= (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad + (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - (\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_j \\
&\quad + (\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_j \\
&\quad + (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j \\
&\quad + \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I(\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j \\
&\quad + (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j \\
&\quad - (\Sigma_{II})^{-1}\Sigma_{Ij}((\Sigma_{I'I'})^{-1})_{jj}\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&= (\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&\quad - (\Sigma_{II})^{-1}\mathbb{1}_I(\mu - r\mathbb{1})_j
\end{aligned}$$

$$
\begin{aligned}
&+ (\Sigma_{II})^{-1}\mathbb{1}_I\Sigma_{jI}(\Sigma_{II})^{-1}(\mu - r\mathbb{1})_I \\
&= \pi_I(1 - \Sigma_{jI}(\Sigma_{II})^{-1}\mathbb{1}_I)
\end{aligned}
$$

Hence $\pi'_I = \pi_I$. Also

$$
\begin{aligned}
\lambda_{J'} &= -(\mu - r\mathbb{1})_{J'} + \Sigma_{J'I}\pi_I \\
\text{and} \quad \lambda'_{J'} &= -(\mu - r\mathbb{1})_{J'} + \Sigma_{J'I'}\pi'_{I'} \\
&= -(\mu - r\mathbb{1})_{J'} + \Sigma_{J'I}\pi'_I = \lambda_{J'}
\end{aligned}
$$

This proofs the assumption.

□

A.4 Proofs from Chapter 4

A.4.1 Martingale approach from [BDL10]:

Consider the local martingale Z_t (which is a martingale by assumption)

$$dZ_t = -Z_t\theta_t^\top dW_t, \quad Z_0 = 1$$

the discount process $\beta_t := e^{-rt}$ such that $\tilde{X}_t := \beta_t X_t$ and the wealth equations

$$\begin{aligned} dX_t &= X_t\left(\left(r + \pi_t^\top(\mu_t - r)\right)dt + \pi_t^\top\sigma_t dW_t\right) \\ &= X_t\left(rdt + \pi_t^\top\sigma_t dW_t^{\mathbb{Q}}\right) \\ d\tilde{X}_t &= \tilde{X}_t\pi_t^\top\left((\mu_t - r)dt + \sigma dW_t\right) \\ &= \tilde{X}_t\pi_t^\top\sigma dW_t^{\mathbb{Q}} \qquad \text{for} \quad dW_t^{\mathbb{Q}} := dW_t + \theta dt \end{aligned}$$

By Girsanov $W_t^{\mathbb{Q}}$ is a Brownian motion under Q where $\frac{dQ}{dP} = Z_T, \quad Z_t = \frac{dQ}{dP}\Big|_{\mathcal{F}_t}$.
Define the P-martingale Z^0 via its dynamics $dZ_t^0 = Z_t^0\frac{\alpha}{1-\alpha}\theta_t^\top dW_t$
and the corresponding measure Q^0 via $\frac{dQ^0}{dP} = Z_t^0$ on $\mathcal{F}_t$.
Under power utility $U(x) = \frac{1}{\alpha}x^\alpha$,
we get $X_T^* = I(y\beta_T Z_T) = x_0 H_0^{-1}\beta_T^{\frac{1}{\alpha-1}}Z_T^{\frac{1}{\alpha-1}}$ where $H_0 := \mathbb{E}\left[\beta_T^{\frac{\alpha}{\alpha-1}}Z_T^{\frac{\alpha}{\alpha-1}}\right]$.
Note:

$$\begin{aligned} Z_t^0 &= \exp\int_0^t \frac{\alpha}{1-\alpha}\theta_s^\top dW_s - \frac{1}{2}\int_0^t\left(\frac{\alpha}{1-\alpha}\right)^2\theta_s^\top\theta_s ds \\ Z_T^{\frac{\alpha}{\alpha-1}} &= \exp\int_0^T \frac{\alpha}{1-\alpha}\theta_t^\top dW_t + \frac{1}{2}\int_0^T\frac{\alpha}{1-\alpha}\theta_t^\top\theta_t dt \\ &= Z_T^0\exp\frac{1}{2}\int_0^T\frac{\alpha}{(1-\alpha)^2}\theta_t^\top\theta_t dt \\ \Rightarrow H_0 &= \mathbb{E}^0\left[\left(Z_T^0\right)^{-1}\beta_T^{\frac{\alpha}{\alpha-1}}Z_T^{\frac{\alpha}{\alpha-1}}\right] \\ &= \mathbb{E}^0\left[\exp\frac{\alpha}{1-\alpha}\int_0^T r + \frac{1}{2}\frac{1}{1-\alpha}\theta_t^\top\theta_t dt\right] \\ \text{and define: } H_t &:= \mathbb{E}^0\left[\exp\frac{\alpha}{1-\alpha}\int_t^T r + \frac{1}{2}\frac{1}{1-\alpha}\theta_s^\top\theta_s ds\,\middle|\,\mathcal{F}_t\right] \end{aligned}$$

Then the optimal wealth process evolves as $X_t^* = x_0\frac{H_t}{H_0}\beta_t^{\frac{1}{\alpha-1}}Z_t^{\frac{1}{\alpha-1}}$.

Proof. This is just calculation and a change of measure via Bayes theorem. □

[BDL10] also show that the process H_t has a stochastic differential representation as

$$dH_t = H_t\big(\mu_H(t)dt + \sigma_H(t)dW_t\big)$$

Then the optimal portfolio process reads as

$$\pi_t^* = \frac{1}{1-\alpha}\Sigma_t^{-1}(\mu_t - r) + \sigma_t^{-\top}\sigma_H^\top(t)$$

Proof. $\tilde{X}_t^* = \frac{x_0}{H_0} H_t \beta_t^{\frac{\alpha}{\alpha-1}} Z_t^{\frac{1}{\alpha-1}}$ hence by Itô

$$d\tilde{X}_t^* = \tilde{X}_t^* \left((...)dt + \left(\frac{1}{1-\alpha}\theta_t + \sigma_H(t) \right) dW_t \right)$$

And from the wealth equation we know $d\tilde{X}_t^* = \tilde{X}_t^* \left((...)dt + \pi_t^{*\top} \sigma dW_t \right)$ □

Note that in general the determination of σ_H is barely possible.

Example:
We start with the observable form of our model:

$$dS_t = S_t(\hat{\mu}_t dt + \sigma dV_t)$$

where V_t is an $\mathcal{F}^S$-Brownian motion and $R_t, \hat{\mu}_t$ and V_t are observable. We have

$$\frac{dQ}{dP} = Z_T, \quad \text{where } dZ_t = -Z_t \hat{\theta}_t^\top dV_t$$
$$\frac{dQ^0}{dP} = Z_T^0, \quad \text{where } dZ_t^0 = Z_t^0 \frac{\alpha}{1-\alpha}\hat{\theta}_t^\top dV_t$$

Hence we get

$$\begin{aligned}
H_t &= \mathbb{E}^0 \left[\exp \frac{\alpha}{1-\alpha} \int_t^T r + \frac{1}{2}\frac{1}{1-\alpha}\hat{\theta}_s^\top \hat{\theta}_s ds \,\middle|\, \mathcal{F}_t^S \right] \\
&= \frac{1}{Z_t^0} \mathbb{E} \left[Z_T^0 \exp \frac{\alpha}{1-\alpha} \int_t^T r + \frac{1}{2}\frac{1}{1-\alpha}\hat{\theta}_s^\top \hat{\theta}_s ds \,\middle|\, \mathcal{F}_t^S \right] \\
&= \mathbb{E} \left[\exp \frac{\alpha}{1-\alpha} \int_t^T r + \frac{1}{2}\hat{\theta}_s^\top \hat{\theta}_s ds + \frac{\alpha}{1-\alpha}\int_t^T \hat{\theta}_s^\top dV_s \,\middle|\, \mathcal{F}_t^S \right] \\
&= Z_t \mathbb{E}^Q \left[Z_T^{-1} \exp \frac{\alpha}{1-\alpha} \int_t^T r + \frac{1}{2}\hat{\theta}_s^\top \hat{\theta}_s ds + \frac{\alpha}{1-\alpha}\int_t^T \hat{\theta}_s^\top dV_s \,\middle|\, \mathcal{F}_t^S \right] \\
&= \mathbb{E}^Q \left[\exp \frac{\alpha}{1-\alpha}(T-t)r + \frac{1}{1-\alpha} \int_t^T \frac{1}{2}\hat{\theta}_s^\top \hat{\theta}_s ds + \frac{1}{1-\alpha}\int_t^T \hat{\theta}_s^\top dV_s \,\middle|\, \mathcal{F}_t^S \right] \\
&= \exp\left(\frac{\alpha}{1-\alpha}(T-t)r \right) \mathbb{E}^Q \left[\exp \frac{1}{\alpha-1} \int_t^T \frac{1}{2}\hat{\theta}_s^\top \hat{\theta}_s ds - \frac{1}{\alpha-1}\int_t^T \hat{\theta}_s^\top dW_s^Q \,\middle|\, \mathcal{F}_t^S \right]
\end{aligned}$$

Now consider the orthogonal matrix P from the derivation of the filter that diagonalizes the covariance matrix of the prior of θ and the following definitions: $\tilde{\theta}_t := P\hat{\theta}_t, \tilde{W}_t^Q := PW_t^Q$. Then the latter is a Brownian motion with respect to Q. Hence

$$H_t = \exp\left(\frac{\alpha}{1-\alpha}(T-t)r \right) \mathbb{E}^Q \left[\exp \frac{1}{\alpha-1} \int_t^T \frac{1}{2}\tilde{\theta}_s^\top \tilde{\theta}_s ds - \frac{1}{\alpha-1}\int_t^T \hat{\theta}_s^\top d\tilde{W}_s^Q \,\middle|\, \mathcal{F}_t^S \right]$$

and on the other hand we get via the filter representations from Section 2.2

$$\tilde{\theta}_t = \bar{D}_t\left(\tilde{W}_t^Q + D^{-1}\tilde{\theta}_0\right)$$
$$\text{with}\quad \tilde{\theta}_i(t) = \delta_i(t)\tilde{W}_i^Q(t) + \frac{\delta_i(t)}{d_i}\tilde{\theta}_i(0)$$
$$\text{where } \delta_i(t) = \frac{d_i}{1+d_it} \text{ with } d\delta_i(t) = -\delta_i(t)^2dt$$
$$\text{where diag}\,(...d_i...) = D = P\sigma^{-1}\Sigma_0\sigma^{-\top}P^\top$$

i.e. the $\left(\tilde{\theta}_i(t), \tilde{W}_i^Q(t)\right)$ are independent of $\left(\tilde{\theta}_j(t), \tilde{W}_j^Q(t)\right)$ and we get.

$$\begin{aligned}
H_t &= \exp\left(\frac{\alpha}{1-\alpha}(T-t)r\right)\mathbb{E}^Q\left[\exp\frac{1}{\alpha-1}\int_t^T\frac{1}{2}\tilde{\theta}_s^\top\tilde{\theta}_s ds - \frac{1}{\alpha-1}\int_t^T\hat{\theta}_s^\top d\tilde{W}_s^Q\middle|\mathcal{F}_t^S\right]\\
&= \exp\left(\frac{\alpha}{1-\alpha}(T-t)r\right)\mathbb{E}^Q\left[\left(\frac{Z_T}{Z_t}\right)^{\frac{1}{\alpha-1}}\middle|\mathcal{F}_t^S\right]\\
&= \exp\left(\frac{\alpha}{1-\alpha}(T-t)r\right)\prod_{i=1}^d\mathbb{E}^Q\left[\left(\frac{Z_i(T)}{Z_i(t)}\right)^{\frac{1}{\alpha-1}}\middle|\mathcal{F}_t^S\right]\\
\text{where } Z_i(t) &= \exp\left(\int_0^t\frac{1}{2}\tilde{\theta}_i^2(s)ds - \int_0^t\tilde{\theta}_i(s)d\tilde{W}_i^Q(s)\right)
\end{aligned}$$

This can be solved like in [CLMZ06, Appendix A.1], resulting in

$$\begin{aligned}
H_t &= \exp\left(\frac{\alpha}{1-\alpha}(T-t)r\right)\prod_{i=1}^d\sqrt{\frac{\left(1+\delta_i(t)(T-t)\right)^{\frac{\alpha}{\alpha-1}}}{1+\delta_i(t)(T-t)\frac{\alpha}{\alpha-1}}}\exp\frac{\frac{1}{\alpha-1}\frac{\alpha}{\alpha-1}\tilde{\theta}_i^2(t)(T-t)}{2\left(1+\delta_i(t)(T-t)\frac{\alpha}{\alpha-1}\right)}\\
&=: f(t)\prod_{i=1}^d Y_i(t) =: f(t)\prod_{i=1}^d f_i(t)\exp c_i(t)\frac{\tilde{\theta}_i^2(t)}{2}
\end{aligned}$$

Via Itô's product-formula we get

$$\begin{aligned}
dY_i(t) &= (...)dt + Y_i(t)c_i(t)\tilde{\theta}_i(t)d\tilde{\theta}_i(t)\\
&= (...)dt + Y_i(t)c_i(t)\tilde{\theta}_i(t)\delta_i(t)d\tilde{W}_i^Q(t)\\
\text{hence}\quad dH_t &= (...)dt + H_t\sum_{i=1}^d\frac{\sigma_{Y_i}}{Y_i}d\tilde{W}_i^Q(t)\\
&= (...)dt + H_t\sum_{i=1}^d c_i(t)\tilde{\theta}_i(t)\delta_i(t)\tilde{W}_i^Q(t)
\end{aligned}$$

where

$$c_i(t)\delta_i(t) = \frac{\frac{1}{\alpha-1}\frac{\alpha}{\alpha-1}(T-t)\delta_i(t)}{1+\delta_i(t)(T-t)\frac{\alpha}{\alpha-1}}$$

$$
\begin{aligned}
&= \frac{1}{\alpha-1}\frac{\alpha d_i(T-t)}{(\alpha-1)(1+d_it)+\alpha d_i(T-t)} \\
&= \frac{1}{1-\alpha}\left(\frac{(1-\alpha)(1+d_it)}{(1-\alpha)(1+d_it)-\alpha d_i(T-t)}-1\right) \\
&=: \frac{1}{1-\alpha}\left(d_i'(t)-1\right) \\
&=: e_i'(t)
\end{aligned}
$$

hence

$$
\begin{aligned}
dH_t &= (...)dt + H_t\sum_{i=1}^{d} e_i'(t)\tilde{\theta}_i(t)\tilde{W}_i^Q(t) \\
&= (...)dt + H_t\tilde{\theta}_t^\top E'(t)d\tilde{W}_t^Q \\
&= (...)dt + H_t\hat{\theta}_t^\top P^\top E'(t)PdW_t^{\mathbb{Q}} \\
&= (...)dt + H_t\hat{\theta}_t^\top P^\top E'(t)PdV_t \\
&= (...)dt + H_t\sigma_H dV_t \\
\text{and } \sigma_H &= \hat{\theta}_t^\top P^\top E'(t)P \\
&= \hat{\theta}_t^\top P^\top \frac{1}{1-\alpha}\left(D'(t)-I\right)P
\end{aligned}
$$

Hence we get for the optimal solution

$$
\begin{aligned}
\pi_t^* &= \frac{1}{1-\alpha}\Sigma^{-1}(\hat{\mu}_t-r)+\sigma^{-\top}\sigma_H^\top(t) \\
&= \frac{1}{1-\alpha}\Sigma^{-1}(\hat{\mu}_t-r)+\sigma^{-\top}P^\top\frac{1}{1-\alpha}\left(D'(t)-I\right)P\hat{\theta}_t \\
&= \sigma^{-\top}P^\top\frac{1}{1-\alpha}D'(t)P\hat{\theta}_t \\
\text{where}\quad D'(t) &= \operatorname{diag}\left(...,\frac{(1-\alpha)(1+d_it)}{(1-\alpha)(1+d_it)-\alpha d_i(T-t)},...\right) \\
\text{and}\quad D &= \operatorname{diag}(...d_i...) = P\sigma^{-1}\Sigma_0\sigma^{-\top}P^\top
\end{aligned}
$$

A.4.2 Martingale approach from [Lak98]

For $dZ_t = -\theta^\top Z_t dW_t$ define the *optional projection* $\zeta_t := \mathbb{E}\left[Z_t \middle| \mathcal{F}_t^S\right]$ (if μ is observable via $\mathcal{F}_t^S$, then $\zeta_t = Z_t$). As $\mathbb{E}\left[Z_t \middle| \mathcal{F}_t^S\right] = \mathbb{E}\left[Z_T \middle| \mathcal{F}_t^S\right]$, ζ_t is a $\left(P,\mathcal{F}^S\right)$-martinagale.

Then we observe for $t \le u$: (by cond. exp. and the tower property)

$$
\begin{aligned}
&\text{let X be } \mathcal{F}_u^S\text{-measurable, then} \quad \mathbb{E}^Q[X] = \mathbb{E}[Z_uX] = \mathbb{E}[\zeta_uX] \\
&\text{let X be } \mathcal{F}_u^S\text{-measurable, then} \quad \mathbb{E}^Q\left[X \middle| \mathcal{F}_t^S\right] = \frac{1}{\zeta_t}\mathbb{E}\left[\zeta_uX \middle| \mathcal{F}_t^S\right] \\
&\text{let X be } \mathcal{F}_u\text{-measurable, then} \quad \mathbb{E}^Q\left[X \middle| \mathcal{F}_t^S\right] = \frac{1}{\zeta_t}\mathbb{E}\left[Z_uX \middle| \mathcal{F}_t^S\right]
\end{aligned}
$$

In particular $1/\zeta$ is a $\big(Q, \mathcal{F}^S\big)$-martinagale, i.e. $\mathbb{E}^Q\left[1/\zeta_u \middle| \mathcal{F}_t^S\right] = 1/\zeta_t$.

[Lak95, Thm. 6.6] and [Lak98] proofs:

Theorem A.2
Let $I(y)$ be the pseudo-inverse of $U'(x)$ for the utility function $U(x)$.
Assume $\mathbb{E}^Q\left[I(x\zeta_T)\right] < \infty\ \forall$ constant x
and y being determined by $\mathbb{E}\left[\tilde{\zeta}_T I(y\tilde{\zeta}_T)\right] = \mathbb{E}^Q\left[e^{-rT} I(ye^{-rT}\zeta_T)\right] = x_0$.
Then the optimal terminal wealth is $X_T^* = I\big(y\tilde{\zeta}_T\big)$ and the wealth process satisfies

$$x_0 + \int_0^T \tilde{X}_t\, \pi^{*\top} \sigma dW_t^{\mathbb{Q}} = \tilde{X}_t^* = \mathbb{E}^Q\left[\tilde{X}_T^* \middle| \mathcal{F}_t^S\right] = \frac{1}{\zeta_t}\mathbb{E}\left[\zeta_T \tilde{X}_T^* \middle| \mathcal{F}_t^S\right]$$

In particular [Lak98] show that ζ_t is explicitly given via $\hat{\mu}_t$

$$\zeta_t = \exp\left(-\int_0^t \hat{\theta}_u^\top dW_u^Q + \frac{1}{2}\int_0^t \hat{\theta}_u^\top \hat{\theta}_u du\right)$$
$$\text{where } \hat{\theta}_t := \sigma^{-1}(\hat{\mu}_t - r) = \mathbb{E}\left[\theta_t \middle| \mathcal{F}_t^S\right]$$

note: $d\big(\zeta_t^{-1}\big) = \zeta_t^{-1}\hat{\theta}_t^\top dW_t^Q$.

Example: power utility for $\alpha < 0$
$U(x) = \frac{1}{\alpha}x^\alpha$ and $I(y) = y^{\frac{1}{\alpha-1}}$
$x_0 = \mathbb{E}\left[\tilde{\zeta}_T I(y\tilde{\zeta}_T)\right] = y^{\frac{1}{\alpha-1}}\mathbb{E}\left[\tilde{\zeta}_T^{\frac{\alpha}{\alpha-1}}\right]$, therefore $y = x_0^{\alpha-1}\mathbb{E}\left[\tilde{\zeta}_T^{\frac{\alpha}{\alpha-1}}\right]^{1-\alpha}$
hence

$$X_T^* = I\big(y\tilde{\zeta}_T\big) = x_0 \mathbb{E}\left[\tilde{\zeta}_T^{\frac{\alpha}{\alpha-1}}\right]^{-1} \tilde{\zeta}_T^{\frac{1}{\alpha-1}}$$

Note:

$$\begin{aligned}
\mathbb{E}\left[\tilde{\zeta}_T^{\frac{\alpha}{\alpha-1}}\right] &= \mathbb{E}^Q\left[\zeta_T^{-1}\tilde{\zeta}_T^{\frac{\alpha}{\alpha-1}}\right] \\
&= \beta_T^{\frac{\alpha}{\alpha-1}} \cdot \mathbb{E}^Q\left[\exp\int_0^T \frac{1}{2}\frac{1}{\alpha-1}\hat{\theta}_s^\top\hat{\theta}_s ds - \int_0^T \frac{1}{\alpha-1}\hat{\theta}_s^\top dW_s^Q\right] \\
&= \beta_T^{\frac{\alpha}{\alpha-1}} \cdot \mathbb{E}^Q\left[\exp\frac{1}{\alpha-1}\int_0^T \frac{1}{2}\hat{\theta}_s^\top\hat{\theta}_s ds - \frac{1}{\alpha-1}\int_0^T \hat{\theta}_s^\top dW_s^Q \middle| \mathcal{F}_0^S\right]
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}^Q\left[\tilde{\zeta}_T^{\frac{1}{\alpha-1}} \middle| \mathcal{F}_t^S\right] &= \beta_T^{\frac{1}{\alpha-1}}\mathbb{E}^Q\left[\exp\int_0^T \frac{1}{2}\frac{1}{\alpha-1}\hat{\theta}_s^\top\hat{\theta}_s ds - \int_0^T \frac{1}{\alpha-1}\hat{\theta}_s^\top dW_s^Q \middle| \mathcal{F}_t^S\right] \\
&= \beta_T^{\frac{1}{\alpha-1}} \exp\frac{1}{\alpha-1}\int_0^t \frac{1}{2}\hat{\theta}_s^\top\hat{\theta}_s ds - \frac{1}{\alpha-1}\int_0^t \hat{\theta}_s^\top dW_s^Q \\
&\quad \cdot \mathbb{E}^Q\left[\exp\frac{1}{\alpha-1}\int_t^T \frac{1}{2}\hat{\theta}_s^\top\hat{\theta}_s ds - \frac{1}{\alpha-1}\int_t^T \hat{\theta}_s^\top dW_s^Q \middle| \mathcal{F}_t^S\right]
\end{aligned}$$

These expectation values get calculated via the same procedure as above from [CLMZ06].

Consider the orthogonal matrix P from the derivation of the filter that diagonalizes the covariance matrix of the prior of θ and the following definitions: $\tilde{\theta}_t := P\hat{\theta}_t, \tilde{W}_t^Q := PW_t^Q$. Then the latter is a Brownian motion with respect to Q and

$$\begin{aligned}
&\mathbb{E}^Q\left[\exp\frac{1}{\alpha-1}\int_t^T\frac{1}{2}\hat{\theta}_s^\top\hat{\theta}_s ds - \frac{1}{\alpha-1}\int_t^T\hat{\theta}_s^\top dW_s^Q\middle|\mathcal{F}_t^S\right]\\
=&\mathbb{E}^Q\left[\exp\frac{1}{\alpha-1}\int_t^T\frac{1}{2}\tilde{\theta}_s^\top\tilde{\theta}_s ds - \frac{1}{\alpha-1}\int_t^T\hat{\theta}_s^\top d\tilde{W}_s^Q\middle|\mathcal{F}_t^S\right]
\end{aligned}$$

and on the other hand we get via the filter representations from Remark 2.4:

$$\begin{aligned}
\tilde{\theta}_t &= \bar{D}_t\left(\tilde{W}_t^Q + D^{-1}\tilde{\theta}_0\right)\\
\text{with}\quad \tilde{\theta}_i(t) &= \delta_i(t)\tilde{W}_i^Q(t) + \frac{\delta_i(t)}{d_i}\tilde{\theta}_i(0)\\
\text{where } \delta_i(t) &= \frac{d_i}{1+d_i t} \text{ with } d\delta_i(t) = -\delta_i(t)^2 dt\\
\text{where } \operatorname{diag}(...d_i...) &= D = P\sigma^{-1}\Sigma_0\sigma^{-\top}P^\top
\end{aligned}$$

i.e. the $\left(\tilde{\theta}_i(t), \tilde{W}_i^Q(t)\right)$ are independent of $\left(\tilde{\theta}_j(t), \tilde{W}_j^Q(t)\right)$ and we get.

$$\begin{aligned}
&\mathbb{E}^Q\left[\exp\frac{1}{\alpha-1}\int_t^T\frac{1}{2}\tilde{\theta}_s^\top\tilde{\theta}_s ds - \frac{1}{\alpha-1}\int_t^T\hat{\theta}_s^\top d\tilde{W}_s^Q\middle|\mathcal{F}_t^S\right]\\
&=\prod_{i=1}^d\sqrt{\frac{\left(1+\delta_i(t)(T-t)\right)^{\frac{\alpha}{\alpha-1}}}{1+\delta_i(t)(T-t)\frac{\alpha}{\alpha-1}}}\exp\frac{\frac{1}{\alpha-1}\frac{\alpha}{\alpha-1}\tilde{\theta}_i^2(t)(T-t)}{2\left(1+\delta_i(t)(T-t)\frac{\alpha}{\alpha-1}\right)}\\
&=:\prod_{i=1}^d Y_i(t) =: \prod_{i=1}^d f_i(t)\exp c_i(t)\frac{\tilde{\theta}_i^2(t)}{2}
\end{aligned}$$

This leads to

$$\begin{aligned}
\tilde{X}_t^* &= \mathbb{E}^Q\left[\tilde{X}_T^*\middle|\mathcal{F}_t^S\right]\\
&= x_0\beta_T\mathbb{E}\left[\tilde{\zeta}_T^{\frac{\alpha}{\alpha-1}}\right]^{-1}\mathbb{E}^Q\left[\tilde{\zeta}_T^{\frac{1}{\alpha-1}}\middle|\mathcal{F}_t^S\right]\\
&= x_0\beta_T\beta_T^{\frac{-\alpha}{\alpha-1}}\prod_{i=1}^d f_i(0)^{-1}\exp -c_i(0)\frac{\tilde{\theta}_i^2(0)}{2}\\
&\quad\cdot\beta_T^{\frac{1}{\alpha-1}}\exp\frac{1}{\alpha-1}\int_0^t\frac{1}{2}\tilde{\theta}_s^\top\tilde{\theta}_s ds - \frac{1}{\alpha-1}\int_0^t\tilde{\theta}_s^\top d\tilde{W}_s^Q\cdot\prod_{i=1}^d f_i(t)\exp c_i(t)\frac{\tilde{\theta}_i^2(t)}{2}\\
&= x_0\prod_{i=1}^d f_i(0)^{-1}\exp -c_i(0)\frac{\tilde{\theta}_i^2(0)}{2}\cdot\exp\frac{1}{\alpha-1}\int_0^t\frac{1}{2}\tilde{\theta}_s^\top\tilde{\theta}_s ds
\end{aligned}$$

$$
\begin{aligned}
&\cdot \exp -\frac{1}{\alpha-1}\int_0^t \tilde{\theta}_s^\top d\tilde{W}_s^Q \cdot \prod_{i=1}^d f_i(t)\exp c_i(t)\frac{\tilde{\theta}_i^2(t)}{2} \\
&=: f(t)\cdot Z_t \cdot Y_t \\
\text{with } dZ_t &= (...)dt - Z_t\frac{1}{\alpha-1}\tilde{\theta}_t^\top d\tilde{W}_t^Q \\
\text{and } dY_t &= (...)dt + Y_t\sum_{i=1}^d c_i(t)\tilde{\theta}_i(t)d\tilde{\theta}_i(t) = (...)dt + Y_t\sum_{i=1}^d c_i(t)\delta_i(t)\tilde{\theta}_i(t)d\tilde{W}_i^Q(t)
\end{aligned}
$$

where (as above)

$$
\begin{aligned}
c_i(t)\delta_i(t) &= \frac{1}{1-\alpha}\left(\frac{(1-\alpha)(1+d_it)}{(1-\alpha)(1+d_it)-\alpha d_i(T-t)} - 1\right) \\
&=: \frac{1}{1-\alpha}\left(d_i'(t)-1\right) \\
&=: e_i'(t)
\end{aligned}
$$

such that

$$
\begin{aligned}
dZ_t &= (...)dt - Z_t\frac{1}{\alpha-1}\hat{\theta}_t^\top dW_t^Q \\
\text{and } dY_t &= (...)dt + Y_t\tilde{\theta}_t^\top E_t' d\tilde{W}_t^Q = (...)dt + Y_t\hat{\theta}_t^\top P^\top E_t' P dW_t^Q
\end{aligned}
$$

hence

$$
\begin{aligned}
d\tilde{X}_t &= (...)dt + \tilde{X}_t\left(-\frac{1}{\alpha-1}\hat{\theta}_t^\top + \hat{\theta}_t^\top P^\top E_t' P\right)dW_t^Q \\
\Rightarrow \quad \tilde{X}_t^*\, \pi_t^{*\,\top}\sigma dW_t^{\mathbb{Q}} &= (...)dt + \tilde{X}_t\frac{1}{1-\alpha}\hat{\theta}_t^\top P^\top D_t' P dW_t^Q \\
\Rightarrow \quad \pi_t^* &= \frac{1}{1-\alpha}\sigma^{-\top}P^\top D_t' P\hat{\theta}_t
\end{aligned}
$$

A.4.3 Another stochastic control approach

In the stochastic control approach in Section 4.1.2 we need to solve in particular the following ODE

$$
A'(t) = -\frac{\alpha}{1-\alpha}\left(A(t)^\top\gamma_t + Id\right)\Sigma^{-1}\left(\gamma_t A(t) + Id\right) - A(t)^\top\gamma_t\Sigma^{-1}\gamma_t A(t)
$$

With the ansatz $A(t) = \gamma_t^{-1}(B(t) - Id)$ (and boundary condition $B(T) = Id$) we get

$$
\gamma_t^{-1}B'(t) = -\frac{\alpha}{1-\alpha}B(t)^\top\Sigma^{-1}B(t) - B(t)^\top\Sigma^{-1}(B(t)-Id)
$$

And with $B(t) = \sigma C(t)\sigma^{-1}$ (and boundary condition $C(T) = Id$) we get

$$
\gamma_t^{-1}\sigma C'(t)\sigma^{-1} = -\frac{\alpha}{1-\alpha}\sigma^{-\top}C(t)^\top C(t)\sigma^{-1} - \sigma^{-\top}C(t)^\top(C(t)-Id)\sigma^{-1}
$$

$$\Rightarrow \qquad \gamma_t^{-1}\sigma C'(t) = -\frac{\alpha}{1-\alpha}\sigma^{-\top}C(t)^\top C(t) - \sigma^{-\top}C(t)^\top(C(t)-Id)$$

If $C(t) = P^\top D(t)P$ (where P is an orthogonal matrix and D is diagonal) we get

$$\begin{aligned}
\gamma_t^{-1}\sigma P^\top D'(t)P &= -\frac{\alpha}{1-\alpha}\sigma^{-\top}P^\top D^2(t)P - \sigma^{-\top}P^\top D(t)(D(t)-Id)P \\
\Rightarrow \qquad \gamma_t^{-1}\sigma P^\top D'(t) &= -\frac{\alpha}{1-\alpha}\sigma^{-\top}P^\top D^2(t) - \sigma^{-\top}P^\top D(t)(D(t)-Id) \\
&= -\frac{1}{1-\alpha}\sigma^{-\top}P^\top D^2(t) + \sigma^{-\top}P^\top D(t) \\
\Rightarrow \quad (\Sigma_0^{-1} + t\Sigma^{-1})\sigma P^\top D'(t) &= -\frac{1}{1-\alpha}\sigma^{-\top}P^\top D^2(t) + \sigma^{-\top}P^\top D(t) \\
\Rightarrow \quad P\sigma^\top(\Sigma_0^{-1} + t\Sigma^{-1})\sigma P^\top D'(t) &= -\frac{1}{1-\alpha}D^2(t) + D(t) \\
\Rightarrow \quad \left(P\sigma^\top\Sigma_0^{-1}\sigma P^\top + tId\right)D'(t) &= -\frac{1}{1-\alpha}D^2(t) + D(t)
\end{aligned}$$

Hence P needs to be chosen such that $P\sigma^\top\Sigma_0^{-1}\sigma P^\top$ is diagonal with entries p_i^{-1}. Then $D(t) = \operatorname{diag}(...\delta_i(t)...)$ where

$$d_i'(t) = \frac{1}{p_i^{-1}+t}\left(d_i(t) - \frac{1}{1-\alpha}d_i^2(t)\right) = \frac{p_i}{1+p_it}\left(d_i(t) - \frac{1}{1-\alpha}d_i^2(t)\right)$$

with boundary condition $d_i(T) = 1 \ \forall i$

Therefore we get the following solution:

$$D(t) = \operatorname{diag}\left(..., \frac{(1-\alpha)(1+p_it)}{(1-\alpha)(1+p_iT) - p_i(T-t)}, ...\right)$$

Proof. consider $f_i(t) := \frac{1}{d_i(t)}$. Then

$$f_i'(t) = -\frac{d_i'(t)}{d_i^2(t)} = -\frac{p_i}{1+p_it}f_i(t) + \frac{1}{1-\alpha}\frac{p_i}{1+p_it} \qquad \text{with } f_i'(T) = 1$$

variation of constants yields the homogeneous solution

$$f_i^H(t) = c \cdot \exp\int_0^t -\frac{p_i}{1+p_is}ds = c \cdot \frac{1}{1+p_it}$$

hence with $f_i(t) = c(t) \cdot \frac{1}{1+p_it}$ we get

$$\begin{aligned}
-\frac{p_i}{1+p_it}f_i(t) + \frac{1}{1-\alpha}\frac{p_i}{1+p_it} = f_i'(t) &= c'(t)\cdot\frac{1}{1+p_it} - c(t)\cdot\frac{p_i}{1+p_it}\frac{1}{1+p_it} \\
\Rightarrow \quad c'(t) &= \frac{1}{1-\alpha}p_i
\end{aligned}$$

with boundary condition $1 = f(T) = c(T)\frac{1}{1+p_i T} \;\Rightarrow\; c(T) = 1 + p_i T$. This leads to

$$\begin{aligned}
c(t) &= -\frac{1}{1-\alpha}p_i(T-t) + 1 + p_i T \\
\Rightarrow \quad f_i(t) &= -\frac{1}{1-\alpha}\frac{p_i}{1+p_i t}(T-t) + \frac{1+p_i T}{1+p_i t} \\
&= \frac{(1-\alpha)(1+p_i T) - p_i(T-t)}{(1-\alpha)(1+p_i t)} \\
\Rightarrow \quad d_i(t) &= \frac{(1-\alpha)(1+p_i t)}{(1-\alpha)(1+p_i T) - p_i(T-t)}
\end{aligned}$$

□

In particular we then get for the optimal trading strategy

$$\begin{aligned}
\pi^*(t,x,\mu) &= -\frac{V_x}{xV_{xx}}\Sigma^{-1}(\mu - r\mathbb{1}) - \frac{1}{xV_{xx}}\Sigma^{-1}\gamma_t V_{x\mu} \\
&= \frac{1}{1-\alpha}\Sigma^{-1}(\mu - r\mathbb{1}) + \frac{1}{f(1-\alpha)}\Sigma^{-1}\gamma_t f_\mu \\
&= \frac{1}{1-\alpha}\Sigma^{-1}(\mu - r\mathbb{1}) + \frac{1}{1-\alpha}\Sigma^{-1}\gamma_t A(t)(\mu - r) \\
&= \frac{1}{1-\alpha}\Sigma^{-1}B(t)(\mu - r) \\
&= \frac{1}{1-\alpha}\sigma^{-\top}P^\top D(t) P \sigma^{-1}(\mu - r)
\end{aligned}$$

Hence $\pi_t^* = \pi^*(t, X_t, \hat{\mu}_t) = \frac{1}{1-\alpha}\sigma^{-\top}P^\top D(t)P\sigma^{-1}(\hat{\mu}_t - r)$.

A.4.4 Choice of I in Case 3

Let I and J be the active and passive dimensions such that (λ, π) are the optimal solution in Case 3 and let (λ', π') be the solution if we chose I' and J' instead.
1) Let $i \in I$ with $\pi_i > 0$ be an active dimension.
If we choose $I' = I \setminus \{i\}$ and $J' = J \cup \{i\}$ then we get $\lambda'_i < 0$.
2) Let $j \in J$ with $\lambda_j > 0$ be a passive dimension (by Proposition 4.26).
If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ then we get $\pi'_j < 0$.
Hence several false choices of active and passive dimensions lead to non-admissible solutions, revealing that the choice was wrong.

Proof. From Proposition 4.26 we know for the optimal solution:

$$\begin{aligned}
0_I \le \pi_I &= \frac{1}{1-\alpha}(D(t)_{II})^{-1}(\mu - r\mathbb{1})_I \\
0_J \le \lambda_J &= -(\mu - r\mathbb{1})_J + D(t)_{JI}(D(t)_{II})^{-1}(\mu - r\mathbb{1})_I
\end{aligned}$$

and for the solution with respect to the active dimensions I':

$$\pi'_{I'} = \frac{1}{1-\alpha}(D(t)_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'}$$
$$\lambda'_{J'} = -(\mu - r\mathbb{1})_{J'} + D(t)_{J'I'}(D(t)_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'}$$

First consider $i \in I$ with $\pi_i > 0$ and $I' := I \setminus \{i\}$ and $J' = J \cup \{i\}$. Then

$$\begin{aligned}
0 < (1-\alpha)\pi_i &= ((D(t)_{II})^{-1})_{iI}(\mu - r\mathbb{1})_I \\
&= ((D(t)_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i - ((D(t)_{II})^{-1})_{iI'}(\mu - r\mathbb{1})_{I'} \\
&= ((D(t)_{II})^{-1})_{ii}(\mu - r\mathbb{1})_i - ((D(t)_{II})^{-1})_{ii}D(t)_{iI'}(D(t)_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
\Leftrightarrow \quad & 0 < (\mu - r\mathbb{1})_i - D(t)_{iI'}(D(t)_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} \\
\Rightarrow \quad \lambda'_i = (\lambda'_{J'})_i &= -(\mu - r\mathbb{1})_i + D(t)_{iI'}(D(t)_{I'I'})^{-1}(\mu - r\mathbb{1})_{I'} < 0
\end{aligned}$$

hence we have shown the first part.

For the second part consider $j \in J$ with $\lambda_j > 0$ and $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$. Then

$$0 < \lambda_j = -(\mu - r\mathbb{1})_j + D(t)_{jI}(D(t)_{II})^{-1}(\mu - r\mathbb{1})_I$$

On the other hand

$$\begin{aligned}
\pi'_j = (\pi'_{I'})_j &= \frac{1}{1-\alpha}((D(t)_{I'I'})^{-1})_{jI'}(\mu - r\mathbb{1})_{I'} \\
&= \frac{1}{1-\alpha}((D(t)_{I'I'})^{-1})_{jj}(\mu - r\mathbb{1})_j + \frac{1}{1-\alpha}((D(t)_{I'I'})^{-1})_{jI}(\mu - r\mathbb{1})_I \\
&= \frac{1}{1-\alpha}((D(t)_{I'I'})^{-1})_{jj}\left((\mu - r\mathbb{1})_j - D(t)_{jI}(D(t)_{II})^{-1}(\mu - r\mathbb{1})_I\right) \\
&= -\frac{1}{1-\alpha}((D(t)_{I'I'})^{-1})_{jj}\lambda_j \\
&< 0
\end{aligned}$$

hence we have shown the second part. □

A.4.5 Dimensions on the boundary in Case 3

If $j \in J$ is a passive dimensions that is almost invested in, than j can also be considered an active dimension.

In particular:
Let I and J be the active and passive dimensions such that (λ, π) are the optimal solution in Case 3 and let (λ', π') be the solution if we chose I' and J' instead.
Let $j \in J$ with $\pi_j = 0$ and $\lambda_j = 0$ be a passive dimension, that is almost invested in.
Then: If we choose $I' = I \cup \{j\}$ and $J' = J \setminus \{j\}$ we still get $\lambda' = \lambda$ and $\pi' = \pi$.

Proof. Note the matrix inversion formula (3.8) applied to $D(t)_{I'I'}$.

$$((D_{I'I'})^{-1})_{II} = (D_{II})^{-1} + (D_{II})^{-1}D_{Ij}((D_{I'I'})^{-1})_{jj}D_{jI}(D_{II})^{-1}$$

$$
\begin{aligned}
((D_{I'I'})^{-1})_{jI} &= -((D_{I'I'})^{-1})_{jj} D_{jI} (D_{II})^{-1} \\
((D_{I'I'})^{-1})_{Ij} &= -(D_{II})^{-1} D_{Ij} ((D_{I'I'})^{-1})_{jj} \\
((D_{I'I'})^{-1})_{jj} &= \left(D_{jj} - D_{jI}(D_{II})^{-1} D_{Ij}\right)^{-1}
\end{aligned}
$$

We know from Proposition 4.26

$$
\begin{aligned}
& \lambda_j = -(\mu - r\mathbb{1})_j + D(t)_{jI} (D(t)_{II})^{-1} (\mu - r\mathbb{1})_I \\
\Rightarrow \quad & \pi'_j = \frac{1}{1-\alpha} ((D(t)_{I'I'})^{-1})_{jI'} (\mu - r\mathbb{1})_{I'} \\
& = \frac{1}{1-\alpha} ((D(t)_{I'I'})^{-1})_{jI} (\mu - r\mathbb{1})_I + \frac{1}{1-\alpha} ((D(t)_{I'I'})^{-1})_{jj} (\mu - r\mathbb{1})_j \\
& = -\frac{1}{1-\alpha} ((D(t)_{I'I'})^{-1})_{jj} D(t)_{jI} (D(t)_{II})^{-1} (\mu - r\mathbb{1})_I + \frac{1}{1-\alpha} ((D(t)_{I'I'})^{-1})_{jj} (\mu - r\mathbb{1})_j \\
& = -\frac{1}{1-\alpha} ((D(t)_{I'I'})^{-1})_{jj} \lambda_j \\
\Rightarrow \quad & \Big(\quad 0 = \lambda_j \quad \Leftrightarrow \quad \pi'_j = 0 \quad \Big)
\end{aligned}
$$

Now consider $(1-\alpha)\pi_I = (D(t)_{II})^{-1}(\mu - r\mathbb{1})_I$ and

$$
\begin{aligned}
(1-\alpha)\pi'_I &= ((D(t)_{I'I'})^{-1})_{II'} (\mu - r\mathbb{1})_{I'} \\
&= ((D(t)_{I'I'})^{-1})_{II} (\mu - r\mathbb{1})_I + ((D(t)_{I'I'})^{-1})_{Ij} (\mu - r\mathbb{1})_j \\
&= (D(t)_{II})^{-1} (\mu - r\mathbb{1})_I \\
&\quad + (D(t)_{II})^{-1} D(t)_{Ij} ((D(t)_{I'I'})^{-1})_{jj} D(t)_{jI} (D(t)_{II})^{-1} (\mu - r\mathbb{1})_I \\
&\quad - (D(t)_{II})^{-1} D(t)_{Ij} ((D(t)_{I'I'})^{-1})_{jj} (\mu - r\mathbb{1})_j \\
&= (D(t)_{II})^{-1} (\mu - r\mathbb{1})_I + (D(t)_{II})^{-1} D(t)_{Ij} ((D(t)_{I'I'})^{-1})_{jj} \lambda_j \\
&= (D(t)_{II})^{-1} (\mu - r\mathbb{1})_I = (1-\alpha)\pi_I
\end{aligned}
$$

Hence $\pi'_I = \pi_I$ and hence $\pi' = \pi$. Also $\lambda' = \lambda$ since

$$
\begin{aligned}
\lambda_{J'} &= -(\mu - r\mathbb{1})_{J'} + D(t)_{J'I} (D(t)_{II})^{-1} (\mu - r\mathbb{1})_I \\
&= -(\mu - r\mathbb{1})_{J'} + (1-\alpha) D(t)_{J'I} \pi_I \\
\text{and} \quad \lambda'_{J'} &= -(\mu - r\mathbb{1})_{J'} + D(t)_{J'I'} (D(t)_{I'I'})^{-1} (\mu - r\mathbb{1})_{I'} \\
&= -(\mu - r\mathbb{1})_{J'} + (1-\alpha) D(t)_{J'I'} \pi'_{I'} \\
&= -(\mu - r\mathbb{1})_{J'} + (1-\alpha) D(t)_{J'I} \pi'_I = \lambda_{J'}
\end{aligned}
$$

□

A.4.6 The optimal solutions in Case 4

Let I and J be the (correct) choices for active and passive dimensions. To simplify notation we use following abbreviations:

$$\begin{aligned}
\mathbb{I}d^I &:= Id_I - g^I(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top \\
\mathbb{D}^I(t) &:= \Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}\mathbb{I}d^I \\
(\mu - r\mathbb{1})_I^{+\mathbb{I}} &:= (\mu - r\mathbb{1})_I + \alpha(T-t)g^I(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I
\end{aligned}$$

Then for $\lambda = \lambda^{(4,I)}(t,\mu)$ and $\pi = \pi^{(4,I)}(t,\mu)$ we get:

$$\begin{aligned}
\pi_I &= \frac{1}{1-\alpha}\mathbb{I}d^I\ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}}\ + g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\
\pi_J &= 0
\end{aligned}$$

with $\mathbb{1}_I^\top \pi_I = 1$ and

$$\begin{aligned}
\bar{\lambda} &= -g^I\ \mathbb{1}_I^\top\ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}}\ + g^I(1-\alpha) \\
\bar{\lambda}\mathbb{1}_I &= (1-\alpha)\Sigma_{II}\pi_I - \Sigma_{II}\ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
\lambda_J &= -(\mu - r\mathbb{1})_J + D(t)_{JI}(1-\alpha)\pi_I
\end{aligned}$$

Proof. The optimal strategy in Case 4 is given by

$$\pi = \frac{1}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)}\left(C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) + \alpha(T-t)g^I C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}\right) + g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}$$

Hence we get $\pi_J = 0_J$ and

$$\begin{aligned}
\pi_I = &\frac{1}{1-\alpha}\begin{pmatrix}(\Sigma_{II})^{-1}(Id - \mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1}g^I) & 0_{IJ}\end{pmatrix} \\
&\cdot\left(C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1}) + \alpha(T-t)g^I C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}\right) + (\Sigma_{II})^{-1}\mathbb{1}_I g^I \\
\Rightarrow\ \mathbb{1}_I^\top\pi_I = &\frac{1}{1-\alpha}\begin{pmatrix}0_I^\top & 0_J^\top\end{pmatrix}C^{(4,I)}(t)\left(\gamma_t^{-1}(\mu - r\mathbb{1}) + \alpha(T-t)g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}\right) + \mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I g^I = 1
\end{aligned}$$

Now define

$$\begin{aligned}
B(t) := C^{(4,I)}(t)\gamma_t^{-1} &= \left(Id - \frac{\alpha}{1-\alpha}(T-t)\gamma_t\Sigma_{\mathbb{J}}^{(4,I)}\right)^{-1} \\
b(t) := B(t)^{-1} &= \begin{pmatrix}Id_I - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}\bar{\Sigma}^I & 0_{IJ} \\ -\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{JI}\bar{\Sigma}^I & Id_J\end{pmatrix} \\
\Rightarrow\quad B(t)_{II} = (b^{-1})_{II} &= \left(b_{II} - b_{IJ}(b_{JJ})^{-1}b_{JI}\right)^{-1} \\
&= (b_{II})^{-1} \\
B(t)_{JJ} = (b^{-1})_{JJ} &= (b_{JJ})^{-1} + (b_{JJ})^{-1}b_{JI}(b^{-1})_{II}b_{IJ}(b_{JJ})^{-1}
\end{aligned}$$

$$
\begin{aligned}
& & &= Id_J \\
B(t)_{JI} &= (b^{-1})_{JI} & &= -(b_{JJ})^{-1} b_{JI} (b^{-1})_{II} \\
& & &= -b_{JI}(b_{II})^{-1} \\
B(t)_{IJ} &= (b^{-1})_{IJ} & &= -(b^{-1})_{II} b_{IJ} (b_{JJ})^{-1} \\
& & &= 0_{IJ}
\end{aligned}
$$

hence we get

$$
\begin{aligned}
\pi_I &= \frac{1}{1-\alpha}\Big(\big(Id_I - g^I(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top\big)(\Sigma_{II})^{-1} \quad 0_{IJ}\Big) \\
&\qquad \cdot \big(B(t)(\mu - r\mathbb{1}) + \alpha(T-t)g^I C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}\big) + (\Sigma_{II})^{-1}\mathbb{1}_I g^I \\
&= \frac{1}{1-\alpha}\mathbb{I}d^I(\Sigma_{II})^{-1}\Big(B(t)_{II}(\mu - r\mathbb{1})_I + \alpha(T-t)g^I B(t)_{II}(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I\Big) + g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\
&= \frac{1}{1-\alpha}\mathbb{I}d^I(\Sigma_{II})^{-1}\left(Id_I - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}\mathbb{I}d^I(\Sigma_{II})^{-1}\right)^{-1} \\
&\qquad \cdot \Big((\mu - r\mathbb{1})_I + \alpha(T-t)g^I(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I\Big) + g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\
&= \frac{1}{1-\alpha}\mathbb{I}d^I \ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} + g^I(\Sigma_{II})^{-1}\mathbb{1}_I
\end{aligned}
$$

For the optimal dual process λ we get

$$
\begin{aligned}
\lambda = \begin{pmatrix}\bar{\lambda}\mathbb{1}_I \\ \lambda_J\end{pmatrix} &= \mathbb{J}^{(4,I)}B(t)(\mu - r\mathbb{1}) + (1-\alpha)g^I(\mathbb{J}^I + Id)\mathbb{1} \\
&\quad + \alpha(T-t)g^I\mathbb{J}^{(4,I)}C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} \\
\text{where} \quad \mathbb{J}^{(4,I)} &= \begin{pmatrix}-g^I\mathbb{1}_I\mathbb{1}_I^\top(\Sigma_{II})^{-1} & 0_{IJ} \\ \Sigma_{JI}\mathbb{I}d^I(\Sigma_{II})^{-1} & -Id_J\end{pmatrix}
\end{aligned}
$$

and

$$
\begin{aligned}
\bar{\lambda} &= g^I(1-\alpha) - g^I\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}^\top B(t)(\mu - r\mathbb{1}) \\
&\quad - \alpha(T-t)(g^I)^2\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix}^\top C^{(4,I)}(t)\begin{pmatrix}(\Sigma_{II})^{-1}\mathbb{1}_I \\ 0_J\end{pmatrix} \\
&= g^I(1-\alpha) - g^I\mathbb{1}_I^\top(\Sigma_{II})^{-1}B(t)_{II}(\mu - r\mathbb{1})_I \\
&\quad - \alpha(T-t)(g^I)^2\mathbb{1}_I^\top(\Sigma_{II})^{-1}C^{(4,I)}(t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I \\
&= g^I(1-\alpha) - g^I\mathbb{1}_I^\top(\Sigma_{II})^{-1}B(t)_{II}(\mu - r\mathbb{1})_I \\
&\quad - \alpha(T-t)(g^I)^2\mathbb{1}_I^\top(\Sigma_{II})^{-1}B(t)_{II}(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I \\
&= g^I(1-\alpha) - g^I\mathbb{1}_I^\top(\Sigma_{II})^{-1}B(t)_{II}\Big((\mu - r\mathbb{1})_I + \alpha(T-t)g^I(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I\Big)
\end{aligned}
$$

$$= g^I(1-\alpha) - g^I \mathbb{1}_I^\top \left(\Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(Id_I - g^I(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top)\right)^{-1} (\mu - r\mathbb{1})_I^{+\mathbb{I}}$$
$$= g^I(1-\alpha) - g^I \mathbb{1}_I^\top \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}}$$

Hence

$$\begin{aligned}(\Sigma_{II})^{-1}\bar{\lambda}\mathbb{1}_I &= (1-\alpha)g^I(\Sigma_{II})^{-1}\mathbb{1}_I - (\Sigma_{II})^{-1}g^I\mathbb{1}_I\mathbb{1}_I^\top\mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\ &= (1-\alpha)g^I(\Sigma_{II})^{-1}\mathbb{1}_I + \mathbb{I}d^I\mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} - \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\ &= (1-\alpha)\pi_I - \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}}\end{aligned}$$

and

$$\begin{aligned}\lambda_J = &\ \mathbb{J}_{JI}^{(4,I)}B(t)_{II}(\mu - r\mathbb{1})_I - Id_J B(t)_{JI}(\mu - r\mathbb{1})_I \\ &- Id_J B(t)_{JJ}(\mu - r\mathbb{1})_J + (1-\alpha)\Sigma_{JI}g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\ &+ \alpha(T-t)\Sigma_{JI}\mathbb{I}d^I(\Sigma_{II})^{-1}C^{(4,I)}(t)_{II}g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\ &- \alpha(T-t)Id_J C^{(4,I)}(t)_{JI}g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\ = &-(\mu - r\mathbb{1})_J \\ &+ \Sigma_{JI}\mathbb{I}d^I(\Sigma_{II})^{-1}(b(t)_{II})^{-1}(\mu - r\mathbb{1})_I \\ &+ b(t)_{JI}(b(t)_{II})^{-1}(\mu - r\mathbb{1})_I + (1-\alpha)\Sigma_{JI}g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\ &+ \alpha(T-t)g^I\Sigma_{JI}\mathbb{I}d^I(\Sigma_{II})^{-1}B(t)_{II}(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I \\ &- \alpha(T-t)g^I B(t)_{JI}(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I - \alpha(T-t)g^I B(t)_{JJ}(\gamma_t)_{JI}(\Sigma_{II})^{-1}\mathbb{1}_I \\ = &-(\mu - r\mathbb{1})_J \\ &+ \Sigma_{JI}\mathbb{I}d^I(\Sigma_{II})^{-1}(b(t)_{II})^{-1}(\mu - r\mathbb{1})_I \\ &+ b(t)_{JI}\Sigma_{II}(\Sigma_{II})^{-1}(b(t)_{II})^{-1}(\mu - r\mathbb{1})_I \\ &+ \alpha(T-t)g^I\Sigma_{JI}\mathbb{I}d^I(\Sigma_{II})^{-1}(b(t)_{II})^{-1}(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I \\ &+ \alpha(T-t)g^I b(t)_{JI}\Sigma_{II}(\Sigma_{II})^{-1}(b(t)_{II})^{-1}(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I \\ &- \alpha(T-t)(\gamma_t)_{JI}g^I(\Sigma_{II})^{-1}\mathbb{1}_I + (1-\alpha)\Sigma_{JI}g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\ = &-(\mu - r\mathbb{1})_J - \alpha(T-t)(\gamma_t)_{JI}g^I(\Sigma_{II})^{-1}\mathbb{1}_I + (1-\alpha)\Sigma_{JI}g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\ &+ \left(\Sigma_{JI}\mathbb{I}d^I + b(t)_{JI}\Sigma_{II}\right)(\Sigma_{II})^{-1}(b(t)_{II})^{-1} \\ &\quad \cdot \left((\mu - r\mathbb{1})_I + \alpha(T-t)g^I(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I\right) \\ = &-(\mu - r\mathbb{1})_J + (1-\alpha)\left(\Sigma_{JI} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{JI}\right)g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\ &+ \left(\Sigma_{JI}\mathbb{I}d^I - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{JI}\mathbb{I}d^I\right)\left(\Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}\mathbb{I}d^I\right)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\ = &-(\mu - r\mathbb{1})_J + (1-\alpha)D(t)_{JI}g^I(\Sigma_{II})^{-1}\mathbb{1}_I + D(t)_{JI}\mathbb{I}d^I\ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\ = &-(\mu - r\mathbb{1})_J + D(t)_{JI}(1-\alpha)\pi_I\end{aligned}$$

□

A.4.7 Shifted unconstrained solution

The optimal constrained solution π^λ in all cases is equal to the optimal unconstrained solution with $\mu + \lambda$ plugged in instead of μ, i.e.

$$\pi^\lambda = \pi_{\text{unc}}^{+\lambda} := \frac{1}{1-\alpha}\Sigma^{-1}C(t)\gamma_t^{-1}(\mu - r\mathbb{1} + \lambda) = \frac{1}{1-\alpha}D(t)^{-1}(\mu - r\mathbb{1} + \lambda)$$

In particular all solutions in the domains $\mathbb{D}^{(\cdot,\cdot)}$ can be extended continuously to the boundary of their respective domain and coincide at the boundary points, such that π^λ is continuous in μ.

Proof. In Case 1 with $\lambda = 0$ and in Case 2 with $\lambda = -(\mu - r\mathbb{1})$ this is trivial.
In Case $(3, I)$ we have

$$\begin{aligned}\pi^\lambda &= \frac{1}{1-\alpha}\Sigma_{\mathbb{J}}^{(3,I)}C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1})\\ \lambda &= \mathbb{J}^I C^{(3,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1})\end{aligned}$$

By Proposition 4.26 we get

$$\begin{aligned}\pi_I &= \frac{1}{1-\alpha}\left(D(t)_{II}\right)^{-1}(\mu - r\mathbb{1})_I \qquad \text{and} \qquad \pi_J = 0\\ \lambda_J &= -(\mu - r\mathbb{1})_J + D(t)_{JI}(1-\alpha)\pi_I \quad \text{and} \qquad \lambda_I = 0\\ \Rightarrow \quad (\mu - r\mathbb{1} + \lambda)_I &= (\mu - r\mathbb{1})_I\\ (\mu - r\mathbb{1} + \lambda)_J &= (D(t)_{JI})(D(t)_{II})^{-1}(\mu - r\mathbb{1})_I\end{aligned}$$

Hence the updated unconstrained solution satisfies:

$$\begin{aligned}(\pi_{\text{unc}}^{+\lambda})_J &= \frac{1}{1-\alpha}(D(t)^{-1})_{JI}(\mu - r\mathbb{1} + \lambda)_I + \frac{1}{1-\alpha}(D(t)^{-1})_{JJ}(\mu - r\mathbb{1} + \lambda)_J\\ &= \frac{1}{1-\alpha}(D(t)^{-1})_{JI}(\mu - r\mathbb{1})_I + \frac{1}{1-\alpha}(D(t)^{-1})_{JJ}(D(t)_{JI})(D(t)_{II})^{-1}(\mu - r\mathbb{1})_I\\ &= 0_J\\ (\pi_{\text{unc}}^{+\lambda})_I &= \frac{1}{1-\alpha}(D(t)^{-1})_{II}(\mu - r\mathbb{1} + \lambda)_I + \frac{1}{1-\alpha}(D(t)^{-1})_{IJ}(\mu - r\mathbb{1} + \lambda)_J\\ &= \frac{1}{1-\alpha}(D(t)^{-1})_{II}(\mu - r\mathbb{1})_I + \frac{1}{1-\alpha}(D(t)^{-1})_{IJ}(D(t)_{JI})(D(t)_{II})^{-1}(\mu - r\mathbb{1})_I\\ &= \frac{1}{1-\alpha}(D(t)_{II})^{-1}(\mu - r\mathbb{1})_I\\ &= \pi_I^\lambda\\ \Rightarrow \quad \pi_{\text{unc}}^{+\lambda} &= \pi^\lambda\end{aligned}$$

In Case $(4, I)$ we have

$$\pi^\lambda = \frac{1}{1-\alpha}\Sigma_{\mathbb{J}}^{(4,I)}C^{(4,I)}(t)\gamma_t^{-1}(\mu - r\mathbb{1})$$

$$
\begin{aligned}
&+ \frac{\alpha}{1-\alpha}(T-t) g^I \Sigma_{\mathbb{J}}^{(4,I)} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1} \mathbb{1}_I \\ 0_J \end{pmatrix} + g^I \begin{pmatrix} (\Sigma_{II})^{-1} \mathbb{1}_I \\ 0_J \end{pmatrix} \\
&= \frac{1}{1-\alpha} \Sigma_{\mathbb{J}}^{(4,I)} C^{(4,I)}(t) \left(\gamma_t^{-1}(\mu - r\mathbb{1}) + \alpha(T-t) g^I \begin{pmatrix} (\Sigma_{II})^{-1} \mathbb{1}_I \\ 0_J \end{pmatrix} \right) + g^I \begin{pmatrix} (\Sigma_{II})^{-1} \mathbb{1}_I \\ 0_J \end{pmatrix} \\
\lambda &= \mathbb{J}^{(4,I)} C^{(4,I)}(t) \gamma_t^{-1}(\mu - r\mathbb{1}) + (1-\alpha) g^I (\mathbb{J}^I + Id)\mathbb{1} \\
&+ \alpha(T-t) g^I \mathbb{J}^{(4,I)} C^{(4,I)}(t) \begin{pmatrix} (\Sigma_{II})^{-1} \mathbb{1}_I \\ 0_J \end{pmatrix}
\end{aligned}
$$

where $\mathbb{J}^{(4,I)} = \begin{pmatrix} -\mathbb{1}_I \mathbb{1}_I^\top (\Sigma_{II})^{-1} g^I & 0_{IJ} \\ \Sigma_{JI}(\Sigma_{II})^{-1}(Id - \mathbb{1}_I \mathbb{1}_I^\top (\Sigma_{II})^{-1} g^I) & -Id_J \end{pmatrix}$

By Proposition 4.26 we get

$$
\begin{aligned}
\pi_J^\lambda &= 0 \\
\pi_I^\lambda &= \frac{1}{1-\alpha}(Id_I - g^I (\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top) \left(\Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(Id_I - g^I(\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top) \right)^{-1} \\
&\quad \cdot \left((\mu - r\mathbb{1})_I + \alpha(T-t)(\gamma_t)_{II} g^I (\Sigma_{II})^{-1} \mathbb{1}_I \right) + g^I (\Sigma_{II})^{-1} \mathbb{1}_I \\
\bar{\lambda} &= g^I(1-\alpha) - g^I \mathbb{1}_I^\top \left(\Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(Id_I - g^I(\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top) \right)^{-1} \\
&\quad \cdot \left((\mu - r\mathbb{1})_I + \alpha(T-t)(\gamma_t)_{II} g^I (\Sigma_{II})^{-1} \mathbb{1}_I \right) \\
\bar{\lambda}\mathbb{1} &= (1-\alpha)\Sigma_{II}\pi_I - \Sigma_{II} \left(\Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(Id_I - g^I(\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top) \right)^{-1} \\
&\quad \cdot \left((\mu - r\mathbb{1})_I + \alpha(T-t)(\gamma_t)_{II} g^I (\Sigma_{II})^{-1} \mathbb{1}_I \right) \\
\lambda_J &= -(\mu - r\mathbb{1})_J - \alpha(T-t)(\gamma_t)_{JI} g^I (\Sigma_{II})^{-1} \mathbb{1}_I + (1-\alpha)\Sigma_{JI} g^I (\Sigma_{II})^{-1} \mathbb{1}_I \\
&\quad + \left(\Sigma_{JI} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{JI} \right) \left(Id - g^I (\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top \right) \\
&\quad \cdot \left(\Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(Id_I - g^I(\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top) \right)^{-1} \\
&\quad \cdot \left((\mu - r\mathbb{1})_I + \alpha(T-t)(\gamma_t)_{II} g^I (\Sigma_{II})^{-1} \mathbb{1}_I \right) \\
&= -(\mu - r\mathbb{1})_J + \left(\Sigma_{JI} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{JI} \right)(1-\alpha)\pi_I^\lambda
\end{aligned}
$$

abbreviations:

$$
\begin{aligned}
\mathbb{I}d^I &:= Id_I - g^I (\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top \\
\mathbb{D}^I(t) &:= \Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II} \mathbb{I}d^I \\
(\mu - r\mathbb{1})_I^{+\mathbb{I}} &:= (\mu - r\mathbb{1})_I + \alpha(T-t) g^I (\gamma_t)_{II} (\Sigma_{II})^{-1} \mathbb{1}_I
\end{aligned}
$$

Then with $\mathbb{1}_I^\top \pi_I^\lambda = 1$

$$
\pi_I^\lambda = \frac{1}{1-\alpha} \mathbb{I}d^I \; \mathbb{D}^I(t)^{-1} (\mu - r\mathbb{1})_I^{+\mathbb{I}} \; + g^I (\Sigma_{II})^{-1} \mathbb{1}_I
$$

$$\pi_J^\lambda = 0$$

and

$$\begin{aligned}
\bar{\lambda} &= -g^I \ \mathbb{1}_I^\top \ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \ + g^I(1-\alpha) \\
\bar{\lambda}\mathbb{1}_I &= (1-\alpha)\Sigma_{II}\pi_I^\lambda - \Sigma_{II} \ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
\lambda_J &= -(\mu - r\mathbb{1})_J + D(t)_{JI}(1-\alpha)\pi_I^\lambda
\end{aligned}$$

Hence we get for $\mu - r\mathbb{1} + \lambda$:

$$\begin{aligned}
(\mu - r\mathbb{1} + \lambda)_J &= D(t)_{JI}(1-\alpha)\pi_I^\lambda \\
(\mu - r\mathbb{1} + \lambda)_I &= (\mu - r\mathbb{1})_I + \bar{\lambda}\mathbb{1}_I \\
&= (\mu - r\mathbb{1})_I + (1-\alpha)\Sigma_{II}\pi_I^\lambda - \Sigma_{II} \ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}}
\end{aligned}$$

Hence the updated unconstrained solution satisfies:

$$\begin{aligned}
(1-\alpha)(\pi_{\text{unc}}^{+\lambda})_J &= (D(t)^{-1})_{JI}(\mu - r\mathbb{1} + \lambda)_I + (D(t)^{-1})_{JJ}(\mu - r\mathbb{1} + \lambda)_J \\
&= (D(t)^{-1})_{JI}(\mu - r\mathbb{1})_I + (D(t)^{-1})_{JI}\bar{\lambda}\mathbb{1}_I + (D(t)^{-1})_{JJ}D(t)_{JI}(1-\alpha)\pi_I^\lambda \\
&= (D(t)^{-1})_{JI}(\mu - r\mathbb{1})_I + (D(t)^{-1})_{JI}\Sigma_{II}(1-\alpha)\pi_I^\lambda - (D(t)^{-1})_{JI}D(t)_{II}(1-\alpha)\pi_I^\lambda \\
&\quad - (D(t)^{-1})_{JI}\Sigma_{II} \ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&= (D(t)^{-1})_{JI}(\mu - r\mathbb{1})_I + (D(t)^{-1})_{JI}\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(1-\alpha)\pi_I^\lambda \\
&\quad - (D(t)^{-1})_{JI}\Sigma_{II} \ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&= (D(t)^{-1})_{JI}(\mu - r\mathbb{1})_I \\
&\quad + (D(t)^{-1})_{JI}\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}\mathbb{I}d^I \ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&\quad + (D(t)^{-1})_{JI}\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(1-\alpha)g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\
&\quad - (D(t)^{-1})_{JI}\Sigma_{II} \ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&= (D(t)^{-1})_{JI}(\mu - r\mathbb{1})_I \\
&\quad + (D(t)^{-1})_{JI}\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(1-\alpha)g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\
&\quad - (D(t)^{-1})_{JI}\left(\Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}\mathbb{I}d^I\right)\mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&= (D(t)^{-1})_{JI}(\mu - r\mathbb{1})_I \\
&\quad + (D(t)^{-1})_{JI}\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(1-\alpha)g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\
&\quad - (D(t)^{-1})_{JI}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&= 0_J
\end{aligned}$$

and

$$(1-\alpha)(\pi_{\text{unc}}^{+\lambda})_I = (D(t)^{-1})_{II}(\mu - r\mathbb{1} + \lambda)_I + (D(t)^{-1})_{IJ}(\mu - r\mathbb{1} + \lambda)_J$$

$$
\begin{aligned}
&= (D(t)^{-1})_{II}(\mu - r\mathbb{1})_I + (D(t)^{-1})_{II}\lambda\mathbb{1}_I + (D(t)^{-1})_{IJ}D(t)_{JI}(1-\alpha)\pi_I^\lambda \\
&= (D(t)^{-1})_{II}(\mu - r\mathbb{1})_I \\
&\quad + (D(t)^{-1})_{II}(1-\alpha)\Sigma_{II}\pi_I^\lambda - (D(t)^{-1})_{II}\Sigma_{II}\mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&\quad + ((D(t)_{II})^{-1} - (D(t)^{-1})_{II})D(t)_{II}(1-\alpha)\pi_I^\lambda \\
&= (D(t)^{-1})_{II}(\mu - r\mathbb{1})_I + (1-\alpha)\pi_I^\lambda \\
&\quad - (D(t)^{-1})_{II}\Sigma_{II}\mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&\quad + (D(t)^{-1})_{II}\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(1-\alpha)\pi_I^\lambda \\
&= (D(t)^{-1})_{II}(\mu - r\mathbb{1})_I + (1-\alpha)\pi_I^\lambda \\
&\quad - (D(t)^{-1})_{II}\Sigma_{II}\mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&\quad + (D(t)^{-1})_{II}\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}\mathbb{I}d^I\ \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&\quad + (D(t)^{-1})_{II}\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(1-\alpha)g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\
&= (D(t)^{-1})_{II}(\mu - r\mathbb{1})_I + (1-\alpha)\pi_I^\lambda \\
&\quad - (D(t)^{-1})_{II}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&\quad + (D(t)^{-1})_{II}\frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}(1-\alpha)g^I(\Sigma_{II})^{-1}\mathbb{1}_I \\
&= (1-\alpha)\pi_I^\lambda
\end{aligned}
$$

hence $\pi_{\text{unc}}^{+\lambda} = \pi^\lambda$. □

A.4.8 Exact shifting in Case 4

Let $\rho \in \partial K$ be arbitrary with $\mathbb{1}_I^\top \rho_I = 1$ and

$$
\begin{aligned}
\rho_I &> 0_I \quad \text{where } \emptyset \neq I \subset \{1, ..., d\} \\
\rho_J &= 0_J \quad \text{where } J = \{1, ..., d\} \setminus I
\end{aligned}
$$

and consider the parameter μ' on the boundary of $\mathbb{D}^{(1,\cdot)}$

$$\mu' := r\mathbb{1} + (1-\alpha)D(t) \cdot \rho \quad \in \partial\mathbb{D}^{(1,\cdot)}$$

Let $\mu := \mu' + \begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix}$ be an arbitrary input parameter for any $c_J \leq c\mathbb{1}_J$ and $c > 0$. Then the optimal dual solution for the constrained portfolio optimization problem with input parameter μ is in Case $(4, I)$ with $\lambda = -\begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix}$ and hence $\mu + \lambda = \mu' \in \partial\mathbb{D}^{(1,\cdot)}$.

Proof. We get the optimal dual solution with Algorithm 4.34:
Case 1 does not fit:

$$(1-\alpha)\pi_{\text{unc}} = \Sigma^{-1}C(t)\gamma_t^{-1}(\mu - r\mathbb{1})$$

$$
\begin{aligned}
&= D(t)^{-1}\left(\mu' - r\mathbb{1} + \begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix}\right) \\
&= (1-\alpha)\rho + D(t)^{-1}\begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix} \\
\Rightarrow \quad \begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix}^\top \pi_{\text{unc}} &= \begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix}^\top \rho + \frac{1}{1-\alpha}\begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix}^\top D(t)^{-1}\begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix} \\
&= c + \frac{1}{1-\alpha}\begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix}^\top D(t)^{-1}\begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix} \\
&> c \\
\Rightarrow \quad 1 &< \mathbb{1}_I^\top(\pi_{\text{unc}})_I + \frac{1}{c}c_J^\top(\pi_{\text{unc}})_J
\end{aligned}
$$

For Case 1 to fit, we need $\pi_{\text{unc}} \geq 0$ and $\mathbb{1}^\top\pi_{\text{unc}} \leq 1$. Hence

$$
\begin{aligned}
\mathbb{1}_I^\top(\pi_{\text{unc}})_I + \mathbb{1}_J^\top(\pi_{\text{unc}})_J \leq 1 &< \mathbb{1}_I^\top(\pi_{\text{unc}})_I + \frac{1}{c}c_J^\top(\pi_{\text{unc}})_J \\
\Rightarrow \qquad \mathbb{1}_J^\top(\pi_{\text{unc}})_J &< \frac{1}{c}c_J^\top(\pi_{\text{unc}})_J \\
\Rightarrow \qquad (c\mathbb{1}_J - c_J)^\top(\pi_{\text{unc}})_J &< 0
\end{aligned}
$$

But $c\mathbb{1}_J - c_J \geq 0$ and $\pi_{\text{unc}} \geq 0$, hence Case 1 doesn't fit.

Also Case 2 does not fit:

$$
\begin{aligned}
\mu - r\mathbb{1} = \mu' - r\mathbb{1} + \begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix} &= (1-\alpha)D(t)\cdot\rho + \begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix} \\
\Rightarrow \quad (\mu - r\mathbb{1})_I &= (1-\alpha)D(t)_{II}\rho_I + c\mathbb{1}_I
\end{aligned}
$$

Again $D(t)_{II}$ is positive definite and $\rho_I \neq 0$, hence

$$
\rho_I^\top(\mu - r\mathbb{1})_I = \rho_I^\top(1-\alpha)D(t)_{II}\rho_I + c\rho_I^\top\mathbb{1}_I > 1
$$

But with $\rho_I > 0_I$ and $(\mu - r\mathbb{1})_I < 0_I$ this cannot work, hence $(\mu - r\mathbb{1})_I \nless 0_I$ hence Case 2 does not fit.

Also Case 3 does not fit:
Assume Case 3 would work. Then there was some $\lambda \geq 0$ such that the optimal constrained solution satisfies $\pi^\lambda \in K$ and $\lambda^\top\pi^\lambda = 0$. But by Lemma 4.39 π^λ is equal to the unconstrained solution $\pi_{\text{unc}}^{+\lambda}$ that starts in $\mu - r\mathbb{1} + \lambda$, i.e.

$$
\begin{aligned}
(1-\alpha)\pi^\lambda &= (1-\alpha)\pi_{\text{unc}}^{+\lambda} \\
&= D(t)^{-1}(\mu - r\mathbb{1} + \lambda) \\
&= D(t)^{-1}(\mu' - r\mathbb{1} + \begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix} + \lambda) \\
&= (1-\alpha)\rho + D(t)^{-1}\begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix} + D(t)^{-1}\lambda
\end{aligned}
$$

$$\begin{aligned}
\Rightarrow\quad \begin{pmatrix} c\mathbb{1}_I \\ c_J \end{pmatrix}^\top \pi^\lambda &= \begin{pmatrix} c\mathbb{1}_I + \lambda_I \\ c_J + \lambda_J \end{pmatrix}^\top \pi^\lambda \\
&= \begin{pmatrix} c\mathbb{1}_I + \lambda_I \\ c_J + \lambda_J \end{pmatrix}^\top \rho + \frac{1}{1-\alpha} \begin{pmatrix} c\mathbb{1}_I + \lambda_I \\ c_J + \lambda_J \end{pmatrix}^\top D(t)^{-1} \begin{pmatrix} c\mathbb{1}_I + \lambda_I \\ c_J + \lambda_J \end{pmatrix} \\
&= c + \lambda_I^\top \rho_I + \frac{1}{1-\alpha} \begin{pmatrix} c\mathbb{1}_I + \lambda_I \\ c_J + \lambda_J \end{pmatrix}^\top D(t)^{-1} \begin{pmatrix} c\mathbb{1}_I + \lambda_I \\ c_J + \lambda_J \end{pmatrix} \\
&> c + \lambda_I^\top \rho_I \geq c \\
\Rightarrow\quad 1 &< \mathbb{1}_I^\top \pi_I^\lambda + \frac{1}{c} c_J^\top \pi_J^\lambda
\end{aligned}$$

For Case 3 to fit, we need $\pi^\lambda \geq 0$ and $\mathbb{1}^\top \pi^\lambda \leq 1$. Hence

$$\begin{aligned}
\mathbb{1}_I^\top \pi_I^\lambda + \mathbb{1}_J^\top \pi_J^\lambda \leq 1 &< \mathbb{1}_I^\top \pi_I^\lambda + \frac{1}{c} c_J^\top \pi_J^\lambda \\
\Rightarrow\quad \mathbb{1}_J^\top \pi_J^\lambda &< \frac{1}{c} c_J^\top \pi_J^\lambda \\
\Rightarrow\quad (c\mathbb{1}_J - c_J)^\top \pi_J^\lambda &< 0
\end{aligned}$$

But $c\mathbb{1}_J - c_J \geq 0$ and $\pi_{\text{unc}} \geq 0$, hence Case 3 doesn't fit.

Now we consider Case $(4, I)$ and get by Proposition 4.26:

$$\begin{aligned}
\pi_J^\lambda &= 0_J \\
\pi_I^\lambda &= \frac{1}{1-\alpha} \mathbb{I}d^I \; \mathbb{D}^I(t)^{-1} (\mu - r\mathbb{1})_I^{+\mathbb{I}} \; + g^I (\Sigma_{II})^{-1} \mathbb{1}_I \\
&= \mathbb{I}d^I \left((1-\alpha)\Sigma_{II} - \alpha(T-t)(\gamma_t)_{II} \mathbb{I}d^I\right)^{-1} \\
&\qquad \cdot \Big((\mu - r\mathbb{1})_I + \alpha(T-t) g^I (\gamma_t)_{II} (\Sigma_{II})^{-1} \mathbb{1}_I\Big) + g^I (\Sigma_{II})^{-1} \mathbb{1}_I \\
&= \mathbb{I}d^I \Big((1-\alpha)\Sigma_{II} - \alpha(T-t)(\gamma_t)_{II} (Id_I - g^I (\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top)\Big)^{-1} \\
&\qquad \cdot \Big((\mu' - r\mathbb{1})_I + c\mathbb{1}_I + \alpha(T-t) g^I (\gamma_t)_{II} (\Sigma_{II})^{-1} \mathbb{1}_I\Big) + g^I (\Sigma_{II})^{-1} \mathbb{1}_I \\
&= \mathbb{I}d^I \Big((1-\alpha) D(t)_{II} + \alpha(T-t)(\gamma_t)_{II} g^I (\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top\Big)^{-1} \\
&\qquad \cdot \Big((1-\alpha) D(t)_{II} \rho_I + c\mathbb{1}_I + \alpha(T-t) g^I (\gamma_t)_{II} (\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top \rho_I\Big) + g^I (\Sigma_{II})^{-1} \mathbb{1}_I \\
&= \mathbb{I}d^I \rho_I + g^I (\Sigma_{II})^{-1} \mathbb{1}_I \\
&\quad + \mathbb{I}d^I \left((1-\alpha)\Sigma_{II} - \alpha(T-t)(\gamma_t)_{II} \mathbb{I}d^I\right)^{-1} c\mathbb{1}_I \qquad (*) \\
&= (Id_I - g^I (\Sigma_{II})^{-1} \mathbb{1}_I \mathbb{1}_I^\top) \rho_I + g^I (\Sigma_{II})^{-1} \mathbb{1}_I \\
&= \rho_I > 0_I
\end{aligned}$$

Here line $(*)$ is equal to zero by the matrix inversion lemma (3.8) in the following form:

$$(A + BC)^{-1} = A^{-1} + A^{-1} B (Id - CA^{-1} B)^{-1} C A^{-1}$$

$$\begin{aligned}
&\text{with} \quad A = (1-\alpha)\Sigma_{II} \quad , \quad B = -\alpha(T-t)(\gamma_t)_{II} \quad , \quad C = \mathbb{I}d^I \\
\Rightarrow \quad & CA^{-1}\mathbb{1}_I = \frac{1}{1-\alpha}(Id_I - g^I(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top)(\Sigma_{II})^{-1}\mathbb{1}_I = 0_I \\
\Rightarrow \quad & (*) = C(A+BC)^{-1}\mathbb{1}_I c \\
&= CA^{-1}\mathbb{1}_I c + CA^{-1}B(Id - CA^{-1}B)^{-1}CA^{-1}\mathbb{1}_I c \\
&= 0_I
\end{aligned}$$

In particular $\mathbb{1}^\top \pi^\lambda = \mathbb{1}_I^\top \pi_I^\lambda = \mathbb{1}_I^\top \rho_I = 1$ and $\pi^\lambda \in K$.
On the other hand we get for λ by the same arguments:

$$\begin{aligned}
\lambda_J &= -(\mu - r\mathbb{1})_J + D(t)_{JI}(1-\alpha)\pi_I^\lambda \\
&= -c_J - (\mu' - r\mathbb{1})_J + (1-\alpha)D(t)_{JI}\rho_I \\
&= -c_J \\
\bar{\lambda} &= (1-\alpha)g^I - g^I \; \mathbb{1}_I^\top \; \mathbb{D}^I(t)^{-1}(\mu - r\mathbb{1})_I^{+\mathbb{I}} \\
&= (1-\alpha)g^I - g^I\mathbb{1}_I^\top \left(\Sigma_{II} - \frac{\alpha}{1-\alpha}(T-t)(\gamma_t)_{II}\mathbb{I}d^I\right)^{-1} \\
&\qquad \Big((\mu - r\mathbb{1})_I + \alpha(T-t)g^I(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I\Big) \\
&= (1-\alpha)g^I - (1-\alpha)g^I\mathbb{1}_I^\top \left((1-\alpha)\Sigma_{II} - \alpha(T-t)(\gamma_t)_{II}\mathbb{I}d^I\right)^{-1} \\
&\qquad \Big(\big((1-\alpha)\Sigma_{II} - \alpha(T-t)(\gamma_t)_{II}\big)\rho_I + c\mathbb{1}_I + \alpha(T-t)g^I(\gamma_t)_{II}(\Sigma_{II})^{-1}\mathbb{1}_I\mathbb{1}_I^\top\rho_I\Big) \\
&= (1-\alpha)g^I - (1-\alpha)g^I\mathbb{1}_I^\top\rho_I \\
&\quad - (1-\alpha)g^I\mathbb{1}_I^\top \left((1-\alpha)\Sigma_{II} - \alpha(T-t)(\gamma_t)_{II}\mathbb{I}d^I\right)^{-1}\mathbb{1}_I c \\
&= -c
\end{aligned}$$

Again the last equality holds due to the above matrix inversion formula:

$$\begin{aligned}
&(1-\alpha)g^I\mathbb{1}_I^\top(A+BC)^{-1}\mathbb{1}_I c \\
&\qquad = (1-\alpha)g^I\mathbb{1}_I^\top A^{-1}\mathbb{1}_I c + (1-\alpha)g^I\mathbb{1}_I^\top A^{-1}B(Id - CA^{-1}B)^{-1}CA^{-1}\mathbb{1}_I c \\
&\qquad = g^I\mathbb{1}_I^\top(\Sigma_{II})^{-1}\mathbb{1}_I c \\
&\qquad = c
\end{aligned}$$

Hence these solutions are admissible and therefore they are already the unique optimal solutions. □

A.5 Notations

This is a brief overview of most of the used notation:

- *Technical Notations:*

$\mathbb{1}$	the d-dimensional vector of ones
Id	the $d \times d$-dimensional unit matrix
e_i	a d-dimensional unit-vector
I, J	various collections of dimensions
M_{IJ}	the part of matrix M with rows collected in I and columns collected in J
$\mathbb{1}_X$	the one-function if X is true
$(x)^-$	the negative part of x
$\mathrm{diag}\,(S)$	a matrix with S on its diagonal
$<, \leq, >, \geq$	those are meant component-wisely when applied to vectors

- *The market:*

$(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$	the usual filtered probability space
$(\mathcal{F}_t^S)$	the observed filtration, augmented by the stock prices
$[0, T]$	the investment horizon
d	the number of stocks
α	the risk-averseness parameter (in case of power utility)
B_t	the risk-free bond
S_t	the stocks (risky assets)
R_t	the returns of the stocks
W_t	the Brownian motion in the market
V_t	the Innovations Process (Brownian motion)
r	the interest rate of the bond
μ	the average growth rate of the stocks
$\Sigma = \sigma\sigma^\top$	the Covariance-matrix of the stocks
θ	the market price of risk
$\hat{\mu}_t$	the filter for μ
γ_t	the variance of the filter $\hat{\mu}_t$
X_t	various wealth processes
H_t	the 'state of the world'-process

- *Constrained optimization:*

K	various admissibility sets for the optimal strategies
δ	the support function of $-K$ (from convex analysis)
$\tilde{K}$	the domain of the support function
B^ν, S^ν	the bond and stocks of the auxiliary market
ν_t	the processes that describe the changes of the auxiliary markets
λ_t	the optimal choice for ν_t
$\mathcal{D}$	the set of admissible ν
π_t^ν	the optimal portfolio strategy in the auxiliary market
X_t^ν	the wealth process in the auxiliary market
H_t^ν	the 'state of the world'-process in the auxiliary market

- *Portfolio strategies:*

π^*	the optimal portfolio strategy of the respective current setting
π^M	the Merton (plug-in) strategy
π_{unc}	the optimal unconstrained portfolio strategy
π^λ	the optimal portfolio strategy in the auxiliary market
$\pi_{\text{unc}}^{+\lambda}$	the λ-augmented unconstrained portfolio strategy
$\pi^{(i,I)}$	the optimal portfolio strategy in Case $C = (i, I)$ in Chapter 4

B Bibliography

[BC09] A. Bain and D. Crisan. *Fundamentals of Stochastic Filtering.* Stochastic Modelling and Applied Probability 60, Springer, New York, 2009.

[BDL10] T. Björk, M. Davis, and C. Landén. Optimal investment under partial information. *Mathematical Methods of Operations Research*, 71(2):371–399, 2010.

[BJ97] X. Bardina and M. Jolis. An extension of ito's formula for elliptic diffusion processes. *Stochastic Processes and their Applications*, 69:83–109, 1997.

[Bre06] S. Brendle. Portfolio selection under incomplete information. *Stochastic Processes and their Applications*, 116(5):701–823, 2006.

[BS73] F. Black and M Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.

[BUV12] N. Bäuerle, S. Urban, and L. Veraart. The relaxed investor with partial information. *SIAM Journal on Financial Mathematics*, 3(1):304–327, 2012.

[BW96] S. Browne and W. Whitt. Portfolio choice and the bayesian kelly criterion. *Advances in Applied Probability*, 28(4):1145–1176, 1996.

[CK92] J. Cvitanic and I. Karatzas. Convex duality in constrained portfolio optimization. *The Annals of Applied Probability*, 2(4):767–818, 1992.

[CLMZ06] J. Cvitanic, A. Lazrak, L. Martellini, and F. Zapatero. Dynamic portfolio choice with parameter uncertainty and the economic value of analysts' recommendations. *The Review of Financial Studies*, 19(4):1113–1156, 2006.

[Cvi97] J. Cvitanic. Optimal trading under constraints. *Lecture Notes in Mathematics*, 1656:123–190, 1997. Springer, New York.

[DL13] M. Davis and S. Lleo. Black-litterman in continuous time: the case for filtering. *Quantitative Finance Letters*, 1:30–35, 2013.

[DV15] M. Dubois and L. Veraart. Optimal diversification in the presence of parameter uncertainty for a risk averse investor. *SIAM Journal on Financial Mathematics*, 6(1):201–241, 2015.

[EAM95] R. Elliott, L. Aggoun, and J. Moore. *Hidden Markov Models. Estimation and control.* Applications of Mathematics 29, Springer, New York, 1995.

[Eis00] N. Eisenbaum. Integration with respect to local time. *Potential Analysis*, 13:303–328, 2000.

[Eis01] N. Eisenbaum. On itô's formula of föllmer and protter. *Séminaire de Probabilités*, 35:390–395, 2001.

[ER94] R. Elliott and R. Rishel. Estimating the implicit interest rate of a risky asset. *Stochastic Processes and their Applications*, 49(2):199–206, 1994.

[FP00] H. Föllmer and P. Protter. On itô's formula for multidimensional brownian motion. *Probability Theory ans Related Fields*, 116:1–20, 2000.

[FPS95] H. Föllmer, P. Protter, and A. Shiryayev. Quadratic covariation and an extension of itô's formula. *Bernoulli*, 1(1):149–169, 1995.

[GKSW14] A. Gabih, H. Kondakji, J. Sass, and R. Wunderlich. Expert opinions and logarithmis utility maximization in a market with gaussian drift. *Communications on Stochastic Analysis*, 8(1):27–47, 2014.

[GV13] A. Gandy and L. Veraart. The effect of estimation in high-dimensional portfolios. *Mathematical Finance*, 23(3):531–559, 2013.

[HPS07] M. Hahn, W. Putschoegl, and J. Sass. Portfolio optimization with nonconstant volatility and partial information. *Brazilian Journal of Probability and Statistics*, 21(1):27–61, 2007.

[IS06] A. Irle and J. Sass. Optimal portfolio policies under fixed and proportional transaction costs. *Advances in Applied Probability 38*, pages 916–942, 2006.

[KK01] R. Korn and H. Kraft. A stochastic control approach to portfolio problems with stochastic interest rates. *SIAM Journal on Control and Optimization*, 40(12):1250–1269, 2001.

[KLS87] I. Karatzas, J. Lehoczky, and S. Shreve. Optimal portfolio and consumption decisions for a small investor on a finite horizon. *SIAM J. Control and Optimization*, 25(6):1557–1586, 1987.

[KLSX91] I. Karatzas, J. Lehoczky, S. Shreve, and G. Xu. Martingale and duality methods for utility maximization in an imcomplete market. *SIAM J. Control and Optimization*, 29(3):702–730, 1991.

[Kru12] S. Krumke. *Nonlinear Optimization.* Lecture notes, Technical University of Kaiserslautern, 2012.

[KZ07] R. Kan and G. Zhou. Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantative Analysis*, 42(3):621–656, 2007.

[Lak95] P. Lakner. Utility maximization with partial information. *Stochastic Processes and their Applications*, 56:247–273, 1995.

[Lak98] P. Lakner. Optimal trading strategy for an investor: the case of partial information. *Stochastic Processes and their Applications*, 76:77–97, 1998.

[Low10] G. Lowther. Nondifferentiable functions of one-dimensional semimartingales. *The Annals of Probability*, 38(1):76–101, 2010.

[LS01] R. Liptser and A. Shiryayev. *Statistics of Random Processes I, General Theory.* Springer, New York, 2001.

[Mar52] H. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

[Mer71] R. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3:373–413, 1971.

[Mer80] R. Merton. On estimating the expected return on the market. *Journal of Financial Economics*, 8:323–361, 1980.

[MN00] S. Moret and D. Nualart. Quadratic covariation and itô's formula for smooth nondegenerate martingales. *Journal of Theoretical Probability*, 13(1):193–224, 2000.

[MN01] S. Moret and D. Nualart. Generalization of itô's formula for smooth nondegenerate martingales. *Stochastic Processes and their Applications*, 91:115–149, 2001.

[NK11] E. Noh and J. Kim. An optimal portfolio model with stochastic volatility and stochastic interest rate. *Journal of Mathematical Analysis and Applications*, 375(2):510–522, 2011.

[PS08] W. Putschoegl and J. Sass. Optimal consumption and investment under partial information. *Decisions in Economics and Finance*, 31(2):137–170, 2008.

[PS11] W. Putschoegl and J. Sass. Optimal investment under dynamic risk constraints and partial information. *Quantitative Finance*, 11(10):1547–1564, 2011.

[Roc70] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[Rog01] L. Rogers. The relaxed investor and parameter uncertainty. *Finance and Stochastics*, 5(2):131–154, 2001.

[Sas07] J. Sass. Utility maximization with convex constraints and partial information. *Acta Applicandae Mathematicae*, 97:221–238, 2007.

[Saß12] J. Saß. *Continuous-Time Portfolio Optimization.* Lecture notes, Technical University of Kaiserslautern, 2012.

[SH04] J. Sass and U. Haussmann. Optimizing the terminal wealth under partial information: The drift process as a continuous time markov chain. *Finance and Stochastics*, 8(4):553–577, 2004.

[Shi80] A. Shiryaev. *Probability.* Graduate Texts in Mathematics, 95, Springer-Verlag, 1980.

[SS94] S. Shreve and H. Soner. Optimal investment and consumption with transaction costs. *The Annals of Applied Probability*, 4(3):609–692, 1994.

[SWW17] J. Sass, D. Westphal, and R. Wunderlich. Expert opinions and logarithmic utility maximization for multivariate stock returns with gaussian drift. 2017. Preprint, arXiv:1601.08155.

[Von14] C. Vonwirth. Expected utility maximization for high-dimensional portfolios: The effect of estimation. Mastersthesis, Technical University of Kaiserslautern, 2014.

C Scientific Career

August 1999 – March 2008	High School: Leininger Gymnasium Grünstadt, Degree: Abitur
April 2008 – December 2008	Compulsory Civilian Service at Ebertsheimer Bildungsinitiative e.V.
April 2009 – August 2011	Bachelor of Mathematics at TU Kaiserslautern, Specialization: Optimization and Stochastic
October 2010 – February 2011	Internship at Fraunhofer ITWM (Institute for technical and economical mathematics) Kaiserslautern, department of financial mathematics
September 2011 – February 2012	ERASMUS-Semester at University of Sheffield (United Kingdom), Department of Mathematics and Statistics
October 2011 – February 2014	Master Mathematics International at TU Kaiserslautern, Specialization: financial mathematics
April 2013 – July 2013	Internship at DWS (Deutsche Bank Group) in Frankfurt am Main, Chief Investment Office (Asset Management)
April 2014 – March 2017	Dissertation in financial mathematics at TU Kaiserslautern
May 2014 – February 2015	academic assistant at TU Kaiserslautern: research and teaching

Wissenschaftlicher Werdegang

August 1999 – März 2008	Leininger Gymnasium Grünstadt, Abitur
April 2008 – Dezember 2008	Zivildienst: Ebertsheimer Bildungsinitiative e.V.
April 2009 – August 2011	Bachelor of Mathematics, TU Kaiserslautern, Vertiefung: Optimierung und Stochastik
Oktober 2010 – Februar 2011	Fachpraktikum am Fraunhofer ITWM Kaiserslautern, Abteilung Finanzmathematik
September 2011 – Februar 2012	ERASMUS-Semester: University of Sheffield (United Kingdom), Department of Mathematics and Statistics
Oktober 2011 – Februar 2014	Master Mathematics International, TU Kaiserslautern, Vertiefung: Finanzmathematik
April 2013 – Juli 2013	Praktikum bei der DWS (Deutsche Bank Group) in Frankfurt am Main, Chief Investment Office (Asset Management)
April 2014 – März 2017	Promotion in Finanzmathematik an der TU Kaiserslautern
Mai 2014 – Februar 2015	wissenschaftlicher Mitarbeiter in Forschung und Lehre der TU Kaiserslautern

D Declaration of Academic Honesty

I, Christian Vonwirth, hereby declare that the work presented in this Dissertation has been performed and interpreted by myself except where I have explicitly referred to another source. Information derived from the work of others has been acknowledged in a list of references in the bibliography.

I confirm that this Dissertation is completely done to achieve the Philosophers Degree (Dr. rer. nat.) in Mathematics at Technische Universität Kaiserslautern and has not been submitted elsewhere.

(Christian Vonwirth)