Model Uncertainty and Expert Opinions in Continuous-Time Financial Markets

DOROTHEE WESTPHAL

1. Gutachter:	PROF. DR. JÖRN SASS Technische Universität Kaiserslautern
2. Gutachter:	PROF. DR. SÉBASTIEN LLEO NEOMA Business School Reims

Datum der Disputation: 31. Oktober 2019

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

D 386

Acknowledgments

First and foremost, I want to express my gratitude to Prof. Dr. Jörn Sass for his constant support and guidance in supervising this thesis. He has always been available to discuss problems and ideas, and his encouraging advice and endless motivation during the past years have been indispensable for developing this work. Moreover, I am very grateful for the opportunity to participate in various research conferences during my time as a Ph.D. student.

Special thanks are due to Prof. Dr. Ralf Wunderlich for introducing me to the concept of diffusion approximations and for countless valuable discussions during the development of our joint papers. I feel very grateful for his constant interest in my work and for his enthusiastic encouragement as well as for the invitations to Cottbus and the kind hospitality.

I would like to thank all members of the Financial Mathematics Group at the Department of Mathematics for providing such a pleasant working environment. My thanks also go to my friends and fellow Ph.D. students in the department for accompanying me on some part of the way. Our joint lunch breaks were always something I was looking forward to.

I thank my parents and my brother for their unconditional support, seeing that I never managed to explain what this thesis was about. Finally, I wish to thank Felix Riemann for teaching me how to use LaTeX years ago and for patiently listening to every up and down.

Contents

1. Introduction

Model uncertainty is a challenge that is inherent in many applications of mathematical models in various areas, for instance in mathematical finance and stochastic control. Optimization procedures in general take place under a particular model. This model, however, might be misspecified due to statistical estimation errors, incomplete information, biases, and for various other reasons. In that sense, any specified model must be understood as an approximation of the unknown "true" model. Difficulties arise since a strategy which is optimal under the approximating model might perform rather badly for the true model specifications.

To overcome these problems, it is crucial to improve the approximation of the true model by including any available source of information. On the other hand, one needs to find robust strategies, i.e. strategies that are less vulnerable to the specific choice of the model. A natural way to achieve this goal is to consider worst-case optimization problems. Instead of working with just one particular model, one specifies a range of possible models and tries to optimize the objective, given that for any chosen strategy the worst of all possible models will occur. Another way of looking at this problem is to see it as a two-player game in which one player tries to maximize and the other player simultaneously tries to minimize a given objective.

The optimization problems that we consider in this thesis are utility maximization problems in continuous-time financial markets. The most basic utility maximization problem in a Black–Scholes market is the Merton problem of maximizing expected utility of terminal wealth. It can be written in the form

$$V(x_0) = \sup_{\pi \in \mathcal{A}(x_0)} \mathbb{E}\big[U(X_T^\pi)\big],$$

where $U\colon \mathbb{R}_+ \to \mathbb{R}$ is a utility function, X_T^π denotes the terminal wealth achieved when using the self-financing strategy π and $\mathcal{A}(x_0)$ is the class of admissible strategies starting with initial capital x_0. An optimal strategy π^* for this problem is one that satisfies

$$V(x_0) = \mathbb{E}\big[U(X_T^{\pi^*})\big].$$

Merton [43] solves this problem for power and logarithmic utility functions and gives a corresponding optimal strategy. However, the setup of the problem assumes that an investor knows the market parameters, in particular the constant drift μ of asset returns. This is a rather unrealistic assumption since drift parameters are notoriously difficult to estimate from historical asset prices. At the same time, a misspecification of the drift has a massive effect on the optimal portfolio choice, see Chopra and Ziemba [10].

Model uncertainty (Part I)

To obtain strategies that are robust with respect to a possible misspecification of the drift parameter we consider in Part I of this thesis the worst-case optimization problem

$$\overline{V}(x_0) = \sup_{\pi \in \mathcal{A}(x_0)} \inf_{\mu \in K} \mathbb{E}_\mu\big[U(X_T^\pi)\big].$$

Here, we write $\mathbb{E}_\mu[\cdot]$ for the expectation with respect to a measure $\mathbb{P}^\mu$ under which the drift of the asset returns is $\mu \in \mathbb{R}^d$, with d denoting the number of risky assets in the market. The set $K \subseteq \mathbb{R}^d$ is called the *uncertainty set* or *ambiguity set.* An investor who faces uncertainty about the true drift parameter can choose K based on historical observations or external sources of information.

Our aim is to study the structure of optimal strategies for the robust utility maximization problem above in a continuous-time Black–Scholes market, as well as their asymptotic behavior as the degree of uncertainty increases, i.e. as the uncertainty set K becomes large. Since for large uncertainty investors usually do not invest in the risky assets at all, we restrict the class of admissible strategies by imposing a constraint that prevents a pure bond investment. Our focus is on ellipsoidal uncertainty sets K and on investigating what happens when increasing the radius of the uncertainty ellipsoid.

The main result in the first part of this thesis is an explicit representation of the optimal strategy and the worst-case drift parameter for the robust utility maximization problem with constrained strategies and ellipsoidal uncertainty sets. Moreover, a minimax theorem of the form

$$\sup_{\pi\in\mathcal{A}(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U(X_T^\pi)\big] = \inf_{\mu\in K} \sup_{\pi\in\mathcal{A}(x_0)} \mathbb{E}_\mu\big[U(X_T^\pi)\big]$$

is proven. The optimal strategy and worst-case drift parameter therefore constitute a saddle point of the problem. Using the explicit representation of the solution enables us to also study in detail the asymptotic behavior as uncertainty increases. We prove that the optimal strategy converges to a generalized uniform diversification strategy. In that sense, our results help to explain the popularity of uniform diversification strategies by the presence of uncertainty in the model.

Model uncertainty, also called *Knightian uncertainty* in reference to the seminal book by Knight [36], has been addressed in numerous papers. Gilboa and Schmeidler [28] and Schmeidler [59] formulate rigorous axioms on preference relations that account for risk aversion as well as uncertainty aversion. A robust utility functional in their sense is a mapping

$$X \mapsto \inf_{Q\in\mathcal{Q}} \mathbb{E}_Q\big[U(X)\big]$$

where U is a utility function and $\mathcal{Q}$ a convex set of probability measures. Chen and Epstein [9] give a continuous-time extension of this multiple-priors utility. Optimal investment decisions under such preferences are investigated in Quenez [50] and Schied [57], building up on results by Kramkov and Schachermayer [37, 38]. An extension of those results by means of a duality approach is given in Schied [58].

Papers addressing drift uncertainty in a financial market are Garlappi et al. [27] and Biagini and Pınar [4], among others. The latter focuses on ellipsoidal uncertainty sets like in this thesis. Uncertainty about both drift and volatility is investigated in a recent paper by Pham et al. [48]. In comparison to the literature mentioned above, we include in our setting a constraint on the admissible strategies that prevents investors from solely investing in the bond. We then derive explicit solutions to our robust utility maximization problem.

In a discrete-time setting, robust control problems have been studied by Favero and Runggaldier [20]. They give bounds for the suboptimality of strategies within a given class of models. In Dai Pra et al. [12] the authors consider a mixture of stochastic and robust optimization in a discrete-time setup.

Expert opinions (Part II)

The uncertainty set K in the robust utility maximization problem described above can in general be defined freely by the investor. To choose K in a reasonable way it makes sense to estimate the drift of asset returns and build the uncertainty set centered around the estimated value, with the shape and size of the set reflecting the reliability of the estimation. The ability to choose good trading strategies depends on how well the unobserved drift can be estimated. However, drift processes tend to fluctuate randomly over time and even if they were constant, long time series would be needed to estimate this parameter from asset returns with a satisfactory degree of precision. For these reasons, practitioners also incorporate external sources of information such as news, company reports, ratings or their own intuitive views when estimating the drift. These external sources of information are called *expert opinions*. In the context of the classical one-period Markowitz model this leads to the well-known Black–Litterman approach, where return predictions are improved by means of views formulated by securities analysts, see Black and Litterman [6].

In Part II of this thesis we consider a Black–Scholes type financial market where returns depend on an underlying drift process, which is unobservable due to additional noise coming from a Brownian motion. Investors in the market typically observe the return process. An additional source of information about the drift is provided by expert opinions, which we model as unbiased estimates of the drift that arrive at discrete points in time. Investors who have access to these expert opinions update their current drift estimates at each arrival time of such an expert opinion.

If the information of an investor is described by the filtration $(\mathcal{F}_t^H)_{t\in[0,T]}$, then the so-called *filter* is the conditional distribution of the drift μ_t given $\mathcal{F}_t^H$. The best estimate for the hidden drift process in a mean-squared sense is given by the conditional mean $m_t^H = \mathbb{E}[\mu_t \,|\, \mathcal{F}_t^H]$. A measure for the goodness of this estimator is its conditional covariance matrix

$$Q_t^H := \mathbb{E}\big[(\mu_t - m_t^H)(\mu_t - m_t^H)^\top \,\big|\, \mathcal{F}_t^H\big].$$

The aim in the second part of this thesis is to investigate in detail the influence that expert opinions have on investors' estimates of the drift process by analyzing the filter for different investor filtrations.

For investors who observe only the return process, the filter is the classical Kalman filter, see for example Liptser and Shiryaev [42] or Davis [13]. Observing in addition also discrete-time expert opinions leads to updates of the filter at each information date. These updates decrease the conditional covariance, hence they yield better estimates. This can be seen as a continuous-time version of the above mentioned static Black–Litterman approach.

Our main results in this part are concerned with the asymptotic behavior of the filter as the arrival frequency of expert opinions goes to infinity on a finite time horizon. If expert opinions have some minimal level of accuracy, characterized by bounded experts' covariances, then the conditional covariance of the drift estimate goes to zero as the arrival frequency goes to infinity. This implies that the conditional mean converges to the true drift process, i.e. in the limit investors have full information about the drift.

We further study a situation in which a higher frequency of expert opinions is only available at the cost of accuracy. In other words, as the frequency of expert opinions increases, their variance becomes larger. For properly scaled variance of expert opinions we derive limit theorems which state that the information obtained from observing the discrete-time expert opinions is asymptotically the same as that from observing a certain diffusion process, which is why we speak of *diffusion approximations*. The limit process can be interpreted as a continuous-time expert who permanently delivers noisy information about the drift.

The diffusion approximations that we obtain from our main results are useful since the asymptotic filter is easy to compute whereas the updates for the discrete-time expert opinions lead to a computationally involved filter. Numerical simulations show that the approximation is very accurate even for a small number of expert opinions. Our convergence results for the filter also carry over to convergence of the value function in a portfolio optimization problem with logarithmic utility. This allows to find approximate solutions of utility maximization problems by replacing the filter of an investor who observes the discrete-time expert opinions by the asymptotic filter. This filter, corresponding to an investor who observes a continuous-time expert, is much easier to handle numerically.

Without expert opinions, our utility maximization problem is a classical optimization problem under partial information. Under suitable integrability assumptions the existence of optimal trading strategies can be shown, see Björk et al. [5] and Lakner [40]. The computation of optimal strategies is possible if the model allows for finite-dimensional filters. This is the case if the drift process is an Ornstein–Uhlenbeck process or a continuous-time Markov chain. In these two models, the solution of the utility maximization problem is known, see Brendle [7], Lakner [41] and Putschögl and Sass [49], as well as Honda [31], Rieder and Bäuerle [51] and Sass and Haussmann [54]. Fouque et al. [21, 22] also model the drift as an Ornstein–Uhlenbeck process and analyze the loss of utility due to partial information.

Including unbiased expert opinions reduces the variance of the filter. The better estimate then improves the expected utility. In a static model, the Black–Litterman approach combines an estimate of the asset returns with expert opinions on the performance of the assets, see Black and Litterman [6]. The idea of a continuous-time expert in this thesis is in line with Davis and Lleo [14] where such an expert is introduced as an approximation of discrete-time experts, allowing for more explicit solutions in portfolio optimization problems. Davis and Lleo [14] term that approach "Black–Litterman in Continuous Time". First papers addressing this approach are Frey et al. [23, 24]. They consider a continuous-time Markov chain for the drift and expert opinions arriving at the jump times of a Poisson process and study the maximization of expected power utility of terminal wealth. An Ornstein–Uhlenbeck drift process is considered in Gabih et al. [25] for a one-dimensional financial market. Part II of this thesis builds up on the Master's thesis Westphal [64], in which many results from Gabih et al. [25] are carried over to a financial market with multivariate stock returns. Some results from the Master's thesis [64] are repeated in Part II to give a complete picture.

In the literature, diffusion approximations also appear in other contexts. They are well-known in operations research and actuarial mathematics. The basic idea is to replace a complicated stochastic process by an appropriate diffusion process which is more analytically tractable than the original process. The approach is comparable with the normal approximation of sums of random variables following from the Central Limit Theorem. For an introduction to diffusion approximations based on the theory of weak convergence and applications to queueing systems in heavy traffic we refer to the survey article by Glynn [29]. In risk theory the application of diffusion approximations for computing ruin probabilities goes back to Iglehart [33]. We also refer to Grandell [30, Sec. 1.2], Schmidli [60, Sec. 5.10 and 6.5] and Asmussen and Albrecher [3, Sec. V.5], as well as the references therein. Convergence of a discrete-time Kalman filter to the continuous-time equivalent has been addressed e.g. by Salgado et al. [53] or Aalto [1] for the case of deterministic information dates.

Robust optimization with expert opinions (Part III)

In Part III we combine our results for the structure of the robust utility maximization problem with the observations about how expert opinions improve drift estimates. The main idea is that the uncertainty set K can be defined based on a drift estimate and hence a better estimate due to the observation of expert opinions should be reflected in a smaller uncertainty set.

To be able to capture a change in information about the drift over time we generalize our financial market model to one with non-constant drift. We then carry over our results for the robust utility maximization problem to the more general model where we also introduce time-dependence in the uncertainty set. The computation of the optimal strategy carries over when assuming in the worst-case problem that the drift process μ_t at time t can take any value in an ellipsoid K_t, where K_t is known at time t.

We then show how the time-dependent uncertainty set $(K_t)_{t\in[0,T]}$ can be defined based on the filter $\mathbb{E}[\mu_t \,|\, \mathcal{F}_t^H]$ for various investor filtrations $(\mathcal{F}_t^H)_{t\in[0,T]}$. It becomes clear that expert opinions decrease the size of the uncertainty set, reflecting the better estimations. We also investigate which effect expert opinions have on the robust strategies and compare them with the non-robust strategies that rely on the drift estimation only.

For detailed outlines we refer to the beginning of each of the three parts, see pp. 9, 77 and 145. Appendix A provides some auxiliary results needed for proving our main theorems in Chapter 8. In Appendix B we extend the well-known result from Merton [43] by showing that for a power utility maximization problem we obtain the same structure of the optimal strategy if the risk-free interest rate as well as drift and volatility of the stocks are not necessarily constant but independent of the driving Brownian motions.

Part I.

Model Uncertainty

Outline and Notation

In this part of the thesis we investigate optimal trading strategies for a robust utility maximization problem in a continuous-time Black–Scholes type financial market under a constraint that prevents a pure bond investment. We deduce a minimax theorem for our robust optimization problem and show that, as the degree of model uncertainty increases, a generalized uniform portfolio diversification strategy outperforms more sophisticated strategies.

This part of the thesis is organized as follows. In Chapter 2 we state our financial market model and introduce the robust utility maximization problem. Chapter 3 addresses a special case of that problem for uncertainty sets that are balls in $\mathbb{R}^d$ and investors with logarithmic utility. In this setting, we can carry over the approach from a one-period risk minimization problem by Pflug et al. [47] to our continuous-time setting.

The main results of this part are given in Chapter 4. Here we use a duality approach to solve our robust utility maximization problem for ellipsoidal uncertainty sets and power or logarithmic utility. The main idea is to solve the dual problem explicitly and show that the solution forms a saddle point of the problem. We give representations of the optimal strategy and the worst-case drift parameter and provide a minimax theorem.

In Chapter 5 we study the asymptotic behavior of the optimal strategy and the worst-case parameter as the degree of uncertainty goes to infinity. We show that the optimal strategy converges to a generalized uniform diversification strategy, where by uniform diversification we mean the equal weight or $1/d$ strategy for the investment in the risky assets. Furthermore, we analyze the influence of the investor's risk aversion on the speed of convergence and investigate measures for the performance of the optimal robust strategies.

Notation. Throughout this part, we use the notation I_d for the identity matrix in $\mathbb{R}^{d\times d}$ as well as e_i, $i = 1,\dots,d$, for the i-th standard unit vector in $\mathbb{R}^d$, and $\mathbf{1}_d$ for the vector in $\mathbb{R}^d$ containing a one in every component. For a vector $x \in \mathbb{R}^d$ we denote with $\operatorname{diag}(x)$ the diagonal matrix $A = (a_{ij}) \in \mathbb{R}^{d\times d}$ with $a_{ii} = x_i$ for all $i = 1,\dots,d$, $a_{ij} = 0$ if $i \neq j$. We shortly write $\mathbb{R}_+ = (0,\infty)$. By $\langle\cdot,\cdot\rangle$ we denote the scalar product on $\mathbb{R}^d \times \mathbb{R}^d$ with $\langle x,y\rangle = x^\top y$ for $x,y \in \mathbb{R}^d$.

If $x \in \mathbb{R}^d$, then for any $p \in [1,\infty]$ we write $\|x\|_p$ for the p-norm of x on $\mathbb{R}^d$. When using the duality $\frac{1}{p} + \frac{1}{q} = 1$ we stick to the convention that $\frac{1}{\infty} = 0$. Unless stated otherwise, whenever $x \in \mathbb{R}^d$ is a vector, $\|x\|$ denotes the Euclidean norm of x, whenever A is a matrix, $\|A\|$ denotes the spectral norm of A.

2. A Robust Utility Maximization Problem

2.1. Financial market model

We consider a continuous-time financial market with one risk-free and various risky assets. By $T > 0$ we denote some finite investment horizon. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions. All processes are assumed to be $\mathbb{F}$-adapted.

The dynamics of the risk-free asset S^0 are given by

$$\mathrm{d}S_t^0 = S_t^0 r \,\mathrm{d}t, \quad S_0^0 = 1,$$

where $r \in \mathbb{R}$ is the deterministic risk-free interest rate. Note that we can write S_t^0 in explicit form as $S_t^0 = \mathrm{e}^{rt}$, $t \in [0, T]$. Aside from the risk-free asset, investors can also invest in $d \geq 2$ risky assets $S^1, \dots, S^d$ where the dynamics of $S = (S^1, \dots, S^d)^\top$ are given by

$$\mathrm{d}S_t = \operatorname{diag}(S_t)\big(\mu \,\mathrm{d}t + \sigma \,\mathrm{d}W_t\big), \quad S_0 = s_0.$$

Here, $W = (W_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion with $m \geq d$. Further, $\mu \in \mathbb{R}^d$ and $\sigma \in \mathbb{R}^{d\times m}$, where we assume that σ has full rank equal to d. The vector $s_0 \in \mathbb{R}^d$ of initial asset prices is assumed to have strictly positive entries. In explicit form, S^i can be written as

$$S_t^i = s_0^i \exp\Bigg(\bigg(\mu^i - \frac{1}{2}\sum_{j=1}^m \sigma_{ij}^2\bigg)t + \sum_{j=1}^m \sigma_{ij} W_t^j\Bigg), \quad i = 1, \dots, d.$$

Rather than working directly with the price process $S = (S^1, \dots, S^d)^\top$ we consider the return process $R = (R^1, \dots, R^d)^\top$, defined by

$$\mathrm{d}R_t = \mu \,\mathrm{d}t + \sigma \,\mathrm{d}W_t, \quad R_0 = 0.$$

Note that we can write

$$R_t^i = \log(S_t^i) - \log(s_0^i) + \frac{1}{2}\sum_{j=1}^m \sigma_{ij}^2 t$$

for $i = 1, \dots, d$, which implies that the filtration generated by the stock price process is the same as the filtration generated by the return process. Therefore, we assume in the following that investors observe stock returns directly.

An investor's trading decisions in this market are described by a self-financing trading strategy $(\pi_t)_{t\in[0,T]}$ with values in $\mathbb{R}^d$. The entry π_t^i, $i = 1, \dots, d$, is the proportion of wealth that is invested in asset i at time t. Consequently, the proportion $1 - \mathbf{1}_d^\top \pi_t$ is invested in the risk-free asset. The corresponding wealth process $(X_t^\pi)_{t\in[0,T]}$ can then be described by the stochastic differential equation

$$\mathrm{d}X_t^\pi = X_t^\pi\Big(r \,\mathrm{d}t + \pi_t^\top(\mu - r\mathbf{1}_d)\,\mathrm{d}t + \pi_t^\top \sigma \,\mathrm{d}W_t\Big), \quad X_0^\pi = x_0.$$

Here, $x_0 > 0$ denotes the initial wealth of an investor. The solution to this SDE can be written as

$$X_t^\pi = x_0 \exp\bigg(\int_0^t \Big(r + \pi_s^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \pi_s\|^2\Big)\mathrm{d}s + \int_0^t \pi_s^\top \sigma \, \mathrm{d}W_s\bigg).$$

We assume that investors do not observe the Brownian motion $(W_t)_{t\in[0,T]}$ but the returns $(R_t)_{t\in[0,T]}$, and that any trading strategy has to be based on the current information of an investor. We therefore require trading strategies to be $\mathbb{F}^R$-adapted, where $\mathbb{F}^R = (\mathcal{F}_t^R)_{t\in[0,T]}$ for $\mathcal{F}_t^R = \sigma((R_s)_{s\in[0,t]})$. Hence, the class of admissible trading strategies has the form

$$\mathcal{A}(x_0) = \bigg\{\pi = (\pi_t)_{t\in[0,T]} \;\bigg|\; \pi \text{ is } \mathbb{F}^R\text{-adapted, } X_0^\pi = x_0,\ \mathbb{E}\bigg[\int_0^T \|\sigma^\top \pi_t\|^2 \,\mathrm{d}t\bigg] < \infty\bigg\}.$$

To start with, we consider an investor who seeks to maximize expected utility of terminal wealth. The value function of that investor is given by

$$V(x_0) = \sup_{\pi\in\mathcal{A}(x_0)} \mathbb{E}\big[U(X_T^\pi)\big], \tag{2.1}$$

where $U\colon \mathbb{R}_+ \to \mathbb{R}$ is a utility function, i.e. a strictly concave and continuously differentiable function satisfying the so-called Inada conditions

$$\lim_{x\downarrow 0} U'(x) = \infty \quad \text{and} \quad \lim_{x\to\infty} U'(x) = 0.$$

It is well known from the paper by Merton [43] that for logarithmic utility $U = \log$, the optimal strategy for this portfolio optimization problem is $(\pi_t^*)_{t\in[0,T]}$ with

$$\pi_t^* = (\sigma\sigma^\top)^{-1}(\mu - r\mathbf{1}_d),$$

and that for power utility $U(x) = \frac{x^\gamma}{\gamma}$ with $\gamma \in (-\infty, 1)$, $\gamma \neq 0$, one obtains

$$\pi_t^* = \frac{1}{1-\gamma}(\sigma\sigma^\top)^{-1}(\mu - r\mathbf{1}_d).$$

Hence, the portfolio optimization problem (2.1) can be solved explicitly for both logarithmic and power utility. However, the above setting has a serious drawback. It relies on the assumption that the parameters μ and σ of the stock price dynamics are known to the investor. This assumption is, at least for the drift, rather unrealistic since drift parameters are notoriously difficult to estimate from historical asset return data. For investors to be able to estimate μ with a reasonable degree of precision, they would need to observe very long time series. For this reason, we introduce in the following a robust version of the above portfolio optimization problem in which there is uncertainty about the drift parameter μ.

2.2. Introducing model uncertainty

We take the model from the preceding section as a starting point, but now assume that the true drift of the stocks is only known to be an element of some set $K \subseteq \mathbb{R}^d$ and that investors want to maximize their worst-case expected utility

$$\inf\Big\{\mathbb{E}\big[U(X_T^\pi)\big] \;\Big|\; \text{drift of the stocks is } \mu \in K\Big\}$$

for some utility function $U\colon \mathbb{R}_+ \to \mathbb{R}$. We formalize this in the following by means of a change of measure. Let us fix some element $\nu \in K$ as a reference parameter and say that, under $\mathbb{P}$, the drift of the stock returns is equal to ν, i.e.

$$\mathrm{d}R_t = \nu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \quad R_0 = 0.$$

The value ν can be thought of as an estimate for the drift that was obtained by observing the historical stock prices or by external sources of information. Changing the drift from ν to some $\mu \in K$ can then be expressed by a change of measure. For this purpose, define the process $(Z_t^\mu)_{t\in[0,T]}$ by

$$Z_t^\mu = \exp\Big(\theta(\mu)^\top W_t - \frac{1}{2}\|\theta(\mu)\|^2 t\Big),$$

where $\theta(\mu) = \sigma^\top(\sigma\sigma^\top)^{-1}(\mu - \nu)$. We can then define a new measure $\mathbb{P}^\mu$ by setting

$$\frac{\mathrm{d}\mathbb{P}^\mu}{\mathrm{d}\mathbb{P}} = Z_T^\mu.$$

Note that since $\theta(\mu)$ is a constant, the process $(Z_t^\mu)_{t\in[0,T]}$ is a strictly positive martingale. Therefore, $\mathbb{P}^\mu$ is a probability measure that is equivalent to $\mathbb{P}$ and we obtain from Girsanov's Theorem that the process $(W_t^\mu)_{t\in[0,T]}$, defined by

$$W_t^\mu = W_t - \theta(\mu)t,$$

is a Brownian motion under $\mathbb{P}^\mu$. We can thus rewrite the return dynamics as

$$\mathrm{d}R_t = \nu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t = \nu\,\mathrm{d}t + \sigma\big(\mathrm{d}W_t^\mu + \theta(\mu)\,\mathrm{d}t\big) = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t^\mu,$$

and see that a change of measure from $\mathbb{P}$ to $\mathbb{P}^\mu$ corresponds to changing the drift in the return dynamics from ν to μ. We thus shortly write $\mathbb{E}_\mu[\cdot]$ for the expectation under measure $\mathbb{P}^\mu$ and $\mathbb{E}[\cdot] = \mathbb{E}_\nu[\cdot]$ for the expectation under our reference measure $\mathbb{P} = \mathbb{P}^\nu$.

For a trading strategy $(\pi_t)_{t\in[0,T]}$ we require admissibility under any of the measures $\mathbb{P}^\mu$ for $\mu \in K$. We thus modify our notion of admissibility and let

$$\mathcal{A}(x_0) = \left\{(\pi_t)_{t\in[0,T]} \,\middle|\, \pi \text{ is } \mathbb{F}^R\text{-adapted, } X_0^\pi = x_0,\ \mathbb{E}_\mu\left[\int_0^T \|\sigma^\top \pi_t\|^2\,\mathrm{d}t\right] < \infty \text{ for all } \mu \in K\right\}.$$

Then admissibility in the robust context means admissibility under any of the measures $\mathbb{P}^\mu$ for $\mu \in K$ and our robust portfolio optimization problem for a generic utility function $U\colon \mathbb{R}_+ \to \mathbb{R}$ can be formulated as

$$\overline{V}(x_0) = \sup_{\pi\in\mathcal{A}(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U(X_T^\pi)\big]. \tag{2.2}$$

2.3. Constraint on the admissible strategies

In the next chapters, we investigate the robust utility maximization problem (2.2) for power and logarithmic utility. We use the notation $U_\gamma\colon \mathbb{R}_+ \to \mathbb{R}$ for $\gamma \in (-\infty, 1)$, where $U_\gamma(x) = \frac{x^\gamma}{\gamma}$ for $\gamma \neq 0$ denotes power utility and $U_0(x) = \log(x)$ is the logarithmic utility function.

First, we make the observation that for a large degree of model uncertainty the trivial strategy $\pi \equiv 0$ becomes optimal both for logarithmic and for power utility.

Proposition 2.1. *Let $\gamma \in (-\infty, 1)$ and $K \subseteq \mathbb{R}^d$. If $r\mathbf{1}_d \in K$, then the strategy $(\pi_t)_{t\in[0,T]}$ with $\pi_t = 0$ for all $t \in [0,T]$ is optimal for the optimization problem*

$$\sup_{\pi\in\mathcal{A}(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]. \tag{2.3}$$

Proof. Recall that under $\mathbb{P}^\mu$, the drift of the stock returns is the vector μ, and that the wealth process corresponding to a self-financing trading strategy $\pi \in \mathcal{A}(x_0)$ thus satisfies

$$X_t^\pi = x_0 \exp\biggl(\int_0^t \Bigl(r + \pi_s^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_s\|^2\Bigr)\mathrm{d}s + \int_0^t \pi_s^\top\sigma\,\mathrm{d}W_s^\mu\biggr),$$

where W^μ is a Brownian motion under $\mathbb{P}^\mu$. We consider the case $\gamma = 0$ first. When applying the logarithm $U_0 = \log$ to terminal wealth X_T^π, we obtain

$$\log(X_T^\pi) = \log(x_0) + \int_0^T \Bigl(r + \pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_t\|^2\Bigr)\mathrm{d}t + \int_0^T \pi_t^\top\sigma\,\mathrm{d}W_t^\mu.$$

For any admissible π, the stochastic integral in the above equation is a martingale under $\mathbb{P}^\mu$, hence it vanishes in expectation. The expected logarithmic utility of terminal wealth under measure $\mathbb{P}^\mu$ is then

$$\mathbb{E}_\mu\big[\log(X_T^\pi)\big] = \log(x_0) + \mathbb{E}_\mu\biggl[\int_0^T \Bigl(r + \pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_t\|^2\Bigr)\mathrm{d}t\biggr].$$

Now since the vector $r\mathbf{1}_d$ is an element of the set K of possible drift parameters, we immediately see that

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^\pi)\big] \le \mathbb{E}_{r\mathbf{1}_d}\big[\log(X_T^\pi)\big] \le \log(x_0) + rT,$$

so we can deduce that the trivial strategy $\pi \equiv 0$ is optimal for (2.3), since $\pi \equiv 0$ leads to expected utility of terminal wealth $\log(x_0) + rT$ under each of the measures $\mathbb{P}^\mu$.

For power utility, i.e. $\gamma \neq 0$, the argumentation is similar. Since $r\mathbf{1}_d \in K$, we have

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \le \frac{x_0^\gamma}{\gamma}\mathrm{e}^{\gamma rT}\,\mathbb{E}_{r\mathbf{1}_d}\biggl[\exp\biggl(-\frac{\gamma}{2}\int_0^T \|\sigma^\top\pi_t\|^2\,\mathrm{d}t + \gamma\int_0^T \pi_t^\top\sigma\,\mathrm{d}W_t^{r\mathbf{1}_d}\biggr)\biggr]$$

and we can rewrite

$$\begin{aligned}
&\mathbb{E}_{r\mathbf{1}_d}\biggl[\exp\biggl(-\frac{\gamma}{2}\int_0^T \|\sigma^\top\pi_t\|^2\,\mathrm{d}t + \gamma\int_0^T \pi_t^\top\sigma\,\mathrm{d}W_t^{r\mathbf{1}_d}\biggr)\biggr] \\
&= \mathbb{E}_{r\mathbf{1}_d}\biggl[\exp\biggl(\gamma\int_0^T \pi_t^\top\sigma\,\mathrm{d}W_t^{r\mathbf{1}_d} - \frac{1}{2}\gamma^2\int_0^T \|\sigma^\top\pi_t\|^2\,\mathrm{d}t\biggr)\exp\biggl(-\frac{1}{2}\gamma(1-\gamma)\int_0^T \|\sigma^\top\pi_t\|^2\,\mathrm{d}t\biggr)\biggr].
\end{aligned}$$

Note that the term

$$\exp\biggl(-\frac{1}{2}\gamma(1-\gamma)\int_0^T \|\sigma^\top\pi_t\|^2\,\mathrm{d}t\biggr)$$

is less or equal than one if $\gamma > 0$ and greater or equal than one if $\gamma < 0$. Thus, in both cases,

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \le \frac{x_0^\gamma}{\gamma}\mathrm{e}^{\gamma rT}\,\mathbb{E}_{r\mathbf{1}_d}\biggl[\exp\biggl(\gamma\int_0^T \pi_t^\top\sigma\,\mathrm{d}W_t^{r\mathbf{1}_d} - \frac{1}{2}\gamma^2\int_0^T \|\sigma^\top\pi_t\|^2\,\mathrm{d}t\biggr)\biggr].$$

But the exponential local martingale in the expression above has expectation less or equal than one, so

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \leq \frac{x_0^\gamma}{\gamma}\mathrm{e}^{\gamma rT}.$$

So again, as for logarithmic utility, the trivial strategy $\pi \equiv 0$ is optimal for (2.3) if $r\mathbf{1}_d \in K$, since the zero strategy leads exactly to expected power utility $\frac{x_0^\gamma}{\gamma}\mathrm{e}^{\gamma rT}$. $\square$

The observations above imply that when we take K to be for example some ball with respect to a norm on $\mathbb{R}^d$, say centered around our reference parameter ν, then as the radius of that ball and thus the level of uncertainty about the true drift parameter exceeds a certain threshold, it will be optimal for investors, both with logarithmic and with power utility, to not invest anything in the stocks and everything in the risk-free asset.

Remark 2.2. The result from Proposition 2.1 is in line with a result in Biagini and Pınar [4] where the authors also consider an increasing degree of uncertainty. In Øksendal and Sulem [44, 45] the authors obtain a similar result for optimality of $\pi \equiv 0$. They consider a jump diffusion model with a worst-case approach where the market chooses a scenario from a fixed but very comprehensive set of probability measures. In contrast, it is shown in Zawisza [67] that, if the model allows for stochastic interest rate r, the previous result does not hold in general. In particular, the author shows that the optimal investment strategy in a model with stochastic interest rate r does not invest exclusively in the bond.

Investing everything in the risk-free asset is, of course, a very extreme reaction to model uncertainty. We are interested in finding less conservative strategies that still take into account the increasing risk coming with a higher degree of model uncertainty. For that purpose, we introduce another constraint on our strategies that prevents investors from solely investing in the bond. Consider therefore the admissibility set

$$\mathcal{A}_h(x_0) = \big\{\pi \in \mathcal{A}(x_0) \,\big|\, \langle \pi_t, \mathbf{1}_d\rangle = h \text{ for all } t \in [0,T]\big\},$$

where $h > 0$. Taking $h = 1$ would imply that investors are not allowed to invest anything in the risk-free asset. They must then distribute all of their wealth so as to invest everything in the risky assets. Any value $h > 1$ forces investors to have a negative position in the risk-free asset. For instance, a constraint of the form $\langle \pi_t, \mathbf{1}_d\rangle = h > 0$ typically applies for some mutual funds when investors are required to invest a certain amount in risky assets.

Remark 2.3. The admissibility set $\mathcal{A}_h(x_0)$ might seem unnecessarily restrictive at first glance. Instead of fixing $\langle \pi_t, \mathbf{1}_d\rangle = h$ for any t one might also want to consider utility maximization among the larger class of strategies π with $\langle \pi_t, \mathbf{1}_d\rangle \geq h$. However, we are mainly interested in the asymptotic behavior of the optimal strategies as the level of uncertainty increases. From that point of view it is intuitively clear that, when uncertainty is large, investors seek to invest as little as possible in the risky assets. Therefore, enlarging the class of admissible strategies asymptotically does not change the value of the optimization problem, which is why we consider optimization among strategies in $\mathcal{A}_h(x_0)$. Our results can then be used to show that one obtains the same value asymptotically when allowing for a larger range of strategies, see Section 5.2 for the exact statements.

In the following chapters we solve the optimization problem

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \inf_{\mu \in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

for specific uncertainty sets K and analyze the asymptotic behavior of the optimal strategy as the degree of uncertainty goes to infinity.

3. Robust Optimization in Uncertainty Balls

In this chapter we consider our continuous-time utility maximization problem

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \inf_{\mu \in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

for $\gamma = 0$, i.e. logarithmic utility, and for uncertainty sets K which are balls in some p-norm. For solving the problem it turns out that one can carry over results from a discrete-time risk minimization problem as introduced in Pflug et al. [47]. In the following, we first repeat the approach and results obtained in the aforementioned paper.

3.1. Risk minimization in a one-period model

Pflug et al. [47] fix a probability space $(\Omega, \mathcal{F}, P)$ and consider for $1 \leq p < \infty$ random variables $X \in \mathrm{L}^p(\Omega, \mathcal{F}, P)$ representing random future losses. For measuring the riskiness of such losses, they apply risk measures $\mathcal{R}\colon \mathrm{L}^p(\Omega, \mathcal{F}, P) \to \mathbb{R}$. The authors analyze the optimization problem for an investor who wants to invest a fixed amount of money in a combination of $d \geq 2$ risky assets for one period of time. The investment decision is modelled as a vector $w \in \mathbb{R}^d$ of portfolio weights.

Now let $X\colon \Omega \to \mathbb{R}^d$ be a random variable representing future losses of d risky assets. If an investor was sure about the correct distribution of the future losses, then her risk minimization problem could be defined as

$$\begin{aligned} &\inf_{w \in \mathbb{R}^d} \mathcal{R}(\langle X, w\rangle) \\ &\quad \text{s.t. } \langle w, \mathbf{1}_d\rangle = 1. \end{aligned} \tag{3.1}$$

However, in many situations investors face uncertainty about the true distribution of the losses. For Q any Borel measure on $\mathbb{R}^d$ we denote by X^Q an $\mathbb{R}^d$-valued random variable with image measure Q. Let us denote by $\widehat{Q}$ one specific Borel measure on $\mathbb{R}^d$ that can be thought of as an estimation for the true distribution of the losses. To account for model uncertainty it is reasonable to define a set of possible loss distributions centered around the reference measure $\widehat{Q}$. This set is called *ambiguity set* in Pflug et al. [47] and defined as

$$B_\kappa(\widehat{Q}) = \big\{Q \in \mathcal{P}(\mathbb{R}^d) \,\big|\, d_p(\widehat{Q}, Q) \leq \kappa\big\}$$

for some $\kappa > 0$. Here, $\mathcal{P}(\mathbb{R}^d)$ denotes the space of all Borel probability measures on $\mathbb{R}^d$ and d_p is the *Wasserstein metric* with parameter $p \in [1, \infty)$, i.e. $d_p\colon \mathcal{P}(\mathbb{R}^d) \times \mathcal{P}(\mathbb{R}^d) \to [0, \infty)$ with

$$d_p(\widehat{Q}, Q) = \inf_{\eta \in H(\widehat{Q}, Q)} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|_p^p \,\mathrm{d}\eta(x, y)\right)^{1/p},$$

where $H(\widehat{Q}, Q)$ is the set of all measures on $\mathbb{R}^d \times \mathbb{R}^d$ for which the marginal distribution of the first d components is $\widehat{Q}$ and the marginal distribution of the last d components is Q. Note that κ is the radius of the ambiguity set, hence an increase in κ corresponds to a higher level of uncertainty about the true distribution. A robust version of (3.1) can then be formulated as

$$\begin{aligned} \inf_{w \in \mathbb{R}^d} \sup_{Q \in B_\kappa(\widehat{Q})} \; & \mathcal{R}(\langle X^Q, w \rangle) \\ \text{s.t. } & \langle w, \mathbf{1}_d \rangle = 1. \end{aligned} \tag{3.2}$$

This is a worst-case approach where the investor wants to minimize the risk of her investment under the worst possible probability measure $Q \in B_\kappa(\widehat{Q})$. A resulting optimal strategy will therefore be robust with respect to the model uncertainty.

We now recap the main results in Pflug et al. [47]. Assume that $\mathcal{R}$ is a convex, law invariant risk measure on $\mathrm{L}^p(\Omega, \mathcal{F}, P)$, $1 \leq p < \infty$, which admits a dual characterization of the form

$$\mathcal{R}(X) = \max_{Z \in \mathrm{L}^q} \{\mathbb{E}[XZ] - \alpha(Z)\},$$

where q is such that $\frac{1}{p} + \frac{1}{q} = 1$ and $\alpha \colon \mathrm{L}^q(\Omega, \mathcal{F}, P) \to \mathbb{R}$ is convex. The key observation is that, under suitable assumptions on the risk measure $\mathcal{R}$, there exists a constant $C > 0$ such that

$$\sup_{Q \in B_\kappa(\widehat{Q})} \mathcal{R}(\langle X^Q, w \rangle) = \mathcal{R}(\langle X^{\widehat{Q}}, w \rangle) + C\kappa \|w\|_q \tag{3.3}$$

for all $\kappa > 0$ and $w \in \mathbb{R}^d$. Equation (3.3) shows that the investor's objective in the robust problem (3.2) can be decomposed into two components. On the one hand it is beneficial to minimize the risk $\mathcal{R}(\langle X^{\widehat{Q}}, w \rangle)$ under the reference measure $\widehat{Q}$. On the other hand, investors also seek to minimize $\|w\|_q$ among the admissible weights $w \in \mathbb{R}^d$. These two effects are antithetic and from (3.3) we see that for large levels of uncertainty κ, the latter will dominate. This is an intuitive explanation for why uniform diversification $w^u = (\frac{1}{d}, \ldots, \frac{1}{d})^\top$ will be optimal for large levels of uncertainty, since w^u minimizes $\|w\|_q$ subject to the constraint $\langle w, \mathbf{1}_d \rangle = 1$. This result is formalized in Pflug et al. [47, Prop. 3] which we repeat in the following.

Proposition 3.1. *Let $p \in [1, \infty)$ and $q \in (1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$, and let $\mathcal{R}$ be a convex risk measure on $\mathrm{L}^p(\Omega, \mathcal{F}, P)$. Then, under some additional assumptions on $\mathcal{R}$, it holds:*

(i) For any $\varepsilon > 0$ there exists a κ_ε such that for all $\kappa > \kappa_\varepsilon$ the optimal solution w^* for (3.2) *satisfies $\|w^* - w^u\|_q < \varepsilon$.*

(ii) If $p = 1$, then w^u is optimal for (3.2) *for $\kappa > \kappa^*$, where the threshold κ^* can be stated explicitly contingent on the first moment of $\|X^{\widehat{Q}}\|_1$.*

(iii) If $p = 2$, then the κ_ε from (i) can be stated explicitly contingent on the second moment of $\|X^{\widehat{Q}}\|_2$.

The authors give two examples of risk measures $\mathcal{R}$ that satisfy the assumptions of Proposition 3.1. For $p = 1$ one example is the *Conditional Value-at-Risk*, defined for some $\gamma \in (0, 1)$ as

$$\mathrm{CVaR}_\gamma(X) = \frac{1}{1-\gamma} \int_\gamma^1 F_X^{-1}(t) \, \mathrm{d}t.$$

The second example is for the case $p = 2$ the *Markowitz functional*

$$M_\gamma(X) = \mathbb{E}[X] + \gamma\sqrt{\mathrm{var}(X)},$$

where $\gamma > 0$. Hence, in both cases the optimal solution to the robust problem (3.2) will converge to a uniform diversification strategy as the level of uncertainty κ goes to infinity.

3.2. Logarithmic utility maximization

We now want to carry over the methods from Pflug et al. [47], presented in the preceding section, to our continuous-time utility maximization problem for logarithmic utility. To be able to do so we restrict our class of admissible trading strategies to those that are deterministic and define for $h > 0$ the class

$$\widetilde{\mathcal{A}}_h(x_0) = \big\{\pi \in \mathcal{A}_h(x_0) \,\big|\, \pi_t \text{ is deterministic for all } t \in [0,T]\big\}.$$

Here, we only investigate our optimization problem among these deterministic strategies, i.e. we solve the problem

$$\sup_{\pi\in\widetilde{\mathcal{A}}_h(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^\pi)\big], \tag{3.4}$$

where we also assume a specific form of the uncertainty set K. However, in Chapter 4 we follow another approach to solve the same robust optimization problem both for power and logarithmic utility without imposing this preliminary restriction to deterministic strategies. It will turn out that the overall optimal strategy is actually an element of $\widetilde{\mathcal{A}}_h(x_0)$. Hence, our restriction here is one that will not change the value of our optimization problem.

To begin with, we show that within the class of strategies $\widetilde{\mathcal{A}}_h(x_0)$ we can restrict attention to those strategies that are constant in time.

Lemma 3.2. *For any strategy $\pi \in \widetilde{\mathcal{A}}_h(x_0)$ there exists a strategy $\widetilde{\pi} \in \widetilde{\mathcal{A}}_h(x_0)$ that is constant in time, defined by*

$$\widetilde{\pi}_t = \widetilde{\pi}_0 = \frac{1}{T}\int_0^T \pi_s \,\mathrm{d}s$$

for all $t \in [0,T]$, such that

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^\pi)\big] \le \inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^{\widetilde{\pi}})\big].$$

Proof. Let $\pi \in \widetilde{\mathcal{A}}_h(x_0)$ be arbitrary. Define $(\widetilde{\pi}_t)_{t\in[0,T]}$ by

$$\widetilde{\pi}_t = \widetilde{\pi}_0 = \frac{1}{T}\int_0^T \pi_s \,\mathrm{d}s$$

for all $t \in [0,T]$. Then $\widetilde{\pi}$ is by definition constant in time. Since $\pi \in \widetilde{\mathcal{A}}_h(x_0)$ we also obtain

$$\sum_{i=1}^d \widetilde{\pi}_t^i = \sum_{i=1}^d \frac{1}{T}\int_0^T \pi_s^i \,\mathrm{d}s = \frac{1}{T}\int_0^T \sum_{i=1}^d \pi_s^i \,\mathrm{d}s = \frac{1}{T}\int_0^T h \,\mathrm{d}s = h,$$

hence $\widetilde{\pi} \in \widetilde{\mathcal{A}}_h(x_0)$. We now show that in our worst-case context $\widetilde{\pi}$ leads to an objective value that is at least as good as that from using strategy π. Note that

$$\int_0^T \widetilde{\pi}_t \,\mathrm{d}t = T\widetilde{\pi}_0 = \int_0^T \pi_t \,\mathrm{d}t. \tag{3.5}$$

Also, due to Jensen's inequality we have

$$\int_0^T \|\sigma^\top \pi_t\|^2 \,\mathrm{d}t = T\frac{1}{T}\int_0^T \|\sigma^\top \pi_t\|^2 \,\mathrm{d}t \geq T\left\|\frac{1}{T}\int_0^T \sigma^\top \pi_t \,\mathrm{d}t\right\|^2 = \frac{1}{T}\left\|\int_0^T \sigma^\top \pi_t \,\mathrm{d}t\right\|^2,$$

which yields

$$\int_0^T \|\sigma^\top \widetilde{\pi}_t\|^2 \,\mathrm{d}t = T\|\sigma^\top \widetilde{\pi}_0\|^2 = T\left\|\sigma^\top \frac{1}{T}\int_0^T \pi_t \,\mathrm{d}t\right\|^2 = \frac{1}{T}\left\|\int_0^T \sigma^\top \pi_t \,\mathrm{d}t\right\|^2 \leq \int_0^T \|\sigma^\top \pi_t\|^2 \,\mathrm{d}t.$$

By combining the above inequality and (3.5) we thus obtain

$$\int_0^T \Big(\pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \pi_t\|^2\Big)\mathrm{d}t \leq \int_0^T \Big(\widetilde{\pi}_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \widetilde{\pi}_t\|^2\Big)\mathrm{d}t$$

for any $\mu \in K$ and therefore

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^\pi)\big] \leq \inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^{\widetilde{\pi}})\big].$$

Hence, the constant strategy $\widetilde{\pi}$ is at least as good as the strategy π for the worst-case optimization problem. $\square$

The preceding lemma ensures that the optimal strategy for the worst-case optimization problem in (3.4) will be one that is constant in time. For these strategies, we prove a useful representation of the influence that model uncertainty has on the objective when considering a specific uncertainty set K. From now on let

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, \|\mu - \nu\|_p \leq \kappa\big\}$$

for some $\kappa > 0$ and some $p \in [1, \infty]$. Here, $\|\cdot\|_p$ denotes the p-norm on $\mathbb{R}^d$. The set K is simply a ball with radius κ around our reference parameter ν. One can think of ν as an estimation for the drift that might for example be obtained from historical data. By increasing κ we model a higher degree of uncertainty about the true drift. A small value of κ corresponds to high accuracy of our drift estimate.

Proposition 3.3. *Let $p \in [1, \infty]$ and*

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, \|\mu - \nu\|_p \leq \kappa\big\}.$$

For any strategy $\pi \in \widetilde{\mathcal{A}}_h(x_0)$ that is constant in time, i.e. with $\pi_t = \pi_0$ for all $t \in [0, T]$, it holds

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^\pi)\big] = \mathbb{E}_\nu\big[\log(X_T^\pi)\big] - \kappa T\|\pi_0\|_q,$$

where $q \in [1, \infty]$ is such that $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. Let $\mu \in K$ and recall that

$$\mathbb{E}_\mu\big[\log(X_T^\pi)\big] = \log(x_0) + \mathbb{E}_\mu\bigg[\int_0^T \Big(r + \pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \pi_t\|^2\Big)\mathrm{d}t\bigg].$$

Since $\pi \in \widetilde{\mathcal{A}}_h(x_0)$ is deterministic and assumed to be constant in time,

$$\mathbb{E}_\mu\big[\log(X_T^\pi)\big] - \mathbb{E}_\nu\big[\log(X_T^\pi)\big] = T\pi_0^\top(\mu - \nu).$$

From Hölder's inequality we know that

$$\big|\pi_0^\top(\mu - \nu)\big| = \bigg|\sum_{i=1}^d \pi_0^i(\mu^i - \nu^i)\bigg| \le \sum_{i=1}^d \big|\pi_0^i(\mu^i - \nu^i)\big| \le \|\pi_0\|_q\|\mu - \nu\|_p \le \kappa\|\pi_0\|_q, \tag{3.6}$$

where $\frac{1}{p} + \frac{1}{q} = 1$. We can construct a parameter $\mu \in K$ such that the inequality in (3.6) becomes an equality. In the case $p \in (1, \infty)$ let therefore

$$\beta = -\frac{\kappa}{\|\pi_0\|_q^{q/p}}.$$

Note that the norm of π_0 is strictly positive since $\pi_0 = 0$ is not admissible. Then define

$$\mu^i = \begin{cases} \nu^i + \beta(\pi_0^i)^{q/p} & \text{, if } \pi_0^i \ge 0, \\ \nu^i - \beta(-\pi_0^i)^{q/p} & \text{, if } \pi_0^i < 0. \end{cases}$$

Now we easily calculate

$$\begin{aligned} \pi_0^\top(\mu - \nu) &= \sum_{i=1}^d \pi_0^i(\mu^i - \nu^i) = \sum_{i=1}^d |\pi_0^i|\beta|\pi_0^i|^{q/p} \\ &= \beta\sum_{i=1}^d |\pi_0^i|^{1+q/p} = \beta\sum_{i=1}^d |\pi_0^i|^q = \beta\|\pi_0\|_q^q = -\kappa\|\pi_0\|_q. \end{aligned}$$

If $p = 1$, let $j \in \{1, \ldots, d\}$ denote an index with

$$|\pi_0^j| = \max_{i=1,\ldots,d} |\pi_0^i|$$

and define

$$\mu^i = \begin{cases} \nu^j - \kappa\,\mathrm{sgn}(\pi_0^j) & , i = j, \\ \nu^i & , i \ne j. \end{cases}$$

Then also

$$\pi_0^\top(\mu - \nu) = \sum_{i=1}^d \pi_0^i(\mu^i - \nu^i) = -\kappa\,\mathrm{sgn}(\pi_0^j)\pi_0^j = -\kappa \max_{i=1,\ldots,d}|\pi_0^i| = -\kappa\|\pi_0\|_\infty.$$

In the case $p = \infty$ define $\mu \in K$ by

$$\mu^i = \nu^i - \kappa\,\mathrm{sgn}(\pi_0^i)$$

for all $i = 1, \dots, d$. Then

$$\pi_0^\top(\mu - \nu) = \sum_{i=1}^d \pi_0^i(\mu^i - \nu^i) = -\kappa \sum_{i=1}^d |\pi_0^i| = -\kappa \|\pi_0\|_1.$$

In conclusion, there always exists a parameter $\mu \in K$ such that the inequality in (3.6) becomes an equality. Therefore, it follows that

$$\inf_{\mu \in K} \mathbb{E}_\mu\big[\log(X_T^\pi)\big] = \mathbb{E}_\nu\big[\log(X_T^\pi)\big] + T \inf_{\mu \in K} \pi_0^\top(\mu - \nu) = \mathbb{E}_\nu\big[\log(X_T^\pi)\big] - \kappa T \|\pi_0\|_q,$$

and the claim is proven. □

The preceding proposition gives a useful characterization of the worst-case objective. Recall that an investor wants to maximize

$$\inf_{\mu \in K} \mathbb{E}_\mu\big[\log(X_T^\pi)\big]$$

over admissible strategies π. Proposition 3.3 shows that this objective can be split into two antithetic components. The first summand is the expected utility under the reference parameter ν. The second summand is a negative multiple of $\|\pi_0\|_q$. So on the one hand, there is a natural incentive to maximize expected utility under the reference parameter that we think of as an estimation for the true drift. On the other hand, it is beneficial to keep $\|\pi_0\|_q$ as small as possible which suggests to use diversified strategies. What we also see is that the larger κ is, the more important this latter intention becomes. The vector $u \in \mathbb{R}^d$ that minimizes $\|u\|_q$ subject to $\langle u, \mathbf{1}_d \rangle = h$ is $u = \frac{h}{d}\mathbf{1}_d$. This observation suggests that a uniform diversification strategy of the form $\pi^u = (\pi_t^u)_{t \in [0,T]}$ with

$$\pi_t^u = \frac{h}{d}\mathbf{1}_d$$

for all $t \in [0, T]$ might be a good choice for an investor who faces a high degree of model uncertainty. This conjecture is in line with the results in Pflug et al. [47] for a one-period setting and with the work by DeMiguel et al. [17] who show that the uniform diversification strategy as a benchmark outperforms other more involved strategies in terms of various performance criteria. In the following we analyze the performance of the uniform diversification strategy π^u in our continuous-time setting. From now on, let $p \in [1, \infty)$ and $q \in (1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$.

Theorem 3.4. *For any $\kappa > 0$ let $\pi^*(\kappa)$ denote an optimal strategy for the constrained worst-case optimization problem (3.4) with $K = \{\mu \in \mathbb{R}^d \,|\, \|\mu - \nu\|_p \leq \kappa\}$. Then for every $\varepsilon > 0$ there exists a $\kappa_0 > 0$ such that for all $\kappa \geq \kappa_0$ the strategy $\pi^*(\kappa)$ satisfies*

$$\left\| \frac{1}{T} \int_0^T \big(\pi_s^*(\kappa) - \pi_0^u\big)\,\mathrm{d}s \right\|_q < \varepsilon.$$

In particular, if an optimal strategy exists, then there is an optimal strategy that is constant in time and any such strategy satisfies

$$\|\pi_0^*(\kappa) - \pi_0^u\|_q < \varepsilon$$

for all $\kappa \geq \kappa_0$.

Proof. The proof goes along the lines of the proof in Pflug et al. [47, Prop. 3]. We define for each $n \in \mathbb{N}$ the set

$$A_n := \Big\{ \pi \in \mathbb{R}^d \,\Big|\, \langle \pi, \mathbf{1}_d \rangle = h \text{ and } \mathbb{E}_\nu\big[\log(X_T^{\pi^u})\big] - nT\|\pi_0^u\|_q \leq \mathbb{E}_\nu\big[\log(X_T^{\pi})\big] - nT\|\pi\|_q \Big\}.$$

By X_T^π we here mean terminal wealth when using the strategy $(\pi_t)_{t\in[0,T]}$ with $\pi_t = \pi$ for all $t \in [0,T]$. This constant strategy is then admissible in the sense of $\widetilde{\mathcal{A}}_h(x_0)$. Since $\|\pi_0^u\|_q \leq \|\pi\|_q$ for each $\pi \in \mathbb{R}^d$ with $\langle \pi, \mathbf{1}_d \rangle = h$ it is easy to see that $A_{n+1} \subseteq A_n$ for each $n \in \mathbb{N}$. Also, it follows that

$$\bigcap_{n=1}^{\infty} A_n = \{\pi_0^u\}$$

since $\|\pi_0^u\|_q < \|\pi\|_q$ for $q \in (1,\infty]$ if $\pi \neq \pi_0^u$. Furthermore, the sets A_n are bounded since for each $\pi \in A_n$ it holds

$$nT\|\pi\|_q \leq \mathbb{E}_\nu\big[\log(X_T^{\pi})\big] - \mathbb{E}_\nu\big[\log(X_T^{\pi^u})\big] + nT\|\pi_0^u\|_q,$$

and the term $\mathbb{E}_\nu[\log(X_T^\pi)]$ is bounded from above by the finite value that is attained in the corresponding unconstrained portfolio optimization problem. Due to continuity of

$$\mathbb{E}_\nu\big[\log(X_T^{\pi})\big] - nT\|\pi\|_q$$

in π, the sets A_n are also closed. Hence, as subsets of $\mathbb{R}^d$, they are compact.

Now let $\varepsilon > 0$. Then the sets

$$B_n^\varepsilon := A_n \setminus \big\{ \pi \in \mathbb{R}^d \,\big|\, \|\pi - \pi_0^u\|_q < \varepsilon \big\}$$

are also compact with $B_{n+1}^\varepsilon \subseteq B_n^\varepsilon$ for each $n \in \mathbb{N}$ and

$$\bigcap_{n=1}^{\infty} B_n^\varepsilon = \varnothing.$$

Hence, there exists an $M_\varepsilon \in \mathbb{N}$ such that

$$\bigcap_{n=1}^{M_\varepsilon} B_n^\varepsilon = \varnothing.$$

On the other hand, however,

$$\bigcap_{n=1}^{M_\varepsilon} B_n^\varepsilon = \bigcap_{n=1}^{M_\varepsilon} A_n \setminus \big\{ \pi \in \mathbb{R}^d \,\big|\, \|\pi - \pi_0^u\|_q < \varepsilon \big\} = A_{M_\varepsilon} \setminus \big\{ \pi \in \mathbb{R}^d \,\big|\, \|\pi - \pi_0^u\|_q < \varepsilon \big\}.$$

So it follows that

$$A_{M_\varepsilon} \subseteq \big\{ \pi \in \mathbb{R}^d \,\big|\, \|\pi - \pi_0^u\|_q < \varepsilon \big\}.$$

Now let $\kappa \geq M_\varepsilon$. Then, if $\|\pi - \pi_0^u\|_q \geq \varepsilon$ for some $\pi \in \mathbb{R}^d$ with $\langle \pi, \mathbf{1}_d \rangle = h$, we know $\pi \notin A_{M_\varepsilon}$, hence in particular

$$\mathbb{E}_\nu\big[\log(X_T^{\pi^u})\big] - \kappa T\|\pi_0^u\|_q > \mathbb{E}_\nu\big[\log(X_T^{\pi})\big] - \kappa T\|\pi\|_q.$$

But then it follows from Proposition 3.3 that the strategy $(\pi_t)_{t\in[0,T]}$ with $\pi_t = \pi$ for all $t \in [0,T]$ leads to a worse expected utility than the uniform investment strategy π^u and cannot be optimal.

Now let $(\pi_t)_{t\in[0,T]} \in \widetilde{\mathcal{A}}_h(x_0)$ be an arbitrary strategy. From Lemma 3.2 we know that the constant strategy $(\widetilde{\pi}_t)_{t\in[0,T]}$ with

$$\widetilde{\pi}_t = \widetilde{\pi}_0 = \frac{1}{T}\int_0^T \pi_s \,\mathrm{d}s$$

for all $t \in [0,T]$ satisfies

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^\pi)\big] \le \inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^{\widetilde{\pi}})\big].$$

By Proposition 3.3, the right-hand side is equal to

$$\mathbb{E}_\nu\big[\log(X_T^{\widetilde{\pi}})\big] - \kappa T \|\widetilde{\pi}_0\|_q.$$

We have shown above that this expression is strictly smaller than

$$\mathbb{E}_\nu\big[\log(X_T^{\pi^u})\big] - \kappa T \|\pi_0^u\|_q = \inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^{\pi^u})\big]$$

for all $\kappa \ge M_\varepsilon$ if $\|\widetilde{\pi}_0 - \pi_0^u\|_q \ge \varepsilon$. So, in that case, the strategy $(\pi_t)_{t\in[0,T]}$ cannot be optimal. Therefore, if $\pi^*(\kappa)$ denotes an optimal strategy for the optimization problem with degree of uncertainty κ we must have

$$\left\| \frac{1}{T}\int_0^T \big(\pi_s^*(\kappa) - \pi_0^u\big)\,\mathrm{d}s \right\|_q = \left\| \frac{1}{T}\int_0^T \pi_s^*(\kappa)\,\mathrm{d}s - \pi_0^u \right\|_q < \varepsilon$$

for $\kappa \ge M_\varepsilon$. □

The previous theorem shows that any optimal strategy for the constrained worst-case optimization problem (3.4) with $K = \{\mu \in \mathbb{R}^d \,|\, \|\mu - \nu\|_p \le \kappa\}$ for $p \in [1,\infty)$ converges, as model uncertainty increases, to the uniform diversification strategy $\pi^u = (\pi_t^u)_{t\in[0,T]}$ with

$$\pi_t^u = \frac{h}{d}\mathbf{1}_d.$$

This implies that, as uncertainty about the true drift parameter goes to infinity, investors who are forced to invest by the constraint $\langle \pi_t, \mathbf{1}_d\rangle = h$ for all $t \in [0,T]$ will split the proportion h of their money more and more evenly among all risky assets. This uniform diversification is hence a robust response to model uncertainty. The above convergence result holds for all $p \in [1,\infty)$. For $p = 1$ and $p = 2$ we can show even stronger results. This is done in the following subsections. Both for $p = 1$ and $p = 2$ we need the following lemma that gives an estimation for the influence that the choice of strategy has on the objective, given that one optimizes under the reference measure $\mathbb{P} = \mathbb{P}^\nu$.

Lemma 3.5. *There exists a constant $C_p > 0$ such that*

$$\mathbb{E}_\nu\big[\log(X_T^\pi) - \log(X_T^{\pi^u})\big] \le C_p \|\pi_0 - \pi_0^u\|_q$$

for all strategies $\pi \in \widetilde{\mathcal{A}}_h(x_0)$ that are constant in time.

Proof. Observe that since both π and π^u are deterministic and constant in time we can rewrite

$$\begin{aligned}
&\mathbb{E}_\nu\big[\log(X_T^\pi) - \log(X_T^{\pi^u})\big] \\
&= T\Big(\pi_0^\top(\nu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_0\|^2\Big) - T\Big((\pi_0^u)^\top(\nu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_0^u\|^2\Big) \\
&= T(\pi_0 - \pi_0^u)^\top(\nu - r\mathbf{1}_d) - \frac{T}{2}\big(\|\sigma^\top\pi_0\|^2 - \|\sigma^\top\pi_0^u\|^2\big) \\
&= T(\pi_0 - \pi_0^u)^\top(\nu - r\mathbf{1}_d) - \frac{T}{2}\big(\|\sigma^\top\pi_0\| + \|\sigma^\top\pi_0^u\|\big)\big(\|\sigma^\top\pi_0\| - \|\sigma^\top\pi_0^u\|\big).
\end{aligned}$$

If π is such that $\|\sigma^\top\pi_0\| \geq \|\sigma^\top\pi_0^u\|$, then clearly the right-hand side of the above equation is bounded by

$$T(\pi_0 - \pi_0^u)^\top(\nu - r\mathbf{1}_d) \leq T\|\pi_0 - \pi_0^u\|_q\|\nu - r\mathbf{1}_d\|_p \tag{3.7}$$

by Hölder's inequality. If on the other hand $\|\sigma^\top\pi_0\| < \|\sigma^\top\pi_0^u\|$, then

$$\begin{aligned}
\Big|\frac{T}{2}\big(\|\sigma^\top\pi_0\| + \|\sigma^\top\pi_0^u\|\big)\big(\|\sigma^\top\pi_0\| - \|\sigma^\top\pi_0^u\|\big)\Big| &\leq T\|\sigma^\top\pi_0^u\|\,\big|\|\sigma^\top\pi_0\| - \|\sigma^\top\pi_0^u\|\big| \\
&\leq T\|\sigma^\top\pi_0^u\|\,\|\sigma^\top(\pi_0 - \pi_0^u)\|.
\end{aligned} \tag{3.8}$$

In the second step we have used the reverse triangle inequality. Using submultiplicativity of the spectral norm we get

$$\|\sigma^\top(\pi_0 - \pi_0^u)\| \leq \|\sigma\|\,\|\pi_0 - \pi_0^u\|.$$

Due to the equivalence of the Euclidean norm and the q-norm there exists a constant $C_{2,q} > 0$ such that

$$\|\pi_0 - \pi_0^u\| \leq C_{2,q}\|\pi_0 - \pi_0^u\|_q.$$

Plugging these estimations back into (3.8) yields

$$\Big|\frac{T}{2}\big(\|\sigma^\top\pi_0\| + \|\sigma^\top\pi_0^u\|\big)\big(\|\sigma^\top\pi_0\| - \|\sigma^\top\pi_0^u\|\big)\Big| \leq T\|\sigma^\top\pi_0^u\|\,\|\sigma\|\,C_{2,q}\|\pi_0 - \pi_0^u\|_q.$$

Hence,

$$\begin{aligned}
\mathbb{E}_\nu\big[\log(X_T^\pi) - \log(X_T^{\pi^u})\big] &\leq T(\pi_0 - \pi_0^u)^\top(\nu - r\mathbf{1}_d) + T\|\sigma^\top\pi_0^u\|\,\|\sigma\|\,C_{2,q}\|\pi_0 - \pi_0^u\|_q \\
&\leq T\|\pi_0 - \pi_0^u\|_q\|\nu - r\mathbf{1}_d\|_p + T\|\sigma^\top\pi_0^u\|\,\|\sigma\|\,C_{2,q}\|\pi_0 - \pi_0^u\|_q \\
&= T\Big(\|\nu - r\mathbf{1}_d\|_p + C_{2,q}\|\sigma\|\,\|\sigma^\top\pi_0^u\|\Big)\|\pi_0 - \pi_0^u\|_q.
\end{aligned}$$

When comparing this upper bound with the one in (3.7) we see that for all $\pi \in \widetilde{\mathcal{A}}_h(x_0)$ that are constant in time we obtain

$$\mathbb{E}_\nu\big[\log(X_T^\pi) - \log(X_T^{\pi^u})\big] \leq C_p\|\pi_0 - \pi_0^u\|_q,$$

where

$$C_p = T\Big(\|\nu - r\mathbf{1}_d\|_p + C_{2,q}\|\sigma\|\,\|\sigma^\top\pi_0^u\|\Big)$$

is a positive constant that does not depend on π. $\square$

3.2.1. Special case: uncertainty ball in 1-norm

With the help of Lemma 3.5 we now prove a result for the special case where

$$K = \{\mu \in \mathbb{R}^d \mid \|\mu - \nu\|_1 \le \kappa\},$$

i.e. for $p = 1$. We show below that there exists some threshold κ^* such that π^u will be optimal for the optimization problem (3.4) if $\kappa \ge \kappa^*$. The proof relies on the following relation.

Lemma 3.6. *Let $u \in \mathbb{R}^d$ be some vector with $\langle u, \mathbf{1}_d\rangle = h$ and $v = \frac{h}{d}\mathbf{1}_d$. Then*

$$\|u - v\|_\infty \le (d-1)\big(\|u\|_\infty - \|v\|_\infty\big).$$

Proof. The statement is already proven in Pflug et al. [47, Prop. 3]. We repeat the proof here for completeness.

Note that

$$\|u - v\|_\infty = \max_{i=1,\dots,d}\Big|u^i - \frac{h}{d}\Big| = \Big|u^j - \frac{h}{d}\Big|$$

for some $j \in \{1,\dots,d\}$. If $u^j \ge \frac{h}{d}$, then $u^j = \max_{i=1,\dots,d}|u^i| = \|u\|_\infty$ and

$$\|u - v\|_\infty = u^j - \frac{h}{d} = \|u\|_\infty - \|v\|_\infty \le (d-1)\big(\|u\|_\infty - \|v\|_\infty\big).$$

If $u^j < \frac{h}{d}$, then $u^j = \min_{i=1,\dots,d} u^i$ and

$$(d-1)\max_{i=1,\dots,d} u^i + u^j \ge \sum_{i=1}^d u^i = h,$$

hence

$$\max_{i=1,\dots,d} u^i - \frac{h}{d} \ge \frac{h-u^j}{d-1} - \frac{h}{d} = \frac{dh - du^j - (d-1)h}{d(d-1)} = \frac{\frac{h}{d} - u^j}{d-1}.$$

It follows that

$$\|u - v\|_\infty = \frac{h}{d} - u^j \le (d-1)\Big(\max_{i=1,\dots,d} u^i - \frac{h}{d}\Big) \le (d-1)\big(\|u\|_\infty - \|v\|_\infty\big).$$

Hence, in both cases the claim is proven. □

Now we can show optimality of π^u for high degree of model uncertainty κ.

Proposition 3.7. *Let $p = 1$, i.e.*

$$K = \{\mu \in \mathbb{R}^d \mid \|\mu - \nu\|_1 \le \kappa\}.$$

Then $\pi^u = (\pi^u_t)_{t\in[0,T]}$ is optimal for the optimization problem (3.4) if $\kappa \ge \kappa^*$, where

$$\kappa^* = \frac{C_1}{T}(d-1)$$

and $C_1 > 0$ is the constant from Lemma 3.5 for $p = 1$.

Proof. By combining the results from Lemma 3.5 for $p = 1$, $q = \infty$ and from Lemma 3.6 we deduce that for any strategy $\pi \in \widetilde{\mathcal{A}}_h(x_0)$ that is constant in time it holds

$$\mathbb{E}_\nu\big[\log(X_T^\pi) - \log(X_T^{\pi^u})\big] \leq C_1\|\pi_0 - \pi_0^u\|_\infty \leq C_1(d-1)\big(\|\pi_0\|_\infty - \|\pi_0^u\|_\infty\big).$$

So if $\kappa \geq \kappa^* = \frac{C_1}{T}(d-1)$ then

$$\mathbb{E}_\nu\big[\log(X_T^\pi) - \log(X_T^{\pi^u})\big] \leq \kappa T\big(\|\pi_0\|_\infty - \|\pi_0^u\|_\infty\big),$$

hence by Proposition 3.3 it holds

$$\begin{aligned}\inf_{\mu\in K}\mathbb{E}_\mu\big[\log(X_T^\pi)\big] &= \mathbb{E}_\nu\big[\log(X_T^\pi)\big] - \kappa T\|\pi_0\|_\infty \\ &\leq \mathbb{E}_\nu\big[\log(X_T^{\pi^u})\big] - \kappa T\|\pi_0^u\|_\infty = \inf_{\mu\in K}\mathbb{E}_\mu\big[\log(X_T^{\pi^u})\big].\end{aligned}$$

We can conclude that π^u is at least as good as the strategy π. Since this holds for any $\pi \in \widetilde{\mathcal{A}}_h(x_0)$ that is constant in time, it follows from Lemma 3.2 that π^u is optimal among all strategies in $\widetilde{\mathcal{A}}_h(x_0)$. □

The preceding proposition shows that in the case $p = 1$, i.e. for an uncertainty ball of the form

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, \|\mu - \nu\|_1 \leq \kappa\big\},$$

we do not only have convergence of the optimal solution to π^u, as shown in Theorem 3.4. The strategy π^u is even optimal as soon as the degree of uncertainty κ exceeds a certain threshold. As can be seen from the proof of Lemma 3.5, this threshold κ^* can be computed explicitly, given that all model parameters are known.

3.2.2. Special case: uncertainty ball in 2-norm

Another special case is an uncertainty ball of the form

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, \|\mu - \nu\|_2 \leq \kappa\big\},$$

i.e. $p = 2$. In terms of the interpretation of the uncertainty set this is perhaps the most natural model for uncertainty about the drift parameter. In this case as well we can prove a stronger result than that in Theorem 3.4. Firstly, we state the following technical lemma.

Lemma 3.8. *Let $v = \frac{h}{d}\mathbf{1}_d$ and $u \in \mathbb{R}^d$, $u \neq v$, some vector with $\langle u, \mathbf{1}_d\rangle = h$. Then*

$$\frac{\|u - v\|_2}{\|u\|_2 - \|v\|_2} = \left(\frac{h^2}{d\|u - v\|_2^2} + 1\right)^{1/2} + \frac{h}{\sqrt{d}\|u - v\|_2}.$$

Proof. The proof goes along the lines of Pflug et al. [47, Prop. 3]. Firstly, observe that if $\langle u, \mathbf{1}_d\rangle = h$ and $u \neq v$ then $\|u\|_2 > \|v\|_2$. Let $v_1 = v$ and extend to an orthogonal basis $(v_1, v_2, \ldots, v_d)$ of $\mathbb{R}^d$ where $\|v_i\|_2 = 1$ for all $i = 2, \ldots, d$. Then we can write u as

$$u = v + \sum_{i=2}^d c_i v_i$$

for some $c_2, \dots, c_d \in \mathbb{R}$, where we have used $\langle u, \mathbf{1}_d \rangle = h$. Hence,

$$\frac{\|u-v\|_2}{\|u\|_2 - \|v\|_2} = \frac{\|u-v\|_2}{\left(\frac{h^2}{d} + \sum_{i=2}^d c_i^2\right)^{1/2} - \frac{h}{\sqrt{d}}} = \frac{\|u-v\|_2}{\left(\frac{h^2}{d} + \|u-v\|_2^2\right)^{1/2} - \frac{h}{\sqrt{d}}}.$$

When using the third binomial formula we see that this expression is equal to

$$\frac{\|u-v\|_2\Big(\left(\frac{h^2}{d} + \|u-v\|_2^2\right)^{1/2} + \frac{h}{\sqrt{d}}\Big)}{\frac{h^2}{d} + \|u-v\|_2^2 - \frac{h^2}{d}} = \bigg(\frac{h^2}{d\|u-v\|_2^2} + 1\bigg)^{1/2} + \frac{h}{\sqrt{d}\|u-v\|_2},$$

which proves the claim. ☐

This lemma can be used to show that for the case $p = 2$ the optimal strategy gets arbitrarily close to the uniform diversification strategy for a high degree of model uncertainty. In addition to the general convergence result of Theorem 3.4 we also gain insights into the threshold value κ^* for the uncertainty radius.

Proposition 3.9. *Let $p = 2$, i.e. $K = \{\mu \in \mathbb{R}^d \,|\, \|\mu - \nu\|_2 \le \kappa\}$. Then for any $\varepsilon > 0$ an optimal solution $\pi^*(\kappa)$ to the optimization problem* (3.4) *satisfies*

$$\bigg\|\frac{1}{T}\int_0^T \left(\pi_s^*(\kappa) - \pi_0^u\right) \mathrm{d}s\bigg\|_2 < \varepsilon$$

if $\kappa > \kappa^*$, where

$$\kappa^* = \frac{C_2}{T}\bigg(\Big(\frac{h^2}{d\varepsilon^2} + 1\Big)^{1/2} + \frac{h}{\sqrt{d}\varepsilon}\bigg),$$

and where $C_2 > 0$ is the constant from Lemma 3.5 for $p = 2$. In particular, if an optimal strategy exists, then there is an optimal strategy that is constant in time and any such strategy satisfies

$$\|\pi_0^*(\kappa) - \pi_0^u\|_2 < \varepsilon$$

for all $\kappa > \kappa^*$.

Proof. Let $\varepsilon > 0$. Let $v = \frac{h}{d}\mathbf{1}_d$ and $u \in \mathbb{R}^d$ with $\langle u, \mathbf{1}_d \rangle = h$ and $\|u - v\|_2 \ge \varepsilon$. Then it follows from Lemma 3.8 that

$$\|u-v\|_2 \le \bigg(\Big(\frac{h^2}{d\varepsilon^2} + 1\Big)^{1/2} + \frac{h}{\sqrt{d}\varepsilon}\bigg)\big(\|u\|_2 - \|v\|_2\big). \tag{3.9}$$

Now let $\pi \in \widetilde{\mathcal{A}}_h(x_0)$ be an arbitrary admissible strategy. From Lemma 3.2 we know that the constant strategy $(\widetilde{\pi}_t)_{t\in[0,T]} \in \widetilde{\mathcal{A}}_h(x_0)$, defined by

$$\widetilde{\pi}_t = \widetilde{\pi}_0 = \frac{1}{T}\int_0^T \pi_s \,\mathrm{d}s$$

for all $t \in [0, T]$, satisfies

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^\pi)\big] \le \inf_{\mu\in K} \mathbb{E}_\mu\big[\log(X_T^{\widetilde{\pi}})\big].$$

Since $\widetilde{\pi}$ is constant, we can conclude from Lemma 3.5 for $p = q = 2$ together with (3.9) that if $\|\widetilde{\pi}_0 - \pi_0^u\|_2 \geq \varepsilon$, then

$$\mathbb{E}_\nu\big[\log(X_T^{\widetilde{\pi}}) - \log(X_T^{\pi^u})\big] \leq C_2\|\widetilde{\pi}_0 - \pi_0^u\|_2 \leq C_2\bigg(\Big(\frac{h^2}{d\varepsilon^2}+1\Big)^{1/2} + \frac{h}{\sqrt{d}\varepsilon}\bigg)\big(\|\widetilde{\pi}_0\|_2 - \|\pi_0^u\|_2\big).$$

So if $\kappa > \kappa^*$, where

$$\kappa^* = \frac{C_2}{T}\bigg(\Big(\frac{h^2}{d\varepsilon^2}+1\Big)^{1/2} + \frac{h}{\sqrt{d}\varepsilon}\bigg),$$

then

$$\mathbb{E}_\nu\big[\log(X_T^{\widetilde{\pi}}) - \log(X_T^{\pi^u})\big] < \kappa T\big(\|\widetilde{\pi}_0\|_2 - \|\pi_0^u\|_2\big).$$

Then by Proposition 3.3 it follows that

$$\inf_{\mu\in K}\mathbb{E}_\mu\big[\log(X_T^{\widetilde{\pi}})\big] = \mathbb{E}_\nu\big[\log(X_T^{\widetilde{\pi}})\big] - \kappa T\|\widetilde{\pi}_0\|_2 < \mathbb{E}_\nu[\log(X_T^{\pi^u})] - \kappa T\|\pi_0^u\|_2 = \inf_{\mu\in K}\mathbb{E}_\mu\big[\log(X_T^{\pi^u})\big].$$

Therefore, neither π itself nor $\widetilde{\pi}$ are optimal if $\kappa > \kappa^*$. Consequently, any optimal strategy $\pi^*(\kappa)$ must fulfill

$$\bigg\|\frac{1}{T}\int_0^T \big(\pi_s^*(\kappa) - \pi_0^u\big)\,\mathrm{d}s\bigg\|_2 = \bigg\|\frac{1}{T}\int_0^T \pi_s^*(\kappa)\,\mathrm{d}s - \pi_0^u\bigg\|_2 < \varepsilon$$

if $\kappa > \kappa^*$. □

The preceding proposition shows convergence of the optimal strategy to the uniform diversification strategy $(\pi_t^u)_{t\in[0,T]}$ for the uncertainty set $K = \{\mu \in \mathbb{R}^d \mid \|\mu - \nu\|_2 \leq \kappa\}$. This result is already stated in Theorem 3.4 in a more general form. The additional benefit in Proposition 3.9 over Theorem 3.4 lies in the explicit form of the threshold value κ^*. Vice versa, if a certain level of uncertainty is assumed for the model, investors can investigate how far from the optimum they will end up when using the uniform diversification strategy π^u. The larger the value for κ, the closer the optimal strategy will be to the uniform one.

4. A Duality Approach

In the preceding chapter we have seen how methods from a one-period risk minimization problem can be carried over to our continuous-time robust utility maximization problem. However, this approach has several drawbacks. Firstly, it is restricted to logarithmic utility and to uncertainty sets K that are balls in some p-norm. Already when using power utility functions one is confronted with various problems that make it more difficult to carry over the results immediately. Secondly, we have seen that we have to restrict to the class of deterministic strategies to be able to use the methods from Pflug et al. [47]. However, it is by no means clear in the first place that an optimal strategy to our worst-case optimization problem should be a deterministic one. In fact, in many worst-case optimization problems it is even beneficial to use randomized strategies, see for example Delage et al. [15]. And lastly, the results in general do not yield the solution to the robust optimization problems in explicit form. They only give asymptotic results for large levels of uncertainty, but for a fixed level κ it is not clear what a good strategy for the optimization problem looks like. To overcome these problems we follow here a different approach that works for both power and logarithmic utility and that results in an explicit solution of the optimization problem.

Throughout this chapter, we denote with $U_\gamma\colon \mathbb{R}_+ \to \mathbb{R}$, $\gamma \in (-\infty, 1)$, the power utility function $U_\gamma(x) = \frac{x^\gamma}{\gamma}$ if $\gamma \neq 0$, and the logarithmic utility function $U_0(x) = \log(x)$ if $\gamma = 0$.

4.1. Ellipsoidal uncertainty sets

Recall that our robust constrained utility maximization problem reads

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \inf_{\mu \in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big], \tag{4.1}$$

where the admissibility set is given by

$$\mathcal{A}_h(x_0) = \big\{\pi \in \mathcal{A}(x_0) \,\big|\, \langle \pi_t, \mathbf{1}_d\rangle = h \text{ for all } t \in [0,T]\big\}$$

for some $h > 0$. In this section we study the case where the uncertainty set K is an ellipsoid in $\mathbb{R}^d$ that is centered around some reference parameter. We therefore consider

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, (\mu - \nu)^\top \Gamma^{-1}(\mu - \nu) \leq \kappa^2\big\}.$$

Here, $\kappa > 0$, $\nu \in \mathbb{R}^d$ and $\Gamma \in \mathbb{R}^{d\times d}$ is symmetric and positive definite. Hence, K is centered around ν and shaped by Γ. For $\Gamma = I_d$ we simply get a ball in Euclidean norm that has radius κ and center ν. By means of Γ, however, we can model that some (linear combinations of) drifts are known at a higher degree of accuracy than others. One interesting special case that is discussed in the literature is $\Gamma = \sigma\sigma^\top$, see for example Biagini and Pınar [4]. The value of κ determines the size of the ellipsoid. Higher values of κ correspond to a higher degree of uncertainty about the true drift parameter.

4.1.1. Solution of the non-robust problem

To solve the optimization problem in (4.1) we first address the non-robust constrained utility maximization problem under a fixed parameter $\mu \in \mathbb{R}^d$. We repeatedly make use of a specific matrix that we introduce in the following lemma.

Lemma 4.1. *Define the matrix $D \in \mathbb{R}^{(d-1)\times d}$ as*

$$D = \begin{pmatrix} 1 & & 0 & -1 \\ & \ddots & & \vdots \\ 0 & & 1 & -1 \end{pmatrix} \in \mathbb{R}^{(d-1)\times d}.$$

Then, given that $\sigma \in \mathbb{R}^{d\times m}$ has rank d, $D\sigma$ has rank $d-1$.

Proof. Since $d \leq m$ and $\sigma \in \mathbb{R}^{d\times m}$ has rank d, the rows of σ are independent vectors in $\mathbb{R}^m$. Now $D\sigma \in \mathbb{R}^{(d-1)\times m}$ and due to the specific form of D, the i-th row of $D\sigma$ is $\sigma_{i,\cdot} - \sigma_{d,\cdot}$, $i = 1, \dots, d-1$. Here, $\sigma_{i,\cdot}$ denotes the i-th row of matrix σ. Now from the independence of $\sigma_{1,\cdot}, \dots, \sigma_{d,\cdot}$ it follows for any $a_1, \dots, a_{d-1} \in \mathbb{R}$ that if

$$0 = \sum_{i=1}^{d-1} a_i(\sigma_{i,\cdot} - \sigma_{d,\cdot}) = \sum_{i=1}^{d-1} a_i\sigma_{i,\cdot} - \sum_{i=1}^{d-1} a_i\sigma_{d,\cdot},$$

then $a_1 = \dots = a_{d-1} = 0$. Hence, the rows of $D\sigma$ are independent, and therefore it holds $\operatorname{rank}(D\sigma) = d-1$. □

The matrix D defined in the lemma above comes up naturally in the following calculations when using the constraint $\langle \pi_t, \mathbf{1}_d \rangle = h$ in the form

$$\pi_t^d = h - \sum_{i=1}^{d-1} \pi_t^i.$$

This can be seen as a reduction of the problem from d dimensions to $d-1$ dimensions. Of course, one could just as well write any other component of π_t as a function of the remaining components. In the calculations below, certain matrix-vector products permanently appear. For the sake of better readability it makes sense to introduce the following notation.

Definition 4.2. We define the matrix $A \in \mathbb{R}^{d\times d}$ and the vector $c \in \mathbb{R}^d$ by

$$\begin{aligned} A &= D^\top(D\sigma\sigma^\top D^\top)^{-1}D, \\ c &= e_d - D^\top(D\sigma\sigma^\top D^\top)^{-1}D\sigma\sigma^\top e_d = (I_d - A\sigma\sigma^\top)e_d, \end{aligned}$$

where $D \in \mathbb{R}^{(d-1)\times d}$ is the matrix from Lemma 4.1 and e_d is the d-th standard unit vector in $\mathbb{R}^d$.

Note that we assume $\sigma \in \mathbb{R}^{d\times m}$ to have full rank, hence by the previous lemma we know that $D\sigma$ has full rank, in particular $D\sigma\sigma^\top D^\top = D\sigma(D\sigma)^\top$ is nonsingular. Using this notation we now give the optimal strategy for the non-robust constrained optimization problem, i.e. given a fixed drift μ.

Proposition 4.3. *Let $\mu \in \mathbb{R}^d$. Then the optimal strategy for the optimization problem*

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

is the strategy $(\pi_t)_{t\in[0,T]}$ with

$$\pi_t = \frac{1}{1-\gamma} A\mu + hc$$

for all $t \in [0,T]$, with A and c as in Definition 4.2.

Proof. Let $(\pi_t)_{t\in[0,T]} \in \mathcal{A}_h(x_0)$ be an arbitrary strategy. Then $\pi_t^d = h - \sum_{i=1}^{d-1} \pi_t^i$ for all $t \in [0,T]$. Recall that the terminal wealth under strategy π can be written as

$$X_T^\pi = x_0 \exp\biggl(rT + \int_0^T \Bigl(\pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \pi_t\|^2\Bigr)\mathrm{d}t + \int_0^T \pi_t^\top \sigma \,\mathrm{d}W_t^\mu\biggr).$$

Now note that

$$\begin{aligned}\pi_t^\top(\mu - r\mathbf{1}_d) &= \sum_{i=1}^{d-1} \pi_t^i(\mu^i - r) + \biggl(h - \sum_{i=1}^{d-1}\pi_t^i\biggr)(\mu^d - r) \\ &= h(\mu^d - r) + \sum_{i=1}^{d-1}\pi_t^i(\mu^i - \mu^d) = h(e_d^\top \mu - r) + \widetilde{\pi}_t^\top D\mu,\end{aligned} \tag{4.2}$$

where $\widetilde{\pi}_t := \pi_t^{1:d-1}$ for all $t \in [0,T]$. With the same notation we can also rewrite

$$\pi_t^\top \sigma = \sum_{i=1}^{d-1}\pi_t^i \sigma_{i,\cdot} + \biggl(h - \sum_{i=1}^{d-1}\pi_t^i\biggr)\sigma_{d,\cdot} = h\sigma_{d,\cdot} + \sum_{i=1}^{d-1}\pi_t^i(\sigma_{i,\cdot} - \sigma_{d,\cdot}) = he_d^\top\sigma + \widetilde{\pi}_t^\top D\sigma, \tag{4.3}$$

where $\sigma_{i,\cdot}$ denotes the i-th row of matrix σ.

In the case $\gamma \neq 0$ we now apply the power function to terminal wealth and get

$$\mathbb{E}_\mu\big[(X_T^\pi)^\gamma\big] = x_0^\gamma e^{\gamma r T}\,\mathbb{E}_\mu\biggl[\exp\biggl(\gamma\int_0^T\Bigl(\pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_t\|^2\Bigr)\mathrm{d}t + \gamma\int_0^T \pi_t^\top\sigma\,\mathrm{d}W_t^\mu\biggr)\biggr]. \tag{4.4}$$

Here, we can plug in (4.3) in the stochastic integral. The integral then splits up into

$$\int_0^T \gamma\pi_t^\top\sigma\,\mathrm{d}W_t^\mu = \int_0^T \gamma h e_d^\top \sigma\,\mathrm{d}W_t^\mu + \int_0^T \gamma\widetilde{\pi}_t^\top D\sigma\,\mathrm{d}W_t^\mu.$$

We then perform a change of measure

$$\frac{\mathrm{d}\widetilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}^\mu} = Z_T = \exp\biggl(\int_0^T \gamma h e_d^\top\sigma\,\mathrm{d}W_t^\mu - \frac{1}{2}\int_0^T \|\gamma h\sigma^\top e_d\|^2\,\mathrm{d}t\biggr).$$

With all these considerations, (4.4) becomes

$$\begin{aligned}\mathbb{E}_\mu\big[(X_T^\pi)^\gamma\big] &= x_0^\gamma e^{\gamma r T}\,\mathbb{E}_\mu\biggl[\exp\biggl(\gamma\int_0^T\Bigl(\pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_t\|^2\Bigr)\mathrm{d}t + \gamma\int_0^T\pi_t^\top\sigma\,\mathrm{d}W_t^\mu\biggr)\biggr] \\ &= x_0^\gamma e^{\gamma r T}\widetilde{\mathbb{E}}\biggl[\exp\biggl(\gamma\int_0^T\Bigl(\pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_t\|^2 + \frac{1}{2}\gamma\|h\sigma^\top e_d\|^2\Bigr)\mathrm{d}t + \int_0^T\gamma\widetilde{\pi}_t^\top D\sigma\,\mathrm{d}W_t^\mu\biggr)\biggr].\end{aligned}$$

Note that, under $\widetilde{\mathbb{P}}$, the process $(\widetilde{W}_t^\mu)_{t\in[0,T]}$ with

$$\widetilde{W}_t^\mu = W_t^\mu - \int_0^t \gamma h\sigma^\top e_d \,\mathrm{d}s$$

is a Brownian motion by Girsanov's Theorem. Hence, we substitute

$$\int_0^T \gamma\widetilde{\pi}_t^\top D\sigma \,\mathrm{d}W_t^\mu = \int_0^T \gamma\widetilde{\pi}_t^\top D\sigma \,\mathrm{d}\widetilde{W}_t^\mu + \int_0^T \gamma^2 h\widetilde{\pi}_t^\top D\sigma\sigma^\top e_d \,\mathrm{d}t$$

and rearrange to obtain

$$\begin{aligned}
&\gamma\int_0^T \Big(\pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_t\|^2 + \frac{1}{2}\gamma\|h\sigma^\top e_d\|^2\Big)\mathrm{d}t + \int_0^T \gamma\widetilde{\pi}_t^\top D\sigma \,\mathrm{d}W_t^\mu \\
&= \gamma\int_0^T \Big(\pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_t\|^2 + \frac{1}{2}\gamma\|h\sigma^\top e_d\|^2 + \gamma h\widetilde{\pi}_t^\top D\sigma\sigma^\top e_d\Big)\mathrm{d}t + \int_0^T \gamma\widetilde{\pi}_t^\top D\sigma \,\mathrm{d}\widetilde{W}_t^\mu.
\end{aligned}$$

By using (4.2) and (4.3) the integrand in the Lebesgue integral above can be written as

$$\begin{aligned}
&he_d^\top\mu - hr + \widetilde{\pi}_t^\top D\mu - \frac{1}{2}\|h\sigma^\top e_d + (D\sigma)^\top\widetilde{\pi}_t\|^2 + \frac{1}{2}\gamma\|h\sigma^\top e_d\|^2 + \gamma h\widetilde{\pi}_t^\top D\sigma\sigma^\top e_d \\
&= he_d^\top\mu - hr + \widetilde{\pi}_t^\top\big(D\mu + \gamma hD\sigma\sigma^\top e_d\big) - \frac{1}{2}(1-\gamma)\|h\sigma^\top e_d\|^2 - h\widetilde{\pi}_t^\top D\sigma\sigma^\top e_d - \frac{1}{2}\|(D\sigma)^\top\widetilde{\pi}_t\|^2 \\
&= \widetilde{\pi}_t^\top\big(D\mu - h(1-\gamma)D\sigma\sigma^\top e_d\big) - \frac{1}{2}\|(D\sigma)^\top\widetilde{\pi}_t\|^2 + he_d^\top\mu - hr - \frac{1}{2}(1-\gamma)\|h\sigma^\top e_d\|^2.
\end{aligned}$$

If we now substitute

$$\begin{aligned}
\widetilde{\sigma} &= D\sigma, \\
\widetilde{r} &= (1-h)r + he_d^\top\mu - \frac{1}{2}(1-\gamma)\|h\sigma^\top e_d\|^2, \qquad (4.5)\\
\widetilde{\mu} &= D\mu - h(1-\gamma)D\sigma\sigma^\top e_d + \widetilde{r}\mathbf{1}_{d-1},
\end{aligned}$$

then the expected utility of terminal wealth is given by

$$\begin{aligned}
&\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \\
&= \frac{x_0^\gamma}{\gamma}\widetilde{\mathbb{E}}\bigg[\exp\bigg(\gamma\int_0^T\Big(\widetilde{r} + \widetilde{\pi}_t^\top(\widetilde{\mu} - \widetilde{r}\mathbf{1}_{d-1}) - \frac{1}{2}\|\widetilde{\sigma}^\top\widetilde{\pi}_t\|^2\Big)\mathrm{d}t + \gamma\int_0^T \widetilde{\pi}_t^\top\widetilde{\sigma}\,\mathrm{d}\widetilde{W}_t\bigg)\bigg]. \qquad (4.6)
\end{aligned}$$

In the case $\gamma = 0$ we apply the logarithm to terminal wealth and get

$$\mathbb{E}_\mu\big[\log(X_T^\pi)\big] = \log(x_0) + rT + \mathbb{E}_\mu\bigg[\int_0^T\Big(\pi_t^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_t\|^2\Big)\mathrm{d}t\bigg].$$

Like in the case for power utility, we see that we can rewrite this expression as

$$\mathbb{E}_\mu\big[\log(X_T^\pi)\big] = \log(x_0) + \widetilde{r}T + \mathbb{E}\bigg[\int_0^T\Big(\widetilde{\pi}_t^\top\big(\widetilde{\mu} - \widetilde{r}\mathbf{1}_{d-1}\big) - \frac{1}{2}\|\widetilde{\sigma}^\top\widetilde{\pi}_t\|^2\Big)\mathrm{d}t\bigg], \qquad (4.7)$$

where we use the same substitution with $\widetilde{r}$, $\widetilde{\mu}$ and $\widetilde{\sigma}$ as in (4.5) for $\gamma = 0$.

In both cases $\gamma \neq 0$ and $\gamma = 0$ we realize that the expressions in (4.6) and (4.7) are again the expected utility of terminal wealth in a financial market with $d-1$ risky assets where the risk-free interest rate is $\widetilde{r}$, the drift of the $d-1$ risky assets is given by $\widetilde{\mu} \in \mathbb{R}^{d-1}$, and the volatility matrix is $\widetilde{\sigma} \in \mathbb{R}^{(d-1)\times m}$. Note that $\widetilde{\sigma} = D\sigma$ has full row rank equal to $d-1$ by Lemma 4.1, in particular $\widetilde{\sigma}\widetilde{\sigma}^\top$ is nonsingular. So we have reduced the d-dimensional constrained problem to a $(d-1)$-dimensional unconstrained problem. When trying to maximize the right-hand side of (4.6), respectively (4.7), over all admissible strategies $\widetilde{\pi}$ with values in $\mathbb{R}^{d-1}$, we know that the optimal strategy is constant in time and has the form

$$\widetilde{\pi}_t = \frac{1}{1-\gamma}(\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}(\widetilde{\mu} - \widetilde{r}\mathbf{1}_{d-1}) = \frac{1}{1-\gamma}(D\sigma\sigma^\top D^\top)^{-1}\big(D\mu - h(1-\gamma)D\sigma\sigma^\top e_d\big). \tag{4.8}$$

Now note that

$$\pi_t = \sum_{i=1}^{d} \pi_t^i e_i = \sum_{i=1}^{d-1} \pi_t^i e_i + \left(h - \sum_{i=1}^{d-1} \pi_t^i\right) e_d = \sum_{i=1}^{d-1} \pi_t^i (e_i - e_d) + h e_d = D^\top \widetilde{\pi}_t + h e_d.$$

Plugging in the optimal $\widetilde{\pi}_t$ from (4.8) then yields

$$\begin{aligned}
\pi_t &= D^\top \frac{1}{1-\gamma}(D\sigma\sigma^\top D^\top)^{-1}\big(D\mu - h(1-\gamma)D\sigma\sigma^\top e_d\big) + h e_d \\
&= \frac{1}{1-\gamma} D^\top (D\sigma\sigma^\top D^\top)^{-1} D\mu + h\big(I_d - D^\top (D\sigma\sigma^\top D^\top)^{-1} D\sigma\sigma^\top\big) e_d \\
&= \frac{1}{1-\gamma} A\mu + hc
\end{aligned}$$

for all $t \in [0,T]$. □

In the preceding proposition we have calculated the optimal strategy when maximizing utility under the constraint that $\langle \pi_t, \mathbf{1}_d \rangle = h$ for all $t \in [0,T]$, given that the drift parameter $\mu \in \mathbb{R}^d$ is known. In the proof we have seen that the d-dimensional constrained problem could be reduced to a $(d-1)$-dimensional unconstrained problem. This is also useful for determining the optimal expected utility from terminal wealth.

Corollary 4.4. *Let $\mu \in \mathbb{R}^d$. Then the optimal expected utility from terminal wealth is*

$$\begin{aligned}
&\sup_{\pi \in \mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \\
&\quad = \begin{cases} \dfrac{x_0^\gamma}{\gamma} \exp\Big(\gamma T\Big(\widetilde{r} + \dfrac{1}{2(1-\gamma)}\big(\widetilde{\mu} - \widetilde{r}\mathbf{1}_{d-1}\big)^\top (\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}\big(\widetilde{\mu} - \widetilde{r}\mathbf{1}_{d-1}\big)\Big)\Big), & \gamma \neq 0, \\ \log(x_0) + \Big(\widetilde{r} + \dfrac{1}{2}\big(\widetilde{\mu} - \widetilde{r}\mathbf{1}_{d-1}\big)^\top (\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}\big(\widetilde{\mu} - \widetilde{r}\mathbf{1}_{d-1}\big)\Big)T, & \gamma = 0, \end{cases}
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{\sigma} &= D\sigma, \\
\widetilde{r} &= (1-h)r + h e_d^\top \mu - \frac{1}{2}(1-\gamma)\|h\sigma^\top e_d\|^2, \\
\widetilde{\mu} &= D\mu - h(1-\gamma)D\sigma\sigma^\top e_d + \widetilde{r}\mathbf{1}_{d-1}.
\end{aligned} \tag{4.9}$$

Proof. In the proof of Proposition 4.3 we have seen that for any admissible strategy π we can write $\mathbb{E}_\mu[U_\gamma(X_T^\pi)]$ as the expected utility of terminal wealth in a market with $d-1$ risky assets where the constraint on the strategies vanishes and the parameters of the market are given as in (4.9). We have also seen that the optimal strategy in this $(d-1)$-dimensional market fulfills

$$\widetilde{\pi}_t = \frac{1}{1-\gamma}(\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}(\widetilde{\mu} - \widetilde{r}\mathbf{1}_{d-1}).$$

Plugging this optimal strategy in yields the expression from the corollary. $\square$

The previous results give a representation of the optimal strategy and the optimal expected utility of terminal wealth under the constraint $\langle \pi_t, \mathbf{1}_d \rangle = h$ for all $t \in [0,T]$, given that the drift parameter μ is known. Of course, both the strategy and the terminal wealth then depend on μ. However, we aim at solving the robust utility maximization problem

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \inf_{\mu \in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big],$$

where investors maximize their worst-case expected utility among all $\mu \in K$. For that purpose, we address in a next step the question what the worst possible parameter μ would be for the investor, given that she reacts optimally, i.e. by applying the strategy from Proposition 4.3. This corresponds to solving the dual problem

$$\inf_{\mu \in K} \sup_{\pi \in \mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

where, in comparison to our problem, the infimum and supremum are interchanged. Note here that we do not know yet whether the equality

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \inf_{\mu \in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] = \inf_{\mu \in K} \sup_{\pi \in \mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

holds. So, in general the solution of the dual problem may not be of great help as long as we cannot show equality. In the following we will first derive the solution to the dual problem and then prove a minimax theorem that establishes the equality above. This will ensure that the solution to the dual problem also solves our original problem.

4.1.2. The worst-case parameter

From Corollary 4.4 we have a representation of the optimal expected utility of terminal wealth, depending on the transformed parameters $\widetilde{r}$, $\widetilde{\mu}$ and $\widetilde{\sigma}$. Note that due to the representation from the corollary, for any $\gamma \in (-\infty, 1)$, minimizing this expression in μ is equivalent to minimizing

$$\widetilde{r} + \frac{1}{2(1-\gamma)}\big(\widetilde{\mu} - \widetilde{r}\mathbf{1}_{d-1}\big)^\top(\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}\big(\widetilde{\mu} - \widetilde{r}\mathbf{1}_{d-1}\big).$$

We now plug in the representations of $\widetilde{r}$, $\widetilde{\mu}$ and $\widetilde{\sigma}$ from the corollary and obtain

$$\begin{aligned}&(1-h)r + he_d^\top\mu - \frac{1}{2}(1-\gamma)\|h\sigma^\top e_d\|^2 \\ &\qquad + \frac{1}{2(1-\gamma)}\big(D\mu - h(1-\gamma)D\sigma\sigma^\top e_d\big)^\top(D\sigma\sigma^\top D^\top)^{-1}\big(D\mu - h(1-\gamma)D\sigma\sigma^\top e_d\big).\end{aligned}$$

Our aim is to minimize the above expression in μ. We see that many terms do not depend on μ. The minimization is therefore equivalent to the minimization of

$$\begin{aligned}
&he_d^\top \mu + \frac{1}{2(1-\gamma)}\Big(\mu^\top D^\top (D\sigma\sigma^\top D^\top)^{-1} D\mu - 2h(1-\gamma)(D\sigma\sigma^\top e_d)^\top (D\sigma\sigma^\top D^\top)^{-1} D\mu\Big) \\
&= \frac{1}{2(1-\gamma)}\mu^\top D^\top (D\sigma\sigma^\top D^\top)^{-1} D\mu + h\Big(e_d^\top \mu - (D\sigma\sigma^\top e_d)^\top (D\sigma\sigma^\top D^\top)^{-1} D\mu\Big) \qquad (4.10) \\
&= \frac{1}{2(1-\gamma)}\mu^\top A\mu + hc^\top \mu,
\end{aligned}$$

where A and c were introduced in Definition 4.2. Recall that the domain on which we have to minimize (4.10) is

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, (\mu-\nu)^\top \Gamma^{-1}(\mu-\nu) \le \kappa^2\big\}.$$

To make this minimization problem easier to handle, we apply an easy transformation to the elements $\mu \in K$. For that purpose, note that since $\Gamma \in \mathbb{R}^{d\times d}$ is assumed to be symmetric and positive definite, there exists a decomposition of the form $\Gamma = \tau\tau^\top$ where $\tau \in \mathbb{R}^{d\times d}$ is nonsingular. Then we can rewrite the constraint $(\mu-\nu)^\top\Gamma^{-1}(\mu-\nu) \le \kappa^2$ as

$$\kappa^2 \ge (\mu-\nu)^\top(\tau\tau^\top)^{-1}(\mu-\nu) = (\mu-\nu)^\top(\tau^\top)^{-1}\tau^{-1}(\mu-\nu) = \big(\tau^{-1}(\mu-\nu)\big)^\top\big(\tau^{-1}(\mu-\nu)\big).$$

Hence, for an arbitrary $\mu \in K$ we define $\rho := \tau^{-1}(\mu-\nu)$ so that $\mu = \nu + \tau\rho$ and $\|\rho\| \le \kappa$. We can then rewrite (4.10) as

$$\begin{aligned}
\frac{1}{2(1-\gamma)}\mu^\top A\mu + hc^\top\mu &= \frac{1}{2(1-\gamma)}\big((\tau\rho)^\top A\tau\rho + 2\nu^\top A\tau\rho + \nu^\top A\nu\big) + hc^\top\tau\rho + hc^\top\nu \\
&= \frac{1}{2(1-\gamma)}\rho^\top\tau^\top A\tau\rho + \Big(\frac{1}{1-\gamma}A\nu + hc\Big)^\top\tau\rho + \frac{1}{2(1-\gamma)}\nu^\top A\nu + hc^\top\nu.
\end{aligned}$$

Minimizing (4.10) in $\mu \in K$ is therefore equivalent to minimizing $g\colon B_\kappa(0) \to \mathbb{R}$ with

$$g(\rho) = \frac{1}{2(1-\gamma)}\rho^\top\tau^\top A\tau\rho + \Big(hc + \frac{1}{1-\gamma}A\nu\Big)^\top\tau\rho$$

in ρ and then setting $\mu = \nu + \tau\rho$.

The behavior of the function g is, especially for large values of κ, determined to a large extent by the matrix A. We collect some useful results about D and A in the following lemmas.

Lemma 4.5. *For the matrix D from Lemma 4.1 it holds*

(i) $\ker(D) = \operatorname{span}(\{\mathbf{1}_d\})$*;*

(ii) $\ker(D^\top) = \{0\}$.

Proof. (i) For any $x \in \mathbb{R}^d$ we have that $Dx \in \mathbb{R}^{d-1}$ where the i-th component of Dx is $x_i - x_d$, $i = 1, \dots, d-1$. Hence, $Dx = 0$ if and only if $x_1 = \cdots = x_d$.

(ii) From Lemma 4.1 we know that D has rank $d-1$, hence $\ker(D^\top) = \{0\}$.

□

Lemma 4.6. *The matrix A from Definition 4.2 is symmetric and positive semidefinite with* $\ker(A) = \operatorname{span}(\{\mathbf{1}_d\})$.

Proof. Note that $D\sigma\sigma^\top D^\top$ is symmetric. Hence, the same is true for its inverse and thus for $D^\top(D\sigma\sigma^\top D^\top)^{-1}D$. Also, $D\sigma\sigma^\top D^\top = (D\sigma)(D\sigma)^\top$ is positive definite since $\sigma \in \mathbb{R}^{d\times m}$ has rank d and therefore by Lemma 4.1, $D\sigma$ has full row rank $d-1$. It follows that also the inverse $(D\sigma\sigma^\top D^\top)^{-1}$ is positive definite. So since

$$x^\top Ax = x^\top D^\top(D\sigma\sigma^\top D^\top)^{-1}Dx = (Dx)^\top(D\sigma\sigma^\top D^\top)^{-1}(Dx) \geq 0$$

for any $x \in \mathbb{R}^d$, the matrix A is positive semidefinite. From Lemma 4.5 it follows that

$$Ax = D^\top(D\sigma\sigma^\top D^\top)^{-1}Dx = 0$$

if and only if $(D\sigma\sigma^\top D^\top)^{-1}Dx = 0$, which is equivalent to $Dx = 0$. Hence we can deduce $\ker(A) = \ker(D) = \operatorname{span}(\{\mathbf{1}_d\})$. □

From Lemma 4.6 we can immediately deduce that also the matrix $\tau^\top A\tau \in \mathbb{R}^{d\times d}$ is symmetric and positive semidefinite with

$$\ker(\tau^\top A\tau) = \operatorname{span}(\{\tau^{-1}\mathbf{1}_d\}).$$

Having collected these properties of the matrix A and of $\tau^\top A\tau$ enables us to find the parameter ρ that minimizes

$$g(\rho) = \frac{1}{2(1-\gamma)}\rho^\top\tau^\top A\tau\rho + \Big(hc + \frac{1}{1-\gamma}A\nu\Big)^\top\tau\rho$$

on the set $B_\kappa(0) = \{\rho \in \mathbb{R}^d \mid \|\rho\| \leq \kappa\}$. The following lemma identifies the minimizer.

Lemma 4.7. *Let $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_d$ denote the eigenvalues of $\tau^\top A\tau$, and let*

$$v_1 = \frac{1}{\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}\mathbf{1}_d, v_2, \ldots, v_d \in \mathbb{R}^d$$

denote the respective orthogonal eigenvectors with $\|v_i\| = 1$ for all $i = 1, \ldots, d$. Then the minimum of the function $g\colon B_\kappa(0) \to \mathbb{R}$ with

$$g(\rho) = \frac{1}{2(1-\gamma)}\rho^\top\tau^\top A\tau\rho + \Big(hc + \frac{1}{1-\gamma}A\nu\Big)^\top\tau\rho$$

on the domain $B_\kappa(0) = \{\rho \in \mathbb{R}^d \mid \|\rho\| \leq \kappa\}$ is attained by the vector

$$\rho^* = -\sum_{i=1}^d\bigg(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\bigg)^{-1}\Big\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\Big\rangle v_i,$$

where $\psi(\kappa) \in (0, \kappa]$ is uniquely determined by $\|\rho^*\| = \kappa$.

Proof. Recall that $\tau^\top A\tau$ has eigenvalue $\lambda_1 = 0$ with a corresponding normed eigenvector of the form $v_1 = \frac{1}{\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}\mathbf{1}_d$. Also, the other eigenvalues of $\tau^\top A\tau$ are positive, and due to the symmetry of the matrix we can assume that $v_1, \dots, v_d$ are orthogonal and form a basis of $\mathbb{R}^d$.

Firstly, we show that the minimum of g is attained on the boundary of $B_\kappa(0)$. For that purpose, we observe that the gradient of g is

$$\begin{aligned}
\nabla g(\rho) &= \frac{1}{2(1-\gamma)}2\tau^\top A\tau\rho + \tau^\top\Big(hc + \frac{1}{1-\gamma}A\nu\Big) \\
&= \frac{1}{1-\gamma}\tau^\top A\tau\rho + h\tau^\top(I_d - A\sigma\sigma^\top)e_d + \frac{1}{1-\gamma}\tau^\top A\nu \\
&= \tau^\top\bigg(A\Big(\frac{1}{1-\gamma}(\tau\rho+\nu) - h\sigma\sigma^\top e_d\Big) + he_d\bigg) \\
&= \tau^\top\bigg(D^\top(D\sigma\sigma^\top D^\top)^{-1}D\Big(\frac{1}{1-\gamma}(\tau\rho+\nu) - h\sigma\sigma^\top e_d\Big) + he_d\bigg).
\end{aligned}$$

From the last representation of the gradient it becomes apparent that there is no $\rho \in B_\kappa(0)$ with $\nabla g(\rho) = 0$, since $\tau^\top$ is nonsingular and the vector he_d is not in the range of $D^\top$. Hence, there is no critical point of the function g. The minimum of the function on the bounded, closed set $B_\kappa(0)$ is therefore attained on the boundary. In the following, we compute that minimizer.

Let $\rho \in B_\kappa(0)$ be arbitrary. Since $v_1, \dots, v_d$ form a basis of $\mathbb{R}^d$, we can write

$$\rho = \sum_{i=1}^d a_i v_i,$$

where $a_1, \dots, a_d \in \mathbb{R}$ are uniquely determined. The minimization of g in ρ is then equivalent to a minimization in the coefficients $a_1, \dots, a_d$. Since we know that a minimizer of the function g must lie on the boundary of $B_\kappa(0)$ and therefore have norm $\|\rho\| = \kappa$, we obtain the constraint

$$\kappa^2 = \|\rho\|^2 = \sum_{i=1}^d a_i^2 \tag{4.11}$$

on the coefficients. Before doing the minimization, we first notice that for our minimizer, the coefficient a_1 will be less or equal than zero. This is because

$$\begin{aligned}
g\Big(\sum_{i=1}^d a_i v_i\Big) &= \frac{1}{2(1-\gamma)}\Big(\sum_{i=1}^d a_i v_i\Big)^\top \tau^\top A\tau\Big(\sum_{i=1}^d a_i v_i\Big) + \Big(hc + \frac{1}{1-\gamma}A\nu\Big)^\top \tau\Big(\sum_{i=1}^d a_i v_i\Big) \\
&= \frac{1}{2(1-\gamma)}\sum_{i=1}^d\sum_{j=1}^d a_i a_j v_i^\top \tau^\top A\tau v_j + \sum_{i=1}^d a_i hc^\top \tau v_i + \frac{1}{1-\gamma}\sum_{i=1}^d a_i (A\nu)^\top \tau v_i \\
&= \frac{1}{2(1-\gamma)}\sum_{i=1}^d a_i^2\lambda_i + \sum_{i=1}^d a_i hc^\top\tau v_i + \frac{1}{1-\gamma}\sum_{i=1}^d a_i \nu^\top \lambda_i (\tau^\top)^{-1} v_i \\
&= \frac{1}{2(1-\gamma)}\sum_{i=2}^d a_i^2\lambda_i + \sum_{i=2}^d a_i\Big(hc + \frac{\lambda_i}{1-\gamma}\Gamma^{-1}\nu\Big)^\top \tau v_i + a_1 hc^\top \tau v_1.
\end{aligned}$$

For the third equality we have used that v_i is an eigenvector of $\tau^\top A\tau$ to eigenvalue λ_i and that $v_1, \dots, v_d$ are orthogonal. In the last step we have used $\lambda_1 = 0$. Next, one easily sees that

$$c^\top \tau v_1 = e_d^\top (I_d - A\sigma\sigma^\top)^\top \tau \frac{1}{\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}\mathbf{1}_d = \frac{1}{\|\tau^{-1}\mathbf{1}_d\|} e_d^\top (\mathbf{1}_d - \sigma\sigma^\top A\mathbf{1}_d) = \frac{1}{\|\tau^{-1}\mathbf{1}_d\|}, \tag{4.12}$$

since $A\mathbf{1}_d = 0$ by Lemma 4.6. By plugging in this representation we deduce that, when looking for the minimizer of g, we can restrict to the parameters ρ with coefficient $a_1 \le 0$. Hence, we can rewrite the constraint (4.11) as

$$a_1 = -\sqrt{\kappa^2 - \sum_{i=2}^d a_i^2}.$$

We plug this representation of a_1, as well as (4.12), back in to obtain

$$\begin{aligned}\widetilde{g}(a_2, \dots, a_d) &:= g\Big(\sum_{i=1}^d a_i v_i\Big)\\ &= \frac{1}{2(1-\gamma)}\sum_{i=2}^d a_i^2\lambda_i + \sum_{i=2}^d a_i\Big(hc + \frac{\lambda_i}{1-\gamma}\Gamma^{-1}\nu\Big)^\top \tau v_i - \frac{h}{\|\tau^{-1}\mathbf{1}_d\|}\sqrt{\kappa^2 - \sum_{i=2}^d a_i^2},\end{aligned}$$

and minimize this expression in $a_2, \dots, a_d$. Note that the domain of $\widetilde{g}$ is $\{x \in \mathbb{R}^{d-1} \mid \|x\| \le \kappa\}$. In the interior of this domain, the partial derivative of $\widetilde{g}$ with respect to a_k, $k = 2, \dots, d$, is given by

$$\begin{aligned}\frac{\partial \widetilde{g}}{\partial a_k}(a_2, \dots, a_d) &= \frac{2a_k\lambda_k}{2(1-\gamma)} + \Big(hc + \frac{\lambda_k}{1-\gamma}\Gamma^{-1}\nu\Big)^\top \tau v_k - \frac{h}{2\|\tau^{-1}\mathbf{1}_d\|\sqrt{\kappa^2 - \sum_{i=2}^d a_i^2}}(-2a_k)\\ &= \left(\frac{\lambda_k}{1-\gamma} + \frac{h}{\|\tau^{-1}\mathbf{1}_d\|\sqrt{\kappa^2 - \sum_{i=2}^d a_i^2}}\right)a_k + \Big(hc + \frac{\lambda_k}{1-\gamma}\Gamma^{-1}\nu\Big)^\top \tau v_k.\end{aligned}$$

When setting this expression equal to zero, we obtain

$$\begin{aligned}a_k &= -\left(\frac{\lambda_k}{1-\gamma} + \frac{h}{\|\tau^{-1}\mathbf{1}_d\|\sqrt{\kappa^2 - \sum_{i=2}^d a_i^2}}\right)^{-1}\Big(hc + \frac{\lambda_k}{1-\gamma}\Gamma^{-1}\nu\Big)^\top \tau v_k\\ &= -\left(\frac{\lambda_k}{1-\gamma} - \frac{h}{\|\tau^{-1}\mathbf{1}_d\| a_1}\right)^{-1}\Big\langle h\tau^\top c + \frac{\lambda_k}{1-\gamma}\tau^{-1}\nu, v_k\Big\rangle.\end{aligned} \tag{4.13}$$

Note that this representation does not provide the coefficients a_k explicitly since a_1 here is a function of $(a_2, \dots, a_d)$. However, the function

$$[-\kappa, 0) \ni a_1 \mapsto a_1^2 + \sum_{i=2}^d \left(\frac{\lambda_i}{1-\gamma} - \frac{h}{\|\tau^{-1}\mathbf{1}_d\| a_1}\right)^{-2}\Big\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\Big\rangle^2$$

has the derivative

$$
\begin{aligned}
&2a_1 + \sum_{i=2}^{d}(-2)\Big(\frac{\lambda_i}{1-\gamma} - \frac{h}{\|\tau^{-1}\mathbf{1}_d\|a_1}\Big)^{-3}\Big(-\frac{h}{\|\tau^{-1}\mathbf{1}_d\|}\Big)\Big(-\frac{1}{a_1^2}\Big)\Big\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\Big\rangle^2 \\
&= 2a_1 - \frac{2h}{\|\tau^{-1}\mathbf{1}_d\|a_1^2}\sum_{i=2}^{d}\Big(\frac{\lambda_i}{1-\gamma} - \frac{h}{\|\tau^{-1}\mathbf{1}_d\|a_1}\Big)^{-3}\Big\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\Big\rangle^2,
\end{aligned}
$$

which is strictly negative on $[-\kappa, 0)$. For $a_1 = -\kappa$, the value of the function is greater or equal κ^2, for a_1 tending to zero from below it converges to zero, hence there is a unique value of $a_1 \in [-\kappa, 0)$ where the function has value κ^2. So (4.13) together with (4.11) uniquely determines $a_1, \ldots, a_d$.

Moreover, the second partial derivatives of $\widetilde{g}$ have the form

$$
\begin{aligned}
\frac{\partial^2 \widetilde{g}}{\partial a_k^2}(a_2,\ldots,a_d) &= \frac{\lambda_k}{1-\gamma} + \frac{h}{\|\tau^{-1}\mathbf{1}_d\|\sqrt{\kappa^2 - \sum_{i=2}^d a_i^2}} - a_k \frac{h}{2\|\tau^{-1}\mathbf{1}_d\|\big(\kappa^2 - \sum_{i=2}^d a_i^2\big)^{3/2}}(-2a_k) \\
&= \frac{\lambda_k}{1-\gamma} + \frac{h}{\|\tau^{-1}\mathbf{1}_d\|\sqrt{\kappa^2 - \sum_{i=2}^d a_i^2}} + \frac{h a_k^2}{\|\tau^{-1}\mathbf{1}_d\|\big(\kappa^2 - \sum_{i=2}^d a_i^2\big)^{3/2}}
\end{aligned}
$$

for $k = 2, \ldots, d$, and for $k, l = 1, \ldots, d$ with $k \neq l$ we obtain

$$
\frac{\partial^2 \widetilde{g}}{\partial a_l \partial a_k}(a_2,\ldots,a_d) = -\frac{h a_k}{2\|\tau^{-1}\mathbf{1}_d\|\big(\kappa^2 - \sum_{i=2}^d a_i^2\big)^{3/2}}(-2a_l) = \frac{h a_k a_l}{\|\tau^{-1}\mathbf{1}_d\|\big(\kappa^2 - \sum_{i=2}^d a_i^2\big)^{3/2}}.
$$

Hence, the Hessian of $\widetilde{g}$ is of the form

$$
\begin{aligned}
H\widetilde{g}(a_2,\ldots,a_d) = \frac{1}{1-\gamma}\widetilde{\Lambda} &+ \frac{h}{\|\tau^{-1}\mathbf{1}_d\|\sqrt{\kappa^2 - \sum_{i=2}^d a_i^2}} I_{d-1} \\
&+ \frac{h}{\|\tau^{-1}\mathbf{1}_d\|\big(\kappa^2 - \sum_{i=2}^d a_i^2\big)^{3/2}}(a_2,\ldots,a_d)^\top (a_2,\ldots,a_d),
\end{aligned}
$$

where $\widetilde{\Lambda} \in \mathbb{R}^{(d-1)\times(d-1)}$ is a diagonal matrix with diagonal entries $\lambda_2, \ldots, \lambda_d > 0$. Obviously, the first two summands on the right-hand side are positive-definite matrices. The last summand is positive semidefinite. So we conclude that $H\widetilde{g}$ is positive definite on the whole interior of the domain of $\widetilde{g}$. In particular, in the point $(a_2, \ldots, a_d)$ defined via (4.13) together with (4.11), there is a global minimum of the function $\widetilde{g}$.

To conclude with, the minimum of the function g on $B_\kappa(0)$ is attained by the vector

$$
\rho^* = \sum_{i=1}^{d} a_i v_i,
$$

where

$$
a_i = -\Big(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\Big)^{-1}\Big\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\Big\rangle \tag{4.14}
$$

for $i = 1, \ldots, d$, and where $\psi(\kappa) = -a_1 \in (0, \kappa]$ is uniquely determined by $\|\rho^*\| = \kappa$. Note that (4.14) also holds for $i = 1$ since $\lambda_1 = 0$ and $c^\top \tau v_1 = \frac{1}{\|\tau^{-1}\mathbf{1}_d\|}$ by (4.12). $\square$

The previous lemma now yields the solution of the dual problem

$$\inf_{\mu\in K}\sup_{\pi\in\mathcal{A}_h(x_0)}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

to our original optimization problem.

Theorem 4.8. *Let $0=\lambda_1<\lambda_2\leq\cdots\leq\lambda_d$ denote the eigenvalues of $\tau^\top A\tau$, and let*

$$v_1=\frac{1}{\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}\mathbf{1}_d, v_2,\ldots,v_d\in\mathbb{R}^d$$

denote the respective orthogonal eigenvectors with $\|v_i\|=1$ for all $i=1,\ldots,d$. Then

$$\inf_{\mu\in K}\sup_{\pi\in\mathcal{A}_h(x_0)}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]=\mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big],$$

where

$$\mu^*=\nu-\tau\sum_{i=1}^d\bigg(\frac{\lambda_i}{1-\gamma}+\frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\bigg)^{-1}\Big\langle h\tau^\top c+\frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\Big\rangle v_i$$

for $\psi(\kappa)\in(0,\kappa]$ that is uniquely determined by $\|\tau^{-1}(\mu^*-\nu)\|=\kappa$, and where $(\pi_t^*)_{t\in[0,T]}$ is defined by

$$\pi_t^*=\frac{1}{1-\gamma}A\mu^*+hc$$

for all $t\in[0,T]$.

Proof. For any fixed parameter $\mu\in\mathbb{R}^d$, Proposition 4.3 gives the optimal strategy for the optimization problem

$$\sup_{\pi\in\mathcal{A}_h(x_0)}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big].$$

With the help of Corollary 4.4 we have seen that minimizing the above expression in μ on the set $K=\big\{\mu\in\mathbb{R}^d\,\big|\,(\mu-\nu)^\top\Gamma^{-1}(\mu-\nu)\leq\kappa^2\big\}$ is equivalent to minimizing the function $g\colon B_\kappa(0)\to\mathbb{R}$ with

$$g(\rho)=\frac{1}{2(1-\gamma)}\rho^\top\tau^\top A\tau\rho+\Big(hc+\frac{1}{1-\gamma}A\nu\Big)^\top\tau\rho$$

in ρ and then setting $\mu=\nu+\tau\rho$. The claim now follows from Lemma 4.7 together with the representation in Proposition 4.3. □

The preceding theorem solves the problem

$$\inf_{\mu\in K}\sup_{\pi\in\mathcal{A}_h(x_0)}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big], \tag{4.15}$$

where $K=\big\{\mu\in\mathbb{R}^d\,\big|\,(\mu-\nu)^\top\Gamma^{-1}(\mu-\nu)\leq\kappa^2\big\}$. This is the corresponding dual problem to our original optimization problem

$$\sup_{\pi\in\mathcal{A}_h(x_0)}\inf_{\mu\in K}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big], \tag{4.16}$$

but in general the values of these two problems do not coincide since the supremum and the infimum do not interchange. There are, of course, special cases in which the supremum and the infimum do interchange. Those results are called *minimax theorems* in the literature. In the context of our portfolio optimization problem, a minimax theorem has been shown in Quenez [50], building up on the theory by Kramkov and Schachermayer [37]. However, the setting in Quenez [50] does not include any additional constraints on the trading strategies. Due to our constraint $\langle \pi_t, \mathbf{1}_d \rangle = h$ for all $t \in [0,T]$ we cannot carry over these results directly. In the following, we will however use our knowledge about the optimal strategy for (4.15) to show that it indeed also solves (4.16) and that in this case, the supremum and the infimum can be interchanged.

4.1.3. A minimax theorem

The following representation of π^* is useful for establishing duality and proving our minimax theorem.

Lemma 4.9. *The strategy π^* from Theorem 4.8 satisfies*

$$\pi_t^* = -\frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\Gamma^{-1}(\mu^* - \nu)$$

for all $t \in [0,T]$.

Proof. Throughout the proof, let

$$a_i = -\Big(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\Big)^{-1}\Big\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\Big\rangle$$

for $i = 1, \dots, d$, so that

$$\tau^{-1}(\mu^* - \nu) = \sum_{i=1}^d a_i v_i.$$

Due to the form of the a_i we can write

$$\sum_{i=1}^d \Big(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\Big) a_i v_i = -\sum_{i=1}^d \Big\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\Big\rangle v_i.$$

Since the vectors $v_1, \dots, v_d$ form an orthonormal basis of $\mathbb{R}^d$ and are eigenvectors to the eigenvalues $\lambda_1, \dots, \lambda_d$ of the symmetric matrix $\tau^\top A\tau$, the right-hand side equals

$$\begin{aligned}
-h\tau^\top c - \frac{1}{1-\gamma}\sum_{i=1}^d \langle \tau^{-1}\nu, \lambda_i v_i\rangle v_i &= -h\tau^\top c - \frac{1}{1-\gamma}\sum_{i=1}^d \langle \tau^{-1}\nu, \tau^\top A\tau v_i\rangle v_i \\
&= -h\tau^\top c - \frac{1}{1-\gamma}\sum_{i=1}^d \langle \tau^\top A\nu, v_i\rangle v_i \\
&= -h\tau^\top c - \frac{1}{1-\gamma}\tau^\top A\nu.
\end{aligned}$$

On the other hand, we get

$$\begin{aligned}\sum_{i=1}^{d}\Big(\frac{\lambda_i}{1-\gamma}+\frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\Big)a_i v_i &= \frac{1}{1-\gamma}\sum_{i=1}^{d} a_i\lambda_i v_i + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\sum_{i=1}^{d} a_i v_i\\ &= \frac{1}{1-\gamma}\sum_{i=1}^{d} a_i\tau^\top A\tau v_i + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}(\mu^*-\nu)\\ &= \frac{1}{1-\gamma}\tau^\top A(\mu^*-\nu) + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}(\mu^*-\nu).\end{aligned}$$

We have used here that v_i is an eigenvector of $\tau^\top A\tau$ to the eigenvalue λ_i for each $i = 1,\dots,d$. In conclusion,

$$\frac{1}{1-\gamma}\tau^\top A\mu^* = -\frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}(\mu^*-\nu) - h\tau^\top c.$$

Hence, by using the representation of π^* from Theorem 4.8 we obtain

$$\begin{aligned}\pi_t^* = \frac{1}{1-\gamma}A\mu^* + hc &= (\tau^\top)^{-1}\Big(\frac{1}{1-\gamma}\tau^\top A\mu^* + h\tau^\top c\Big)\\ &= -\frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}(\tau\tau^\top)^{-1}(\mu^*-\nu)\\ &= -\frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\Gamma^{-1}(\mu^*-\nu)\end{aligned}$$

for all $t\in[0,T]$. □

The preceding lemma characterizes the strategy π^* that is optimal for the parameter μ^*. In the following we show that, vice versa, μ^* is also the worst possible drift parameter, given that an investor applies strategy π^*. It then follows that the point (π^*,μ^*) is a *saddle point* of our problem, i.e. it holds

$$\mathbb{E}_{\mu^*}\big[U_\gamma(X_T^\pi)\big] \le \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] \le \mathbb{E}_{\mu}\big[U_\gamma(X_T^{\pi^*})\big]$$

for all $\mu\in K$ and $\pi\in\mathcal{A}_h(x_0)$. This property of (π^*,μ^*) is essential for proving that the value of our original optimization problem

$$\sup_{\pi\in\mathcal{A}_h(x_0)}\inf_{\mu\in K}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \tag{4.17}$$

equals the value of the corresponding dual problem

$$\inf_{\mu\in K}\sup_{\pi\in\mathcal{A}_h(x_0)}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big], \tag{4.18}$$

i.e. that the supremum and the infimum interchange. Note that the inequality

$$\sup_{\pi\in\mathcal{A}_h(x_0)}\inf_{\mu\in K}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \le \inf_{\mu\in K}\sup_{\pi\in\mathcal{A}_h(x_0)}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

always holds when interchanging supremum and infimum, see for example Ekeland and Temam [18, Ch. VI, Prop. 1.1]. For the reverse inequality the saddle point property is needed.

Proposition 4.10. *The parameter μ that attains the minimum in*

$$\inf_{\mu \in K} \mathbb{E}_\mu\big[U_\gamma(X_T^{\pi^*})\big]$$

is μ^*, i.e. μ^* is the worst possible parameter, given that an investor chooses strategy π^*.

Proof. Since π^* is a strategy that is constant in time and deterministic, we can rewrite the expected utility of terminal wealth in the case $\gamma \neq 0$ as

$$\begin{aligned}\mathbb{E}_\mu\big[U_\gamma(X_T^{\pi^*})\big] &= \frac{x_0^\gamma}{\gamma}\,\mathbb{E}_\mu\Big[\exp\Big(\gamma r T + \gamma T\Big((\pi_0^*)^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \pi_0^*\|^2\Big) + \gamma(\pi_0^*)^\top \sigma W_T\Big)\Big]\\ &= \frac{x_0^\gamma}{\gamma}\exp\Big(\gamma r T + \gamma T\Big((\pi_0^*)^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \pi_0^*\|^2\Big) + \frac{1}{2}\gamma^2 T\|\sigma^\top \pi_0^*\|^2\Big).\end{aligned}$$

In the case $\gamma = 0$ we have

$$\mathbb{E}_\mu\big[\log(X_T^{\pi^*})\big] = \log(x_0) + rT + T\Big((\pi_0^*)^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \pi_0^*\|^2\Big).$$

Obviously, for any $\gamma \in (-\infty, 1)$ the parameter $\mu \in K$ that minimizes the expressions above is the parameter that minimizes the term $(\pi_0^*)^\top \mu$.

For an arbitrary $\theta \in \mathbb{R}^d$, $\theta \neq 0$, it holds that the parameter $\mu \in \mathbb{R}^d$ that minimizes $\theta^\top \mu$ such that $(\mu - \nu)^\top \Gamma^{-1}(\mu - \nu) \le \kappa^2$ has the form

$$\widetilde{\mu} = \nu - \frac{\kappa}{\sqrt{\theta^\top \Gamma \theta}}\Gamma\theta. \tag{4.19}$$

To verify this, note that for any μ with $\|\tau^{-1}(\mu - \nu)\| \le \kappa$ we can rewrite

$$\theta^\top \mu = \theta^\top \nu + \theta^\top \tau\tau^{-1}(\mu - \nu),$$

so that Hölder's inequality implies

$$\theta^\top \mu \ge \theta^\top \nu - \|\tau^\top \theta\|\,\|\tau^{-1}(\mu - \nu)\| \ge \theta^\top \nu - \kappa\|\tau^\top \theta\| = \theta^\top \nu - \kappa\sqrt{\theta^\top \Gamma \theta}. \tag{4.20}$$

For the parameter $\widetilde{\mu}$ from (4.19) we obtain

$$\theta^\top \widetilde{\mu} = \theta^\top \nu - \frac{\kappa}{\sqrt{\theta^\top \Gamma \theta}}\theta^\top \Gamma \theta = \theta^\top \nu - \kappa\sqrt{\theta^\top \Gamma \theta},$$

so that the lower bound in (4.20) is attained. Now it is sufficient to show that the parameter μ^* is equal to $\widetilde{\mu}$ from (4.19) for $\theta = \pi_0^*$.

From Lemma 4.9 we recall

$$\pi_t^* = -\frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\Gamma^{-1}(\mu^* - \nu). \tag{4.21}$$

Hence,

$$(\pi_0^*)^\top \Gamma \pi_0^* = \frac{h^2}{\psi(\kappa)^2\|\tau^{-1}\mathbf{1}_d\|^2}(\mu^* - \nu)^\top \Gamma^{-1}(\mu^* - \nu) = \frac{h^2\kappa^2}{\psi(\kappa)^2\|\tau^{-1}\mathbf{1}_d\|^2}$$

and

$$\sqrt{(\pi_0^*)^\top \Gamma \pi_0^*} = \frac{h\kappa}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}. \tag{4.22}$$

When rearranging (4.21) for μ^* and plugging in (4.22) we obtain

$$\mu^* = \nu - \frac{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}{h}\Gamma\pi_0^* = \nu - \frac{\kappa}{\sqrt{(\pi_0^*)^\top \Gamma \pi_0^*}}\Gamma\pi_0^*.$$

Comparing with $\widetilde{\mu}$ in (4.19) we conclude that μ^* is the parameter that minimizes $(\pi_0^*)^\top\mu$ over all $\mu \in K$, and is therefore also the parameter that attains the minimum in

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^{\pi^*})\big].$$

We conclude that μ^* is the worst possible parameter for the strategy π^*. $\square$

The above proposition establishes an equilibrium result and a direct connection between the optimization problems (4.17) and (4.18). The strategy π^* is the best strategy that an investor can choose when the drift of stocks is known to be μ^*. On the other hand, μ^* is also the parameter the market has to choose to minimize the investor's expected utility of terminal wealth, given that the investor applies strategy π^*. This can also be seen in the context of a two-player game where player one (the investor) tries to maximize $\mathbb{E}_\mu[U_\gamma(X_T^\pi)]$ over strategies π and player two (the market) tries to minimize the same expression over parameters μ. In general, it makes a difference which player is the first one to make a choice in such a two-player game. The point (π^*, μ^*) however constitutes a saddle point, which enables us to show that in our setting the solution to both optimization problems (4.17) and (4.18) is the same. Put differently, if both players behave optimally, it does not make any difference whether player one or player two is the first one to make a choice.

Theorem 4.11. *Let $K = \{\mu \in \mathbb{R}^d \mid (\mu-\nu)^\top\Gamma^{-1}(\mu-\nu) \le \kappa^2\}$. Then*

$$\sup_{\pi\in\mathcal{A}_h(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] = \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] = \inf_{\mu\in K} \sup_{\pi\in\mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big],$$

where μ^* and π^* are defined as in Theorem 4.8.

Proof. For an arbitrary parameter $\mu \in K$, let $\pi(\mu) = (\pi_t(\mu))_{t\in[0,T]}$ denote the strategy from $\mathcal{A}_h(x_0)$ that is optimal, given that the drift parameter is μ. Recall that we have found a representation for $\pi(\mu)$ in Proposition 4.3. Then we know from Theorem 4.8 that

$$\inf_{\mu\in K} \sup_{\pi\in\mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] = \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^{\pi(\mu)})\big] = \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big]. \tag{4.23}$$

On the other hand, Proposition 4.10 yields

$$\mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] = \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^{\pi^*})\big] \le \sup_{\pi\in\mathcal{A}_h(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]. \tag{4.24}$$

Furthermore, we also have

$$\sup_{\pi\in\mathcal{A}_h(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \le \inf_{\mu\in K} \sup_{\pi\in\mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

since the inequality always holds when interchanging supremum and infimum, see for example Ekeland and Temam [18, Ch. VI, Prop. 1.1]. Hence, combining (4.23) and (4.24) yields

$$\begin{aligned}\inf_{\mu\in K}\sup_{\pi\in\mathcal{A}_h(x_0)}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] &= \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big]\\ &\le \sup_{\pi\in\mathcal{A}_h(x_0)}\inf_{\mu\in K}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \le \inf_{\mu\in K}\sup_{\pi\in\mathcal{A}_h(x_0)}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big].\end{aligned} \tag{4.25}$$

Consequently, all inequalities in (4.25) are equalities and the claim follows. □

The previous theorem establishes duality between our original robust utility maximization problem (4.17) and the dual problem (4.18) where supremum and infimum are interchanged. Additionally, we now also know the solution to our original problem since the theorem states

$$\sup_{\pi\in\mathcal{A}_h(x_0)}\inf_{\mu\in K}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] = \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big],$$

where both μ^* and π^* are given in Theorem 4.8. This gives the optimal solution to our constrained robust utility maximization problem in a nearly explicit way. Note that the parameter μ^* in Theorem 4.8 is not given explicitly since the parameter $\psi(\kappa)$ is defined in an implicit way. However, finding $\psi(\kappa)$ numerically can be done in a straightforward way by a numerical root search of a monotone function. For this reason, determining μ^* and π^* numerically does not pose any problems.

From the representation of π^* in Theorem 4.8 we immediately see that the optimal strategy is deterministic and constant in time. It is therefore optimal for investors to hold a certain constant fraction of wealth in every asset. The degree of uncertainty κ only appears in the parameter $\psi(\kappa)$. In Chapter 5 we study the asymptotic behavior of the strategy as uncertainty becomes large.

4.2. Alternative uncertainty sets

In the preceding section we have modelled K as an ellipsoid with some arbitrary radius κ. A special case of such an ellipsoid is an uncertainty ball. Of course, one can think of other reasonable sets for modelling uncertainty about the drift parameter μ. In this section we want to apply our duality approach to the optimization problem

$$\sup_{\pi\in\mathcal{A}_h(x_0)}\inf_{\mu\in K}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big], \tag{4.26}$$

where again

$$\mathcal{A}_h(x_0) = \big\{\pi\in\mathcal{A}(x_0)\,\big|\,\langle\pi_t,\mathbf{1}_d\rangle = h \text{ for all } t\in[0,T]\big\}$$

for some $h>0$ but with an alternative form of K. We now instead define

$$K = \big\{\mu\in\mathbb{R}^d\,\big|\,\mathbf{1}_d^\top\mu = b\big\}$$

for some $b\in\mathbb{R}$. The motivation for this uncertainty set is that one has an estimate for the performance of a stock index, and therefore for the overall average performance of the stocks, but not for the single stocks themselves. Note that, in contrast to the ellipsoidal uncertainty set, this set K is unbounded and we do not have any radius or level of uncertainty.

Recall that Proposition 4.3 gives the optimal strategy for an investor who knows the true drift parameter μ. Further, we have seen that minimizing

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

in μ is equivalent to minimizing $g\colon K \to \mathbb{R}$ with

$$g(\mu) = \frac{1}{2(1-\gamma)}\mu^\top A\mu + hc^\top \mu.$$

The following theorem gives an explicit representation of the minimizer μ^* on the set K defined above.

Theorem 4.12. *Let $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_d$ denote the eigenvalues of the matrix A, and let $v_1, \dots, v_d \in \mathbb{R}^d$ denote the respective orthogonal eigenvectors with $\|v_i\| = 1$ for all $i = 1, \dots, d$. Then*

$$\inf_{\mu \in K} \sup_{\pi \in \mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] = \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big],$$

where

$$\mu^* = \frac{b}{d}\mathbf{1}_d - h(1-\gamma)\sum_{i=2}^d \lambda_i^{-1}\langle c, v_i\rangle\, v_i,$$

and where $(\pi_t^*)_{t\in[0,T]}$ is defined by $\pi_t^* = \frac{h}{d}\mathbf{1}_d$ for all $t \in [0,T]$.

Proof. Let $\mu \in K$. We know that we can take $v_1 = \frac{1}{\sqrt{d}}\mathbf{1}_d$. Then we can write $\mu = \sum_{i=1}^d a_i v_i$ for some $a_1, \dots, a_d \in \mathbb{R}$. Due to our assumption $\mu \in K$ we have

$$b = \mathbf{1}_d^\top \mu = \mathbf{1}_d^\top \sum_{i=1}^d a_i v_i = \sum_{i=1}^d a_i \mathbf{1}_d^\top v_i = a_1 \mathbf{1}_d^\top v_1 = a_1\sqrt{d}.$$

The second but last equality follows from orthogonality of the v_i. For any $\mu \in K$ we must therefore have $a_1 = \frac{b}{\sqrt{d}}$. Now we plug in the representation of μ into g and obtain

$$\begin{aligned}
g(\mu) &= \frac{1}{2(1-\gamma)}\left(\sum_{i=1}^d a_i v_i\right)^\top A\left(\sum_{i=1}^d a_i v_i\right) + hc^\top \sum_{i=1}^d a_i v_i \\
&= \frac{1}{2(1-\gamma)}\sum_{i=1}^d a_i^2\lambda_i + h\sum_{i=1}^d a_i c^\top v_i \\
&= \frac{1}{2(1-\gamma)}\sum_{i=2}^d a_i^2\lambda_i + h\sum_{i=2}^d a_i c^\top v_i + ha_1 c^\top v_1 \\
&= \frac{1}{2(1-\gamma)}\sum_{i=2}^d a_i^2\lambda_i + h\sum_{i=2}^d a_i c^\top v_i + h\frac{b}{d}.
\end{aligned} \tag{4.27}$$

In the last step we have used that $a_1 = \frac{b}{\sqrt{d}}$ and

$$c^\top v_1 = \frac{1}{\sqrt{d}}e_d^\top (I_d - \sigma\sigma^\top A)\mathbf{1}_d = \frac{1}{\sqrt{d}}e_d^\top \mathbf{1}_d = \frac{1}{\sqrt{d}}$$

due to $A\mathbf{1}_d = 0$. The expression in (4.27) is now just a function of the parameters $a_2, \dots, a_d$. So we define $\widetilde{g}\colon \mathbb{R}^{d-1} \to \mathbb{R}$ by

$$\widetilde{g}(a_2, \dots, a_d) = \frac{1}{2(1-\gamma)} \sum_{i=2}^{d} a_i^2 \lambda_i + h \sum_{i=2}^{d} a_i c^\top v_i + h\frac{b}{d}$$

and perform a minimization of $\widetilde{g}$ in $a_2, \dots, a_d$. Note that

$$\frac{\partial \widetilde{g}}{\partial a_k}(a_2, \dots, a_d) = \frac{\lambda_k}{1-\gamma} a_k + h c^\top v_k$$

for any $k = 2, \dots, d$. The first-order condition thus yields $a_k = -\frac{1-\gamma}{\lambda_k} h c^\top v_k$, $k = 2, \dots, d$, as a candidate. The second derivatives are

$$\frac{\partial^2 \widetilde{g}}{\partial a_k^2} = \frac{\lambda_k}{1-\gamma} > 0 \quad \text{and} \quad \frac{\partial^2 \widetilde{g}}{\partial a_l \partial a_k} = 0$$

for all $k, l = 2, \dots, d$ with $k \neq l$. Hence, our candidate is really a minimizer of the function $\widetilde{g}$. It follows that the minimum of

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

over $\mu \in K$ is attained by

$$\mu^* = \sum_{i=1}^{d} a_i v_i = \frac{b}{\sqrt{d}} v_1 - h(1-\gamma) \sum_{i=2}^{d} \frac{1}{\lambda_i} c^\top v_i v_i = \frac{b}{d}\mathbf{1}_d - h(1-\gamma) \sum_{i=2}^{d} \lambda_i^{-1} \langle c, v_i \rangle \, v_i.$$

From Proposition 4.3 we know that the optimal strategy given parameter μ^* is $(\pi_t^*)_{t\in[0,T]}$ with

$$\pi_t^* = \frac{1}{1-\gamma} A\mu^* + hc$$

for all $t \in [0, T]$. Now note that

$$\sum_{i=1}^{d} \lambda_i a_i v_i = -h(1-\gamma) \sum_{i=2}^{d} \langle c, v_i \rangle v_i = -h(1-\gamma)(c - \langle c, v_1 \rangle v_1) = -h(1-\gamma)\Big(c - \frac{1}{d}\mathbf{1}_d\Big). \tag{4.28}$$

On the other hand

$$\sum_{i=1}^{d} \lambda_i a_i v_i = \sum_{i=1}^{d} a_i A v_i = A \sum_{i=1}^{d} a_i v_i = A\mu^*. \tag{4.29}$$

Combining (4.28) and (4.29) yields

$$A\mu^* = -h(1-\gamma)\Big(c - \frac{1}{d}\mathbf{1}_d\Big).$$

This implies for the optimal strategy that

$$\pi_t^* = -h\Big(c - \frac{1}{d}\mathbf{1}_d\Big) + hc = \frac{h}{d}\mathbf{1}_d$$

for any $t \in [0, T]$. □

The above theorem gives the solution to the problem

$$\inf_{\mu \in K} \sup_{\pi \in \mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \tag{4.30}$$

for our index-based uncertainty set K. Note that the representation of μ^* is explicit given the eigenvalues and eigenvectors of the matrix A. The strategy π^* is constant in time and equal to

$$\pi_t^* = \frac{h}{d}\mathbf{1}_d$$

for any $t \in [0,T]$. Due to our constraint $\langle \pi_t, \mathbf{1}_d\rangle = h$ we realize that the optimal strategy is just a uniform diversification strategy taking the constraint on the bond investment into account. What is left to show to establish duality between (4.30) and our original optimization problem (4.26) is that μ^* is the worst possible parameter in K, given that an investor chooses strategy π^*. The parameter μ that minimizes $\mathbb{E}_\mu[U_\gamma(X_T^{\pi^*})]$ is, as we have seen before, the one that minimizes $(\pi_0^*)^\top \mu$. But

$$(\pi_0^*)^\top \mu = \frac{h}{d}\mathbf{1}_d^\top \mu = \frac{h}{d} b$$

for any $\mu \in K$. Thus, any $\mu \in K$ is a minimizer. In particular, we can deduce the following.

Corollary 4.13. *Let* $K = \{\mu \in \mathbb{R}^d \,|\, \mathbf{1}_d^\top \mu = b\}$. *Then*

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \inf_{\mu \in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] = \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] = \inf_{\mu \in K} \sup_{\pi \in \mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big],$$

where μ^* *and* π^* *are defined as in Theorem 4.12.*

Together with our previous considerations, the proof of the above corollary is analogous to the one of Theorem 4.11.

In conclusion, we have shown that our duality approach and the corresponding minimax theorem are not restricted to ellipsoidal uncertainty sets. In this section our approach has been applied to an uncertainty set of the form

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, \mathbf{1}_d^\top \mu = b\big\}$$

for some $b \in \mathbb{R}$, i.e. where the drift of the single assets is unknown, but one has knowledge about the overall performance of the collection of assets via the sum of the single drifts. Here we get an explicit representation of the worst-case parameter and obtain that the optimal robust strategy is the uniform diversification strategy. We will see in Chapter 5 how this fits into the framework of our results for ellipsoidal uncertainty sets when we let the degree of uncertainty κ, i.e. the radius of the uncertainty ellipsoid, go to infinity. Furthermore, in Corollary 4.13 we have also established duality between our original optimization problem and the dual one with supremum and infimum interchanged.

5. Asymptotic Behavior as Uncertainty Increases

In this chapter we consider once more the setting from Section 4.1 with ellipsoidal uncertainty sets and investigate what happens as the degree of uncertainty increases. Recall that we consider power or logarithmic utility $U_\gamma \colon \mathbb{R}_+ \to \mathbb{R}$, $\gamma \in (-\infty, 1)$, with $U_\gamma(x) = \frac{x^\gamma}{\gamma}$ if $\gamma \neq 0$ and $U_0(x) = \log(x)$. Our robust utility maximization problem has the form

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \inf_{\mu \in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big], \tag{5.1}$$

where

$$\mathcal{A}_h(x_0) = \big\{\pi \in \mathcal{A}(x_0) \,\big|\, \langle \pi_t, \mathbf{1}_d \rangle = h \text{ for all } t \in [0,T]\big\}$$

for some $h > 0$ and where we set

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, (\mu - \nu)^\top \Gamma^{-1}(\mu - \nu) \leq \kappa^2\big\}$$

for some $\kappa > 0$, $\nu \in \mathbb{R}^d$ and $\Gamma = \tau\tau^\top \in \mathbb{R}^{d\times d}$ a symmetric and positive-definite matrix. The set K is an ellipsoid with radius κ. By increasing κ we increase the degree of uncertainty about the true drift parameter.

5.1. Limit of worst-case parameter and optimal strategy

In the following, we analyze the optimal strategy π^* for problem (5.1) and the corresponding worst-case parameter μ^* in more detail and investigate their behavior as the degree of uncertainty about the drift parameter grows, i.e. when increasing the value κ in the uncertainty ellipsoid $K = \{\mu \in \mathbb{R}^d \,|\, (\mu - \nu)^\top \Gamma^{-1}(\mu - \nu) \leq \kappa^2\}$. In Theorem 4.8 we have stated representations of both μ^* and π^*. Theorem 4.11 then ensures that they really form a solution to (5.1). We use the notation from Section 4.1 again, in particular the matrix $A \in \mathbb{R}^{d\times d}$ and the vector $c \in \mathbb{R}^d$ are given as in Definition 4.2. Recall that for μ^* we have the representation

$$\mu^* = \nu - \tau \sum_{i=1}^d \left(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\right)^{-1} \left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i \right\rangle v_i, \tag{5.2}$$

where $\lambda_1 = 0, \lambda_2, \dots, \lambda_d$ are eigenvalues to $\tau^\top A \tau$ and $v_1, \dots, v_d$ are corresponding orthogonal eigenvectors. The value $\psi(\kappa) \in (0, \kappa]$ is such that $\|\tau^{-1}(\mu^* - \nu)\| = \kappa$. For the optimal strategy π^* we have shown that

$$\pi_t^* = \frac{1}{1-\gamma} A\mu^* + hc$$

for all $t \in [0,T]$.

Neither the matrix $\tau^\top A\tau$ nor the vector c depend on κ. The only quantity in representation (5.2) of the parameter μ^* that does depend on κ is the value $\psi(\kappa)$. To underline the dependence of the worst-case parameter and the optimal strategy on κ, we write $\mu^* = \mu^*(\kappa)$ and $\pi^* = \pi^*(\kappa)$ in the following. The next lemma characterizes the asymptotic behavior of $\psi(\kappa)$ as κ goes to infinity.

Lemma 5.1. *It holds*

$$\lim_{\kappa\to\infty} \frac{\psi(\kappa)}{\kappa} = 1.$$

Proof. As before, by acknowledging the dependence on κ, we write

$$a_i(\kappa) = -\left(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\right)^{-1} \left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\right\rangle$$

for $i = 1, \dots, d$, so that

$$\tau^{-1}\big(\mu^*(\kappa) - \nu\big) = \sum_{i=1}^d a_i(\kappa) v_i.$$

We have already seen in the proof of Lemma 4.7 that $a_1(\kappa) = -\psi(\kappa)$. Hence, the constraint $\|\tau^{-1}(\mu^* - \nu)\| = \kappa$ implies

$$\kappa^2 = \|\tau^{-1}(\mu^* - \nu)\|^2 = \sum_{i=1}^d a_i(\kappa)^2 = \psi(\kappa)^2 + \sum_{i=2}^d a_i(\kappa)^2 \tag{5.3}$$

due to orthonormality of the vectors $v_1, \dots, v_d$. We rewrite (5.3) as

$$\left(\frac{\psi(\kappa)}{\kappa}\right)^2 = 1 - \sum_{i=2}^d \left(\frac{a_i(\kappa)}{\kappa}\right)^2. \tag{5.4}$$

In the following, we show that the sum in the expression above goes to zero as κ goes to infinity. To prove this, take some $i \in \{2, \dots, d\}$. We know that

$$\left(\frac{a_i(\kappa)}{\kappa}\right)^2 = \frac{1}{\kappa^2}\left(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\right)^{-2} \left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\right\rangle^2,$$

where the expression in the inner product does not depend on κ. For the other factor, recall that $\psi(\kappa) > 0$ and $\lambda_i > 0$. Hence,

$$\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|} > \frac{\lambda_i}{1-\gamma} > 0$$

and therefore

$$\frac{1}{\kappa^2}\left(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\right)^{-2} \le \frac{1}{\kappa^2}\left(\frac{\lambda_i}{1-\gamma}\right)^{-2},$$

where the upper bound goes to zero as κ goes to infinity. Now we can deduce that

$$\lim_{\kappa\to\infty} \left(\frac{a_i(\kappa)}{\kappa}\right)^2 = 0,$$

which together with (5.4) implies

$$\lim_{\kappa\to\infty}\left(\frac{\psi(\kappa)}{\kappa}\right)^2 = 1.$$

The claim now follows from the fact that $\psi(\kappa)$ is positive for each κ. □

From the preceding lemma we gain insights into the asymptotic behavior of the parameter $\mu^*(\kappa)$, too. For big values of κ, one sees that $\mu^*(\kappa) - \nu$ is essentially a multiple of τv_1 where v_1 is an eigenvector to eigenvalue zero of $\tau^\top A\tau$. The contribution of the other eigenvectors $v_2, \dots, v_d$ becomes negligible. This observation is formalized in the following proposition.

Proposition 5.2. *It holds*

$$\lim_{\kappa\to\infty}\frac{1}{\kappa}\tau^{-1}\big(\mu^*(\kappa)-\nu\big) = -v_1 = -\frac{1}{\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}\mathbf{1}_d$$

and

$$\lim_{\kappa\to\infty}\frac{1}{\kappa}\mu^*(\kappa) = -\tau v_1 = -\frac{1}{\|\tau^{-1}\mathbf{1}_d\|}\mathbf{1}_d.$$

Proof. As in the previous proof, let

$$a_i(\kappa) = -\left(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\right)^{-1}\left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\right\rangle$$

for $i = 1, \dots, d$, so that

$$\tau^{-1}\big(\mu^*(\kappa)-\nu\big) = \sum_{i=1}^{d} a_i(\kappa)v_i.$$

Like in the previous proof, and using the result from the previous lemma, we can now deduce that

$$\frac{1}{\kappa}\tau^{-1}\big(\mu^*(\kappa)-\nu\big) = \frac{a_1(\kappa)}{\kappa}v_1 + \sum_{i=2}^{d}\frac{a_i(\kappa)}{\kappa}v_i = -\frac{\psi(\kappa)}{\kappa}v_1 + \sum_{i=2}^{d}\frac{a_i(\kappa)}{\kappa}v_i$$

goes to $-v_1$ as κ goes to infinity. Now we easily see

$$\lim_{\kappa\to\infty}\frac{1}{\kappa}\mu^*(\kappa) = \lim_{\kappa\to\infty}\frac{1}{\kappa}\tau\tau^{-1}\big(\mu^*(\kappa)-\nu\big) = -\tau v_1,$$

which proves the second claim. □

We have seen before that the worst possible parameter $\mu^*(\kappa)$ always lies on the boundary of the domain $K = \{\mu \in \mathbb{R}^d \,|\, (\mu-\nu)^\top\Gamma^{-1}(\mu-\nu) \le \kappa^2\}$, i.e. $\|\tau^{-1}(\mu^*(\kappa)-\nu)\| = \kappa$. In the above proposition we have seen that $\frac{1}{\kappa}\mu^*(\kappa)$ tends to

$$-\frac{1}{\|\tau^{-1}\mathbf{1}_d\|}\mathbf{1}_d.$$

So, asymptotically the direction of the worst-case parameter is simply $-\mathbf{1}_d$. That means that as κ tends to infinity, the worst drift that the market can choose for an investor who applies the optimal strategy π^*, is a drift vector where all entries are the same and negative. Since the constraint $\langle\pi_t, \mathbf{1}_d\rangle = h$ for all $t \in [0,T]$ forces the investor to have a positive position in at least some of the stocks it is intuitive that the market chooses negative drift in response to that. We have the following result for the asymptotic behavior of the investor's optimal strategy.

Corollary 5.3. *For any $t \in [0,T]$ it holds*

$$\lim_{\kappa\to\infty} \pi_t^*(\kappa) = \frac{h}{\mathbf{1}_d^\top \Gamma^{-1}\mathbf{1}_d}\Gamma^{-1}\mathbf{1}_d.$$

Proof. Recall that by Lemma 4.9 we can write

$$\pi_t^*(\kappa) = -\frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\Gamma^{-1}\big(\mu^*(\kappa)-\nu\big) = -\frac{h}{\|\tau^{-1}\mathbf{1}_d\|}\frac{\kappa}{\psi(\kappa)}\frac{1}{\kappa}\Gamma^{-1}\big(\mu^*(\kappa)-\nu\big)$$

for any $t \in [0,T]$. We then obtain

$$\lim_{\kappa\to\infty} \pi_t^*(\kappa) = \frac{h}{\|\tau^{-1}\mathbf{1}_d\|}(\tau^\top)^{-1}v_1 = \frac{h}{\|\tau^{-1}\mathbf{1}_d\|^2}(\tau\tau^\top)^{-1}\mathbf{1}_d = \frac{h}{\mathbf{1}_d^\top \Gamma^{-1}\mathbf{1}_d}\Gamma^{-1}\mathbf{1}_d$$

by combining the results from Lemma 5.1 and Proposition 5.2. □

The above corollary shows that for an investor who wants to maximize

$$\inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

over all strategies $\pi \in \mathcal{A}_h(x_0)$, the optimal strategy $\pi^*(\kappa)$ tends to a strategy that is deterministic and constant in time as the degree of uncertainty κ goes to infinity. This limit strategy consists in investing a certain constant fraction in the assets at any point in time. Note that the constraint $\langle \pi_t, \mathbf{1}_d\rangle = h$ for all $t \in [0,T]$ is fulfilled by the limit strategy since

$$\left\langle \frac{h}{\mathbf{1}_d^\top \Gamma^{-1}\mathbf{1}_d}\Gamma^{-1}\mathbf{1}_d, \mathbf{1}_d \right\rangle = \frac{h}{\mathbf{1}_d^\top \Gamma^{-1}\mathbf{1}_d}\mathbf{1}_d^\top \Gamma^{-1}\mathbf{1}_d = h.$$

Interestingly, the above strategy does not depend on the volatility matrix σ. So for a sufficiently high degree of uncertainty, the influence of the volatility matrix σ is negligible for the optimal strategy as it is dominated by the uncertainty present in the model. For any fixed degree of uncertainty κ, the volatility matrix σ still enters the optimal strategy via the matrix $A = D^\top(D\sigma\sigma^\top D^\top)^{-1}D$.

5.2. Relaxing the investment constraint

The above results about the asymptotic behavior of the worst-case parameter and the optimal strategy can also be used to show that, as uncertainty κ goes to infinity, our robust optimization problem yields the same optimal value as a slightly different optimization problem with a more general class of admissible strategies. Recall that we have so far considered for $h > 0$ the set

$$\mathcal{A}_h(x_0) = \big\{\pi \in \mathcal{A}(x_0) \,\big|\, \langle \pi_t, \mathbf{1}_d\rangle = h \text{ for all } t \in [0,T]\big\}$$

as the class of admissible strategies. Requiring $\langle \pi_t, \mathbf{1}_d\rangle \geq h$ instead of $\langle \pi_t, \mathbf{1}_d\rangle = h$ obviously enlarges this set. In the following, we show for logarithmic utility that maximizing the worst-case expected utility among bounded strategies in this larger set asymptotically leads to the same value as our original problem. We write $K = K(\kappa)$ for the uncertainty ellipsoid with radius κ to underline the dependence on κ.

Proposition 5.4. *Define for $h > 0$ the admissibility set*

$$\mathcal{A}'_h(x_0) = \big\{\pi \in \mathcal{A}(x_0) \,\big|\, \langle \pi_t, \mathbf{1}_d\rangle \geq h \text{ for all } t \in [0,T]\big\}$$

and let $M > 0$. Then there exists a $\kappa_M > 0$ such that for all $\kappa \geq \kappa_M$ it holds

$$\sup_{\substack{\pi\in\mathcal{A}'_h(x_0)\\ \|\pi\|\leq M}} \inf_{\mu\in K(\kappa)} \mathbb{E}_\mu\big[\log(X_T^\pi)\big] \leq \sup_{\pi\in\mathcal{A}_h(x_0)} \inf_{\mu\in K(\kappa)} \mathbb{E}_\mu\big[\log(X_T^\pi)\big].$$

Here we use $\|\pi\| \leq M$ as a short notation for $\|\pi_t\| \leq M$ for all $t \in [0,T]$.

Proof. Let $\pi' \in \mathcal{A}'_h(x_0)$ with $\|\pi'\| \leq M$. Then π' can be decomposed as

$$\pi'_t = \pi_t + \varepsilon_t \mathbf{1}_d$$

for all $t \in [0,T]$, where $\pi = (\pi_t)_{t\in[0,T]} \in \mathcal{A}_h(x_0)$ and $\varepsilon_t \geq 0$ for all $t \in [0,T]$. For any fixed $\mu \in K(\kappa)$ we rewrite the expected logarithmic utility given strategy π' as

$$\begin{aligned}\mathbb{E}_\mu\big[\log(X_T^{\pi'})\big] &= \log(x_0) + rT + \mathbb{E}_\mu\bigg[\int_0^T \Big((\pi'_t)^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi'_t\|^2\Big)\mathrm{d}t\bigg] \\ &= \mathbb{E}_\mu\big[\log(X_T^\pi)\big] + \mathbb{E}_\mu\bigg[\int_0^T \varepsilon_t\Big(\mathbf{1}_d^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\varepsilon_t\|\sigma^\top\mathbf{1}_d\|^2 - \mathbf{1}_d^\top\sigma\sigma^\top\pi_t\Big)\mathrm{d}t\bigg].\end{aligned}$$

In particular, we have

$$\begin{aligned}&\inf_{\mu\in K(\kappa)} \mathbb{E}_\mu\big[\log(X_T^{\pi'})\big] \leq \mathbb{E}_{\mu^*}\big[\log(X_T^{\pi'})\big] \\ &= \mathbb{E}_{\mu^*}\big[\log(X_T^\pi)\big] + \mathbb{E}_{\mu^*}\bigg[\int_0^T \varepsilon_t\Big(\mathbf{1}_d^\top(\mu^*(\kappa) - r\mathbf{1}_d) - \frac{1}{2}\varepsilon_t\|\sigma^\top\mathbf{1}_d\|^2 - \mathbf{1}_d^\top\sigma\sigma^\top\pi_t\Big)\mathrm{d}t\bigg],\end{aligned} \tag{5.5}$$

where $\mu^* = \mu^*(\kappa)$ is the worst-case parameter from Theorem 4.8. Our assumption $\|\pi'\| \leq M$ implies that also $\|\pi_t\|$ is bounded for every $t \in [0,T]$, and so is $\mathbf{1}_d^\top\sigma\sigma^\top\pi_t$. Hence the second summand in (5.5) becomes non-positive when κ is big enough (depending on M). That is because $\varepsilon_t \geq 0$ for all $t \in [0,T]$ and

$$\lim_{\kappa\to\infty} \mathbf{1}_d^\top\mu^*(\kappa) = \mathbf{1}_d^\top\nu - \lim_{\kappa\to\infty}\psi(\kappa)\mathbf{1}_d^\top\tau v_1 = \mathbf{1}_d^\top\nu - \lim_{\kappa\to\infty}\psi(\kappa)\frac{d}{\|\tau^{-1}\mathbf{1}_d\|} = -\infty.$$

So there exists a $\kappa_M > 0$ such that

$$\inf_{\mu\in K(\kappa)} \mathbb{E}_\mu\big[\log(X_T^{\pi'})\big] \leq \mathbb{E}_{\mu^*}\big[\log(X_T^\pi)\big]$$

for all $\kappa \geq \kappa_M$. Since κ_M depends only on M but not on the strategy π' or its decomposition, we can further deduce

$$\sup_{\substack{\pi\in\mathcal{A}'_h(x_0)\\ \|\pi\|\leq M}} \inf_{\mu\in K(\kappa)} \mathbb{E}_\mu\big[\log(X_T^\pi)\big] \leq \sup_{\pi\in\mathcal{A}_h(x_0)} \mathbb{E}_{\mu^*}\big[\log(X_T^\pi)\big] = \sup_{\pi\in\mathcal{A}_h(x_0)} \inf_{\mu\in K(\kappa)} \mathbb{E}_\mu\big[\log(X_T^\pi)\big]$$

for all $\kappa \geq \kappa_M$, which completes the proof. □

For power utility, the result is slightly weaker. We first give a lemma that states some useful equalities concerning the matrix A and vector c. These will be useful later as well.

Lemma 5.5. *For the matrix A and the vector c we have*

$$A\sigma\sigma^\top A = A, \quad c^\top \sigma\sigma^\top A = 0 \quad \text{and} \quad c^\top \mathbf{1}_d = 1.$$

Proof. Using the definition of A in Definition 4.2 we see that

$$A\sigma\sigma^\top A = D^\top(D\sigma\sigma^\top D^\top)^{-1}D\sigma\sigma^\top D^\top(D\sigma\sigma^\top D^\top)^{-1}D = D^\top(D\sigma\sigma^\top D^\top)^{-1}D = A,$$

and hence in particular

$$c^\top\sigma\sigma^\top A = e_d^\top(I_d - \sigma\sigma^\top A)\sigma\sigma^\top A = e_d^\top(\sigma\sigma^\top A - \sigma\sigma^\top A) = 0.$$

Further, we also have

$$c^\top\mathbf{1}_d = e_d^\top(I_d - \sigma\sigma^\top A)\mathbf{1}_d = e_d^\top\mathbf{1}_d = 1$$

due to $A\mathbf{1}_d = 0$. □

The next proposition gives a result similar to that in Proposition 5.4 for power utility. We define a different enlarged admissibility set $\overline{\mathcal{A}}_h(x_0)$ in this case. The reason is that, in contrast to the logarithmic utility case, we cannot ensure that we can restrict to deterministic strategies in $\mathcal{A}_h'(x_0)$.

Proposition 5.6. *Let $\gamma \neq 0$ and $h > 0$ and define the admissibility set*

$$\overline{\mathcal{A}}_h(x_0) = \bigcup_{h' \geq h} \mathcal{A}_{h'}(x_0).$$

Then there exists a $\kappa' > 0$ such that for all $\kappa \geq \kappa'$ it holds

$$\sup_{\pi\in\overline{\mathcal{A}}_h(x_0)} \inf_{\mu\in K(\kappa)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] = \sup_{\pi\in\mathcal{A}_h(x_0)} \inf_{\mu\in K(\kappa)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big].$$

Proof. Take an arbitrary strategy $\pi \in \overline{\mathcal{A}}_h(x_0)$. Then there exists some $h' \geq h$ such that $\pi \in \mathcal{A}_{h'}(x_0)$ and we know that

$$\inf_{\mu\in K(\kappa)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \leq \inf_{\mu\in K(\kappa)} \mathbb{E}_\mu\big[U_\gamma(X_T^{\pi'})\big] = \mathbb{E}_{\mu'}\big[U_\gamma(X_T^{\pi'})\big],$$

where $\mu' = \mu'(\kappa)$ is the minimizer of the function

$$\mu \mapsto \frac{1}{2(1-\gamma)}\mu^\top A\mu + h'c^\top\mu$$

on the uncertainty set $K(\kappa)$ and $\pi' = \pi'(\kappa) \equiv \frac{1}{1-\gamma}A\mu' + h'c$. In the following we show that for sufficiently large level of uncertainty

$$\mathbb{E}_{\mu'}\big[U_\gamma(X_T^{\pi'})\big] \leq \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] \tag{5.6}$$

where $\mu^* = \mu^*(\kappa)$ and $\pi^* = \pi^*(\kappa)$ are the worst-case parameter and the optimal strategy for the utility maximization among strategies in $\mathcal{A}_h(x_0)$. Note that for strategies π that are deterministic and constant in time we can write

$$\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] = \frac{x_0^\gamma}{\gamma}\exp\Big(\gamma T\Big(r + \pi_0^\top(\mu - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top \pi_0\|^2\Big)\Big)$$

for any $\mu \in K(\kappa)$, hence for showing (5.6) it is sufficient to prove

$$(\pi_0')^\top(\mu' - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top \pi_0'\|^2 \le (\pi_0^*)^\top(\mu^* - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top \pi_0^*\|^2. \tag{5.7}$$

Using the representation of π' we obtain

$$\begin{aligned}
&(\pi_0')^\top(\mu' - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top \pi_0'\|^2 \\
&= \frac{1}{1-\gamma}(\mu')^\top A\mu' + h'c^\top(\mu' - r\mathbf{1}_d) - \frac{1}{2(1-\gamma)}(\mu')^\top A\mu' - \frac{1-\gamma}{2}(h')^2 c^\top\sigma\sigma^\top c \\
&= \frac{1}{2(1-\gamma)}(\mu')^\top A\mu' + h'c^\top\mu' - h'r - \frac{1-\gamma}{2}(h')^2 c^\top\sigma\sigma^\top c.
\end{aligned}$$

In the first step we have used $A\mathbf{1}_d = 0$, $A\sigma\sigma^\top A = A$ and $c^\top\sigma\sigma^\top A = 0$, in the second step $c^\top\mathbf{1}_d = 1$, see Lemma 5.5. An analogous computation can be done for π^* and μ^*. We then see that, since μ' minimizes

$$\mu \mapsto \frac{1}{2(1-\gamma)}\mu^\top A\mu + h'c^\top\mu$$

on $K(\kappa)$, in particular it holds

$$\begin{aligned}
\frac{1}{2(1-\gamma)}(\mu')^\top A\mu' + h'c^\top\mu' &\le \frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^* + h'c^\top\mu^* \\
&= \frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^* + hc^\top\mu^* + (h'-h)c^\top\mu^*.
\end{aligned}$$

Using again $c^\top\mathbf{1}_d = 1$ it is easy to show that $c^\top\mu^* = c^\top\mu^*(\kappa)$ goes to minus infinity as κ goes to infinity. Hence we can choose $\kappa' > 0$ such that $c^\top\mu^* \le 0$ for all $\kappa \ge \kappa'$. Note that κ' does not depend on π'. For all $\kappa \ge \kappa'$ we then have

$$\begin{aligned}
&(\pi_0')^\top(\mu' - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top \pi_0'\|^2 \\
&\le \frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^* + hc^\top\mu^* + (h'-h)c^\top\mu^* - h'r - \frac{1-\gamma}{2}(h')^2 c^\top\sigma\sigma^\top c \\
&\le \frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^* + hc^\top\mu^* - hr - \frac{1-\gamma}{2}h^2 c^\top\sigma\sigma^\top c \\
&= (\pi_0^*)^\top(\mu^* - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top \pi_0^*\|^2,
\end{aligned}$$

which proves (5.7) and hence (5.6). Since κ' was chosen independent of h' or π', we deduce in particular

$$\sup_{\pi\in\overline{\mathcal{A}}_h(x_0)}\inf_{\mu\in K(\kappa)}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] \le \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] = \sup_{\pi\in\mathcal{A}_h(x_0)}\inf_{\mu\in K(\kappa)}\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]$$

for all $\kappa \ge \kappa'$. The reverse inequality holds trivially. □

The previous propositions show that as uncertainty increases it is reasonable for investors to choose strategies π with $\langle \pi_t, \mathbf{1}_d \rangle$ as small as possible. Even if the class of admissible strategies is enlarged, the optimal value will for large uncertainty be attained by a strategy from $\mathcal{A}_h(x_0)$. This is in line with the intuition from Proposition 2.1, where we have seen that as uncertainty exceeds a certain threshold, investors prefer to not invest anything into the risky assets. Here we have seen that when the level of uncertainty κ exceeds a threshold, investors seek to invest as little as possible in the risky assets, meaning that $\langle \pi_t, \mathbf{1}_d \rangle$ is chosen as small as possible in a given class of admissible strategies.

5.3. Special case: uncertainty ball

We now take again

$$\mathcal{A}_h(x_0) = \big\{\pi \in \mathcal{A}(x_0) \,\big|\, \langle \pi_t, \mathbf{1}_d \rangle = h \text{ for all } t \in [0,T]\big\}$$

for some $h > 0$ as the class of admissible strategies. So far, for the asymptotic analysis we have considered general uncertainty ellipsoids for the drift parameter μ , i.e. sets of the form

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, (\mu - \nu)^\top \Gamma^{-1} (\mu - \nu) \le \kappa^2\big\}$$

for some $\nu \in \mathbb{R}^d$ and a matrix $\Gamma \in \mathbb{R}^{d \times d}$ that is symmetric and positive definite. The set K is thus an ellipsoid where the shape is determined by the matrix Γ. An interesting special case is $\Gamma = I_d$. In that case, the uncertainty set is simply a ball

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, \|\mu - \nu\| \le \kappa\big\}.$$

Hence, the degree of uncertainty is the same for any component of the drift parameter μ. Every parameter within Euclidean distance κ of the center ν is deemed possible.

Corollary 5.7. *If the uncertainty set is of the form*

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, \|\mu - \nu\| \le \kappa\big\},$$

then the parameter $\mu^*(\kappa)$ fulfills

$$\lim_{\kappa \to \infty} \frac{1}{\kappa} \big(\mu^*(\kappa) - \nu\big) = -\frac{1}{\sqrt{d}} \mathbf{1}_d$$

and the optimal strategy fulfills

$$\lim_{\kappa \to \infty} \pi_t^*(\kappa) = \frac{h}{d} \mathbf{1}_d$$

for any $t \in [0,T]$.

Proof. When noting that $\tau = I_d$ for the given set K and $\|\mathbf{1}_d\| = \sqrt{d}$, the claims follow directly from Proposition 5.2, respectively from Corollary 5.3. $\square$

The above corollary implies that if the uncertainty set for the drift parameter μ is a ball with radius κ, then the optimal strategy converges to a uniform diversification strategy, given by $\frac{h}{d}\mathbf{1}_d$ at each point in time. Hence, when forced to invest a total fraction of $h > 0$ in the risky assets, then in the limit for κ going to infinity investors will diversify their portfolio uniformly, thus investing a constant fraction of $\frac{h}{d}$ in each risky asset. We illustrate this convergence to a uniform diversification strategy by an example.

Example 5.8. We consider a market with $d = 8$ risky assets. The volatility matrix is

$$\sigma = \begin{pmatrix} 0.3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.2 & 0.3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.2 & 0.3 & 0 & 0 & 0 & 0 & 0 \\ 0.3 & 0.2 & 0 & 0.4 & 0 & 0 & 0 & 0 \\ 0.2 & 0.3 & 0 & 0.1 & 0.3 & 0 & 0 & 0 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.2 & 0.2 & 0 & 0 \\ 0.2 & 0.1 & 0.2 & 0.1 & 0.2 & 0.2 & 0.4 & 0 \\ 0.1 & 0 & 0 & 0.2 & 0.1 & 0.1 & 0.2 & 0.4 \end{pmatrix}.$$

Investors use strategies from $\mathcal{A}_h(x_0)$ with $h = 1$, i.e. strategies π with $\langle \pi_t, \mathbf{1}_d \rangle = 1$ for all $t \in [0, T]$. The coefficient in the utility function U_γ is chosen to be $\gamma = \frac{1}{2}$. We take $\Gamma = I_d$ and $\nu = \frac{3}{10}\mathbf{1}_d$. We then compute the optimal strategy $\pi^*(\kappa)$ that is given in Theorem 4.8 as

$$\pi^*(\kappa) \equiv \frac{1}{1-\gamma} A\mu^*(\kappa) + hc,$$

where

$$\mu^*(\kappa) = \nu - \tau \sum_{i=1}^{d} \left(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|} \right)^{-1} \left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i \right\rangle v_i$$

for $\psi(\kappa) \in (0, \kappa]$ that is uniquely determined by $\|\tau^{-1}(\mu^* - \nu)\| = \kappa$. We calculate the resulting constant optimal composition for all $\kappa \in (0, 0.5)$ and plot the result in Figure 5.1 against κ. For any fixed level of uncertainty κ, the optimal composition $\pi^*(\kappa)$ is plotted as a stacked plot where every color corresponds to one stock. For small values of κ, the optimal π^* is negative in some components. This leads to an overall investment larger than one on the positive side. For $\kappa > 0.05$ the strategy $\pi^*(\kappa)$ has only positive entries. As κ becomes larger, the composition gets closer and closer to the uniform diversification vector.

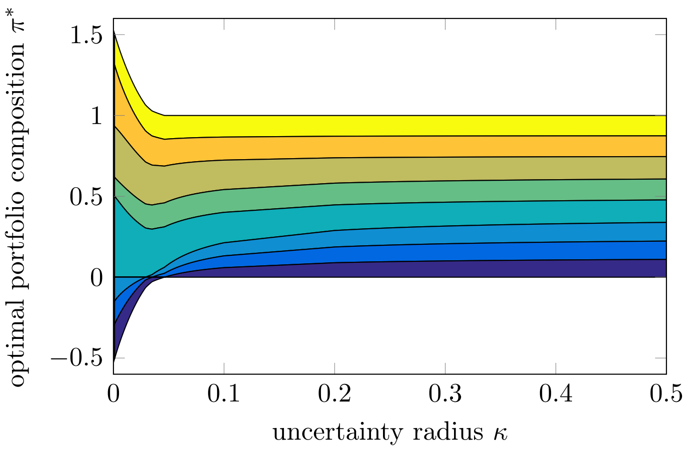


Figure 5.1.: The optimal portfolio composition π^* plotted against uncertainty radius κ for a market with $d = 8$ risky assets with parameters given in Example 5.8. The portfolio composition approaches a uniform diversification strategy for large values of κ.

5.4. Risk aversion and speed of convergence

We have seen in Section 5.1 that the optimal strategy $\pi^*(\kappa)$ for the robust optimization problem

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \inf_{\mu \in K} \mathbb{E}_\mu \big[U_\gamma(X_T^\pi) \big]$$

with ellipsoidal uncertainty sets

$$K = \big\{ \mu \in \mathbb{R}^d \,\big|\, (\mu - \nu)^\top \Gamma^{-1} (\mu - \nu) \le \kappa^2 \big\}$$

converges as the level of uncertainty κ goes to infinity. Recall from Corollary 5.3 that

$$\lim_{\kappa \to \infty} \pi_t^*(\kappa) = \frac{h}{\mathbf{1}_d^\top \Gamma^{-1} \mathbf{1}_d} \Gamma^{-1} \mathbf{1}_d$$

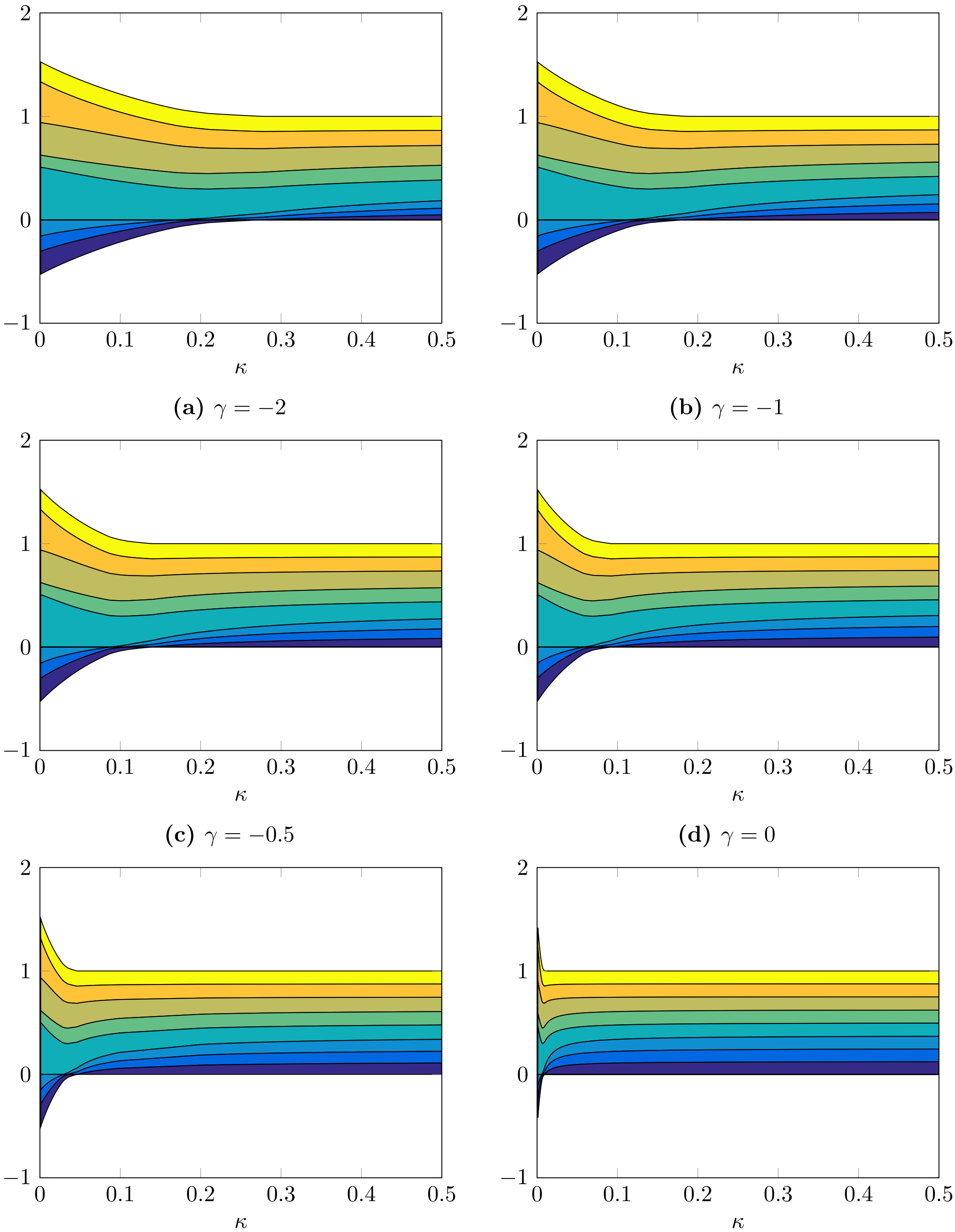
for any $t \in [0, T]$. If the uncertainty set K is a ball around ν, meaning that the matrix $\Gamma \in \mathbb{R}^{d \times d}$ is the identity matrix, then the above expression for the limit simplifies to $\frac{h}{d} \mathbf{1}_d$. So in the limit, the optimal strategy is a uniform diversification strategy.

In the following, we investigate which influence the risk aversion parameter γ has on the speed of the convergence. Recall that we consider the utility function $U_\gamma \colon \mathbb{R}_+ \to \mathbb{R}$ with $U_\gamma(x) = \frac{x^\gamma}{\gamma}$ if $\gamma \neq 0$ and $U_0(x) = \log(x)$. For this class of utility functions, the value $1 - \gamma$ is equal to the Arrow–Pratt measure of relative risk aversion. Note that for positive γ, the function U_γ is non-negative whereas it takes only negative values and is unbounded from below if γ is negative. These observations help to illustrate that the smaller γ is the more risk-averse is the investor.

In Figure 5.2 we plot the constant optimal portfolio composition $\pi^*(\kappa)$ against κ for various values of γ. For this purpose, we consider a market with $d = 8$ risky assets. As model parameters we take those from Example 5.8. In particular, we assume $\Gamma = I_d$, hence our convergence results ensure that $\pi^*(\kappa)$ will converge to a uniform diversification strategy.

The figure illustrates the convergence against the uniform diversification strategy. For each fixed value of γ one sees that for small values of κ the strategy is quite different from a uniform one. In particular, some entries of $\pi^*(\kappa)$ are negative. This changes for increasing value of κ.

When comparing the different subplots one sees that the value of γ seems to have a direct influence on the speed of convergence to the uniform strategy. Interestingly, the convergence is faster for higher values of γ, i.e. for less risk-averse investors. This might be surprising at first glance since one expects a more risk-averse investor to choose a "safer" strategy sooner than a less risk-averse investor does. However, the effect becomes more intuitive when keeping in mind that we address a robust optimization problem where an investor is confronted with the worst possible drift parameter in the uncertainty set. An investor with a high, positive value of γ would, in the non-robust problem, invest in the assets with the allegedly highest drift. In the worst-case market this undiversified strategy would allow the market to choose a very extreme drift parameter with high absolute values for exactly these assets. This implies that a less risk-averse investor is much more prone to the market's choice of a drift parameter. To make up for this, the optimal robust strategy converges very fast, so that even for small values of uncertainty κ, the investor is already driven into the diversified uniform strategy.



(a) $\gamma = -2$

(b) $\gamma = -1$

(c) $\gamma = -0.5$

(d) $\gamma = 0$

(e) $\gamma = 0.5$

(f) $\gamma = 0.9$

Figure 5.2.: Optimal portfolio composition π^* plotted against κ for different values of γ. The model parameters are those from Example 5.8. For any γ, we observe convergence against a uniform diversification strategy. For larger values of γ, convergence appears to take place faster than for smaller values of γ.

Remark 5.9. To see how the speed of convergence is influenced by γ, recall that the optimal strategy of an investor has the form

$$\pi_t^*(\kappa) = \frac{1}{1-\gamma} A\mu^*(\kappa) + hc$$

for all $t \in [0, T]$, where the worst-case parameter $\mu^*(\kappa)$ is given by

$$\mu^*(\kappa) = \nu + \tau \sum_{i=1}^{d} a_i(\kappa) v_i$$

with

$$a_i(\kappa) = -\left(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|} \right)^{-1} \left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i \right\rangle$$

for $i = 1, \ldots, d$. It follows that

$$\pi_t^*(\kappa) = \frac{1}{1-\gamma} A\nu + hc + \frac{1}{1-\gamma} A\tau \sum_{i=1}^{d} a_i(\kappa) v_i = \frac{1}{1-\gamma} A\nu + hc + \frac{1}{1-\gamma} A\tau \sum_{i=2}^{d} a_i(\kappa) v_i,$$

where we have used that $A\tau v_1 = 0$. Note that only the last term in the above expression depends on κ. When rewriting it we obtain

$$\begin{aligned}
\frac{1}{1-\gamma} A\tau \sum_{i=2}^{d} a_i(\kappa) v_i &= -\frac{1}{1-\gamma} A\tau \sum_{i=2}^{d} \left(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|} \right)^{-1} \left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i \right\rangle v_i \\
&= -A\tau \sum_{i=2}^{d} \left(\lambda_i + \frac{h(1-\gamma)}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|} \right)^{-1} \left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i \right\rangle v_i.
\end{aligned}$$

Recall that the coefficients all converge due to $\lambda_i > 0$ and

$$\lim_{\kappa \to \infty} \frac{\psi(\kappa)}{\kappa} = 1.$$

From the above representation we see that convergence takes place faster if $1 - \gamma$ is small, hence convergence is faster for large values of γ. This is in line with our observations from Figure 5.2.

5.5. Measures of robustness performance

We have seen in the previous sections that introducing uncertainty in our utility maximization problem leads to more diversified strategies. The question naturally arises what an investor gains from this robust behavior. On the other hand, one may also be interested in the loss in utility coming from behaving in a robust way in situations where it would not be necessary. These two antithetic effects can be measured by the performance measures *cost of ambiguity* and *reward for distributional robustness* that have already been studied in a different context in Analui [2, Sec. 3.4] for multistage stochastic optimization problems. We adapt these definitions to our setting.

Recall that our maximization problem is

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \inf_{\mu \in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big], \tag{5.8}$$

where the uncertainty set is an ellipsoid of the form

$$K = \big\{\mu \in \mathbb{R}^d \,\big|\, (\mu - \nu)^\top \Gamma^{-1}(\mu - \nu) \le \kappa^2\big\}.$$

The vector ν as the center of the uncertainty ellipsoid can be seen as an estimation for the true drift of the stocks. If there was no uncertainty present in the model and an investor was sure that the estimation ν was correct, then she would simply maximize $\mathbb{E}_\nu[U_\gamma(X_T^\pi)]$ over the admissible strategies, which corresponds to solving the problem

$$\sup_{\pi \in \mathcal{A}_h(x_0)} \mathbb{E}_\nu\big[U_\gamma(X_T^\pi)\big].$$

From Proposition 4.3 we know that the optimal strategy is then of the form $(\hat{\pi}_t)_{t\in[0,T]}$ with

$$\hat{\pi}_t = \frac{1}{1-\gamma} A\nu + hc \tag{5.9}$$

for all $t \in [0,T]$. In the presence of uncertainty, the solution to (5.8) is given by the strategy $(\pi_t^*)_{t\in[0,T]}$ with

$$\pi_t^* = \frac{1}{1-\gamma} A\mu^* + hc \tag{5.10}$$

for all $t \in [0,T]$, see Theorem 4.11. We have an implicit representation of the parameter μ^* in Theorem 4.8. Recall that μ^*, and hence also π^*, depends on the radius κ of the uncertainty set and that μ^* is the worst parameter in K, given that an investor chooses strategy π^*. We are now able to define the following measures for robustness performance.

Definition 5.10. We define the *cost of ambiguity* as

$$\mathrm{COA} = \mathbb{E}_\nu\big[U_\gamma(X_T^{\hat{\pi}})\big] - \mathbb{E}_\nu\big[U_\gamma(X_T^{\pi^*})\big]$$

and the *reward for distributional robustness* as

$$\mathrm{RDR} = \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] - \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\hat{\pi}})\big].$$

The cost of ambiguity captures how big the loss in utility is when using the robust strategy π^*, given that the estimation ν for the drift was actually correct. Note that $\hat{\pi}$ is the best strategy given drift ν, hence COA is always non-negative. In the reward for distributional robustness on the other hand, it is reflected how much an investor is rewarded when using the robust strategy π^* compared to the "naive" strategy $\hat{\pi}$, assuming that indeed the worst possible drift parameter μ^* is the true one. We see that also RDR is non-negative since π^* maximizes expected utility given the parameter μ^*.

In the following we investigate the qualitative behavior of the performance measures COA and RDR. In particular, we analyze the influence of the risk aversion parameter γ and of the level of uncertainty κ.

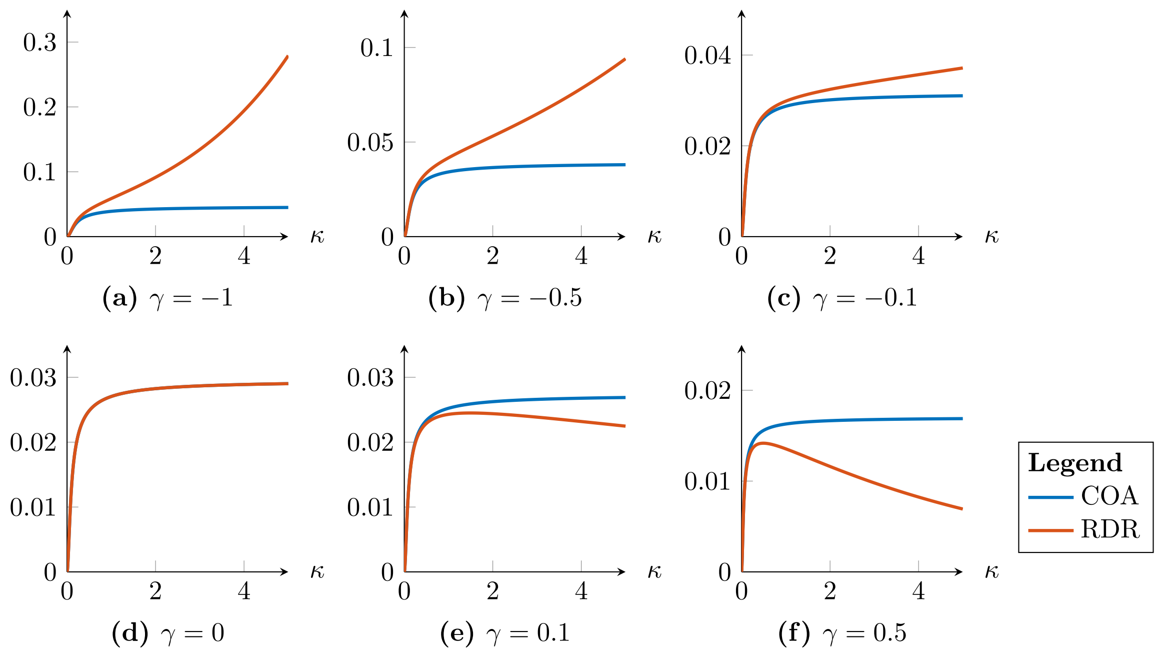


Figure 5.3.: The behavior of COA and RDR plotted against uncertainty radius κ for different values of the risk aversion coefficient γ. The parameters are those from Example 5.8.

In Figure 5.3 we plot COA and RDR against κ for different values of the risk aversion parameter γ to give a first impression. The model parameters are the same as in Example 5.8, in particular the number of stocks is $d = 8$. Note that the scaling is different among the single subplots and that the absolute values of both COA and RDR become smaller with an increasing value of γ.

We see a qualitative difference between COA and RDR. Whereas COA is in our example always increasing in κ, the monotonicity behavior of RDR seems to depend on γ. For negative values of γ we observe that RDR is increasing in κ. For positive values of γ, however, RDR is increasing for small values of κ and decreasing for bigger values of κ. What also changes with γ is the relation between COA and RDR. Note that for $\gamma = 0$, corresponding to the logarithmic utility case, the cost of ambiguity seems to equal the reward for distributional robustness for any level of uncertainty κ. For $\gamma < 0$ we observe COA $\leq$ RDR, for $\gamma > 0$ on the other hand COA $\geq$ RDR. In the following, we verify these conjectures. It is helpful to write COA and RDR in a more explicit way.

Lemma 5.11. *If $\gamma = 0$, then*

$$\mathrm{COA} = \frac{T}{2}(\mu^* - \nu)^\top A(\mu^* - \nu) = \mathrm{RDR}\,.$$

If $\gamma \neq 0$, then

$$\mathrm{COA} = \frac{M(\gamma)}{\gamma} L(\gamma, \kappa) \exp\Big(\gamma T \Big(hc^\top \nu + \frac{1}{2(1-\gamma)} \nu^\top A \nu\Big)\Big)$$

and

$$\mathrm{RDR} = \frac{M(\gamma)}{\gamma} L(\gamma, \kappa) \exp\Big(\gamma T \Big(hc^\top \mu^* + \frac{1}{2(1-\gamma)} (\mu^*)^\top A \mu^*\Big)\Big),$$

where

$$M(\gamma) = x_0^\gamma \exp\Big(\gamma T\Big((1-h)r - \frac{1-\gamma}{2} h^2 c^\top \sigma\sigma^\top c\Big)\Big)$$

and

$$L(\gamma,\kappa) = 1 - \exp\Big(-\frac{\gamma T}{2(1-\gamma)}(\mu^* - \nu)^\top A(\mu^* - \nu)\Big).$$

Proof. We address the case $\gamma = 0$ first. Note that both $\hat{\pi}$ and π^* are constant in time and deterministic. Therefore, the expected logarithmic utility of terminal wealth can be calculated explicitly for both strategies and we get

$$\begin{aligned}
\mathrm{COA} &= \mathbb{E}_\nu\big[\log(X_T^{\hat{\pi}})\big] - \mathbb{E}_\nu\big[\log(X_T^{\pi^*})\big] \\
&= T\Big(\hat{\pi}_0^\top(\nu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \hat{\pi}_0\|^2\Big) - T\Big((\pi_0^*)^\top(\nu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \pi_0^*\|^2\Big) \\
&= T\Big((\hat{\pi}_0 - \pi_0^*)^\top(\nu - r\mathbf{1}_d) - \frac{1}{2}\big(\|\sigma^\top\hat{\pi}_0\|^2 - \|\sigma^\top \pi_0^*\|^2\big)\Big).
\end{aligned}$$

Next, we plug in the representations for $\hat{\pi}$ and π^* from (5.9) and (5.10) and obtain

$$\begin{aligned}
\mathrm{COA} &= T\Big((\nu - \mu^*)^\top A(\nu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top A\nu + h\sigma^\top c\|^2 + \frac{1}{2}\|\sigma^\top A\mu^* + h\sigma^\top c\|^2\Big) \\
&= T\Big(\nu^\top A\nu - (\mu^*)^\top A\nu - \frac{1}{2}\big(\nu^\top A\sigma\sigma^\top A\nu + 2hc^\top\sigma\sigma^\top A\nu + h^2c^\top\sigma\sigma^\top c\big) \\
&\quad + \frac{1}{2}\big((\mu^*)^\top A\sigma\sigma^\top A\mu^* + 2hc^\top\sigma\sigma^\top A\mu^* + h^2c^\top\sigma\sigma^\top c\big)\Big) \qquad (5.11)\\
&= T\Big(\nu^\top A\nu - (\mu^*)^\top A\nu - \frac{1}{2}\nu^\top A\nu + \frac{1}{2}(\mu^*)^\top A\mu^*\Big) \\
&= \frac{T}{2}(\mu^* - \nu)^\top A(\mu^* - \nu).
\end{aligned}$$

In the second step we have used that $A\mathbf{1}_d = 0$ and in the third step the properties $A\sigma\sigma^\top A = A$ and $c^\top\sigma\sigma^\top A = 0$ from Lemma 5.5. The calculation of RDR is very similar. We first see

$$\mathrm{RDR} = \mathbb{E}_{\mu^*}\big[\log(X_T^{\pi^*})\big] - \mathbb{E}_{\mu^*}\big[\log(X_T^{\hat{\pi}})\big] = T\Big((\pi_0^* - \hat{\pi}_0)^\top(\mu^* - r\mathbf{1}_d) - \frac{1}{2}\big(\|\sigma^\top\pi_0^*\|^2 - \|\sigma^\top\hat{\pi}_0\|^2\big)\Big).$$

Plugging in the representations of $\hat{\pi}$ and π^* and using $A\mathbf{1}_d = 0$ then yields

$$\mathrm{RDR} = T\Big((\mu^* - \nu)^\top A\mu^* - \frac{1}{2}\|\sigma^\top A\mu^* + h\sigma^\top c\|^2 + \frac{1}{2}\|\sigma^\top A\nu + h\sigma^\top c\|^2\Big).$$

Now, in analogy to (5.11) we make use of $A\sigma\sigma^\top A = A$ and $c^\top\sigma\sigma^\top A = 0$ and finally get

$$\mathrm{RDR} = T\Big((\mu^*)^\top A\mu^* - \nu^\top A\mu^* - \frac{1}{2}(\mu^*)^\top A\mu^* + \frac{1}{2}\nu^\top A\nu\Big) = \frac{T}{2}(\mu^* - \nu)^\top A(\mu^* - \nu) = \mathrm{COA}\,.$$

Next, we assume $\gamma \neq 0$. Again, we can make use of the fact that $\hat{\pi}$ and π^* are deterministic strategies and constant in time. We thus get

$$\begin{aligned}
\mathrm{COA} &= \mathbb{E}_\nu\big[U_\gamma(X_T^{\hat{\pi}})\big] - \mathbb{E}_\nu\big[U_\gamma(X_T^{\pi^*})\big] \\
&= \frac{x_0^\gamma}{\gamma}\exp\Big(\gamma T\Big(r + \hat{\pi}_0^\top(\nu - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\hat{\pi}_0\|^2\Big)\Big) \\
&\quad - \frac{x_0^\gamma}{\gamma}\exp\Big(\gamma T\Big(r + (\pi_0^*)^\top(\nu - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\pi_0^*\|^2\Big)\Big).
\end{aligned}$$

Next, we plug in (5.9) and (5.10) and obtain

$$\begin{aligned}\mathrm{COA} = \frac{x_0^\gamma \mathrm{e}^{\gamma r T}}{\gamma}\Big(&\exp\Big(\gamma T\Big(\frac{1}{1-\gamma}\nu^\top A\nu + hc^\top(\nu - r\mathbf{1}_d) - \frac{1-\gamma}{2}\Big\|\frac{1}{1-\gamma}\sigma^\top A\nu + h\sigma^\top c\Big\|^2\Big)\Big) \\ &- \exp\Big(\gamma T\Big(\frac{1}{1-\gamma}(\mu^*)^\top A\nu + hc^\top(\nu - r\mathbf{1}_d) - \frac{1-\gamma}{2}\Big\|\frac{1}{1-\gamma}\sigma^\top A\mu^* + h\sigma^\top c\Big\|^2\Big)\Big)\Big).\end{aligned}$$

In the next step, just like for $\gamma = 0$, we split the terms with the squared Euclidean norm into summands. We use again $A\sigma\sigma^\top A = A$ and $c^\top\sigma\sigma^\top A = 0$ to cancel some terms. Additionally, we recall that $c^\top\mathbf{1}_d = 1$ due to Lemma 5.5. Then, by using the notation given in the lemma and completing the square, the above expression becomes

$$\begin{aligned}\mathrm{COA} &= \frac{M(\gamma)}{\gamma}\Big(\exp\Big(\gamma T\Big(\frac{1}{1-\gamma}\nu^\top A\nu + hc^\top\nu - \frac{1}{2(1-\gamma)}\nu^\top A\nu\Big)\Big) \\ &\qquad\qquad - \exp\Big(\gamma T\Big(\frac{1}{1-\gamma}(\mu^*)^\top A\nu + hc^\top\nu - \frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^*\Big)\Big)\Big) \\ &= \frac{M(\gamma)}{\gamma}\exp\Big(\gamma T\Big(hc^\top\nu + \frac{1}{2(1-\gamma)}\nu^\top A\nu\Big)\Big)\Big(1 - \exp\Big(-\frac{\gamma T}{2(1-\gamma)}(\mu^* - \nu)^\top A(\mu^* - \nu)\Big)\Big) \\ &= \frac{M(\gamma)}{\gamma}L(\gamma,\kappa)\exp\Big(\gamma T\Big(hc^\top\nu + \frac{1}{2(1-\gamma)}\nu^\top A\nu\Big)\Big).\end{aligned}$$

For the reward of distributional robustness the procedure is again similar. Firstly, we see

$$\begin{aligned}\mathrm{RDR} &= \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] - \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\hat{\pi}})\big] \\ &= \frac{x_0^\gamma}{\gamma}\exp\Big(\gamma T\Big(r + (\pi_0^*)^\top(\mu^* - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\pi_0^*\|^2\Big)\Big) \\ &\quad - \frac{x_0^\gamma}{\gamma}\exp\Big(\gamma T\Big(r + \hat{\pi}_0^\top(\mu^* - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\hat{\pi}_0\|^2\Big)\Big).\end{aligned}$$

Plugging in the representations from (5.9) and (5.10) and simplifying with the same tools as before then gives

$$\begin{aligned}\mathrm{RDR} &= \frac{M(\gamma)}{\gamma}\Big(\exp\Big(\gamma T\Big(\frac{1}{1-\gamma}(\mu^*)^\top A\mu^* + hc^\top\mu^* - \frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^*\Big)\Big) \\ &\qquad\qquad - \exp\Big(\gamma T\Big(\frac{1}{1-\gamma}\nu^\top A\mu^* + hc^\top\mu^* - \frac{1}{2(1-\gamma)}\nu^\top A\nu\Big)\Big)\Big) \\ &= \frac{M(\gamma)}{\gamma}L(\gamma,\kappa)\exp\Big(\gamma T\Big(hc^\top\mu^* + \frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^*\Big)\Big).\end{aligned}$$

This gives the representations in the lemma for $\gamma \neq 0$. □

Lemma 5.11 gives COA and RDR as explicit functions of the parameters ν and μ^*. Note that the expression $M(\gamma)$ appearing in the case $\gamma \neq 0$ is strictly positive for any γ. Further, it does not depend on μ^* and hence not on the degree of uncertainty κ.

The lemma already gives a relation between cost of ambiguity and reward for distributional robustness. For $\gamma = 0$, i.e. logarithmic utility, the two measures are equal, independent of the degree of model uncertainty. Recall that we already noted this in our numerical example from Figure 5.3. That figure also suggests that $\mathrm{COA} \geq \mathrm{RDR}$ if γ is positive and $\mathrm{COA} \leq \mathrm{RDR}$ if γ is negative. We prove this statement in the following proposition.

Proposition 5.12. *The following relation holds between cost of ambiguity and reward for distributional robustness:*

(i) $\mathrm{COA} \geq \mathrm{RDR}$ *if* $\gamma > 0$;

(ii) $\mathrm{COA} = \mathrm{RDR}$ *if* $\gamma = 0$;

(iii) $\mathrm{COA} \leq \mathrm{RDR}$ *if* $\gamma < 0$.

Proof. Note that the equality in the case $\gamma = 0$ is already proven in Lemma 5.11. Now take $\gamma \neq 0$. By using the explicit representation of cost of ambiguity and reward for distributional robustness from Lemma 5.11 we obtain

$$\frac{\mathrm{COA}}{\mathrm{RDR}} = \exp\Big(\gamma T \Big(hc^\top \nu + \frac{1}{2(1-\gamma)}\nu^\top A \nu - hc^\top \mu^* - \frac{1}{2(1-\gamma)}(\mu^*)^\top A \mu^*\Big)\Big). \tag{5.12}$$

When reconsidering our method to find the worst-case parameter in Section 4.1.2 we recall that $\mu^* \in K$ is the parameter that minimizes the function

$$K \ni \mu \mapsto \frac{1}{2(1-\gamma)}\mu^\top A \mu + hc^\top \mu.$$

Since $\nu \in K$, it follows in particular that

$$\frac{1}{2(1-\gamma)}(\mu^*)^\top A \mu^* + hc^\top \mu^* \leq \frac{1}{2(1-\gamma)}\nu^\top A \nu + hc^\top \nu.$$

Therefore, we obtain

$$hc^\top \nu + \frac{1}{2(1-\gamma)}\nu^\top A \nu - hc^\top \mu^* - \frac{1}{2(1-\gamma)}(\mu^*)^\top A \mu^* \geq 0. \tag{5.13}$$

Combining (5.13) and (5.12) implies

$$\frac{\mathrm{COA}}{\mathrm{RDR}} \geq 1 \text{ if } \gamma > 0, \quad \frac{\mathrm{COA}}{\mathrm{RDR}} \leq 1 \text{ if } \gamma < 0,$$

which proves our claim. □

The previous proposition contains information about the relation between cost of ambiguity and reward for distributional robustness. It turns out that the risk aversion parameter γ in the investor's utility function plays a crucial role here. For positive γ the cost of ambiguity is larger than the reward for distributional robustness, for negative γ the relation is the other way around. Logarithmic utility is the boundary case where the two measures of robustness performance are equal.

Figure 5.3 suggests that the parameter γ also has an effect on the monotonicity behavior of RDR in dependence on κ. Our next result helps to verify this by investigating the asymptotic behavior of cost of ambiguity and reward for distributional robustness as the level of uncertainty tends to infinity. Since we want to study the influence of the level of uncertainty on COA and RDR we write $\mathrm{COA}(\kappa)$ and $\mathrm{RDR}(\kappa)$ in the following to emphasize the dependence on the parameter κ.

Proposition 5.13. *(i) If* $h\tau^\top c + \frac{1}{1-\gamma}A\tau^{-1}\nu \in \operatorname{span}(\{v_1\})$*, then* $\mathrm{COA}(\kappa) = 0 = \mathrm{RDR}(\kappa)$ *for any* $\kappa > 0$.

(ii) If $h\tau^\top c + \frac{1}{1-\gamma}A\tau^{-1}\nu \notin \operatorname{span}(\{v_1\})$*, then* $\mathrm{COA}(\kappa)$ *converges for* κ *to infinity to a positive real value and*

$$\lim_{\kappa\to\infty} \mathrm{RDR}(\kappa) = \begin{cases} 0, & \gamma > 0, \\ C, & \gamma = 0, \\ \infty & \gamma < 0, \end{cases}$$

where $C = \frac{T}{2}\sum_{i=2}^{d} \frac{1}{\lambda_i}\langle h\tau^\top c + \lambda_i \tau^{-1}\nu, v_i\rangle^2 > 0$.

Proof. Firstly, note that from Theorem 4.8 we know that the worst-case parameter μ^* has the form

$$\mu^*(\kappa) = \nu + \tau \sum_{i=1}^{d} a_i(\kappa) v_i,$$

where

$$a_i(\kappa) = -\left(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\right)^{-1} \left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\right\rangle$$

for $i = 1, \dots, d$. Recall from Lemma 5.1 that

$$\lim_{\kappa\to\infty} \frac{\psi(\kappa)}{\kappa} = 1.$$

Since $v_1, \dots, v_d$ form an orthogonal basis of $\mathbb{R}^d$, it holds $h\tau^\top c + \frac{1}{1-\gamma}A\tau^{-1}\nu \in \operatorname{span}(\{v_1\})$ if and only if

$$\left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\right\rangle = \left\langle h\tau^\top c + \frac{1}{1-\gamma}A\tau^{-1}\nu, v_i\right\rangle = 0$$

for all $i = 2, \dots, d$, which implies that $a_2(\kappa) = \cdots = a_d(\kappa) = 0$. Therefore, we see that $h\tau^\top c + \frac{1}{1-\gamma}A\tau^{-1}\nu \in \operatorname{span}(\{v_1\})$ is equivalent to $\mu^* = \nu + a_1(\kappa)\tau v_1 = \nu - \kappa\tau v_1$. In that case, we immediately see

$$\pi_t^*(\kappa) = \frac{1}{1-\gamma}A\mu^*(\kappa) + hc = \frac{1}{1-\gamma}A(\nu - \kappa\tau v_1) + hc = \frac{1}{1-\gamma}A\nu + hc = \hat{\pi}_t$$

for all $t \in [0, T]$ and $\kappa > 0$. It follows that $\mathrm{COA}(\kappa) = \mathrm{RDR}(\kappa) = 0$ for all $\kappa > 0$.

Now suppose that $h\tau^\top c + \frac{1}{1-\gamma}A\tau^{-1}\nu \notin \operatorname{span}(\{v_1\})$. Then there exists an $i \in \{2, \dots, d\}$ such that $a_i(\kappa) \neq 0$ for all $\kappa > 0$. We can rewrite

$$\begin{aligned}(\mu^*(\kappa) - \nu)^\top A(\mu^*(\kappa) - \nu) &= \left(\sum_{i=1}^{d} a_i(\kappa)v_i\right)^\top \tau^\top A\tau \left(\sum_{i=1}^{d} a_i(\kappa)v_i\right) \\ &= \sum_{i,j=1}^{d} a_i(\kappa)a_j(\kappa)v_i^\top \tau^\top A\tau v_j = \sum_{i=2}^{d} a_i(\kappa)^2\lambda_i,\end{aligned}$$

using firstly that v_i is an eigenvector to eigenvalue λ_i of the matrix $\tau^\top A\tau$ and secondly that the v_i are orthonormal and $\lambda_1 = 0$. For any $i = 2, \dots, d$ it holds

$$\begin{aligned}\lim_{\kappa\to\infty} a_i(\kappa) &= -\lim_{\kappa\to\infty}\left(\frac{\lambda_i}{1-\gamma} + \frac{h}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}\right)^{-1} \left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\right\rangle \\ &= -\frac{1-\gamma}{\lambda_i}\left\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\right\rangle\end{aligned} \tag{5.14}$$

and our assumption ensures that there is at least one index in $\{2, \ldots, d\}$ such that the limit is not equal to zero. Consequently,

$$\lim_{\kappa \to \infty} (\mu^*(\kappa) - \nu)^\top A(\mu^*(\kappa) - \nu) = \sum_{i=2}^{d} \frac{(1-\gamma)^2}{\lambda_i} \Big\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i \Big\rangle^2 > 0. \tag{5.15}$$

For $\gamma = 0$ we have shown in Lemma 5.11 that

$$\mathrm{COA}(\kappa) = \mathrm{RDR}(\kappa) = \frac{T}{2}(\mu^*(\kappa) - \nu)^\top A(\mu^*(\kappa) - \nu),$$

so from (5.15) it follows that

$$\lim_{\kappa \to \infty} \mathrm{COA}(\kappa) = \lim_{\kappa \to \infty} \mathrm{RDR}(\kappa) = \frac{T}{2} \sum_{i=2}^{d} \frac{1}{\lambda_i} \langle h\tau^\top c + \lambda_i \tau^{-1}\nu, v_i \rangle^2 = C > 0.$$

If $\gamma \neq 0$, then

$$\mathrm{COA}(\kappa) = \frac{M(\gamma)}{\gamma} L(\gamma, \kappa) \exp\Big(\gamma T \Big(hc^\top \nu + \frac{1}{2(1-\gamma)}\nu^\top A\nu\Big)\Big).$$

Using (5.15) together with the representation

$$L(\gamma, \kappa) = 1 - \exp\Big(-\frac{\gamma T}{2(1-\gamma)}(\mu^* - \nu)^\top A(\mu^* - \nu)\Big),$$

as well as the fact that $M(\gamma) > 0$, it is easy to see that $\mathrm{COA}(\kappa)$ converges as κ goes to infinity and that the limit is a positive real value. For the reward for distributional robustness we have the representation

$$\mathrm{RDR}(\kappa) = \frac{M(\gamma)}{\gamma} L(\gamma, \kappa) \exp\Big(\gamma T \Big(hc^\top \mu^*(\kappa) + \frac{1}{2(1-\gamma)}(\mu^*(\kappa))^\top A\mu^*(\kappa)\Big)\Big).$$

Note that, again by (5.15), the term

$$\frac{M(\gamma)}{\gamma} L(\gamma, \kappa)$$

converges to a positive limit as κ goes to infinity. It remains to study the asymptotic behavior of the exponential term. For that purpose we rewrite

$$\begin{aligned}
&hc^\top \mu^*(\kappa) + \frac{1}{2(1-\gamma)}(\mu^*(\kappa))^\top A\mu^*(\kappa) \\
&= hc^\top \Big(\nu + \tau \sum_{i=1}^{d} a_i(\kappa) v_i\Big) + \frac{1}{2(1-\gamma)} \Big(\nu + \tau \sum_{i=1}^{d} a_i(\kappa) v_i\Big)^\top A \Big(\nu + \tau \sum_{i=1}^{d} a_i(\kappa) v_i\Big) \\
&= hc^\top \nu + hc^\top \tau \sum_{i=2}^{d} a_i(\kappa) v_i + hc^\top \tau a_1(\kappa) v_1 \\
&\quad + \frac{1}{2(1-\gamma)} \Big(\nu^\top A\nu + 2\nu^\top A\tau \sum_{i=2}^{d} a_i(\kappa) v_i + (\mu^*(\kappa) - \nu)^\top A(\mu^*(\kappa) - \nu)\Big).
\end{aligned}$$

Here, due to (5.14) we have

$$\lim_{\kappa\to\infty}\sum_{i=2}^{d} a_i(\kappa)v_i = -\sum_{i=2}^{d}\frac{1-\gamma}{\lambda_i}\Big\langle h\tau^\top c + \frac{\lambda_i}{1-\gamma}\tau^{-1}\nu, v_i\Big\rangle v_i$$

and the expression $(\mu^*(\kappa)-\nu)^\top A(\mu^*(\kappa)-\nu)$ converges due to (5.15). However,

$$hc^\top \tau a_1(\kappa)v_1 = -\psi(\kappa)\frac{h}{\|\tau^{-1}\mathbf{1}_d\|}$$

goes to minus infinity as κ goes to infinity. Putting these results together, we obtain that

$$hc^\top \mu^*(\kappa) + \frac{1}{2(1-\gamma)}(\mu^*(\kappa))^\top A\mu^*(\kappa)$$

goes to minus infinity and in conclusion

$$\lim_{\kappa\to\infty}\mathrm{RDR}(\kappa) = \begin{cases} 0, & \gamma > 0, \\ \infty, & \gamma < 0. \end{cases}$$

Putting the various cases together proves the claim of the proposition. □

The preceding proposition analyzes the asymptotic behavior of $\mathrm{COA}(\kappa)$ and $\mathrm{RDR}(\kappa)$ as κ goes to infinity. The first case is a special case corresponding to such a combination of model parameters that makes the optimal robust strategy $\pi^*(\kappa)$ equal to $\hat{\pi}$ for any κ. In this special case the increase of model uncertainty affects the robust and the non-robust investor in the same way.

Apart from this special situation the measure $\mathrm{COA}(\kappa)$ converges to a positive value as κ goes to infinity. In particular, the cost of ambiguity is bounded and for high levels of uncertainty, the influence of an increase in κ is neglectable. For the reward for distributional robustness, the situation is different. Here, the asymptotic behavior is determined by the parameter γ of the investor's risk aversion. For logarithmic utility, i.e. $\gamma = 0$, we know already from Lemma 5.11 that reward for distributional robustness equals cost of ambiguity. Hence, in this case $\mathrm{RDR}(\kappa)$ converges as well and the limit is a positive value. For negative γ, $\mathrm{RDR}(\kappa)$ goes to infinity whereas for positive γ, $\mathrm{RDR}(\kappa)$ converges to zero.

This qualitative difference is due to the essential difference between the utility functions U_γ for either positive or negative γ. For positive γ, the function U_γ is non-negative on $\mathbb{R}_+$, hence expected utility is bounded from below by zero. That implies that the worst-case expected utility corresponding to κ going to infinity will go to zero independently of the strategy that an investor uses. This explains the convergence of $\mathrm{RDR}(\kappa)$ to zero in the case $\gamma > 0$. In contrast, for negative γ, the function U_γ is unbounded from below, leading to arbitrarily high rewards for distributional robustness.

The above observations suggest to use a different measure for the reward for distributional robustness that is less affected by the specific form of the investor's utility function. To reach this goal one might not want to consider the difference in expected utility but the difference in the corresponding certainty equivalents, for example. As the certainty equivalent of the random terminal wealth X_T^π we denote the deterministic amount of money that leads to a utility value equal to the expected utility of X_T^π. Since our utility function U_γ is in any case invertible we can compute the certainty equivalent of a terminal wealth X_T^π as

$$U_\gamma^{-1}\big(\mathbb{E}\big[U_\gamma(X_T^\pi)\big]\big).$$

The inverse of the utility function is given by

$$U_\gamma^{-1}(y) = \begin{cases} \exp(y), & \gamma = 0, \\ (\gamma y)^{1/\gamma}, & \gamma \neq 0. \end{cases}$$

Note that the domain of U_γ^{-1} is $\mathbb{R}$ if $\gamma = 0$, $\mathbb{R}_+$ for $\gamma > 0$ and $\mathbb{R}_-$ for $\gamma < 0$. We now define a modification of COA and RDR as follows.

Definition 5.14. We define the *certainty-equivalent-based cost of ambiguity* as

$$\overline{\mathrm{COA}} = U_\gamma^{-1}\big(\mathbb{E}_\nu\big[U_\gamma(X_T^{\hat{\pi}})\big]\big) - U_\gamma^{-1}\big(\mathbb{E}_\nu\big[U_\gamma(X_T^{\pi^*})\big]\big)$$

and the *certainty-equivalent-based reward for distributional robustness* as

$$\overline{\mathrm{RDR}} = U_\gamma^{-1}\big(\mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big]\big) - U_\gamma^{-1}\big(\mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\hat{\pi}})\big]\big).$$

Note that since U_γ^{-1} is a strictly increasing function, the non-negativity $\overline{\mathrm{COA}} \geq 0$ and $\overline{\mathrm{RDR}} \geq 0$ persists. Further, it always holds $\overline{\mathrm{COA}} \geq \overline{\mathrm{RDR}}$ which we show in the following proposition.

Proposition 5.15. *Independently of $\gamma \in (-\infty, 1)$ it always holds $\overline{\mathrm{COA}} \geq \overline{\mathrm{RDR}}$.*

Proof. Recall that we can write

$$\mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big] = \begin{cases} \log(x_0) + rT + T\Big(\pi_0^\top(\mu - r\mathbf{1}_d) - \dfrac{1}{2}\|\sigma^\top \pi_0\|^2\Big), & \gamma = 0, \\ \dfrac{x_0^\gamma}{\gamma}\exp\Big(\gamma T\Big(r + \pi_0^\top(\mu - r\mathbf{1}_d) - \dfrac{1-\gamma}{2}\|\sigma^\top \pi_0\|^2\Big)\Big), & \gamma \neq 0, \end{cases}$$

if $(\pi_t)_{t\in[0,T]}$ is constant in time and deterministic. Hence for $\gamma \neq 0$ we have

$$\begin{aligned} \overline{\mathrm{COA}} = x_0 \mathrm{e}^{rT}\bigg(&\exp\Big(T\Big((\hat{\pi}_0)^\top(\nu - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top \hat{\pi}_0\|^2\Big)\Big) \\ &- \exp\Big(T\Big((\pi_0^*)^\top(\nu - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top \pi_0^*\|^2\Big)\Big)\bigg) \end{aligned} \tag{5.16}$$

and

$$\begin{aligned} \overline{\mathrm{RDR}} = x_0 \mathrm{e}^{rT}\bigg(&\exp\Big(T\Big((\pi_0^*)^\top(\mu^* - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top \pi_0^*\|^2\Big)\Big) \\ &- \exp\Big(T\Big((\hat{\pi}_0)^\top(\mu^* - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top \hat{\pi}_0\|^2\Big)\Big)\bigg). \end{aligned} \tag{5.17}$$

For $\gamma = 0$ we obtain

$$\overline{\mathrm{COA}} = x_0 \mathrm{e}^{rT}\Big(\exp\Big(T\big((\hat{\pi}_0)^\top(\nu - r\mathbf{1}_d) - \tfrac{1}{2}\|\sigma^\top \hat{\pi}_0\|^2\big)\Big) - \exp\Big(T\big((\pi_0^*)^\top(\nu - r\mathbf{1}_d) - \tfrac{1}{2}\|\sigma^\top \pi_0^*\|^2\big)\Big)\Big)$$

and

$$\overline{\mathrm{RDR}} = x_0 \mathrm{e}^{rT}\Big(\exp\Big(T\big((\pi_0^*)^\top(\mu^*-r\mathbf{1}_d)-\tfrac{1}{2}\|\sigma^\top\pi_0^*\|^2\big)\Big)-\exp\Big(T\big((\hat\pi_0)^\top(\mu^*-r\mathbf{1}_d)-\tfrac{1}{2}\|\sigma^\top\hat\pi_0\|^2\big)\Big)\Big)$$

which are the same representations as in (5.16) and (5.17) with $\gamma = 0$. We now plug in the representations from (5.9), respectively (5.10) of the strategies π^* and $\hat\pi$ and use the properties $A\mathbf{1}_d = 0$, $c^\top\sigma\sigma^\top A = 0$ and $A\sigma\sigma^\top A = A$, see Lemma 5.5. We obtain

$$\begin{aligned}\frac{\overline{\mathrm{COA}}}{x_0\mathrm{e}^{rT}} &= \exp\Big(T\Big(hc^\top(\nu - r\mathbf{1}_d) + \frac{1}{1-\gamma}\nu^\top A\nu - \frac{1-\gamma}{2}h^2c^\top\sigma\sigma^\top c - \frac{1}{2(1-\gamma)}\nu^\top A\nu\Big)\Big)\\ &\quad - \exp\Big(T\Big(hc^\top(\nu - r\mathbf{1}_d) + \frac{1}{1-\gamma}(\mu^*)^\top A\nu - \frac{1-\gamma}{2}h^2c^\top\sigma\sigma^\top c - \frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^*\Big)\Big)\\ &= \overline{L}(\gamma,\kappa)\exp\Big(T\Big(-hr - \frac{1-\gamma}{2}h^2c^\top\sigma\sigma^\top c + hc^\top\nu + \frac{1}{2(1-\gamma)}\nu^\top A\nu\Big)\Big),\end{aligned}$$

where

$$\overline{L}(\gamma,\kappa) = 1 - \exp\Big(-\frac{T}{2(1-\gamma)}(\mu^* - \nu)^\top A(\mu^* - \nu)\Big).$$

Analogously we get

$$\frac{\overline{\mathrm{RDR}}}{x_0\mathrm{e}^{rT}} = \overline{L}(\gamma,\kappa)\exp\Big(T\Big(-hr - \frac{1-\gamma}{2}h^2c^\top\sigma\sigma^\top c + hc^\top\mu^* + \frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^*\Big)\Big).$$

Hence, we can deduce in particular that

$$\frac{\overline{\mathrm{COA}}}{\overline{\mathrm{RDR}}} = \frac{\exp\Big(T\Big(\frac{1}{2(1-\gamma)}\nu^\top A\nu + hc^\top\nu\Big)\Big)}{\exp\Big(T\Big(\frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^* + hc^\top\mu^*\Big)\Big)} \geq 1,$$

since μ^* minimizes the function

$$\mu \mapsto \frac{1}{2(1-\gamma)}\mu^\top A\mu + hc^\top\mu$$

on the set K. $\square$

Additionally, the certainty-equivalent-based definitions of cost of ambiguity and reward for distributional robustness show an asymptotic behavior that is independent of the investor's risk aversion. We write $\overline{\mathrm{COA}}(\kappa)$ and $\overline{\mathrm{RDR}}(\kappa)$ to emphasize the underlying dependence on the degree of uncertainty.

Proposition 5.16. *As κ goes to infinity, $\overline{\mathrm{COA}}(\kappa)$ converges to a non-negative limit and $\overline{\mathrm{RDR}}(\kappa)$ goes to zero.*

Proof. Firstly, note that by the same reasoning as in the proof of Proposition 4.10 we have

$$(\hat\pi_0)^\top\mu^* \leq (\pi_0^*)^\top\mu^* = (\pi_0^*)^\top\nu - \kappa\sqrt{(\pi_0^*)^\top\Gamma\pi_0^*},$$

and that the right-hand side goes to $-\infty$ as κ goes to infinity. It follows that

$$\lim_{\kappa\to\infty} \mathbb{E}_{\mu^*}\left[U_\gamma(X_T^{\hat{\pi}})\right] = \lim_{\kappa\to\infty} \mathbb{E}_{\mu^*}\left[U_\gamma(X_T^{\pi^*})\right] = \begin{cases} -\infty, & \gamma \leq 0, \\ 0, & \gamma > 0, \end{cases}$$

and therefore $\lim_{\kappa\to\infty} \overline{\mathrm{RDR}}(\kappa) = 0$.

For $\overline{\mathrm{COA}}$ we observe that $\mathbb{E}_\nu[U_\gamma(X_T^{\pi^*})]$ converges to a finite value as κ goes to infinity, with that limit being different from zero if $\gamma \neq 0$. It follows that $U_\gamma^{-1}(\mathbb{E}_\nu[U_\gamma(X_T^{\pi^*})])$ also converges. We thus deduce convergence of $\overline{\mathrm{COA}}(\kappa)$. Since $\overline{\mathrm{COA}}(\kappa) \geq 0$ for any κ, we know that the limit is non-negative. □

In Figure 5.4 we illustrate the behavior of $\overline{\mathrm{COA}}$ and $\overline{\mathrm{RDR}}$ in dependence on the level of uncertainty κ and on the risk aversion coefficient γ. We consider a market with $d = 8$ stocks, where the underlying market parameters are again those from Example 5.8. The figure shows $\overline{\mathrm{COA}}$ and $\overline{\mathrm{RDR}}$ plotted against κ for different values of γ. Note that the scaling in the second row of subfigures is different from the scaling in the first row. The absolute values of $\overline{\mathrm{COA}}$ and $\overline{\mathrm{RDR}}$ become smaller as γ increases.

We observe that the qualitative behavior of $\overline{\mathrm{COA}}$ and $\overline{\mathrm{RDR}}$ is the same for any value of the risk aversion coefficient γ. For any fixed γ and κ, the value of the certainty-equivalent-based reward for distributional robustness is always less than the value of the certainty-equivalent-based cost of ambiguity, a property that we have proven in Proposition 5.15. As κ increases, $\overline{\mathrm{COA}}$ goes to a finite positive limit, whereas $\overline{\mathrm{RDR}}$ tends to zero. This is due to Proposition 5.16.

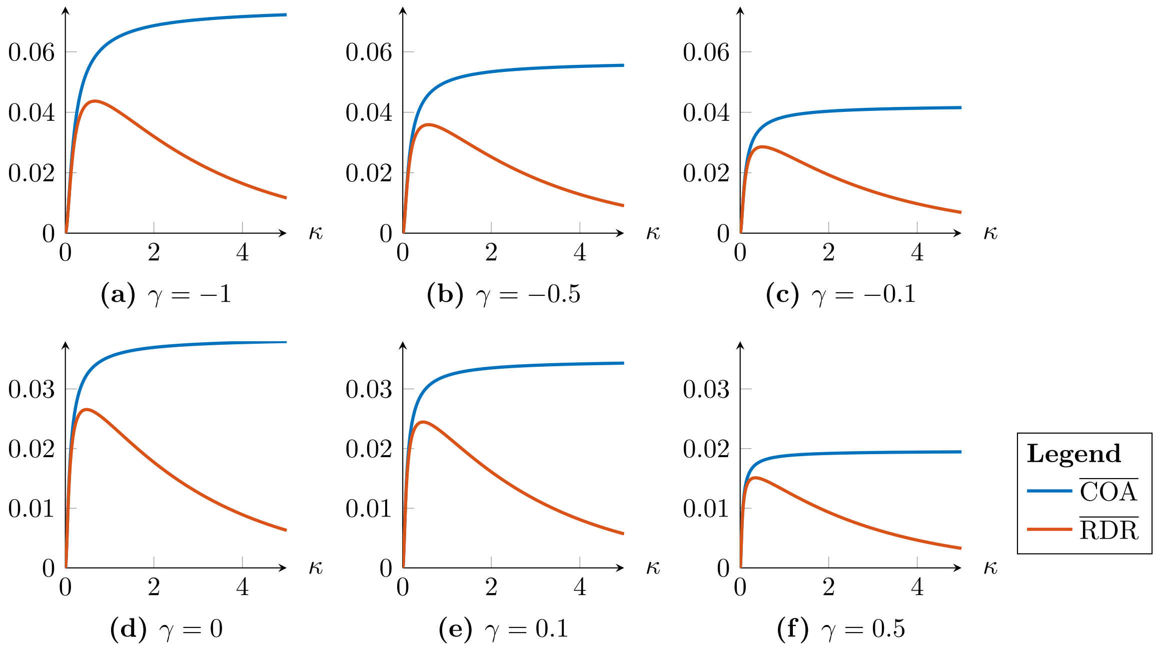


Figure 5.4.: The behavior of $\overline{\mathrm{COA}}$ and $\overline{\mathrm{RDR}}$ plotted against uncertainty radius κ for different values of the risk aversion coefficient γ. The parameters are those from Example 5.8.

Part II.

Expert Opinions

Outline and Notation

In this part of the thesis we consider a financial market in which returns are driven by an unobservable Gaussian drift process. We assume that investors observe the return process and additionally expert opinions that arrive at discrete time points and that are modelled as unbiased estimates of the drift at that time point. The aim is to analyze the influence of expert opinions on the filter for different investor filtrations. In particular, we focus on the asymptotic behavior of the filter as the frequency of expert opinions goes to infinity.

In detail, we proceed as follows. In Chapter 6 we introduce the model for our financial market including expert opinions and define different information regimes for investors with different sources of information. For each of those information regimes, we state the dynamics of the corresponding conditional mean and conditional covariance matrix of the filter.

Chapter 7 analyzes the asymptotic behavior of the conditional covariance matrices on an infinite time horizon with regularly arriving expert opinions. We show convergence for the situation where an investor observes the return process only. For investors who also have access to the expert opinions we provide convergence results under assumptions on the model parameters.

Our main results are given in Chapter 8 and focus on the asymptotic behavior of the filter as the arrival frequency of expert opinions goes to infinity on a finite time horizon. Section 8.1 addresses the situation where experts' covariance matrices are bounded, corresponding to some minimal level of reliability. In that case, we show that in the limit investors have full information about the drift process. In Section 8.2 we study a setting where the experts' covariance matrices are not bounded but grow linearly in the frequency of expert opinions. For an investor observing returns and discrete-time expert opinions we show convergence of the corresponding conditional mean and conditional covariance matrix to those of an investor observing the returns and a continuous-time expert. We consider two different situations, one with deterministic equidistant information dates and one with information dates that arrive randomly as the jump times of a Poisson process, i.e. with exponentially distributed waiting times between information dates.

Chapter 9 provides an application of the convergence results to a utility maximization problem. For investors who maximize expected logarithmic utility of terminal wealth the optimal trading strategy depends on the conditional mean of the drift, and the corresponding optimal terminal wealth is a functional of the conditional covariance matrices. That is why the convergence results from Chapter 8 carry over to convergence of the corresponding value functions. We also provide simulations and numerical calculations to illustrate our theoretical results.

This part of the thesis builds up on Gabih et al. [25] and on the Master's thesis Westphal [64] in which results from Gabih et al. [25] are carried over from a one-dimensional to a multivariate financial market. Parts of the Master's thesis [64] are repeated in this thesis for completeness. The financial market model given in Chapter 6 and the filtering equations are already given in the Master's thesis, with the exception of the D-investor. Large parts of Chapter 7 can be found in the Master's thesis already and are repeated since they are needed for later conclusions. The asymptotic results from Theorem 7.8 and Proposition 7.9 are a new contribution, however. Whereas Theorem 8.1 was already proven in Westphal [64], the whole Section 8.2 containing our main results is new. Parts of this work and the Master's thesis are published in Sass et al. [55], the follow-up paper [56] contains the new part on diffusion approximations.

Notation. Throughout this part, we use the notation I_d for the identity matrix in $\mathbb{R}^{d\times d}$ and $\mathbf{0}_d$ for the matrix in $\mathbb{R}^{d\times d}$ containing only zeros. For a symmetric and positive-semidefinite matrix $A \in \mathbb{R}^{d\times d}$ we call a symmetric and positive-semidefinite matrix $B \in \mathbb{R}^{d\times d}$ the *square root* of A if $B^2 = A$. The square root is unique and will be denoted by $A^{1/2}$.

For a matrix A we denote with $\|A\|$ the spectral norm of A. If A is a symmetric matrix, we write $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ for the minimal and maximal eigenvalue of A, respectively. Furthermore, since $\|A\|$ is the square root of the maximal eigenvalue of $A^\top A$, note that for symmetric positive-semidefinite matrices A it holds $\|A\| = \lambda_{\max}(A)$.

We also use a partial ordering of symmetric matrices. For symmetric matrices $A, B \in \mathbb{R}^{d\times d}$ we write $A \preceq B$ if $B - A$ is positive semidefinite. Note that $A \preceq B$ in particular implies $\|A\| \leq \|B\|$.

If $\mathcal{F}$ and $\mathcal{G}$ are σ-algebras, we write $\mathcal{F} \vee \mathcal{G}$ for the smallest σ-algebra containing $\mathcal{F} \cup \mathcal{G}$.

6. Market Model and Filtering

6.1. Financial market model

We consider a financial market with one risk-free and various risky assets. The basic model is the same as in the Master's thesis [64], respectively in the papers by Sass et al. [55, 56]. In the following, we use the notation $\mathbb{T}$ for the time interval and assume that either $\mathbb{T} = [0, \infty)$ or $\mathbb{T} = [0, T]$ for some finite $T > 0$. We fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ satisfies the usual conditions. All processes are assumed to be $\mathbb{F}$-adapted. The market consists of one risk-free bond with constant deterministic interest rate $r \in \mathbb{R}$ and d risky assets such that the d-dimensional return process follows the stochastic differential equation

$$\mathrm{d}R_t = \mu_t \,\mathrm{d}t + \sigma_R \,\mathrm{d}W_t^R.$$

Here, $W^R = (W_t^R)_{t\in\mathbb{T}}$ is an m-dimensional Brownian motion with $m \geq d$ and we assume that $\sigma_R \in \mathbb{R}^{d\times m}$ has full rank. The drift μ is an Ornstein–Uhlenbeck process and follows the dynamics

$$\mathrm{d}\mu_t = \alpha(\delta - \mu_t)\,\mathrm{d}t + \beta\,\mathrm{d}B_t,$$

where α and $\beta \in \mathbb{R}^{d\times d}$, $\delta \in \mathbb{R}^d$, and where $B = (B_t)_{t\in\mathbb{T}}$ is a d-dimensional Brownian motion independent of W^R. We assume that α and $\beta\beta^\top$ are symmetric and positive-definite matrices. The initial drift μ_0 is multivariate normally distributed, $\mu_0 \sim \mathcal{N}(m_0, \Sigma_0)$, for some $m_0 \in \mathbb{R}^d$ and some $\Sigma_0 \in \mathbb{R}^{d\times d}$ which is symmetric and positive semidefinite. We assume that μ_0 is independent of B and W^R, so also μ is independent of W^R. The drift process μ can then be written in explicit form as

$$\mu_t = \delta + \mathrm{e}^{-\alpha t}\left(\mu_0 - \delta + \int_0^t \mathrm{e}^{\alpha s}\beta\,\mathrm{d}B_s\right).$$

We further denote $m_t := \mathbb{E}[\mu_t]$ and $\Sigma_t := \mathrm{cov}(\mu_t)$. It holds

$$\begin{aligned} m_t &= \delta + \mathrm{e}^{-\alpha t}(m_0 - \delta), \\ \Sigma_t &= \mathrm{e}^{-\alpha t}\left(\Sigma_0 + \int_0^t \mathrm{e}^{\alpha s}\beta\beta^\top \mathrm{e}^{\alpha s}\,\mathrm{d}s\right)\mathrm{e}^{-\alpha t}. \end{aligned}$$

Investors in this market are able to observe the return process R. They neither observe the underlying drift process μ nor the Brownian motion W^R. However, information about μ can be drawn from observing R. Additionally, we include expert opinions in our model. These expert opinions arrive at discrete time points and give an unbiased estimate of the state of the drift at that time point. Let $(T_k)_{k\in I}$ be an increasing sequence with values in $\mathbb{T} \setminus \{0\}$, where we allow for index sets $I = \mathbb{N}$ or $I = \{1, \dots, N\}$ for some $N \in \mathbb{N}$. The T_k, $k \in I$, are the time points at which expert opinions arrive. For the sake of convenience we also write $T_0 = 0$ although there is no expert opinion arriving at time zero.

The expert view at time T_k is modelled as an $\mathbb{R}^d$-valued random vector

$$Z_k = \mu_{T_k} + (\Gamma_k)^{1/2}\varepsilon_k,$$

where the matrix $\Gamma_k \in \mathbb{R}^{d\times d}$ is symmetric and positive definite and ε_k is multivariate $\mathcal{N}(0, I_d)$-distributed. We assume that the sequence $(\varepsilon_k)_{k\in I}$ is independent and also that it is independent of both μ_0 and the Brownian motions B and W^R. Given μ_{T_k}, the expert opinion Z_k is multivariate $\mathcal{N}(\mu_{T_k}, \Gamma_k)$-distributed. That means that the expert view at time T_k gives an unbiased estimate of the state of the drift at that time. The matrix Γ_k reflects the reliability of the expert.

Note that the time points T_k do not need to be deterministic. However, we impose the additional assumption that the sequence $(T_k)_{k\in I}$ is independent of the $(\varepsilon_k)_{k\in I}$ and also of the Brownian motions in the market and of μ_0. This essentially says that the timing of information dates carries no additional information about the drift μ. Nevertheless, information on the sequence $(T_k)_{k\in I}$ may be important for optimal portfolio decisions. In Chapter 8 we consider on the one hand the situation with deterministic information dates and on the other hand a case where information dates are the jump times of a Poisson process.

It is possible to allow relative expert views in the sense that an expert may give an estimate for the difference in drift of two stocks instead of absolute views. See Schöttle et al. [61] for how to switch between these two models for expert opinions by means of a pick matrix.

Our main results in Chapter 8 address the question how to obtain rigorous convergence results when the number of information dates increases. If experts have a minimal level of reliability, one obtains full information in the limit. For other sequences of expert opinions, where the expert's variance grows linearly in the number of expert opinions, the information drawn from these expert opinions is asymptotically the same as the information one gets from observing yet another diffusion process. This diffusion process can then be interpreted as an expert who gives a continuous-time estimation about the state of the drift. Let this estimate be given by the diffusion process

$$\mathrm{d}J_t = \mu_t\,\mathrm{d}t + \sigma_J\,\mathrm{d}W_t^J, \tag{6.1}$$

where W^J is an l-dimensional Brownian motion with $l \geq d$ that is independent of all other Brownian motions in the model, of μ_0 and of the information dates T_k. The matrix $\sigma_J \in \mathbb{R}^{d\times l}$ has full rank equal to d.

6.2. Filtering for different information regimes

For an investor in the financial market defined above, the ability to choose good trading strategies is based heavily on which information is available about the unknown drift process μ. To be able to assess the value of information coming from observing expert opinions, we consider various types of investors with different sources of information. This follows the approach in Gabih et al. [25] and in Sass et al. [55]. The information available to an investor can be described by the investor filtration $\mathbb{F}^H = (\mathcal{F}_t^H)_{t\in\mathbb{T}}$ where H serves as a placeholder for the various information regimes. We work with filtrations that are augmented by $\mathcal{N}_\mathbb{P}$, the set of null sets under measure $\mathbb{P}$. We consider the cases

$$\begin{aligned}
\mathbb{F}^R &= (\mathcal{F}_t^R)_{t\in\mathbb{T}} \text{ where } \mathcal{F}_t^R = \sigma((R_s)_{s\in[0,t]}) \vee \sigma(\mathcal{N}_{\mathbb{P}}),\\
\mathbb{F}^E &= (\mathcal{F}_t^E)_{t\in\mathbb{T}} \text{ where } \mathcal{F}_t^E = \sigma((T_k, Z_k)_{T_k\le t}) \vee \sigma(\mathcal{N}_{\mathbb{P}}),\\
\mathbb{F}^C &= (\mathcal{F}_t^C)_{t\in\mathbb{T}} \text{ where } \mathcal{F}_t^C = \sigma((R_s)_{s\in[0,t]}) \vee \sigma((T_k, Z_k)_{T_k\le t}) \vee \sigma(\mathcal{N}_{\mathbb{P}}),\\
\mathbb{F}^D &= (\mathcal{F}_t^D)_{t\in\mathbb{T}} \text{ where } \mathcal{F}_t^D = \sigma((R_s)_{s\in[0,t]}) \vee \sigma((J_s)_{s\in[0,t]}) \vee \sigma(\mathcal{N}_{\mathbb{P}}),\\
\mathbb{F}^F &= (\mathcal{F}_t^F)_{t\in\mathbb{T}} \text{ where } \mathcal{F}_t^F = \sigma((R_s)_{s\in[0,t]}) \vee \sigma((\mu_s)_{s\in[0,t]}) \vee \sigma(\mathcal{N}_{\mathbb{P}}).
\end{aligned}$$

When speaking of the H-investor we mean the investor with investor filtration $\mathbb{F}^H = (\mathcal{F}_t^H)_{t\in\mathbb{T}}$, $H \in \{R, E, C, D, F\}$. Note that the R-investor observes only the return process, the E-investor only the incoming expert opinions at information dates. The C-investor can combine the information from observing both the return process and the expert opinions and the D-investor combines return observations with continuous-time expert opinions, i.e. the D-investor observes the two diffusion processes R and J. The F-investor has full information about the drift in the sense that she can observe the drift process directly. This case is included as a benchmark.

As already mentioned, the investors in our financial market make trading decisions based on available information about the drift process μ. Only the F-investor can observe the drift, the other investors have to estimate it. The conditional distribution of the drift under partial information is called the *filter*. In the mean-squared sense, an optimal estimator for the drift at time t given the available information is then the *conditional mean* $m_t^H := \mathbb{E}[\mu_t \,|\, \mathcal{F}_t^H]$. How close this estimator is to the true state of the drift can be assessed by looking at the corresponding *conditional covariance matrix*

$$Q_t^H := \mathbb{E}\big[(\mu_t - m_t^H)(\mu_t - m_t^H)^\top \,\big|\, \mathcal{F}_t^H\big].$$

Note that since we deal with Gaussian distributions here, the conditional mean and conditional covariance matrix completely characterize the filter since the filter is also Gaussian. We state in the following the dynamics of the filters for the various investors defined above. With the exception of the D-investor and with a restriction to deterministic information dates T_k, the following results have been derived in the Master's thesis Westphal [64] already. For the R-investor, we are in the setting of the well-known Kalman filter.

Lemma 6.1. *The filter of the R-investor is Gaussian. The conditional mean m^R follows the dynamics*

$$\mathrm{d}m_t^R = \alpha(\delta - m_t^R)\,\mathrm{d}t + Q_t^R(\sigma_R\sigma_R^\top)^{-1}(\mathrm{d}R_t - m_t^R\,\mathrm{d}t),$$

where Q^R is the solution of the ordinary Riccati differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}Q_t^R = -\alpha Q_t^R - Q_t^R\alpha + \beta\beta^\top - Q_t^R(\sigma_R\sigma_R^\top)^{-1}Q_t^R.$$

The initial values are $m_0^R = m_0$ and $Q_0^R = \Sigma_0$.

This lemma follows directly from the Kalman filter theory, see for example Liptser and Shiryaev [42, Thm. 10.3]. Note that Q_t^R follows an ordinary differential equation, called Riccati equation, and is hence deterministic.

Next, we consider the D-investor. Recall that this investor observes both the diffusion processes R and J. These observations can be written in a combined $2d$-dimensional process

$$\mathrm{d}D_t = \begin{pmatrix} \mathrm{d}R_t \\ \mathrm{d}J_t \end{pmatrix} = \begin{pmatrix} I_d \\ I_d \end{pmatrix} \mu_t \,\mathrm{d}t + \begin{pmatrix} \sigma_R & 0 \\ 0 & \sigma_J \end{pmatrix} \mathrm{d}W_t^D,$$

where

$$W^D = \begin{pmatrix} W^R \\ W^J \end{pmatrix}$$

is an $(m+l)$-dimensional Brownian motion. Now we can easily deduce the dynamics of m^D and Q^D.

Lemma 6.2. *The filter of the D-investor is Gaussian. The conditional mean m^D follows the dynamics*

$$\mathrm{d}m_t^D = \alpha(\delta - m_t^D)\,\mathrm{d}t + Q_t^D \begin{pmatrix} (\sigma_R\sigma_R^\top)^{-1} \\ (\sigma_J\sigma_J^\top)^{-1} \end{pmatrix}^\top \left(\mathrm{d}D_t - \begin{pmatrix} m_t^D \\ m_t^D \end{pmatrix} \mathrm{d}t \right),$$

where Q^D is the solution of the ordinary Riccati differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_t^D = -\alpha Q_t^D - Q_t^D \alpha + \beta\beta^\top - Q_t^D (\sigma_D\sigma_D^\top)^{-1} Q_t^D \tag{6.2}$$

with $(\sigma_D\sigma_D^\top)^{-1} = (\sigma_R\sigma_R^\top)^{-1} + (\sigma_J\sigma_J^\top)^{-1}$. The initial values are $m_0^D = m_0$ and $Q_0^D = \Sigma_0$.

Proof. Firstly, note that the matrix $(\sigma_R\sigma_R^\top)^{-1} + (\sigma_J\sigma_J^\top)^{-1} \in \mathbb{R}^{d\times d}$ is symmetric and positive definite, and hence nonsingular. Let $\sigma_D \in \mathbb{R}^{d\times d}$ denote the unique symmetric and positive-definite square root of the inverse. Then it holds $(\sigma_D\sigma_D^\top)^{-1} = (\sigma_R\sigma_R^\top)^{-1} + (\sigma_J\sigma_J^\top)^{-1}$. The distribution of the filter as well as the dynamics of m^D and Q^D then follow immediately from the Kalman filter theory, see again Liptser and Shiryaev [42, Thm. 10.3]. $\square$

Note that, just like in the case for the R-investor, the conditional covariance matrix is deterministic. Next, we address the E-investor who observes the expert opinions Z_k at the (possibly random) information dates T_k. The dynamics of conditional mean and conditional covariance matrix are given in the following lemma.

Lemma 6.3. *Given a sequence of information dates T_k, the filter of the E-investor is Gaussian. The dynamics of the conditional mean and conditional covariance matrix are given as follows:*

(i) Between the information dates T_{k-1} and T_k, $k \in I$, it holds

$$\begin{aligned} m_t^E &= \mathrm{e}^{-\alpha(t-T_{k-1})} m_{T_{k-1}}^E + \big(I_d - \mathrm{e}^{-\alpha(t-T_{k-1})}\big)\delta, \\ Q_t^E &= \mathrm{e}^{-\alpha(t-T_{k-1})} \left(Q_{T_{k-1}}^E + \int_{T_{k-1}}^t \mathrm{e}^{\alpha(s-T_{k-1})} \beta\beta^\top \mathrm{e}^{\alpha(s-T_{k-1})}\,\mathrm{d}s \right) \mathrm{e}^{-\alpha(t-T_{k-1})} \end{aligned}$$

for $t \in [T_{k-1}, T_k)$. It holds $m_0^E = m_0$ and $Q_0^E = \Sigma_0$.

(ii) The update formulas at information dates T_k, $k \in I$, are

$$\begin{aligned} m^E_{T_k} &= \rho_k(Q^E_{T_k-})m^E_{T_k-} + \big(I_d - \rho_k(Q^E_{T_k-})\big)Z_k \\ &= m^E_{T_k-} + \big(I_d - \rho_k(Q^E_{T_k-})\big)\big(Z_k - m^E_{T_k-}\big) \end{aligned}$$

and

$$\begin{aligned} Q^E_{T_k} &= \rho_k(Q^E_{T_k-})Q^E_{T_k-} \\ &= Q^E_{T_k-} + \big(\rho_k(Q^E_{T_k-}) - I_d\big)Q^E_{T_k-}, \end{aligned}$$

where $\rho_k(Q) = \Gamma_k(Q + \Gamma_k)^{-1}$.

Proof. We give a proof for deterministic time points T_k first.

(i) Note that we can write the drift μ_t at time t as

$$\mu_t = \delta + \mathrm{e}^{-\alpha(t-T_{k-1})}\bigg(\mu_{T_{k-1}} - \delta + \int_{T_{k-1}}^t \mathrm{e}^{\alpha(s-T_{k-1})}\beta\,\mathrm{d}B_s\bigg).$$

Also, there is no incoming information between T_{k-1} and t, so $\mathcal{F}^E_t = \mathcal{F}^E_{T_{k-1}}$. Hence,

$$\begin{aligned} m^E_t &= \mathbb{E}\big[\mu_t|\mathcal{F}^E_t\big] = \mathbb{E}\big[\mu_t|\mathcal{F}^E_{T_{k-1}}\big] \\ &= \delta + \mathrm{e}^{-\alpha(t-T_{k-1})}\bigg(m^E_{T_{k-1}} - \delta + \mathbb{E}\bigg[\int_{T_{k-1}}^t \mathrm{e}^{\alpha(s-T_{k-1})}\beta\,\mathrm{d}B_s\bigg]\bigg) \\ &= \mathrm{e}^{-\alpha(t-T_{k-1})}m^E_{T_{k-1}} + \big(I_d - \mathrm{e}^{-\alpha(t-T_{k-1})}\big)\delta, \end{aligned} \tag{6.3}$$

where we have used that the stochastic integral is independent of $\mathcal{F}^E_{T_{k-1}}$ and that it has expectation zero. For the conditional covariance matrix we get

$$\begin{aligned} Q^E_t &= \mathbb{E}\big[(\mu_t - m^E_t)(\mu_t - m^E_t)^\top \,\big|\, \mathcal{F}^E_t\big] \\ &= \mathbb{E}\bigg[\bigg(\delta + \mathrm{e}^{-\alpha(t-T_{k-1})}\bigg(\mu_{T_{k-1}} - \delta + \int_{T_{k-1}}^t \mathrm{e}^{\alpha(s-T_{k-1})}\beta\,\mathrm{d}B_s\bigg) - m^E_t\bigg) \\ &\qquad \cdot \bigg(\delta + \mathrm{e}^{-\alpha(t-T_{k-1})}\bigg(\mu_{T_{k-1}} - \delta + \int_{T_{k-1}}^t \mathrm{e}^{\alpha(s-T_{k-1})}\beta\,\mathrm{d}B_s\bigg) - m^E_t\bigg)^\top \,\bigg|\, \mathcal{F}^E_t\bigg]. \end{aligned}$$

When inserting (6.3) into the above representation, the terms involving δ cancel. The remaining conditional expectation can then be written as

$$\begin{aligned} &\mathrm{e}^{-\alpha(t-T_{k-1})}\,\mathbb{E}\bigg[\bigg(\mu_{T_{k-1}} - m^E_{T_{k-1}} + \int_{T_{k-1}}^t \mathrm{e}^{\alpha(s-T_{k-1})}\beta\,\mathrm{d}B_s\bigg) \\ &\qquad\qquad \cdot \bigg(\mu_{T_{k-1}} - m^E_{T_{k-1}} + \int_{T_{k-1}}^t \mathrm{e}^{\alpha(s-T_{k-1})}\beta\,\mathrm{d}B_s\bigg)^\top \,\bigg|\, \mathcal{F}^E_t\bigg]\mathrm{e}^{-\alpha(t-T_{k-1})}. \end{aligned}$$

An expansion of the product inside the conditional expectation gives

$$\begin{aligned} &Q^E_{T_{k-1}} + \mathbb{E}\bigg[\Big(\int_{T_{k-1}}^t \mathrm{e}^{\alpha(s-T_{k-1})}\beta\,\mathrm{d}B_s\Big)\Big(\int_{T_{k-1}}^t \mathrm{e}^{\alpha(s-T_{k-1})}\beta\,\mathrm{d}B_s\Big)^\top\bigg] \\ &= Q^E_{T_{k-1}} + \int_{T_{k-1}}^t \mathrm{e}^{\alpha(s-T_{k-1})}\beta\beta^\top \mathrm{e}^{\alpha(s-T_{k-1})}\,\mathrm{d}s, \end{aligned}$$

where the mixed terms cancel because of independence.

(ii) For the update formulas at information dates we interpret the situation as a degenerate discrete-time Kalman filter with time points T_k- and T_k. From Elliott et al. [19, Eq. (5.12) and (5.13)] we get for the conditional expectation

$$\begin{aligned} m^E_{T_k} &= m^E_{T_k-} + Q^E_{T_k-}(Q^E_{T_k-} + \Gamma_k)^{-1}(Z_k - m^E_{T_k-}) \\ &= \big(I_d - Q^E_{T_k-}(Q^E_{T_k-} + \Gamma_k)^{-1}\big) m^E_{T_k-} + Q^E_{T_k-}(Q^E_{T_k-} + \Gamma_k)^{-1} Z_k \\ &= \rho_k(Q^E_{T_k-}) m^E_{T_k-} + \big(I_d - \rho_k(Q^E_{T_k-})\big) Z_k, \end{aligned}$$

where $\rho_k(Q) = \Gamma_k(Q + \Gamma_k)^{-1}$, and for the conditional covariance matrix

$$\begin{aligned} Q^E_{T_k} &= \mathbb{E}\big[(\mu_{T_k} - m^E_{T_k})(\mu_{T_k} - m^E_{T_k})^\top \,\big|\, \mathcal{F}^E_{T_k}\big] \\ &= Q^E_{T_k-} - Q^E_{T_k-}(Q^E_{T_k-} + \Gamma_k)^{-1} Q^E_{T_k-} \\ &= \big(I_d - Q^E_{T_k-}(Q^E_{T_k-} + \Gamma_k)^{-1}\big) Q^E_{T_k-} \\ &= \rho_k(Q^E_{T_k-}) Q^E_{T_k-}. \end{aligned}$$

These are the update formulas for the filter at information dates. Alternatively, we can also compute the estimator $m^E_{T_k}$ and its conditional covariance matrix as a Bayesian update of $m^E_{T_k-}$ given the $\mathcal{N}(\mu_{T_k}, \Gamma_k)$-distributed expert opinion Z_k, see for example Shiryaev [62, Thm. II.8.2].

For the more general case where the T_k do not need to be deterministic, recall that we have made the assumption that the sequence $(T_k)_{k\in I}$ is independent of the other random variables in the market. In particular, $(T_k)_{k\in I}$ and the drift process μ are independent. Because of that, the dynamics of the conditional mean and conditional covariance matrix are the same as for deterministic information dates and we get the same update formulas, the only difference being that the update times might now be non-deterministic.

The Gaussian distribution of the filter between information dates follows as in the previous lemmas from the Kalman filter theory. Since the updates at information dates can be seen as a degenerate discrete-time Kalman filter the distribution of the filter at information dates remains Gaussian after the Bayesian update. □

From the second part of the lemma we see that the conditional mean $m^E_{T_k}$ at information date T_k is a weighted mean of the conditional mean $m^E_{T_k-}$ before the update and the expert opinion Z_k. The corresponding weights depend on the reliability of the expert at time T_k, more precisely on the covariance matrix Γ_k of the expert's view. The more reliable the expert is, i.e. the smaller Γ_k, the more weight is put on the expert's view. For the conditional covariance matrices we have a similar update formula.

If we have non-deterministic information dates T_k, then in contrast to both the R-investor and the D-investor, the conditional covariance matrices Q^E_t of the E-investor for fixed t are random matrices since updates take place at random times. However, the timing of the updates is the only randomness in the conditional covariance matrices. Given a sequence of information dates $(T_k)_{k\in I}$, the dynamics of Q^E are deterministic.

Let us now come to the C-investor. Recall that this investor observes the return process R continuously in time and at (possibly random) information dates T_k the expert opinions Z_k. We state the dynamics of m^C and Q^C in the following lemma.

Lemma 6.4. *Given a sequence of information dates T_k, the filter of the C-investor is Gaussian. The dynamics of the conditional mean and conditional covariance matrix are given as follows:*

(i) Between the information dates T_{k-1} and T_k, $k \in I$, it holds

$$\mathrm{d}m_t^C = \alpha(\delta - m_t^C)\,\mathrm{d}t + Q_t^C(\sigma_R\sigma_R^\top)^{-1}(\mathrm{d}R_t - m_t^C\,\mathrm{d}t)$$

for $t \in [T_{k-1}, T_k)$, where Q^C follows the ordinary Riccati differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}Q_t^C = -\alpha Q_t^C - Q_t^C\alpha + \beta\beta^\top - Q_t^C(\sigma_R\sigma_R^\top)^{-1}Q_t^C$$

for $t \in [T_{k-1}, T_k)$. The initial values are $m_{T_{k-1}}^C$ and $Q_{T_{k-1}}^C$, respectively, with $m_0^C = m_0$ and $Q_0^C = \Sigma_0$.

(ii) The update formulas at information dates T_k, $k \in I$, are

$$\begin{aligned} m_{T_k}^C &= \rho_k(Q_{T_k-}^C)m_{T_k-}^C + \big(I_d - \rho_k(Q_{T_k-}^C)\big)Z_k \\ &= m_{T_k-}^C + \big(I_d - \rho_k(Q_{T_k-}^C)\big)\big(Z_k - m_{T_k-}^C\big) \end{aligned}$$

and

$$\begin{aligned} Q_{T_k}^C &= \rho_k(Q_{T_k-}^C)Q_{T_k-}^C \\ &= Q_{T_k-}^C + \big(\rho_k(Q_{T_k-}^C) - I_d\big)Q_{T_k-}^C, \end{aligned}$$

where $\rho_k(Q) = \Gamma_k(Q + \Gamma_k)^{-1}$.

Proof. Note that between two subsequent information dates, no additional expert opinions arrive. Therefore, only return observations contribute to the filtration which implies that $\mathcal{F}_t^C = \mathcal{F}_{T_{k-1}}^C \vee \sigma(R_s \,|\, T_{k-1} < s \le t)$. Hence, in $[T_{k-1}, T_k)$ we are in the standard situation of the Kalman filter. The dynamics then follow as in Lemma 6.1.

At the information dates T_k we use, as in the proof of Lemma 6.3, the degenerate discrete-time Kalman filter or a Bayesian update formula. $\square$

Note that the dynamics of m^C and Q^C between information dates are the same as for the R-investor, see Lemma 6.1. The update formulas at information dates T_k, on the other hand, are the same as for the E-investor. If we have non-deterministic information dates T_k, then also the conditional covariance matrices Q^C of the C-investor are non-deterministic since updates take place at random times. Again, the timing of the information dates is the only source of randomness in the dynamics of Q^C.

For the sake of completeness we address as a last case the situation of full information, i.e. where the investor filtration is $\mathbb{F}^F$. This case corresponds to an investor who is able to observe the drift process directly. We consider it as a reference case to compare it to the other settings of information. It is clear that in this situation $m_t^F = \mathbb{E}[\mu_t \,|\, \mathcal{F}_t^F] = \mu_t$ and $Q_t^F = \mathbb{E}[(\mu_t - m_t^F)(\mu_t - m_t^F)^\top \,|\, \mathcal{F}_t^F] = \mathbf{0}_d$ for all $t \in \mathbb{T}$.

6.3. Properties of the conditional covariance matrix

In the preceding section we have stated the filtering equations for the various investors. We now deduce some properties of the conditional covariance matrices. Firstly, it is straightforward to show that an update caused by an incoming expert opinion decreases the conditional covariance matrices Q^E and Q^C of the E- and C-investor. In this respect, it is useful to consider the partial ordering $\preceq$ of symmetric matrices. Recall that for symmetric matrices $A, B \in \mathbb{R}^{d\times d}$ we write $A \preceq B$ if $B - A$ is positive semidefinite and that $A \preceq B$ in particular implies $\|A\| \leq \|B\|$.

Proposition 6.5. *It holds $Q^E_{T_k} \preceq Q^E_{T_k-}$ and $Q^E_{T_k} \preceq \Gamma_k$ as well as $Q^C_{T_k} \preceq Q^C_{T_k-}$ and $Q^C_{T_k} \preceq \Gamma_k$ for any $k \in I$.*

Proof. Let $k \in I$. From the update formula in Lemma 6.3 we know that

$$Q^E_{T_k} = Q^E_{T_k-} + \big(\rho_k(Q^E_{T_k-}) - I_d\big)Q^E_{T_k-},$$

where $\rho_k(Q) = \Gamma_k(Q + \Gamma_k)^{-1}$. Hence, we can rewrite

$$\begin{aligned} Q^E_{T_k-} - Q^E_{T_k} &= -\big(\rho_k(Q^E_{T_k-}) - I_d\big)Q^E_{T_k-} \\ &= \big(I_d - \Gamma_k(Q^E_{T_k-} + \Gamma_k)^{-1}\big)Q^E_{T_k-} \\ &= Q^E_{T_k-}(Q^E_{T_k-} + \Gamma_k)^{-1}Q^E_{T_k-}. \end{aligned}$$

Since the matrix $Q^E_{T_k-} + \Gamma_k$ is symmetric and positive definite, so is its inverse and hence also the matrix

$$Q^E_{T_k-}(Q^E_{T_k-} + \Gamma_k)^{-1}Q^E_{T_k-} = Q^E_{T_k-}(Q^E_{T_k-} + \Gamma_k)^{-1}(Q^E_{T_k-})^\top$$

by symmetry of $Q^E_{T_k-}$. It follows that $Q^E_{T_k} \preceq Q^E_{T_k-}$.

For the second assertion, we use

$$Q^E_{T_k} = \rho_k(Q^E_{T_k-})Q^E_{T_k-} = \Gamma_k(Q^E_{T_k-} + \Gamma_k)^{-1}Q^E_{T_k-} = \Gamma_k - \Gamma_k(Q^E_{T_k-} + \Gamma_k)^{-1}\Gamma_k.$$

Again, due to $Q^E_{T_k-}$ being symmetric and positive semidefinite and Γ_k being symmetric and positive definite, we can deduce that $\Gamma_k(Q^E_{T_k-} + \Gamma_k)^{-1}\Gamma_k$ is symmetric and positive semidefinite. Hence, $Q^E_{T_k} \preceq \Gamma_k$. For the C-investor, the proof is completely analogous. □

The previous proposition shows that for the E- and C-investor, the information from an incoming expert opinion decreases the conditional covariance. Also, the covariance matrix of the expert's view, Γ_k, forms an upper bound for the investors' covariance matrix after the update.

The partial ordering introduced above also proves useful for comparing the different investors among each other. Recall that the conditional covariance matrix is a means of measuring the goodness of the filter. With that interpretation, we show in the following proposition that the additional information that both the C-investor and the D-investor have in comparison to the R-investor results in a more precise estimate of the drift. Also, the conditional covariance matrix of the C-investor is smaller than that of the E-investor.

Proposition 6.6. *For any sequence $(T_k, Z_k)_{k\in I}$ we have $Q^C_t \preceq Q^R_t$ and $Q^D_t \preceq Q^R_t$ as well as $Q^C_t \preceq Q^E_t$ for all $t \in \mathbb{T}$.*

Proof. Let $(T_k, Z_k)_{k\in I}$ be any sequence of expert opinions and $(Q^C_t)_{t\in\mathbb{T}}$ the conditional covariance matrices of the corresponding filter for the C-investor. Every update decreases the covariance in the sense that

$$Q^C_{T_k} \preceq Q^C_{T_k-},$$

see Proposition 6.5. Furthermore, if $(P_t)_{t\in\mathbb{T}}$ and $(\tilde{P}_t)_{t\in\mathbb{T}}$ are solutions of the same Riccati differential equation, where the initial values fulfill $P_0 \preceq \tilde{P}_0$, then $P_t \preceq \tilde{P}_t$ for all $t \in \mathbb{T}$, see for example Kučera [39, Thm. 10]. Inductively, we can deduce that in our setting $Q^C_t \preceq Q^R_t$ for all $t \in \mathbb{T}$.

To compare the conditional covariance matrices of the D-investor and of the R-investor, we use the fact that for any random variable X and σ-algebra $\mathcal{G}$ the conditional expectation $\mathbb{E}[X \mid \mathcal{G}]$ is the best mean-square estimate for X, meaning that

$$\mathbb{E}\big[(X - \mathbb{E}[X \mid \mathcal{G}])^2\big] \le \mathbb{E}\big[(X - Y)^2\big] \tag{6.4}$$

for all $\mathcal{G}$-measurable random variables Y. Now, take an arbitrary $x \in \mathbb{R}^d$ and $t \in \mathbb{T}$. Then we can write

$$\begin{aligned} x^\top Q^D_t x &= \mathbb{E}\big[x^\top(\mu_t - m^D_t)(\mu_t - m^D_t)^\top x \mid \mathcal{F}^D_t\big] \\ &= \mathbb{E}\big[\big(x^\top(\mu_t - m^D_t)\big)^2 \mid \mathcal{F}^D_t\big] \\ &= \mathbb{E}\big[\big(x^\top\mu_t - \mathbb{E}[x^\top\mu_t \mid \mathcal{F}^D_t]\big)^2 \mid \mathcal{F}^D_t\big]. \end{aligned}$$

Since $\mathcal{F}^R_t \subseteq \mathcal{F}^D_t$ for all $t \in \mathbb{T}$, it follows from (6.4) that

$$\mathbb{E}[x^\top Q^D_t x] = \mathbb{E}\big[\big(x^\top\mu_t - \mathbb{E}[x^\top\mu_t \mid \mathcal{F}^D_t]\big)^2\big] \le \mathbb{E}\big[\big(x^\top\mu_t - \mathbb{E}[x^\top\mu_t \mid \mathcal{F}^R_t]\big)^2\big] = \mathbb{E}[x^\top Q^R_t x].$$

We already know that Q^R_t and Q^D_t are deterministic, so the above inequality simplifies to

$$x^\top Q^D_t x \le x^\top Q^R_t x.$$

Since $x \in \mathbb{R}^d$ was arbitrary, it follows that $Q^D_t \preceq Q^R_t$.

The comparison of Q^C_t and Q^E_t also works inductively. Note that $Q^C_0 = Q^E_0 = \Sigma_0$. Given that $Q^C_{T_{k-1}} \preceq Q^E_{T_{k-1}}$ for some $k \in I$, we have $Q^C_t \preceq Q^E_t$ for all $t \in [T_{k-1}, T_k)$ due to the additional information as when comparing the D- and R-investor. In particular, we deduce $Q^C_{T_k-} \preceq Q^E_{T_k-}$ and hence the matrix

$$\begin{aligned} Q^E_{T_k} - Q^C_{T_k} &= \Gamma_k - \Gamma_k(Q^E_{T_k-} + \Gamma_k)^{-1}\Gamma_k - \Gamma_k + \Gamma_k(Q^C_{T_k-} + \Gamma_k)^{-1}\Gamma_k \\ &= \Gamma_k\big((Q^C_{T_k-} + \Gamma_k)^{-1} - (Q^E_{T_k-} + \Gamma_k)^{-1}\big)\Gamma_k \end{aligned}$$

is also positive semidefinite. So $Q^C_{T_k} \preceq Q^E_{T_k}$. Inductively, it follows that $Q^C_t \preceq Q^E_t$ for all $t \in \mathbb{T}$. □

The preceding proposition gives a relation between the conditional covariance matrices of the different investors in terms of the partial ordering of symmetric matrices. Recall that $Q^F_t = \mathbf{0}_d$ for all $t \in \mathbb{T}$. Therefore, in addition to the results from the proposition, it also trivially holds $Q^F_t \preceq Q^H_t$ for all $H \in \{R, E, C, D\}$.

Figure 6.1 illustrates how the various conditional covariances Q^H behave in the course of time. For illustration purposes we consider a market with $d = 1$ stock and choose $\mathbb{T} = [0, T]$ for a time horizon $T = 1$. We assume that there are $n = 10$ equidistant information dates $T_k = t_k = \frac{k}{n}$, $k = 1, \ldots, n$, on $[0, T]$ and that the expert's variance Γ is the same at each information date.

When looking at Q_t^R and Q_t^D we observe that they are smooth and decreasing functions in time and seem to converge, as t increases, to some finite value. One can also see that $Q_t^D \le Q_t^R$ for any $t \in \mathbb{T}$, which has been proven in Proposition 6.6. For the E-investor, the conditional variance decreases at each information date, which is due to Proposition 6.5. It is also striking that $Q_{t_k}^E$ and $Q_{t_k-}^E$ seem to converge to some finite value for large values of k. The conditional variance of the C-investor always lies below the minimum of Q_t^R and Q_t^E, again due to Proposition 6.6. As in the case with expert opinions only, the conditional variance decreases at information dates and $Q_{t_k}^C$ and $Q_{t_k-}^C$ seem to converge.

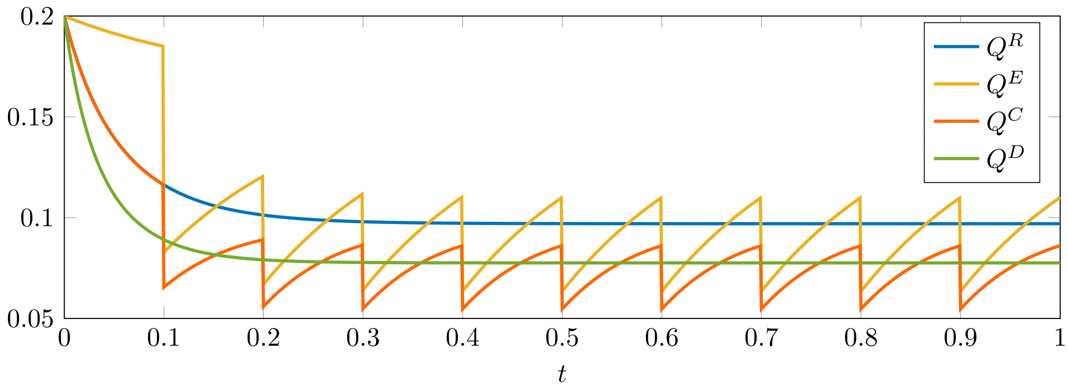


Figure 6.1.: An example for the behavior of the conditional variances in a market with $d = 1$ stock, time horizon $T = 1$ and $n = 10$ equidistant information dates. Further parameters are $\alpha = 3$, $\beta = 1$, $\sigma_R = \sigma_J = 0.15$, $\Gamma = 0.15$, $\Sigma_0 = 0.2$.

The figure illustrates the properties of the conditional covariance matrices that we have shown in this section. It also reveals an interesting asymptotic behavior of the conditional covariance matrices as t increases. The aim of the next chapter will be to analyze this asymptotic behavior in more detail and to investigate which of the properties that we have observed in the one-dimensional example hold in general in multivariate financial markets.

7. Asymptotic Behavior on an Infinite Time Horizon

In this chapter we consider an infinite time horizon, i.e. $\mathbb{T} = [0, \infty)$. Our aim is to derive convergence results for the conditional covariance matrices Q_t^H, $H \in \{R, D, E, C\}$, as t goes to infinity. Many results from this chapter, aside from Proposition 7.7(ii), Theorem 7.8 and Proposition 7.9, are already proven in the Master's thesis Westphal [64] and repeated here for completeness and since they are needed for later computations.

7.1. Observation of diffusions only

We first consider the conditional covariance matrix of the R-investor. The following definition can be found for example in Wonham [65] and Kučera [39] and proves to be useful for analyzing the asymptotic behavior of Q_t^R as t goes to infinity.

Definition 7.1. A quadratic matrix is called *stable* if all its eigenvalues have negative real parts. A pair (A, B) of matrices $A, B \in \mathbb{R}^{n \times n}$ is called *stabilizable* if there exists some matrix $L \in \mathbb{R}^{n \times n}$ such that $A + BL$ is stable. It is called *detectable* if there exists some matrix $F \in \mathbb{R}^{n \times n}$ such that $FA + B$ is stable.

With the help of the above definition we now prove that, as t tends to infinity, Q_t^R converges. We provide a proof that makes use of results from Kučera [39].

Theorem 7.2. *Starting with any initial covariance matrix Σ_0 it holds*

$$\lim_{t \to \infty} Q_t^R = Q_\infty^R$$

for a positive-semidefinite matrix Q_∞^R. Furthermore, Q_∞^R is the unique positive-semidefinite solution of the algebraic Riccati equation

$$-\alpha Q - Q\alpha + \beta\beta^\top - Q(\sigma_R \sigma_R^\top)^{-1} Q = \mathbf{0}_d.$$

Proof. We make use of the results in the review paper on matrix Riccati equations by Kučera [39]. After applying a simple time reversion to the differential equation considered in that paper, Theorem 17 therein states that the solution P of the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} P(t) = -P(t)BB^\top P(t) + P(t)A + A^\top P(t) + C^\top C, \qquad P(t_0) = P_0, \tag{7.1}$$

satisfies

$$\lim_{t \to \infty} P(t) = P_\infty$$

under the assumption that (A, B) is stabilizable and (C, A) is detectable. Theorem 5 in the aforementioned paper ensures that P_∞ is the unique positive-semidefinite solution of the quadratic algebraic Riccati equation

$$-PBB^\top P + PA + A^\top P + C^\top C = \mathbf{0}_d. \tag{7.2}$$

In our model, Q_t^R follows the dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_t^R = -\alpha Q_t^R - Q_t^R \alpha + \beta\beta^\top - Q_t^R (\sigma_R \sigma_R^\top)^{-1} Q_t^R, \qquad Q_0^R = \Sigma_0.$$

Let τ denote the symmetric positive-definite root of $(\sigma_R\sigma_R^\top)^{-1}$, i.e. $\tau^2 = (\sigma_R\sigma_R^\top)^{-1}$. Comparing with (7.1) and (7.2), it is sufficient to show that $(-\alpha, \tau)$ is stabilizable and $(\beta^\top, -\alpha)$ is detectable. Note that $(-\alpha) + \tau(-I_d) = -(\alpha + \tau)$ is symmetric which implies that all its eigenvalues are real. Now

$$\lambda_{\max}\big(-(\alpha+\tau)\big) = -\,\lambda_{\min}(\alpha+\tau) \le -\big(\lambda_{\min}(\alpha) + \lambda_{\min}(\tau)\big) < 0,$$

where we have used Weyl's inequality, see Horn and Johnson [32, Thm. 4.3.1], and the fact that both α and τ are positive definite. Hence, the pair $(-\alpha, \tau)$ is stabilizable. Furthermore, the matrix $(-\beta)\beta^\top + (-\alpha) = -(\beta\beta^\top + \alpha)$ is also symmetric and

$$\lambda_{\max}\big(-(\beta\beta^\top+\alpha)\big) = -\,\lambda_{\min}(\beta\beta^\top+\alpha) \le -\big(\lambda_{\min}(\beta\beta^\top) + \lambda_{\min}(\alpha)\big) < 0,$$

where we have used positive definiteness of α and $\beta\beta^\top$. Hence, $(\beta^\top, -\alpha)$ is detectable. □

In a one-dimensional market, we even obtain an explicit formula for the matrix Q_∞^R, see Gabih et al. [25, Prop. 4.6]. In the multivariate case, one has to calculate Q_∞^R numerically.

The above asymptotic result is also helpful because it ensures boundedness of the conditional covariance matrices of not only the R-investor but also the C- and D-investor. For proving our results in the subsequent chapter we need to find upper bounds for various expressions that involve the conditional covariance matrices Q^D and Q^C. The following lemma states the required boundedness.

Lemma 7.3. *There exists a constant $C_Q > 0$ such that*

$$\|Q_t^R\| \le C_Q, \quad \|Q_t^C\| \le C_Q \quad \textit{and} \quad \|Q_t^D\| \le C_Q$$

for all $t \in \mathbb{T}$.

Proof. From Proposition 6.6 we know that $Q_t^C \preceq Q_t^R$ and $Q_t^D \preceq Q_t^R$ for any $t \in \mathbb{T}$. In particular, it follows that

$$\|Q_t^C\| \le \|Q_t^R\| \quad \text{and} \quad \|Q_t^D\| \le \|Q_t^R\|.$$

Therefore, it is enough to show boundedness of Q_t^R. By Theorem 7.2 there exists a positive-semidefinite matrix Q_∞^R such that

$$\lim_{t\to\infty} Q_t^R = Q_\infty^R.$$

Hence, $\|Q_t^R\|$ is for all $t \in \mathbb{T}$ bounded by some constant $C_Q > 0$. □

For the D-investor we recall the dynamics of the filter, in particular the matrix Riccati differential equation for Q_t^D, from Lemma 6.2. We can prove convergence of Q_t^D as t goes to infinity as in the situation with return observations only.

Theorem 7.4. *Starting with any initial covariance matrix Σ_0 it holds*

$$\lim_{t\to\infty} Q_t^D = Q_\infty^D$$

for a positive-semidefinite matrix Q_∞^D. Furthermore, Q_∞^D is the unique positive-semidefinite solution of the algebraic Riccati equation

$$-\alpha Q - Q\alpha + \beta\beta^\top - Q(\sigma_D\sigma_D^\top)^{-1}Q = \mathbf{0}_d.$$

Proof. The proof is analogous to the proof of Theorem 7.2. □

7.2. Observation of discrete expert opinions

After having analyzed in the preceding section the asymptotic behavior of Q_t^R and Q_t^D as t goes to infinity we now address the E- and C-investor. We assume throughout this section that the expert opinions arrive at deterministic equidistant time points $t_k = k\Delta$, $k \in \mathbb{N}$, for some $\Delta > 0$, and that $\Gamma_k = \Gamma$ is some constant positive-definite matrix. Our goal is to establish conditions that ensure convergence of $(Q_{t_k}^H)_{k\in\mathbb{N}}$ where $H \in \{E, C\}$. The following lemma identifies conditions under which these sequences are monotone in the positive-semidefinite ordering.

Lemma 7.5. *Let $H \in \{E, C\}$. If $\Sigma_0 \preceq Q_{t_1}^H$, then the sequences*

$$(Q_{t_k}^H)_{k\in\mathbb{N}} \quad \text{and} \quad (Q_{t_k-}^H)_{k\in\mathbb{N}}$$

are monotonically non-decreasing. If $\Sigma_0 \succeq Q_{t_1}^H$, then they are monotonically non-increasing.

Proof. We consider first $H = E$ and show the claim by induction. Suppose for some $k \geq 1$ that $Q_{t_{k-1}}^E \preceq Q_{t_k}^E$. Then by Lemma 6.3 we can write

$$\begin{aligned}
Q_{t_{k+1}-}^E - Q_{t_k-}^E &= \Big(\mathrm{e}^{-\alpha\Delta} Q_{t_k}^E \mathrm{e}^{-\alpha\Delta} + \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)}\beta\beta^\top \mathrm{e}^{-\alpha(t_{k+1}-s)}\,\mathrm{d}s\Big) \\
&\quad - \Big(\mathrm{e}^{-\alpha\Delta} Q_{t_{k-1}}^E \mathrm{e}^{-\alpha\Delta} + \int_{t_{k-1}}^{t_k} \mathrm{e}^{-\alpha(t_k-s)}\beta\beta^\top \mathrm{e}^{-\alpha(t_k-s)}\,\mathrm{d}s\Big) \\
&= \mathrm{e}^{-\alpha\Delta}(Q_{t_k}^E - Q_{t_{k-1}}^E)\mathrm{e}^{-\alpha\Delta} + \int_0^\Delta \mathrm{e}^{\alpha s}\beta\beta^\top \mathrm{e}^{\alpha s}\,\mathrm{d}s - \int_0^\Delta \mathrm{e}^{\alpha s}\beta\beta^\top \mathrm{e}^{\alpha s}\,\mathrm{d}s \\
&= \mathrm{e}^{-\alpha\Delta}(Q_{t_k}^E - Q_{t_{k-1}}^E)\mathrm{e}^{-\alpha\Delta},
\end{aligned}$$

which is positive semidefinite by assumption. It follows that $Q_{t_k-}^E \preceq Q_{t_{k+1}-}^E$. But then also $Q_{t_k-}^E + \Gamma \preceq Q_{t_{k+1}-}^E + \Gamma$ and hence $(Q_{t_k-}^E + \Gamma)^{-1} \succeq (Q_{t_{k+1}-}^E + \Gamma)^{-1}$. So we can deduce that

$$\begin{aligned}
Q_{t_{k+1}}^E - Q_{t_k}^E &= \Gamma(Q_{t_{k+1}-}^E + \Gamma)^{-1}Q_{t_{k+1}-}^E - \Gamma(Q_{t_k-}^E + \Gamma)^{-1}Q_{t_k-}^E \\
&= \big(\Gamma - \Gamma(Q_{t_{k+1}-}^E + \Gamma)^{-1}\Gamma\big) - \big(\Gamma - \Gamma(Q_{t_k-}^E + \Gamma)^{-1}\Gamma\big) \\
&= \Gamma\Big((Q_{t_k-}^E + \Gamma)^{-1} - (Q_{t_{k+1}-}^E + \Gamma)^{-1}\Big)\Gamma
\end{aligned}$$

is positive semidefinite which yields $Q^E_{t_k} \preceq Q^E_{t_{k+1}}$. Inductively, the claim follows. In case that $\Sigma_0 \succeq Q^E_{t_1}$ one simply has to interchange $\preceq$ and $\succeq$ throughout the proof.

Secondly, we consider the case $H = C$ and assume again for some $k \geq 1$ that $Q^C_{t_{k-1}} \preceq Q^C_{t_k}$. In Lemma 6.4 we have seen that between two information dates Q^C_t follows the dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t} Q^C_t = -\alpha Q^C_t - Q^C_t \alpha + \beta\beta^\top - Q^C_t (\sigma_R \sigma_R^\top)^{-1} Q^C_t. \tag{7.3}$$

We consider the intervals $[t_{k-1}, t_k)$ and $[t_k, t_{k+1})$. In both intervals, Q^C_t evolves with the same dynamics, but for the initial values we have $Q^C_{t_{k-1}} \preceq Q^C_{t_k}$. Since the differential equation (7.3) is a Riccati equation, it follows from Kučera [39, Thm. 10] that $Q^C_{t_{k-1}+h} \preceq Q^C_{t_k+h}$ for any $h \in [0, \Delta)$, and in particular $Q^C_{t_k-} \preceq Q^C_{t_{k+1}-}$. As above for the E-investor, it follows from the update formula that also $Q^C_{t_k} \preceq Q^C_{t_{k+1}}$. Inductively, the claim has been shown. The proof in the other case is again completely analogous. ☐

If either of the inequalities in the assumption of the lemma holds, we can deduce convergence of the sequences

$$(Q^H_{t_k})_{k\in\mathbb{N}} \quad \text{and} \quad (Q^H_{t_k-})_{k\in\mathbb{N}}$$

as k goes to infinity. This is shown in the next proposition.

Proposition 7.6. *Let $H \in \{E, C\}$ and assume that either $\Sigma_0 \preceq Q^H_{t_1}$ or $\Sigma_0 \succeq Q^H_{t_1}$. Then there exist symmetric positive-semidefinite matrices L^H and $U^H \in \mathbb{R}^{d\times d}$ such that*

$$\lim_{k\to\infty} Q^H_{t_k} = L^H \quad \textit{and} \quad \lim_{k\to\infty} Q^H_{t_k-} = U^H.$$

Proof. By Lemma 7.5 the sequences

$$(Q^H_{t_k})_{k\in\mathbb{N}} \quad \text{and} \quad (Q^H_{t_k-})_{k\in\mathbb{N}}$$

are monotone. Next, we show that they are bounded. For the C-investor, recall from Lemma 7.3 that Q^C_t is bounded. For the E-investor, note from Lemma 6.3 that we are able to rewrite the dynamics of Q^E_t between two information dates t_{k-1} and t_k in the form

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} Q^E_t &= -\alpha \mathrm{e}^{-\alpha(t-t_{k-1})} \left(Q^E_{t_{k-1}} + \int_{t_{k-1}}^t \mathrm{e}^{\alpha(s-t_{k-1})} \beta\beta^\top \mathrm{e}^{\alpha(s-t_{k-1})}\,\mathrm{d}s \right) \mathrm{e}^{-\alpha(t-t_{k-1})} \\
&\quad - \mathrm{e}^{-\alpha(t-t_{k-1})} \left(Q^E_{t_{k-1}} + \int_{t_{k-1}}^t \mathrm{e}^{\alpha(s-t_{k-1})} \beta\beta^\top \mathrm{e}^{\alpha(s-t_{k-1})}\,\mathrm{d}s \right) \mathrm{e}^{-\alpha(t-t_{k-1})} \alpha \\
&\quad + \mathrm{e}^{-\alpha(t-t_{k-1})} \mathrm{e}^{\alpha(t-t_{k-1})} \beta\beta^\top \mathrm{e}^{\alpha(t-t_{k-1})} \mathrm{e}^{-\alpha(t-t_{k-1})} \\
&= -\alpha Q^E_t - Q^E_t \alpha + \beta\beta^\top.
\end{aligned}$$

This is a degenerate Riccati differential equation in which the quadratic term vanishes. From Definition 7.1 it follows immediately that $(-\alpha, \mathbf{0}_d)$ is stabilizable. By Kučera [39, Thm. 11] the solution of this differential equation is bounded. Since at each information date t_k Proposition 6.5 ensures $Q^E_{t_k} \preceq Q^E_{t_k-}$, and by applying again Kučera [39, Thm. 10], we can conclude that there is some $M \in \mathbb{R}^{d\times d}$ such that $x^\top Q^E_t x \leq x^\top M x$ for all $x \in \mathbb{R}^d$ and $t \geq 0$.

Hence, for $H \in \{E, C\}$, the sequences $(Q^H_{t_k})_{k\in\mathbb{N}}$ and $(Q^H_{t_k-})_{k\in\mathbb{N}}$ are monotone and bounded sequences of symmetric matrices. It follows that the limits

$$\lim_{k\to\infty} Q^H_{t_k} = L^H \quad \text{and} \quad \lim_{k\to\infty} Q^H_{t_k-} = U^H$$

exist for real-valued matrices L^H and $U^H \in \mathbb{R}^{d\times d}$, see Wonham [65, Lem. 3.1] for instance. □

The condition $\Sigma_0 \preceq Q^H_{t_1}$ is always fulfilled if $\Sigma_0 = \mathbf{0}_d$, i.e. if we assume that the initial drift μ_0 is known. In that case, we obtain from Lemma 7.5 and Proposition 7.6 that $(Q^H_{t_k})_{k\in\mathbb{N}}$ and $(Q^H_{t_k-})_{k\in\mathbb{N}}$ are monotonically increasing and converge to L^H, respectively U^H. Note however that the relation $\preceq$ is not a total order on the space of symmetric matrices. Therefore it might be the case that neither $\Sigma_0 \preceq Q^H_{t_1}$ nor $\Sigma_0 \succeq Q^H_{t_1}$. In the following proposition we identify conditions on the set of model parameters that nevertheless guarantee convergence.

Proposition 7.7. *(i) Suppose that the model parameters α and Γ and the interval size Δ are chosen in such a way that*

$$\frac{\lambda_{\max}(\Gamma)}{\lambda_{\min}(\Gamma)} < \mathrm{e}^{\lambda_{\min}(\alpha)\Delta}.$$

Then there exist symmetric positive-semidefinite matrices L^E and $U^E \in \mathbb{R}^{d\times d}$ such that

$$\lim_{k\to\infty} Q^E_{t_k} = L^E \quad \textit{and} \quad \lim_{k\to\infty} Q^E_{t_k-} = U^E,$$

and L^E and U^E are independent of the choice of Σ_0.

(ii) There exists a $\Delta_0 > 0$ such that if $\Delta \geq \Delta_0$ then there exist symmetric positive-semidefinite matrices L^C and $U^C \in \mathbb{R}^{d\times d}$ such that

$$\lim_{k\to\infty} Q^C_{t_k} = L^C \quad \textit{and} \quad \lim_{k\to\infty} Q^C_{t_k-} = U^C,$$

and L^C and U^C are independent of the choice of Σ_0.

Proof. (i) For the proof we denote with Q^E the conditional covariance matrices of the E-investor starting in $Q^E_0 = \Sigma_0$ and with $Q^{E,0}$ the conditional covariance matrices when starting with initial covariance $Q^{E,0}_0 = \mathbf{0}_d$. In the following, we show that

$$\|Q^E_{t_k} - Q^{E,0}_{t_k}\| \leq q\|Q^E_{t_{k-1}} - Q^{E,0}_{t_{k-1}}\| \tag{7.4}$$

for all $k \geq 1$ and some $q < 1$. From (7.4) it then follows that

$$\|Q^E_{t_k} - Q^{E,0}_{t_k}\|$$

goes to zero as k goes to infinity. Since $(Q^{E,0}_{t_k})_{k\in\mathbb{N}}$ converges to some matrix L^E by Proposition 7.6, we then deduce that $(Q^E_{t_k})_{k\in\mathbb{N}}$ converges to the same limit.

Let $k \geq 1$. A calculation similar to the one in the proof of Lemma 7.5 yields

$$\begin{aligned} Q^E_{t_k} - Q^{E,0}_{t_k} &= \Gamma\big((Q^{E,0}_{t_k-} + \Gamma)^{-1} - (Q^E_{t_k-} + \Gamma)^{-1}\big)\Gamma \\ &= \Gamma(Q^{E,0}_{t_k-} + \Gamma)^{-1}(Q^E_{t_k-} - Q^{E,0}_{t_k-})(Q^E_{t_k-} + \Gamma)^{-1}\Gamma. \end{aligned} \tag{7.5}$$

Here we have used that

$$A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1}$$

for quadratic, nonsingular matrices A and B of the same size. Further, we calculate again similarly to the proof of Lemma 7.5 that

$$Q^E_{t_k-} - Q^{E,0}_{t_k-} = \mathrm{e}^{-\alpha\Delta}(Q^E_{t_{k-1}} - Q^{E,0}_{t_{k-1}})\mathrm{e}^{-\alpha\Delta}. \tag{7.6}$$

Plugging this representation into (7.5) yields

$$Q^E_{t_k} - Q^{E,0}_{t_k} = \Gamma(Q^{E,0}_{t_k-} + \Gamma)^{-1}\mathrm{e}^{-\alpha\Delta}(Q^E_{t_{k-1}} - Q^{E,0}_{t_{k-1}})\mathrm{e}^{-\alpha\Delta}(Q^E_{t_k-} + \Gamma)^{-1}\Gamma.$$

We now apply the spectral norm and use submultiplicativity to obtain

$$\big\|Q^E_{t_k} - Q^{E,0}_{t_k}\big\| \le \big\|\Gamma(Q^{E,0}_{t_k-} + \Gamma)^{-1}\mathrm{e}^{-\alpha\Delta}\big\| \, \big\|Q^E_{t_{k-1}} - Q^{E,0}_{t_{k-1}}\big\| \, \big\|\mathrm{e}^{-\alpha\Delta}(Q^E_{t_k-} + \Gamma)^{-1}\Gamma\big\|. \tag{7.7}$$

We then derive the upper bound

$$\begin{aligned}\big\|\Gamma(Q^{E,0}_{t_k-} + \Gamma)^{-1}\mathrm{e}^{-\alpha\Delta}\big\| &\le \|\Gamma\| \, \|(Q^{E,0}_{t_k-} + \Gamma)^{-1}\| \, \|\mathrm{e}^{-\alpha\Delta}\| \\ &= \frac{\lambda_{\max}(\Gamma)}{\lambda_{\min}(Q^{E,0}_{t_k-} + \Gamma)}\mathrm{e}^{-\lambda_{\min}(\alpha)\Delta} \le \frac{\lambda_{\max}(\Gamma)}{\lambda_{\min}(\Gamma)}\mathrm{e}^{-\lambda_{\min}(\alpha)\Delta},\end{aligned}$$

and the same upper bound holds for the last factor in (7.7). Putting these results together results in

$$\big\|Q^E_{t_k} - Q^{E,0}_{t_k}\big\| \le q\big\|Q^E_{t_{k-1}} - Q^{E,0}_{t_{k-1}}\big\|,$$

where

$$q = \left(\frac{\lambda_{\max}(\Gamma)}{\lambda_{\min}(\Gamma)}\mathrm{e}^{-\lambda_{\min}(\alpha)\Delta}\right)^2 < 1$$

due to our assumptions on the parameters. This proves (7.4) and hence convergence of $(Q^E_{t_k})_{k\in\mathbb{N}}$ to L^E. Convergence of $(Q^E_{t_k-})_{k\in\mathbb{N}}$ to U^E then follows immediately from (7.6).

(ii) Here we follow the same idea as in (i) for the E-investor. Denote with Q^C the conditional covariance matrices of the C-investor with initial value $Q^C_0 = \Sigma_0$ and with $Q^{C,0}$ those with initial value $Q^{C,0}_0 = \mathbf{0}_d$. By Proposition 7.6 the sequence $(Q^{C,0}_{t_k})_{k\in\mathbb{N}}$ converges to a symmetric positive-semidefinite matrix $L^C \in \mathbb{R}^{d\times d}$.

For any $k \ge 1$ we have, as in (i), the update step

$$Q^C_{t_k} - Q^{C,0}_{t_k} = \Gamma(Q^{C,0}_{t_k-} + \Gamma)^{-1}(Q^C_{t_k-} - Q^{C,0}_{t_k-})(Q^C_{t_k-} + \Gamma)^{-1}\Gamma. \tag{7.8}$$

Furthermore, Bucy and Joseph [8, Thm. 5.4], see also Zabczyk [66], state that solutions to Riccati differential equations are exponentially stable. In particular, there exist constants $C > 0$ and $\gamma > 0$ such that

$$\big\|Q^C_{t_k-} - Q^{C,0}_{t_k-}\big\| \le C\big\|Q^C_{t_{k-1}} - Q^{C,0}_{t_{k-1}}\big\|\,\mathrm{e}^{-2\gamma\Delta} \tag{7.9}$$

for all $k \ge 1$. Plugging (7.9) into (7.8) and using the same estimations as in (i) yields

$$\big\|Q^C_{t_k} - Q^{C,0}_{t_k}\big\| \le C\left(\frac{\lambda_{\max}(\Gamma)}{\lambda_{\min}(\Gamma)}\mathrm{e}^{-\gamma\Delta}\right)^2\big\|Q^C_{t_{k-1}} - Q^{C,0}_{t_{k-1}}\big\|.$$

Hence, there exists $\Delta_0 > 0$ such that if $\Delta \geq \Delta_0$ we have

$$C\left(\frac{\lambda_{\max}(\Gamma)}{\lambda_{\min}(\Gamma)}\mathrm{e}^{-\gamma\Delta}\right)^2 < 1.$$

In that case, we can deduce that the sequence $(Q^C_{t_k})_{k\in\mathbb{N}}$ converges to L^C. Convergence of $(Q^C_{t_k-})_{k\in\mathbb{N}}$ to the matrix U^C then follows immediately from (7.9). $\square$

The preceding proposition imposes conditions on the model parameters that ensure convergence of the sequences

$$(Q^H_{t_k})_{k\in\mathbb{N}} \quad \text{and} \quad (Q^H_{t_k-})_{k\in\mathbb{N}}$$

for $H \in \{E, C\}$. Note that convergence holds given that the interval size Δ between the arrivals of expert opinions is sufficiently large. For the E-investor we have an explicit relation between Δ on the one hand and α and Γ on the other hand that yields a sufficient condition for convergence. It can also be seen from part (i) of the preceding proposition that convergence always holds for the E-investor if Γ is a multiple of the identity matrix. For the C-investor we only have existence of a threshold Δ_0, but no explicit representation. The result in part (ii) of the proposition is a new extension to the results in the Master's thesis [64], and the same is true for the subsequent results in this chapter.

Given that the limits of the sequences $(Q^E_{t_k})_{k\in\mathbb{N}}$ and $(Q^E_{t_k-})_{k\in\mathbb{N}}$ exist for the E-investor we can show that they yield asymptotic upper and lower bounds for the trace of Q^E_t.

Theorem 7.8. *Consider the case with expert opinions only and assume that the limits L^E and U^E of the sequences $(Q^E_{t_k})_{k\in\mathbb{N}}$, respectively $(Q^E_{t_k-})_{k\in\mathbb{N}}$ exist. Then*

$$\liminf_{t\to\infty} \operatorname{tr}(Q^E_t) = \operatorname{tr}(L^E) \quad \textit{and} \quad \limsup_{t\to\infty} \operatorname{tr}(Q^E_t) = \operatorname{tr}(U^E).$$

Proof. Throughout this proof, let $(Q_t)_{t\geq 0}$ be the solution of

$$\frac{\mathrm{d}}{\mathrm{d}t}Q_t = -\alpha Q_t - Q_t\alpha + \beta\beta^\top, \quad Q_0 = \Sigma_0.$$

Note that Q_t follows the same dynamics as Q^E_t, but no updates take place. This corresponds to an investor who observes neither expert opinions nor the return diffusion. In particular, it holds $Q^E_t \preceq Q_t$ for all $t \geq 0$. As in Theorem 7.2 it follows that $\lim_{t\to\infty} Q_t = Q_\infty$ where $Q_\infty \in \mathbb{R}^{d\times d}$ is a symmetric positive-semidefinite matrix solving

$$-\alpha Q_\infty - Q_\infty\alpha + \beta\beta^\top = \mathbf{0}_d,$$

in particular

$$-2\operatorname{tr}(\alpha Q_\infty) + \operatorname{tr}(\beta\beta^\top) = 0. \tag{7.10}$$

In the following, we prove an asymptotic bound for the minimal eigenvalue of $Q_\infty - Q^E_{t_k+h}$ where $h \in [0, \Delta)$. For that purpose, define

$$G_h = \mathrm{e}^{-\alpha h}\left(L^E + \int_0^h \mathrm{e}^{\alpha s}\beta\beta^\top\mathrm{e}^{\alpha s}\,\mathrm{d}s\right)\mathrm{e}^{-\alpha h}$$

for any $h \in [0,\Delta)$ and note that

$$\big\|G_h - Q^E_{t_k+h}\big\| = \big\|\mathrm{e}^{-\alpha h}\big(L^E - Q^E_{t_k}\big)\mathrm{e}^{-\alpha h}\big\| \leq \big\|L^E - Q^E_{t_k}\big\|,$$

in particular the sequences $(G_h - Q^E_{t_k+h})_{k\in\mathbb{N}}$ converge uniformly to zero. Using Weyl's inequality, see Horn and Johnson [32, Thm. 4.3.1], we get

$$\begin{aligned}\inf_{h\in[0,\Delta)} \lambda_{\min}(Q_\infty - Q^E_{t_k+h}) &= \inf_{h\in[0,\Delta)} \lambda_{\min}(Q_\infty - G_h + G_h - Q^E_{t_k+h}) \\ &\geq \inf_{h\in[0,\Delta)} \big\{\lambda_{\min}(Q_\infty - G_h) + \lambda_{\min}(G_h - Q^E_{t_k+h})\big\} \\ &\geq \inf_{h\in[0,\Delta)} \lambda_{\min}(Q_\infty - G_h) + \inf_{h\in[0,\Delta)} \lambda_{\min}(G_h - Q^E_{t_k+h})\end{aligned} \tag{7.11}$$

for any $k \in \mathbb{N}$. Here for any $h \in [0,\Delta)$ it holds that

$$Q_\infty - G_h = \lim_{t\to\infty} Q_t - \lim_{k\to\infty} Q^E_{t_k+h} = \lim_{k\to\infty}\big(Q_{t_k+h} - Q^E_{t_k+h}\big)$$

is positive semidefinite since $Q^E_t \preceq Q_t$ for any $t \geq 0$. This implies

$$\inf_{h\in[0,\Delta)} \lambda_{\min}(Q_\infty - G_h) \geq 0. \tag{7.12}$$

The uniform convergence of $(G_h - Q^E_{t_k+h})_{k\in\mathbb{N}}$ implies

$$\lim_{k\to\infty} \inf_{h\in[0,\Delta)} \lambda_{\min}(G_h - Q^E_{t_k+h}) = 0. \tag{7.13}$$

Plugging (7.12) and (7.13) into (7.11) yields

$$\lim_{k\to\infty} \inf_{h\in[0,\Delta)} \lambda_{\min}(Q_\infty - Q^E_{t_k+h}) \geq 0.$$

By applying a trace inequality from Wang et al. [63, Lem. 1] we now conclude

$$\lim_{k\to\infty} \inf_{h\in[0,\Delta)} \operatorname{tr}\big(\alpha(Q_\infty - Q^E_{t_k+h})\big) \geq \lim_{k\to\infty} \inf_{h\in[0,\Delta)} \operatorname{tr}(\alpha)\,\lambda_{\min}(Q_\infty - Q^E_{t_k+h}) \geq 0,$$

and therefore

$$\lim_{k\to\infty} \sup_{h\in[0,\Delta)} \operatorname{tr}(\alpha Q^E_{t_k+h}) \leq \operatorname{tr}(\alpha Q_\infty) = \frac{1}{2}\operatorname{tr}(\beta\beta^\top), \tag{7.14}$$

where we have used (7.10).

For the remainder of the proof, suppose there exists some $\varepsilon > 0$ and $0 < h_1 < h_2 < \Delta$ such that

$$\operatorname{tr}(Q^E_{t_{k_n}+h_2}) < \operatorname{tr}(Q^E_{t_{k_n}+h_1}) - \varepsilon$$

for an increasing sequence $(k_n)_{n\in\mathbb{N}}$. Define the functions $g_n\colon (0,\Delta) \to \mathbb{R}$, $h \mapsto \operatorname{tr}(Q^E_{t_{k_n}+h})$ and note that due to linearity the trace of Q^E_t is between information dates differentiable with

$$\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{tr}(Q^E_t) = \operatorname{tr}\Big(\frac{\mathrm{d}}{\mathrm{d}t}Q^E_t\Big) = \operatorname{tr}(-\alpha Q^E_t - Q^E_t\alpha + \beta\beta^\top) = -2\operatorname{tr}(\alpha Q^E_t) + \operatorname{tr}(\beta\beta^\top).$$

Hence, for all $n \in \mathbb{N}$ we have

$$\begin{aligned}-\varepsilon &> \operatorname{tr}(Q^E_{t_{k_n}+h_2}) - \operatorname{tr}(Q^E_{t_{k_n}+h_1}) = g_n(h_2) - g_n(h_1) = g_n'(h_n^*)(h_2 - h_1) \\ &= \big(-2\operatorname{tr}(\alpha Q^E_{t_{k_n+h_n^*}}) + \operatorname{tr}(\beta\beta^\top)\big)(h_2 - h_1)\end{aligned}$$

for some $h_n^* \in (h_1, h_2)$, so

$$-2\operatorname{tr}(\alpha Q^E_{t_{k_n+h_n^*}}) + \operatorname{tr}(\beta\beta^\top) < -\frac{\varepsilon}{h_2 - h_1}$$

for all $n \in \mathbb{N}$. But this is a contradiction to (7.14). Hence, the assumption was wrong. In particular, we can conclude that

$$\liminf_{t\to\infty} \operatorname{tr}(Q^E_t) = \lim_{k\to\infty} \operatorname{tr}(Q^E_{t_k}) = \operatorname{tr}(L^E) \quad \text{and} \quad \limsup_{t\to\infty} \operatorname{tr}(Q^E_t) = \lim_{k\to\infty} \operatorname{tr}(Q^E_{t_k-}) = \operatorname{tr}(U^E),$$

which finishes the proof. □

It can be shown that the statement of the previous theorem is not true when considering the C-investor instead of the E-investor, i.e. when there are additionally also return observations. When imposing the constraint that $\sigma_R\sigma_R^\top$ is a multiple of the identity matrix, however, the result carries over. This is shown in the following proposition.

Proposition 7.9. *Consider the case with expert opinions and return observations. Under the assumption that $\sigma_R\sigma_R^\top = sI_d$ for some $s > 0$ and that the limits L^C and U^C of the sequences $(Q^C_{t_k})_{k\in\mathbb{N}}$, respectively $(Q^C_{t_k-})_{k\in\mathbb{N}}$ exist it holds*

$$\liminf_{t\to\infty} \operatorname{tr}(Q^C_t) = \operatorname{tr}(L^C) \quad \textit{and} \quad \limsup_{t\to\infty} \operatorname{tr}(Q^C_t) = \operatorname{tr}(U^C).$$

Proof. Firstly, we note that between information dates the trace of Q^C_t is differentiable with

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{tr}(Q^C_t) &= \operatorname{tr}\Big(\frac{\mathrm{d}}{\mathrm{d}t}Q^C_t\Big) = \operatorname{tr}(-\alpha Q^C_t - Q^C_t\alpha + \beta\beta^\top - Q^C_t(\sigma_R\sigma_R^\top)^{-1}Q^C_t) \\ &= -2\operatorname{tr}(\alpha Q^C_t) + \operatorname{tr}(\beta\beta^\top) - \operatorname{tr}(Q^C_t(\sigma_R\sigma_R^\top)^{-1}Q^C_t).\end{aligned}$$

Hence, it holds $\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{tr}(Q^C_t) \ge 0$ if and only if

$$2\operatorname{tr}(\alpha Q^C_t) + \operatorname{tr}(Q^C_t(\sigma_R\sigma_R^\top)^{-1}Q^C_t) \le \operatorname{tr}(\beta\beta^\top).$$

Further, we have shown in Theorem 7.2 that $\lim_{t\to\infty} Q^R_t = Q^R_\infty$ where

$$-\alpha Q^R_\infty - Q^R_\infty\alpha + \beta\beta^\top - Q^R_\infty(\sigma_R\sigma_R^\top)^{-1}Q^R_\infty = \mathbf{0}_d,$$

in particular

$$2\operatorname{tr}(\alpha Q^R_\infty) + \operatorname{tr}(Q^R_\infty(\sigma_R\sigma_R^\top)^{-1}Q^R_\infty) = \operatorname{tr}(\beta\beta^\top). \tag{7.15}$$

From Proposition 6.6 we additionally know that $Q^C_t \preceq Q^R_t$ for all $t \ge 0$.

For finding an asymptotic bound for the minimal eigenvalue of $Q^R_\infty - Q^C_{t_k+h}$, $h \in [0, \Delta)$, we first define $(G^C_h)_{h\in[0,\Delta)}$ by

$$\frac{\mathrm{d}}{\mathrm{d}h}G^C_h = -\alpha G^C_h - G^C_h\alpha + \beta\beta^\top - G^C_h(\sigma_R\sigma_R^\top)^{-1}G^C_h, \quad G^C_0 = L^C,$$

i.e. the dynamics of G_h^C only differ from those of Q_t^C in the starting value. In Bucy and Joseph [8, Thm. 5.4], see also Zabczyk [66], the authors show exponential stability of the solution to Riccati differential equations. In particular, it follows that there exist constants $C > 0$ and $\gamma > 0$ such that

$$\big\|G_h^C - Q_{t_k+h}^C\big\| \le C\big\|L^C - Q_{t_k}^C\big\| \, \mathrm{e}^{-2\gamma h} \le C\big\|L^C - Q_{t_k}^C\big\|$$

for any $h \in [0, \Delta)$ and $k \in \mathbb{N}$. Since the right-hand side goes to zero as k goes to infinity, we deduce in particular that the sequences $(G_h^C - Q_{t_k+h}^C)_{k\in\mathbb{N}}$ converge uniformly to zero.

Like in the proof of the previous theorem we now deduce with Weyl's inequality

$$\begin{aligned} \inf_{h\in[0,\Delta)} \lambda_{\min}(Q_\infty^R - Q_{t_k+h}^C) &= \inf_{h\in[0,\Delta)} \lambda_{\min}(Q_\infty^R - G_h^C + G_h^C - Q_{t_k+h}^C) \\ &\ge \inf_{h\in[0,\Delta)} \big\{\lambda_{\min}(Q_\infty^R - G_h^C) + \lambda_{\min}(G_h^C - Q_{t_k+h}^C)\big\} \\ &\ge \inf_{h\in[0,\Delta)} \lambda_{\min}(Q_\infty^R - G_h^C) + \inf_{h\in[0,\Delta)} \lambda_{\min}(G_h^C - Q_{t_k+h}^C) \end{aligned} \tag{7.16}$$

for any $k \in \mathbb{N}$. Here for any $h \in [0, \Delta)$ it holds that

$$Q_\infty^R - G_h^C = \lim_{t\to\infty} Q_t^R - \lim_{k\to\infty} Q_{t_k+h}^C = \lim_{k\to\infty} \big(Q_{t_k+h}^R - Q_{t_k+h}^C\big)$$

is positive semidefinite since $Q_t^C \preceq Q_t^R$ for any $t \ge 0$. This implies

$$\inf_{h\in[0,\Delta)} \lambda_{\min}(Q_\infty^R - G_h^C) \ge 0. \tag{7.17}$$

The uniform convergence of $(G_h^C - Q_{t_k+h}^C)_{k\in\mathbb{N}}$ implies

$$\lim_{k\to\infty} \inf_{h\in[0,\Delta)} \lambda_{\min}(G_h^C - Q_{t_k+h}^C) = 0. \tag{7.18}$$

We plug (7.17) and (7.18) into (7.16) and obtain

$$\lim_{k\to\infty} \inf_{h\in[0,\Delta)} \lambda_{\min}(Q_\infty^R - Q_{t_k+h}^C) \ge 0. \tag{7.19}$$

In order to prove the claim of the proposition we need to show that

$$\lim_{k\to\infty} \sup_{h\in[0,\Delta)} 2\,\mathrm{tr}(\alpha Q_{t_k+h}^C) + \mathrm{tr}(Q_{t_k+h}^C(\sigma_R\sigma_R^\top)^{-1}Q_{t_k+h}^C) \le \mathrm{tr}(\beta\beta^\top),$$

where the right-hand side is equal to $2\,\mathrm{tr}(\alpha Q_\infty^R) + \mathrm{tr}(Q_\infty^R(\sigma_R\sigma_R^\top)^{-1}Q_\infty^R)$ by (7.15). Using cyclicity of the trace we see that the inequality above is equivalent to

$$\lim_{k\to\infty} \inf_{h\in[0,\Delta)} 2\,\mathrm{tr}\big(\alpha(Q_\infty^R - Q_{t_k+h}^C)\big) + \mathrm{tr}\big((\sigma_R\sigma_R^\top)^{-1}((Q_\infty^R)^2 - (Q_{t_k+h}^C)^2)\big) \ge 0. \tag{7.20}$$

For the first summand, it follows from Wang et al. [63, Lem. 1] that

$$\lim_{k\to\infty} \inf_{h\in[0,\Delta)} 2\,\mathrm{tr}\big(\alpha(Q_\infty^R - Q_{t_k+h}^C)\big) \ge \lim_{k\to\infty} \inf_{h\in[0,\Delta)} 2\,\mathrm{tr}(\alpha)\,\lambda_{\min}(Q_\infty^R - Q_{t_k+h}^C) \ge 0$$

by (7.19). For the second summand, we need our assumption $\sigma_R \sigma_R^\top = sI_d$. Under this assumption we obtain

$$\operatorname{tr}\big((\sigma_R \sigma_R^\top)^{-1}((Q_\infty^R)^2 - (Q_{t_k+h}^C)^2)\big) = \frac{1}{s} \operatorname{tr}\big((Q_\infty^R)^2 - (Q_{t_k+h}^C)^2\big).$$

Similarly to above, we write

$$\inf_{h\in[0,\Delta)} \operatorname{tr}\big((Q_\infty^R)^2 - (Q_{t_k+h}^C)^2\big) = \inf_{h\in[0,\Delta)} \operatorname{tr}\big((Q_\infty^R)^2 - (G_h^C)^2\big) + \inf_{h\in[0,\Delta)} \operatorname{tr}\big((G_h^C)^2 - (Q_{t_k+h}^C)^2\big).$$

Due to uniform convergence, the second summand above goes to zero as k goes to infinity. For the first summand, we recall $G_h^C \preceq Q_\infty^R$ which implies

$$\operatorname{tr}\big((G_h^C)^2\big) \leq \operatorname{tr}\big((Q_\infty^R)^2\big).$$

Plugging these results together we obtain

$$\lim_{k\to\infty} \inf_{h\in[0,\Delta)} \operatorname{tr}\big((\sigma_R \sigma_R^\top)^{-1}((Q_\infty^R)^2 - (Q_{t_k+h}^C)^2)\big) \geq 0.$$

Now, the inequality in (7.20) has been shown and the proof is complete. □

8. Asymptotic Behavior for an Increasing Number of Expert Opinions

We now address the question what happens when the number of dates at which expert opinions arrive goes to infinity on a finite time horizon. For that purpose, we fix $\mathbb{T} = [0, T]$ throughout this chapter, where $T > 0$ is some finite time horizon.

In the following, we deduce convergence results for both the conditional means and the conditional covariance matrices of an investor who observes discrete-time expert opinions that arrive more and more frequently on a finite time horizon. Note that convergence of discrete-time filters is addressed e.g. by Salgado et al. [53] or Aalto [1]. In those works the authors show convergence of the discrete-time Kalman filter to the continuous-time equivalent. In Aalto [1] the discrete-time filter is based on discrete-time observations of the continuous-time observation process whereas in Salgado et al. [53] the authors approximate both the continuous-time signal and observation by discrete-time processes. Neither of these assumptions match our model for the discrete-time expert opinions which is why we need to prove convergence in the following.

8.1. Experts with bounded covariance matrices

It seems likely that when increasing the number of expert opinions such that the time between any two information dates goes to zero and such that there is some minimal level of reliability of the experts, we should get an arbitrarily accurate estimate of the drift process μ. The corresponding statement in a financial market with one stock is proven in Gabih et al. [25, Prop. 4.3]. The result in a market with d stocks is formalized in the following theorem that is already proven in the Master's thesis Westphal [64].

Theorem 8.1. *Let $0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{n-1}^{(n)} < t_n^{(n)} = T$ be a sequence of deterministic partitions of the interval $[0, T]$. Assume that for the mesh size*

$$\Delta_n = \max_{k=1,\dots,n} \big(t_k^{(n)} - t_{k-1}^{(n)}\big)$$

we have $\lim_{n\to\infty} \Delta_n = 0$. Denote by $\Gamma_k^{(n)}$, $k = 1, \dots, n$, the covariance matrices of the expert opinions at time $t_k^{(n)}$, and assume that there exists some $C > 0$ such that for all $n \in \mathbb{N}$, $k = 1, \dots, n$, it holds $\|\Gamma_k^{(n)}\| \le C$.

Then for all $u \in (0, T]$ the conditional covariance matrices $Q_u^{E,n}$ and $Q_u^{C,n}$ that correspond to these n expert opinions fulfill

$$\lim_{n\to\infty} Q_u^{E,n} = \lim_{n\to\infty} Q_u^{C,n} = \mathbf{0}_d.$$

Proof. Firstly, we note that by Proposition 6.6 it holds $\|Q_u^{C,n}\| \le \|Q_u^{E,n}\|$ for any $u \in [0,T]$. Therefore, and by the equivalence of norms, it suffices to show that the spectral norm of $Q_u^{E,n}$ goes to zero. To shorten notation we write Q_u^n for $Q_u^{E,n}$ in the following. We also write t_k for time points $t_k^{(n)}$, keeping the dependency on n in mind.

Let $n \in \mathbb{N}$ and $k \in \{0,\dots,n-1\}$. For any $t \in [t_k, t_{k+1})$ we have shown in Lemma 6.3 that

$$Q_t^n = \mathrm{e}^{-\alpha(t-t_k)} Q_{t_k}^n \mathrm{e}^{-\alpha(t-t_k)} + \int_{t_k}^t \mathrm{e}^{-\alpha(t-s)} \beta\beta^\top \mathrm{e}^{-\alpha(t-s)}\,\mathrm{d}s. \tag{8.1}$$

At the information dates the update is given by

$$Q_{t_k}^n = \rho_k^{(n)}(Q_{t_k-}^n) Q_{t_k-}^n, \qquad \text{where} \qquad \rho_k^{(n)}(Q) = \Gamma_k^{(n)}(Q + \Gamma_k^{(n)})^{-1}.$$

The spectral norm of the first summand in (8.1) fulfills due to submultiplicativity

$$\big\|\mathrm{e}^{-\alpha(t-t_k)} Q_{t_k}^n \mathrm{e}^{-\alpha(t-t_k)}\big\| \le \big\|\mathrm{e}^{-\alpha(t-t_k)}\big\|\,\big\|Q_{t_k}^n\big\|\,\big\|\mathrm{e}^{-\alpha(t-t_k)}\big\|. \tag{8.2}$$

Now since α is symmetric positive definite, and for the spectrum of a matrix exponential it holds $\sigma(\mathrm{e}^\alpha) = \{\mathrm{e}^\lambda \mid \lambda \in \sigma(\alpha)\}$, we can conclude that e^α is also symmetric positive definite. Hence,

$$\big\|\mathrm{e}^{-\alpha(t-t_k)}\big\| = \frac{1}{\lambda_{\min}(\mathrm{e}^{\alpha(t-t_k)})} = \frac{1}{\min_{\lambda\in\sigma(\alpha)} \mathrm{e}^{\lambda(t-t_k)}} = \frac{1}{\mathrm{e}^{\lambda_{\min}(\alpha)(t-t_k)}} \le 1. \tag{8.3}$$

Combining (8.3) with (8.2) yields

$$\big\|\mathrm{e}^{-\alpha(t-t_k)} Q_{t_k}^n \mathrm{e}^{-\alpha(t-t_k)}\big\| \le \big\|Q_{t_k}^n\big\|. \tag{8.4}$$

By the same argument, we can conclude for the norm of the second summand in (8.1) that

$$\begin{aligned}\Big\|\int_{t_k}^t \mathrm{e}^{-\alpha(t-s)} \beta\beta^\top \mathrm{e}^{-\alpha(t-s)}\,\mathrm{d}s\Big\| &\le \int_{t_k}^t \big\|\mathrm{e}^{-\alpha(t-s)}\big\|\,\big\|\beta\beta^\top\big\|\,\big\|\mathrm{e}^{-\alpha(t-s)}\big\|\,\mathrm{d}s \\ &\le \big\|\beta\beta^\top\big\|(t-t_k) \le \big\|\beta\beta^\top\big\|\Delta_n.\end{aligned}$$

This, together with (8.4), yields for any $t \in (0,T]$ with $t \in [t_k, t_{k+1})$ that

$$\big\|Q_t^n\big\| \le \big\|Q_{t_k}^n\big\| + \Delta_n \big\|\beta\beta^\top\big\|. \tag{8.5}$$

By our assumption on the mesh size we can conclude for any $t \in (0,T]$ that $t \ge t_1 = t_1^{(n)}$ for all n large enough. Note that since $\beta\beta^\top$ is positive definite the matrices $Q_{t_k-}^n$ are nonsingular for all $k \ge 1$. The first summand in (8.5) can then be written as

$$\begin{aligned}\big\|\rho_k^{(n)}(Q_{t_k-}^n)Q_{t_k-}^n\big\| &= \Big\|\Gamma_k^{(n)}(Q_{t_k-}^n + \Gamma_k^{(n)})^{-1}Q_{t_k-}^n\Big\| = \Big\|\Gamma_k^{(n)}\big(I_d + (Q_{t_k-}^n)^{-1}\Gamma_k^{(n)}\big)^{-1}\Big\| \\ &= \Big\|\big((\Gamma_k^{(n)})^{-1} + (Q_{t_k-}^n)^{-1}\big)^{-1}\Big\| = \Big(\lambda_{\min}\big((\Gamma_k^{(n)})^{-1} + (Q_{t_k-}^n)^{-1}\big)\Big)^{-1}.\end{aligned}$$

Weyl's theorem, see for example Horn and Johnson [32, Thm 4.3.1], implies that

$$\begin{aligned}\Big(\lambda_{\min}\big((\Gamma_k^{(n)})^{-1} + (Q_{t_k-}^n)^{-1}\big)\Big)^{-1} &\le \Big(\lambda_{\min}\big((\Gamma_k^{(n)})^{-1}\big) + \lambda_{\min}\big((Q_{t_k-}^n)^{-1}\big)\Big)^{-1} \\ &= \bigg(\frac{1}{\|\Gamma_k^{(n)}\|} + \frac{1}{\|Q_{t_k-}^n\|}\bigg)^{-1} = \frac{\|\Gamma_k^{(n)}\|\|Q_{t_k-}^n\|}{\|\Gamma_k^{(n)}\| + \|Q_{t_k-}^n\|} \\ &\le \bigg(\frac{C}{C + \|Q_{t_k-}^n\|}\bigg)\|Q_{t_k-}^n\|,\end{aligned}$$

where we have used that $\|\Gamma_k^{(n)}\| \leq C$. Inserting this into (8.5), we get

$$\|Q_t^n\| \leq \left(\frac{C}{C + \|Q_{t_k-}^n\|}\right)\|Q_{t_k-}^n\| + \Delta_n\|\beta\beta^\top\|. \tag{8.6}$$

Next, we iterate (8.6) to obtain

$$\|Q_t^n\| \leq \prod_{j=1}^{k}\left(\frac{C}{C + \|Q_{t_j-}^n\|}\right)\|\Sigma_0\| + \Delta_n\|\beta\beta^\top\|\sum_{j=0}^{k}\prod_{l=1}^{j}\left(\frac{C}{C + \|Q_{t_{k+1-l}-}^n\|}\right).$$

By setting

$$L_k^n = \max_{j=1,\dots,k}\left(\frac{C}{C + \|Q_{t_j-}^n\|}\right)$$

we can conclude

$$\|Q_t^n\| \leq (L_k^n)^k\|\Sigma_0\| + \Delta_n\|\beta\beta^\top\|\sum_{j=0}^{k}(L_k^n)^j. \tag{8.7}$$

Now let $u \in (0, T]$ and $\varepsilon > 0$. For all $n \in \mathbb{N}$ let k_n denote the index for which $u \in [t_{k_n}, t_{k_n+1})$, or, in the case $u = T$, let $k_n = n$. Suppose that for all $n_0 \in \mathbb{N}$ there is some $n \geq n_0$ such that

$$\|Q_{t_1-}^n\|, \dots, \|Q_{t_{k_n}-}^n\| \geq \varepsilon/2.$$

Then for all $j = 1, \dots, k_n$ it holds

$$\frac{C}{C + \|Q_{t_j-}^n\|} \leq \frac{C}{C + \varepsilon/2}, \qquad \text{hence} \qquad L_{k_n}^n \leq \frac{C}{C + \varepsilon/2}.$$

Now, equation (8.7) implies

$$\begin{aligned}\|Q_{t_{k_n}-}^n\| &\leq \left(\frac{C}{C + \varepsilon/2}\right)^{k_n-1}\|\Sigma_0\| + \Delta_n\|\beta\beta^\top\|\sum_{j=0}^{k_n-1}\left(\frac{C}{C + \varepsilon/2}\right)^j \\ &\leq \left(\frac{C}{C + \varepsilon/2}\right)^{k_n-1}\|\Sigma_0\| + \Delta_n\|\beta\beta^\top\|\frac{2C + \varepsilon}{\varepsilon}.\end{aligned} \tag{8.8}$$

Since our assumption on the mesh size implies $\lim_{n\to\infty} k_n = \infty$, the right-hand side of (8.8) goes to zero as n tends to infinity. So there is some $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ it holds $\|Q_{t_{k_n}-}^n\| < \varepsilon/2$. This is a contradiction to our assumption. Hence, there is some $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ there exists some index $1 \leq l_n \leq k_n$ with $\|Q_{t_{l_n}-}^n\| < \varepsilon/2$. We denote by l_n the maximal index less or equal k_n with that property. If $l_n = k_n$, then

$$\|Q_u^n\| \leq \|\rho_{k_n}^{(n)}(Q_{t_{k_n}-}^n)Q_{t_{k_n}-}^n\| + \Delta_n\|\beta\beta^\top\| \leq \|Q_{t_{k_n}-}^n\| + \Delta_n\|\beta\beta^\top\| < \varepsilon/2 + \Delta_n\|\beta\beta^\top\|.$$

If $l_n < k_n$, then for $j = l_n + 1, \dots, k_n$ it holds $\|Q_{t_j-}^n\| \geq \varepsilon/2$. As above, one gets

$$\begin{aligned}\|Q_u^n\| &\leq \left(\frac{C}{C + \varepsilon/2}\right)^{k_n-l_n}\|\rho_{l_n}^{(n)}(Q_{t_{l_n}-}^n)Q_{t_{l_n}-}^n\| + \Delta_n\|\beta\beta^\top\|\frac{2C + \varepsilon}{\varepsilon} \\ &\leq \|Q_{t_{l_n}-}^n\| + \Delta_n\|\beta\beta^\top\|\frac{2C + \varepsilon}{\varepsilon} < \varepsilon/2 + \Delta_n\|\beta\beta^\top\|\frac{2C + \varepsilon}{\varepsilon}.\end{aligned}$$

We can choose $n_1 \geq n_0$ such that $\Delta_n\|\beta\beta^\top\|\frac{2C+\varepsilon}{\varepsilon} < \varepsilon/2$ for all $n \geq n_1$. Then $\|Q_u^n\| < \varepsilon$ for all $n \geq n_1$. $\square$

Recalling that $Q_t^F = \mathbf{0}_d$ for all $t \in [0,T]$, the above theorem shows that the covariance matrices $Q_t^{E,n}$ and $Q_t^{C,n}$ converge to the covariance matrix in the case of full information as the number of expert opinions on $[0,T]$ tends to infinity. Since the covariance matrices contain information about the quality of the drift estimators, this means that we get an arbitrarily good estimator by increasing the number of expert opinions.

Corollary 8.2. *Under the assumptions of Theorem 8.1 it holds*

$$\lim_{n\to\infty} \mathbb{E}\Big[\big\|m_t^{E,n} - \mu_t\big\|^2\Big] = \lim_{n\to\infty} \mathbb{E}\Big[\big\|m_t^{C,n} - \mu_t\big\|^2\Big] = 0$$

for any $t \in (0,T]$.

Proof. For the E-investor, due to

$$\big\|m_t^{E,n} - \mu_t\big\|^2 = (m_t^{E,n} - \mu_t)^\top (m_t^{E,n} - \mu_t) = \operatorname{tr}\big((m_t^{E,n} - \mu_t)(m_t^{E,n} - \mu_t)^\top\big)$$

and linearity of the trace we can deduce that

$$\mathbb{E}\Big[\big\|m_t^{E,n} - \mu_t\big\|^2\Big] = \operatorname{tr}\Big(\mathbb{E}\big[(m_t^{E,n} - \mu_t)(m_t^{E,n} - \mu_t)^\top\big]\Big) = \operatorname{tr}\big(\mathbb{E}\big[Q_t^{E,n}\big]\big).$$

The claim now follows directly from the preceding theorem since $Q_t^{E,n}$ is deterministic for every $t \in (0,T]$. For the C-investor the proof is analogous. $\square$

Hence, the investor with access to expert opinions essentially approximates the fully informed F-investor. In this context it does not matter whether we have an investor who observes stock returns as well as expert opinions or an investor whose only source of information are the expert opinions.

This result can be seen as follows: The law of large numbers for the increasing number of expert opinions dominates the uncertainty (worsening the filters) between the arrivals of expert opinions. Note that the assumption $\|\Gamma_k^{(n)}\| \le C$ for all $n \in \mathbb{N}$ and $k = 0, \dots, n-1$ is a way of ensuring that the experts' estimates of the drift do not become arbitrarily bad. Instead one assumes some minimal level of reliability of the experts.

In the following, we illustrate the convergence results by a numerical example. For illustration purposes, we consider a one-dimensional financial market. We list the model parameters in Table 8.1. These parameters will also be used for later simulations.

investment horizon	T	$=$	1
interest rate	r	$=$	0
mean reversion speed of drift process	α	$=$	3
volatility of drift process	β	$=$	1
mean reversion level of drift process	δ	$=$	0.05
initial mean of drift process	m_0	$=$	0.05
initial variance of drift process	Σ_0	$=$	0.2
volatility of returns	σ_R	$=$	0.25

Table 8.1.: Model parameters for numerical examples.

Example 8.3. Figure 8.1 illustrates the above convergence results for the C-investor. We consider here a financial market with $d = 1$ stock on an investment horizon of one year. The parameters of our model are as given in Table 8.1. We have equidistant information dates and choose the expert's variance $\Gamma = 0.05$ as a constant. We consider the cases $n = 10$, 100, 1000. In the upper subplot we see the conditional variances Q^R and $Q^{C,n}$ as well as $Q^F = 0$ plotted against time. The lower subplot shows a realization of the conditional means for the same parameters. Recall that $m_t^F = \mu_t$ for all $t \in [0, T]$.

We observe that for any time $t \in (0, T]$ the values of $Q_t^{C,n}$ are decreasing in n and get closer and closer to zero. This illustrates our convergence result from Theorem 8.1. The lower subplot shows that the conditional mean m_t^R is rather far away from the true drift process $\mu_t = m_t^F$. When expert opinions are included, one can see the jumps of the conditional mean at information times when an update takes place. By increasing the number n of equidistant expert opinions the conditional mean of the C-investor gets closer and closer to the true drift process. For $n = 1000$ expert opinions the estimate is most of the time very close to the true drift.

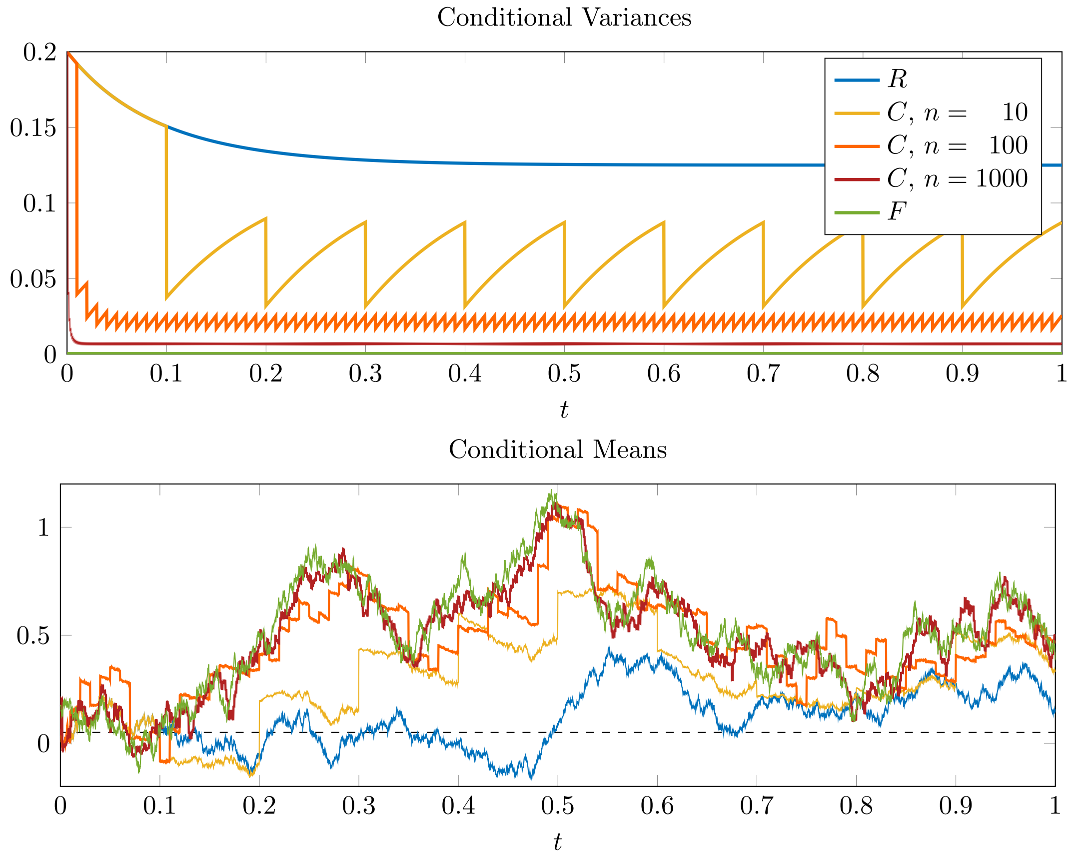


Figure 8.1.: A simulation of the filters for deterministic equidistant information dates and constant expert variance Γ. The upper subplot shows the conditional variances of the R-, F- and C-investor for various values of n, the lower subplot shows a realization of the corresponding conditional means. The dashed black line is the mean reversion level δ of the drift.

8.2. Diffusion approximations of filters

In this section we investigate the asymptotic behavior of the filters for a C-investor when the frequency of expert opinion arrivals goes to infinity but the expert's covariance matrices are not bounded. These results are the main ones in this second part of the thesis and have not been an element of the Master's thesis [64].

8.2.1. Deterministic information dates

We consider first the case for deterministic and equidistant information dates. For that purpose, let $n \in \mathbb{N}$ and $\Delta_n = \frac{T}{n}$. Now assume that $T_k = t_k$ for every $k = 1, \dots, n$, where $(t_k)_{k=1,\dots,n}$ is the sequence of deterministic time points $t_k = k\Delta_n$. So there are n expert opinions that arrive equidistantly in the time interval $[0, T]$, the distance between two information dates being equal to Δ_n. We use an additional superscript n to underline dependence on the number of expert opinions, writing for example $(Q_t^{C,n})_{t\in[0,T]}$ for the conditional covariance matrix of the filter corresponding to these n expert opinions.

In the previous section we have stated a convergence result for the case where the expert's covariance matrices $\Gamma_k^{(n)}$ are bounded. In that case, $Q_t^{C,n}$ converges to the zero matrix. This result heavily relies on the boundedness of the expert's covariance matrices, meaning that the expert's views possess some minimal level of reliability. Here, we study a different situation where more frequent expert opinions are only available at the cost of accuracy. In other words, we assume that, as Δ_n goes to zero, the variance of the expert opinions $Z_k^{(n)}$ increases. This is done for the purpose of approximating $m^{C,n}$ and $Q^{C,n}$ for large $n \in \mathbb{N}$ and large $\Gamma_k^{(n)}$.

In the following we assume for the sake of simplicity that $\Gamma_k^{(n)} = \Gamma^{(n)}$ is not time-dependent. We then show that for properly scaled $\Gamma^{(n)}$ which grows linearly in n, the information obtained from observing the discrete-time expert opinions is asymptotically the same as that from observing another diffusion process. This will be the diffusion process J that we have defined in (6.1).

Assumption 8.4. Let $(T_k^{(n)})_{k=1,\dots,n} = (t_k^{(n)})_{k=1,\dots,n}$, where $t_k^{(n)} = k\Delta_n$ for $k = 1, \dots, n$, and let the experts' covariance matrices be given by

$$\Gamma_k^{(n)} = \Gamma^{(n)} = \frac{1}{\Delta_n}\sigma_J\sigma_J^\top$$

for $k = 1, \dots, n$. Further, we assume that the expert opinions are given as

$$Z_k^{(n)} = \mu_{t_k^{(n)}} + \frac{1}{\Delta_n}\sigma_J \int_{t_k^{(n)}}^{t_{k+1}^{(n)}} \mathrm{d}W_s^J \tag{8.9}$$

for $k = 1, \dots, n$.

Recall that the matrix $\sigma_J \in \mathbb{R}^{d\times l}$ is exactly the volatility of the diffusion process J, where J was defined via the dynamics

$$\mathrm{d}J_t = \mu_t\,\mathrm{d}t + \sigma_J\,\mathrm{d}W_t^J,$$

and that σ_J has full rank. With $Z_k^{(n)}$ as defined above the discrete-time expert opinions and the continuous-time expert J are obviously correlated. In fact, it holds

$$Z_k^{(n)} \approx \frac{1}{\Delta_n} \int_{t_k^{(n)}}^{t_{k+1}^{(n)}} \mathrm{d}J_s = \frac{1}{\Delta_n} \Big(J_{t_{k+1}^{(n)}} - J_{t_k^{(n)}} \Big).$$

One can easily show by applying Donsker's Theorem that the piecewise constant process $(\widetilde{J}_t)_{t\in[0,T]}$, defined by

$$\widetilde{J}_t := \Delta_n \sum_{k=1}^{\lfloor t/\Delta_n \rfloor} Z_k^{(n)}$$

for all $t \in [0,T]$, converges in distribution to J_t as n goes to infinity. For our main convergence results we however require stronger notions of convergence. The following theorem states uniform convergence of $Q_t^{C,n}$ to Q_t^D on $[0,T]$ as n goes to infinity.

Theorem 8.5. *Under Assumption 8.4 there exists a constant $K_1 > 0$ such that*

$$\big\| Q_t^{C,n} - Q_t^D \big\| \le K_1 \Delta_n$$

for all $t \in [0,T]$. In particular,

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \big\| Q_t^{C,n} - Q_t^D \big\| = 0.$$

For the sake of better readability the proof of Theorem 8.5 is given in Section 8.3. It makes use of a discrete version of Gronwall's Lemma for error accumulation, see Lemma A.1 in Appendix A.

Using the uniform convergence of the conditional covariance matrices $Q^{C,n}$ to Q^D we can also deduce convergence of the corresponding conditional mean $m^{C,n}$ to m^D in an L^2-sense.

Theorem 8.6. *Under Assumption 8.4 there exists a constant $K_2 > 0$ such that*

$$\mathbb{E}\Big[\big\| m_t^{C,n} - m_t^D \big\|^2 \Big] \le K_2 \Delta_n$$

for all $t \in [0,T]$. In particular,

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \mathbb{E}\Big[\big\| m_t^{C,n} - m_t^D \big\|^2 \Big] = 0.$$

The proof of Theorem 8.6 can also be found in Section 8.3. The previous two theorems state that in the setting of Assumption 8.4 the filter of a C-investor who observes n equidistant expert opinions on $[0,T]$ converges to the filter of the D-investor. Recalling that the D-investor observes the diffusion processes R and J, this implies that the information obtained from observing the discrete-time expert opinions is for large n arbitrarily close to the information that comes with observing the continuous-time diffusion-type expert J.

This diffusion approximation of the discrete expert opinions is useful since the filter of the D-investor is much easier to compute than the filter of the C-investor observing n expert opinions, since no updates take place for the D-investor. We will see in Chapter 9 that the convergence carries over to convergence of the value function in a portfolio optimization problem.

Remark 8.7. Note that for the convergence of the conditional covariance matrices $Q^{C,n}$ to Q^D in Theorem 8.5 we do not need the assumption that the expert opinions are given as

$$Z_k^{(n)} = \mu_{t_k^{(n)}} + \frac{1}{\Delta_n}\sigma_J \int_{t_k^{(n)}}^{t_{k+1}^{(n)}} \mathrm{d}W_s^J$$

as in (8.9). This is because the conditional covariance matrices $Q_t^{C,n}$ do not depend on the actual form of the expert opinions, see Lemma 6.4. Hence, it would be sufficient to assume that the experts' covariance matrices are given by

$$\Gamma_k^{(n)} = \Gamma^{(n)} = \frac{1}{\Delta_n}\sigma_J\sigma_J^\top .$$

The assumption on the form of $Z_k^{(n)}$ is only needed in Theorem 8.6 where the conditional mean $m_t^{C,n}$ is considered.

Remark 8.8. Theorem 8.6 states L^2-convergence of the conditional mean $m_t^{C,n}$ to m_t^D for any $t \in [0,T]$. The joint distribution of the conditional means is Gaussian. A classical result, see for example Rosiński and Suchanecki [52, Lem. 2.1], hence yields that in this case L^2-convergence implies L^p-convergence for any $1 \le p < \infty$.

The convergence results from the previous theorems are illustrated in the following by a numerical example.

Example 8.9. We consider as in Example 8.3 a financial market with investment horizon $T = 1$ and with $d = 1$ stock. As model parameters, we take again those from Table 8.1, additionally specifying the volatility of the continuous expert as $\sigma_J = 0.2$. In Figure 8.2 we plot the filters of the R-, D- and C-investor against time. For the C-investor we consider the cases $n = 10, 20, 100$. In the upper plot one sees the conditional variances Q^R and Q^D as well as $Q^{C,n}$ plotted against time. The lower plot shows a realization of the conditional means m^R, m^D and $m^{C,n}$ for the same parameters.

Recall that Q^R and Q^D as well as $Q^{C,n}$ for any $n \in \mathbb{N}$ are deterministic. In the upper plot of Figure 8.2 one sees that for any fixed $t \in [0,T]$, the value of Q_t^D as well as the value of $Q_t^{C,n}$ for any n is less or equal than the value of Q_t^R, as has been shown in Proposition 6.6. For the C-investors one sees that the updates at information dates lead to a decrease in the conditional variance. As the number n increases, the conditional variances $Q_t^{C,n}$ approach Q_t^D for any $t \in [0,T]$. This is due to what has been shown in Theorem 8.5.

Note that for t going to infinity, Q_t^R and Q_t^D approach a finite value which follows from Theorem 7.2, respectively Theorem 7.4. For $(Q_t^{C,n})_{t\ge 0}$ we observe a periodic behavior with asymptotic upper and lower bounds in the limit. This has been studied in Section 7.2 and in more detail in Sass et al. [55, Sec. 4.2].

In the lower subplot we show a realization of the various conditional means. For $m^{C,n}$ the updating steps at information dates are visible. In general, we observe that when increasing the value of n, the distance between the paths of m^D and $m^{C,n}$ becomes smaller, as shown in Theorem 8.6.

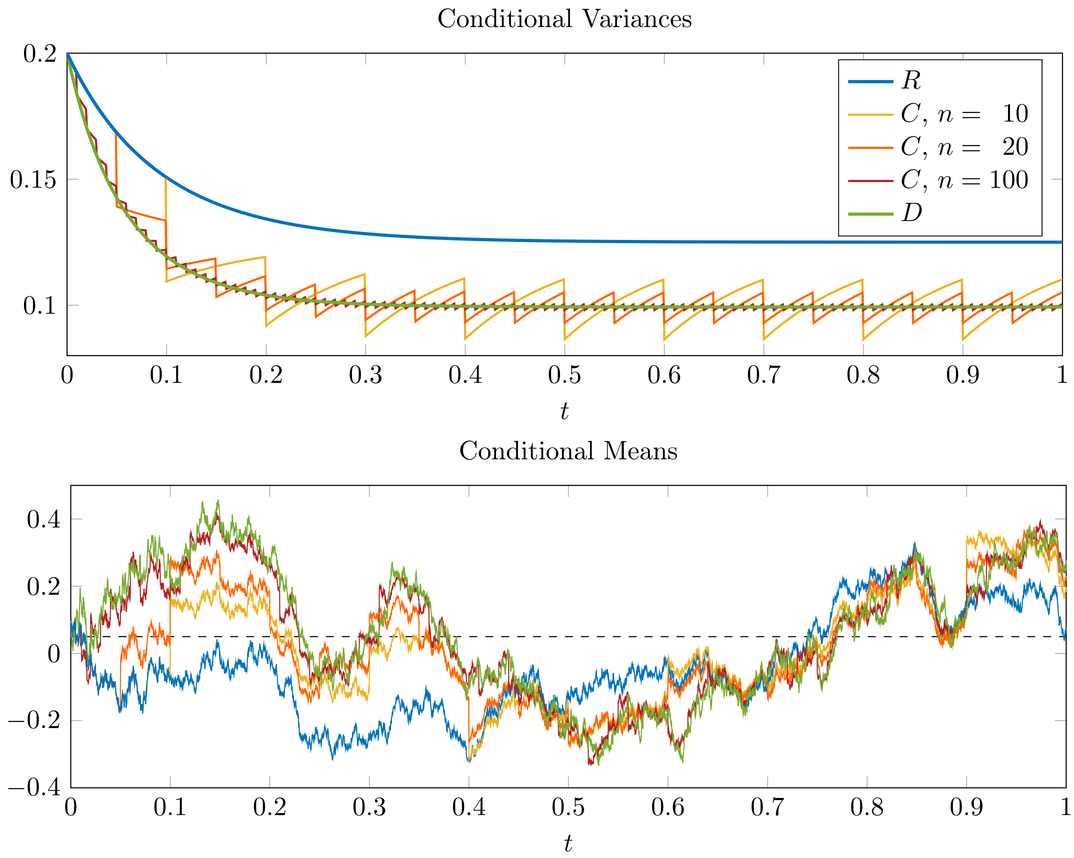


Figure 8.2.: A simulation of the filters for deterministic equidistant information dates and experts' variances growing linearly in the number of expert opinions. The upper subplot shows the conditional variances of the R-, D- and C-investor for various values of n, the lower subplot shows a realization of the corresponding conditional means. The dashed black line is the mean reversion level δ of the drift.

8.2.2. Random information dates

In this section we consider the situation where the experts' opinions do not arrive at deterministic time points but at random information dates T_k, where the waiting times $T_{k+1} - T_k$ between information dates are independent and exponentially distributed with rate $\lambda > 0$. Recall that we have set $T_0 = 0$ for ease of notation. The information dates can therefore be seen as the jump times of a standard Poisson process with intensity λ.

In this situation, the total number of expert opinions arriving in $[0, T]$ is no longer deterministic. However, as the intensity λ increases, expert opinions will arrive more and more frequently. So the question we address in this subsection is, in analogy to sending n to infinity in the preceding subsection, what happens when λ goes to infinity. We use a superscript λ to underline the dependence on the intensity. The expert opinions are of the form

$$Z_k^{(\lambda)} = \mu_{T_k^{(\lambda)}} + (\Gamma_k^{(\lambda)})^{1/2} \varepsilon_k^{(\lambda)}.$$

For constant variances $\Gamma_k^{(\lambda)} = \Gamma$, i.e. when there is some constant level of the expert's reliability which does not depend on the arrival intensity λ, one can derive a similar result for the convergence to full information as in Theorem 8.1 for the case of deterministic information dates. This result implies that for large λ the C-investor approximates the fully informed investor. More precisely, it holds

$$\lim_{\lambda\to\infty} \mathbb{E}\big[Q_t^{C,\lambda}\big] = \mathbf{0}_d \quad \text{and} \quad \lim_{\lambda\to\infty} \mathbb{E}\Big[\big\|m_t^{C,\lambda} - \mu_t\big\|^2\Big] = 0$$

for all $t \in (0, T]$, see Gabih et al. [26]. In contrast to that, we now again assume that, as the frequency of expert opinions increases, the variance of the expert opinions $Z_k^{(\lambda)}$ also increases. As in Section 8.2.1 it will turn out that letting $\Gamma_k^{(\lambda)}$ grow linearly in λ is the proper scaling for deriving diffusion limits.

Assumption 8.10. Let $(N_t^{(\lambda)})_{t\in[0,T]}$ be a standard Poisson process with intensity $\lambda > 0$ that is independent of the Brownian motions in the model. Define the information dates $(T_k^{(\lambda)})_{k=0,\dots,N_T^{(\lambda)}}$ as the jump times of that process. Furthermore, let the experts' covariance matrices be given as

$$\Gamma_k^{(\lambda)} = \Gamma^{(\lambda)} = \lambda\sigma_J\sigma_J^\top$$

for $k = 1, \dots, N_T^{(\lambda)}$. Further, we assume that

$$Z_k^{(\lambda)} = \mu_{T_k^{(\lambda)}} + \lambda\sigma_J \int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J \tag{8.10}$$

is the expert opinion at information date $T_k^{(\lambda)}$. Note that for defining the $Z_k^{(\lambda)}$, the Brownian motion W^J has to be extended to a Brownian motion on $[0, \infty)$.

Given a realization of the drift process at the random information date $T_k^{(\lambda)}$, the only randomness in the expert opinion comes from the Brownian motion W^J between the deterministic times $\frac{k-1}{\lambda}$ and $\frac{k}{\lambda}$. Recall that W^J is the Brownian motion that drives the diffusion J which we interpret as our continuous-time expert. Hence there is a direct connection between the discrete expert opinions $Z_k^{(\lambda)}$ and the continuous expert.

In the following, we will omit the superscript λ at the time points $T_k^{(\lambda)}$ for better readability, keeping the dependence on the intensity in mind.

Remark 8.11. At first glance, it seems more intuitive to construct the expert opinions as

$$\widetilde{Z}_k^{(\lambda)} = \mu_{T_k} + \sqrt{\lambda}\sigma_J \frac{1}{\sqrt{T_k - T_{k-1}}} \int_{T_{k-1}}^{T_k} \mathrm{d}W_s^J$$

rather than in (8.10). However, we later want to prove convergence of $m_t^{C,\lambda}$ to m_t^D in L^2, which requires to look at the difference of a weighted sum of $\frac{1}{\lambda}(Z_k^{(\lambda)} - \mu_{T_k})$ and $\int_0^t Q_s^D \,\mathrm{d}W_s^J$. It turns out that when replacing $Z_k^{(\lambda)}$ with $\widetilde{Z}_k^{(\lambda)}$, this leads to an integral where the integrand is defined piecewisely as

$$\bigg(\frac{1}{\sqrt{\lambda(T_k - T_{k-1})}} - 1\bigg) Q_s^D.$$

However, the term in brackets does not have a finite variance. This carries over to the weighted sum mentioned above. The core result here is that for $X \sim \mathrm{Exp}(\lambda)$, the expectation of $\frac{1}{X}$ is not finite. When considering $Z_k^{(\lambda)}$ instead, the difference that appears has finite variance since the additional randomness from the information dates is missing. Intuitively, the problem with the $\widetilde{Z}_k^{(\lambda)}$ is that the expert opinions of this form put different weight on the paths of the Brownian motion W^J in different intervals. This is in contrast to the continuous expert whose information comes from observing the diffusion J, driven by the Brownian motion W^J, continuously in time. Therefore, in terms of information about the Brownian motion W^J, the $Z_k^{(\lambda)}$ modelled as in (8.10) are closer to the continuous expert than the $\widetilde{Z}_k^{(\lambda)}$. Hence, we work with expert opinions defined via $Z_k^{(\lambda)}$ as above.

The aim of this section is to determine the behavior of the conditional covariance matrix $Q^{C,\lambda}$ and of the conditional mean $m^{C,\lambda}$ under Assumption 8.10 when λ goes to infinity, i.e. when expert opinions arrive more and more frequently, becoming at the same time less and less reliable. Here, it is useful to express the dynamics of $Q^{C,\lambda}$ and $m^{C,\lambda}$ in a way that comprises both the behavior between information dates and the jumps at times T_k. For this purpose, we work with a representation using a Poisson random measure as given in Cont and Tankov [11, Sec. 2.6].

Definition 8.12. Let $(\Omega_0, \mathcal{A}, \mathbb{Q})$ be a probability space and let ν be a measure on some measurable space $(E, \mathcal{E})$. A *Poisson random measure* with intensity measure ν is a function $N\colon \Omega_0 \times \mathcal{E} \to \mathbb{N}_0$ such that:

1. For each $\omega \in \Omega_0$, $N(\omega, \cdot)$ is a measure on $(E, \mathcal{E})$.
2. For every $B \in \mathcal{E}$, $N(\cdot, B)$ is a Poisson random variable with parameter $\nu(B)$.
3. For disjoint $E_1, \ldots, E_p \in \mathcal{E}$, the random variables $N(\cdot, E_1), \ldots, N(\cdot, E_p)$ are independent.

For a Poisson random measure N, the *compensated measure* $\tilde{N}$ is defined by $\tilde{N}\colon \Omega_0 \times \mathcal{E} \to \mathbb{R}$ with $\tilde{N}(\omega, B) = N(\omega, B) - \nu(B)$.

The following proposition states the results we will need in the following. For a proof, see Cont and Tankov [11, Sec. 2.6.3].

Proposition 8.13. *Let $E = [0,T] \times \mathbb{R}^d$. Let $(T_k)_{k\geq 1}$ be the jump times of a Poisson process with intensity $\lambda > 0$ and let U_k, $k = 1, 2, \ldots,$ be a sequence of independent multivariate standard Gaussian random variables on $\mathbb{R}^d$. For any $I \in \mathcal{B}([0,T])$ and $B \in \mathcal{B}(\mathbb{R}^d)$ let*

$$N(I \times B) = \sum_{k\colon T_k \in I} \mathbb{1}_{\{U_k \in B\}}$$

denote the number of jump times in I where U_k takes a value in B. Then N defines a Poisson random measure and it holds:

(i) The corresponding intensity measure ν satisfies

$$\nu([t_1, t_2] \times B) = \int_{[t_1,t_2]} \lambda \, \mathrm{d}t \int_B \varphi(u) \, \mathrm{d}u$$

for $0 \leq t_1 \leq t_2 \leq T$, where φ is the multivariate standard normal density on $\mathbb{R}^d$.

(ii) For Borel-measurable functions g defined on $\mathbb{R}^d$ it holds

$$\sum_{k\colon T_k\in[0,t]} g(U_k) = \int_{[0,t]}\int_{\mathbb{R}^d} g(u)\, N(\mathrm{d}s,\mathrm{d}u).$$

Now we can use the Poisson random measure for reformulating the dynamics of $Q^{C,\lambda}$.

Proposition 8.14. *Let $L\colon \mathbb{R}^{d\times d}\to\mathbb{R}^{d\times d}$ denote the function with*

$$L(Q) = -\alpha Q - Q\alpha + \beta\beta^\top - Q(\sigma_R\sigma_R^\top)^{-1}Q.$$

Then under Assumption 8.10 we can write

$$Q_t^D = \Sigma_0 + \int_0^t \big(L(Q_s^D) - Q_s^D(\sigma_J\sigma_J^\top)^{-1}Q_s^D\big)\,\mathrm{d}s$$

and

$$\begin{aligned} Q_t^{C,\lambda} = \Sigma_0 &+ \int_0^t \big(L(Q_s^{C,\lambda}) - \lambda Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda}\big)\,\mathrm{d}s\\ &- \int_0^t\int_{\mathbb{R}^d} Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda}+\lambda\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda}\,\tilde N(\mathrm{d}s,\mathrm{d}u)\end{aligned}$$

for any $t\in[0,T]$.

Proof. From Lemma 6.2 one directly obtains

$$\frac{\mathrm{d}}{\mathrm{d}t}Q_t^D = L(Q_t^D) - Q_t^D(\sigma_J\sigma_J^\top)^{-1}Q_t^D,$$

and the representation of Q_t^D follows immediately. From Lemma 6.4 recall that between information dates the matrix differential equation for $Q^{C,\lambda}$ reads

$$\frac{\mathrm{d}}{\mathrm{d}t}Q_t^{C,\lambda} = L(Q_t^{C,\lambda}).$$

Now we can use Proposition 8.13 to include the updates of $Q^{C,\lambda}$ at information dates and write

$$\mathrm{d}Q_t^{C,\lambda} = L(Q_t^{C,\lambda})\,\mathrm{d}t + \int_{\mathbb{R}^d}\big(\rho^{(\lambda)}(Q_{t-}^{C,\lambda}) - I_d\big)Q_{t-}^{C,\lambda}\,N(\mathrm{d}t,\mathrm{d}u) \tag{8.11}$$

for $\rho^{(\lambda)}(Q) = \Gamma^{(\lambda)}(Q+\Gamma^{(\lambda)})^{-1}$. Note that the integrand is matrix-valued and the integral is defined componentwise. By (8.11) we can write

$$\begin{aligned} Q_t^{C,\lambda} &= \Sigma_0 + \int_0^t L(Q_s^{C,\lambda})\,\mathrm{d}s + \int_0^t\int_{\mathbb{R}^d}\big(\rho^{(\lambda)}(Q_{s-}^{C,\lambda}) - I_d\big)Q_{s-}^{C,\lambda}\,N(\mathrm{d}s,\mathrm{d}u)\\ &= \Sigma_0 + \int_0^t L(Q_s^{C,\lambda})\,\mathrm{d}s + \int_0^t\int_{\mathbb{R}^d}\big(\rho^{(\lambda)}(Q_{s-}^{C,\lambda}) - I_d\big)Q_{s-}^{C,\lambda}\,\tilde N(\mathrm{d}s,\mathrm{d}u) \qquad (8.12)\\ &\quad+ \int_0^t\int_{\mathbb{R}^d}\big(\rho^{(\lambda)}(Q_{s-}^{C,\lambda}) - I_d\big)Q_{s-}^{C,\lambda}\,\nu(\mathrm{d}s,\mathrm{d}u).\end{aligned}$$

We see that

$$\big(\rho^{(\lambda)}(Q)-I_d\big)Q=\big(\Gamma^{(\lambda)}(Q+\Gamma^{(\lambda)})^{-1}-I_d\big)Q=-Q(Q+\Gamma^{(\lambda)})^{-1}Q=-Q(Q+\lambda\sigma_J\sigma_J^\top)^{-1}Q.$$

Therefore, the last integral in (8.12) can be written as

$$\begin{aligned}\int_0^t\int_{\mathbb{R}^d}\big(\rho^{(\lambda)}(Q^{C,\lambda}_{s-})-I_d\big)Q^{C,\lambda}_{s-}\,\nu(\mathrm{d}s,\mathrm{d}u)&=-\int_0^t\int_{\mathbb{R}^d}Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}\,\nu(\mathrm{d}s,\mathrm{d}u)\\&=-\int_0^t\int_{\mathbb{R}^d}Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}\varphi(u)\lambda\,\mathrm{d}u\,\mathrm{d}s\\&=-\int_0^t\lambda Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}\,\mathrm{d}s,\end{aligned}$$

where the second equality follows from Proposition 8.13 and the last equality is due to φ being a probability density. Plugging back in into (8.12) yields

$$\begin{aligned}Q^{C,\lambda}_t=\Sigma_0&+\int_0^t\big(L(Q^{C,\lambda}_s)-\lambda Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}\big)\,\mathrm{d}s\\&-\int_0^t\int_{\mathbb{R}^d}Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}\,\tilde N(\mathrm{d}s,\mathrm{d}u),\end{aligned}$$

and the representation of $Q^{C,\lambda}_t$ is also proven. □

The following theorem now states uniform convergence of $Q^{C,\lambda}$ to Q^D on $[0,T]$ as λ goes to infinity.

Theorem 8.15. *Under Assumption 8.10 there exists a constant $K_3>0$ and a $\lambda_0>0$ such that*

$$\mathbb{E}\Big[\sup_{t\in[0,T]}\big\|Q^{C,\lambda}_t-Q^D_t\big\|^2\Big]\leq\frac{K_3}{\lambda}$$

for all $\lambda\geq\lambda_0$. In particular,

$$\lim_{\lambda\to\infty}\mathbb{E}\Big[\sup_{t\in[0,T]}\big\|Q^{C,\lambda}_t-Q^D_t\big\|^2\Big]=0.$$

For better readability, also the proof of Theorem 8.15 is given in Section 8.3. It is based on applying Gronwall's Lemma in integral form which we recall in Lemma A.5. As in the situation with deterministic time points we can now also prove L^2-convergence of the conditional mean for the setting with random information dates T_k.

Theorem 8.16. *Under Assumption 8.10 there exists a constant $K_4>0$ and a $\tilde\lambda_0>0$ such that*

$$\mathbb{E}\Big[\big\|m^{C,\lambda}_t-m^D_t\big\|^2\Big]\leq\frac{K_4}{\sqrt{\lambda}}$$

for all $t\in[0,T]$ and $\lambda\geq\tilde\lambda_0$. In particular,

$$\lim_{\lambda\to\infty}\sup_{t\in[0,T]}\mathbb{E}\Big[\big\|m^{C,\lambda}_t-m^D_t\big\|^2\Big]=0.$$

The proof of Theorem 8.16 can also be found in Section 8.3. The previous two theorems show that under Assumption 8.10, the filter of the C-investor converges to the filter of the D-investor. These are the analogous results to those in Section 8.2.1 where we have assumed deterministic and equidistant information dates. Here, we see that the convergence result also holds for non-deterministic information dates T_k being defined as the jump times of a standard Poisson process, i.e. where the time between information dates is exponentially distributed with parameter $\lambda > 0$. When sending λ to infinity, the frequency of expert opinions goes to infinity.

As for the case with deterministic information dates, the assumption that $Z_k^{(\lambda)}$ is given as in (8.10) is only needed for the proof of Theorem 8.16. For Theorem 8.15 it is sufficient to assume that the experts' covariance matrices are of the form $\Gamma_k^{(\lambda)} = \Gamma^{(\lambda)} = \lambda \sigma_J \sigma_J^\top$.

Remark 8.17. When comparing the convergence results from Theorems 8.5 and 8.15 for the conditional covariance matrices, respectively those from Theorems 8.6 and 8.16 for the conditional means, there is a difference in the speed of convergence that we have shown. For deterministic equidistant information dates, the speed of convergence of

$$\left\| Q_t^{C,n} - Q_t^D \right\|^2$$

to zero is of the order $\frac{1}{n^2}$. For random information dates, however, we only get a speed of $\frac{1}{\lambda}$ for the convergence of

$$\mathbb{E}\left[\sup_{t \in [0,T]} \left\| Q_t^{C,\lambda} - Q_t^D \right\|^2 \right]$$

to zero. This can be explained by the additional randomness coming from the Poisson process that determines the information dates T_k in this situation.

Example 8.18. The analogous simulation as in Example 8.9 can be done for random information dates T_k that are defined as the jump times of a Poisson process. We again suppose that the model parameters are as given in Table 8.1, additionally specifying the volatility of the continuous expert as $\sigma_J = 0.2$.

Figure 8.3 shows, in addition to the filters of the R- and D-investor, the filters of the C-investor for different intensities λ. Note that the conditional variances of the filter in the case of the C-investor behave qualitatively much like in the situation with deterministic information dates. The time at which the expert opinions arrive is now random, however. The waiting times between two information dates are exponentially distributed with parameter λ. As a consequence, the updates for the C-investor do not take place as regularly as in Figure 8.2.

The upper subplot of Figure 8.3 shows realizations for $\lambda = 10, 100, 1000$. In general, by increasing the value of λ, one can increase the frequency of information dates, causing convergence of $Q_t^{C,\lambda}$ to Q_t^D for any $t \in [0,T]$, as shown in Theorem 8.15. In the lower subplot, we see the corresponding realizations of $m^{C,\lambda}$, in addition to m^R and m^D. Again, the updates in the conditional mean of the C-investor are visible.

What is also striking is that, when we consider the C-investor with intensity $\lambda = 10$, there are times where the distance between two subsequent information dates is rather big. During those times, the conditional mean of the C-investor comes closer to the conditional mean of the R-investor who does not observe any expert opinion. When the intensity λ is increased, however, the conditional mean of the C-investor approaches the conditional mean of the D-investor. For $\lambda = 1000$, the conditional means $m^{C,\lambda}$ and m^D already behave quite similarly. Note, however, that for fixed information dates $m^{C,n}$ is rather close to m^D for $n = 100$ already. The difference in the speed of convergence when comparing the situation with equidistant information dates to the situation with random information dates is discussed in the preceding remark.

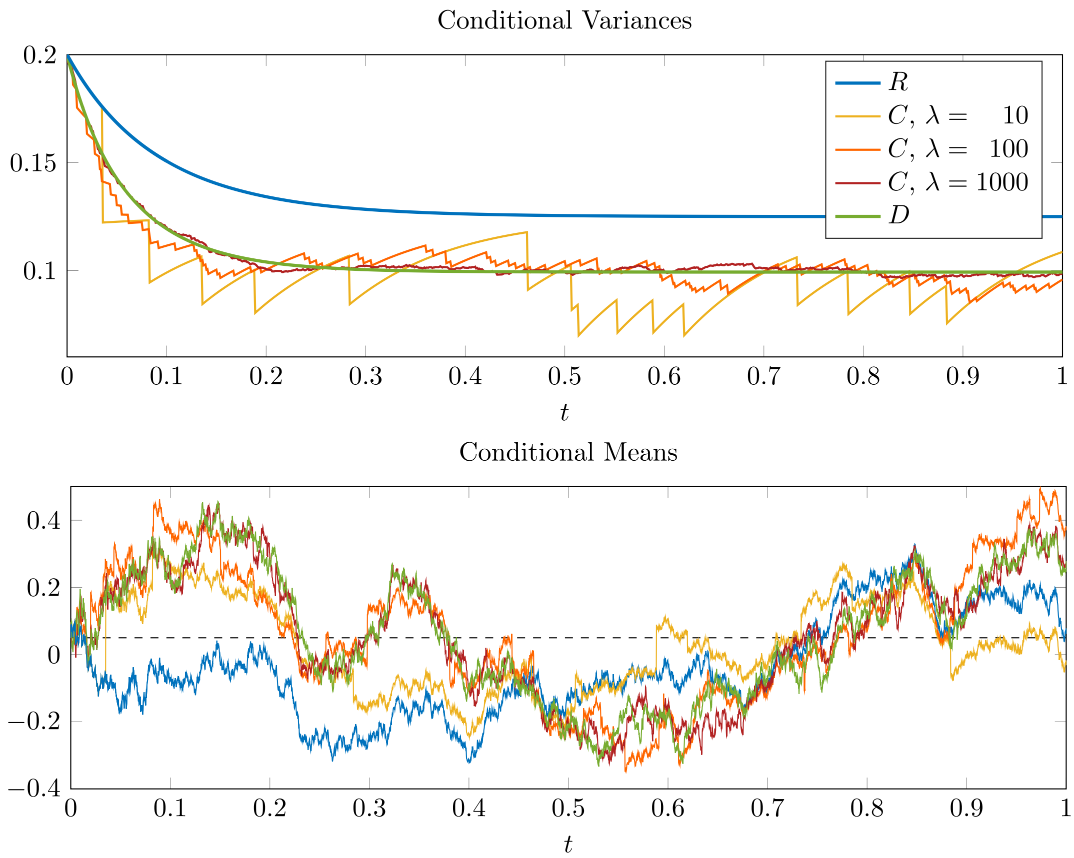


Figure 8.3.: A simulation of the filters for random information dates coming as jump times of a Poisson process and experts' variances growing linearly in the intensity of the Poisson process. The upper subplot shows the conditional variances of the R- and D-investor as well as realizations of $Q^{C,\lambda}$ for various intensities λ, the lower subplot shows a realization of the corresponding conditional means. The dashed black line is the mean reversion level δ of the drift.

We will see in the next chapter that the convergence results of this chapter carry over to convergence of the value function in a portfolio optimization problem for an investor with logarithmic utility. In that respect, the above theorems provide a useful diffusion approximation since for large intensity λ one can work with the filters of the D-investor instead of the C-investor. This is a big advantage from a numerical point of view since the filter of the D-investor is much easier to compute than the filter of the C-investor. For the conditional covariance matrices Q^C one needs to update the value at each information date. In contrast, for computing Q^D it suffices to solve just one matrix Riccati differential equation.

8.3. Proofs of main results

For better readability we give the proofs of our main results from Section 8.2 for the diffusion approximations with deterministic or random information dates in this separate section.

Proof of Theorem 8.5

Throughout the proof, we omit the superscript n at information dates $t_k^{(n)}$ for the sake of better readability, keeping the dependence on n in mind. The proof is based on finding a recursive formula for the distance between $Q^{C,n}_{t_k-}$ and $Q^D_{t_k}$ where we make use of an Euler approximation of Q^D.

Euler scheme approximation of $\boldsymbol{Q^D}$. Recall the dynamics of Q^D from Lemma 6.2. To shorten notation, let $G\colon \mathbb{R}^{d\times d}\to\mathbb{R}^{d\times d}$ with

$$G(Q) = -\alpha Q - Q\alpha + \beta\beta^\top - Q(\sigma_D\sigma_D^\top)^{-1}Q$$

denote the right-hand side of the differential equation (6.2). Then (6.2) reads as

$$\frac{\mathrm{d}}{\mathrm{d}t}Q^D_t = G(Q^D_t).$$

The first step is to approximate Q^D by an Euler scheme. Therefore, define $Q^{D,n}$ by setting

$$Q^{D,n}_t := Q^D_{t_k} + G(Q^D_{t_k})(t-t_k) \tag{8.13}$$

for all $t\in[t_k,t_{k+1})$. From a Taylor expansion we get that

$$Q^D_t = Q^D_{t_k} + G(Q^D_{t_k})(t-t_k) + \xi_t(t-t_k)^2$$

where ξ is a matrix-valued function involving the second derivative of Q^D_t. Since Q^D and its derivatives are bounded on $[0,T]$, see Lemma 7.3, the matrices ξ_t are bounded, hence the local truncation error is proportional to Δ_n^2. In other words, there exists some $C_{\mathrm{Euler}}>0$ such that

$$\big\|Q^D_t - Q^{D,n}_t\big\| \le C_{\mathrm{Euler}}\Delta_n^2 \tag{8.14}$$

for all $t\in[0,T]$.

Estimation of the error in $\boldsymbol{G}$. Let C_e, $C_Q>0$ and let $\varepsilon\in\mathbb{R}^{d\times d}$ with $\|\varepsilon\|\le C_e$, $Q\in\mathbb{R}^{d\times d}$ with $\|Q\|\le C_Q$. Then

$$\begin{aligned}
G(Q+\varepsilon) &= -\alpha(Q+\varepsilon) - (Q+\varepsilon)\alpha + \beta\beta^\top - (Q+\varepsilon)(\sigma_D\sigma_D^\top)^{-1}(Q+\varepsilon)\\
&= (-\alpha Q - Q\alpha + \beta\beta^\top - Q(\sigma_D\sigma_D^\top)^{-1}Q) - \alpha\varepsilon - \varepsilon\alpha\\
&\quad - \varepsilon(\sigma_D\sigma_D^\top)^{-1}Q - Q(\sigma_D\sigma_D^\top)^{-1}\varepsilon - \varepsilon(\sigma_D\sigma_D^\top)^{-1}\varepsilon\\
&= G(Q) - \varepsilon\big(\alpha + (\sigma_D\sigma_D^\top)^{-1}Q + (\sigma_D\sigma_D^\top)^{-1}\varepsilon\big) - \big(\alpha + Q(\sigma_D\sigma_D^\top)^{-1}\big)\varepsilon.
\end{aligned}$$

Hence,

$$\|G(Q+\varepsilon)-G(Q)\| \le \|\varepsilon\|\big(2\|\alpha\| + 2\|(\sigma_D\sigma_D^\top)^{-1}\|\|Q\| + \|(\sigma_D\sigma_D^\top)^{-1}\|\|\varepsilon\|\big).$$

This implies that there exists a constant $C_G>0$ such that for all ε, $Q\in\mathbb{R}^{d\times d}$ with $\|\varepsilon\|\le C_e$ and $\|Q\|\le C_Q$ it holds

$$\|G(Q+\varepsilon)-G(Q)\| \le C_G\|\varepsilon\|. \tag{8.15}$$

Dynamics of $Q^{C,n}$. Next, we take a look at the dynamics of $Q^{C,n}$, i.e. of the covariance matrix corresponding to the investor who observes the stock returns and the opinions of the discrete expert. We know that at information dates t_k, $k=1,\dots,n$, we have the update formula

$$Q^{C,n}_{t_k} = \Gamma^{(n)}\big(Q^{C,n}_{t_k-} + \Gamma^{(n)}\big)^{-1} Q^{C,n}_{t_k-}.$$

Observe that

$$\Gamma^{(n)}\big(Q^{C,n}_{t_k-} + \Gamma^{(n)}\big)^{-1} = \big(I_d + Q^{C,n}_{t_k-}(\Gamma^{(n)})^{-1}\big)^{-1} = \big(I_d + \Delta_n Q^{C,n}_{t_k-}(\sigma_J\sigma_J^\top)^{-1}\big)^{-1}$$

which can be written as the Neumann series

$$\sum_{i=0}^{\infty}\big(-\Delta_n Q^{C,n}_{t_k-}(\sigma_J\sigma_J^\top)^{-1}\big)^i = I_d - \Delta_n Q^{C,n}_{t_k-}(\sigma_J\sigma_J^\top)^{-1} + \sum_{i=2}^{\infty}\big(-\Delta_n Q^{C,n}_{t_k-}(\sigma_J\sigma_J^\top)^{-1}\big)^i.$$

It follows that

$$Q^{C,n}_{t_k} = Q^{C,n}_{t_k-} - \Delta_n Q^{C,n}_{t_k-}(\sigma_J\sigma_J^\top)^{-1} Q^{C,n}_{t_k-} + \bar{R}^n \tag{8.16}$$

where $\|\bar{R}^n\| \le r\Delta_n^2$, since $Q^{C,n}_{t_k-}$ is bounded. Between information dates, the matrix $Q^{C,n}$ follows the dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t} Q^{C,n}_t = -\alpha Q^{C,n}_t - Q^{C,n}_t\alpha + \beta\beta^\top - Q^{C,n}_t(\sigma_R\sigma_R^\top)^{-1}Q^{C,n}_t$$

for $t\in[t_k,t_{k+1})$.

One time step for $Q^{C,n}$. In the following, we construct a formula that connects $Q^{C,n}_{t_{k+1}-}$ with $Q^{C,n}_{t_k-}$. Firstly, by making a Taylor expansion we see that

$$Q^{C,n}_{t_{k+1}-} = Q^{C,n}_{t_k} + \big(-\alpha Q^{C,n}_{t_k} - Q^{C,n}_{t_k}\alpha + \beta\beta^\top - Q^{C,n}_{t_k}(\sigma_R\sigma_R^\top)^{-1}Q^{C,n}_{t_k}\big)\Delta_n + L^n,$$

where $\|L^n\| \le C_L\Delta_n^2$. Now, when inserting the representation of $Q^{C,n}_{t_k}$ from (8.16) and rearranging terms we can conclude that

$$Q^{C,n}_{t_{k+1}-} = Q^{C,n}_{t_k-} + \Delta_n G(Q^{C,n}_{t_k-}) + R^n, \tag{8.17}$$

where R^n is a matrix with $\|R^n\| \le C_{\text{Taylor}}\Delta_n^2$ for $C_{\text{Taylor}}>0$.

Recursive formula for estimation error. For $k=0,\dots,n$, define $A_k = Q^{C,n}_{t_k-} - Q^D_{t_k}$ and $a_k = \|A_k\|$. Our aim is to find a recursive formula that yields an upper bound for these estimation errors. Let $k\ge 0$. Then we have by (8.17) that

$$a_{k+1} = \|A_{k+1}\| = \|Q^{C,n}_{t_{k+1}-} - Q^D_{t_{k+1}}\| = \|Q^{C,n}_{t_k-} + \Delta_n G(Q^{C,n}_{t_k-}) + R^n - Q^D_{t_{k+1}}\|.$$

Thus, by definition of A_k and $Q^{D,n}$ as given in (8.13),

$$\begin{aligned}
a_{k+1} &= \|(Q^D_{t_k} + A_k) + \Delta_n G(Q^D_{t_k} + A_k) + R^n - Q^D_{t_{k+1}}\| \\
&= \|Q^D_{t_k} + \Delta_n\big(G(Q^D_{t_k}) + G(Q^D_{t_k} + A_k) - G(Q^D_{t_k})\big) + A_k + R^n - Q^D_{t_{k+1}}\| \\
&= \|Q^{D,n}_{t_{k+1}-} + \Delta_n\big(G(Q^D_{t_k} + A_k) - G(Q^D_{t_k})\big) + A_k + R^n - Q^D_{t_{k+1}}\|.
\end{aligned}$$

Now, the estimations from (8.14), (8.15) and (8.17) yield

$$a_{k+1} \leq C_{\text{Euler}}\Delta_n^2 + \Delta_n C_G \|A_k\| + \|A_k\| + C_{\text{Taylor}}\Delta_n^2 = (1+\Delta_n C_G)a_k + (C_{\text{Euler}} + C_{\text{Taylor}})\Delta_n^2.$$

By a discrete version of Gronwall's Lemma, see Lemma A.1 in the appendix, this implies

$$a_k \leq \frac{\mathrm{e}^{C_G k \Delta_n} - 1}{C_G}(C_{\text{Euler}} + C_{\text{Taylor}})\Delta_n \leq \frac{\mathrm{e}^{C_G T} - 1}{C_G}(C_{\text{Euler}} + C_{\text{Taylor}})\Delta_n =: \tilde{C}\Delta_n.$$

Therefore, for all $k = 0, \dots, n$ we have

$$\|Q^{C,n}_{t_k-} - Q^D_{t_k}\| \leq \tilde{C}\Delta_n. \tag{8.18}$$

Difference of $Q_t^{C,n}$ and Q_t^D for arbitrary t. We now show that there exists some $K_1 > 0$ such that $\|Q^{C,n}_t - Q^D_t\| \leq K_1\Delta_n$ for all $t \in [0,T]$. Let $t \in [0,T]$ with $t \in [t_k, t_{k+1})$. We can write

$$Q^{C,n}_t - Q^D_t = (Q^{C,n}_t - Q^{C,n}_{t_k-}) + (Q^{C,n}_{t_k-} - Q^D_{t_k}) + (Q^D_{t_k} - Q^D_t),$$

and hence

$$\|Q^{C,n}_t - Q^D_t\| \leq \|Q^{C,n}_t - Q^{C,n}_{t_k-}\| + \|Q^{C,n}_{t_k-} - Q^D_{t_k}\| + \|Q^D_{t_k} - Q^D_t\|.$$

By (8.18), the second summand is bounded by $\tilde{C}\Delta_n$. We now take a look at the other two summands. By definition of $Q^{D,n}$ we can write the third summand as

$$\begin{aligned}\|Q^D_{t_k} - Q^D_t\| &= \|Q^{D,n}_t - G(Q^D_{t_k})(t-t_k) - Q^D_t\| \\ &\leq \|Q^{D,n}_t - Q^D_t\| + \Delta_n\|G(Q^D_{t_k})\| \\ &\leq C_{\text{Euler}}\Delta_n^2 + \Delta_n\|G(Q^D_{t_k})\|\end{aligned}$$

where the second inequality is due to (8.14). Since G and Q^D are continuous, the function $t \mapsto \|G(Q^D_t)\|$ is bounded by some $\tilde{C}_G$ on $[0,T]$. Hence,

$$\|Q^D_{t_k} - Q^D_t\| \leq C_{\text{Euler}}\Delta_n^2 + \tilde{C}_G\Delta_n.$$

For the first summand we observe that, like in (8.17), we get the representation

$$\|Q^{C,n}_t - Q^{C,n}_{t_k-}\| = \|(t-t_k)G(Q^{C,n}_{t_k-}) + R^n\|$$

for some matrix R^n with $\|R^n\| \leq C_{\text{Taylor}}(t-t_k)^2$. Then the right-hand side is bounded by $\Delta_n\|G(Q^{C,n}_{t_k-})\| + C_{\text{Taylor}}\Delta_n^2$. Also, we have

$$\|G(Q^{C,n}_{t_k-})\| = \|G(Q^D_{t_k} + Q^{C,n}_{t_k-} - Q^D_{t_k})\| \leq \|G(Q^D_{t_k})\| + C_G\|Q^{C,n}_{t_k-} - Q^D_{t_k}\|$$

by (8.15). Again by continuity, $\|G(Q^D_{t_k})\| \leq \tilde{C}_G$, and $\|Q^{C,n}_{t_k-} - Q^D_{t_k}\| \leq \tilde{C}\Delta_n$ by (8.18).

Putting these results together we obtain that there exists a constant $K_1 > 0$ such that

$$\|Q^{C,n}_t - Q^D_t\| \leq K_1\Delta_n$$

for all $t \in [0,T]$. $\square$

Proof of Theorem 8.6

We omit the superscript n at information dates $t_k^{(n)}$ for the sake of better readability. The idea of the proof is to find a recursion for

$$\mathbb{E}\Big[\big\|m^{C,n}_{t_k-} - m^D_{t_k}\big\|^2\Big]$$

and to apply the discrete version of Gronwall's Lemma from Lemma A.1 to derive an appropriate upper bound.

For the proof we introduce the notation

$$L_k^{(n)} := Q^{C,n}_{t_k-}\big(Q^{C,n}_{t_k-} + \Gamma^{(n)}\big)^{-1}\Gamma^{(n)}$$

for $k = 1, \ldots, n$. Then Lemma A.4 in particular implies that

$$\|Q^{C,n}_{t_k-} - L_k^{(n)}\| \leq \bar{C}\Delta_n$$

for some constant $\bar{C} > 0$.

Recursive formulas for $\boldsymbol{m^D}$ and $\boldsymbol{m^{C,n}}$. The representation of m^D via the stochastic differential equation in Lemma 6.2 yields the recursion

$$\begin{aligned} m^D_{t_{k+1}} &= \mathrm{e}^{-\alpha\Delta_n} m^D_{t_k} + (I_d - \mathrm{e}^{-\alpha\Delta_n})\delta + \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q^D_s (\sigma_R\sigma_R^\top)^{-1}\sigma_R \,\mathrm{d}V^{D,1}_s \\ &\quad + \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q^D_s (\sigma_J\sigma_J^\top)^{-1}\sigma_J \,\mathrm{d}V^{D,2}_s, \end{aligned} \tag{8.19}$$

where

$$\begin{aligned} \sigma_R \,\mathrm{d}V^{D,1}_t &= \mathrm{d}R_t - m^D_t \,\mathrm{d}t, \\ \sigma_J \,\mathrm{d}V^{D,2}_t &= \mathrm{d}J_t - m^D_t \,\mathrm{d}t, \end{aligned}$$

and $V^D = (V^{D,1}, V^{D,2})^\top$, the innovation process corresponding to the investor filtration $\mathcal{F}^D$, is an $(m+l)$-dimensional $\mathcal{F}^D$-Brownian motion. Similarly, we get for the conditional mean $m^{C,n}$ the recursion

$$m^{C,n}_{t_{k+1}-} = \mathrm{e}^{-\alpha\Delta_n} m^{C,n}_{t_k} + (I_d - \mathrm{e}^{-\alpha\Delta_n})\delta + \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q^{C,n}_s (\sigma_R\sigma_R^\top)^{-1}\sigma_R \,\mathrm{d}V^R_s, \tag{8.20}$$

where

$$\sigma_R \,\mathrm{d}V^R_t = \mathrm{d}R_t - m^{C,n}_t \,\mathrm{d}t,$$

and V^R, the innovation process corresponding to investor filtration $\mathcal{F}^{C,n}$, is an m-dimensional $\mathcal{F}^{C,n}$-Brownian motion. Furthermore, the update formula for $m^{C,n}$ yields

$$\begin{aligned} m^{C,n}_{t_k} &= m^{C,n}_{t_k-} + \big(I_d - \Gamma^{(n)}\big(Q^{C,n}_{t_k-} + \Gamma^{(n)}\big)^{-1}\big)\big(Z_k^{(n)} - m^{C,n}_{t_k-}\big) \\ &= m^{C,n}_{t_k-} + Q^{C,n}_{t_k-}\big(Q^{C,n}_{t_k-} + \Gamma^{(n)}\big)^{-1}\bigg(\mu_{t_k} + \frac{1}{\Delta_n}\sigma_J \int_{t_k}^{t_{k+1}} \mathrm{d}W^J_s - m^{C,n}_{t_k-}\bigg) \\ &= m^{C,n}_{t_k-} + \Delta_n L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}\bigg(\mu_{t_k} + \frac{1}{\Delta_n}\sigma_J \int_{t_k}^{t_{k+1}} \mathrm{d}W^J_s - m^{C,n}_{t_k-}\bigg). \end{aligned} \tag{8.21}$$

When looking at the difference between m^D and $m^{C,n}$ it is convenient to work with representations that use the same Brownian motions.

Relation between the innovation processes. Note that

$$\sigma_R \,\mathrm{d}V_t^{D,1} = \mathrm{d}R_t - m_t^D \,\mathrm{d}t = \sigma_R \,\mathrm{d}V_t^R + (m_t^{C,n} - m_t^D)\,\mathrm{d}t$$

and

$$\sigma_J \,\mathrm{d}V_t^{D,2} = \mathrm{d}J_t - m_t^D \,\mathrm{d}t = \sigma_J \,\mathrm{d}W_t^J + (\mu_t - m_t^D)\,\mathrm{d}t.$$

Using this connection between the innovation processes, we obtain from (8.19) that

$$\begin{aligned} m_{t_{k+1}}^D = \mathrm{e}^{-\alpha\Delta_n} m_{t_k}^D + (I_d - \mathrm{e}^{-\alpha\Delta_n})\delta + \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q_s^D (\sigma_R\sigma_R^\top)^{-1}\sigma_R \,\mathrm{d}V_s^R \\ + \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q_s^D (\sigma_R\sigma_R^\top)^{-1}(m_s^{C,n} - m_s^D)\,\mathrm{d}s \\ + \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q_s^D (\sigma_J\sigma_J^\top)^{-1}\sigma_J \,\mathrm{d}W_s^J \\ + \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q_s^D (\sigma_J\sigma_J^\top)^{-1}(\mu_s - m_s^D)\,\mathrm{d}s. \end{aligned} \tag{8.22}$$

Also, plugging (8.21) into (8.20) yields

$$\begin{aligned} m_{t_{k+1}-}^{C,n} = \mathrm{e}^{-\alpha\Delta_n} m_{t_k-}^{C,n} + (I_d - \mathrm{e}^{-\alpha\Delta_n})\delta + \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q_s^{C,n} (\sigma_R\sigma_R^\top)^{-1}\sigma_R \,\mathrm{d}V_s^R \\ + \mathrm{e}^{-\alpha\Delta_n} L_k^{(n)} (\sigma_J\sigma_J^\top)^{-1}\sigma_J \int_{t_k}^{t_{k+1}} \mathrm{d}W_s^J + \mathrm{e}^{-\alpha\Delta_n}\Delta_n L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}(\mu_{t_k} - m_{t_k-}^{C,n}). \end{aligned}$$

Splitting the difference of m^D and $m^{C,n}$ into summands. Combining (8.22) with the above representation of $m_{t_{k+1}-}^{C,n}$ yields after a slight rearrangement of terms

$$m_{t_{k+1}}^D - m_{t_{k+1}-}^{C,n} = A^n + B^n + C^n + D^n + E^n + F^n,$$

where

$$\begin{aligned} A^n &= \mathrm{e}^{-\alpha\Delta_n}(m_{t_k}^D - m_{t_k-}^{C,n}), \\ B^n &= \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} (Q_s^D - Q_s^{C,n})(\sigma_R\sigma_R^\top)^{-1}\sigma_R \,\mathrm{d}V_s^R, \\ C^n &= \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q_s^D (\sigma_R\sigma_R^\top)^{-1}(m_s^{C,n} - m_s^D)\,\mathrm{d}s, \\ D^n &= \int_{t_k}^{t_{k+1}} \Big(\mathrm{e}^{-\alpha(t_{k+1}-s)} Q_s^D (\sigma_J\sigma_J^\top)^{-1} - \mathrm{e}^{-\alpha\Delta_n} L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}\Big)\sigma_J \,\mathrm{d}W_s^J, \\ E^n &= \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q_s^D (\sigma_J\sigma_J^\top)^{-1}\mu_s \,\mathrm{d}s - \mathrm{e}^{-\alpha\Delta_n}\Delta_n L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}\mu_{t_k}, \\ F^n &= \mathrm{e}^{-\alpha\Delta_n}\Delta_n L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1} m_{t_k-}^{C,n} - \int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)} Q_s^D (\sigma_J\sigma_J^\top)^{-1} m_s^D \,\mathrm{d}s. \end{aligned}$$

Application of the discrete Gronwall Lemma. The idea is now to apply the discrete Gronwall Lemma from Lemma A.1 to the estimation

$$\begin{aligned}\mathbb{E}\Big[\big\|m^D_{t_{k+1}} - m^{C,n}_{t_{k+1}-}\big\|^2\Big] &= \mathbb{E}\Big[\big\|A^n + B^n + C^n + D^n + E^n + F^n\big\|^2\Big] \\ &\leq \mathbb{E}\Big[\big\|A^n\big\|^2\Big] + 5\,\mathbb{E}\Big[\big\|B^n\big\|^2 + \big\|C^n\big\|^2 + \big\|D^n\big\|^2 + \big\|E^n\big\|^2 + \big\|F^n\big\|^2\Big] \\ &\quad + 2\,\mathbb{E}\Big[(A^n)^\top(E^n + F^n)\Big].\end{aligned} \tag{8.23}$$

In the inequality we have used that $(a_1 + \dots + a_p)^2 \leq p(a_1^2 + \dots + a_p^2)$, and the fact that $B^n + C^n + D^n$ can be written as a sum of stochastic integrals over Brownian motions between t_k and t_{k+1}. Since $A^n = \mathrm{e}^{-\alpha\Delta_n}(m^D_{t_k} - m^{C,n}_{t_k-})$ is independent of these stochastic integrals, the term $\mathbb{E}[(A^n)^\top(B^n + C^n + D^n)]$ vanishes.

Finding upper estimates for the single summands. We now show how to find upper estimates for the single summands in the decomposition above. First of all,

$$\mathbb{E}\Big[\big\|A^n\big\|^2\Big] = \mathbb{E}\Big[\big\|\mathrm{e}^{-\alpha\Delta_n}(m^D_{t_k} - m^{C,n}_{t_k-})\big\|^2\Big] \leq \mathbb{E}\Big[\big\|m^D_{t_k} - m^{C,n}_{t_k-}\big\|^2\Big]$$

by properties of the spectral norm and positive definiteness of α. By using the multidimensional Itô isometry from Lemma A.3 we deduce

$$\begin{aligned}\mathbb{E}\Big[\big\|B^n\big\|^2\Big] &\leq C_{\mathrm{norm}}\,\mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} \|\mathrm{e}^{-\alpha(t_{k+1}-s)}(Q^D_s - Q^{C,n}_s)(\sigma_R\sigma_R^\top)^{-1}\sigma_R\|^2\,\mathrm{d}s\bigg] \\ &\leq C_{\mathrm{norm}}\|(\sigma_R\sigma_R^\top)^{-1}\sigma_R\|^2 \int_{t_k}^{t_{k+1}} \|Q^D_s - Q^{C,n}_s\|^2\,\mathrm{d}s \\ &\leq C_{\mathrm{norm}}\|(\sigma_R\sigma_R^\top)^{-1}\sigma_R\|^2 \int_{t_k}^{t_{k+1}} (K_1\Delta_n)^2\,\mathrm{d}s =: C_B\Delta_n^3.\end{aligned}$$

Note that $\|Q^D_s - Q^{C,n}_s\| \leq K_1\Delta_n$ by Theorem 8.5. Now for the term C^n we use the Cauchy–Schwarz inequality from Lemma A.2 to see that

$$\begin{aligned}\mathbb{E}\Big[\big\|C^n\big\|^2\Big] &= \mathbb{E}\bigg[\Big\|\int_{t_k}^{t_{k+1}} \mathrm{e}^{-\alpha(t_{k+1}-s)}Q^D_s(\sigma_R\sigma_R^\top)^{-1}(m^{C,n}_s - m^D_s)\,\mathrm{d}s\Big\|^2\bigg] \\ &\leq \Delta_n \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|\mathrm{e}^{-\alpha(t_{k+1}-s)}Q^D_s(\sigma_R\sigma_R^\top)^{-1}(m^{C,n}_s - m^D_s)\big\|^2\Big]\,\mathrm{d}s \\ &\leq \Delta_n C_Q^2\|(\sigma_R\sigma_R^\top)^{-1}\|^2 \int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|m^{C,n}_s - m^D_s\big\|^2\Big]\,\mathrm{d}s.\end{aligned}$$

We then apply the mean value theorem for estimating the integral to see that

$$\begin{aligned}\int_{t_k}^{t_{k+1}} \mathbb{E}\Big[\big\|m^{C,n}_s - m^D_s\big\|^2\Big]\,\mathrm{d}s &\leq \Delta_n\,\mathbb{E}\Big[\big\|m^{C,n}_{t_k} - m^D_{t_k}\big\|^2\Big] + C_{\mathrm{mvt}}\Delta_n^2 \\ &\leq \Delta_n\Big(2\,\mathbb{E}\Big[\big\|m^{C,n}_{t_k-} - m^D_{t_k}\big\|^2\Big] + 2\,\mathbb{E}\Big[\big\|m^{C,n}_{t_k} - m^{C,n}_{t_k-}\big\|^2\Big]\Big) + C_{\mathrm{mvt}}\Delta_n^2.\end{aligned}$$

The jump size of $m^{C,n}$ at an information date is bounded, hence all in all we obtain

$$\mathbb{E}\Big[\big\|C^n\big\|^2\Big] \le C_{C,1}\Delta_n^2\,\mathbb{E}\Big[\big\|m^{C,n}_{t_k-} - m^D_{t_k}\big\|^2\Big] + C_{C,2}\Delta_n^2$$

for constants $C_{C,1}$, $C_{C,2} > 0$.

For the term D^n we use again the multidimensional Itô isometry from Lemma A.3 and get

$$\begin{aligned}\mathbb{E}\Big[\big\|D^n\big\|^2\Big] &\le C_{\mathrm{norm}}\,\mathbb{E}\Big[\int_{t_k}^{t_{k+1}} \big\|\big(\mathrm{e}^{-\alpha(t_{k+1}-s)}Q^D_s - \mathrm{e}^{-\alpha\Delta_n}L_k^{(n)}\big)(\sigma_J\sigma_J^\top)^{-1}\sigma_J\big\|^2\,\mathrm{d}s\Big]\\ &\le C_{\mathrm{norm}}\|(\sigma_J\sigma_J^\top)^{-1}\sigma_J\|^2\int_{t_k}^{t_{k+1}} \big\|\mathrm{e}^{-\alpha(t_{k+1}-s)}Q^D_s - \mathrm{e}^{-\alpha\Delta_n}L_k^{(n)}\big\|^2\,\mathrm{d}s.\end{aligned}$$

For the integral above we first use a mean value theorem argument and then Lemma A.4 for the estimation of $\|Q^D_{t_k} - L_k^{(n)}\|^2$ to obtain

$$\begin{aligned}\int_{t_k}^{t_{k+1}} \big\|\mathrm{e}^{-\alpha(t_{k+1}-s)}Q^D_s - \mathrm{e}^{-\alpha\Delta_n}L_k^{(n)}\big\|^2\,\mathrm{d}s &\le \Delta_n\big\|\mathrm{e}^{-\alpha\Delta_n}Q^D_{t_k} - \mathrm{e}^{-\alpha\Delta_n}L_k^{(n)}\big\|^2 + C_{\mathrm{mvt}}\Delta_n^2\\ \le \Delta_n\|Q^D_{t_k} - L_k^{(n)}\|^2 + C_{\mathrm{mvt}}\Delta_n^2 &\le 2\Delta_n\big(\|Q^D_{t_k} - Q^{C,n}_{t_k-}\|^2 + \bar{C}^2\Delta_n^2\big) + C_{\mathrm{mvt}}\Delta_n^2.\end{aligned}$$

Putting these estimations together yields the existence of a constant $C_D > 0$ such that

$$\mathbb{E}\Big[\big\|D^n\big\|^2\Big] \le C_D\Delta_n^2.$$

By writing the next summand E^n as one integral, we can again apply the Cauchy–Schwarz inequality from Lemma A.2 and get

$$\begin{aligned}\mathbb{E}\Big[\big\|E^n\big\|^2\Big] &= \mathbb{E}\Big[\Big\|\int_{t_k}^{t_{k+1}}\Big(\mathrm{e}^{-\alpha(t_{k+1}-s)}Q^D_s(\sigma_J\sigma_J^\top)^{-1}\mu_s - \mathrm{e}^{-\alpha\Delta_n}L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}\mu_{t_k}\Big)\mathrm{d}s\Big\|^2\Big]\\ &\le \Delta_n\int_{t_k}^{t_{k+1}}\mathbb{E}\Big[\Big\|\mathrm{e}^{-\alpha(t_{k+1}-s)}Q^D_s(\sigma_J\sigma_J^\top)^{-1}\mu_s - \mathrm{e}^{-\alpha\Delta_n}L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}\mu_{t_k}\Big\|^2\Big]\mathrm{d}s.\end{aligned}$$

When using again the mean value theorem and the same argumentation as before we see that the integral is bounded by

$$\begin{aligned}&\Delta_n\,\mathbb{E}\Big[\Big\|\mathrm{e}^{-\alpha\Delta_n}\big(Q^D_{t_k} - L_k^{(n)}\big)(\sigma_J\sigma_J^\top)^{-1}\mu_{t_k}\Big\|^2\Big] + C_{\mathrm{mvt}}\Delta_n^2\\ &\le \Delta_n\big\|Q^D_{t_k} - L_k^{(n)}\big\|^2\|(\sigma_J\sigma_J^\top)^{-1}\|^2\,\mathbb{E}[\|\mu_{t_k}\|^2] + C_{\mathrm{mvt}}\Delta_n^2\\ &\le \Delta_n C_\mu\|(\sigma_J\sigma_J^\top)^{-1}\|^2\Big(2\|Q^D_{t_k} - Q^{C,n}_{t_k-}\|^2 + 2\bar{C}^2\Delta_n^2\Big) + C_{\mathrm{mvt}}\Delta_n^2.\end{aligned}$$

In conclusion, we have a constant $C_E > 0$ with

$$\mathbb{E}\Big[\big\|E^n\big\|^2\Big] \le C_E\Delta_n^3.$$

In a similar way, F^n can be treated. By first writing F^n as a single integral and applying the Cauchy–Schwarz inequality from Lemma A.2 as well as the mean value theorem we get

$$\begin{aligned}\mathbb{E}\Big[\big\|F^n\big\|^2\Big] &= \mathbb{E}\Big[\Big\|\int_{t_k}^{t_{k+1}}\Big(\mathrm{e}^{-\alpha\Delta_n}L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}m^{C,n}_{t_k-} - \mathrm{e}^{-\alpha(t_{k+1}-s)}Q^D_s(\sigma_J\sigma_J^\top)^{-1}m^D_s\Big)\mathrm{d}s\Big\|^2\Big]\\ &\le \Delta_n\int_{t_k}^{t_{k+1}}\mathbb{E}\Big[\Big\|\mathrm{e}^{-\alpha\Delta_n}L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}m^{C,n}_{t_k-} - \mathrm{e}^{-\alpha(t_{k+1}-s)}Q^D_s(\sigma_J\sigma_J^\top)^{-1}m^D_s\Big\|^2\Big]\mathrm{d}s\\ &\le \Delta_n^2\,\mathbb{E}\Big[\Big\|\mathrm{e}^{-\alpha\Delta_n}\Big(L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}m^{C,n}_{t_k-} - Q^D_{t_k}(\sigma_J\sigma_J^\top)^{-1}m^D_{t_k}\Big)\Big\|^2\Big] + C_{\mathrm{mvt}}\Delta_n^3.\end{aligned}$$

The expectation above is bounded by

$$\begin{aligned}
&2\,\mathbb{E}\big[\big\|\big(L_k^{(n)}-Q_{t_k}^D\big)(\sigma_J\sigma_J^\top)^{-1}m_{t_k-}^{C,n}\big\|^2\big]+2\,\mathbb{E}\big[\big\|Q_{t_k}^D(\sigma_J\sigma_J^\top)^{-1}\big(m_{t_k-}^{C,n}-m_{t_k}^D\big)\big\|^2\big]\\
&\le 2\|(\sigma_J\sigma_J^\top)^{-1}\|^2\,\mathbb{E}\big[\|m_{t_k-}^{C,n}\|^2\big]\big\|L_k^{(n)}-Q_{t_k}^D\big\|^2+2C_Q^2\|(\sigma_J\sigma_J^\top)^{-1}\|^2\,\mathbb{E}\big[\|m_{t_k-}^{C,n}-m_{t_k}^D\|^2\big].
\end{aligned}$$

By the same reasons as in the calculations above we obtain all in all that there exist constants $C_{F,1}$ and $C_{F,2}>0$ such that

$$\mathbb{E}\Big[\big\|F^n\big\|^2\Big]\le C_{F,1}\Delta_n^2\,\mathbb{E}\big[\|m_{t_k-}^{C,n}-m_{t_k}^D\|^2\big]+C_{F,2}\Delta_n^3.$$

We have now found upper bounds for all quadratic terms in (8.23). Only the mixed terms $(A^n)^\top E^n$ and $(A^n)^\top F^n$ remain to be considered. Firstly, we again rewrite E^n as one integral

$$E^n=\int_{t_k}^{t_{k+1}}\Big(\mathrm{e}^{-\alpha(t_{k+1}-s)}Q_s^D(\sigma_J\sigma_J^\top)^{-1}\mu_s-\mathrm{e}^{-\alpha\Delta_n}L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}\mu_{t_k}\Big)\mathrm{d}s.$$

We see that

$$\begin{aligned}
&\mathbb{E}\big[(A^n)^\top E^n\big]\\
&=\int_{t_k}^{t_{k+1}}\mathbb{E}\big[(m_{t_k}^D-m_{t_k-}^{C,n})^\top\mathrm{e}^{-\alpha\Delta_n}\big(\mathrm{e}^{-\alpha(t_{k+1}-s)}Q_s^D(\sigma_J\sigma_J^\top)^{-1}\mu_s-\mathrm{e}^{-\alpha\Delta_n}L_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}\mu_{t_k}\big)\big]\,\mathrm{d}s\\
&=\int_{t_k}^{t_{k+1}}\mathbb{E}\big[(m_{t_k}^D-m_{t_k-}^{C,n})^\top\mathrm{e}^{-\alpha\Delta_n}\mathrm{e}^{-\alpha(t_{k+1}-s)}Q_s^D(\sigma_J\sigma_J^\top)^{-1}\mu_s\big]\,\mathrm{d}s\\
&\qquad-\mathbb{E}\big[(m_{t_k}^D-m_{t_k-}^{C,n})^\top\mathrm{e}^{-2\alpha\Delta_n}\Delta_nL_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}\mu_{t_k}\big].
\end{aligned}$$

By using the mean value theorem and sublinearity of the spectral norm we obtain

$$\begin{aligned}
\big|\mathbb{E}\big[(A^n)^\top E^n\big]\big|&\le\Big|\Delta_n\,\mathbb{E}\big[(m_{t_k}^D-m_{t_k-}^{C,n})^\top\mathrm{e}^{-2\alpha\Delta_n}Q_{t_k}^D(\sigma_J\sigma_J^\top)^{-1}\mu_{t_k}\big]\\
&\qquad-\mathbb{E}\big[(m_{t_k}^D-m_{t_k-}^{C,n})^\top\mathrm{e}^{-2\alpha\Delta_n}\Delta_nL_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}\mu_{t_k}\big]\Big|+C_{\mathrm{mvt}}\Delta_n^2\\
&=\Delta_n\Big|\mathbb{E}\Big[(m_{t_k}^D-m_{t_k-}^{C,n})^\top\mathrm{e}^{-2\alpha\Delta_n}\big(Q_{t_k}^D-L_k^{(n)}\big)(\sigma_J\sigma_J^\top)^{-1}\mu_{t_k}\Big]\Big|+C_{\mathrm{mvt}}\Delta_n^2\\
&\le\Delta_n\big\|(\sigma_J\sigma_J^\top)^{-1}\big\|\big\|Q_{t_k}^D-L_k^{(n)}\big\|\,\mathbb{E}\Big[\big\|m_{t_k}^D-m_{t_k-}^{C,n}\big\|\big\|\mu_{t_k}\big\|\Big]+C_{\mathrm{mvt}}\Delta_n^2\\
&\le C_{A,E}\Delta_n^2.
\end{aligned}$$

The last inequality is due to boundedness of $\mathbb{E}[\|m_{t_k}^D-m_{t_k-}^{C,n}\|\|\mu_{t_k}\|]$ together with the fact that $\|Q_{t_k}^D-L_k^{(n)}\|$ is bounded by a constant times Δ_n, see Lemma A.4.

The mixed term $(A^n)^\top F^n$ can be handled in a similar way. It holds that

$$\begin{aligned}
(A^n)^\top F^n&=(m_{t_k}^D-m_{t_k-}^{C,n})^\top\mathrm{e}^{-2\alpha\Delta_n}\Delta_nL_k^{(n)}(\sigma_J\sigma_J^\top)^{-1}m_{t_k-}^{C,n}\\
&\qquad-\int_{t_k}^{t_{k+1}}(m_{t_k}^D-m_{t_k-}^{C,n})^\top\mathrm{e}^{-\alpha\Delta_n}\mathrm{e}^{-\alpha(t_{k+1}-s)}Q_s^D(\sigma_J\sigma_J^\top)^{-1}m_s^D\,\mathrm{d}s
\end{aligned}$$

and hence by another application of the mean value theorem

$$\begin{aligned}
&\big|\mathbb{E}\big[(A^n)^\top F^n\big]\big| \\
&\le \bigg|\mathbb{E}\Big[(m^D_{t_k} - m^{C,n}_{t_k-})^\top \mathrm{e}^{-2\alpha\Delta_n}\Delta_n L^{(n)}_k (\sigma_J\sigma_J^\top)^{-1} m^{C,n}_{t_k-}\Big] \\
&\quad - \Delta_n\,\mathbb{E}\big[(m^D_{t_k} - m^{C,n}_{t_k-})^\top \mathrm{e}^{-2\alpha\Delta_n} Q^D_{t_k}(\sigma_J\sigma_J^\top)^{-1} m^D_{t_k}\big]\bigg| + C_{\mathrm{mvt}}\Delta_n^2 \\
&= \Delta_n\bigg|\mathbb{E}\Big[(m^D_{t_k} - m^{C,n}_{t_k-})^\top \mathrm{e}^{-2\alpha\Delta_n}\Big(L^{(n)}_k(\sigma_J\sigma_J^\top)^{-1} m^{C,n}_{t_k-} - Q^D_{t_k}(\sigma_J\sigma_J^\top)^{-1} m^D_{t_k}\Big)\Big]\bigg| + C_{\mathrm{mvt}}\Delta_n^2.
\end{aligned}$$

The absolute value of the expectation is split into two summands as

$$\begin{aligned}
&\Big|\mathbb{E}\Big[(m^D_{t_k} - m^{C,n}_{t_k-})^\top \mathrm{e}^{-2\alpha\Delta_n}\Big(L^{(n)}_k(\sigma_J\sigma_J^\top)^{-1} m^{C,n}_{t_k-} - Q^D_{t_k}(\sigma_J\sigma_J^\top)^{-1} m^D_{t_k}\Big)\Big]\Big| \\
&\le \Big|\mathbb{E}\Big[(m^D_{t_k} - m^{C,n}_{t_k-})^\top \mathrm{e}^{-2\alpha\Delta_n}\big(L^{(n)}_k - Q^D_{t_k}\big)(\sigma_J\sigma_J^\top)^{-1} m^{C,n}_{t_k-}\Big]\Big| \\
&\quad + \Big|\mathbb{E}\Big[(m^D_{t_k} - m^{C,n}_{t_k-})^\top \mathrm{e}^{-2\alpha\Delta_n} Q^D_{t_k}(\sigma_J\sigma_J^\top)^{-1}\big(m^{C,n}_{t_k-} - m^D_{t_k}\big)\Big]\Big| \\
&\le \big\|(\sigma_J\sigma_J^\top)^{-1}\big\|\bigg(\mathbb{E}\Big[\big\|m^D_{t_k} - m^{C,n}_{t_k-}\big\|\big\|m^{C,n}_{t_k-}\big\|\Big]\big\|L^{(n)}_k - Q^D_{t_k}\big\| + C_Q\,\mathbb{E}\Big[\big\|m^D_{t_k} - m^{C,n}_{t_k-}\big\|^2\Big]\bigg).
\end{aligned}$$

From the same argumentations as above we deduce that there exist constants $C_{A,F,1}$ and $C_{A,F,2} > 0$ with

$$\big|\mathbb{E}\big[(A^n)^\top F^n\big]\big| \le C_{A,F,1}\Delta_n\,\mathbb{E}\big[\big\|m^D_{t_k} - m^{C,n}_{t_k-}\big\|^2\big] + C_{A,F,2}\Delta_n^2.$$

Conclusion with discrete Gronwall Lemma. Now we plug all these upper bounds into (8.23) and obtain that there exist constants $L_1, L_2 > 0$ such that

$$\mathbb{E}\Big[\big\|m^D_{t_{k+1}} - m^{C,n}_{t_{k+1}-}\big\|^2\Big] \le \big(1 + L_1\Delta_n\big)\,\mathbb{E}\Big[\big\|m^D_{t_k} - m^{C,n}_{t_k-}\big\|^2\Big] + L_2\Delta_n^2.$$

Setting $a_k := \mathbb{E}\Big[\big\|m^D_{t_k} - m^{C,n}_{t_k-}\big\|^2\Big]$ in the discrete version of Gronwall's Lemma, see Lemma A.1, we can conclude that

$$\mathbb{E}\Big[\big\|m^D_{t_k} - m^{C,n}_{t_k-}\big\|^2\Big] \le \frac{\mathrm{e}^{L_1 T} - 1}{L_1} L_2\Delta_n =: \tilde{C}\Delta_n$$

which proves the claim for $t = t_k$. To find an upper bound that is valid for arbitrary time $t \in [0, T]$ with $t \in [t_k, t_{k+1})$, we observe that

$$\begin{aligned}
\mathbb{E}\Big[\big\|m^D_t - m^{C,n}_t\big\|^2\Big] &= \mathbb{E}\Big[\big\|m^D_t - m^D_{t_k} + m^D_{t_k} - m^{C,n}_{t_k-} + m^{C,n}_{t_k-} - m^{C,n}_t\big\|^2\Big] \\
&\le 3\Big(\mathbb{E}\Big[\big\|m^D_t - m^D_{t_k}\big\|^2\Big] + \mathbb{E}\Big[\big\|m^D_{t_k} - m^{C,n}_{t_k-}\big\|^2\Big] + \mathbb{E}\Big[\big\|m^{C,n}_{t_k-} - m^{C,n}_t\big\|^2\Big]\Big).
\end{aligned}$$

The first summand is bounded by a constant times Δ_n which can be seen from the representation in Lemma 6.2. From (8.21) we can deduce the same for the third summand. Hence, all in all there exists a constant $K_2 > 0$ such that

$$\mathbb{E}\Big[\big\|m^{C,n}_t - m^D_t\big\|^2\Big] \le K_2\Delta_n$$

which proves the claim of the theorem. □

Proof of Theorem 8.15

First of all, we use the representations from Proposition 8.14 to see that

$$Q_t^{C,\lambda} - Q_t^D = \int_0^t \Big(L(Q_s^{C,\lambda}) - L(Q_s^D) - \lambda Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1} Q_{s-}^{C,\lambda} + Q_s^D(\sigma_J\sigma_J^\top)^{-1}Q_s^D \Big) \mathrm{d}s$$
$$+ \int_0^t \int_{\mathbb{R}^d} -Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1} Q_{s-}^{C,\lambda}\, \tilde{N}(\mathrm{d}s, \mathrm{d}u).$$

Denote the first integral by A_t^λ and the one with respect to the compensated measure by X_t^λ. Now for $r \in [0,T]$ let

$$u_r^\lambda := \mathbb{E}\Big[\sup_{t \le r} \|Q_t^{C,\lambda} - Q_t^D\|^2\Big] = \mathbb{E}\Big[\sup_{t \le r} \|A_t^\lambda + X_t^\lambda\|^2\Big].$$

In this notation we want to show that $u_T^\lambda \le \frac{K_3}{\lambda}$ for some constant $K_3 > 0$. It holds that

$$u_r^\lambda \le 2\,\mathbb{E}\Big[\sup_{t \le r} \|A_t^\lambda\|^2\Big] + 2\,\mathbb{E}\Big[\sup_{t \le r} \|X_t^\lambda\|^2\Big]. \tag{8.24}$$

In the following, we find upper bounds for both summands.

Estimate for the martingale term X^λ. Firstly, note that every component of the matrix-valued process $(X_t^\lambda)_{t\ge 0}$ is a martingale since we integrate with respect to the compensated measure $\tilde{N}$. For being able to use Lemma A.6 we replace the spectral norm with the Frobenius norm. By equivalence of norms there is a constant $C_{\mathrm{norm}} > 0$ such that

$$\begin{aligned}\mathbb{E}\Big[\sup_{t \le r} \|X_t^\lambda\|^2\Big] &\le C_{\mathrm{norm}}\,\mathbb{E}\Big[\sup_{t \le r} \|X_t^\lambda\|_F^2\Big] = C_{\mathrm{norm}}\,\mathbb{E}\Big[\sup_{t \le r} \sum_{i,j=1}^d (X_t^\lambda(i,j))^2\Big] \\ &\le C_{\mathrm{norm}} \sum_{i,j=1}^d \mathbb{E}\Big[\sup_{t \le r} (X_t^\lambda(i,j))^2\Big] \le C_{\mathrm{norm}} \sum_{i,j=1}^d 4\,\mathbb{E}\Big[(X_r^\lambda(i,j))^2\Big].\end{aligned} \tag{8.25}$$

The last inequality follows from Doob's inequality for martingales. Next, we can apply Lemma A.6 to the definition of X^λ and get

$$\begin{aligned}\mathbb{E}\Big[(X_r^\lambda(i,j))^2\Big] &= \mathbb{E}\Big[\int_0^r \int_{\mathbb{R}^d} \Big(\big(-Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1} Q_{s-}^{C,\lambda}\big)(i,j)\Big)^2 \lambda\varphi(u)\,\mathrm{d}u\,\mathrm{d}s\Big] \\ &= \lambda\,\mathbb{E}\Big[\int_0^r \Big(\big(-Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1} Q_{s-}^{C,\lambda}\big)(i,j)\Big)^2 \mathrm{d}s\Big],\end{aligned}$$

since the integrand does not depend on u, and φ is a density. Plugging back into (8.25), we get

$$\begin{aligned}\mathbb{E}\Big[\sup_{t \le r} \|X_t^\lambda\|^2\Big] &\le 4C_{\mathrm{norm}}\lambda \int_0^r \mathbb{E}\Big[\|-Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1} Q_{s-}^{C,\lambda}\|_F^2\Big]\mathrm{d}s \\ &\le 4C_{\mathrm{norm}}^2\lambda \int_0^r \mathbb{E}\Big[\|-Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1} Q_{s-}^{C,\lambda}\|^2\Big]\mathrm{d}s,\end{aligned} \tag{8.26}$$

again by equivalence of norms. We now take a closer look at the remaining expectation term in the integral. Since the spectral norm of the matrices $Q^{C,\lambda}$ is bounded by C_Q, see Lemma 7.3, we obtain

$$\begin{aligned}\mathbb{E}\Big[\|-Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}\|^2\Big] &\le C_Q^4\,\mathbb{E}\Big[\|(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}\|^2\Big]\\ &= C_Q^4\,\mathbb{E}\Big[\big(\lambda_{\min}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)\big)^{-2}\Big] \le C_Q^4\,\mathbb{E}\Big[\big(\lambda_{\min}(\lambda\sigma_J\sigma_J^\top)\big)^{-2}\Big]\\ &= \frac{C_Q^4}{\lambda^2}\|(\sigma_J\sigma_J^\top)^{-1}\|^2.\end{aligned}\tag{8.27}$$

When reinserting this upper bound into (8.26), we can conclude that

$$\mathbb{E}\Big[\sup_{t\le r}\|X^\lambda_t\|^2\Big] \le \frac{4C^2_{\mathrm{norm}}C_Q^4\|(\sigma_J\sigma_J^\top)^{-1}\|^2 r}{\lambda} \le \frac{4C^2_{\mathrm{norm}}C_Q^4\|(\sigma_J\sigma_J^\top)^{-1}\|^2 T}{\lambda}.\tag{8.28}$$

Estimate for the finite variation term A^λ. Next, we address the other summand in (8.24). Note that when shortly writing g for the integrand of A^λ_t we get

$$\sup_{t\le r}\|A^\lambda_t\|^2 = \sup_{t\le r}\left\|\int_0^t g(s)\,\mathrm{d}s\right\|^2 \le \sup_{t\le r} t\int_0^t\|g(s)\|^2\,\mathrm{d}s \le r\int_0^r\|g(s)\|^2\,\mathrm{d}s\tag{8.29}$$

by the Cauchy–Schwarz inequality in Lemma A.2. We now address the integrand of A^λ. We can write

$$\begin{aligned}g(s) = &-\alpha(Q^{C,\lambda}_s-Q^D_s)-(Q^{C,\lambda}_s-Q^D_s)\alpha-Q^{C,\lambda}_s(\sigma_R\sigma_R^\top)^{-1}(Q^{C,\lambda}_s-Q^D_s)\\ &-(Q^{C,\lambda}_s-Q^D_s)(\sigma_R\sigma_R^\top)^{-1}Q^D_s-\lambda Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}+Q^D_s(\sigma_J\sigma_J^\top)^{-1}Q^D_s\end{aligned}$$

and hence

$$\begin{aligned}\|g(s)\| \le\;& \|Q^{C,\lambda}_s-Q^D_s\|\big(2\|\alpha\|+2C_Q\|(\sigma_R\sigma_R^\top)^{-1}\|\big)\\ &+\|\lambda Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}-Q^D_s(\sigma_J\sigma_J^\top)^{-1}Q^D_s\|.\end{aligned}$$

So by (8.29) we obtain

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\le r}\|A^\lambda_t\|^2\Big] &\le r\int_0^r 2\big(2\|\alpha\|+2C_Q\|(\sigma_R\sigma_R^\top)^{-1}\|\big)^2\,\mathbb{E}\big[\|Q^{C,\lambda}_s-Q^D_s\|^2\big]\,\mathrm{d}s\\ &\quad+r\int_0^r 2\,\mathbb{E}\big[\|\lambda Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}-Q^D_s(\sigma_J\sigma_J^\top)^{-1}Q^D_s\|^2\big]\,\mathrm{d}s\\ &\le 2T\big(2\|\alpha\|+2C_Q\|(\sigma_R\sigma_R^\top)^{-1}\|\big)^2\int_0^r u^\lambda_s\,\mathrm{d}s\\ &\quad+2T\int_0^r\mathbb{E}\big[\|\lambda Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}-Q^D_s(\sigma_J\sigma_J^\top)^{-1}Q^D_s\|^2\big]\,\mathrm{d}s.\end{aligned}$$

The first term is equal to $\int_0^r u^\lambda_s\,\mathrm{d}s$, multiplied by a constant. We analyze the second summand in more detail. For that purpose, we decompose

$$\begin{aligned}&\lambda Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}-Q^D_s(\sigma_J\sigma_J^\top)^{-1}Q^D_s\\ &=\lambda Q^{C,\lambda}_{s-}(Q^{C,\lambda}_{s-}+\lambda\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}-Q^{C,\lambda}_{s-}(\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}\\ &\quad+Q^{C,\lambda}_{s-}(\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_{s-}-Q^{C,\lambda}_s(\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_s\\ &\quad+Q^{C,\lambda}_s(\sigma_J\sigma_J^\top)^{-1}Q^{C,\lambda}_s-Q^D_s(\sigma_J\sigma_J^\top)^{-1}Q^D_s\end{aligned}\tag{8.30}$$

and find upper bounds for the three summands. For the third summand we observe that

$$\begin{aligned}
&\mathbb{E}\big[\|Q_s^{C,\lambda}(\sigma_J\sigma_J^\top)^{-1}Q_s^{C,\lambda} - Q_s^D(\sigma_J\sigma_J^\top)^{-1}Q_s^D\|^2\big] \\
&= \mathbb{E}\big[\|Q_s^{C,\lambda}(\sigma_J\sigma_J^\top)^{-1}(Q_s^{C,\lambda} - Q_s^D) + (Q_s^{C,\lambda} - Q_s^D)(\sigma_J\sigma_J^\top)^{-1}Q_s^D\|^2\big] \\
&\leq \big(2C_Q\|(\sigma_J\sigma_J^\top)^{-1}\|\big)^2\,\mathbb{E}\big[\|Q_s^{C,\lambda} - Q_s^D\|^2\big] \leq \big(2C_Q\|(\sigma_J\sigma_J^\top)^{-1}\|\big)^2 u_s^\lambda.
\end{aligned} \tag{8.31}$$

We find an upper bound for the second summand in (8.30) by

$$\begin{aligned}
&\mathbb{E}\big[\|Q_{s-}^{C,\lambda}(\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda} - Q_s^{C,\lambda}(\sigma_J\sigma_J^\top)^{-1}Q_s^{C,\lambda}\|^2\big] \\
&= \mathbb{E}\big[\|Q_{s-}^{C,\lambda}(\sigma_J\sigma_J^\top)^{-1}(Q_{s-}^{C,\lambda} - Q_s^{C,\lambda}) + (Q_{s-}^{C,\lambda} - Q_s^{C,\lambda})(\sigma_J\sigma_J^\top)^{-1}Q_s^{C,\lambda}\|^2\big] \\
&\leq \big(2C_Q\|(\sigma_J\sigma_J^\top)^{-1}\|\big)^2\,\mathbb{E}\big[\|Q_{s-}^{C,\lambda} - Q_s^{C,\lambda}\|^2\big] \\
&\leq \big(2C_Q\|(\sigma_J\sigma_J^\top)^{-1}\|\big)^2\,\mathbb{E}\big[\|Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda}\|^2\big] \\
&\leq \big(2C_Q\|(\sigma_J\sigma_J^\top)^{-1}\|\big)^2\frac{C_Q^4}{\lambda^2}\|(\sigma_J\sigma_J^\top)^{-1}\|^2 = \frac{4C_Q^6\|(\sigma_J\sigma_J^\top)^{-1}\|^4}{\lambda^2}
\end{aligned} \tag{8.32}$$

as in (8.27). The first summand in (8.30) can be bounded by

$$\begin{aligned}
&\mathbb{E}\Big[\|\lambda Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda} - Q_{s-}^{C,\lambda}(\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda}\|^2\Big] \\
&= \mathbb{E}\Big[\|\lambda Q_{s-}^{C,\lambda}\Big((Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1} - (\lambda\sigma_J\sigma_J^\top)^{-1}\Big)Q_{s-}^{C,\lambda}\|^2\Big] \\
&= \mathbb{E}\Big[\|\lambda Q_{s-}^{C,\lambda}\Big((Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1}(\lambda\sigma_J\sigma_J^\top) - I_d\Big)(\lambda\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda}\|^2\Big] \\
&= \mathbb{E}\Big[\|Q_{s-}^{C,\lambda}\Big((Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1}(\lambda\sigma_J\sigma_J^\top - Q_{s-}^{C,\lambda} - \lambda\sigma_J\sigma_J^\top)\Big)(\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda}\|^2\Big] \\
&= \mathbb{E}\Big[\|-Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda}(\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda}\|^2\Big] \\
&\leq C_Q^2\|(\sigma_J\sigma_J^\top)^{-1}\|^2\frac{C_Q^4}{\lambda^2}\|(\sigma_J\sigma_J^\top)^{-1}\|^2 = \frac{C_Q^6\|(\sigma_J\sigma_J^\top)^{-1}\|^4}{\lambda^2},
\end{aligned} \tag{8.33}$$

again as in (8.27). We now use (8.31), (8.32) and (8.33) as well as (8.30) and obtain

$$\begin{aligned}
&\mathbb{E}\big[\|\lambda Q_{s-}^{C,\lambda}(Q_{s-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1}Q_{s-}^{C,\lambda} - Q_s^D(\sigma_J\sigma_J^\top)^{-1}Q_s^D\|^2\big] \\
&\leq 3\frac{C_Q^6\|(\sigma_J\sigma_J^\top)^{-1}\|^4}{\lambda^2} + 3\frac{4C_Q^6\|(\sigma_J\sigma_J^\top)^{-1}\|^4}{\lambda^2} + 3\big(2C_Q\|(\sigma_J\sigma_J^\top)^{-1}\|\big)^2 u_s^\lambda \\
&= \frac{15C_Q^6\|(\sigma_J\sigma_J^\top)^{-1}\|^4}{\lambda^2} + 12\big(C_Q\|(\sigma_J\sigma_J^\top)^{-1}\|\big)^2 u_s^\lambda.
\end{aligned}$$

Hence we can write

$$\begin{aligned}
\mathbb{E}\Big[\sup_{t\leq r}\|A_t^\lambda\|^2\Big] &\leq 2T\big(2\|\alpha\| + 2C_Q\|(\sigma_R\sigma_R^\top)^{-1}\|\big)^2\int_0^r u_s^\lambda\,\mathrm{d}s \\
&\quad + 2T\int_0^r\Big(\frac{15C_Q^6\|(\sigma_J\sigma_J^\top)^{-1}\|^4}{\lambda^2} + 12\big(C_Q\|(\sigma_J\sigma_J^\top)^{-1}\|\big)^2 u_s^\lambda\Big)\mathrm{d}s \\
&\leq 2T\Big(\big(2\|\alpha\| + 2C_Q\|(\sigma_R\sigma_R^\top)^{-1}\|\big)^2 + 12\big(C_Q\|(\sigma_J\sigma_J^\top)^{-1}\|\big)^2\Big)\int_0^r u_s^\lambda\,\mathrm{d}s \\
&\quad + 2T^2\frac{15C_Q^6\|(\sigma_J\sigma_J^\top)^{-1}\|^4}{\lambda^2}.
\end{aligned} \tag{8.34}$$

Conclusion with Gronwall's Lemma. We have found upper bounds for

$$\mathbb{E}\Big[\sup_{t\le r}\|X_t^\lambda\|^2\Big] \quad\text{and}\quad \mathbb{E}\Big[\sup_{t\le r}\|A_t^\lambda\|^2\Big]$$

in (8.28) and (8.34), respectively. Plugging into (8.24) yields

$$\begin{aligned} u_r^\lambda &\le 4T\Big(\big(2\|\alpha\| + 2C_Q\|(\sigma_R\sigma_R^\top)^{-1}\|\big)^2 + 12\big(C_Q\|(\sigma_J\sigma_J^\top)^{-1}\|\big)^2\Big)\int_0^r u_s^\lambda\,\mathrm{d}s \\ &\quad + 4T^2\frac{15C_Q^6\|(\sigma_J\sigma_J^\top)^{-1}\|^4}{\lambda^2} + \frac{8C_{\mathrm{norm}}^2C_Q^4\|(\sigma_J\sigma_J^\top)^{-1}\|^2T}{\lambda} \\ &\le \frac{C_1}{\lambda} + C_2\int_0^r u_s^\lambda\,\mathrm{d}s \end{aligned}$$

for all $\lambda \ge \lambda_0$ for some $\lambda_0 > 0$ and constants $C_1,\ C_2 > 0$. By Gronwall's Lemma, see Lemma A.5, it follows

$$u_r^\lambda \le \frac{C_1}{\lambda}\mathrm{e}^{C_2 r},$$

and in particular

$$\mathbb{E}\Big[\sup_{t\le T}\|Q_t^{C,\lambda} - Q_t^D\|^2\Big] = u_T^\lambda \le \frac{C_1}{\lambda}\mathrm{e}^{C_2T} = \frac{K_3}{\lambda}$$

where $K_3 = C_1\mathrm{e}^{C_2T} > 0$. □

Proof of Theorem 8.16

Throughout the proof, we omit the superscript λ at time points $T_k^{(\lambda)}$ and at the Poisson process $(N_t^{(\lambda)})_{t\ge0}$ for better readability. The proof uses again Gronwall's Lemma, see Lemma A.5. For this purpose, define

$$v_t^\lambda := \mathbb{E}\Big[\big\|m_t^{C,\lambda} - m_t^D\big\|^2\Big]$$

for $t\in[0,T]$. The filtering equations from Lemma 6.4 yield that we can write the conditional mean $m^{C,\lambda}$ as

$$m_t^{C,\lambda} = \int_0^t \alpha(\delta - m_s^{C,\lambda})\,\mathrm{d}s + \int_0^t Q_s^{C,\lambda}(\sigma_R\sigma_R^\top)^{-1}\sigma_R\,\mathrm{d}V_s^R + \sum_{k=1}^{N_t}\frac{1}{\lambda}P_k^\lambda\big(Z_k^{(\lambda)} - m_{T_k-}^{C,\lambda}\big), \quad (8.35)$$

where $\mathrm{d}R_s - m_s^{C,\lambda}\,\mathrm{d}s = \sigma_R\,\mathrm{d}V_s^R$ defines the innovations process V^R which is an m-dimensional $\mathcal{F}^{C,\lambda}$-Brownian motion, and where

$$P_k^\lambda = \lambda\big(I_d - \rho^{(\lambda)}(Q_{T_k-}^{C,\lambda})\big) = \lambda Q_{T_k-}^{C,\lambda}(Q_{T_k-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1}.$$

Note that P_k^λ is bounded since

$$\begin{aligned} \|\lambda Q_{T_k-}^{C,\lambda}(Q_{T_k-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1}\| &= \|Q_{T_k-}^{C,\lambda}(Q_{T_k-}^{C,\lambda} + \lambda\sigma_J\sigma_J^\top)^{-1}\lambda\sigma_J\sigma_J^\top(\sigma_J\sigma_J^\top)^{-1}\| \\ &\le C_Q\|(\sigma_J\sigma_J^\top)^{-1}\| =: C_P. \end{aligned}$$

The conditional mean m^D can be written as

$$m_t^D = \int_0^t \alpha(\delta - m_s^D)\,\mathrm{d}s + \int_0^t Q_s^D(\sigma_R\sigma_R^\top)^{-1}(\mathrm{d}R_s - m_s^D\,\mathrm{d}s) + \int_0^t Q_s^D(\sigma_J\sigma_J^\top)^{-1}(\mathrm{d}J_s - m_s^D\,\mathrm{d}s). \tag{8.36}$$

Note that

$$\mathrm{d}R_s - m_s^D\,\mathrm{d}s = \sigma_R\,\mathrm{d}V_s^R + (m_t^{C,\lambda} - m_s^D)\,\mathrm{d}s$$

and

$$\mathrm{d}J_s - m_s^D\,\mathrm{d}s = \sigma_J\,\mathrm{d}W_s^J + (\mu_s - m_s^D)\,\mathrm{d}s.$$

Plugging this into (8.36) and combining with (8.35) yields that the difference of the conditional means equals

$$\begin{aligned} m_t^{C,\lambda} - m_t^D &= -\alpha\int_0^t (m_s^{C,\lambda} - m_s^D)\mathrm{d}s + \int_0^t (Q_s^{C,\lambda} - Q_s^D)(\sigma_R\sigma_R^\top)^{-1}\sigma_R\,\mathrm{d}V_s^R \\ &\quad - \int_0^t Q_s^D(\sigma_R\sigma_R^\top)^{-1}(m_s^{C,\lambda} - m_s^D)\mathrm{d}s + \sum_{k=1}^{N_t}\frac{1}{\lambda}P_k^\lambda(Z_k^{(\lambda)} - m_{T_k-}^{C,\lambda}) \\ &\quad - \int_0^t Q_s^D(\sigma_J\sigma_J^\top)^{-1}\sigma_J\,\mathrm{d}W_s^J - \int_0^t Q_s^D(\sigma_J\sigma_J^\top)^{-1}(\mu_s - m_s^D)\,\mathrm{d}s \\ &= A_t^\lambda + B_t^\lambda + C_t^\lambda + D_t^\lambda + E_t^\lambda + F_t^\lambda, \end{aligned}$$

where

$$\begin{aligned} A_t^\lambda &= -\alpha\int_0^t (m_s^{C,\lambda} - m_s^D)\,\mathrm{d}s, \\ B_t^\lambda &= \int_0^t (Q_s^{C,\lambda} - Q_s^D)(\sigma_R\sigma_R^\top)^{-1}\sigma_R\,\mathrm{d}V_s^R, \\ C_t^\lambda &= \int_0^t Q_s^D(\sigma_R\sigma_R^\top)^{-1}(m_s^D - m_s^{C,\lambda})\,\mathrm{d}s, \\ D_t^\lambda &= \sum_{k=1}^{N_t} P_k^\lambda\sigma_J\int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}}\mathrm{d}W_s^J - \int_0^t Q_s^D(\sigma_J\sigma_J^\top)^{-1}\sigma_J\,\mathrm{d}W_s^J, \\ E_t^\lambda &= \sum_{k=1}^{N_t}\frac{1}{\lambda}P_k^\lambda\mu_{T_k} - \int_0^t Q_s^D(\sigma_J\sigma_J^\top)^{-1}\mu_s\,\mathrm{d}s, \\ F_t^\lambda &= \int_0^t Q_s^D(\sigma_J\sigma_J^\top)^{-1}m_s^D\,\mathrm{d}s - \sum_{k=1}^{N_t}\frac{1}{\lambda}P_k^\lambda m_{T_k-}^{C,\lambda}. \end{aligned}$$

Hence

$$\begin{aligned} v_t^\lambda &= \mathbb{E}\Big[\big\|A_t^\lambda + B_t^\lambda + C_t^\lambda + D_t^\lambda + E_t^\lambda + F_t^\lambda\big\|^2\Big] \\ &\le 6\,\mathbb{E}\Big[\big\|A_t^\lambda\big\|^2 + \big\|B_t^\lambda\big\|^2 + \big\|C_t^\lambda\big\|^2 + \big\|D_t^\lambda\big\|^2 + \big\|E_t^\lambda\big\|^2 + \big\|F_t^\lambda\big\|^2\Big]. \end{aligned} \tag{8.37}$$

For the various summands on the right-hand side of (8.37) we derive suitable upper bounds in the following.

Estimate for A^λ. Firstly, by using the Cauchy–Schwarz inequality from Lemma A.2 we have

$$\begin{aligned}\mathbb{E}\Big[\|A_t^\lambda\|^2\Big] &= \mathbb{E}\Big[\Big\|-\alpha\int_0^t (m_s^{C,\lambda} - m_s^D)\,\mathrm{d}s\Big\|^2\Big] \\ &\le \|\alpha\|^2 t\int_0^t \mathbb{E}\big[\|m_s^{C,\lambda} - m_s^D\|^2\big]\,\mathrm{d}s \le \|\alpha\|^2 T\int_0^t v_s^\lambda\,\mathrm{d}s =: C_A\int_0^t v_s^\lambda\,\mathrm{d}s.\end{aligned} \tag{8.38}$$

Estimate for B^λ. For the summand B_t^λ we use the multivariate version of Itô's isometry from Lemma A.3 and get

$$\begin{aligned}\mathbb{E}\Big[\|B_t^\lambda\|^2\Big] &= \mathbb{E}\Big[\Big\|\int_0^t (Q_s^{C,\lambda} - Q_s^D)(\sigma_R\sigma_R^\top)^{-1}\sigma_R\,\mathrm{d}V_s^R\Big\|^2\Big] \\ &\le C_{\mathrm{norm}}\int_0^t \mathbb{E}\big[\|(Q_s^{C,\lambda} - Q_s^D)(\sigma_R\sigma_R^\top)^{-1}\sigma_R\|^2\big]\,\mathrm{d}s \\ &\le C_{\mathrm{norm}}\|(\sigma_R\sigma_R^\top)^{-1}\sigma_R\|^2\int_0^t \mathbb{E}\big[\|Q_s^{C,\lambda} - Q_s^D\|^2\big]\,\mathrm{d}s \\ &\le C_{\mathrm{norm}}\|(\sigma_R\sigma_R^\top)^{-1}\sigma_R\|^2 T\frac{K_3}{\lambda} =: \frac{C_B}{\lambda}.\end{aligned} \tag{8.39}$$

The last inequality is due to Theorem 8.15.

Estimate for C^λ. For the summand C_t^λ we can argue similarly as for A_t^λ and get

$$\begin{aligned}\mathbb{E}\Big[\|C_t^\lambda\|^2\Big] &= \mathbb{E}\Big[\Big\|\int_0^t Q_s^D(\sigma_R\sigma_R^\top)^{-1}(m_s^D - m_s^{C,\lambda})\,\mathrm{d}s\Big\|^2\Big] \\ &\le t\int_0^t \mathbb{E}\big[\|Q_s^D(\sigma_R\sigma_R^\top)^{-1}(m_s^{C,\lambda} - m_s^D)\|^2\big]\,\mathrm{d}s \\ &\le t\int_0^t \|Q_s^D(\sigma_R\sigma_R^\top)^{-1}\|^2\,\mathbb{E}\big[\|m_s^{C,\lambda} - m_s^D\|^2\big]\,\mathrm{d}s \\ &\le C_Q^2\|(\sigma_R\sigma_R^\top)^{-1}\|^2 T\int_0^t v_s^\lambda\,\mathrm{d}s =: C_C\int_0^t v_s^\lambda\,\mathrm{d}s.\end{aligned} \tag{8.40}$$

Estimate for D^λ. The estimation of the terms containing D_t^λ, E_t^λ and F_t^λ is more involved. Recall that

$$D_t^\lambda = \sum_{k=1}^{N_t} P_k^\lambda \sigma_J\int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J - \int_0^t Q_s^D(\sigma_J\sigma_J^\top)^{-1}\sigma_J\,\mathrm{d}W_s^J.$$

The main difficulty in estimating this expression arises from the fact that the integrals

$$\int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J$$

in the sum do not align well with the integral over W^J from 0 to t. Since N_t is a random variable that can be smaller or larger than λt, it is necessary to distinguish various cases. Therefore, we define the integer-valued random variable $n_t := \min\{N_t, \lfloor\lambda t\rfloor\}$. Note also that

$N_t > \lfloor \lambda t \rfloor$ if and only if $N_t > \lambda t$, since N_t is integer-valued. This leads to the representation of D_t^λ as

$$D_t^{1,\lambda} + D_t^{2,\lambda} + D_t^{3,\lambda} + D_t^{4,\lambda},$$

where

$$\begin{aligned}
D_t^{1,\lambda} &= \sum_{k=1}^{n_t} P_k^\lambda \sigma_J \int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J - \int_0^{\frac{n_t}{\lambda}} Q_s^D (\sigma_J \sigma_J^\top)^{-1} \sigma_J \,\mathrm{d}W_s^J, \\
D_t^{2,\lambda} &= \mathbb{1}_{\{N_t > \lambda t\}} \sum_{k=\lfloor \lambda t \rfloor + 1}^{N_t} P_k^\lambda \sigma_J \int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J, \\
D_t^{3,\lambda} &= -\mathbb{1}_{\{N_t > \lambda t\}} \int_{\frac{\lfloor \lambda t \rfloor}{\lambda}}^{t} Q_s^D (\sigma_J \sigma_J^\top)^{-1} \sigma_J \,\mathrm{d}W_s^J, \\
D_t^{4,\lambda} &= -\mathbb{1}_{\{N_t \le \lambda t\}} \int_{\frac{N_t}{\lambda}}^{t} Q_s^D (\sigma_J \sigma_J^\top)^{-1} \sigma_J \,\mathrm{d}W_s^J.
\end{aligned}$$

Here, $D_t^{1,\lambda}$ can be written as

$$D_t^{1,\lambda} = \int_0^{\frac{n_t}{\lambda}} \big(H_s^\lambda - Q_s^D (\sigma_J \sigma_J^\top)^{-1} \big) \sigma_J \,\mathrm{d}W_s^J,$$

where $H_s^\lambda = P_k^\lambda$ for $s \in [\frac{k-1}{\lambda}, \frac{k}{\lambda})$. Therefore,

$$\begin{aligned}
\mathbb{E}\Big[\big\| D_t^{1,\lambda} \big\|^2 \Big] &= \mathbb{E}\Big[\Big\| \int_0^{\frac{n_t}{\lambda}} \big(H_s^\lambda - Q_s^D (\sigma_J \sigma_J^\top)^{-1} \big) \sigma_J \,\mathrm{d}W_s^J \Big\|^2 \Big] \\
&= \mathbb{E}\Big[\Big\| \int_0^{t} \mathbb{1}_{\{s \le \frac{n_t}{\lambda}\}} \big(H_s^\lambda - Q_s^D (\sigma_J \sigma_J^\top)^{-1} \big) \sigma_J \,\mathrm{d}W_s^J \Big\|^2 \Big] \\
&\le C_{\mathrm{norm}} \,\mathbb{E}\Big[\int_0^{t} \big\| \mathbb{1}_{\{s \le \frac{n_t}{\lambda}\}} \big(H_s^\lambda - Q_s^D (\sigma_J \sigma_J^\top)^{-1} \big) \sigma_J \big\|^2 \,\mathrm{d}s \Big] \\
&= C_{\mathrm{norm}} \big\| \sigma_J \big\|^2 \,\mathbb{E}\Big[\int_0^{\frac{n_t}{\lambda}} \big\| H_s^\lambda - Q_s^D (\sigma_J \sigma_J^\top)^{-1} \big\|^2 \,\mathrm{d}s \Big].
\end{aligned} \tag{8.41}$$

We take a closer look at the integrand inside the expectation in (8.41). Let $k \le n_t$ and $s \in [\frac{k-1}{\lambda}, \frac{k}{\lambda})$. Then

$$\begin{aligned}
H_s^\lambda - Q_s^D (\sigma_J \sigma_J^\top)^{-1} &= P_k^\lambda - Q_s^D (\sigma_J \sigma_J^\top)^{-1} \\
&= \lambda Q_{T_k-}^{C,\lambda} (Q_{T_k-}^{C,\lambda} + \lambda \sigma_J \sigma_J^\top)^{-1} - Q_s^D (\sigma_J \sigma_J^\top)^{-1} \\
&= \big(Q_{T_k-}^{C,\lambda} (Q_{T_k-}^{C,\lambda} + \lambda \sigma_J \sigma_J^\top)^{-1} \lambda \sigma_J \sigma_J^\top - Q_s^D \big) (\sigma_J \sigma_J^\top)^{-1}.
\end{aligned}$$

Hence, we can deduce that

$$\begin{aligned}
\big\| H_s^\lambda - Q_s^D (\sigma_J \sigma_J^\top)^{-1} \big\|^2 &\le \big\| (\sigma_J \sigma_J^\top)^{-1} \big\|^2 \big\| Q_{T_k-}^{C,\lambda} (Q_{T_k-}^{C,\lambda} + \lambda \sigma_J \sigma_J^\top)^{-1} \lambda \sigma_J \sigma_J^\top - Q_s^D \big\|^2 \\
&\le \big\| (\sigma_J \sigma_J^\top)^{-1} \big\|^2 \Big(2 \| Q_s^D - Q_{T_k-}^{C,\lambda} \|^2 + \frac{2\bar{C}^2}{\lambda^2} \Big) \\
&\le 2 \big\| (\sigma_J \sigma_J^\top)^{-1} \big\|^2 \Big(2 \| Q_s^D - Q_{T_k}^D \|^2 + 2 \| Q_{T_k}^D - Q_{T_k-}^{C,\lambda} \|^2 + \frac{\bar{C}^2}{\lambda^2} \Big)
\end{aligned}$$

by means of Lemma A.4. Since Q_s^D is differentiable in s with bounded derivative we can deduce that

$$\|Q_s^D - Q_{T_k}^D\|^2 \leq \tilde{C}_Q^2 (T_k - s)^2.$$

By means of Theorem 8.15 and plugging back into (8.41) this implies that

$$\mathbb{E}\Big[\|D_t^{1,\lambda}\|^2\Big] \leq 2TC_{\text{norm}}\|\sigma_J\|^2\|(\sigma_J\sigma_J^\top)^{-1}\|^2\Big(2\frac{\tilde{C}_Q^2}{\lambda} + \frac{2K_3}{\lambda} + \frac{\bar{C}^2}{\lambda^2}\Big) \leq \frac{C_{D,1}}{\lambda} \tag{8.42}$$

for all $\lambda \geq \tilde{\lambda}_0$ and some $\tilde{\lambda}_0 > 0$ and where $C_{D,1} > 0$ is a suitable constant. Next, we consider $D_t^{2,\lambda}$. Note that

$$\begin{aligned}
\mathbb{E}\Big[\|D_t^{2,\lambda}\|^2\Big] &= \mathbb{E}\Bigg[\mathbb{1}_{\{N_t > \lambda t\}}\Bigg\|\sum_{k=\lfloor \lambda t\rfloor+1}^{N_t} P_k^\lambda \sigma_J \int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J\Bigg\|^2\Bigg] \\
&= \mathbb{E}\Bigg[\mathbb{1}_{\{N_t > \lambda t\}}\sum_{i=1}^{d}\Bigg(\sum_{k=\lfloor \lambda t\rfloor+1}^{\infty} \mathbb{1}_{\{N_t \geq k\}}\Big(P_k^\lambda \sigma_J \int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J\Big)^i\Bigg)^2\Bigg] \\
&= \sum_{i=1}^{d}\sum_{k,n=\lfloor \lambda t\rfloor+1}^{\infty} \mathbb{E}\Bigg[\mathbb{1}_{\{N_t \geq k,n\}}\Big(P_k^\lambda \sigma_J \int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J\Big)^i\Big(P_n^\lambda \sigma_J \int_{\frac{n-1}{\lambda}}^{\frac{n}{\lambda}} \mathrm{d}W_s^J\Big)^i\Bigg].
\end{aligned}$$

Since W^J is independent of N_t and of the matrices P^λ, this expression equals

$$\begin{aligned}
&\sum_{i=1}^{d}\sum_{k=\lfloor \lambda t\rfloor+1}^{\infty} \mathbb{E}\Bigg[\mathbb{1}_{\{N_t \geq k\}}\Bigg(\Big(P_k^\lambda \sigma_J \int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J\Big)^i\Bigg)^2\Bigg] \\
&= \sum_{k=\lfloor \lambda t\rfloor+1}^{\infty} \mathbb{E}\Bigg[\mathbb{1}_{\{N_t \geq k\}}\Bigg\|P_k^\lambda \sigma_J \int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J\Bigg\|^2\Bigg] \\
&\leq C_P^2\|\sigma_J\|^2 \sum_{k=\lfloor \lambda t\rfloor+1}^{\infty} \mathbb{E}\Bigg[\mathbb{1}_{\{N_t \geq k\}}\Bigg\|\int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J\Bigg\|^2\Bigg],
\end{aligned}$$

where

$$\begin{aligned}
\sum_{k=\lfloor \lambda t\rfloor+1}^{\infty} \mathbb{E}\Bigg[\mathbb{1}_{\{N_t \geq k\}}\Bigg\|\int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J\Bigg\|^2\Bigg] &= \sum_{k=\lfloor \lambda t\rfloor+1}^{\infty} \mathbb{E}\big[\mathbb{1}_{\{N_t \geq k\}}\big]\,\mathbb{E}\Bigg[\Bigg\|\int_{\frac{k-1}{\lambda}}^{\frac{k}{\lambda}} \mathrm{d}W_s^J\Bigg\|^2\Bigg] \\
&\leq C_{\text{norm}}\frac{1}{\lambda}\sum_{k=\lfloor \lambda t\rfloor+1}^{\infty} \mathbb{E}\big[\mathbb{1}_{\{N_t \geq k\}}\big] \leq C_{\text{norm}}\frac{1}{\lambda}C_N\sqrt{\lambda}
\end{aligned}$$

due to Lemma A.7. All in all, we therefore find a constant $C_{D,2} > 0$ with

$$\mathbb{E}\Big[\|D_t^{2,\lambda}\|^2\Big] \leq \frac{C_{D,2}}{\sqrt{\lambda}}. \tag{8.43}$$

For $D_t^{3,\lambda}$ and $D_t^{4,\lambda}$ the estimations lead to similar expressions. For $D_t^{3,\lambda}$ we get, again using the Itô isometry from Lemma A.3,

$$\begin{aligned}
\mathbb{E}\Big[\big\|D_t^{3,\lambda}\big\|^2\Big] &= \mathbb{E}\Big[\Big\|\int_{\frac{\lfloor\lambda t\rfloor}{\lambda}}^t \mathbb{1}_{\{N_t>\lambda t\}} Q_s^D(\sigma_J\sigma_J^\top)^{-1}\sigma_J\,\mathrm{d}W_s^J\Big\|^2\Big] \\
&\le C_{\mathrm{norm}} \int_{\frac{\lfloor\lambda t\rfloor}{\lambda}}^t \mathbb{E}\big[\big\|\mathbb{1}_{\{N_t>\lambda t\}} Q_s^D(\sigma_J\sigma_J^\top)^{-1}\sigma_J\big\|^2\big]\,\mathrm{d}s \\
&\le C_{\mathrm{norm}} C_Q^2 \big\|(\sigma_J\sigma_J^\top)^{-1}\sigma_J\big\|^2\Big(t-\frac{\lfloor\lambda t\rfloor}{\lambda}\Big)
\end{aligned}$$

and observe that

$$t-\frac{\lfloor\lambda t\rfloor}{\lambda} = \Big(\lambda t-\lfloor\lambda t\rfloor\Big)\frac{1}{\lambda} \le \frac{1}{\lambda}.$$

In conclusion, we have a constant $C_{D,3}>0$ with

$$\mathbb{E}\Big[\big\|D_t^{3,\lambda}\big\|^2\Big] \le \frac{C_{D,3}}{\lambda}. \tag{8.44}$$

For $D_t^{4,\lambda}$ we get by the Itô isometry from Lemma A.3 that

$$\begin{aligned}
\mathbb{E}\Big[\big\|D_t^{4,\lambda}\big\|^2\Big] &= \mathbb{E}\Big[\Big\|\mathbb{1}_{\{N_t\le\lambda t\}}\int_{\frac{N_t}{\lambda}}^t Q_s^D(\sigma_J\sigma_J^\top)^{-1}\sigma_J\,\mathrm{d}W_s^J\Big\|^2\Big] \\
&\le C_{\mathrm{norm}}\,\mathbb{E}\Big[\int_{\frac{N_t}{\lambda}}^t \mathbb{1}_{\{N_t\le\lambda t\}}\big\|Q_s^D(\sigma_J\sigma_J^\top)^{-1}\sigma_J\big\|^2\,\mathrm{d}s\Big] \\
&\le C_{\mathrm{norm}} C_Q^2 \big\|(\sigma_J\sigma_J^\top)^{-1}\sigma_J\big\|^2\,\mathbb{E}\Big[\mathbb{1}_{\{N_t\le\lambda t\}}\Big(t-\frac{N_t}{\lambda}\Big)\Big].
\end{aligned}$$

Here, we see that

$$\mathbb{E}\Big[\mathbb{1}_{\{N_t\le\lambda t\}}\Big(t-\frac{N_t}{\lambda}\Big)\Big] \le \frac{1}{\lambda}\,\mathbb{E}\big[(\lambda t-N_t)^+\big] \le \frac{C_N}{\sqrt{\lambda}}$$

due to Lemma A.7. Hence, it holds

$$\mathbb{E}\Big[\big\|D_t^{4,\lambda}\big\|^2\Big] \le C_{\mathrm{norm}} C_Q^2 \big\|(\sigma_J\sigma_J^\top)^{-1}\sigma_J\big\|^2\frac{C_N}{\sqrt{\lambda}} =: \frac{C_{D,4}}{\sqrt{\lambda}}. \tag{8.45}$$

Now, combining the estimations from (8.42), (8.43), (8.44) and (8.45), we obtain

$$\begin{aligned}
\mathbb{E}\Big[\big\|D_t^{\lambda}\big\|^2\Big] &= \mathbb{E}\Big[\big\|D_t^{1,\lambda}+D_t^{2,\lambda}+D_t^{3,\lambda}+D_t^{4,\lambda}\big\|^2\Big] \\
&\le 4\Big(\mathbb{E}\Big[\big\|D_t^{1,\lambda}\big\|^2\Big]+\mathbb{E}\Big[\big\|D_t^{2,\lambda}\big\|^2\Big]+\mathbb{E}\Big[\big\|D_t^{3,\lambda}\big\|^2\Big]+\mathbb{E}\Big[\big\|D_t^{4,\lambda}\big\|^2\Big]\Big) \\
&\le 4\Big(\frac{C_{D,1}}{\lambda}+\frac{C_{D,2}}{\sqrt{\lambda}}+\frac{C_{D,3}}{\lambda}+\frac{C_{D,4}}{\sqrt{\lambda}}\Big) \\
&\le 4(C_{D,1}+C_{D,2}+C_{D,3}+C_{D,4})\frac{1}{\sqrt{\lambda}} =: \frac{C_D}{\sqrt{\lambda}}
\end{aligned} \tag{8.46}$$

for all $\lambda\ge 1$.

Estimates for $\boldsymbol{E^\lambda}$ and $\boldsymbol{F^\lambda}$. Finding upper bounds for the terms $\mathbb{E}[\|E_t^\lambda\|^2]$ and $\mathbb{E}[\|F_t^\lambda\|^2]$ works by the same approach as for $\mathbb{E}[\|D_t^\lambda\|^2]$, i.e. by splitting up into different parts by means of the random variable $n_t = \min\{N_t, \lfloor \lambda t \rfloor\}$. The result is that there exist constants C_E and $C_F > 0$ such that

$$\mathbb{E}\Big[\big\|E_t^\lambda\big\|^2\Big] \leq \frac{C_E}{\sqrt{\lambda}} \quad \text{and} \quad \mathbb{E}\Big[\big\|F_t^\lambda\big\|^2\Big] \leq \frac{C_F}{\sqrt{\lambda}}.$$

Conclusion with Gronwall's Lemma. These upper bounds, as well as those in (8.38), (8.39), (8.40) and (8.46) can now be used in (8.37) to obtain

$$\begin{aligned}
v_t^\lambda &\leq 6\,\mathbb{E}\Big[\big\|A_t^\lambda\big\|^2 + \big\|B_t^\lambda\big\|^2 + \big\|C_t^\lambda\big\|^2 + \big\|D_t^\lambda\big\|^2 + \big\|E_t^\lambda\big\|^2 + \big\|F_t^\lambda\big\|^2\Big] \\
&\leq 6(C_A + C_C)\int_0^t v_s^\lambda \,\mathrm{d}s + \frac{6C_B}{\lambda} + \frac{6(C_D + C_E + C_F)}{\sqrt{\lambda}} \\
&\leq 6(C_A + C_C)\int_0^t v_s^\lambda \,\mathrm{d}s + \frac{6(C_B + C_D + C_E + C_F)}{\sqrt{\lambda}}
\end{aligned}$$

for all $\lambda \geq 1$. Now Gronwall's Lemma, see Lemma A.5, implies

$$\begin{aligned}
v_t^\lambda &\leq \frac{6(C_B + C_D + C_E + C_F)}{\sqrt{\lambda}} \exp\big(6(C_A + C_C)t\big) \\
&\leq 6(C_B + C_D + C_E + C_F)\exp\big(6(C_A + C_C)T\big)\frac{1}{\sqrt{\lambda}} =: \frac{K_4}{\sqrt{\lambda}}.
\end{aligned}$$

This concludes the proof. □

9. Application to Utility Maximization

As an application of the convergence results from Chapter 8 we now consider a portfolio optimization problem in our financial market. We take again $\mathbb{T} = [0,T]$ where $T > 0$ is some finite investment horizon. For convenience, we assume here that the interest rate r of the risk-free asset is equal to zero. However, the results below can easily be extended to a market model with $r \neq 0$.

9.1. Optimal strategy and value function

An investor's trading in the market can be described by a self-financing trading strategy $(\pi_t)_{t\in[0,T]}$ with values in $\mathbb{R}^d$. Here, π_t^i, $i = 1,\dots,d$, is the proportion of wealth that is invested in asset i at time t. The corresponding wealth process $(X_t^\pi)_{t\in[0,T]}$ then follows the stochastic differential equation

$$\mathrm{d}X_t^\pi = X_t^\pi \pi_t^\top \big(\mu_t \,\mathrm{d}t + \sigma_R \,\mathrm{d}W_t^R\big)$$

with initial capital $X_0^\pi = x_0 > 0$. An investor's trading strategy has to be adapted to her investor filtration. To ensure strictly positive wealth, we also impose some integrability constraint on the trading strategies. Then we denote by

$$\mathcal{A}^H(x_0) = \left\{ \pi = (\pi_t)_{t\in[0,T]} \;\middle|\; \pi \text{ is } \mathbb{F}^H\text{-adapted, } X_0^\pi = x_0,\ \mathbb{E}\left[\int_0^T \|\sigma_R^\top \pi_t\|^2 \,\mathrm{d}t\right] < \infty \right\}$$

the class of admissible trading strategies for the H-investor. The optimization problem we address is a utility maximization problem where investors want to maximize expected logarithmic utility of terminal wealth. Hence,

$$V^H(x_0) = \sup\Big\{ \mathbb{E}\big[\log(X_T^\pi)\big] \;\Big|\; \pi \in \mathcal{A}^H(x_0) \Big\} \tag{9.1}$$

is the value function of our optimization problem. This utility maximization problem under partial information has been solved in Brendle [7] for the case of power utility. Karatzas and Zhao [35] address also the case with logarithmic utility. In the Master's thesis Westphal [64] the optimization problem has been solved for a general H-investor in the context of the different information regimes addressed in this thesis. We recall the result in the proposition below.

Proposition 9.1. *The optimal strategy for the optimization problem* (9.1) *is* $(\pi_t^{H,*})_{t\in[0,T]}$ *with* $\pi_t^{H,*} = (\sigma_R\sigma_R^\top)^{-1} m_t^H$, *and the optimal value is*

$$\begin{aligned} V^H(x_0) &= \log(x_0) + \frac{1}{2}\int_0^T \operatorname{tr}\Big((\sigma_R\sigma_R^\top)^{-1}\, \mathbb{E}\big[m_t^H (m_t^H)^\top\big]\Big)\mathrm{d}t \\ &= \log(x_0) + \frac{1}{2}\int_0^T \operatorname{tr}\Big((\sigma_R\sigma_R^\top)^{-1} \big(\Sigma_t + m_t m_t^\top - \mathbb{E}\big[Q_t^H\big]\big)\Big)\mathrm{d}t. \end{aligned}$$

Proof. As a preliminary result, note that

$$\begin{aligned} Q_t^H &= \mathbb{E}\big[(\mu_t - m_t^H)(\mu_t - m_t^H)^\top \,\big|\, \mathcal{F}_t^H\big] \\ &= \mathbb{E}\big[\mu_t\mu_t^\top - m_t^H\mu_t^\top - \mu_t(m_t^H)^\top + m_t^H(m_t^H)^\top \,\big|\, \mathcal{F}_t^H\big] \\ &= \mathbb{E}\big[\mu_t\mu_t^\top \,\big|\, \mathcal{F}_t^H\big] - m_t^H(m_t^H)^\top. \end{aligned}$$

Therefore, by taking expectation on both sides,

$$\mathbb{E}\big[m_t^H(m_t^H)^\top\big] = \mathbb{E}\big[\mu_t\mu_t^\top\big] - \mathbb{E}\big[Q_t^H\big] = \Sigma_t + m_t m_t^\top - \mathbb{E}\big[Q_t^H\big]. \tag{9.2}$$

From the dynamics of the wealth process we get for any $\pi \in \mathcal{A}^H(x_0)$ that

$$\log(X_T^\pi) = \log(x_0) + \int_0^T \Big(\pi_t^\top \mu_t - \frac{1}{2}\|\sigma_R^\top \pi_t\|^2\Big)\mathrm{d}t + \int_0^T \pi_t^\top \sigma_R \,\mathrm{d}W_t^R.$$

By using that the stochastic integral has expectation zero and applying Fubini we deduce

$$\begin{aligned} \mathbb{E}\big[\log(X_T^\pi)\big] &= \log(x_0) + \int_0^T \mathbb{E}\Big[\pi_t^\top \mu_t - \frac{1}{2}\|\sigma_R^\top \pi_t\|^2\Big]\mathrm{d}t \\ &= \log(x_0) + \int_0^T \mathbb{E}\Big[\mathbb{E}\Big[\pi_t^\top \mu_t - \frac{1}{2}\|\sigma_R^\top \pi_t\|^2 \,\Big|\, \mathcal{F}_t^H\Big]\Big]\mathrm{d}t \\ &= \log(x_0) + \int_0^T \mathbb{E}\Big[\pi_t^\top m_t^H - \frac{1}{2}\|\sigma_R^\top \pi_t\|^2\Big]\mathrm{d}t. \end{aligned} \tag{9.3}$$

Now we fix some $t \in [0,T]$. Following a pointwise maximization, we formally take the derivative of the expression inside the expectation with respect to π_t. Using the first-order condition, we set the derivative equal to zero, which means setting $m_t^H - \sigma_R\sigma_R^\top \pi_t$ equal to the zero vector. Since we have assumed $\sigma_R\sigma_R^\top$ to be positive definite, this implies that $\pi_t^{H,*} = (\sigma_R\sigma_R^\top)^{-1} m_t^H$ maximizes the above integrand pointwise. It remains to check that $(\pi_t^{H,*})_{t\in[0,T]}$ is indeed admissible. Firstly, we note that

$$\int_0^T \|\sigma_R^\top \pi_t^{H,*}\|^2 \,\mathrm{d}t = \int_0^T \big\|\sigma_R^\top (\sigma_R\sigma_R^\top)^{-1} m_t^H\big\|^2 \,\mathrm{d}t = \int_0^T (m_t^H)^\top (\sigma_R\sigma_R^\top)^{-1} m_t^H \,\mathrm{d}t. \tag{9.4}$$

Taking the expectation in (9.4) and applying Fubini we get

$$\begin{aligned} \mathbb{E}\Big[\int_0^T \|\sigma_R^\top \pi_t^{H,*}\|^2 \,\mathrm{d}t\Big] &= \int_0^T \mathbb{E}\big[(m_t^H)^\top (\sigma_R\sigma_R^\top)^{-1} m_t^H\big] \,\mathrm{d}t \\ &= \int_0^T \mathrm{tr}\Big((\sigma_R\sigma_R^\top)^{-1}\,\mathbb{E}\big[m_t^H(m_t^H)^\top\big]\Big)\mathrm{d}t, \end{aligned}$$

where the second equality follows from cyclicity of the trace. Additionally, by (9.2) we have $\mathbb{E}[m_t^H(m_t^H)^\top] = \Sigma_t + m_t m_t^\top - \mathbb{E}[Q_t^H]$, so

$$\mathbb{E}\Big[\int_0^T \|\sigma_R^\top \pi_t^{H,*}\|^2 \,\mathrm{d}t\Big] = \int_0^T \Big(\mathrm{tr}\big((\sigma_R\sigma_R^\top)^{-1}(\Sigma_t + m_t m_t^\top)\big) - \mathrm{tr}\big((\sigma_R\sigma_R^\top)^{-1}\,\mathbb{E}\big[Q_t^H\big]\big)\Big)\mathrm{d}t.$$

Recall that $(\sigma_R\sigma_R^\top)^{-1}$ is symmetric positive definite and Q_t^H is symmetric positive semidefinite. By Wang et al. [63, Lem. 1] the product $(\sigma_R\sigma_R^\top)^{-1}Q_t^H$ has a non-negative trace, hence

$$\mathbb{E}\Big[\int_0^T \|\sigma_R^\top \pi_t^{H,*}\|^2 \,\mathrm{d}t\Big] \le \int_0^T \operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}(\Sigma_t + m_t m_t^\top)\big)\,\mathrm{d}t < \infty,$$

where finiteness follows from continuity. It follows that $(\pi_t^{H,*})_{t\in[0,T]}$ is an admissible strategy. As in (9.3) we then get for the value function

$$\begin{aligned}
V^H(x_0) &= \log(x_0) + \int_0^T \mathbb{E}\Big[(\pi_t^{H,*})^\top m_t^H - \frac{1}{2}\|\sigma_R^\top \pi_t^{H,*}\|^2\Big]\mathrm{d}t \\
&= \log(x_0) + \int_0^T \mathbb{E}\Big[(m_t^H)^\top(\sigma_R\sigma_R^\top)^{-1}m_t^H - \frac{1}{2}\big\|(m_t^H)^\top(\sigma_R\sigma_R^\top)^{-1}\sigma_R\big\|^2\Big]\mathrm{d}t \\
&= \log(x_0) + \int_0^T \mathbb{E}\Big[\frac{1}{2}(m_t^H)^\top(\sigma_R\sigma_R^\top)^{-1}m_t^H\Big]\mathrm{d}t \\
&= \log(x_0) + \frac{1}{2}\int_0^T \operatorname{tr}\Big((\sigma_R\sigma_R^\top)^{-1}\,\mathbb{E}\big[m_t^H(m_t^H)^\top\big]\Big)\mathrm{d}t,
\end{aligned}$$

again by cyclicity of the trace. The second representation of the value function then follows directly from (9.2). □

Note that under full information the optimal strategy is $(\sigma_R\sigma_R^\top)^{-1}\mu_t$. This implies that for our portfolio optimization problem under partial information, the *certainty equivalence principle* holds, meaning that the drift μ_t in the optimal strategy is replaced by the conditional mean m_t^H.

9.2. Properties of the value function

The value function of the H-investor is an integral functional of the expectation of $(Q_t^H)_{t\in[0,T]}$. This makes it possible to deduce properties of the value function from properties of the conditional covariance matrices. Firstly, we prove the following intuitive relation between the value functions of different investors.

Corollary 9.2. *For any $x_0 > 0$ it holds*

$$\max\big\{V^R(x_0), V^E(x_0)\big\} \le V^C(x_0) \le V^F(x_0) \quad \text{and} \quad V^R(x_0) \le V^D(x_0) \le V^F(x_0).$$

Proof. By Proposition 6.6 we know that $Q_t^C \preceq Q_t^R$ for any $t \in [0,T]$. By assumption, $(\sigma_R\sigma_R^\top)^{-1}$ is positive definite. Hence,

$$\operatorname{tr}((\sigma_R\sigma_R^\top)^{-1}(Q_t^R - Q_t^C)) \ge 0$$

and therefore

$$\operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}\,\mathbb{E}[Q_t^R]\big) \ge \operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}\,\mathbb{E}[Q_t^C]\big),$$

which implies by the previous proposition that $V^C(x_0) \ge V^R(x_0)$. The same holds for $H = E$ instead of R or $H = D$ instead of C. Since $Q_t^F = \mathbf{0}_d$ we also have $V^F(x_0) \ge V^C(x_0)$ and $V^F(x_0) \ge V^D(x_0)$. □

From Theorem 8.1 we immediately deduce the following result about the asymptotic behavior of the value function when the number of expert opinions goes to infinity and the covariance matrices of the expert are bounded.

Corollary 9.3. *Let the assumptions of Theorem 8.1 be fulfilled. Denote the value functions corresponding to the n expert opinions by $V^{E,n}(x_0)$ and $V^{C,n}(x_0)$. Then*

$$\lim_{n\to\infty} V^{E,n}(x_0) = \lim_{n\to\infty} V^{C,n}(x_0) = V^F(x_0).$$

Proof. Recall from Theorem 8.1 that

$$\lim_{n\to\infty} Q_u^{E,n} = \lim_{n\to\infty} Q_u^{C,n} = \mathbf{0}_d \tag{9.5}$$

and that $Q_u^F = \mathbf{0}_d$ for all $u \in (0, T]$. We observe for $H \in \{E, C\}$ that

$$\operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}(\Sigma_t + m_t m_t^\top - Q_t^H)\big) \le \operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}(\Sigma_t + m_t m_t^\top)\big)$$

since both $(\sigma_R\sigma_R^\top)^{-1}$ and Q_t^H are positive semidefinite. By using dominated convergence and (9.5) we conclude from the representation of the value function in Proposition 9.1 that $V^{E,n}(x_0)$ and $V^{C,n}(x_0)$ converge to $V^F(x_0)$ when n goes to infinity. □

The convergence results of Theorems 8.5 and 8.15 also carry over to convergence results for the respective value functions. Firstly, we address the situation with deterministic information dates t_k from Section 8.2.1 where we have shown uniform convergence of $Q^{C,n}$ to Q^D.

Corollary 9.4. *Under Assumption 8.4 there exists a constant $K_5 > 0$ such that*

$$\big|V^{C,n}(x_0) - V^D(x_0)\big| \le K_5\Delta_n$$

for any initial wealth $x_0 > 0$. In particular, $\lim_{n\to\infty} V^{C,n}(x_0) = V^D(x_0)$.

Proof. From Proposition 9.1 we deduce

$$\begin{aligned}\big|V^{C,n}(x_0) - V^D(x_0)\big| &= \left|\frac{1}{2}\int_0^T \operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}(Q_t^D - Q_t^{C,n})\big)\,\mathrm{d}t\right| \\ &\le \frac{1}{2}\int_0^T \big|\operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}(Q_t^D - Q_t^{C,n})\big)\big|\,\mathrm{d}t,\end{aligned} \tag{9.6}$$

noting that $Q_t^{C,n}$ and Q_t^D are deterministic for every $t \in [0, T]$. Since $(\sigma_R\sigma_R^\top)^{-1}$ is symmetric and positive definite, and $Q_t^D - Q_t^{C,n}$ is symmetric, it follows from Wang et al. [63, Lem. 1] that

$$\big|\operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}(Q_t^D - Q_t^{C,n})\big)\big| \le \operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}\big)\big\|Q_t^D - Q_t^{C,n}\big\|.$$

Inserting this into (9.6) we then get from Theorem 8.5 that

$$\big|V^{C,n}(x_0) - V^D(x_0)\big| \le \frac{1}{2}T\operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}\big)K_1\Delta_n$$

which proves the claim when setting $K_5 = \frac{1}{2}K_1T\operatorname{tr}((\sigma_R\sigma_R^\top)^{-1})$. □

The analogous result also holds in the setting of Section 8.2.2 where information dates T_k are the jump times of a Poisson process. Recall that in Theorem 8.15 we have shown uniform convergence of $Q^{C,\lambda}$ to Q^D.

Corollary 9.5. *Under Assumption 8.10 there exists a constant $K_6 > 0$ and a $\lambda_0 > 0$ such that*

$$\big|V^{C,\lambda}(x_0) - V^D(x_0)\big| \le \frac{K_6}{\sqrt{\lambda}}$$

for any initial wealth $x_0 > 0$ and all $\lambda \ge \lambda_0$. In particular, $\lim_{\lambda\to\infty} V^{C,\lambda}(x_0) = V^D(x_0)$.

Proof. As in the proof of Corollary 9.4 we first use Proposition 9.1 to obtain

$$\begin{aligned}\big|V^{C,\lambda}(x_0) - V^D(x_0)\big| &= \left|\frac{1}{2}\int_0^T \operatorname{tr}\Big((\sigma_R\sigma_R^\top)^{-1}\big(Q_t^D - \mathbb{E}\big[Q_t^{C,\lambda}\big]\big)\Big)\mathrm{d}t\right| \\ &\le \frac{1}{2}\int_0^T \mathbb{E}\Big[\big|\operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}(Q_t^D - Q_t^{C,\lambda})\big)\big|\Big]\mathrm{d}t.\end{aligned}$$

Since $(\sigma_R\sigma_R^\top)^{-1}$ is symmetric and positive definite, and $Q_t^D - Q_t^{C,\lambda}$ is symmetric, it follows from Wang et al. [63, Lem. 1] that

$$\big|\operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}(Q_t^D - Q_t^{C,\lambda})\big)\big| \le \operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}\big)\big\|Q_t^D - Q_t^{C,\lambda}\big\|.$$

Consequently, by applying the Lyapunov inequality $\mathbb{E}[|X|] \le \sqrt{\mathbb{E}[X^2]}$ and Theorem 8.15 we get

$$\begin{aligned}\big|V^{C,\lambda}(x_0) - V^D(x_0)\big| &\le \frac{1}{2}\int_0^T \mathbb{E}\Big[\operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}\big)\big\|Q_t^D - Q_t^{C,\lambda}\big\|\Big]\mathrm{d}t \\ &\le \frac{1}{2}\int_0^T \operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}\big)\sqrt{\mathbb{E}\big[\big\|Q_t^D - Q_t^{C,\lambda}\big\|^2\big]}\,\mathrm{d}t \\ &\le \frac{1}{2}T\operatorname{tr}\big((\sigma_R\sigma_R^\top)^{-1}\big)\sqrt{\frac{K_3}{\lambda}},\end{aligned}$$

for all $\lambda \ge \lambda_0$, which completes the proof when setting $K_6 = \frac{1}{2}\sqrt{K_3}T\operatorname{tr}((\sigma_R\sigma_R^\top)^{-1})$. $\square$

Corollaries 9.4 and 9.5 show that for both settings, with deterministic information dates as in Assumption 8.4 and with random information dates being the jump times of a Poisson process as in Assumption 8.10, the value function of the C-investor converges to the value function of the D-investor as the frequency of information dates goes to infinity.

The following proposition shows that not only does the value function of the C-investor converge to the value function of the D-investor, but also the absolute difference of the utility attained by $\pi^{C,*}$, respectively $\pi^{D,*}$, goes to zero when increasing the number or the frequency of discrete-time expert opinions. This implies that the utility of the C-investor observing the discrete-time expert opinions also pathwise becomes arbitrarily close to the utility of the D-investor when the number of discrete-time expert opinions becomes large. For this result, we need the strong L^2-convergence of the conditional expectations, convergence in distribution would not be enough here.

Proposition 9.6. *Under Assumption 8.4 it holds*

$$\lim_{n\to\infty} \mathbb{E}\Big[\big|\log(X_T^{\pi^{C,n,*}}) - \log(X_T^{\pi^{D,*}})\big|\Big] = 0,$$

under Assumption 8.10 it holds

$$\lim_{\lambda\to\infty} \mathbb{E}\Big[\big|\log(X_T^{\pi^{C,\lambda,*}}) - \log(X_T^{\pi^{D,*}})\big|\Big] = 0.$$

Proof. Consider the setting of Assumption 8.4. Note that

$$\begin{aligned}
&\log(X_T^{\pi^{C,n,*}}) - \log(X_T^{\pi^{D,*}}) \\
&= \int_0^T \big((\pi_t^{C,n,*} - \pi_t^{D,*})^\top \mu_t - \tfrac{1}{2}\big(\|\sigma_R^\top \pi_t^{C,n,*}\|^2 - \|\sigma_R^\top \pi_t^{D,*}\|^2\big)\big)\,\mathrm{d}t + \int_0^T (\pi_t^{C,n,*} - \pi_t^{D,*})^\top \sigma_R \,\mathrm{d}W_t^R \\
&= \int_0^T \big((m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}\mu_t - \tfrac{1}{2}\big((m_t^{C,n})^\top (\sigma_R\sigma_R^\top)^{-1} m_t^{C,n} - (m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1} m_t^D\big)\big)\mathrm{d}t \\
&\quad + \int_0^T (m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}\sigma_R \,\mathrm{d}W_t^R \\
&= \tfrac{1}{2}\int_0^T (m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}(2\mu_t - m_t^{C,n} - m_t^D)\,\mathrm{d}t + \int_0^T (m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}\sigma_R \,\mathrm{d}W_t^R
\end{aligned}$$

where we have used the representation of the optimal strategies given in Proposition 9.1. After applying the absolute value and the expectation to the expression above, and by using the triangle inequality, we obtain

$$\begin{aligned}
\mathbb{E}\Big[\big|\log(X_T^{\pi^{C,n,*}}) - \log(X_T^{\pi^{D,*}})\big|\Big] &\le \frac{1}{2}\,\mathbb{E}\Big[\Big|\int_0^T (m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}(\mu_t - m_t^{C,n})\,\mathrm{d}t\Big|\Big] \\
&\quad + \frac{1}{2}\,\mathbb{E}\Big[\Big|\int_0^T (m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}(\mu_t - m_t^D)\,\mathrm{d}t\Big|\Big] \qquad (9.7) \\
&\quad + \mathbb{E}\Big[\Big|\int_0^T (m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}\sigma_R\,\mathrm{d}W_t^R\Big|\Big].
\end{aligned}$$

For the first summand in (9.7) we have, due to the Cauchy–Schwarz inequality,

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\int_0^T (m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}(\mu_t - m_t^{C,n})\,\mathrm{d}t\Big|\Big] \\
&\le \mathbb{E}\Big[\int_0^T \big|(m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}(\mu_t - m_t^{C,n})\big|\,\mathrm{d}t\Big] \\
&\le \|(\sigma_R\sigma_R^\top)^{-1}\|\,\mathbb{E}\Big[\int_0^T \|m_t^{C,n} - m_t^D\|\,\|\mu_t - m_t^{C,n}\|\,\mathrm{d}t\Big] \\
&\le \|(\sigma_R\sigma_R^\top)^{-1}\|\,\mathbb{E}\Big[\int_0^T \|m_t^{C,n} - m_t^D\|^2\,\mathrm{d}t\Big]^{1/2}\,\mathbb{E}\Big[\int_0^T \|\mu - m_t^{C,n}\|^2\,\mathrm{d}t\Big]^{1/2}.
\end{aligned}$$

The right-hand side of this expression goes to zero when n goes to infinity by Theorem 8.6 and by boundedness of $Q^{C,n}$, see Lemma 7.3. The second summand in (9.7) goes to zero by an analogous argumentation. For the third summand in (9.7), note that

$$
\begin{aligned}
&\mathbb{E}\bigg[\bigg|\int_0^T (m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}\sigma_R \,\mathrm{d}W_t^R\bigg|\bigg] \\
&\leq \mathbb{E}\bigg[\bigg(\int_0^T (m_t^{C,n} - m_t^D)^\top (\sigma_R\sigma_R^\top)^{-1}\sigma_R \,\mathrm{d}W_t^R\bigg)^2\bigg]^{1/2} \\
&= \mathbb{E}\bigg[\int_0^T \|\sigma_R^\top(\sigma_R\sigma_R^\top)^{-1}(m_t^{C,n} - m_t^D)\|^2 \,\mathrm{d}t\bigg]^{1/2} \\
&\leq \|\sigma_R^\top(\sigma_R\sigma_R^\top)^{-1}\| \,\mathbb{E}\bigg[\int_0^T \|m_t^{C,n} - m_t^D\|^2 \,\mathrm{d}t\bigg]^{1/2}.
\end{aligned}
$$

In the second step we have used the Itô isometry. Again, the right-hand side of the above inequality goes to zero as n goes to infinity by Theorem 8.6. Putting these results together shows

$$\lim_{n\to\infty} \mathbb{E}\Big[\big|\log(X_T^{\pi^{C,n,*}}) - \log(X_T^{\pi^{D,*}})\big|\Big] = 0.$$

The proof for the convergence under Assumption 8.10 is completely analogous. □

Note that the convergence of the value functions could also be deduced directly from the previous proposition. However, the proofs that we have given in Corollaries 9.4 and 9.5 using the convergence of the conditional covariance matrices are more direct and thus yield a sharper bound for the order of convergence than what we would get from the previous proposition.

Remark 9.7. Portfolio problems that consider maximization of expected power utility instead of logarithmic utility are typically more demanding and the above methods cannot be applied directly.

We have seen that for logarithmic utility the value function is given in terms of an integral functional of the expected conditional variance of the filter. The resulting optimal portfolio strategy is myopic and depends on the current drift estimate only. For power utility, the value function can be expressed as the expectation of the exponential of a quite involved integral functional of the conditional mean. It depends on the complete filter distribution and not only on its second-order moments. Further, the optimal strategies do not depend on the current drift estimate only but contain correction terms depending on the distribution of the future drift estimates.

In the portfolio problem one can use the dynamic programming approach for solving the associated stochastic optimal control problem. A solution can usually only be determined numerically. Diffusion approximations for the filter and the value function thus allow to find approximate solutions which can be given in closed form or at least derived with less numerical effort by solving a simplified control problem.

In the following example, we illustrate the convergence results from Corollary 9.4 and Corollary 9.5 by a numerical example. For that purpose, we compare the value function of the D-investor with the value function of the C-investor for various numbers of information dates.

Example 9.8. In Table 9.1a we list the value functions of the R-investor and of the D-investor as well as the value function of the C-investor in the setting with n equidistant information dates for different values of n. We assume that investors have initial capital $x_0 = 1$ and that the model parameters are those from Table 8.1, specifying additionally the volatility of the continuous expert as $\sigma_J = 0.2$. We see that the value functions $V^{C,n}(1)$ are increasing in n and approach the value $V^D(1)$ for large values of n.

Calculating the value function of the C-investor in the situation with non-deterministic information dates is a little more involved. This is because the conditional covariance matrices $(Q_t^{C,\lambda})_{t\in[0,T]}$ are then also non-deterministic. The value function, see Proposition 9.1, depends on the expectation of $Q_t^{C,\lambda}$ for $t \in [0,T]$. This value cannot be calculated easily. To determine the value function numerically we therefore perform for each value of λ a Monte Carlo simulation with 10 000 iterations. In each iteration, we generate a sequence of information dates as jump times of a Poisson process with intensity λ and calculate the corresponding conditional variances $(Q_t^{C,\lambda})_{t\in[0,T]}$. By taking an average of all simulations this leads to a good approximation of $V^{C,\lambda}(1)$. Table 9.1b shows the resulting estimations for $V^{C,\lambda}(1)$ and in brackets the corresponding 95% confidence intervals.

The values $V^{C,\lambda}(1)$ lie between $V^R(1)$ and $V^D(1)$, they are increasing in the intensity λ and for large values of λ they approach the value $V^D(1)$. This is in line with Corollary 9.5. We also observe that $V^{C,\lambda}(1) \le V^{C,n}(1)$ when setting the intensity λ equal to the deterministic number n. Recall that an intensity $\lambda = n$ means that there are on average n information dates in the time interval $[0,1]$. The randomness coming from the Poisson process however leads to a lower value function, compared to $V^{C,n}(1)$. This difference is negligible for large intensities.

H	n	$V^{H,n}(1)$
R		0.3410
C	10	0.5245
C	100	0.5511
C	1000	0.5531
C	10 000	0.5533
D		0.5533

(a) Equidistant information dates

H	λ	$V^{H,\lambda}(1)$
R		0.3410
C	10	0.5221 (0.5211, 0.5230)
C	100	0.5499 (0.5496, 0.5502)
C	1000	0.5530 (0.5529, 0.5531)
C	10 000	0.5533 (0.5533, 0.5533)
D		0.5533

(b) Random information dates

Table 9.1.: Value function for different investors.

This example shows that, when the number of discrete-time expert opinions is large, the value function of the C-investor can be approximated well by the value function of the D-investor. We see that it does not make a big difference whether the information dates are deterministic and equidistant or random as the jump times of a Poisson process. This approximation is useful since calculating the value function of the C-investor is numerically involved due to the updates at information dates. It becomes especially challenging if the information dates are non-deterministic. The value function of the D-investor, on the other hand, can be calculated much easier. Since it can be written as a functional of $(Q_t^D)_{t\in[0,T]}$, the essential part in the calculation is to solve one ordinary matrix Riccati differential equation.

Part III.

Robust Optimization with Expert Opinions

Outline and Notation

In this last part of the thesis we combine our results from the previous parts and show how uncertainty sets K for the robust utility maximization problem in Part I can be defined based on filters for the different investor filtrations that we have considered in Part II. The aim is to investigate the effect of expert opinions on robust strategies for the investors in the market.

The structure of this part is as follows. In Chapter 10 we generalize the financial market model from Part I to one with non-constant drift. We also allow for time-dependent uncertainty sets $(K_t)_{t\in[0,T]}$ then. If the sets K_t are adapted to the investor's filtration and have the form of ellipsoids, we can carry over the results from Chapter 4 and determine optimal trading strategies and worst-case drift processes.

Chapter 11 then explains how one can use filters to set up time-dependent uncertainty sets, motivated by confidence regions. In particular, the various investor filtrations from Part II are addressed. We show how expert opinions decrease the size of the uncertainty sets and investigate their effect on the corresponding robust strategies, compared to strategies that only rely on the respective drift estimation. A short conclusion of our results is given in Chapter 12.

Notation. For this part we adhere to the notation from the previous two parts. In particular, we write I_d for the identity matrix in $\mathbb{R}^{d\times d}$ as well as e_i, $i = 1, \dots, d$, for the i-th standard unit vector in $\mathbb{R}^d$ and $\mathbf{1}_d$ for the vector in $\mathbb{R}^d$ containing a one in every component. By $\langle \cdot, \cdot \rangle$ we denote the scalar product on $\mathbb{R}^d \times \mathbb{R}^d$ with $\langle x, y\rangle = x^\top y$ for $x, y \in \mathbb{R}^d$. Whenever $x \in \mathbb{R}^d$ is a vector, $\|x\|$ denotes the Euclidean norm of x.

For a symmetric and positive-semidefinite matrix $A \in \mathbb{R}^{d\times d}$ we write $A^{1/2}$ for the square root of A.

If $\mathcal{F}$ and $\mathcal{G}$ are σ-algebras, we write $\mathcal{F} \vee \mathcal{G}$ for the smallest σ-algebra containing $\mathcal{F} \cup \mathcal{G}$.

10. Generalized Duality Approach for Non-Constant Drift

In this chapter we generalize the approach from Chapter 4 to a financial market model where the drift is a stochastic process instead of a constant. To account for a change in information about the drift we also introduce time-dependence in the uncertainty set. The basic idea is that the available information in the market, for instance the observed asset returns and expert opinions as in Part II, are used to estimate the true drift based on filtering techniques and to set up a corresponding uncertainty set K_t at any time $t \in [0,T]$. Given K_t, investors then take model uncertainty into account by assuming that in the future the worst possible drift process $(\mu_s^{(t)})_{s\in[t,T]}$ with values in K_t will be realized. In our continuous-time setting the decision about the uncertainty set will be revised as soon as the information about the true drift changes, so in the extreme case continuously in time.

Before stating our generalized financial market, we make an observation that justifies the setup of the model. Suppose that the "true" dynamics of the d-dimensional return process R are given by

$$\mathrm{d}R_t = \mu_t \,\mathrm{d}t + \sigma \,\mathrm{d}W_t, \quad R_0 = 0,$$

for some stochastic drift process $(\mu_t)_{t\in[0,T]}$, an m-dimensional Brownian motion $(W_t)_{t\in[0,T]}$, $m \geq d$, and some $\sigma \in \mathbb{R}^{d\times m}$ with full rank. Assume further that the information of an investor is given by the investor filtration $\mathbb{F}^H = (\mathcal{F}_t^H)_{t\in[0,T]}$. The investor's best estimator for μ is then the conditional mean $\hat{\mu}_t := \mathbb{E}[\mu_t \,|\, \mathcal{F}_t^H]$ and one can rewrite the dynamics of the return process as

$$\mathrm{d}R_t = \hat{\mu}_t \,\mathrm{d}t + \sigma \,\mathrm{d}V_t,$$

where the so-called innovations process $(V_t)_{t\in[0,T]}$ is an $\mathbb{F}^H$-adapted Brownian motion. For instance, in the setting of Part II with $H = R$, the process $(\hat{\mu}_t)_{t\in[0,T]}$ would be the Kalman filter.

In the following, we set up our continuous-time financial market model working directly with the innovations process and therefore assuming an $\mathbb{F}^H$-adapted drift process. The separation principle that we use here by filtering first and then performing the optimization is a common approach for dealing with partial information.

10.1. Generalized financial market model

We fix an investment horizon $T > 0$ and some filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions. All processes are assumed to be $\mathbb{F}$-adapted. We assume that an investor's information is described by the investor filtration $\mathbb{F}^H = (\mathcal{F}_t^H)_{t\in[0,T]}$ with $\mathcal{F}_t^H \subseteq \mathcal{F}_t$ for all $t \in [0,T]$. We consider, as before, a financial market with one risk-free and $d \geq 2$ risky assets. The risk-free asset S^0 evolves as

$$\mathrm{d}S_t^0 = S_t^0 r \,\mathrm{d}t, \quad S_0^0 = 1,$$

where $r > 0$ is the deterministic risk-free interest rate. The returns $R^1, \ldots, R^d$ of the risky assets follow the dynamics

$$\mathrm{d}R_t = \nu_t \,\mathrm{d}t + \sigma \,\mathrm{d}W_t, \quad R_0 = 0, \tag{10.1}$$

where $R = (R^1, \ldots, R^d)^\top$. Here, $(W_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion under $\mathbb{P}$, $m \geq d$. Note that the volatility matrix $\sigma \in \mathbb{R}^{d\times m}$ in (10.1) is constant. Further, we assume that σ has full rank equal to d. In contrast to the volatility, the drift might change in the course of time. We assume that $(\nu_t)_{t\in[0,T]}$ is an $\mathbb{R}^d$-valued $\mathbb{F}^H$-adapted stochastic process and think of $(\nu_t)_{t\in[0,T]}$ as an estimation for the true drift process given all available information. We speak of $(\nu_t)_{t\in[0,T]}$ as the *reference drift*.

As before, we are concerned with investors who are uncertain about the true drift. They are aware that $(\nu_t)_{t\in[0,T]}$ in (10.1) might not be the true drift process. In utility maximization problems they want to maximize their worst-case expected utility, given that the true drift process is in a way “close” to ν. To model the uncertainty about the drift we specify the ellipsoidal sets

$$K_t = \bigl\{\mu \in \mathbb{R}^d \,\big|\, (\mu - \nu_t)^\top \Gamma_t^{-1} (\mu - \nu_t) \leq \kappa_t^2\bigr\}, \quad t \in [0, T],$$

where $(\Gamma_t)_{t\in[0,T]}$ is an $\mathbb{F}^H$-adapted stochastic process of symmetric and positive-definite matrices $\Gamma_t \in \mathbb{R}^{d\times d}$ and $(\kappa_t)_{t\in[0,T]}$ is $\mathbb{F}^H$-adapted with $\kappa_t > 0$ for each $t \in [0,T]$. The set K_t is determined at time $t \in [0,T]$ by taking the available information about the true drift process into account, for example based on filtering techniques. The process $(K_t)_{t\in[0,T]}$ is an $\mathbb{F}^H$-adapted set-valued process, therefore the investor knows the realization of K_t at time $t \in [0,T]$.

Given this K_t, investors then take model uncertainty into account by assuming that in the future the worst possible drift process having values in K_t will be realized. We denote this worst-case future drift by $(\mu_s^{(t),*})_{s\in[t,T]}$. This allows for some deterministic dynamics given K_t, i.e. the $\mu_s^{(t),*}$ for any $s \in [t,T]$ are $\mathcal{F}_t^H$-measurable. The worst-case optimization problem then leads to an optimal strategy $(\pi_s^{(t),*})_{s\in[t,T]}$, determined at time t. In our continuous-time setting this decision will be revised as soon as K_t changes, possibly continuously in time. The realized worst-case drift process $(\mu_t^*)_{t\in[0,T]}$ and optimal strategy $(\pi_t^*)_{t\in[0,T]}$ are then given by

$$\mu_t^* = \mu_t^{(t),*}, \quad \pi_t^* = \pi_t^{(t),*}$$

for any $t \in [0,T]$. If μ^* and π^* are uniquely determined, then they are by construction $\mathbb{F}^H$-adapted.

This is not so much a game setting but rather a way how the investor determines the worst case. It is a mixture of using estimation methods as in Part II and taking model uncertainty as in Part I into account.

The optimization problem can be derived only locally for each $t \in [0,T]$. In detail, the setup looks as follows. At time $t \in [0,T]$ investors assume that the future drift process will be the worst one within the class

$$\mathcal{K}^{(t)} = \bigl\{\mu^{(t)} = (\mu_s^{(t)})_{s\in[t,T]} \,\big|\, \mu_s^{(t)} \in K_t \text{ and } \mu_s^{(t)} \text{ is } \mathcal{F}_t^H\text{-measurable for each } s \in [t,T]\bigr\}.$$

For each $\mu = \mu^{(t)} \in \mathcal{K}^{(t)}$ we can construct a new measure by defining the $\mathbb{R}^m$-valued process $(\theta_s(\mu))_{s\in[0,T]}$ with

$$\theta_s(\mu) = \begin{cases} 0, & s < t, \\ \sigma^\top(\sigma\sigma^\top)^{-1}(\mu_s - \nu_s), & s \geq t, \end{cases}$$

and

$$Z_s^\mu = \exp\Big(\int_0^s \theta_u(\mu)^\top \,\mathrm{d}W_u - \frac{1}{2}\int_0^s \|\theta_u(\mu)\|^2 \,\mathrm{d}u\Big)$$

for $s \in [0,T]$. We then define the new probability measure $\mathbb{P}^\mu$ by

$$\frac{\mathrm{d}\mathbb{P}^\mu}{\mathrm{d}\mathbb{P}} = Z_T^\mu$$

and note that, under $\mathbb{P}^\mu$, the process $(W_s^\mu)_{s\in[0,T]}$ with

$$W_s^\mu = W_s - \int_0^s \theta_u(\mu)\,\mathrm{d}u$$

for $s \in [0,T]$ is a Brownian motion by Girsanov's Theorem. Note that due to boundedness of K_t the process $\theta(\mu)$ is bounded and therefore $(Z_s^\mu)_{s\in[0,T]}$ is a true martingale. The change of measure causes a change in the drift on the interval $[t,T]$ only. For our optimization problems this is the only relevant time interval since we condition on $\mathcal{F}_t^H$. For $s \in [t,T]$ we can rewrite the dynamics of the asset returns as

$$\mathrm{d}R_s = \nu_s\,\mathrm{d}s + \sigma\,\mathrm{d}W_s = \mu_s\,\mathrm{d}s + \sigma\,\mathrm{d}W_s^\mu,$$

which means that under $\mathbb{P}^\mu$ the future drift of the stocks is given by $(\mu_s)_{s\in[t,T]}$. We write $\mathbb{E}_\mu[\cdot] = \mathbb{E}_{\mu^{(t)}}[\cdot]$ for expectation under the measure $\mathbb{P}^\mu$.

An investor's behavior in the time interval $[t,T]$ is described by a self-financing trading strategy $\pi^{(t)} = (\pi_s^{(t)})_{s\in[t,T]}$. The class of admissible trading strategies, given that the investor has wealth $x > 0$ at time t, is

$$\begin{aligned}\mathcal{A}(t,x) = \Big\{\pi^{(t)} = (\pi_s^{(t)})_{s\in[t,T]} \;\Big|\; &\pi^{(t)} \text{ is } \mathbb{F}^H\text{-adapted},\; X_t^\pi = x, \\ &\mathbb{E}_{\mu^{(t)}}\Big[\int_t^T \|\sigma^\top \pi_s^{(t)}\|^2\,\mathrm{d}s\Big] < \infty \text{ for all } \mu^{(t)} \in \mathcal{K}^{(t)}\Big\}.\end{aligned}$$

We will restrict these strategies by imposing, as before, a constraint that prevents a pure bond investment. For any $h > 0$ we define the set

$$\mathcal{A}_h(t,x) = \big\{\pi^{(t)} \in \mathcal{A}(t,x) \,\big|\, \langle \pi_s^{(t)}, \mathbf{1}_d\rangle = h \text{ for all } s \in [t,T]\big\}.$$

For an investor choosing strategy $\pi = \pi^{(t)} \in \mathcal{A}(t, X_t^\pi)$ the terminal wealth can be written as

$$X_T^\pi = X_t^\pi \exp\Big(\int_t^T \Big(r + \pi_s^\top(\mu_s - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_s\|^2\Big)\mathrm{d}s + \int_t^T \pi_s^\top \sigma\,\mathrm{d}W_s^\mu\Big).$$

We are now able to state our utility maximization problem. At time t the local optimization problem reads

$$\sup_{\pi^{(t)}\in\mathcal{A}_h(t,X_t^\pi)} \inf_{\mu^{(t)}\in\mathcal{K}^{(t)}} \mathbb{E}_{\mu^{(t)}}\Big[U_\gamma\big(X_T^{\pi^{(t)}}\big) \,\Big|\, \mathcal{F}_t^H\Big]. \tag{10.2}$$

Here, U_γ with $\gamma \in (-\infty, 1)$ again denotes the power utility function $U_\gamma(x) = \frac{x^\gamma}{\gamma}$ if $\gamma \neq 0$, and logarithmic utility $U_0(x) = \log(x)$ if $\gamma = 0$.

Remark 10.1. In the case where $K_t = \{\mu \in \mathbb{R}^d \,|\, (\mu - \nu)^\top \Gamma^{-1}(\mu - \nu) \leq \kappa^2\}$ for all $t \in [0, T]$, i.e. where our reference drift is simply a constant ν, and also the matrix $\Gamma_t = \Gamma$ as well as the radius $\kappa_t = \kappa$ are constant in time, we obtain the setting from Part I as a special case.

10.2. Solution of the non-robust problem

As a first step towards solving (10.2) we compute the optimal strategy for an investor given a particular future drift $\mu^{(t)} \in \mathcal{K}^{(t)}$.

Proposition 10.2. *Let $t \in [0, T]$ and $\mu^{(t)} \in \mathcal{K}^{(t)}$. Then the optimal strategy for the optimization problem*

$$\sup_{\pi^{(t)} \in \mathcal{A}_h(t, X_t^\pi)} \mathbb{E}_{\mu^{(t)}}\Big[U_\gamma\big(X_T^{\pi^{(t)}}\big) \,\Big|\, \mathcal{F}_t^H\Big]$$

is the strategy $(\pi_s^{(t)})_{s\in[t,T]}$ with

$$\pi_s^{(t)} = \frac{1}{1-\gamma} A\mu_s^{(t)} + hc$$

for all $s \in [t, T]$, where $A \in \mathbb{R}^{d\times d}$ and $c \in \mathbb{R}^d$ are as introduced in Definition 4.2.

Proof. The proof works along the lines of the proof of Proposition 4.3. We take an arbitrary strategy $\pi = \pi^{(t)} \in \mathcal{A}_h(t, X_t^\pi)$ and recall that we can write the terminal wealth under strategy π as

$$X_T^\pi = X_t^\pi \exp\bigg(\int_t^T \Big(r + \pi_s^\top(\mu_s^{(t)} - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top \pi_s\|^2\Big)\mathrm{d}s + \int_t^T \pi_s^\top \sigma \,\mathrm{d}W_s^\mu\bigg).$$

We now proceed exactly as in the proof of Proposition 4.3, replacing the constant μ by the $\mathcal{F}_t^H$-measurable $(\mu_s^{(t)})_{s\in[t,T]}$, and perform the same transformation to a $(d-1)$-dimensional unconstrained financial market.

We can deduce that $\mathbb{E}_{\mu^{(t)}}[U_\gamma(X_T^\pi) \,|\, \mathcal{F}_t^H]$ equals the expected utility of terminal wealth, conditional on $\mathcal{F}_t^H$, in an unconstrained financial market with $d-1$ risky assets, where the future drift process is $(\widetilde{\mu}_s)_{s\in[t,T]}$, the risk-free interest rate is $(\widetilde{r}_s)_{s\in[t,T]}$ and the volatility matrix is $\widetilde{\sigma} \in \mathbb{R}^{(d-1)\times m}$. These transformed market parameters have the form

$$\begin{aligned}
\widetilde{\sigma} &= D\sigma,\\
\widetilde{r}_s &= (1-h)r + he_d^\top \mu_s^{(t)} - \frac{1}{2}(1-\gamma)\|h\sigma^\top e_d\|^2,\\
\widetilde{\mu}_s &= D\mu_s^{(t)} - h(1-\gamma)D\sigma\sigma^\top e_d + \widetilde{r}_s\mathbf{1}_{d-1}.
\end{aligned}$$

Note that since the $(\mu_s^{(t)})_{s\in[t,T]}$ are $\mathcal{F}_t^H$-measurable, so are $(\widetilde{r}_s)_{s\in[t,T]}$ and $(\widetilde{\mu}_s)_{s\in[t,T]}$, in particular the market parameters in the transformed market can be observed by the investor. In this $(d-1)$-dimensional unconstrained financial market we know that the optimal strategy is of the form

$$\widetilde{\pi}_s = \frac{1}{1-\gamma}(\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}(\widetilde{\mu}_s - \widetilde{r}_s\mathbf{1}_{d-1}) = \frac{1}{1-\gamma}(D\sigma\sigma^\top D^\top)^{-1}\big(D\mu_s^{(t)} - h(1-\gamma)D\sigma\sigma^\top e_d\big)$$

for every $s \in [t,T]$. For the logarithmic utility case, this is immediate, for power utility see Appendix B. Now we can return to our original market and obtain that the optimal strategy fulfills

$$\begin{aligned}
\pi_s^{(t)} &= D^\top \widetilde{\pi}_s + he_d \\
&= D^\top \frac{1}{1-\gamma}(D\sigma\sigma^\top D^\top)^{-1}\big(D\mu_s^{(t)} - h(1-\gamma)D\sigma\sigma^\top e_d\big) + he_d \\
&= \frac{1}{1-\gamma} D^\top (D\sigma\sigma^\top D^\top)^{-1} D\mu_s^{(t)} + h\big(I_d - D^\top (D\sigma\sigma^\top D^\top)^{-1} D\sigma\sigma^\top\big)e_d \\
&= \frac{1}{1-\gamma} A\mu_s^{(t)} + hc
\end{aligned}$$

for all $s \in [t,T]$, where we have used the notation for A and c from Definition 4.2. Note that $(\pi_s^{(t)})_{s\in[t,T]}$ is indeed admissible due to boundedness of K_t. $\square$

The preceding proposition states the form of the investor's optimal strategy under the assumption that a specific future drift process $(\mu_s^{(t)})_{s\in[t,T]}$ is given. The explicit form can be used to compute also the expected utility obtained when applying the optimal strategy.

Corollary 10.3. *Let $t \in [0,T]$ and $\mu^{(t)} \in \mathcal{K}^{(t)}$. Then the optimal expected utility from terminal wealth is*

$$\begin{aligned}
&\sup_{\pi^{(t)} \in \mathcal{A}_h(t, X_t^\pi)} \mathbb{E}_{\mu^{(t)}}\Big[U_\gamma\big(X_T^{\pi^{(t)}}\big)\,\Big|\,\mathcal{F}_t^H\Big] \\
&= \begin{cases}
\dfrac{(X_t^\pi)^\gamma}{\gamma}\exp\Big(\gamma\displaystyle\int_t^T\Big(\widetilde{r}_s + \frac{1}{2(1-\gamma)}\big(\widetilde{\mu}_s - \widetilde{r}_s\mathbf{1}_{d-1}\big)^\top(\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}\big(\widetilde{\mu}_s - \widetilde{r}_s\mathbf{1}_{d-1}\big)\Big)\mathrm{d}s\Big), & \gamma \neq 0, \\
\log(X_t^\pi) + \displaystyle\int_t^T\Big(\widetilde{r}_s + \frac{1}{2}\big(\widetilde{\mu}_s - \widetilde{r}_s\mathbf{1}_{d-1}\big)^\top(\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}\big(\widetilde{\mu}_s - \widetilde{r}_s\mathbf{1}_{d-1}\big)\Big)\mathrm{d}s, & \gamma = 0,
\end{cases}
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{\sigma} &= D\sigma, \\
\widetilde{r}_s &= (1-h)r + he_d^\top\mu_s^{(t)} - \frac{1}{2}(1-\gamma)\|h\sigma^\top e_d\|^2, \\
\widetilde{\mu}_s &= D\mu_s^{(t)} - h(1-\gamma)D\sigma\sigma^\top e_d + \widetilde{r}_s\mathbf{1}_{d-1}.
\end{aligned}$$

Proof. The representation in the corollary follows, just like in the proof of Corollary 4.4, by the fact that we have reduced our constrained utility maximization problem to a $(d-1)$-dimensional unconstrained problem where the parameters of our transformed financial market are exactly those that are listed in the corollary. We have seen that the optimal strategy in this $(d-1)$-dimensional market fulfills

$$\widetilde{\pi}_s = \frac{1}{1-\gamma}(\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}(\widetilde{\mu}_s - \widetilde{r}_s\mathbf{1}_{d-1})$$

for all $s \in [t,T]$. Plugging this optimal strategy in yields the expression from the corollary. $\square$

10.3. The worst-case drift process

In the following, we compute the worst-case future drift process that is determined at time $t \in [0,T]$, i.e. the drift process $\mu^{(t)} \in \mathcal{K}^{(t)}$ for which

$$\sup_{\pi^{(t)} \in \mathcal{A}_h(t, X_t^\pi)} \mathbb{E}_{\mu^{(t)}}\Big[U_\gamma\big(X_T^{\pi^{(t)}}\big) \,\Big|\, \mathcal{F}_t^H\Big]$$

is minimized. Due to the previous corollary we see that this is equivalent to the minimization of the integral

$$\int_t^T \Big(\widetilde{r}_s + \frac{1}{2}\big(\widetilde{\mu}_s - \widetilde{r}_s \mathbf{1}_{d-1}\big)^\top (\widetilde{\sigma}\widetilde{\sigma}^\top)^{-1}\big(\widetilde{\mu}_s - \widetilde{r}_s \mathbf{1}_{d-1}\big)\Big)\mathrm{d}s. \tag{10.3}$$

When plugging the representations for $\widetilde{\mu}$, $\widetilde{r}$ and $\widetilde{\sigma}$ back in, we obtain an expression that depends on $(\mu_s^{(t)})_{s\in[t,T]}$ again. By the same calculations as in the setting with constant drift we deduce that minimizing (10.3) is equivalent to minimizing

$$\int_t^T \Big(\frac{1}{2(1-\gamma)}(\mu_s^{(t)})^\top A \mu_s^{(t)} + hc^\top \mu_s^{(t)}\Big)\mathrm{d}s.$$

But the minimization of this integral is equivalent to a pointwise minimization of

$$K_t \ni \mu \mapsto \frac{1}{2(1-\gamma)}\mu^\top A\mu + hc^\top \mu.$$

Now it is straightforward to see that we can use our results from Section 4.1 to obtain the worst-case drift process $(\mu_s^{(t),*})_{s\in[t,T]}$. Here, $\mu_s^{(t),*}$ is for any $s \in [t,T]$ obtained as the minimizer of the above function on K_t. Recall that the uncertainty set is an ellipsoid of the form $K_t = \{\mu \in \mathbb{R}^d \,|\, (\mu - \nu_t)^\top \Gamma_t^{-1}(\mu - \nu_t) \leq \kappa_t^2\}$. We have assumed that Γ_t is a symmetric positive-definite matrix in $\mathbb{R}^{d\times d}$. In the following we use the representation $\Gamma_t = \tau_t \tau_t^\top$ where $\tau_t \in \mathbb{R}^{d\times d}$ is a nonsingular matrix.

Corollary 10.4. *We fix some $t \in [0,T]$ and let $0 = \lambda_{t,1} < \lambda_{t,2} \leq \cdots \leq \lambda_{t,d}$ denote the eigenvalues of $\tau_t^\top A \tau_t$, and*

$$v_{t,1} = \frac{1}{\|\tau_t^{-1}\mathbf{1}_d\|}\tau_t^{-1}\mathbf{1}_d, v_{t,2}, \ldots, v_{t,d} \in \mathbb{R}^d$$

the respective orthogonal eigenvectors with $\|v_{t,i}\| = 1$ for all $i = 1, \ldots, d$. Then

$$\inf_{\mu^{(t)} \in \mathcal{K}^{(t)}} \sup_{\pi^{(t)} \in \mathcal{A}_h(t, X_t^\pi)} \mathbb{E}_{\mu^{(t)}}\Big[U_\gamma\big(X_T^{\pi^{(t)}}\big) \,\Big|\, \mathcal{F}_t^H\Big] = \mathbb{E}_{\mu^{(t),*}}\Big[U_\gamma\big(X_T^{\pi^{(t),*}}\big) \,\Big|\, \mathcal{F}_t^H\Big],$$

where

$$\mu_s^{(t),*} = \nu_t - \tau_t \sum_{i=1}^d \bigg(\frac{\lambda_{t,i}}{1-\gamma} + \frac{h}{\psi_t(\kappa_t)\|\tau_t^{-1}\mathbf{1}_d\|}\bigg)^{-1} \Big\langle h\tau_t^\top c + \frac{\lambda_{t,i}}{1-\gamma}\tau_t^{-1}\nu_t, v_{t,i}\Big\rangle v_{t,i}$$

for all $s \in [t,T]$, and where $\psi_t(\kappa_t) \in (0, \kappa_t]$ is uniquely determined by $\|\tau_t^{-1}(\mu_s^{(t),*} - \nu_t)\| = \kappa_t$. The strategy $(\pi_s^{(t),*})_{s\in[t,T]}$ has the form

$$\pi_s^{(t),*} = \frac{1}{1-\gamma}A\mu_s^{(t),*} + hc$$

for all $s \in [t,T]$.

Proof. We have seen that the worst-case drift process $(\mu_s^{(t),*})_{s\in[t,T]}$ is the one where $\mu_s^{(t),*}$ is for any $s \in [t,T]$ equal to the minimizer of the function

$$\mu \mapsto \frac{1}{2(1-\gamma)}\mu^\top A\mu + hc^\top \mu$$

over all $\mu \in K_t$. So we can do the minimization as in Section 4.1. We know that the matrix $\tau_t^\top A\tau_t \in \mathbb{R}^{d\times d}$ is symmetric and positive definite with

$$\ker(\tau_t^\top A\tau_t) = \operatorname{span}(\{\tau_t^{-1}\mathbf{1}_d\}).$$

Now the representation of $\mu_s^{(t),*}$ follows as in Theorem 4.8 with Lemma 4.7. The form of the optimal strategy $\pi^{(t),*}$ then follows from Proposition 10.2. $\square$

The preceding corollary shows that the problem

$$\inf_{\mu^{(t)}\in\mathcal{K}^{(t)}} \sup_{\pi^{(t)}\in\mathcal{A}_h(t,X_t^\pi)} \mathbb{E}_{\mu^{(t)}}\Big[U_\gamma\big(X_T^{\pi^{(t)}}\big)\,\Big|\,\mathcal{F}_t^H\Big]$$

is solved by drift process $(\mu_s^{(t),*})_{s\in[t,T]}$ and strategy $(\pi_s^{(t),*})_{s\in[t,T]}$. Note that both the worst-case drift process and the optimal strategy are constant on $[t,T]$ and $\mathcal{F}_t^H$-measurable. This is due to the setup of the model in which investors assume that the future drift process will take values in the ellipsoid K_t only.

The problem above is the dual to our original problem

$$\sup_{\pi^{(t)}\in\mathcal{A}_h(t,X_t^\pi)} \inf_{\mu^{(t)}\in\mathcal{K}^{(t)}} \mathbb{E}_{\mu^{(t)}}\Big[U_\gamma\big(X_T^{\pi^{(t)}}\big)\,\Big|\,\mathcal{F}_t^H\Big].$$

To ensure that $\mu^{(t),*}$ and $\pi^{(t),*}$ are also a solution to this problem we have to show that $\mu^{(t),*}$ is the worst drift process in the set $\mathcal{K}^{(t)}$, given that an investor chooses trading strategy $\pi^{(t),*}$. In that case, the infimum and the supremum interchange and we can deduce that $\pi^{(t),*}$ and $\mu^{(t),*}$ also establish a solution to our original robust optimization problem.

10.4. A minimax theorem

We proceed as in Section 4.1 and note that the strategy $\pi^{(t),*}$ from the previous corollary satisfies

$$\pi_s^{(t),*} = -\frac{h}{\psi_t(\kappa_t)\|\tau_t^{-1}\mathbf{1}_d\|}\Gamma_t^{-1}\big(\mu_s^{(t),*} - \nu_t\big)$$

for all $s \in [t,T]$. This can be proven by analogy with Lemma 4.9. This observation helps to prove the following proposition.

Proposition 10.5. *The drift process $(\mu_s^{(t)})_{s\in[t,T]}$ that attains the minimum in*

$$\inf_{\mu^{(t)}\in\mathcal{K}^{(t)}} \mathbb{E}_{\mu^{(t)}}\Big[U_\gamma\big(X_T^{\pi^{(t),*}}\big)\,\Big|\,\mathcal{F}_t^H\Big]$$

is $(\mu_s^{(t),*})_{s\in[t,T]}$, i.e. $\mu^{(t),*}$ is the worst possible drift process, given that an investor chooses the strategy $\pi^{(t),*}$.

Proof. We take an arbitrary $\mu = \mu^{(t)} \in \mathcal{K}^{(t)}$. Note that in case $\gamma \neq 0$ we can write

$$\begin{aligned}
&\mathbb{E}_\mu\Big[U_\gamma\big(X_T^{\pi^{(t),*}}\big)\,\Big|\,\mathcal{F}_t^H\Big] \\
&= \frac{(X_t^\pi)^\gamma}{\gamma}\mathrm{e}^{\gamma r(T-t)}\,\mathbb{E}_\mu\Big[\exp\Big(\gamma\int_t^T\Big((\pi_s^{(t),*})^\top(\mu_s - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_s^{(t),*}\|^2\Big)\mathrm{d}s + \gamma\int_t^T(\pi_s^{(t),*})^\top\sigma\,\mathrm{d}W_s^\mu\Big)\Big] \\
&= \frac{(X_t^\pi)^\gamma}{\gamma}\mathrm{e}^{\gamma r(T-t)}\exp\Big(\gamma\int_t^T\Big((\pi_s^{(t),*})^\top(\mu_s - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\pi_s^{(t),*}\|^2\Big)\mathrm{d}s\Big).
\end{aligned}$$

In case $\gamma = 0$ we have

$$\mathbb{E}_\mu\Big[\log\big(X_T^{\pi^{(t),*}}\big)\,\Big|\,\mathcal{F}_t^H\Big] = \log(X_t^\pi) + r(T-t) + \int_t^T\Big((\pi_s^{(t),*})^\top(\mu_s - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_s^{(t),*}\|^2\Big)\mathrm{d}s.$$

In both cases, the drift process $(\mu_s)_{s\in[t,T]} \in \mathcal{K}^{(t)}$ that minimizes this expression is the one that minimizes

$$\int_t^T (\pi_s^{(t),*})^\top\mu_s\,\mathrm{d}s.$$

Since $(\pi_s^{(t),*})_{s\in[t,T]}$ is constant, we find the minimizer as the minimizer of $(\pi_s^{(t),*})^\top\mu_s$. Recall that

$$\pi_s^{(t),*} = -\frac{h}{\psi_t(\kappa_t)\|\tau_t^{-1}\mathbf{1}_d\|}\Gamma_t^{-1}\big(\mu_s^{(t),*} - \nu_t\big).$$

It follows that

$$(\pi_s^{(t),*})^\top\Gamma_t\pi_s^{(t),*} = \frac{h^2}{\psi_t(\kappa_t)^2\|\tau_t^{-1}\mathbf{1}_d\|^2}\big(\mu_s^{(t),*} - \nu_t\big)^\top\Gamma_t^{-1}\big(\mu_s^{(t),*} - \nu_t\big) = \frac{h^2\kappa_t^2}{\psi_t(\kappa_t)^2\|\tau_t^{-1}\mathbf{1}_d\|^2}.$$

Knowing that $\psi_t(\kappa_t) > 0$ we can deduce

$$\sqrt{(\pi_s^{(t),*})^\top\Gamma_t\pi_s^{(t),*}} = \frac{h\kappa_t}{\psi_t(\kappa_t)\|\tau_t^{-1}\mathbf{1}_d\|}.$$

The drift process $\mu_s^{(t),*}$ at time s can thus be rewritten in the form

$$\mu_s^{(t),*} = \nu_t - \frac{\psi_t(\kappa_t)\|\tau_t^{-1}\mathbf{1}_d\|}{h}\Gamma_t\pi_s^{(t),*} = \nu_t - \frac{\kappa_t}{\sqrt{(\pi_s^{(t),*})^\top\Gamma_t\pi_s^{(t),*}}}\Gamma_t\pi_s^{(t),*}.$$

This is exactly the vector that minimizes $(\pi_s^{(t),*})^\top\mu$ over all $\mu \in K_t$, see the proof of Proposition 4.10. Hence, $\mu^{(t),*}$ is the drift process that minimizes the expected utility of terminal wealth for an investor who chooses strategy $\pi^{(t),*}$. □

The previous proposition establishes an equilibrium result. By definition, the strategy $\pi^{(t),*}$ is optimal for the drift $\mu^{(t),*}$. Due to the proposition, it also holds that $\mu^{(t),*}$ is the worst drift given that an investor chooses strategy $\pi^{(t),*}$. Hence, we see that $(\pi^{(t),*}, \mu^{(t),*})$ is a saddle point of the optimization problem

$$\sup_{\pi^{(t)}\in\mathcal{A}_h(t,X_t^\pi)}\ \inf_{\mu^{(t)}\in\mathcal{K}^{(t)}}\ \mathbb{E}_{\mu^{(t)}}\Big[U_\gamma\big(X_T^{\pi^{(t)}}\big)\,\Big|\,\mathcal{F}_t^H\Big].$$

In particular, the supremum and infimum can be interchanged. We obtain the following minimax theorem.

Theorem 10.6. *Let $t \in [0,T]$. Then*

$$\sup_{\pi^{(t)} \in \mathcal{A}_h(t, X_t^\pi)} \inf_{\mu^{(t)} \in \mathcal{K}^{(t)}} \mathbb{E}_{\mu^{(t)}} \left[U_\gamma\big(X_T^{\pi^{(t)}}\big) \,\middle|\, \mathcal{F}_t^H \right] = \mathbb{E}_{\mu^{(t),*}} \left[U_\gamma\big(X_T^{\pi^{(t),*}}\big) \,\middle|\, \mathcal{F}_t^H \right]$$
$$= \inf_{\mu^{(t)} \in \mathcal{K}^{(t)}} \sup_{\pi^{(t)} \in \mathcal{A}_h(t, X_t^\pi)} \mathbb{E}_{\mu^{(t),*}} \left[U_\gamma\big(X_T^{\pi^{(t),*}}\big) \,\middle|\, \mathcal{F}_t^H \right],$$

where $\mu^{(t),*}$ and $\pi^{(t),*}$ are defined as in Corollary 10.4.

Proof. The proof is analogous to the proof of Theorem 4.11. □

The previous theorem solves our original local optimization problem (10.2) for a fixed time $t \in [0,T]$. It shows that the best strategy for an investor in this robust optimization problem is the strategy $(\pi_s^{(t),*})_{s\in[t,T]}$ with

$$\pi_s^{(t),*} = \frac{1}{1-\gamma} A \mu_s^{(t),*} + hc$$

for all $s \in [t,T]$, where $(\mu_s^{(t),*})_{s\in[t,T]}$ is defined as in Corollary 10.4. The process $(\mu_s^{(t),*})_{s\in[t,T]}$ can be interpreted as the worst possible realization of the future drift process from the investor's point of view at time t. The worst-case drift and optimal strategy in this setting are constant on $[t,T]$. This is due to the assumption of the investor that the future drift will take values in the set K_t only, where K_t is determined at time t using all available information, i.e. K_t is $\mathcal{F}_t^H$-measurable.

In our continuous-time setting it is likely that the information about the unobservable true drift process changes continuously, therefore also the uncertainty set K_t will be updated continuously in time. At each time $t \in [0,T]$, the investor will revise both the uncertainty set and the optimization problem

$$\sup_{\pi^{(t)} \in \mathcal{A}_h(t, X_t^\pi)} \inf_{\mu^{(t)} \in \mathcal{K}^{(t)}} \mathbb{E}_{\mu^{(t)}} \left[U_\gamma\big(X_T^{\pi^{(t)}}\big) \,\middle|\, \mathcal{F}_t^H \right].$$

The strategy that is realized by the investor can then be found as $(\pi_t^*)_{t\in[0,T]}$ with

$$\pi_t^* = \pi_t^{(t),*}$$

for any $t \in [0,T]$. It has the form

$$\pi_t^* = \frac{1}{1-\gamma} A \mu_t^* + hc$$

where $(\mu_t^*)_{t\in[0,T]}$ is constructed via

$$\mu_t^* = \mu_t^{(t),*}$$

for all $t \in [0,T]$. Note that the processes $(\mu_t^*)_{t\in[0,T]}$ and $(\pi_t^*)_{t\in[0,T]}$ are uniquely determined, $\mathbb{F}^H$-adapted and in general non-constant. In the special case where $K_t = K_0$ for all $t \in [0,T]$, i.e. where our reference drift is simply a constant ν, and also the matrix $\Gamma_t = \Gamma$ as well as the radius $\kappa_t = \kappa$ are constant in time, also $(\mu_t^*)_{t\in[0,T]}$ and $(\pi_t^*)_{t\in[0,T]}$ are constant in time. The constant values are the ones that we also get in the setting with constant drift and uncertainty set in Theorem 4.8.

11. Construction of Uncertainty Sets via Filters

11.1. Confidence regions as uncertainty sets

In the preceding chapter we have seen how the duality approach from Chapter 4 carries over to a financial market where the drift is not necessarily constant. The generalized model allows for local uncertainty sets of the form

$$K_t = \big\{\mu \in \mathbb{R}^d \,\big|\, (\mu - \nu_t)^\top \Gamma_t^{-1}(\mu - \nu_t) \le \kappa_t^2\big\}, \quad t \in [0,T].$$

We have fixed an investor filtration $\mathbb{F}^H = (\mathcal{F}_t^H)_{t\in[0,T]}$ describing the investor's information in the course of time. Our model then assumes that the processes $\nu = (\nu_t)_{t\in[0,T]}$, $\Gamma = (\Gamma_t)_{t\in[0,T]}$ and $\kappa = (\kappa_t)_{t\in[0,T]}$ are $\mathbb{F}^H$-adapted. Recall that ν takes values in $\mathbb{R}^d$, Γ in the set of symmetric and positive-definite matrices in $\mathbb{R}^{d\times d}$ and κ on the positive real line.

We motivated the reference drift ν as an estimation for the true drift, based on the information available to the investor. Here we want to make this more specific by considering the filter. Recall that the filter is the conditional distribution of μ given the available information $\mathbb{F}^H$. We take ν to be the conditional expectation of the drift given $\mathbb{F}^H$, i.e. $\nu_t = m_t^H = \mathbb{E}[\mu_t \,|\, \mathcal{F}_t^H]$ for every $t \in [0,T]$. The conditional covariance matrix

$$Q_t^H = \mathbb{E}\big[(\mu_t - m_t^H)(\mu_t - m_t^H)^\top \,\big|\, \mathcal{F}_t^H\big]$$

measures how close the estimator m_t^H is to the true drift. Note that by construction both m^H and Q^H are $\mathbb{F}^H$-adapted processes. The key idea for constructing uncertainty sets based on the filter is to create confidence regions centered around m_t^H, shaped by Q_t^H for every $t \in [0,T]$.

Let us assume that the drift process and the investor filtration are such that the filter is normally distributed, more precisely

$$\mu_t \,|\, \mathcal{F}_t^H \sim \mathcal{N}(m_t^H, Q_t^H).$$

By applying a simple transformation we deduce that

$$(\mu_t - m_t^H)^\top (Q_t^H)^{-1}(\mu_t - m_t^H)$$

given $\mathcal{F}_t^H$ is χ^2-distributed with d degrees of freedom. We fix some $\eta \in (0,1)$ and observe that a $(1-\eta)$-confidence region can be obtained from

$$1 - \eta = \mathbb{P}\Big((\mu_t - m_t^H)^\top (Q_t^H)^{-1}(\mu_t - m_t^H) \le \chi^2_{d,1-\eta} \,\Big|\, \mathcal{F}_t^H\Big).$$

Here, $\chi^2_{d,1-\eta}$ denotes the $(1-\eta)$-quantile of the χ^2-distribution with d degrees of freedom. This motivates the choice of

$$K_t = \{\mu \in \mathbb{R}^d \,|\, (\mu - m_t^H)^\top (Q_t^H)^{-1}(\mu - m_t^H) \leq \chi^2_{d,1-\eta}\}, \quad t \in [0,T],$$

i.e. taking $\nu_t = m_t^H$, $\Gamma_t = Q_t^H$ and $\kappa_t = \sqrt{\chi^2_{d,1-\eta}}$ for every $t \in [0,T]$.

If indeed μ_t given $\mathcal{F}_t^H$ is normally distributed, we additionally know that at any fixed time $t \in [0,T]$ the probability that $\mu_t \in K_t$, conditional on $\mathcal{F}_t^H$, is equal to $1-\eta$. Note that K_t is still a reasonable uncertainty set for μ_t in the case where the assumption about the normal distribution of the filter is not fulfilled.

11.2. Uncertainty sets based on expert opinions

The preceding section explains how time-dependent uncertainty sets can be created based on filters. We now apply this to the various investor filtrations that we have considered in Part II for a model with an unobservable Ornstein–Uhlenbeck drift process and unbiased, normally distributed expert opinions arriving at discrete points in time. Recall that returns in this setting are modelled as

$$\mathrm{d}R_t = \mu_t \,\mathrm{d}t + \sigma_R \,\mathrm{d}W_t^R,$$

where $W^R = (W_t^R)_{t\in[0,T]}$ is an m-dimensional Brownian motion with $m \geq d$ and where we assume that $\sigma_R \in \mathbb{R}^{d\times m}$ has full rank. The drift process μ is defined by the Ornstein–Uhlenbeck dynamics

$$\mathrm{d}\mu_t = \alpha(\delta - \mu_t)\,\mathrm{d}t + \beta \,\mathrm{d}B_t,$$

where α and $\beta \in \mathbb{R}^{d\times d}$, $\delta \in \mathbb{R}^d$ and $B = (B_t)_{t\in[0,T]}$ is a d-dimensional Brownian motion that is independent of W^R. The matrices α and $\beta\beta^\top$ are assumed to be symmetric and positive definite. We further make the assumption that $\mu_0 \sim \mathcal{N}(m_0, \Sigma_0)$ for some $m_0 \in \mathbb{R}^d$ and some symmetric and positive-semidefinite matrix $\Sigma_0 \in \mathbb{R}^{d\times d}$, and that μ_0 is independent of the Brownian motions W^R and B, i.e. μ is independent of W^R.

Recall that the discrete-time expert opinions arrive at the information dates $(T_k)_{k\in I}$ and that an expert opinion at time T_k is of the form

$$Z_k = \mu_{T_k} + (\Gamma_k)^{1/2}\varepsilon_k,$$

where the matrices $\Gamma_k \in \mathbb{R}^{d\times d}$ are symmetric and positive definite and the ε_k are multivariate $\mathcal{N}(0, I_d)$-distributed and independent of the Brownian motions in the market and of μ_0. The sequence of information dates $(T_k)_{k\in I}$ is also independent of the $(\varepsilon_k)_{k\in I}$ and the Brownian motions as well as of μ_0. In particular, given μ_{T_k} the expert opinion is multivariate $\mathcal{N}(\mu_{T_k}, \Gamma_k)$-distributed.

In Section 8.2 we have proven that the information from observing a suitable sequence of increasingly frequent expert opinions converges to the information an investor gets from observing a certain diffusion process, interpreted as a continuous-time expert. Recall that this diffusion is of the form

$$\mathrm{d}J_t = \mu_t \,\mathrm{d}t + \sigma_J \,\mathrm{d}W_t^J,$$

where W^J is an l-dimensional Brownian motion with $l \geq d$ that is independent of all other Brownian motions in the model, of μ_0 and of the information dates T_k, and where the matrix $\sigma_J \in \mathbb{R}^{d\times l}$ has full rank equal to d. We also include this diffusion here and consider the corresponding investor filtration as a limit case.

The model then gives rise to various investor filtrations $\mathbb{F}^H = (\mathcal{F}_t^H)_{t\in[0,T]}$ where H serves as a placeholder for the various information regimes. We consider as before the cases

$$\begin{aligned}
\mathbb{F}^R &= (\mathcal{F}_t^R)_{t\in\mathbb{T}} \text{ where } \mathcal{F}_t^R = \sigma((R_s)_{s\in[0,t]}) \vee \sigma(\mathcal{N}_{\mathbb{P}}),\\
\mathbb{F}^E &= (\mathcal{F}_t^E)_{t\in\mathbb{T}} \text{ where } \mathcal{F}_t^E = \sigma((T_k, Z_k)_{T_k\leq t}) \vee \sigma(\mathcal{N}_{\mathbb{P}}),\\
\mathbb{F}^C &= (\mathcal{F}_t^C)_{t\in\mathbb{T}} \text{ where } \mathcal{F}_t^C = \sigma((R_s)_{s\in[0,t]}) \vee \sigma((T_k, Z_k)_{T_k\leq t}) \vee \sigma(\mathcal{N}_{\mathbb{P}}),\\
\mathbb{F}^D &= (\mathcal{F}_t^D)_{t\in\mathbb{T}} \text{ where } \mathcal{F}_t^D = \sigma((R_s)_{s\in[0,t]}) \vee \sigma((J_s)_{s\in[0,t]}) \vee \sigma(\mathcal{N}_{\mathbb{P}}),\\
\mathbb{F}^F &= (\mathcal{F}_t^F)_{t\in\mathbb{T}} \text{ where } \mathcal{F}_t^F = \sigma((R_s)_{s\in[0,t]}) \vee \sigma((\mu_s)_{s\in[0,t]}) \vee \sigma(\mathcal{N}_{\mathbb{P}})
\end{aligned}$$

for the investor filtrations. Recall that we write $\mathcal{N}_{\mathbb{P}}$ for the set of null sets under $\mathbb{P}$, i.e. we work with the filtrations that are augmented by null sets.

Based on one realization of the model's stochastic processes, fixing one information setting $H \in \{R, E, C, D, F\}$, we obtain one realization of the filter, leading to a time-dependent uncertainty set K^H. In Figure 11.1 various such filters with the resulting uncertainty sets are plotted. For illustration purposes we take a market with $d = 1$ stock here. The market parameters are the same as for our earlier numerical example, given in Table 8.1, together with $\sigma_J = 0.2$ for the volatility of the continuous-time expert, and we choose $\eta = 0.1$ to construct the confidence regions.

The various subplots are all based on the same realization of the drift process μ, returns R and expert opinions Z_k. As a first case we consider in Figure 11.1a the degenerate information setting $H = E$ with $n = 0$ expert opinions, corresponding to an investor who observes neither the diffusion processes nor the discrete-time expert opinions. The only knowledge the investor has about the model are the model parameters. With our choice of $m_0 = \delta$ it follows immediately from Lemma 6.3 that the conditional mean is in this case constantly equal to δ. The resulting uncertainty set converges very fast to a fixed interval centered around δ.

For $H = R$, the uncertainty set moves up and down along with the conditional mean as can be seen in Figure 11.1b. In Figures 11.1c and 11.1d we have $n = 10$ equidistant information dates with expert opinions. The corresponding uncertainty set jumps at information dates along with the conditional mean, due to the updates caused by an incoming expert opinion. It also becomes apparent from the plots that the conditional variance decreases at information dates, leading to a shrinking uncertainty set. The case $H = D$ is depicted in Figure 11.1e. We observe that the uncertainty set is at any fixed point in time smaller than the one for $H = R$ which can be explained by the smaller conditional variance of the filter due to the additional information from observing the diffusion J.

Overall, the uncertainty set of the D-investor seems to be the one that follows the true drift process best in this example. However, neither of the information filtrations leads to a perfect uncertainty set in the sense that the true drift stays in that uncertainty set at any point in time. By the setup of the uncertainty set there is always a positive probability that the true drift process moves out of the uncertainty set at some point in time.

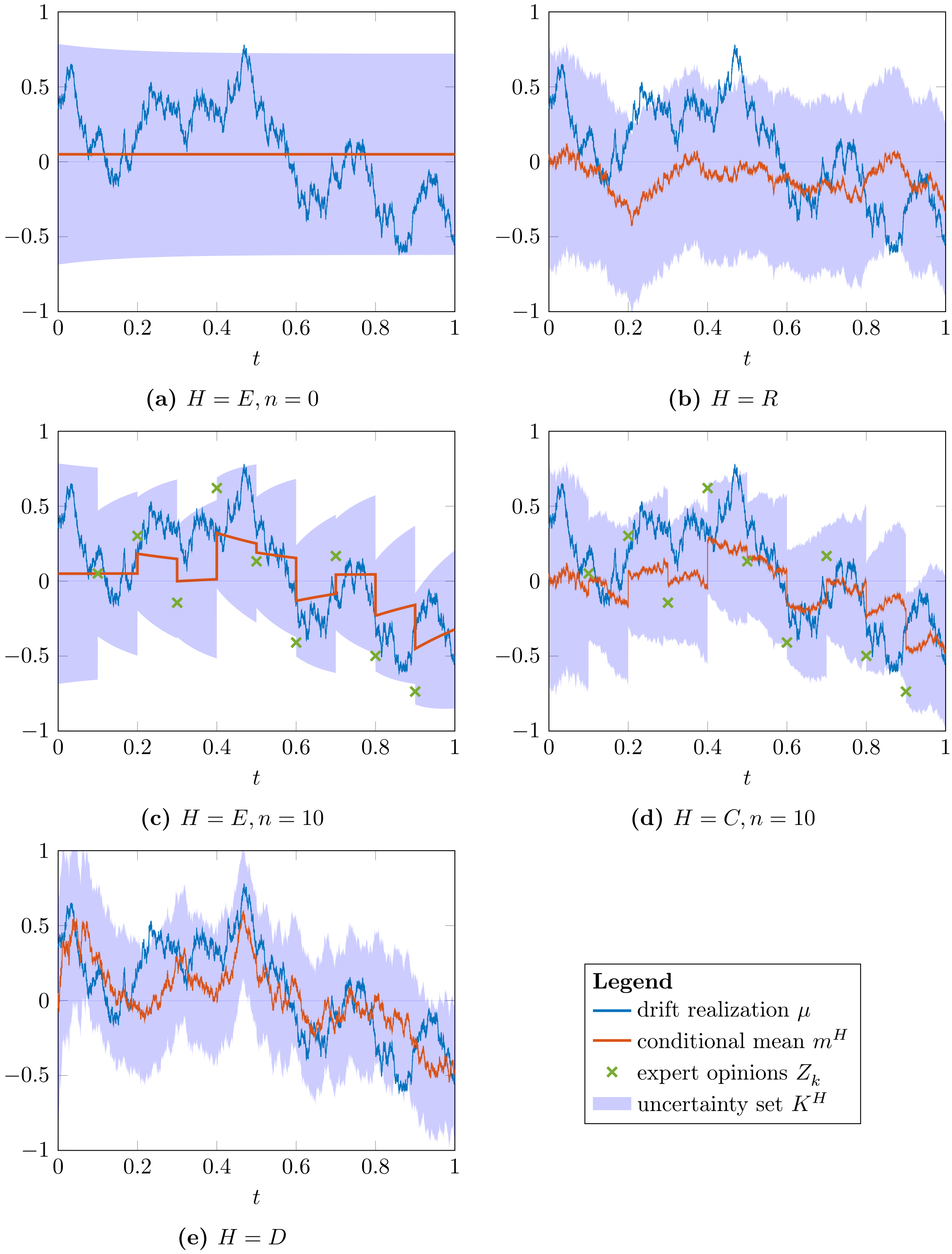


Figure 11.1.: Uncertainty sets based on filters for various investor filtrations $\mathbb{F}^H$. Each subplot is based on the same realization of the drift and return process and expert opinions. Based on this realization, the filter of the H-investor can be computed. The uncertainty set K^H is then determined according to the filter realization.

11.3. Comparison of expected utility for different investors

Lastly, we give a numerical example to illustrate the effect that the worst-case optimization among uncertainty sets created from filters has for the various investor filtrations considered before. Like in the preceding section we create for a fixed realization of the drift process, of the diffusions R and J and the expert opinions Z_k a time-dependent uncertainty set for each of the corresponding filters. The aim is to compare the robust strategies that take into account model uncertainty with the “naive” strategies that rely on the respective drift estimates, only.

We want to apply our worst-case utility maximization problem, in particular also imposing the constraint $\langle \pi_t, \mathbf{1}_d \rangle = h$ on the investor's strategies. For that purpose we take a market with $d = 2$ stocks here. We fix an investment horizon of $T = 1$ and take $h = 1$. Moreover, we assume that investors start with an initial wealth of $x_0 = 1$, use power utility functions U_γ with $\gamma = 0.5$ and a confidence level $\eta = 0.1$ to create their uncertainty sets. Further parameters of the market are given in Table 11.1.

mean reversion speed of drift process	α	$=$	$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$
volatility of drift process	β	$=$	$\begin{pmatrix} 0.50 & 0.25 \\ 0.25 & 0.50 \end{pmatrix}$
mean reversion level of drift process	δ	$=$	$\begin{pmatrix} 0.02 \\ 0.03 \end{pmatrix}$
initial mean of drift process	m_0	$=$	$\begin{pmatrix} 0.02 \\ 0.03 \end{pmatrix}$
initial variance of drift process	Σ_0	$=$	$\begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}$
volatility of returns	σ_R	$=$	$\begin{pmatrix} 0.10 & 0.05 \\ 0.05 & 0.01 \end{pmatrix}$
volatility of continuous expert	σ_J	$=$	$\begin{pmatrix} 0.10 & 0.05 \\ 0.05 & 0.01 \end{pmatrix}$

Table 11.1.: Market parameters for numerical example.

For the given model parameters we simulate a drift process, the diffusion processes R and J and $n = 10$ discrete-time expert opinions arriving at deterministic and equidistant information dates on $[0, T]$. We then obtain a realization of the filters (m^H, Q^H) for any of the information settings H from the preceding section. As before, this leads to one time-dependent uncertainty set for each of the investors.

We can then determine the worst-case drift process $(\mu_t^*)_{t\in[0,T]}$ and the optimal strategy $(\pi_t^*)_{t\in[0,T]}$ that is realized by the investor who solves at each time point the local optimization problem

$$\sup_{\pi^{(t)}\in\mathcal{A}_h(t,X_t^\pi)} \inf_{\mu^{(t)}\in\mathcal{K}^{(t)}} \mathbb{E}_{\mu^{(t)}}\Big[U_\gamma\big(X_T^{\pi^{(t)}}\big)\,\Big|\,\mathcal{F}_t^H\Big].$$

Recall that $(\mu_t^*)_{t\in[0,T]}$ and $(\pi_t^*)_{t\in[0,T]}$ are calculated from the solutions of the local optimization problems via

$$\pi_t^* = \pi_t^{(t),*}, \quad \mu_t^* = \mu_t^{(t),*}$$

for all $t \in [0,T]$. The value of each investor's worst-case optimization is then equal to

$$\mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big]. \tag{11.1}$$

The quantity in (11.1) is the worst-case expected utility from the H-investor's point of view when using the robust strategy π^*. For comparison, we also compute

$$\mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\hat{\pi}})\big], \quad \mathbb{E}_{\nu}\big[U_\gamma(X_T^{\pi^*})\big] \quad \text{and} \quad \mathbb{E}_{\nu}\big[U_\gamma(X_T^{\hat{\pi}})\big],$$

where $\nu = m^H$ is the conditional mean of the H-investor's filter and $\hat{\pi}$ is the corresponding optimal strategy given that the drift equals m^H, i.e.

$$\hat{\pi}_t = \frac{1}{1-\gamma} A m_t^H + hc.$$

We repeat this simulation 10 000 times where in each iteration a new drift process, a new return process and new expert opinions are simulated based on the parameters given above. Table 11.2 gives the sample mean of the various expected utilities over all simulations and in brackets the corresponding sample standard deviation.

H	n	$\mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big]$	$\mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\hat{\pi}})\big]$	$\mathbb{E}_{\nu}\big[U_\gamma(X_T^{\pi^*})\big]$	$\mathbb{E}_{\nu}\big[U_\gamma(X_T^{\hat{\pi}})\big]$
E	0	1.6179 *(0.0000)*	1.5996 *(0.0000)*	2.0196 *(0.0000)*	2.0426 *(0.0000)*
R		1.7086 *(0.1057)*	0.7754 *(0.3737)*	2.2362 *(2.4692)*	25.9029 *(732.4104)*
E	10	1.7055 *(0.1117)*	0.8170 *(0.3870)*	2.2393 *(3.4208)*	21.1610 *(530.6829)*
C	10	1.7854 *(0.4027)*	0.6891 *(0.3752)*	4.5313 *(134.5858)*	264.0838 *(19 288.2826)*
D		1.7888 *(0.4320)*	0.6711 *(0.3692)*	4.6831 *(141.2865)*	267.4413 *(18 827.1094)*

Table 11.2.: Comparison of utility for different investors.

When comparing the worst-case expected utility $\mathbb{E}_{\mu^*}[U_\gamma(X_T^{\pi^*})]$ among the investors we see that the information setting $H = E$, $n = 0$, which corresponds to only knowing the model parameters, gives the lowest value. The observation of returns or of $n = 10$ expert opinions increases this value. The combination of return observation and discrete-time expert opinions yields a considerably larger worst-case expected utility. We also see that the value for $H = D$ is quite close to the value we get for the C-investor with $n = 10$ expert opinions.

In the next column, $\mathbb{E}_{\mu^*}[U_\gamma(X_T^{\hat{\pi}})]$ measures the expected utility when using the strategy $\hat{\pi}$, given that the true drift is actually the worst-case drift μ^*. The values are in any case smaller than the corresponding expected utility when using the robust strategy π^*. What is striking is that the information setting $H = E$ with $n = 0$, i.e. only knowledge of the model parameters, gives the best expected utility here. Adding more information, from return observations or expert opinions, and using the optimal strategy based on the filter leads to a smaller worst-case expected utility. This shows that for the worst-case optimization problem it is dangerous for investors to rely on their estimates of the drift, i.e. the conditional mean of the filter, only. They need to robustify their strategy by taking into account model uncertainty to be able to profit from any additional information.

The last two columns show the expected utility when using strategy π^*, respectively $\hat{\pi}$, given that the true drift was actually the conditional mean $\nu = m^H$. Of course, when compared to the expected utility given the worst-case drift μ^*, the expected utility given ν is much higher. Not surprisingly, the performance of $\hat{\pi}$ given drift ν is on average extremely good. However, we also notice the very large sample standard deviation. In comparison to that, we see that the robust strategies π^* perform reasonably well given drift ν, even though they are tailored for the worst-case drift in the respective uncertainty set. At the same time, the sample standard deviation is much smaller than for strategy $\hat{\pi}$.

In conclusion, we see that a surplus of information, either from return observations or expert opinions, results in better strategies in general. However, investors do need to account for model uncertainty by choosing a robustified strategy π^* instead of relying on the respective filter only. The naive strategy $\hat{\pi}$ performs extremely well if the true drift coincides with the conditional mean m^H, but it is much more vulnerable to model misspecifications than the robust strategy π^*.

12. Conclusion

In this thesis we investigated utility maximization problems in Black–Scholes type financial markets with incomplete information about market parameters. To account for model uncertainty due to statistical estimation errors we considered in Part I robust optimization problems where investors maximize their worst-case expected utility, given that the drift of risky assets can take values in a prespecified uncertainty set. As the degree of uncertainty becomes large, investors usually do not invest in risky assets at all. Therefore, we imposed a constraint that prevents a pure bond investment.

In the logarithmic utility case and with uncertainty sets that are balls in some p-norm we carried over the approach from Pflug et al. [47] for a one-period risk minimization problem to our model. The key result here is that, as the level of uncertainty about the drift goes to infinity, the optimal trading strategy converges to a uniform diversification strategy. However, to be able to apply the methods from Pflug et al. [47] we had to restrict to deterministic strategies in Chapter 3. Also, while we obtain asymptotic results for large levels of uncertainty, this approach does not yield an explicit form of the optimal trading strategy for the problem with a fixed degree of uncertainty.

For these reasons, we came up with a different approach in Chapter 4 that solves our robust utility maximization problem for both power and logarithmic utility without restricting to deterministic strategies and for more general ellipsoidal uncertainty sets. The main idea is to solve the corresponding dual problem explicitly and to prove a minimax theorem to ensure that the solution to the dual also solves our original problem. We used the explicit structure of our solution to derive the asymptotic behavior for large levels of uncertainty. In the limit, as uncertainty goes to infinity, the optimal strategy converges to a generalized uniform diversification strategy.

To come up with a reasonable uncertainty set, it makes sense for investors to estimate the drift of asset returns based on the information that is available in the market. This typically comprises return observations but also external sources of information that are called expert opinions in our context. In Part II of this thesis we dealt with a Black–Scholes type financial market with an underlying Gaussian drift process where investors find estimates about the unobservable drift based on filtering techniques.

We saw that discrete-time expert opinions lead to updates of the filter that decrease the conditional variance, therefore giving better estimates. The conditional covariance matrices of the filter serve for assessing the goodness of the filter. We derived properties of the conditional covariance matrices and studied their asymptotic behavior on an infinite time horizon.

Our focus was on the asymptotic behavior of the filter as the arrival frequency of expert opinions goes to infinity on a finite time horizon. In Chapter 8 we distinguished two cases. If the variance of expert opinions is bounded, then in the limit an investor who observes the discrete-time expert opinions has full information about the drift. In contrast, if the expert's variance grows linearly in the arrival frequency, then the information obtained from observing the discrete-time expert opinions is asymptotically the same as that from observing another diffusion process which we interpret as a continuous-time expert. We showed this by proving convergence of the conditional covariance matrices and the conditional means of the corresponding filters.

Since our convergence results carry over to convergence of the value function in a portfolio optimization problem with logarithmic utility, it is possible to find approximate solutions of utility maximization problems by replacing an investor's filter with the corresponding asymptotic filter which is much easier to handle numerically.

In Part III of this thesis we used our observations about how expert opinions improve drift estimates for the robust utility maximization problem from the first part. We generalized our financial market model from Part I to one with non-constant drift and also allowed for time-dependence in the uncertainty set. The duality approach for finding the optimal trading strategy then carries over from the situation with constant drift.

Finally, we showed how a time-dependent uncertainty set can be defined based on a generic filter. We applied this to the various investor filtrations from Part II to illustrate how expert opinions decrease the size of uncertainty sets and how this surplus of information in general results in better strategies. However, by means of a numerical simulation we also demonstrated that investors need to account for model uncertainty by choosing a robust strategy instead of relying only on the respective drift estimation.

Appendices

A. Auxiliary Results for Diffusion Approximations

Here we collect some auxiliary results that are used in the proofs of our main results from Section 8.2. The following lemma can be interpreted as a discrete version of Gronwall's Lemma for error accumulation. A statement very similar to Lemma A.1 can be found in Demailly [16, Sec. 8.2.4].

Lemma A.1. *Let $(a_j)_{j=0,\dots,n}$, $(h_j)_{j=0,\dots,n}$ be real-valued sequences with $a_j \geq 0$, $h_j > 0$, and let $L > 0$, $b \geq 0$ be real numbers such that*

$$a_{j+1} \leq (1 + h_j L)a_j + h_j b, \quad j = 0, 1, \dots, n-1.$$

Then for all $j = 0, 1, \dots, n$ it holds

$$a_j \leq \frac{\mathrm{e}^{Lt_j} - 1}{L} b + \mathrm{e}^{Lt_j} a_0,$$

where $t_j = \sum_{i=0}^{j-1} h_i$.

Proof. The proof can be done by induction. For $j = 0$ the claim is obvious. For the induction step we observe that $1 + x \leq \mathrm{e}^x$ for all $x \in \mathbb{R}$ and hence

$$a_{j+1} \leq (1 + h_j L)a_j + h_j b \leq \mathrm{e}^{h_j L} a_j + h_j b.$$

Due to the induction hypothesis we therefore have

$$\begin{aligned}
a_{j+1} &\leq \mathrm{e}^{h_j L}\Big(\frac{\mathrm{e}^{Lt_j} - 1}{L} b + \mathrm{e}^{Lt_j} a_0\Big) + h_j b \\
&= \Big(\frac{\mathrm{e}^{L(t_j + h_j)} - \mathrm{e}^{Lh_j} + h_j L}{L}\Big) b + \mathrm{e}^{L(t_j + h_j)} a_0 \\
&\leq \frac{\mathrm{e}^{Lt_{j+1}} - 1}{L} b + \mathrm{e}^{Lt_{j+1}} a_0,
\end{aligned}$$

which completes the proof. □

The next lemmas are used in the proof of Theorem 8.6. Firstly, the following lemma is a Cauchy–Schwarz inequality for multidimensional integrals.

Lemma A.2. *Let $(X_s)_{s\in[0,t]}$ be an $\mathbb{R}^d$-valued stochastic process. Then*

$$\mathbb{E}\Big[\Big\|\int_0^t X_s \,\mathrm{d}s\Big\|^2\Big] \leq t \int_0^t \mathbb{E}\big[\|X_s\|^2\big] \,\mathrm{d}s.$$

Proof. Firstly, pulling the norm into the integral increases the expression on the left-hand side, so

$$\mathbb{E}\left[\left\|\int_0^t X_s \,\mathrm{d}s\right\|^2\right] \le \mathbb{E}\left[\left(\int_0^t \|X_s\| \,\mathrm{d}s\right)^2\right].$$

Now we can apply the usual Cauchy–Schwarz inequality to the one-dimensional integral and get

$$\mathbb{E}\left[\left(\int_0^t \|X_s\| \,\mathrm{d}s\right)^2\right] \le \mathbb{E}\left[t \int_0^t \|X_s\|^2 \,\mathrm{d}s\right] = t \int_0^t \mathbb{E}\big[\|X_s\|^2\big] \,\mathrm{d}s.$$

The last step is due to Fubini. □

A key tool for estimations involving stochastic integrals is the Itô isometry. The following lemma uses the isometry to obtain an estimation for multivariate integrals.

Lemma A.3. *Let $W = (W_s)_{s\in[0,t]}$ be an m-dimensional Brownian motion. Let $(H_s)_{s\in[0,t]}$ be an $\mathbb{R}^{d\times m}$-valued stochastic process that is independent of W, and τ a stopping time that is bounded by t and also independent of W. Then*

$$\mathbb{E}\left[\left\|\int_0^\tau H_s \,\mathrm{d}W_s\right\|^2\right] = \mathbb{E}\left[\int_0^\tau \|H_s\|_F^2 \,\mathrm{d}s\right] \le C_{norm}\, \mathbb{E}\left[\int_0^\tau \|H_s\|^2 \,\mathrm{d}s\right],$$

where $\|\cdot\|_F$ denotes the Frobenius norm and $C_{norm} > 0$ only depends on the dimension $d \times m$ of the integrand H.

Proof. Note that for fixed, deterministic t, the integral $\int_0^t H_s \,\mathrm{d}W_s$ is a random variable with values in $\mathbb{R}^d$. The i-th entry is

$$\sum_{j=1}^m \int_0^t H_s^{ij} \,\mathrm{d}W_s^j.$$

Hence,

$$\left\|\int_0^t H_s \,\mathrm{d}W_s\right\|^2 = \sum_{i=1}^d \left(\sum_{j=1}^m \int_0^t H_s^{ij} \,\mathrm{d}W_s^j\right)^2.$$

When applying the expectation, we get due to independence

$$\begin{aligned}\mathbb{E}\left[\left\|\int_0^t H_s \,\mathrm{d}W_s\right\|^2\right] &= \sum_{i=1}^d \sum_{j,k=1}^m \mathbb{E}\left[\int_0^t H_s^{ij} \,\mathrm{d}W_s^j \int_0^t H_s^{ik} \,\mathrm{d}W_s^k\right] \\ &= \sum_{i=1}^d \sum_{j=1}^m \mathbb{E}\left[\left(\int_0^t H_s^{ij} \,\mathrm{d}W_s^j\right)^2\right].\end{aligned} \tag{A.1}$$

Note that we can consider the filtration $(\mathcal{G}_s)_{s\in[0,t]}$ where $\mathcal{G}_s = \sigma(W_u, u \le s) \vee \sigma(H_u, u \in [0,t])$. Since H and W are independent, W is a Brownian motion with respect to $(\mathcal{G}_s)_{s\in[0,t]}$. Also, H is obviously adapted with respect to $(\mathcal{G}_s)_{s\in[0,t]}$. Hence, we can apply the usual Itô isometry and obtain that the right-hand side of (A.1) equals

$$\sum_{i=1}^d \sum_{j=1}^m \mathbb{E}\left[\int_0^t (H_s^{ij})^2 \,\mathrm{d}s\right] = \mathbb{E}\left[\int_0^t \|H_s\|_F^2 \,\mathrm{d}s\right].$$

Now when taking into account the stopping time τ, we can write

$$\mathbb{E}\left[\left\|\int_0^\tau H_s \,\mathrm{d}W_s\right\|^2\right] = \mathbb{E}\left[\left\|\int_0^t \mathbb{1}_{\{s\leq\tau\}} H_s \,\mathrm{d}W_s\right\|^2\right].$$

Since τ is independent of W we can deduce from the previous part of the proof that

$$\mathbb{E}\left[\left\|\int_0^t \mathbb{1}_{\{s\leq\tau\}} H_s \,\mathrm{d}W_s\right\|^2\right] = \mathbb{E}\left[\int_0^t \|\mathbb{1}_{\{s\leq\tau\}} H_s\|_F^2 \,\mathrm{d}s\right] = \mathbb{E}\left[\int_0^\tau \|H_s\|_F^2 \,\mathrm{d}s\right].$$

Equivalence of norms implies the existence of the constant $C_{\text{norm}} > 0$ with the property that

$$\mathbb{E}\left[\int_0^\tau \|H_s\|_F^2 \,\mathrm{d}s\right] \leq C_{\text{norm}}\, \mathbb{E}\left[\int_0^\tau \|H_s\|^2 \,\mathrm{d}s\right],$$

which concludes the proof. □

Another estimate that is useful in the convergence proofs is given in the following lemma.

Lemma A.4. *Let $\kappa > 0$ and let Q^κ be a symmetric and positive-definite matrix in $\mathbb{R}^{d\times d}$ with $\|Q^\kappa\| \leq C_Q$ for all κ. Then there exists a constant $\bar{C} > 0$ such that*

$$\left\|Q^\kappa - Q^\kappa\big(Q^\kappa + \kappa\sigma_J\sigma_J^\top\big)^{-1}\kappa\sigma_J\sigma_J^\top\right\| \leq \frac{\bar{C}}{\kappa}.$$

Proof. For abbreviation let $A := Q^\kappa$, $B := \sigma_J\sigma_J^\top$. Then we can write

$$\begin{aligned} A - A(A+\kappa B)^{-1}\kappa B &= A(A+\kappa B)^{-1}(A+\kappa B-\kappa B) = A(A+\kappa B)^{-1}A \\ &= \big(A^{-1}(A+\kappa B)A^{-1}\big)^{-1} = \big(A^{-1}+\kappa A^{-1}BA^{-1}\big)^{-1}, \end{aligned}$$

and therefore

$$\begin{aligned} \|A - A(A+\kappa B)^{-1}\kappa B\| &= \|\big(A^{-1}+\kappa A^{-1}BA^{-1}\big)^{-1}\| = \Big(\lambda_{\min}(A^{-1}+\kappa A^{-1}BA^{-1})\Big)^{-1} \\ &\leq \Big(\lambda_{\min}(A^{-1}) + \lambda_{\min}(\kappa A^{-1}BA^{-1})\Big)^{-1} \leq \Big(\lambda_{\min}(\kappa A^{-1}BA^{-1})\Big)^{-1} \\ &= \frac{1}{\kappa}\|AB^{-1}A\|. \end{aligned}$$

Hence, we obtain

$$\left\|Q^\kappa - Q^\kappa\big(Q^\kappa + \kappa\sigma_J\sigma_J^\top\big)^{-1}\kappa\sigma_J\sigma_J^\top\right\| \leq \frac{C_Q^2\|(\sigma_J\sigma_J^\top)^{-1}\|}{\kappa} = \frac{\bar{C}}{\kappa},$$

where $\bar{C} = C_Q^2\|(\sigma_J\sigma_J^\top)^{-1}\|$. □

The next lemma states Gronwall's Lemma in integral form which we use in the proofs of Theorems 8.15 and 8.16. A proof can be found for example in Pachpatte [46, Sec. 1.3].

Lemma A.5 (Gronwall). *Let $I = [a,b]$ be an interval and let u, α and $\beta\colon I \to [0,\infty)$ be continuous functions with*

$$u(t) \leq \alpha(t) + \int_a^t \beta(s)u(s)\,\mathrm{d}s$$

for all $t \in I$. Then

$$u(t) \leq \alpha(t) + \int_a^t \alpha(s)\beta(s)\mathrm{e}^{\int_s^t \beta(r)\,\mathrm{d}r}\,\mathrm{d}s$$

for all $t \in I$.

In Section 8.2.2 we work with a Poisson random measure. An important property of the compensated Poisson measure that we use for the proof of Theorem 8.15 is given in the following lemma, see Cont and Tankov [11, Prop. 2.16].

Lemma A.6. *For an integrable real-valued function $f\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$, the process $(X_t)_{t\geq 0}$ with*

$$X_t = \int_0^t \int_{\mathbb{R}^d} f(s,u)\,\tilde{N}(\mathrm{d}s, \mathrm{d}u)$$

is a martingale with $\mathbb{E}[X_t] = 0$ and

$$\operatorname{var}(X_t) = \mathbb{E}\big[X_t^2\big] = \mathbb{E}\bigg[\int_0^t \int_{\mathbb{R}^d} f^2(s,u)\lambda\varphi(u)\,\mathrm{d}u\,\mathrm{d}s\bigg].$$

Because of the additional randomness from the Poisson process $(N_t^{(\lambda)})_{t\in[0,T]}$ in the situation with random information dates, we also need the estimation from the following lemma in the proof of Theorem 8.16.

Lemma A.7. *Let $(N_t)_{t\in[0,T]}$ be a standard Poisson process with intensity $\lambda > 0$. Then there exists a constant $C_N > 0$ such that*

$$\sum_{k=\lfloor\lambda t\rfloor+1}^{\infty} \mathbb{E}\big[\mathbb{1}_{\{N_t\geq k\}}\big] = \mathbb{E}\big[\big(N_t - \lfloor\lambda t\rfloor\big)^+\big] \leq C_N\sqrt{\lambda}$$

for all $\lambda \geq 1$.

Proof. The first equality holds since

$$\mathbb{E}\Bigg[\sum_{k=\lfloor\lambda t\rfloor+1}^{\infty} \mathbb{1}_{\{N_t\geq k\}}\Bigg] = \mathbb{E}\Big[\mathbb{1}_{\{N_t\geq\lfloor\lambda t\rfloor+1\}}(N_t - \lfloor\lambda t\rfloor)\Big] = \mathbb{E}\Big[\big(N_t - \lfloor\lambda t\rfloor\big)^+\Big].$$

Since $X^+ \leq |X|$ and $\mathbb{E}[|X|] \leq \sqrt{\mathbb{E}[X^2]}$ by the Lyapunov inequality, we find that

$$\mathbb{E}\big[\big(N_t - \lambda t\big)^+\big] \leq \mathbb{E}\big[\big|N_t - \lambda t\big|\big] \leq \sqrt{\mathbb{E}\big[\big(N_t - \lambda t\big)^2\big]} = \sqrt{\operatorname{var}(N_t)} = \sqrt{\lambda t} \leq \sqrt{T}\sqrt{\lambda},$$

since $\mathbb{E}[N_t] = \operatorname{var}(N_t) = \lambda t$. But then also

$$\mathbb{E}\big[\big(N_t - \lfloor\lambda t\rfloor\big)^+\big] \leq \mathbb{E}\big[\big(N_t - \lambda t\big)^+\big] + 1 \leq \sqrt{T}\sqrt{\lambda} + 1 \leq C_N\sqrt{\lambda}$$

for $C_N := \sqrt{T} + 1$ and $\lambda \geq 1$. $\square$

B. Utility Maximization in a Market with Non-Constant Parameters

It is well known since Merton [43] that in a Black–Scholes market with constant parameters the optimal strategy for an investor maximizing expected power utility of terminal wealth has the form

$$\pi_t^* = \frac{1}{1-\gamma}(\sigma\sigma^\top)^{-1}(\mu - r\mathbf{1}_d).$$

In this appendix we slightly extend this result to a market where the risk-free interest rate as well as drift and volatility of the stocks are not necessarily constant parameters. However, they are still assumed to be observable by the investor. A similar result has been proven in Karatzas et al. [34] for complete markets with deterministic market coefficients and for incomplete markets with totally unhedgeable market coefficients.

Let $T > 0$ be a finite time horizon and $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ a filtered probability space where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions. All processes are assumed to be $\mathbb{F}$-adapted. We assume that the risk-free asset S^0 follows the dynamics

$$\mathrm{d}S_t^0 = S_t^0 r_t \,\mathrm{d}t, \quad S_0^0 = 1,$$

and that the risky assets $S = (S^1, \dots, S^d)^\top \in \mathbb{R}^d$ are defined by

$$\mathrm{d}S_t = \operatorname{diag}(S_t)\big(\mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t\big), \quad S_0 = s_0.$$

Here, $(W_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion, $m \geq d$. The processes $r = (r_t)_{t\in[0,T]}$, $\mu = (\mu_t)_{t\in[0,T]}$ and $\sigma = (\sigma_t)_{t\in[0,T]}$ take values in $\mathbb{R}$, $\mathbb{R}^d$ and $\mathbb{R}^{d\times m}$, respectively, and are assumed to be independent of the Brownian motion $(W_t)_{t\in[0,T]}$. We also assume that for any $t \in [0,T]$, the matrix σ_t has full rank. The vector $s_0 \in \mathbb{R}^d$ is assumed to have strictly positive entries.

Now let $\theta_t = \sigma_t^\top(\sigma_t\sigma_t^\top)^{-1}(\mu_t - r_t\mathbf{1}_d)$ for any $t \in [0,T]$ and define the process $(Z_t)_{t\in[0,T]}$ by

$$\mathrm{d}Z_t = -Z_t\theta_t^\top \mathrm{d}W_t.$$

We assume that θ is such that Z is a true martingale. Further, we use the notation

$$\beta_t = \frac{S_0^0}{S_t^0} = \exp\Big(-\int_0^t r_s \,\mathrm{d}s\Big),$$

as well as

$$H_t = \beta_t Z_t = \exp\Big(-\int_0^t r_s \,\mathrm{d}s - \frac{1}{2}\int_0^t \|\theta_s\|^2 \,\mathrm{d}s - \int_0^t \theta_s^\top \mathrm{d}W_s\Big).$$

Then the wealth process for an investor starting with initial wealth x_0 and following the self-financing trading strategy $\pi = (\pi_t)_{t\in[0,T]}$ has the form

$$X_T^\pi = x_0 \exp\Big(\int_0^T r_s \,\mathrm{d}s + \int_0^T \Big(\pi_s^\top(\mu_s - r_s\mathbf{1}_d) - \frac{1}{2}\|\sigma_s^\top\pi_s\|^2\Big)\mathrm{d}s + \int_0^T \pi_s^\top\sigma_s \,\mathrm{d}W_s\Big).$$

The admissible strategies are the elements of

$$\mathcal{A}(x_0) = \left\{ \pi = (\pi_t)_{t\in[0,T]} \;\middle|\; \pi \text{ is } \mathbb{F}^{S,r,\mu,\sigma}\text{-adapted, } X_0^\pi = x_0,\ \mathbb{E}\left[\int_0^T \|\sigma_t^\top \pi_t\|^2 \,\mathrm{d}t\right] < \infty \right\}.$$

Here, $\mathbb{F}^{S,r,\mu,\sigma}$ is the filtration generated by the processes S, r, μ and σ, augmented by the null sets. These strategies correspond to an investor who is able to observe the price process S as well as r, μ and σ.

As in the setting with constant parameters one can now follow the martingale method and separate the problem into the static problem of finding the optimal terminal wealth and the representation problem of finding a corresponding optimal strategy. We let $U\colon \mathbb{R}_+ \to \mathbb{R}$ denote an arbitrary utility function, i.e. a strictly concave and continuously differentiable function satisfying

$$\lim_{x\downarrow 0} U'(x) = \infty \quad \text{and} \quad \lim_{x\to\infty} U'(x) = 0,$$

and use the notation $I = (U')^{-1}$. The following result then characterizes the optimal terminal wealth.

Proposition B.1. *If an admissible trading strategy π^* exists such that $X_T^{\pi^*} = I(\lambda H_T)$ for a stricty positive random variable λ and*

$$\mathbb{E}\big[\lambda H_T X_T^\pi\big] \le \mathbb{E}\big[\lambda H_T I(\lambda H_T)\big] < \infty$$

for all admissible π, then π^* is optimal, i.e.

$$\mathbb{E}\big[U(X_T^{\pi^*})\big] \ge \mathbb{E}\big[U(X_T^\pi)\big]$$

for all admissible π.

Proof. By concavity of U we have

$$U(x) - xy \le U(I(y)) - yI(y)$$

for all $x, y \in \mathbb{R}_+$. For any admissible π we obtain by taking $x = X_T^\pi$ and $y = \lambda H_T$ that

$$\mathbb{E}\big[U(X_T^\pi)\big] \le \mathbb{E}\big[U(I(\lambda H_T))\big] - \mathbb{E}\big[\lambda H_T I(\lambda H_T)\big] + \mathbb{E}\big[\lambda H_T X_T^\pi\big] \le \mathbb{E}\big[U(I(\lambda H_T))\big] = \mathbb{E}\big[U(X_T^{\pi^*})\big].$$

This shows optimality of π^*. □

The result can be applied to the utility maximization problem with power utility. Let $U_\gamma\colon \mathbb{R}_+ \to \mathbb{R}$, $U_\gamma(x) = \frac{x^\gamma}{\gamma}$ for $\gamma \in (-\infty, 1)$, $\gamma \neq 0$.

Theorem B.2. *The strategy $(\pi_t^*)_{t\in[0,T]}$ with*

$$\pi_t^* = \frac{1}{1-\gamma}(\sigma_t\sigma_t^\top)^{-1}(\mu_t - r_t\mathbf{1}_d)$$

is optimal for the problem

$$\sup_{\pi\in\mathcal{A}(x_0)} \mathbb{E}\big[U_\gamma(X_T^\pi)\big].$$

Proof. We first rewrite the terminal wealth when using strategy π^* as

$$\begin{aligned} X_T^{\pi^*} &= x_0 \exp\Big(\int_0^T r_s \,\mathrm{d}s + \int_0^T \Big(\frac{1}{1-\gamma} - \frac{1}{2(1-\gamma)^2}\Big)\theta_s^\top \theta_s \,\mathrm{d}s + \frac{1}{1-\gamma}\int_0^T \theta_s^\top \,\mathrm{d}W_s\Big) \\ &= x_0 \exp\Big(\int_0^T r_s \,\mathrm{d}s + \frac{1}{2}\int_0^T \frac{1-2\gamma}{(1-\gamma)^2}\|\theta_s\|^2 \,\mathrm{d}s + \frac{1}{1-\gamma}\int_0^T \theta_s^\top \,\mathrm{d}W_s\Big) \end{aligned} \tag{B.1}$$

and note that

$$H_T = \exp\Big(-\int_0^T r_s \,\mathrm{d}s - \frac{1}{2}\int_0^T \|\theta_s\|^2 \,\mathrm{d}s - \int_0^T \theta_s^\top \,\mathrm{d}W_s\Big). \tag{B.2}$$

Since $U'(x) = x^{\gamma-1}$ and $I(y) = y^{-\frac{1}{1-\gamma}}$ we obtain for an arbitrary random variable λ that $I(\lambda H_T) = I(\lambda)I(H_T)$ where

$$I(H_T) = \exp\Big(\frac{1}{1-\gamma}\int_0^T r_s \,\mathrm{d}s + \frac{1}{2(1-\gamma)}\int_0^T \|\theta_s\|^2 \,\mathrm{d}s + \frac{1}{1-\gamma}\int_0^T \theta_s^\top \,\mathrm{d}W_s\Big).$$

We also obtain

$$\begin{aligned} \frac{X_T^{\pi^*}}{I(H_T)} &= x_0 \exp\Big(\Big(1 - \frac{1}{1-\gamma}\Big)\int_0^T r_s \,\mathrm{d}s + \frac{1}{2(1-\gamma)}\Big(\frac{1-2\gamma}{1-\gamma} - 1\Big)\int_0^T \|\theta_s\|^2 \,\mathrm{d}s\Big) \\ &= x_0 \exp\Big(-\frac{\gamma}{1-\gamma}\int_0^T r_s \,\mathrm{d}s - \frac{\gamma}{2(1-\gamma)^2}\int_0^T \|\theta_s\|^2 \,\mathrm{d}s\Big). \end{aligned}$$

Hence we have $X_T^{\pi^*} = I(\lambda H_T)$ for the strictly positive random variable

$$\lambda := x_0^{-(1-\gamma)} \exp\Big(\gamma\int_0^T r_s \,\mathrm{d}s + \frac{\gamma}{2(1-\gamma)}\int_0^T \|\theta_s\|^2 \,\mathrm{d}s\Big).$$

Next, we check the condition from Proposition B.1. From (B.1) and (B.2) we deduce

$$\begin{aligned} H_T I(\lambda H_T) = H_T X_T^{\pi^*} &= x_0 \exp\Big(\frac{1}{2}\Big(\frac{1-2\gamma}{(1-\gamma)^2} - 1\Big)\int_0^T \|\theta_s\|^2 \,\mathrm{d}s + \Big(\frac{1}{1-\gamma} - 1\Big)\int_0^T \theta_s^\top \,\mathrm{d}W_s\Big) \\ &= x_0 \exp\Big(-\frac{1}{2}\Big(\frac{\gamma}{1-\gamma}\Big)^2\int_0^T \|\theta_s\|^2 \,\mathrm{d}s + \frac{\gamma}{1-\gamma}\int_0^T \theta_s^\top \,\mathrm{d}W_s\Big), \end{aligned}$$

hence

$$\begin{aligned} &\mathbb{E}\big[\lambda H_T I(\lambda H_T)\big] = \mathbb{E}\Big[\lambda\,\mathbb{E}\big[H_T I(\lambda H_T) \,|\, \mathcal{F}_T^{r,\mu,\sigma}\big]\Big] \\ &= \mathbb{E}\bigg[\lambda x_0\,\mathbb{E}\bigg[\exp\Big(-\frac{1}{2}\Big(\frac{\gamma}{1-\gamma}\Big)^2\int_0^T \|\theta_s\|^2 \,\mathrm{d}s + \frac{\gamma}{1-\gamma}\int_0^T \theta_s^\top \,\mathrm{d}W_s\Big) \,\bigg|\, \mathcal{F}_T^{r,\mu,\sigma}\bigg]\bigg] = x_0\,\mathbb{E}[\lambda]. \end{aligned}$$

Here, $(\mathcal{F}_t^{r,\mu,\sigma})_{t\in[0,T]}$ is the filtration generated by r, μ and σ, augmented by null sets, and we have used that these processes are independent of the Brownian motion $(W_t)_{t\in[0,T]}$ for the fact that the conditional expectation in the expression above is equal to one.

On the other hand, for any admissible strategy π we have

$$
\begin{aligned}
H_T X_T^{\pi} &= x_0 \exp\biggl(\int_0^T \Bigl(\pi_s^\top (\mu_s - r_s \mathbf{1}_d) - \frac{1}{2}\|\sigma_s^\top \pi_s\|^2 - \frac{1}{2}\|\theta_s\|^2 \Bigr) \mathrm{d}s + \int_0^T (\pi_s^\top \sigma_s - \theta_s^\top)\, \mathrm{d}W_s \biggr) \\
&= x_0 \exp\biggl(-\frac{1}{2} \int_0^T \|\sigma_s^\top \pi_s - \theta_s\|^2 \,\mathrm{d}s + \int_0^T (\sigma_s^\top \pi_s - \theta_s)^\top \,\mathrm{d}W_s \biggr),
\end{aligned}
$$

and therefore

$$
\begin{aligned}
&\mathbb{E}\bigl[\lambda H_T X_T^{\pi}\bigr] = \mathbb{E}\Bigl[\lambda\, \mathbb{E}\bigl[H_T X_T^{\pi} \,|\, \mathcal{F}_T^{r,\mu,\sigma}\bigr]\Bigr] \\
&\quad = \mathbb{E}\biggl[\lambda x_0\, \mathbb{E}\biggl[\exp\biggl(-\frac{1}{2} \int_0^T \|\sigma_s^\top \pi_s - \theta_s\|^2 \,\mathrm{d}s + \int_0^T (\sigma_s^\top \pi_s - \theta_s)^\top \,\mathrm{d}W_s \biggr) \,\bigg|\, \mathcal{F}_T^{r,\mu,\sigma}\biggr]\biggr] \le x_0\, \mathbb{E}[\lambda].
\end{aligned}
$$

The last inequality follows from the fact that the expression inside the conditional expectation is a positive local martingale, hence a supermartingale, and therefore the conditional expectation is less or equal than one. In conclusion, we have

$$
\mathbb{E}\bigl[\lambda H_T I(\lambda H_T)\bigr] = x_0\, \mathbb{E}[\lambda] \ge \mathbb{E}\bigl[\lambda H_T X_T^{\pi}\bigr].
$$

Since π was arbitrary, it follows from Proposition B.1 that π^* is optimal. $\square$

Bibliography

[1] A. Aalto, Convergence of discrete-time Kalman filter estimate to continuous time estimate, *International Journal of Control* **89** (2016), no. 4, pp. 668–679.

[2] B. Analui, *Multistage Stochastic Optimization of Energy Portfolios under Model Ambiguity*, Ph.D. thesis, Universität Wien (2014).

[3] S. Asmussen & H. Albrecher, *Ruin Probabilities*, vol. 14 of *Advanced Series on Statistical Science & Applied Probability*, World Scientific (2010).

[4] S. Biagini & M. Ç. Pinar, The robust Merton problem of an ambiguity averse investor, *Mathematics and Financial Economics* **11** (2017), no. 1, pp. 1–24.

[5] T. Björk, M. H. A. Davis & C. Landén, Optimal investment under partial information, *Mathematical Methods of Operations Research* **71** (2010), no. 2, pp. 371–399.

[6] F. Black & R. Litterman, Global portfolio optimization, *Financial Analysts Journal* **48** (1992), no. 5, pp. 28–43.

[7] S. Brendle, Portfolio selection under incomplete information, *Stochastic Processes and their Applications* **116** (2006), no. 5, pp. 701–723.

[8] R. S. Bucy & P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*, vol. 23 of *Interscience Tracts in Pure and Applied Mathematics*, John Wiley & Sons (1968).

[9] Z. Chen & L. Epstein, Ambiguity, risk, and asset returns in continuous time, *Econometrica* **70** (2002), no. 4, pp. 1403–1443.

[10] V. K. Chopra & W. T. Ziemba, The effect of errors in means, variances, and covariances on optimal portfolio choice, *Journal of Portfolio Management* **19** (1993), no. 2, pp. 6–11.

[11] R. Cont & P. Tankov, *Financial Modelling with Jump Processes*, Chapman and Hall/CRC (2004).

[12] P. Dai Pra, W. J. Runggaldier & C. Rudari, On dynamic programming for sequential decision problems under a general form of uncertainty, *Mathematical Methods of Operations Research* **45** (1997), no. 1, pp. 81–107.

[13] M. H. A. Davis, *Linear Estimation and Stochastic Control*, Chapman and Hall (1977).

[14] M. H. A. Davis & S. Lleo, Black–Litterman in continuous time: the case for filtering, *Quantitative Finance Letters* **1** (2013), no. 1, pp. 30–35.

[15] E. Delage, D. Kuhn & W. Wiesemann, "Dice"-sion–making under uncertainty: When can a random decision reduce risk? (2019). To appear in *Management Science*.

[16] J.-P. Demailly, *Gewöhnliche Differentialgleichungen: Theoretische und Numerische Aspekte*, Vieweg+Teubner Verlag (1994).

[17] V. DeMiguel, L. Garlappi & R. Uppal, Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy?, *The Review of Financial Studies* **22** (2009), no. 5, pp. 1915–1953.

[18] I. Ekeland & R. Temam, *Convex Analysis and Variational Problems*, North-Holland Publishing Company (1976).

[19] R. J. Elliott, L. Aggoun & J. B. Moore, *Hidden Markov Models: Estimation and Control*, Springer-Verlag New York (1995).

[20] G. Favero & W. J. Runggaldier, A robustness result for stochastic control, *Systems & Control Letters* **46** (2002), no. 2, pp. 91–97.

[21] J.-P. Fouque, A. Papanicolaou & R. Sircar, Filtering and portfolio optimization with stochastic unobserved drift in asset returns, *Communications in Mathematical Sciences* **13** (2015), no. 4, pp. 935–953.

[22] J.-P. Fouque, A. Papanicolaou & R. Sircar, Perturbation analysis for investment portfolios under partial information with expert opinions, *SIAM Journal on Control and Optimization* **55** (2017), no. 3, pp. 1534–1566.

[23] R. Frey, A. Gabih & R. Wunderlich, Portfolio optimization under partial information with expert opinions, *International Journal of Theoretical and Applied Finance* **15** (2012), no. 1, 1250009. 18 pages.

[24] R. Frey, A. Gabih & R. Wunderlich, Portfolio optimization under partial information with expert opinions: a dynamic programming approach, *Communications on Stochastic Analysis* **8** (2014), no. 1, pp. 49–79.

[25] A. Gabih, H. Kondakji, J. Sass & R. Wunderlich, Expert opinions and logarithmic utility maximization in a market with Gaussian drift, *Communications on Stochastic Analysis* **8** (2014), no. 1, pp. 27–47.

[26] A. Gabih, H. Kondakji & R. Wunderlich, Asymptotic filter behavior for high-frequency expert opinions in a market with Gaussian drift (2018).
Available on arXiv: `https://arxiv.org/abs/1812.03453`.

[27] L. Garlappi, R. Uppal & T. Wang, Portfolio selection with parameter and model uncertainty: A multi-prior approach, *The Review of Financial Studies* **20** (2007), no. 1, pp. 41–81.

[28] I. Gilboa & D. Schmeidler, Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics* **18** (1989), no. 2, pp. 141–153.

[29] P. W. Glynn, Diffusion approximations, in *Stochastic Models*, vol. 2 of *Handbooks in Operations Research and Management Science*, chap. 4, Elsevier (1990), pp. 145–198.

[30] J. Grandell, *Aspects of Risk Theory*, Springer-Verlag New York (1991).

[31] T. Honda, Optimal portfolio choice for unobservable and regime-switching mean returns, *Journal of Economic Dynamics and Control* **28** (2003), no. 1, pp. 45–78.

[32] R. A. Horn & C. R. Johnson, *Matrix Analysis*, Cambridge University Press (1990).

[33] D. L. Iglehart, Diffusion approximations in collective risk theory, *Journal of Applied Probability* **6** (1969), no. 2, pp. 285–292.

[34] I. Karatzas, J. P. Lehoczky, S. E. Shreve & G.-L. Xu, Martingale and duality methods for utility maximization in an incomplete market, *SIAM Journal on Control and Optimization* **29** (1991), no. 3, pp. 702–730.

[35] I. Karatzas & X. Zhao, *Bayesian Adaptive Portfolio Optimization*, chap. 17, Handbooks in Mathematical Finance: Option Pricing, Interest Rates and Risk Management, Cambridge University Press (2001), pp. 632–669.

[36] F. H. Knight, *Risk, Uncertainty and Profit*, Houghton Mifflin, Boston (1921).

[37] D. Kramkov & W. Schachermayer, The asymptotic elasticity of utility functions and optimal investment in incomplete markets, *The Annals of Applied Probability* **9** (1999), no. 3, pp. 904–950.

[38] D. Kramkov & W. Schachermayer, Necessary and sufficient conditions in the problem of optimal investment in incomplete markets, *The Annals of Applied Probability* **13** (2003), no. 4, pp. 1504–1516.

[39] V. Kučera, A review of the matrix Riccati equation, *Kybernetika* **9** (1973), no. 1, pp. 42–61.

[40] P. Lakner, Utility maximization with partial information, *Stochastic Processes and their Applications* **56** (1995), no. 2, pp. 247–273.

[41] P. Lakner, Optimal trading strategy for an investor: the case of partial information, *Stochastic Processes and their Applications* **76** (1998), no. 1, pp. 77–97.

[42] R. S. Liptser & A. N. Shiryaev, *Statistics of Random Processes I: General Theory*, vol. 5 of *Applications of Mathematics*, Springer-Verlag New York (1974).

[43] R. C. Merton, Lifetime portfolio selection under uncertainty: the continuous-time case, *The Review of Economics and Statistics* **51** (1969), no. 3, pp. 247–257.

[44] B. Øksendal & A. Sulem, A game theoretic approach to martingale measures in incomplete markets, *Surveys of Applied and Industrial Mathematics (TVP Publishers, Moscow)* **15** (2008), pp. 18–24.

[45] B. Øksendal & A. Sulem, Robust stochastic control and equivalent martingale measures, in *Stochastic Analysis with Financial Applications*, vol. 65 of *Progress in Probability*, Springer Basel (2011), pp. 179–189.

[46] B. G. Pachpatte, *Inequalities for Differential and Integral Equations*, vol. 197 of *Mathematics in Science and Engineering*, Academic Press (1997).

[47] G. Pflug, A. Pichler & D. Wozabal, The $1/N$ investment strategy is optimal under high model ambiguity, *Journal of Banking & Finance* **36** (2012), no. 2, pp. 410–417.

[48] H. Pham, X. Wei & C. Zhou, Portfolio diversification and model uncertainty: a robust dynamic mean-variance approach (2018).
Available on arXiv: `https://arxiv.org/abs/1809.01464`.

[49] W. Putschögl & J. Sass, Optimal consumption and investment under partial information, *Decisions in Economics and Finance* **31** (2008), no. 2, pp. 137–170.

[50] M.-C. Quenez, Optimal portfolio in a multiple-priors model, in R. C. Dalang, M. Dozzi & F. Russo, eds., *Seminar on Stochastic Analysis, Random Fields and Applications IV*, vol. 58 of *Progress in Probability*, Birkhäuser, Basel (2004), pp. 291–321.

[51] U. Rieder & N. Bäuerle, Portfolio optimization with unobservable Markov-modulated drift process, *Journal of Applied Probability* **42** (2005), no. 2, pp. 362–378.

[52] J. Rosiński & Z. Suchanecki, On the space of vector-valued functions integrable with respect to the white noise, *Colloquium Mathematicae* **43** (1980), no. 1, pp. 183–201.

[53] M. Salgado, R. Middleton & G. C. Goodwin, Connection between continuous and discrete Riccati equations with applications to Kalman filtering, *IEE Proceedings D – Control Theory and Applications* **135** (1988), no. 1, pp. 28–34.

[54] J. Sass & U. G. Haussmann, Optimizing the terminal wealth under partial information: The drift process as a continuous time Markov chain, *Finance and Stochastics* **8** (2004), no. 4, pp. 553–577.

[55] J. Sass, D. Westphal & R. Wunderlich, Expert opinions and logarithmic utility maximization for multivariate stock returns with Gaussian drift, *International Journal of Theoretical and Applied Finance* **20** (2017), no. 4, 1750022. 41 pages.

[56] J. Sass, D. Westphal & R. Wunderlich, Diffusion approximations for expert opinions in a financial market with Gaussian drift (2018).
Available on arXiv: `https://arxiv.org/abs/1807.00568`.

[57] A. Schied, Optimal investments for robust utility functionals in complete market models, *Mathematics of Operations Research* **30** (2005), no. 3, pp. 750–764.

[58] A. Schied, Optimal investments for risk- and ambiguity-averse preferences: a duality approach, *Finance and Stochastics* **11** (2007), no. 1, pp. 107–129.

[59] D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica* **57** (1989), no. 3, pp. 571–587.

[60] H. Schmidli, *Risk Theory*, Springer International Publishing (2017).

[61] K. Schöttle, R. Werner & R. Zagst, Comparison and robustification of Bayes and Black–Litterman models, *Mathematical Methods of Operations Research* **71** (2010), no. 3, pp. 453–475.

[62] A. N. Shiryaev, *Probability*, Springer-Verlag New York, 2nd edn. (1996).

[63] S.-D. Wang, T.-S. Kuo & C.-F. Hsu, Trace bounds on the solution of the algebraic matrix Riccati and Lyapunov equation, *IEEE Transactions on Automatic Control* **31** (1986), no. 7, pp. 654–656.

[64] D. Westphal, *Expert Opinions and Logarithmic Utility Maximization for Multivariate Stock Returns with Gaussian Drift*, Master's thesis, Technische Universität Kaiserslautern (2015).

[65] W. M. Wonham, On a matrix Riccati equation of stochastic control, *SIAM Journal on Control* **6** (1968), no. 4, pp. 681–697.

[66] J. Zabczyk, A note on the exponential stability of a matrix Riccati equation of stochastic control, *Kybernetika* **11** (1975), no. 3, pp. 218–222.

[67] D. Zawisza, A note on the worst case approach for a market with a stochastic interest rate, *Applicationes Mathematicae* **45** (2018), no. 2, pp. 151–160.

Scientific and Professional Career

06/2010	Abitur at the Gymnasium Farmsen, Hamburg
10/2010–03/2014	Bachelor studies in Mathematics with minor Economics at the TU Kaiserslautern (degree: B. Sc.)
07/2013–12/2013	Exchange semester at the University of Auckland, New Zealand
10/2013–09/2015	Master studies in Mathematics with specialization Financial Mathematics at the TU Kaiserslautern (degree: M. Sc.)
since 10/2015	Ph.D. student of Prof. Dr. Jörn Sass at the TU Kaiserslautern
11/2011–02/2013	Undergraduate assistant at the TU Kaiserslautern
04/2014–08/2015	Graduate assistant at the TU Kaiserslautern
since 10/2015	Scientific assistant at the TU Kaiserslautern

Wissenschaftlicher und Beruflicher Werdegang

06/2010	Abitur am Gymnasium Farmsen, Hamburg
10/2010–03/2014	Bachelorstudium in Mathematik mit Nebenfach Wirtschaftswissenschaften an der TU Kaiserslautern (Abschluss: B. Sc.)
07/2013–12/2013	Auslandssemester an der University of Auckland, Neuseeland
10/2013–09/2015	Masterstudium in Mathematik mit Vertiefung Finanzmathematik an der TU Kaiserslautern (Abschluss: M. Sc.)
seit 10/2015	Doktorandin von Prof. Dr. Jörn Sass an der TU Kaiserslautern
11/2011–02/2013	Studentische Hilfskraft an der TU Kaiserslautern
04/2014–08/2015	Wissenschaftliche Hilfskraft an der TU Kaiserslautern
seit 10/2015	Wissenschaftliche Mitarbeiterin an der TU Kaiserslautern