Mathematics and Financial Economics (2022) 16:367–397
https://doi.org/10.1007/s11579-022-00312-w

Robust utility maximizing strategies under model uncertainty and their convergence

Jörn Sass[1] · Dorothee Westphal[1]

Received: 12 September 2020 / Accepted: 22 January 2022 / Published online: 11 March 2022

Abstract
In this paper we investigate a utility maximization problem with drift uncertainty in a multivariate continuous-time Black–Scholes type financial market which may be incomplete. We impose a constraint on the admissible strategies that prevents a pure bond investment and we include uncertainty by means of ellipsoidal uncertainty sets for the drift. Our main results consist firstly in finding an explicit representation of the optimal strategy and the worst-case parameter, secondly in proving a minimax theorem that connects our robust utility maximization problem with the corresponding dual problem. Thirdly, we show that, as the degree of model uncertainty increases, the optimal strategy converges to a generalized uniform diversification strategy.

Keywords Portfolio optimization · Drift uncertainty · Minimax theorems · Diversification

Mathematics Subject Classification 91G10 · 91B16 · 93E20

1 Introduction

Model uncertainty is a challenge that is inherent in many applications of mathematical models. 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. A natural way to deal with model uncertainty is to consider worst-case optimization.

Model uncertainty, also called *Knightian uncertainty* in reference to the seminal book by Knight [10], has been addressed in numerous papers. Gilboa and Schmeidler [9] and

✉ Jörn Sass
sass@mathematik.uni-kl.de

Dorothee Westphal
westphal@mathematik.uni-kl.de

1 Department of Mathematics, Technische Universität Kaiserslautern, Kaiserslautern, Germany

Schmeidler [27] formulate rigorous axioms on preference relations that account for risk aversion and 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 [4] give a continuous-time extension of this multiple-priors utility. In Maccheroni et al. [15] the authors thoroughly axiomatize the robust approach to utility maximization via so-called ambiguity-averse preferences.

Optimal investment decisions under such preferences are investigated in Quenez [23] and Schied [25]. An extension of those results by means of a duality approach is given in Schied [26]. Uncertainty about both drift and volatility in a continuous-time Brownian framework under multiple priors is studied by Lin and Riedel [13]. Further papers addressing drift uncertainty in financial markets are Garlappi et al. [8] and Biagini and Pınar [2]. The latter also focuses on ellipsoidal uncertainty sets, as we do in this work. Neufeld and Nutz [18] incorporate jumps of the price process by considering a Lévy processes setup.

A relation between model uncertainty and portfolio diversification is investigated in a recent paper by Pham et al. [22]. Pflug et al. [21] study a one-period risk minimization problem under model uncertainty and show convergence of the optimal strategy to the uniform diversification strategy. Our results generalize these findings to a continuous-time utility maximization problem and provide an explanation for the good performance of the uniform diversification strategy also in a continuous-time setting.

The optimization problem that we address here is a utility maximization problem in a continuous-time financial market. 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 $U : \mathbb{R}_+ \to \mathbb{R}$ is a utility function, X_T^π denotes the terminal wealth that is achieved when using strategy π, and $\mathcal{A}(x_0)$ is the class of admissible strategies starting with initial capital x_0. Merton [16] solves this problem for power and logarithmic utility in a multivariate financial market model and gives a corresponding optimal strategy. However, the setup of the problem assumes that an investor knows the market parameters, in particular the drift μ of asset returns. This is a rather unrealistic assumption since drift parameters are notoriously difficult to estimate. To obtain strategies that are robust with respect to a possible misspecification of the drift we consider 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*. Our aim is to study the structure of optimal strategies, as well as their asymptotic behavior as the uncertainty set K increases. 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. We focus on ellipsoidal uncertainty sets K, see (4).

Our main results consist firstly in finding 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. Secondly, by using this explicit representation, 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. Thirdly, we show that the optimal strategy converges to a generalized uniform diversification strategy. In case of K being a ball, this is the equal weight strategy, corresponding to uniform diversification. This result is somewhat surprising since in the limit the optimal strategy does not depend on the volatility structure of the assets anymore. In that sense, our results help to explain the popularity of uniform diversification strategies by the presence of uncertainty in the model.

The paper is organized as follows. In Sect. 2 we state our multivariate, possibly incomplete, Black–Scholes type financial market model and introduce the robust utility maximization problem. Our main results are given in Sect. 3, where we solve our optimization problem for power and logarithmic utility. The main idea is to solve the dual problem explicitly and to show then that the solution forms a saddle point of the problem. We give representations of the optimal strategy and the worst-case drift parameter and prove a minimax theorem. In Sect. 4 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. Section 5 gives an outlook on more general financial market models with stochastic drift processes for which we state a suitable problem formulation. Our results can then be used to derive an explicit representation of the optimal strategy as well as a minimax theorem also in the more general model. For better readability, all proofs are collected in Appendix A.
Notation. 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. 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 $x\in\mathbb{R}^d$ is a vector, $\|x\|$ denotes the Euclidean norm of x.

2 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 risk-free asset S^0 is of the form $S_t^0=\mathrm{e}^{rt}$, $t\in[0,T]$, where $r\in\mathbb{R}$ is the constant risk-free interest rate. Aside from the risk-free asset, investors can also invest in $d\geq 2$ risky assets. Their return process $R=(R^1,\dots,R^d)^\top$ is defined by

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

where $W=(W_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion under $\mathbb{P}$ with $m\geq d$, allowing for incomplete markets. Further, $\nu\in\mathbb{R}^d$ and $\sigma\in\mathbb{R}^{d\times m}$, where we assume that σ has full rank equal to d.

We introduce model uncertainty by assuming that the true drift of the stocks is only known to be an element of some set $K \subseteq \mathbb{R}^d$ with $\nu \in K$ and that investors want to maximize their worst-case expected utility when the drift takes values within K. The value ν can be thought of as an estimate for the drift that was for instance obtained from historical stock prices. Changing the drift from ν to some $\mu \in K$ can 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 $\mathbb{P} = \mathbb{P}^\nu$.

An investor's trading decisions 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 invested in asset i at time t. The corresponding wealth process $(X_t^\pi)_{t\in[0,T]}$ given initial wealth $x_0 > 0$ can then be described by the stochastic differential equation

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

for any $\mu \in K$. We require trading strategies to be $\mathbb{F}^R$-adapted, where we have $\mathbb{F}^R = (\mathcal{F}_t^R)_{t\in[0,T]}$ for $\mathcal{F}_t^R = \sigma((R_s)_{s\in[0,t]})$. The admissibility set is defined as

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

Our robust portfolio optimization problem can then be formulated as

$$\overline{V}(x_0) = \sup_{\pi\in\mathcal{A}(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big], \tag{1}$$

where U_γ is a power or logarithmic utility function, i.e. $U_\gamma : \mathbb{R}_+ \to \mathbb{R}$ for any $\gamma \in (-\infty, 1)$, where $U_\gamma(x) = \frac{x^\gamma}{\gamma}$ for $\gamma \neq 0$ denotes power utility, $U_0(x) = \log(x)$ logarithmic utility.

2.2 Constraint on the admissible strategies

In the following, our aim is to investigate problem (1) in detail. 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. This result has been shown in a similar setting by Biagini and Pınar [2, Sec. 3.1–3.2] who address in addition to the finite horizon setting also the case with an infinite time horizon.

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*

$$\sup_{\pi\in\mathcal{A}(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]. \tag{2}$$

This observation implies that as the level of uncertainty about the true drift parameter exceeds a certain threshold, it is optimal for investors to not invest anything in the stocks.

Remark 2.2 Proposition 2.1 could be reformulated in terms of robust utility functionals by assuming only that a martingale measure is in the ambiguity set. The statement of the proposition is in line with Øksendal and Sulem [19,20] where 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 [29] that, if the model allows for stochastic interest rate r, the optimal strategy does not invest exclusively in the bond. Lin and Riedel [14] show that, when there is a large degree of uncertainty about interest rates, the investor even puts all money in the risky assets.

Investing everything in the risk-free asset is a sensible but very extreme reaction to model uncertainty. We are interested in finding out which strategies are reasonable under high model uncertainty if investors still want to invest a part of their wealth into the risky assets, or, alternatively, if they are forced to invest due to some external requirements. For that purpose, we introduce a constraint on our strategies that prevents investors from solely investing in the bond. Consider for some $h > 0$ the admissibility set

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

We do not want to exclude short-selling, so negative entries of π are possible. 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 among the risky assets. 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. Moreover, it has been studied in DeMiguel et al. [6] how constraining the norm of portfolio weight vectors in a one-period model can improve portfolio performance in the presence of estimation errors.

Remark 2.3 The admissibility set $\mathcal{A}_h(x_0)$ might seem unnecessarily restrictive at first glance. Instead of fixing $\langle \pi_t, \mathbf{1}_d\rangle = h$ one might 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. It is intuitively clear that, when uncertainty is large, investors seek to invest as little as possible in the risky assets. Therefore, we consider optimization among strategies in $\mathcal{A}_h(x_0)$ and use our results to show that enlarging the class of admissible strategies asymptotically does not change the value of the optimization problem, see Sect. 4.2.

3 A duality approach

In this section we solve for power or logarithmic utility U_γ and for specific uncertainty sets K the optimization problem

$$\sup_{\pi\in\mathcal{A}_h(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]. \tag{3}$$

Remark 3.1 In the situation with logarithmic utility and uncertainty sets that are balls in some p-norm, $p \in [1, \infty)$, it is possible to carry over methods from a one-period risk minimization problem as in Pflug et al. [21] to our continuous-time robust utility maximization problem. If $K = \{\mu \in \mathbb{R}^d \mid \|\mu - \nu\|_p \leq \kappa\}$, then for every $\varepsilon > 0$ there exists a $\kappa_0 > 0$ such that for all $\kappa \geq \kappa_0$ the strategy $\pi^*(\kappa)$ that is optimal for

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

satisfies

$$\left\| \frac{1}{T} \int_0^T \Big(\pi_s^*(\kappa) - \frac{h}{d}\mathbf{1}_d\Big) \mathrm{d}s \right\|_q < \varepsilon,$$

where $q \in (1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$. See Westphal [28, Thm. 3.4] for a proof. This shows that the optimal strategy among the deterministic ones converges, as uncertainty increases, to a uniform diversification strategy π^u with $\pi_t^u = \frac{h}{d}\mathbf{1}_d$ for every $t \in [0, T]$. Hence, as uncertainty about the true drift parameter goes to infinity, investors split the proportion h of their money more and more evenly among all risky assets.

This approach has several drawbacks. Firstly, we can follow the ideas from Pflug et al. [21] in continuous time only for logarithmic utility and uncertainty sets K that are balls in p-norm. Secondly, we have to restrict to the class of deterministic strategies to be able to use their methods. However, it is by no means clear in the first place that an optimal strategy to our problem should be a deterministic one. In fact, in many worst-case optimization problems it is even beneficial to use randomized strategies, see Delage et al. [5]. And lastly, the above result does not yield an explicit solution to the robust optimization problem, it only gives asymptotic results for large levels of uncertainty. 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.

We study the case where the uncertainty set is an ellipsoid in $\mathbb{R}^d$ centered around the reference parameter ν, i.e.

$$K = \big\{\mu \in \mathbb{R}^d \;\big|\; (\mu - \nu)^\top \Gamma^{-1}(\mu - \nu) \leq \kappa^2\big\}. \tag{4}$$

Here, $\kappa > 0$, $\nu \in \mathbb{R}^d$, and $\Gamma \in \mathbb{R}^{d \times d}$ is symmetric and positive definite. The matrix Γ determines the shape of the ellipsoid, the value of κ its size. Higher values of κ correspond to more uncertainty about the true drift.

By means of Γ we can model that some (linear combinations of) drifts are known at a higher degree of accuracy than others. A special case discussed in the literature is $\Gamma = \sigma\sigma^\top$, see e.g. Biagini and Pınar [2]. But also different forms of Γ can be motivated. For $\Gamma = I_d$ we simply get a ball in the Euclidean norm with radius κ and center ν. By setting Γ equal to a diagonal matrix different from the identity we can give different weights to the uncertainty of the single asset drifts.

More generally, assume that the reference drift parameter ν is obtained as the value of an unbiased estimator $\hat{\mu}$ for the true drift, say from observing historical returns. Then the covariance matrix $\mathrm{cov}(\hat{\mu})$ is a reasonable choice for Γ, because then the uncertainty set K constitutes a natural (asymptotic) confidence region for the true drift. This flexibility in the form of Γ is especially useful for a generalization of our model to a setting with time-dependent drift and uncertainty sets, see Sect. 5, where we give a short outlook on Sass and Westphal [24]. In that follow-up work a time-dependent uncertainty set is constructed based on filtering techniques.

3.1 Solution of the non-robust problem

To solve the optimization problem (3) 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 3.2 *Consider the matrix*

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

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

The matrix D defined in the lemma above comes up naturally in 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. For better readability of the calculations below we introduce the following notation.

Definition 3.3 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 as given in Lemma 3.2 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 give the optimal strategy for the constrained optimization problem given a fixed drift μ. The possible incompleteness of the market does not complicate our approach here. The reason is that, for determining the optimal strategy, we can essentially reduce the problem to an unconstrained less-dimensional financial market where the optimal strategy can be obtained as a classical Merton strategy.

Proposition 3.4 *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 3.3.

In the proof the d-dimensional constrained problem is reduced to a $(d-1)$-dimensional unconstrained problem. Using the form of the optimal strategy in the $(d-1)$-dimensional market which is known from Merton [16] yields the following representation for the optimal expected utility from terminal wealth.

Corollary 3.5 *Let $\mu \in \mathbb{R}^d$. Then the optimal expected utility from terminal wealth is*

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

where

$$\begin{aligned} \widetilde{\sigma} &= D\sigma, \\ \widetilde{r} &= (1-h)r + he_d^\top \mu - \frac{1}{2}(1-\gamma)\|h\sigma^\top e_d\|^2, \\ \widetilde{\mu} &= D\mu - h(1-\gamma)D\sigma\sigma^\top e_d + \widetilde{r}\mathbf{1}_{d-1}. \end{aligned} \tag{5}$$

The previous results give a representation of the optimal strategy and the optimal expected utility of terminal wealth under the constraint $\langle \pi_t, \mathbf{1}_d \rangle = h$, 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

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

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 3.4. This corresponds to solving the dual problem

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

Note here that we do not know yet whether equality holds between our original problem and the corresponding dual problem. In general the solution of the dual problem may not be of great help. In the following, after deriving the solution to the dual problem, we prove a minimax theorem that establishes the desired equality. Results from the literature, e.g. from Quenez [23], do not directly carry over to our setting as we discuss in Remark 3.9 below.

3.2 The worst-case parameter

From Corollary 3.5 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 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

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

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

$$\begin{aligned}&\frac{1}{2(1-\gamma)}\mu^\top D^\top (D\sigma\sigma^\top D^\top)^{-1} D\mu + h\Big(e_d^\top \mu - (D\sigma\sigma^\top e_d)^\top (D\sigma\sigma^\top D^\top)^{-1} D\mu\Big)\\ &= \frac{1}{2(1-\gamma)}\mu^\top A\mu + hc^\top \mu\end{aligned} \tag{6}$$

on the ellipsoid K, where A and c were introduced in Definition 3.3. To make this minimization problem easier, we apply a 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 some non-singular matrix $\tau \in \mathbb{R}^{d\times d}$ such that $\Gamma = \tau\tau^\top$. The matrix τ can be obtained for example by the Cholesky decomposition. Then we can rewrite the constraint $(\mu-\nu)^\top \Gamma^{-1}(\mu-\nu) \le \kappa^2$ as

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

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 (6) as

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

Minimizing (6) in $\mu \in K$ is therefore 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 behavior of g is determined to a large extent by the matrix A from Definition 3.3. So we analyze properties of A next.

Lemma 3.6 *The matrix A is symmetric and positive semidefinite and* $\ker(A) = \operatorname{span}(\{\mathbf{1}_d\})$.

We immediately deduce that also $\tau^\top A\tau \in \mathbb{R}^{d\times d}$ 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 $g(\rho)$ on the set $B_\kappa(0)$.

Lemma 3.7 *Let $0 = \lambda_1 < \lambda_2 \le \cdots \le \lambda_d$ denote the eigenvalues of $\tau^\top A\tau$, and let further $v_1 = \frac{1}{\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}\mathbf{1}_d, v_2, \ldots, v_d \in \mathbb{R}^d$ 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 $B_\kappa(0) = \{\rho \in \mathbb{R}^d \mid \|\rho\| \le \kappa\}$ 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 $\|\rho^*\| = \kappa$.

Note that $\psi(\kappa)$ in the above lemma is the unique value in $(0, \kappa]$ that makes ρ^* lie on the boundary of $B_\kappa(0)$. In the representation $\rho^* = \sum_{i=1}^d a_i v_i$ it holds $a_1 = -\psi(\kappa)$, i.e. $\psi(\kappa)$ is the negative of the coefficient belonging to v_1. Recall that v_1 is the eigenvector to eigenvalue zero of $\tau^\top A\tau$, hence it plays an important role in the minimization of the function g above. In Sect. 4 we will study the asymptotic behavior for large uncertainty κ. It will turn out that asymptotically v_1 will be the dominant component in the representation $\rho^* = \sum_{i=1}^d a_i v_i$, a claim that we show by analyzing the asymptotic behavior of $\psi(\kappa)$. The previous lemma now yields the solution of the dual problem to our original optimization problem.

Theorem 3.8 *Let $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_d$ denote the eigenvalues of $\tau^\top A\tau$, and let further $v_1 = \frac{1}{\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}\mathbf{1}_d, v_2, \ldots, v_d \in \mathbb{R}^d$ 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

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

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

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

Remark 3.9 The preceding theorem solves the problem

$$\inf_{\mu\in K} \sup_{\pi\in\mathcal{A}_h(x_0)} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]. \tag{7}$$

This is the corresponding dual problem to our original optimization problem

$$\sup_{\pi\in\mathcal{A}_h(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big], \tag{8}$$

but in general the values of these two problems do not coincide. 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 a portfolio optimization setting that is similar to ours a minimax theorem has been shown in Quenez [23]. Here, the author applies classical techniques from Kramkov and Schachermayer [11,12] for incomplete markets and embeds them into a multiple-priors framework. However, there are two main points that distinguish our setting from the one in Quenez [23]. Firstly, the results in that paper are only shown for non-negative utility functions and therefore not directly applicable to power utility U_γ with a negative γ. Secondly, the constraint $\langle\pi_t, \mathbf{1}_d\rangle = h$ that we put on the admissible trading strategies alters the structure of attainable terminal wealths so that it would be necessary to adjust the proofs and check several technical assumptions.

In addition, note that a minimax theorem does not endow us with the form of the optimal strategy (or the worst-case drift) yet. To obtain an explicit representation of the same, we would still need to go through the calculations done in this section. In the following, we will use the explicit representation of the optimal strategy for (7) to show that it indeed also solves (8) and that in this case, the supremum and the infimum can be interchanged.

3.3 A minimax theorem

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

Lemma 3.10 *The strategy π^* from Theorem 3.8 satisfies*

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

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

The preceding lemma characterizes the strategy π^*, which is the best strategy an investor can choose when the drift of stocks is μ^*. In the following we show that, vice versa, 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^*. 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 is essential for proving our minimax theorem. 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 Ekeland and Temam [7, Ch. VI, Prop. 1.1], for example. For the reverse inequality the saddle point property is needed.

Theorem 3.11 *Let $K = \{\mu \in \mathbb{R}^d \mid (\mu-\nu)^\top\Gamma^{-1}(\mu-\nu) \le \kappa^2\}$. Then the parameter μ that attains the minimum in*

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

is μ^*, where both μ^* and π^* are defined as in Theorem 3.8. In particular, it follows that

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

The previous theorem establishes duality between our original robust utility maximization problem and the dual problem where supremum and infimum are interchanged. Additionally, we now also know the solution to our original problem. The optimal strategy for our constrained robust utility maximization problem is given in a nearly explicit way. Note that the parameter μ^* in Theorem 3.8 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.

Remark 3.12 One can think of other reasonable sets K for modelling uncertainty about the drift parameter μ. Our duality approach can also be applied to the optimization problem with

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

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 the overall average performance of the stocks, but not for the single stocks themselves. In that case, one can show that the optimal strategy for the optimization problem

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

is $(\pi_t^*)_{t\in[0,T]}$ with $\pi_t^* = \frac{h}{d}\mathbf{1}_d$ for all $t \in [0,T]$. The worst-case parameter μ^* can be determined explicitly given the eigenvalues and eigenvectors of the matrix A. Further, one can show a minimax theorem in analogy to Theorem 3.11. The optimal strategy is here just a uniform diversification strategy given the constraint on the bond investment. In the next section we show how this fits into the framework of our results for ellipsoidal uncertainty sets when we let the degree of uncertainty κ go to infinity.

4 Asymptotic behavior as uncertainty increases

In this section we consider again the setting with ellipsoidal uncertainty sets as in (4) and investigate what happens as the degree of uncertainty changes. Since K is an ellipsoid, we increase the degree of uncertainty about the true drift parameter by increasing the radius κ, a lower value of κ corresponds to a more precise knowledge of the true drift.

4.1 Limit of worst-case parameter and optimal strategy

In the following, we address in detail the asymptotic behavior of the worst-case parameter and the optimal strategy as uncertainty increases, i.e. as κ goes to infinity. To underline the dependence on the degree of uncertainty, we write $\mu^* = \mu^*(\kappa)$ and $\pi^* = \pi^*(\kappa)$ in the following.

Remark 4.1 The other asymptotic regime $\kappa \to 0$ corresponds to a more and more precise knowledge of the true drift. It is easy to see that

$$\lim_{\kappa\to 0}\mu^*(\kappa) = \nu \quad\text{and}\quad \lim_{\kappa\to 0}\pi_t^*(\kappa) = \frac{1}{1-\gamma}A\nu + hc$$

for all $t \in [0,T]$. This means that the worst-case parameter converges to the reference drift ν and the optimal strategy to the best constrained strategy, given that the drift equals ν. So we retrieve in the limit $\kappa \to 0$ the setting without model uncertainty.

We now focus on $\kappa \to \infty$. Note that the only quantity in the representation of μ^* from Theorem 3.8 that depends on κ is $\psi(\kappa)$.

Lemma 4.2 *It holds* $\lim_{\kappa\to\infty}\frac{\psi(\kappa)}{\kappa} = 1$.

From this lemma we gain insights into the asymptotic behavior of μ^*.

Proposition 4.3 *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

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

Hence, asymptotically the direction of the worst-case parameter is $-\mathbf{1}_d$. This means that, as κ tends to infinity, the worst drift which 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. We have the following result for the asymptotic behavior of the investor's optimal strategy.

Theorem 4.4 *For any $t \in [0, T]$ it holds*

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

The theorem shows that the optimal strategy $\pi^*(\kappa)$ converges as the degree of uncertainty κ goes to infinity. If $\Gamma = \sigma\sigma^\top$, then the limit is a multiple of the minimum variance portfolio. Another interesting special case is $\Gamma = I_d$, i.e. when K is simply a ball with radius κ. In that case we have

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

for any $t \in [0, T]$, hence 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. For general Γ we shall speak of a generalized uniform diversification strategy.

This asymptotic behavior of the optimal strategy is striking because the limit is independent of the volatility matrix σ. In combination with the structure of the function g in Lemma 3.7 this indicates that it might also be possible to allow for misspecified volatility. For a high level of uncertainty the optimal strategy is dominated by the matrix Γ shaping the uncertainty ellipsoid whereas both the volatility structure of the assets and the reference drift ν become negligible. This effect is caused by the investor's reaction to the worst-case drift parameter μ^* which, as shown in Proposition 4.3, behaves asymptotically like a multiple of $\mathbf{1}_d$. The best reaction from the investor's point of view is to diversify among all assets, weighted by the uncertainty structure Γ. In the special case where K is a ball, this leads to a uniform diversification strategy. This result is in line with Pflug et al. [21] who show convergence of the optimal strategy to the uniform diversification strategy in a risk minimization setting with increasing model uncertainty.

Remark 4.5 Note that plugging in $\kappa = \infty$ into the definition of the ellipsoid yields the uncertainty set $K = \mathbb{R}^d$, so that in fact every drift parameter $\mu \in \mathbb{R}^d$ is deemed possible by the investor. Then one easily obtains the worst-case utility

$$\inf_{\mu\in\mathbb{R}^d} \mathbb{E}_\mu[U_\gamma(X_T^\pi)] = \begin{cases} 0, & \gamma \in (0,1), \\ -\infty, & \gamma \in (-\infty, 0], \end{cases}$$

for *any* admissible π. Hence, every strategy performs equally bad in the limit case. In particular, plugging in $\kappa = \infty$ into the ellipsoid in the first place does not provide us with the optimal limit strategy of Theorem 4.4.

The intuition is that, as long as the uncertainty set is bounded, there exists a worst-case drift to which the investor can react in an optimal way. Nevertheless, when uncertainty goes to infinity, also the expected utility achieved by the best strategy will be driven to $-\infty$ in case that $\gamma \in (-\infty, 0]$, respectively to zero in case $\gamma \in (0, 1)$.

4.2 Relaxing the investment constraint

We use the above results 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

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

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 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.

Proposition 4.6 *Define for $h > 0$ the admissibility set*

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

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

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

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

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 from Definition 3.3.

Lemma 4.7 *For the matrix A and the vector c we have*

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

The next proposition gives a result similar to Proposition 4.6 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 4.8 *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

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

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.

4.3 Risk aversion and speed of convergence

As the class of admissible strategies we now take again

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

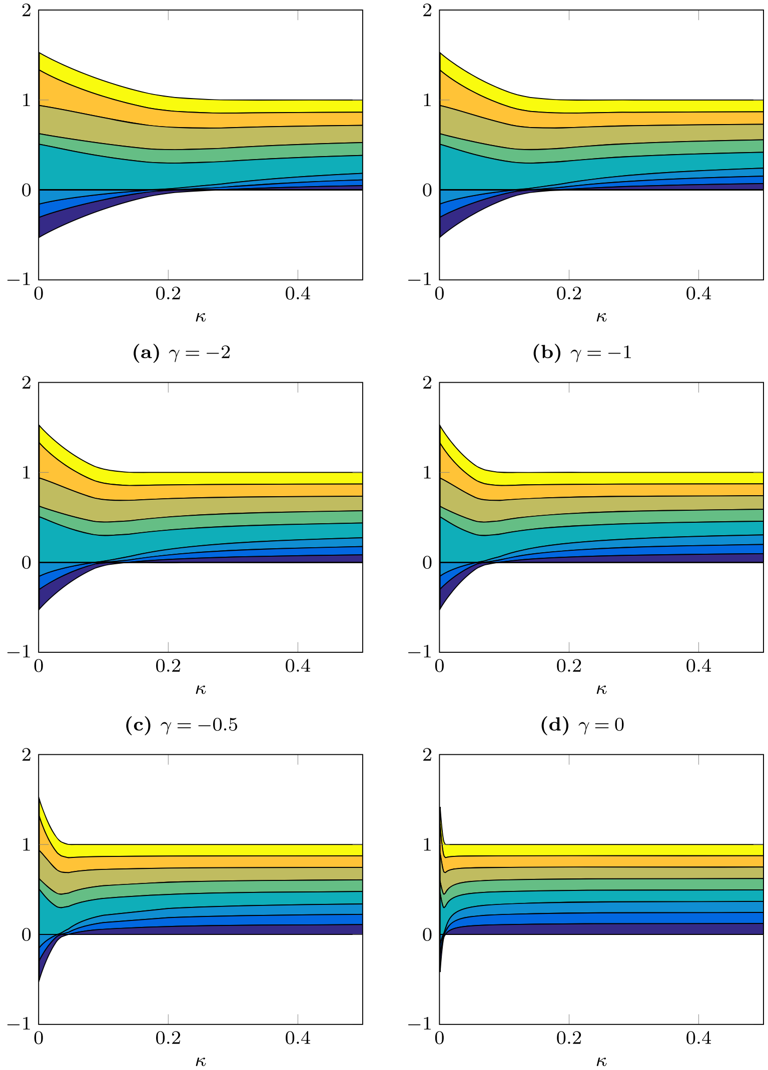
for some $h > 0$. We have seen in Sect. 4.1 that the optimal strategy $\pi^*(\kappa)$ for our robust optimization problem with ellipsoidal uncertainty sets K converges as the level of uncertainty κ goes to infinity. If the uncertainty set K is a ball, then the limit is a uniform diversification strategy $\frac{h}{d}\mathbf{1}_d$. In the following, we illustrate this convergence by an example and investigate which influence the risk aversion parameter γ has on the speed of convergence. Note that for our class of utility functions, the value $1-\gamma$ is equal to the Arrow–Pratt measure of relative risk aversion. The smaller γ is, the more risk-averse is the investor.

Example 4.9 We consider a market with $d = 8$ risky assets. The volatility matrix has the form

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

Investors use strategies from $\mathcal{A}_h(x_0)$ with $h = 1$. Further, we take $\Gamma = I_d$ and $\nu = \frac{3}{10}\mathbf{1}_d$ as parameters of the uncertainty ellipsoid. Note that for this choice of the parameter ν the optimal strategy in the situation without model uncertainty, i.e. with $\kappa = 0$, does not depend on γ. We then compute the constant optimal portfolio composition $\pi^*(\kappa)$ based on different values of γ and for all $\kappa \in (0, 0.5)$, and plot the result in Fig. 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 strategy π^* is negative in some components. This leads to an overall investment larger than one on the positive side. As κ becomes larger, the composition gets closer and closer to the uniform diversification vector. When comparing the different subplots one sees that the convergence is faster for higher values of γ, an effect that has been shown to hold in general, see Westphal [28, Rem. 5.9]. 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, there is more diversification, which can even be amplified by the constraint using $h = 1$, and thus 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)** $\gamma = 0.9$

Fig. 1 Optimal portfolio composition π^* plotted against κ for different values of γ. The model parameters are given in Example 4.9. 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

4.4 Measures of robustness performance

We have seen that introducing uncertainty in our utility maximization problem leads to more diversified strategies. The question arises what an investor gains from using robust strategies and what downside comes with behaving in a robust way in situations where it is not necessary. These two antithetic effects can be rated by the measures *cost of ambiguity* and *reward for distributional robustness* that have been studied in a different context in Analui [1, Sec. 3.4].

For our robust maximization problem, the center ν of the uncertainty ellipsoid can be seen as an estimation for the true drift of the stocks. If an investor was sure that the estimation was correct, she would simply maximize $\mathbb{E}_\nu[U_\gamma(X_T^\pi)]$. From Proposition 3.4 we know that the optimal strategy is then of the form $(\hat{\pi}_t)_{t\in[0,T]}$ with

$$\hat{\pi}_t = \frac{1}{1-\gamma} A\nu + hc \tag{9}$$

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

$$\pi_t^* = \frac{1}{1-\gamma} A\mu^* + hc \tag{10}$$

for all $t \in [0, T]$, see Theorem 3.11. We now define measures for the robustness performance that consider the difference in the corresponding certainty equivalents when using $\hat{\pi}$ or π^*.

Definition 4.10 We define the *cost of ambiguity* as

$$\mathrm{COA} = U_\gamma^{-1}\big(\mathbb{E}_\nu\left[U_\gamma\big(X_T^{\hat{\pi}}\big)\right]\big) - U_\gamma^{-1}\big(\mathbb{E}_\nu\left[U_\gamma\big(X_T^{\pi^*}\big)\right]\big)$$

and the *reward for distributional robustness* as

$$\mathrm{RDR} = U_\gamma^{-1}\big(\mathbb{E}_{\mu^*}\left[U_\gamma\big(X_T^{\pi^*}\big)\right]\big) - U_\gamma^{-1}\big(\mathbb{E}_{\mu^*}\left[U_\gamma\big(X_T^{\hat{\pi}}\big)\right]\big).$$

The cost of ambiguity captures how big the loss in the certainty equivalent 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 and that U_γ^{-1} is a strictly increasing function, hence COA is non-negative. The reward for distributional robustness reflects 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 μ^*.

Remark 4.11 A different definition of COA and RDR is possible where one measures the difference in expected utility rather than the difference of the certainty equivalents. The asymptotic behavior of the reward for distributional robustness for large uncertainty is then heavily affected by the parameter γ of the investor’s utility function. In particular, as κ goes to infinity, the reward for distributional robustness goes to zero if $\gamma > 0$ and to infinity if $\gamma < 0$.

Proposition 4.12 *Independently of* $\gamma \in (-\infty, 1)$ *it always holds* COA $\geq$ RDR.

Furthermore, COA and RDR converge as κ goes to infinity. We write COA(κ) and RDR(κ) to emphasize the dependence on the degree of uncertainty.

Proposition 4.13 *As* κ *goes to infinity,* COA(κ) *converges to a non-negative limit and* RDR(κ) *goes to zero.*

Figure 2 illustrates the behavior of COA and RDR in dependence on the level of uncertainty κ. We consider a market with $d = 8$ stocks, where the underlying market parameters are those from Example 4.9. The figure shows COA and 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 COA and RDR become smaller as γ increases.

We observe that the qualitative behavior of COA and RDR is the same for any value of the risk aversion coefficient γ. For any fixed γ and κ, RDR is always less than COA, a property that we have proven in Proposition 4.12. As κ increases, COA goes to a finite positive limit, whereas RDR tends to zero, as we have shown in Proposition 4.13.

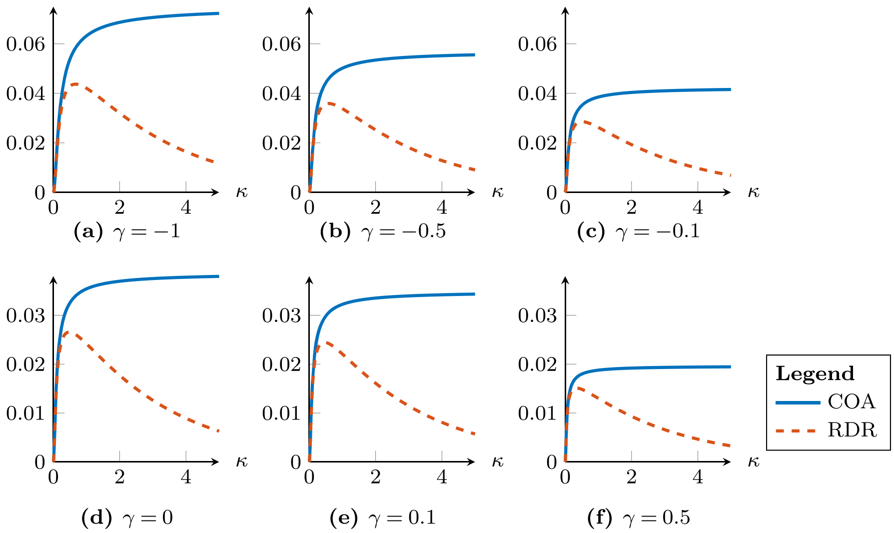


Fig. 2 The behavior of COA and RDR plotted against uncertainty radius κ for different values of the risk aversion coefficient γ. The parameters are those from Example 4.9

5 Outlook on stochastic drift and time-dependent uncertainty sets

In this section we want to give a brief outlook on how the results of this paper can be applied also in more general financial market models with a stochastic drift process. This generalization is the topic of our follow-up work Sass and Westphal [24]. Here we only give a short outline of the setup to illustrate the relevance of this work.

In Sass and Westphal [24] the results of the present paper are generalized to a financial market with a stochastic drift process and time-dependent uncertainty sets $(K_t)_{t\in[0,T]}$. This is motivated by the idea that information about the hidden drift process, as e.g. obtained from filtering techniques, might change over time. A surplus of information should then be reflected in a smaller uncertainty set. More precisely, we assume that under the reference measure returns follow the dynamics

$$\mathrm{d}R_t = \nu_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t,$$

where the reference drift $(\nu_t)_{t\in[0,T]}$ is adapted to the filtration $(\mathcal{G}_t)_{t\in[0,T]}$ representing the investor's information. This is justified by a separation principle where one performs a filtering step before solving the optimization problem, i.e. $(\nu_t)_{t\in[0,T]}$ represents the investor's filter for the drift process. We introduce a time-dependent uncertainty set $(K_t)_{t\in[0,T]}$ that is a set-valued stochastic process adapted to $(\mathcal{G}_t)_{t\in[0,T]}$, meaning that the investor knows the realization of K_t at time t.

It is not obvious how to set up a worst-case optimization problem in this time-dependent setting. The problem lies in the fact that the realization of the uncertainty sets $(K_t)_{t\in[0,T]}$ is not known initially but gets revealed over time. A worst-case drift process $(\mu_t)_{t\in[0,T]}$ is characterized by being the worst one with the property that $\mu_t \in K_t$ for all $t \in [0, T]$. However, optimization with respect to this worst-case drift process is not feasible for an investor since it is not known initially. Instead, it makes sense to consider the following local approach. For any fixed $t \in [0, T]$, the current uncertainty set K_t is known. Given this

K_t, investors take model uncertainty into account by assuming that in the future the worst possible drift process having values in K_t will be realized, i.e. the worst drift process from the class

$$\mathcal{K}^{(t)} = \left\{(\mu_s^{(t)})_{s\in[t,T]} \,\middle|\, \mu_s^{(t)} \in K_t \text{ and } \mu_s^{(t)} \text{ is } \mathcal{G}_t\text{-measurable for each } s \in [t,T]\right\}.$$

Investors then solve at each time $t \in [0, T]$ the local optimization problem

$$\sup_{\pi^{(t)}\in\mathcal{A}_h(t,x)} \inf_{\mu^{(t)}\in\mathcal{K}^{(t)}} \mathbb{E}_{\mu^{(t)}}\left[U\left(X_T^{t,x,\pi^{(t)}}\right)\right]. \tag{11}$$

Here, we write $X_s^{t,x,\pi^{(t)}}$ for the wealth at $s \in [t, T]$ when starting at t with x and using strategy $\pi^{(t)} \in \mathcal{A}_h(t, x)$, where the admissibility set is defined analogously to $\mathcal{A}_h(x_0)$ for strategies starting at t. This leads to an optimal strategy $(\pi_s^{(t),*})_{s\in[t,T]}$. In our continuous-time setting this decision will be revised as soon as K_t changes, possibly continuously in time. The realized optimal strategy of the investor is then given by $\pi_t^* = \pi_t^{(t),*}$ for any $t \in [0, T]$.

This setup of the local optimization problems is reasonable from an investor's point of view. The uncertainty sets K_t change continuously in time due to new incoming information along with return observations, for example. Naturally, the optimal strategy of the investor will then also be adapted continuously. In Sass and Westphal [24] it is shown in detail how the results of this paper can be used to solve the above described more complicated problem. An explicit representation of the optimal strategy and a minimax theorem can be derived. Those results then also apply to much more general financial market models. The convergence results from Sect. 4, however, do not have a straightforward analogon in the setting with time-dependent uncertainty sets.

Remark 5.1 Initially it is not clear whether we have an inconsistent control problem, cf. Björk et al. [3], in our original formulation (3). But for the special case of (11) with a constant uncertainty set K, the results in Sass and Westphal [24] show that one obtains at time t the same optimal risky fractions as when starting at time 0. In combination with the Bellman principle, which implies that at time t we only need the information $X_t = x$, this proves that our robust utility maximization problem with optimal solution π^* obtained in Sect. 3 is time-consistent. A generalization to allowing for more probability measures than those corresponding to a constant drift in a formulation based on a robust utility functional may raise consistency issues and would need assumptions on the structure of this set. This may then be treated as in Müller [17] under appropriate conditions.

Acknowledgements The authors thank two anonymous referees for helpful comments and suggestions that improved this paper.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.The data of our numerical examples are available from the corresponding author upon reasonable request.

A Proofs

For better readability of the paper, all proofs are collected in this appendix.

Proof of Proposition 2.1 Let $\mu \in K$ and $\pi \in \mathcal{A}(x_0)$. We consider the case $\gamma = 0$ first. The expected logarithmic utility of terminal wealth under measure $\mathbb{P}^{\mu}$ is

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

Since the vector $r\mathbf{1}_d$ is an element of the set K, we immediately see that

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

so we can deduce that the trivial strategy $\pi \equiv 0$ is optimal for (2), 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

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

and we can rewrite the expectation on the right-hand side as

$$\mathbb{E}_{r\mathbf{1}_d}\bigg[\exp\Big(\gamma\int_0^T \pi_t^{\top}\sigma\,\mathrm{d}W_t^{r\mathbf{1}_d} - \frac{1}{2}\gamma^2\int_0^T \|\sigma^{\top}\pi_t\|^2\,\mathrm{d}t\Big)\exp\Big(-\frac{1}{2}\gamma(1-\gamma)\int_0^T \|\sigma^{\top}\pi_t\|^2\,\mathrm{d}t\Big)\bigg].$$

Thus,

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

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

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

Again, as for logarithmic utility, the trivial strategy $\pi \equiv 0$ is optimal for (2) if $r\mathbf{1}_d \in K$, since the zero strategy leads exactly to expected power utility $\frac{x_0^{\gamma}}{\gamma}\mathrm{e}^{\gamma r T}$. □

Proof of Lemma 3.2 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 $a_1 = \cdots = a_{d-1} = 0$. So, the rows of $D\sigma$ are independent, and $\operatorname{rank}(D\sigma) = d-1$. □

Proof of Proposition 3.4 Let $\pi \in \mathcal{A}_h(x_0)$. Then $\pi_t^d = h - \sum_{i=1}^{d-1} \pi_t^i$ for all $t \in [0, T]$, therefore we can transform

$$\pi_t^\top(\mu - r\mathbf{1}_d) = \sum_{i=1}^{d-1} \pi_t^i(\mu^i - r) + \left(h - \sum_{i=1}^{d-1} \pi_t^i\right)(\mu^d - r) = h(e_d^\top \mu - r) + \widetilde{\pi}_t^\top D\mu, \tag{A.1}$$

$$\pi_t^\top \sigma = \sum_{i=1}^{d-1} \pi_t^i \sigma_{i,\cdot} + \left(h - \sum_{i=1}^{d-1} \pi_t^i\right)\sigma_{d,\cdot} = he_d^\top \sigma + \widetilde{\pi}_t^\top D\sigma, \tag{A.2}$$

where $\widetilde{\pi}_t := \pi_t^{1:d-1}$ for all $t \in [0, T]$, and where $\sigma_{i,\cdot}$ denotes the i-th row of matrix σ. In the representation of the wealth process we first plug in (A.2) into the stochastic integral. For $\gamma \neq 0$ we perform a change of measure

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

such that under the measure $\widetilde{\mathbb{P}}$, 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. We obtain

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

By straightforward calculations using (A.1) and (A.2) the integrand in the Lebesgue integral above can be rewritten as

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

If we now substitute

$$\begin{aligned}
\widetilde{\sigma} &= D\sigma, \\
\widetilde{r} &= (1-h)r + he_d^\top \mu - \frac{1}{2}(1-\gamma)\|h\sigma^\top e_d\|^2, \\
\widetilde{\mu} &= D\mu - h(1-\gamma)D\sigma\sigma^\top e_d + \widetilde{r}\mathbf{1}_{d-1},
\end{aligned} \tag{A.3}$$

then the expected utility of terminal wealth is given by

$$\begin{aligned}&\mathbb{E}_{\mu}\big[U_{\gamma}(X_T^{\pi})\big]\\&=\frac{x_0^{\gamma}}{\gamma}\widetilde{\mathbb{E}}\Big[\exp\Big(\gamma\int_0^T\Big(\widetilde{r}+\widetilde{\pi}_t^{\top}(\widetilde{\mu}-\widetilde{r}\mathbf{1}_{d-1})-\frac{1}{2}\|\widetilde{\sigma}^{\top}\widetilde{\pi}_t\|^2\Big)\mathrm{d}t+\gamma\int_0^T\widetilde{\pi}_t^{\top}\widetilde{\sigma}\,\mathrm{d}\widetilde{W}_t\Big)\Big].\end{aligned}\tag{A.4}$$

In the case $\gamma=0$, like in the power utility case, we can represent expected utility of terminal wealth as

$$\mathbb{E}_{\mu}\big[\log(X_T^{\pi})\big]=\log(x_0)+\widetilde{r}\,T+\mathbb{E}\Big[\int_0^T\Big(\widetilde{\pi}_t^{\top}\big(\widetilde{\mu}-\widetilde{r}\mathbf{1}_{d-1}\big)-\frac{1}{2}\|\widetilde{\sigma}^{\top}\widetilde{\pi}_t\|^2\Big)\mathrm{d}t\Big],\tag{A.5}$$

where we use the same substitution with $\widetilde{r}$, $\widetilde{\mu}$ and $\widetilde{\sigma}$ as in (A.3) for $\gamma=0$.

In both cases $\gamma\neq 0$ and $\gamma=0$ we realize that the expressions in (A.4) and (A.5) 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 $\widetilde{\sigma}\in\mathbb{R}^{(d-1)\times m}$. 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 (A.4), respectively (A.5), over all admissible strategies $\widetilde{\pi}$ with values in $\mathbb{R}^{d-1}$, we know from Merton [16] that the optimal strategy is constant in time and has the form

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

We now return to our original d-dimensional market, using the relation $\pi_t=D^{\top}\widetilde{\pi}_t+he_d$, giving us the optimal strategy π with

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

□

Proof of Lemma 3.6 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 3.2, $D\sigma$ 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. Furthermore, it is easy to check that $\ker(D)=\operatorname{span}(\{\mathbf{1}_d\})$ and $\ker(D^{\top})=\{0\}$. Hence, it holds $Ax=D^{\top}(D\sigma\sigma^{\top}D^{\top})^{-1}Dx=0$ if and only if $(D\sigma\sigma^{\top}D^{\top})^{-1}Dx=0$, which is equivalent to $Dx=0$. Hence we can deduce $\ker(A)=\ker(D)=\operatorname{span}(\{\mathbf{1}_d\})$. □

Proof of Lemma 3.6 Recall that $\tau^{\top}A\tau$ has an eigenvalue $\lambda_1=0$ with a corresponding normed eigenvector of the form $v_1=\frac{1}{\|\tau^{-1}\mathbf{1}_d\|}\tau^{-1}\mathbf{1}_d$. The other eigenvalues of $\tau^{\top}A\tau$ are positive and due to symmetry we can assume that $v_1,\dots,v_d$ form an orthogonal basis of $\mathbb{R}^d$. Firstly, we observe that the gradient of g is

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

$$= \tau^\top \Big(A \Big(\frac{1}{1-\gamma} (\tau\rho + \nu) - h\sigma\sigma^\top e_d \Big) + h e_d \Big).$$

It follows that there is no $\rho \in B_\kappa(0)$ with $\nabla g(\rho) = 0$, since $\tau^\top$ is nonsingular and he_d is not in the range of $A = D^\top (D\sigma\sigma^\top D^\top)^{-1} D$. The minimum of g on $B_\kappa(0)$ is therefore attained on the boundary.

Let $\rho \in B_\kappa(0)$ be arbitrary. Since $v_1, \ldots, v_d$ form a basis of $\mathbb{R}^d$, we are able to write $\rho = \sum_{i=1}^d a_i v_i$, where $a_1, \ldots, a_d \in \mathbb{R}$ are uniquely determined. Since we know that a minimizer of the function g must lie on the boundary of $B_\kappa(0)$ we obtain the constraint

$$\kappa^2 = \|\rho\|^2 = \sum_{i=1}^d a_i^2 \tag{A.6}$$

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

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

Next, one easily sees that

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

since $A\mathbf{1}_d = 0$ by Lemma 3.6. By plugging in this representation we deduce that, when looking for the minimizer of g, we can restrict to the parameters ρ with coefficient $a_1 \le 0$. We obtain

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

and minimize this expression in $a_2, \ldots, a_d$. 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, $k = 2, \ldots, d$, is given by

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

When setting this expression equal to zero, we obtain

$$a_k = -\left(\frac{\lambda_k}{1-\gamma} - \frac{h}{\|\tau^{-1}\mathbf{1}_d\|a_1}\right)^{-1}\Big\langle h\tau^\top c + \frac{\lambda_k}{1-\gamma}\tau^{-1}\nu, v_k\Big\rangle. \tag{A.7}$$

Note that this representation does not provide the coefficients a_k explicitly since a_1 here is a function of $(a_2, \ldots, a_d)$. However, it is easy to check that 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 a strictly negative derivative 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 (A.7) together with (A.6) uniquely determines $a_1, \ldots, a_d$.

Moreover, by some straightforward calculations we see that the Hessian of $\widetilde{g}$ is of the form

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

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 are positive-definite matrices. The last summand is positive semidefinite. So we conclude that the Hessian of $\widetilde{g}$ 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 (A.7) together with (A.6), there is a global minimum of the function $\widetilde{g}$.

□

Proof of Theorem 3.8 For any fixed parameter $\mu \in \mathbb{R}^d$, Proposition 3.4 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 3.5 we have seen that minimizing the above expression in μ on the set $K = \big\{\mu \in \mathbb{R}^d \,\big|\, (\mu - \nu)^\top \Gamma^{-1}(\mu - \nu) \le \kappa^2\big\}$ is equivalent to minimizing the function $g\colon B_\kappa(0) \to \mathbb{R}$ from Lemma 3.7 in ρ and then setting $\mu = \nu + \tau\rho$. The claim now follows from Lemma 3.7 together with the representation in Proposition 3.4. □

Proof of Lemma 3.10 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

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

In conclusion,

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

and therefore

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

for all $t \in [0, T]$. □

Proof of Theorem 3.11 Since π^* is a strategy that is constant in time and deterministic, we can rewrite the expected utility of terminal wealth as

$$\mathbb{E}_\mu\big[U_\gamma(X_T^{\pi^*})\big] = \begin{cases} \frac{x_0^\gamma}{\gamma}\exp\Big(\gamma T\Big(r + (\pi_0^*)^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_0^*\|^2\Big) + \frac{1}{2}\gamma^2 T\|\sigma^\top\pi_0^*\|^2\Big), & \gamma \neq 0, \\ \log(x_0) + T\Big(r + (\pi_0^*)^\top(\mu - r\mathbf{1}_d) - \frac{1}{2}\|\sigma^\top\pi_0^*\|^2\Big), & \gamma = 0. \end{cases}$$

Obviously, for any $\gamma \in (-\infty, 1)$ the parameter $\mu \in K$ that minimizes the expression above is the parameter that minimizes $(\pi_0^*)^\top\mu$. For an arbitrary $\theta \in \mathbb{R}^d, \theta \neq 0$, an easy calculation shows that the parameter $\mu \in \mathbb{R}^d$ that minimizes $\theta^\top\mu$ such that $(\mu - \nu)^\top\Gamma^{-1}(\mu - \nu) \leq \kappa^2$ has the form

$$\widetilde{\mu} = \nu - \frac{\kappa}{\sqrt{\theta^\top\Gamma\theta}}\Gamma\theta. \tag{A.8}$$

Hence it is sufficient to show that the parameter μ^* is equal to $\widetilde{\mu}$ from (A.8) for $\theta = \pi_0^*$. Using Lemma 3.10 we have

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

and

$$\sqrt{(\pi_0^*)^\top \Gamma \pi_0^*} = \frac{h\kappa}{\psi(\kappa)\|\tau^{-1}\mathbf{1}_d\|}. \tag{A.9}$$

When rearranging the representation in Lemma 3.10 for μ^* and plugging in (A.9) we obtain

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

We conclude that μ^* is the parameter that minimizes $(\pi_0^*)^\top\mu$ over all $\mu \in K$ and therefore the worst possible parameter for the strategy π^*.

Now, 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. Then we know from Theorem 3.8 that

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

On the other hand, the fact that μ^* is the worst parameter for an investor using strategy π^* yields

$$\mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] = \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^{\pi^*})\big] \le \sup_{\pi\in\mathcal{A}_h(x_0)} \inf_{\mu\in K} \mathbb{E}_\mu\big[U_\gamma(X_T^\pi)\big]. \tag{A.10}$$

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 [7, Ch. VI, Prop. 1.1]. Consequently, the inequality in (A.10) is an equality and the claim follows. □

Proof of Lemma 4.2 By acknowledging the dependence on κ, we write $a_i(\kappa)$ for the coefficients of $\rho^* = \tau^{-1}(\mu^* - \nu)$. We have already seen in the proof of Lemma 3.7 that $a_1(\kappa) = -\psi(\kappa)$. Hence, the constraint $\|\tau^{-1}(\mu^* - \nu)\| = \kappa$ implies

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

due to orthonormality of $v_1, \ldots, v_d$. 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 $i \in \{2, \ldots, d\}$. We know that

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

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

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

where the upper bound goes to zero as κ goes to infinity. The claim now follows from the fact that $\psi(\kappa)$ is positive for each κ. □

Proof of Proposition 4.3 Using the same notation as before, as well as the result from the previous lemma, we can 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. The second claim follows immediately. □

Proof of Theorem 4.4 Recall that by Lemma 3.10 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 4.2 and Proposition 4.3. □

Proof of Proposition 4.6 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

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

In particular, we have

$$\begin{aligned}&\inf_{\mu\in K(\kappa)}\mathbb{E}_\mu\big[\log(X_T^{\pi'})\big] \leq \mathbb{E}_{\mu^*}\big[\log(X_T^{\pi'})\big]\\ &= \mathbb{E}_{\mu^*}\big[\log(X_T^{\pi})\big] + \mathbb{E}_{\mu^*}\bigg[\int_0^T \varepsilon_t\Big(\mathbf{1}_d^\top\big(\mu^*(\kappa) - r\mathbf{1}_d\big) - \frac{1}{2}\varepsilon_t\|\sigma^\top\mathbf{1}_d\|^2 - \mathbf{1}_d^\top\sigma\sigma^\top\pi_t\Big)\mathrm{d}t\bigg],\end{aligned} \tag{A.11}$$

where $\mu^* = \mu^*(\kappa)$ is the worst-case parameter from Theorem 3.8. 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 there exists a $\kappa_M > 0$ such that the second summand in (A.11) is non-positive for $\kappa \geq \kappa_M$. That is because $\varepsilon_t \geq 0$ for all $t \in [0, T]$ and

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

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. □

Proof of Lemma 4.7 Using the definition of A in Definition 3.3 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, due to $A\mathbf{1}_d = 0$ we also have

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

□

Proof of Proposition 4.8 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

$$\mathbb{E}_{\mu'}\big[U_\gamma(X_T^{\pi'})\big] \leq \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] \tag{A.12}$$

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 showing (A.12) it is sufficient to prove

$$(\pi_0')^\top(\mu' - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\pi_0'\|^2 \leq (\pi_0^*)^\top(\mu^* - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\pi_0^*\|^2. \tag{A.13}$$

Using the representation of π' we obtain

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

Here we have used the identities from Lemma 4.7. 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

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

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

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

which proves (A.13) and hence (A.12). 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. $\square$

Proof of Proposition 4.12 Since both π^* and $\hat\pi$ are constant in time and deterministic, we can show for $\gamma \neq 0$ that

$$\begin{aligned}
\mathrm{COA} = x_0\mathrm{e}^{rT}\bigg(&\exp\Big(T\Big((\hat\pi_0)^\top(\nu - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\hat\pi_0\|^2\Big)\Big) \\
&- \exp\Big(T\Big((\pi_0^*)^\top(\nu - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\pi_0^*\|^2\Big)\Big)\bigg)
\end{aligned} \tag{A.14}$$

and

$$\begin{aligned}
\mathrm{RDR} = x_0\mathrm{e}^{rT}\bigg(&\exp\Big(T\Big((\pi_0^*)^\top(\mu^* - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\pi_0^*\|^2\Big)\Big) \\
&- \exp\Big(T\Big((\hat\pi_0)^\top(\mu^* - r\mathbf{1}_d) - \frac{1-\gamma}{2}\|\sigma^\top\hat\pi_0\|^2\Big)\Big)\bigg).
\end{aligned} \tag{A.15}$$

For $\gamma = 0$ we obtain the same representations as in (A.14) and (A.15) with $\gamma = 0$. We now plug in the representations from (9), respectively (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 4.7. We obtain

$$\begin{aligned}
\frac{\mathrm{COA}}{x_0\mathrm{e}^{rT}} &= \overline{L}(\gamma,\kappa)\exp\Big(T\Big(-hr - \frac{1-\gamma}{2}h^2c^\top\sigma\sigma^\top c + hc^\top\nu + \frac{1}{2(1-\gamma)}\nu^\top A\nu\Big)\Big), \\
\frac{\mathrm{RDR}}{x_0\mathrm{e}^{rT}} &= \overline{L}(\gamma,\kappa)\exp\Big(T\Big(-hr - \frac{1-\gamma}{2}h^2c^\top\sigma\sigma^\top c + hc^\top\mu^* + \frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^*\Big)\Big),
\end{aligned}$$

where

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

Hence, we can deduce in particular that

$$\frac{\text{COA}}{\text{RDR}} = \frac{\exp\Big(T\Big(\frac{1}{2(1-\gamma)}\nu^\top A\nu + hc^\top \nu\Big)\Big)}{\exp\Big(T\Big(\frac{1}{2(1-\gamma)}(\mu^*)^\top A\mu^* + hc^\top \mu^*\Big)\Big)} \geq 1,$$

since μ^* minimizes the function $\mu \mapsto \frac{1}{2(1-\gamma)}\mu^\top A\mu + hc^\top \mu$ on the set K. □

Proof of Proposition 4.13 Firstly, note that by the same reasoning as in the proof of Theorem 3.11 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

$$\lim_{\kappa\to\infty} \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\hat{\pi}})\big] = \lim_{\kappa\to\infty} \mathbb{E}_{\mu^*}\big[U_\gamma(X_T^{\pi^*})\big] = \begin{cases} -\infty, & \gamma \leq 0, \\ 0, & \gamma > 0, \end{cases}$$

and therefore $\lim_{\kappa\to\infty} \text{RDR}(\kappa) = 0$.

For 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 $\text{COA}(\kappa)$. Since $\text{COA}(\kappa) \geq 0$ for any κ, we know that the limit is non-negative. □

References

1. Analui, B.: Multistage stochastic optimization of energy portfolios under model ambiguity. Ph.D. thesis, Universität Wien (2014)
2. Biagini, S., Pınar, M.Ç.: The robust Merton problem of an ambiguity averse investor. Math. Financ. Econ. **11**(1), 1–24 (2017)
3. Björk, T., Khapko, M., Murgoci, A.: On time-inconsistent stochastic control in continuous time. Finance Stoch. **21**, 331–360 (2017)
4. Chen, Z., Epstein, L.: Ambiguity, risk, and asset returns in continuous time. Econometrica **70**(4), 1403–1443 (2002)
5. Delage, E., Kuhn, D., Wiesemann, W.: “Dice”-sion-making under uncertainty: when can a random decision reduce risk? Manage. Sci. **65**(7), 3282–3301 (2019)
6. DeMiguel, V., Garlappi, L., Nogales, F.J., Uppal, R.: A generalized approach to portfolio optimization: improving performance by constraining portfolio norms. Manage. Sci. **55**(5), 798–812 (2009)
7. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. North-Holland Publishing Company, Amsterdam (1976)
8. Garlappi, L., Uppal, R., Wang, T.: Portfolio selection with parameter and model uncertainty: a multi-prior approach. Rev. Financ. Stud. **20**(1), 41–81 (2007)
9. Gilboa, I., Schmeidler, D.: Maxmin expected utility with non-unique prior. J. Math. Econ. **18**(2), 141–153 (1989)
10. Knight, F.H.: Risk, Uncertainty and Profit. Houghton Mifflin, Boston (1921)
11. Kramkov, D., Schachermayer, W.: The asymptotic elasticity of utility functions and optimal investment in incomplete markets. Ann. Appl. Probab. **9**(3), 904–950 (1999)
12. Kramkov, D., Schachermayer, W.: Necessary and sufficient conditions in the problem of optimal investment in incomplete markets. Ann. Appl. Probab. **13**(4), 1504–1516 (2003)
13. Lin, Q., Riedel, F.: Optimal consumption and portfolio choice with ambiguity (2014). arXiv:1401.1639 [q-fin.PM]

14. Lin, Q., Riedel, F.: Optimal consumption and portfolio choice with ambiguous interest rates and volatility. Econ. Theory **71**, 1189–1202 (2021)
15. Maccheroni, F., Marinacci, M., Rustichini, A.: Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica **74**(6), 1447–1498 (2006)
16. Merton, R.C.: Lifetime portfolio selection under uncertainty: the continuous-time case. Rev. Econ. Stat. **51**(3), 247–257 (1969)
17. Müller, M.: Market completion and robust utility maximization. Ph.D. thesis, Humboldt-Universität zu Berlin (2005)
18. Neufeld, A., Nutz, M.: Robust utility maximization with Lévy processes. Math. Finance **28**(1), 82–105 (2018)
19. Øksendal, B., Sulem, A.: A game theoretic approach to martingale measures in incomplete markets. Surveys of Applied and Industrial Mathematics (TVP Publishers, Moscow) **15**, 18–24 (2008)
20. Øksendal, B., Sulem, A.: Robust stochastic control and equivalent martingale measures. In: Stochastic Analysis with Financial Applications, *Progress in Probability*, vol. 65, pp. 179–189. Springer Basel (2011)
21. Pflug, G., Pichler, A., Wozabal, D.: The $1/N$ investment strategy is optimal under high model ambiguity. J. Bank. Finance **36**(2), 410–417 (2012)
22. Pham, H., Wei, X., Zhou, C.: Portfolio diversification and model uncertainty: a robust dynamic mean-variance approach (2018). arXiv:1809.01464 [q-fin.PM]
23. Quenez, M.C.: Optimal portfolio in a multiple-priors model. In: Dalang, R.C., Dozzi, M., Russo, F. (eds.) Seminar on Stochastic Analysis, Random Fields and Applications IV, Progress in Probability, vol. 58, pp. 291–321. Birkhäuser, Basel (2004)
24. Sass, J., Westphal, D.: Robust utility maximization in a multivariate financial market with stochastic drift. Int. J. Theor. Appl. Finance **24**(4), 28 pages (2021)
25. Schied, A.: Optimal investments for robust utility functionals in complete market models. Math. Oper. Res. **30**(3), 750–764 (2005)
26. Schied, A.: Optimal investments for risk- and ambiguity-averse preferences: a duality approach. Finance Stoch. **11**(1), 107–129 (2007)
27. Schmeidler, D.: Subjective probability and expected utility without additivity. Econometrica **57**(3), 571–587 (1989)
28. Westphal, D.: Model uncertainty and expert opinions in continuous-time financial markets. Ph.D. thesis, Technische Universität Kaiserslautern (2019). https://nbn-resolving.org/urn:nbn:de:hbz:386-kluedo-58414
29. Zawisza, D.: A note on the worst case approach for a market with a stochastic interest rate. Applicationes Mathematicae **45**(2), 151–160 (2018)