Portfolio diversification and model uncertainty: a robust dynamic mean-variance approach

Huyen Pham, Xiaoli Wei, Chao Zhou

Portfolio diversification and model uncertainty: a robust dynamic mean-variance approach *

Huyên PHAM† Xiaoli WEI‡ Chao ZHOU§

September 4, 2018

Abstract

This paper is concerned with a multi-asset mean-variance portfolio selection problem under model uncertainty. We develop a continuous time framework for taking into account ambiguity aversion about both expected return rates and correlation matrix of the assets, and for studying the effects on portfolio diversification. We prove a separation principle for the associated robust control problem, which allows to reduce the determination of the optimal dynamic strategy to the parametric computation of the minimal risk premium function. Our results provide a justification for under-diversification, as documented in empirical studies. We explicitly quantify the degree of under-diversification in terms of correlation and Sharpe ratio ambiguity. In particular, we show that an investor with a poor confidence in the expected return estimation does not hold any risky asset, and on the other hand, trades only one risky asset when the level of ambiguity on correlation matrix is large. This extends to the continuous-time setting the results obtained by Garlappi, Uppal and Wang [13], and Liu and Zeng [24] in a one-period model.

JEL Classification: G11, C61

MSC Classification: 91G10, 91G80, 60H30

Key words: Continuous-time Markowitz problem, model uncertainty, ambiguous drift and correlation, separation principle, portfolio diversification.

*The authors benefited from the research grant for the project "Model uncertainty and stochastic control in financial risk management" under the aegis of the Alliance National University of Singapour (NUS)-Sorbonne Paris Cité (USPC). This work started when the first two authors were visiting the National University of Singapore (NUS), whose hospitality is kindly acknowledged. Part of this work was carried out while the third author was invited to the University Paris Diderot, whose hospitality is kindly thanked.

†LPSM, University Paris Diderot and CREST-ENSAE, pham at math.univ-paris-diderot.fr. The work of this author is supported by the ANR project CAESARS (ANR-15-CE05-0024), and also by FiME and the "Finance and Sustainable Development" EDF - CACIB Chair.

‡LPSM, University Paris Diderot, tyswxl at gmail.com.

§National University of Singapore, matzc at nus.edu.sg. Chao Zhou is supported by Singapore MOE AcRF Grants R-146-000-219-112 and R-146-000-255-114.

1 Introduction

In the Finance and Economics literature, there are many studies on under-diversification of portfolio, i.e., when investors hold only a small part of risky assets among a large number of available risky assets. In the extreme case the anti-diversification effect means that investors hold only a single stock (or even do not hold any risky asset) and exclude many others. Empirical studies reported in numerous papers, see [12], [7], [27], [6], [15], have shown the evidence of portfolio under-diversification in practice. For example, in [12], [7], it is observed that there exists a concentration on (bias towards) domestic assets compared to foreign assets in investors' international equity portfolios. These results are in contrast with the portfolio well-diversification suggested by the classical mean-variance portfolio theory initiated in a single period model in [25], and later in a continuous-time model in [22]. A possible explanation to under-diversification is provided in the Finance and Economics literature by model uncertainty, often also called ambiguity or Knightian uncertainty.

In the classical portfolio theory, the model and the parameters are assumed to be perfectly known. However, in reality, due to statistical estimation issues, there is always uncertainty (ambiguity) about the model or the parameters. In this case, the least favorable model can be used to compute the robust optimal portfolio, i.e., investor take portfolio decisions under the worst-case scenario that corresponds to the least favorable distribution implied by the set of ambiguous parameters. Abundant research has been conducted to tackle different types of model uncertainty. Related works include [13], [30], [19] for uncertainty solely about drift, [26] for ambiguity about volatility (in a probabilistic setup) with a family of nondominated probability measures, [33], [23], [1], [28] for combined uncertainty about both drift and volatility, and also [14] for uncertainty about the probability law generating market data. We are aware of only few results dealing with correlation ambiguity [11], [17]. It is known that the estimation of correlation between assets may be extremely inaccurate, due to the asynchronous data and lead-lag effect, especially when the number of assets is large, see [18], [20]. Besides correlation ambiguity, we are also interested in drift uncertainty as the estimation of expected rate of return is notoriously known to be difficult. In the existing literature, there are two common types of drift uncertainty set: polyhedral set in [33], [23] and ellipsoidal set in [3]. The ellipsoidal representation for the drift ambiguity allows to take into account correlation structure of the assets in the drift uncertainty modelling. Our purpose is to explore the joint effects of ambiguity about drift and correlation on portfolio selection and diversification with mean-variance criterion in continuous time. Notice that in the above cited papers, portfolio selection problems are mainly based on expected utility criterion and effect on portfolio diversification is not really studied.

When it comes to the portfolio diversification, the authors in [5], [34] considered ambiguity about the assets' returns. Their framework includes both uncertainty about the joint distribution of returns for all assets and for different levels of uncertainty for the marginal distribution of returns for any subsets of these assets. They showed that the different levels of uncertainty on different asset subclasses could result in significant under-diversification. They also applied their theoretical results to real data and found consistent results with the empirical studies in [7], [12] among others, showing that international equity portfolios are strongly biased towards domestic stocks, and in [16] and [31], where a similar lack of diversification is revealed on domestic portfolios. The model in [5], [34] offers a partial explanation for the observed under-diversification and bias towards familiar securities.

More recently, in [24], the authors considered the uncertainty about the correlation of the assets. With a static mean-variance investment, they found that the under-diversification

of the robust optimal portfolio depends on the level of correlation ambiguity. They also provided results with market data and showed that using their uncertain (ambiguous) correlation model, the investor only holds less than 20 (17 stocks in average) among 100 stocks randomly selected from about 100 stocks in S&P500. In the two risky assets case, they also found that the degree of diversification depends on the comparison between the ratio of assets' Sharpe ratios and the correlation ambiguity parameters.

A further explanation for under-diversification is that investors can reduce the uncertainty on the model or the parameters through learning. In [35], the authors built a framework to solve jointly for investment and information choices, with general preferences and information cost functions. They showed that, for some special preferences and information acquisition technologies, investors tend to learn more about the assets with which they are more familiar (typically, the domestic assets rather than the foreign ones), and become even more familiar with those assets after learning. As a consequence of this learning procedure, the investors select those assets they have learnt at the expense of others for which they have less information. Their results are consistent with the empirical studies on portfolios of international investors.

The contributions of our paper are threefold: (1) First, we develop a robust model that takes into account uncertainty about both drift and correlation of multi risky assets for $d \geq 2$, in a dynamic mean-variance portfolio setting. Robust mean-variance problem under covariance matrix uncertainty, in particular, correlation ambiguity, has been considered in [17]. However, the authors neither tackle the drift uncertainty nor study the portfolio diversification in detail, and mainly focus in the two asset case $d = 2$. One key assumption in [17] is that one can aggregate a family of processes, however in the case of drift uncertainty, this condition does not hold anymore; (2) Secondly, we derive a separation principle for the associated robust control problem formulated as a mean-field type differential game, which allows us to reduce the original min-max problem to the parametric computation of minimal risk premium. The main methodology for the separation principle is based on a weak version of the martingale optimality principle. We can not apply the classical martingale optimality principle due to the model uncertainty, in particular, due to the drift uncertainty, see Remark 3.3 for details. Indeed, the features of drift uncertainty and covariance matrix uncertainty are different, which explains why in the existing literature, such as [17] and [26], only one type of uncertainty is considered; (3) Finally, we illustrate our results in polyhedral and ellipsoidal uncertainty set and quantify explicitly the diversification effects on the optimal robust portfolio in terms of the ambiguity level. We provide notably a complete picture of the diversification for the optimal robust portfolio in the three risky assets case, which is new to the best of our knowledge. In particular, our findings consist in no trading in assets with large expected return ambiguity and trading only one risky asset with high level of ambiguity about correlation. Both drift uncertainty and correlation uncertainty can result in under-diversification. For our future studies, we may incorporate the different uncertainty levels for return as in [34] or introduce the information acquisition procedure as in [35] in our framework.

The rest of paper is organized as follows. Section 2 presents the formulation of the model uncertainty setting and the robust mean-variance problem. In Section 3, we derive the separation principle and explicit robust solution. Section 4 provides several examples arising from the separation principle, and the implications for the optimal robust portfolio strategy and the portfolio diversification.

2 Problem formulation

2.1 Model uncertainty setting

We consider a financial market with one risk-free asset, assumed to be constant equal to one, and d risky assets on a finite investment horizon $[0,T]$. Model uncertainty is formulated by using a probabilistic setup as in [28]. We define the canonical state space by $\Omega = \{\omega = (\omega(t))_{t\in[0,T]} \in C([0,T],\mathbb{R}^d) : \omega(0) = 0\}$ representing the continuous paths driving the risky assets. We equip Ω with the uniform norm and the corresponding Borel σ-field $\mathcal{F}$. We denote by $B = (B_t)_{t\in[0,T]}$ the canonical process, i.e., $B_t(\omega) = \omega(t)$, and by $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ the canonical filtration, i.e. the natural (raw) filtration generated by B.

We assume that the investor knows the marginal volatilities $\sigma_i > 0$ of each asset $i = 1,\ldots,d$, typically through a quadratic variation estimation of the assets, and we denote by $\mathfrak{S}$ the known constant diagonal matrix with i-th diagonal term equal to σ_i, $i = 1,\ldots,d$. However, there is uncertainty about the drift (expected rate of return) and the correlation between the multi-assets, which are parameters notoriously difficult to estimate in practice.

The ambiguity about drift and correlation matrix is parametrized by a nonempty convex set

$$\Theta \quad \subset \quad \mathbb{R}^d \times \mathbb{C}^d_{>+},$$

where $\mathbb{C}^d_{>+}$ is the subset of all elements $\rho = (\rho_{ij})_{1\le i<j\le d} \in [-1,1]^{d(d-1)/2}$ s.t. the symmetric matrix $C(\rho)$ with diagonal terms 1 and anti-diagonal terms ρ_{ij}:

$$C(\rho) \quad = \quad \begin{pmatrix} 1 & \rho_{12} & \ldots & \rho_{1d} \\ \rho_{12} & 1 & \ldots & . \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1d} & . & \ldots & 1 \end{pmatrix}$$

lies in $\mathbb{S}^d_{>+}$, the set of positive definite symmetric matrices in $\mathbb{R}^{d\times d}$. Notice that $\mathbb{C}^d_{>+}$ is an open convex set of $[-1,1]^{d(d-1)/2}$. The first component set of Θ represent the prior values taken by the (possibly random) drift of the assets, while the matrices $C(\rho)$, when ρ runs in the second component set of Θ, represent the prior correlation matrices of the multi-assets. The prior covariance matrices of the assets are given by

$$\Sigma(\rho) \quad = \quad \mathfrak{S}C(\rho)\mathfrak{S} \quad = \quad \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \ldots & \sigma_1\sigma_d\rho_{1d} \\ \sigma_1\sigma_2\rho_{12} & \sigma_2^2 & \ldots & . \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1\sigma_d\rho_{1d} & . & \ldots & \sigma_d^2 \end{pmatrix},$$

and we denote by $\sigma(\rho) = \Sigma^{\frac{1}{2}}(\rho)$ the square-root matrix, called volatility matrix. Let us also introduce the prior (square) risk premium

$$R(\theta) \quad = \quad b^\intercal\Sigma^{-1}(\rho)b \;=\; \|\sigma(\rho)^{-1}b\|_2^2 \quad \text{for } \theta = (b,\rho) \in \Theta. \tag{2.1}$$

Hereafter, $^\intercal$ denotes the transpose of matrix and $\|\cdot\|_2$ denotes the Euclidean norm in $\mathbb{R}^d$.

Remark 2.1 There exists different conditions for characterizing the positive definiteness of the correlation matrix $C(\rho)$. For example, Sylvester's criterion states that $C(\rho)$ is positive

definite if and only if all the leading principal minors are positive, e.g., in dimension $d = 2$, $\rho \in (-1, 1)$; in dimension $d = 3$, $\rho_{ij} \in (-1, 1)$ $1 \leq i < j \leq 3$ and $\rho_{12}^2 + \rho_{13}^2 + \rho_{23}^2 - 1 - 2\rho_{12}\rho_{13}\rho_{23} < 0$. Alternatively, one can characterize the positive definiteness of $C(\rho)$ using angular coordinates as in [29]. $\Diamond$

In the sequel, we shall focus on the two following cases for the parametrization of the ambiguity set Θ, which are relevant for practical applications:

(HΘ)

(i) *Product set*: $\Theta = \Delta \times \Gamma$, where Δ is a compact convex set of $\mathbb{R}^d$, e.g., in rectangular form $\Delta = \prod_{i=1}^{d} [\underline{b}_i, \bar{b}_i]$, for some constants $\underline{b}_i \leq \bar{b}_i$, $i = 1, \ldots, d$, and Γ is a convex set of $\mathbb{C}^d_{>+}$. In this product formulation, one considers that the uncertainty on drift is independent of the uncertainty on the correlation.

(ii) *Ellipsoidal set*: $\Theta = \{(b, \rho) \in \mathbb{R}^d \times \Gamma : \|\sigma(\rho)^{-1}(b - \hat{b})\|_2 \leq \delta\}$ for some convex set Γ of $\mathbb{C}^d_{>+}$, where $\hat{b}$ is a known vector, representing *a priori* expected rates of return, and $\delta > 0$ represents a level of ambiguity around $\hat{b}$ due to estimation error. It is known from Lemma 2.2 in [10] that Θ is a convex set. This ellipsoidal set in which varies the uncertain drift, for fixed correlation, is used in [2], and allows to take into account the correlation structure of the assets in the drift uncertainty modelling.

We denote by $\mathcal{V}_\Theta$ the set of $\mathbb{F}$-progressively measurable processes $\theta = (\theta_t) = (b_t, \rho_t)_t = (b, \rho)$ valued in Θ, and introduce the set of prior probability measures $\mathcal{P}^\Theta$:

$$\mathcal{P}^\Theta \quad = \quad \{\mathbb{P}^\theta : \theta \in \mathcal{V}_\Theta\},$$

where $\mathbb{P}^\theta$ is the probability measure on $(\Omega, \mathcal{F})$ s.t. B is a semimartingale on $(\Omega, \mathcal{F}, \mathbb{P}^\theta)$ with absolutely continuous characteristics (w.r.t. the Lebesgue measure dt) $(b, \Sigma(\rho))$. The prior probabilities $\mathbb{P}^\theta$ are in general non-equivalent, and actually mutually singular, and we say that a property holds $\mathcal{P}^\Theta$-quasi surely ($\mathcal{P}^\Theta$-q.s. in short) if it holds $\mathbb{P}^\theta$-a.s. for all $\theta \in \mathcal{V}_\Theta$.

The (positive) price process of the d risky assets is given by the dynamics

$$\begin{aligned} dS_t &= \mathrm{diag}(S_t) dB_t, \quad 0 \leq t \leq T, \ \mathcal{P}^\Theta - q.s. \\ &= \mathrm{diag}(S_t)\big(b_t dt + \sigma(\rho_t) dW_t^\theta\big), \quad \mathbb{P}^\theta - a.s., \quad \text{for } \theta \ = \ (b, \rho) \in \mathcal{V}_\Theta, \end{aligned}$$

where W^θ is a d-dimensional Brownian motion under $\mathbb{P}^\theta$. Notice that in this uncertainty modeling, we allow the unknown drift and correlation to be a *priori* random processes, valued in Θ.

2.2 Robust mean-variance problem

An admissible portfolio strategy $\alpha = (\alpha_t)_{0 \leq t \leq T}$ representing the amount invested in the d risky assets, is an $\mathbb{R}^d$-valued $\mathbb{F}$-progressively measurable process, satisfying the integrability condition

$$\sup_{\mathbb{P}^\theta \in \mathcal{P}^\Theta} \mathbb{E}_\theta \big[\int_0^T |\alpha_t^\intercal b_t| dt \ + \ \int_0^T \alpha_t^\intercal \Sigma(\rho_t) \alpha_t dt \big] \ < \ \infty, \tag{2.2}$$

and denoted by $\alpha \in \mathcal{A}$. Hereafter, $\mathbb{E}_\theta$ denotes the expectation under $\mathbb{P}^\theta$. This integrability condition (2.2) ensures that $\mathrm{diag}(S)^{-1}\alpha$ is S-integrable under any $\mathbb{P} \in \mathcal{P}^\Theta$. For a portfolio strategy $\alpha \in \mathcal{A}$, and an initial capital $x_0 \in \mathbb{R}$, the dynamics of the self-financed wealth process is driven by

$$\begin{aligned} dX_t^\alpha &= \alpha_t^\intercal \mathrm{diag}(S_t)^{-1} dS_t \;=\; \alpha_t^\intercal dB_t, \quad 0 \le t \le T,\; X_0^\alpha = x_0,\; \mathcal{P}^\Theta - q.s. \\ &= \alpha_t^\intercal \big(b_t dt + \sigma(\rho_t) dW_t^\theta\big), \;\; 0 \le t \le T,\; X_0^\alpha = x_0 \in \mathbb{R},\;\; \mathbb{P}^\theta - a.s. \end{aligned} \tag{2.3}$$

for all $\theta = (b, \rho) \in \mathcal{V}_\Theta$.

Given a risk aversion parameter $\lambda > 0$, the worst-case mean-variance functional under ambiguous drift and correlation is

$$J_{wc}(\alpha) \;=\; \inf_{\mathbb{P}^\theta \in \mathcal{P}^\Theta} \Big(\mathbb{E}_\theta[X_T^\alpha] - \lambda \mathrm{Var}_\theta(X_T^\alpha) \Big) \;<\; \infty,\; \alpha \in \mathcal{A},$$

where $\mathrm{Var}_\theta(.)$ denotes the variance under $\mathbb{P}^\theta$, and the robust mean-variance portfolio selection is formulated as

$$\begin{cases} V_0 &:= \sup_{\alpha \in \mathcal{A}} J_{wc}(\alpha) \;=\; \sup_{\alpha \in \mathcal{A}} \inf_{\theta \in \mathcal{V}_\Theta} J(\alpha, \theta) \\ J(\alpha, \theta) &:= \mathbb{E}_\theta[X_T^\alpha] - \lambda \mathrm{Var}_\theta(X_T^\alpha), \quad \alpha \in \mathcal{A}, \theta \in \mathcal{V}_\Theta. \end{cases} \tag{2.4}$$

Notice that problem (2.4) is a non standard stochastic differential game due to the presence of the variance term in the criterion, which prevents the use of classical control method by dynamic programming or maximum principle. We end this section by recalling the solution to the mean-variance problem when there is no ambiguity on the model parameters, and which will serve later as benchmark for comparison when studying the uncertainty case.

Remark 2.2 (Case of no uncertainty model) When $\Theta = \{\theta^o = (b^o, \rho^o)\}$ is a singleton, we are reduced to the Black-Scholes model with drift b^o, covariance matrix $\Sigma^o = \Sigma(\rho^o)$, volatility $\sigma = \sigma(\rho^o)$, and risk premium $R^o = R(\theta^o)$. In this case, it is known, see e.g. [22], that the optimal mean-variance strategy is given by

$$\alpha_t^* \;=\; \big[x_0 + \frac{e^{R^o T}}{2\lambda} - X_t^*\big](\Sigma^o)^{-1} b^o \;=:\; \Lambda^o(X_t^*)(\Sigma^o)^{-1} b^o, \quad 0 \le t \le T,$$

where X^* is wealth process associated to α^*, while the optimal performance value is

$$V_0 \;=\; x_0 + \frac{1}{4\lambda}\big[e^{R^o T} - 1\big].$$

The vector $(\Sigma^o)^{-1} b^o$, which depends only on the model parameters of the stock price, determines the allocation in the multi-assets. The above expression of α^* shows that, once we know the exact values of the rate of return and covariance matrix, one diversifies her portfolio among all the assets according to the components of the vector $(\Sigma^o)^{-1} b^o$, and this is weighted by the scalar term $\Lambda^o(X_t^*)$, which depends on the risk aversion of the investor via the parameter λ, on the current wealth but also on the initial capital x_0 (which is sometimes refereed to as the pre-committment of the mean-variance criterion). Notice that $\Lambda^o(X_t^*)$ is positive. Indeed, observe that

$$\begin{aligned} d\Lambda^o(X_t^*) \;=\; -dX_t^* &= -(\alpha_t^*)^\intercal (b^o dt + \sigma^o dW_t^o) \\ &= -\Lambda^o(X_t^*)(R^o dt + (\sigma^o)^{-1} b^o . dW_t^o), \quad 0 \le t \le T, \end{aligned}$$

with $\Lambda^o(X_0^*) = \frac{1}{2\lambda}e^{R^oT} > 0$, which shows clearly that $\Lambda^o(X_t^*) > 0$, $0 \leq t \leq T$, and decreases with λ.

Let us discuss in particular the allocation in the two-asset case. Notice that the vector $(\Sigma^o)^{-1}b^o$ of allocation is then given by

$$(\Sigma^o)^{-1}b^o \;=\; \frac{1}{1-|\rho^o|^2}\begin{pmatrix} \frac{\beta_1^o - \rho^o\beta_2^o}{\sigma_1^o} \\ \frac{\beta_2^o - \rho^o\beta_1^o}{\sigma_2^o} \end{pmatrix} =: \begin{pmatrix} \kappa_1^o \\ \kappa_2^o \end{pmatrix},$$

where $\beta_i^o = b_i^o/\sigma_i^o$ is the sharpe ratio of the i-th asset, $i = 1, 2$. To fix the idea, assume that $\beta_1^o > \beta_2^o > 0$. We then see that $\kappa_1^o > 0$, while $\kappa_2^o \geq 0$ if and only if $\frac{\beta_2^o}{\beta_1^o} \geq \rho^o$. The interpretation is the following: the ratio $\frac{\beta_2^o}{\beta_1^o} \in (0, 1)$ measures the "proximity" in terms of Sharpe ratio between the two assets, and has to be compared with the correlation ρ^o between these assets in order to determine whether it is optimal to invest according to a directional trading, i.e., $\kappa_1^o\kappa_2^o > 0$ (thus here long in both assets) or according to a spread trading, i.e., $\kappa_1^o\kappa_2^o < 0$ (long in the first asset and short in the second one) or according to under-diversification, i.e. $\kappa_1^o\kappa_2^o = 0$ (only long in the first asset). Notice that under-diversification only occurs when $\rho^o = \frac{\beta_2^o}{\beta_1^o}$, a condition "rarely" satisfied in practice. For example, when both assets have close Sharpe ratio, and their correlation is not too high, then one optimally invests in both assets with a directional trading. In contrast, when one asset has a much larger Sharpe ratio than the other one, or when the correlation between the assets is high, then one optimally invests in both assets with a spread trading. $\Diamond$

In the sequel, we study the quantitative impact of the uncertainty model and ambiguity on the drift and correlation, on the optimal robust mean-variance strategy, in particular regarding the portfolio diversification.

3 Separation principle and robust solution

The main result of this section is to state a separation principle for solving the robust dynamic mean-variance problem.

Theorem 3.1 (Separation Principle) *Let us consider a parametric set Θ for model uncertainty as in* **(HΘ)**. *Suppose that there exists a (constant) pair $\theta^* = (b^*, \rho^*) \in \Theta$ solution to $\arg\min_{\theta\in\Theta} R(\theta)$. Then the robust mean-variance problem* (2.4) *admits an optimal portfolio strategy given by*

$$\alpha_t^* \;=\; \Lambda_{\theta^*}(X_t^*)\Sigma(\rho^*)^{-1}b^*, \quad 0 \leq t \leq T, \;\; \mathcal{P}^\Theta - q.s., \tag{3.1}$$

where X^* is the state process associated to α_t^*, and $\Lambda_{\theta^*}(X_t^*) > 0$ with

$$\Lambda_{\theta^*}(x) \;:=\; x_0 + \frac{e^{R(\theta^*)T}}{2\lambda} - x, \;\; x \in \mathbb{R} \tag{3.2}$$

Moreover, the corresponding initial value function is

$$V_0 \;=\; x_0 + \frac{1}{4\lambda}\big[e^{R(\theta^*)T} - 1\big].$$

Interpretation. Theorem 3.1 means that the robust mean-variance problem (2.4) can be solved in two steps according to a separation principle: (i) First, we search for the infimum of the risk premium function $\theta \in \Theta \mapsto R(\theta)$ as defined in (2.1), which depends only on the inputs of the uncertainty model. Existence and explicit determination of an element $\theta^* = (b^*, \rho^*) \in \Theta$ attaining this infimum will be discussed and illustrated all along the paper through several examples. (ii) The solution to (2.4) is then given by the solution to the mean-variance problem in the Black-Scholes model with drift b^* and correlation ρ^*, see Remark 2.2, and the worst-case scenario of the robust dynamic mean-variance problem is simply given by the constant parameter $\theta^* = (b^*, \rho^*)$. Some interesting features show up, especially regarding portfolio diversification, as detailed in the next section. $\Diamond$

The rest of this section is devoted to the proof of Theorem 3.1, and the methodology is based on the following weak version of the martingale optimality principle.

Lemma 3.1 (Weak optimality principle) *Let $\{V_t^{\alpha,\theta}, t \in [0,T], \alpha \in \mathcal{A}, \theta \in \mathcal{V}_\Theta\}$ be a family of real-valued processes in the form*

$$V_t^{\alpha,\theta} : = v_t(X_t^\alpha, \mathbb{E}_\theta[X_t^\alpha]),$$

for some measurable functions v_t on $\mathbb{R} \times \mathbb{R}$, $t \in [0,T]$, such that :

(i) $v_T(x, \bar{x}) = x - \lambda(x - \bar{x})^2$, for all $x, \bar{x} \in \mathbb{R}$,

(ii) the function $t \in [0,T] \mapsto \mathbb{E}_{\theta^*}[V_t^{\alpha,\theta^*}]$ is nonincreasing for all $\alpha \in \mathcal{A}$ and some $\theta^* \in \mathcal{V}_\Theta$,

(iii) $\mathbb{E}_\theta[V_T^{\alpha^*,\theta} - V_0^{\alpha^*,\theta}] \geq 0$, for some $\alpha^* \in \mathcal{A}$ and all $\theta \in \mathcal{V}_\Theta$.

Then, α^* is an optimal portfolio strategy for the robust mean-variance problem (2.4) *with a worst-case scenario θ^*, and*

$$\begin{aligned} V_0 &= J_{wc}(\alpha^*) = \sup_{\alpha \in \mathcal{A}} \inf_{\theta \in \mathcal{V}_\Theta} J(\alpha, \theta) = \inf_{\theta \in \mathcal{V}_\Theta} \sup_{\alpha \in \mathcal{A}} J(\alpha, \theta) = v_0(x_0, x_0) \qquad (3.3) \\ &= J(\alpha^*, \theta^*). \end{aligned}$$

Proof. First, observe that $V_0^{\alpha,\theta} = v_0(x_0, x_0)$ is a constant that does not depend on α, θ and from condition (i) that $\mathbb{E}_\theta[V_T^{\alpha,\theta}] = J(\alpha, \theta)$ for all $\alpha \in \mathcal{A}$, $\theta \in \mathcal{V}_\Theta$. Then, from condition (ii), we see that

$$v_0(x_0, x_0) \; = \; \mathbb{E}_{\theta^*}[V_0^{\alpha,\theta^*}] \; \geq \; \mathbb{E}_{\theta^*}[V_T^{\alpha,\theta^*}] \; = \; J(\alpha, \theta^*),$$

for all $\alpha \in \mathcal{A}$, and thus: $v_0(x_0, x_0) \geq \sup_{\alpha \in \mathcal{A}} J(\alpha, \theta^*) \geq \inf_{\theta \in \mathcal{V}_\Theta} \sup_{\alpha \in \mathcal{A}} J(\alpha, \theta)$. Similarly, from condition (iii), we have: $v_0(x_0, x_0) \leq J(\alpha^*, \theta)$ for all $\theta \in \mathcal{V}_\Theta$, and thus: $v_0(x_0, x_0) \leq \inf_{\theta \in \mathcal{V}_\Theta} J(\alpha^*, \theta) = J_{wc}(\alpha^*) \leq \sup_{\alpha \in \mathcal{A}} \inf_{\theta \in \mathcal{V}_\Theta} J(\alpha, \theta)$. Recalling that we always have $\sup_{\alpha \in \mathcal{A}} \inf_{\theta \in \mathcal{V}_\Theta} J(\alpha, \theta) \leq \inf_{\theta \in \mathcal{V}_\Theta} \sup_{\alpha \in \mathcal{A}} J(\alpha, \theta)$, we obtained the required equality in (3.3). Then, finally, from (ii) with α^* and (iii) with θ^*, we obtain that $v_0(x_0, x_0) = J(\alpha^*, \theta^*)$. $\square$

Remark 3.1 The usual martingale optimality principle for stochastic differential games as in robust portfolio selection problem, and with classical expected utility criterion for some nondecreasing and concave utility function U on $\mathbb{R}$, e.g. $U(x) = -e^{-\eta x}$, $\eta > 0$:

$$\sup_{\alpha \in \mathcal{A}} \inf_{\theta \in \mathcal{V}_\Theta} \mathbb{E}_\theta[U(X_T^\alpha)],$$

would consist in finding a family of processes $V_t^{\alpha,\theta}$ in the form $v_t(X_t^\alpha)$ for some measurable functions v_t on $\mathbb{R}$ s.t. (i) $v_T(x) = U(x)$, (ii') the process $(V_t^{\alpha,\theta^*})_t$ is a supermartingale under $\mathbb{P}_{\theta^*}$ for all α, and some θ^*, and (iii') the process $(V_t^{\alpha^*,\theta})_t$ is a submartingale under $\mathbb{P}_\theta$ for some α^* and all θ. Due to the nonlinear dependence on the law of the state wealth process via the variance term in the mean-variance criterion, making the problem a *priori* time inconsistent, we have to adopt a weaker version of the optimality principle: first, the functions v_t depend not only on the state process X_t^α but also on its mean $\mathbb{E}_\theta[X_t^\alpha]$. Second, we replace condition (ii') by the weaker condition (ii) on the mean in Lemma 3.1, and third, condition (iii') is substituted by the even weaker condition (iii) than (iii") $t \mapsto \mathbb{E}_\theta[V_t^{\alpha^*,\theta}]$ is nondecreasing for some α^* and all θ. This asymmetry of condition between (ii) and (iii) is explained in more detail in Remark 3.3. $\Diamond$

We shall also use the following saddle-point property on the infimum of the prior risk premium function.

Lemma 3.2 (Saddle point property) *Given Θ as in* **(HΘ)**, *and assuming that there exists $\theta^* = (b^*, \rho^*) \in \arg\min_{\theta\in\Theta} R(\theta)$, let us define the function H on Θ by*

$$H(\theta) \;:=\; b^\intercal \Sigma(\rho^*)^{-1}\Sigma(\rho)\Sigma(\rho^*)^{-1}b^*, \quad \text{for } \theta = (b,\rho) \in \Theta. \tag{3.4}$$

Then, we have for all $b \in \Theta_b$, $\rho \in \Gamma$:

$$H(b^*, \rho) \;\le\; H(\theta^*) \;=\; R(\theta^*) \;\le\; H(b, \rho^*). \tag{3.5}$$

where Θ_b is the projection of Θ onto b-plane, in **(HΘ)***(i), $\Theta_b = \Delta$; and in* **(HΘ)***(ii), $\Theta_b = \{b \in \mathbb{R}^d : b \in \Delta_\rho, \rho \in \Gamma\}$ with $\Delta_\rho = \{b \in \mathbb{R}^d : \|\sigma(\rho)^{-1}(b - \hat{b})\|_2 \le \delta\}$.*

Proof. See Section A.1 in Appendix. □

Proof of Theorem 3.1. We aim to construct a family of processes $\{V_t^{\alpha,\theta}, t \in [0,T], \alpha \in \mathcal{A}, \theta \in \mathcal{V}_\Theta\}$ as in Lemma 3.1, and given the linear-quadratic structure of our optimization problem, we look for measurable functions v_t in the form:

$$v_t(x, \bar{x}) \;=\; K_t(x - \bar{x})^2 + Y_t x + \chi_t, \quad t \in [0,T], (x,\bar{x}) \in \mathbb{R}^2, \tag{3.6}$$

for some deterministic processes $(K_t, Y_t, \chi_t)_t$ to be determined. Condition (i) in Lemma 3.1 fixes the terminal condition

$$K_T \;=\; -\lambda, \;\; Y_T \;=\; 1, \;\; \chi_T \;=\; 0. \tag{3.7}$$

We now consider $\theta^* \in \Theta$ as in Theorem 3.1, hence defining in particular a (constant) process $\theta^* \in \mathcal{V}_\Theta$, and α^* given by (3.1). Let us first check that $\alpha^* \in \mathcal{A}$. The corresponding wealth process X^* satisfies under any $\mathbb{P}^\theta$, $\theta = (b,\rho) \in \mathcal{V}_\Theta$, a linear stochastic differential equation with bounded random coefficients (notice that b and $\sigma(\rho)$ are bounded process), and thus by standard estimates: $\mathbb{E}_\theta\big[\sup_{0\le t\le T} |X_t^*|^2\big] \le C(1 + |x_0|^2)$ for some constant C independent of $\theta \in \mathcal{V}_\Theta$. It follows immediately that α^* satisfies the integrability condition in (2.2), i.e., $\alpha^* \in \mathcal{A}$.

The main issue is now to show that such a pair (α^*, θ^*) satisfies conditions (ii)-(iii) of Lemma 3.1.

• *Step 1: condition (ii) of Lemma 3.1.*

For any $\alpha \in \mathcal{A}$, with associated wealth process $X = X^\alpha$, let us compute the derivative of the deterministic function $t \mapsto \mathbb{E}_{\theta^*}[V_t^{\alpha,\theta^*}] = \mathbb{E}_{\theta^*}[v_t(X_t, \mathbb{E}_{\theta^*}[X_t])]$ with v_t as in (3.6). From the dynamics of $X = X_t^\alpha$ in (2.3) under $\mathbb{P}^{\theta^*}$ and by applying Itô's formula, we obtain

$$\begin{aligned} \frac{d\mathbb{E}_{\theta^*}[X_t]}{dt} &= \mathbb{E}_{\theta^*}[\alpha_t^\intercal b^*] \\ \frac{d\mathrm{Var}_{\theta^*}(X_t)}{dt} &= 2\mathrm{Cov}_{\theta^*}(X_t, \alpha_t^\intercal b^*) + \mathbb{E}_{\theta^*}[\alpha_t^\intercal \Sigma(\rho^*)\alpha_t]. \end{aligned}$$

From the quadratic form of v_t in (3.6), with (K, Y, χ) differentiable in time, we then have

$$\begin{aligned} \frac{d\mathbb{E}_{\theta^*}[V_t^{\alpha,\theta^*}]}{dt} &= \frac{d\mathbb{E}_{\theta^*}[v_t(X_t, \mathbb{E}_{\theta^*}[X_t])]}{dt} \\ &= \dot{K}_t \mathrm{Var}_{\theta^*}(X_t) + K_t \frac{d\mathrm{Var}_{\theta^*}(X_t)}{dt} + \dot{Y}_t \mathbb{E}_{\theta^*}[X_t] + Y_t \frac{d\mathbb{E}_{\theta^*}[X_t]}{dt} + \dot{\chi}_t \\ &= \dot{K}_t \mathrm{Var}_{\theta^*}(X_t) + \dot{Y}_t \mathbb{E}_{\theta^*}[X_t] + \dot{\chi}_t + \mathbb{E}_{\theta^*}[G_t(\alpha)] \end{aligned} \tag{3.8}$$

where

$$G_t(\alpha) = \alpha_t^\intercal Q_t \alpha_t + \alpha_t^\intercal \big[2U_t(X_t - \mathbb{E}_{\theta^*}[X_t]) + O_t\big],$$

with the deterministic coefficients

$$Q_t = K_t \Sigma(\rho^*), \quad U_t = K_t b^*, \quad O_t = Y_t b^*.$$

By square completion, we rewrite $G_t(\alpha)$ as

$$G_t(\alpha) = \big(\alpha_t - \hat{a}_t(X_t, \mathbb{E}_{\theta^*}[X_t])\big)^\intercal Q_t \big(\alpha_t - \hat{a}_t(X_t, \mathbb{E}_{\theta^*}[X_t])\big) - \zeta_t,$$

where for $t \in [0,T]$, $x, \bar{x} \in \mathbb{R}^2$,

$$\hat{a}_t(x, \bar{x}) := -Q_t^{-1} U_t (x - \bar{x}) - \frac{1}{2} Q_t^{-1} O_t$$

and

$$\zeta_t = U_t^\intercal Q_t^{-1} U_t \mathrm{Var}_{\theta^*}(X_t) + \frac{1}{4} O_t^\intercal Q_t^{-1} O_t = K_t R(\theta^*) \mathrm{Var}_{\theta^*}(X_t) + \frac{Y_t^2}{4K_t} R(\theta^*).$$

The expression in (3.8) is then rewritten as

$$\begin{aligned} \frac{d\mathbb{E}_{\theta^*}[V_t^{\alpha,\theta^*}]}{dt} &= (\dot{K}_t - K_t R(\theta^*))\mathrm{Var}_{\theta^*}(X_t) + \dot{Y}_t \mathbb{E}_{\theta^*}[X_t] + \dot{\chi}_t - \frac{Y_t^2}{4K_t} R(\theta^*) \\ &\quad + K_t \mathbb{E}_{\theta^*}\big[\big(\alpha_t - \hat{a}_t(X_t, \mathbb{E}_{\theta^*}[X_t])\big)^\intercal \Sigma(\rho^*) \big(\alpha_t - \hat{a}_t(X_t, \mathbb{E}_{\theta^*}[X_t])\big)\big]. \end{aligned} \tag{3.9}$$

Therefore, whenever

$$\begin{cases} \dot{K}_t - K_t R(\theta^*) = 0, \\ \dot{Y}_t = 0, \\ \dot{\chi}_t - \frac{Y_t^2}{4K_t} R(\theta^*) = 0, \end{cases} \tag{3.10}$$

holds for all $t \in [0,T]$, which yields, together with the terminal condition (3.7), the explicit forms:

$$K_t = -\lambda e^{R(\theta^*)(t-T)} < 0, \quad Y_t = 1, \quad \chi_t = \frac{1}{4\lambda}\big[e^{R(\theta^*)(T-t)} - 1\big], \tag{3.11}$$

we have

$$\frac{d\mathbb{E}_{\theta^*}[V_t^{\alpha,\theta^*}]}{dt} \;=\; K_t\mathbb{E}_{\theta^*}\big[\big(\alpha_t-\hat{a}_t(X_t,\mathbb{E}_{\theta^*}[X_t])\big)^\intercal\Sigma(\rho^*)\big(\alpha_t-\hat{a}_t(X_t,\mathbb{E}_{\theta^*}[X_t])\big)\big],$$

which is nonpositive for all $\alpha\in\mathcal{A}$, i.e., the process V_t^{α,θ^*} satisfies the condition (ii) of Lemma 3.1. Moreover, notice that in this case,

$$V_0^{\alpha,\theta^*} \;=\; v_0(x_0,x_0) \;=\; x_0+\frac{1}{4\lambda}\big[e^{R(\theta^*)T}-1\big], \tag{3.12}$$

and

$$\hat{a}_t(x,\bar{x}) \;=\; -\Sigma(\rho^*)^{-1}b^*\big(x-\bar{x}-\frac{1}{2\lambda}e^{R(\theta^*)(T-t)}\big). \tag{3.13}$$

Notice that in this step, we have not yet used the property that θ^* attains the infimum of the *prior* risk premium function. This will be used in the next step.

• *Step 2: Condition (iii) of Lemma 3.1.*
Let us now prove that $V_0^{\alpha^*,\theta}\le\mathbb{E}_\theta[V_T^{\alpha^*,\theta}]$, for all $\theta\in\mathcal{V}_\Theta$. A sufficient condition is the nondecreasing monotonicity of the function $t\mapsto\mathbb{E}_\theta[V_t^{\alpha^*,\theta}]$, by proving that $\frac{d\mathbb{E}_\theta[V_t^{\alpha^*,\theta}]}{dt}$ is nonnegative, for all $\theta\in\mathcal{V}_\Theta$. However, while this nondecreasing property is valid when there is no uncertainty on the drift, this does not hold true in the general uncertainty case as shown in Remark 3.3. We then proceed by computing directly the difference: $\mathbb{E}_\theta[V_T^{\alpha^*,\theta}]-V_0^{\alpha^*,\theta}$. Notice from (3.1), (2.3), that the dynamics of $\Lambda_{\theta^*}(X^*)$, with $\Lambda_{\theta^*}(x)$ defined in (3.2), under $\mathbb{P}^\theta$, $\theta\in\mathcal{V}_\Theta$, is given by

$$d\Lambda_{\theta^*}(X_t^*) \;=\; -\Lambda_{\theta^*}(X_t^*)(b^*)^\intercal\Sigma(\rho^*)^{-1}\big[b_tdt+\sigma(\rho_t)dW_t^\theta\big],$$

with $\Lambda_{\theta^*}(x_0)=\frac{e^{R(\theta^*)T}}{2\lambda}$. By setting $N_t^*:=\frac{2\lambda}{e^{R(\theta^*)T}}\Lambda_{\theta^*}(X_t^*)$, we deduce that

$$\begin{aligned}
N_t^* &= \exp\Big(-\int_0^t\big(b_s^\intercal\Sigma(\rho^*)^{-1}b^*+\frac{1}{2}(b^*)^\intercal\Sigma(\rho^*)^{-1}\Sigma(\rho)\Sigma(\rho^*)^{-1}b^*\big)ds\\
&\qquad -\int_0^t(b^*)^\intercal\Sigma(\rho^*)^{-1}\sigma(\rho_s)dW_s^\theta\Big),\quad 0\le t\le T,\ \mathbb{P}^\theta-a.s.\\
X_t^* &= x_0+\frac{e^{R(\theta^*)T}}{2\lambda}(1-N_t^*),\quad 0\le t\le T,\ \mathcal{P}^\Theta-q.s.
\end{aligned}$$

and thus

$$\mathbb{E}_\theta[X_t^*] \;=\; x_0+\frac{e^{R(\theta^*)T}}{2\lambda}(1-\mathbb{E}_\theta[N_t^*]),\qquad \mathrm{Var}_\theta(X_t^*) \;=\; \frac{e^{2R(\theta^*)T}}{4\lambda^2}\mathrm{Var}_\theta(N_t^*). \tag{3.14}$$

By using the quadratic form (3.6) of v_t, together with the terminal condition (3.7), (3.12), and (3.14), we then obtain for all $\theta\in\mathcal{V}_\Theta$:

$$\begin{aligned}
\mathbb{E}_\theta[V_T^{\alpha^*,\theta}]-V_0^{\alpha^*,\theta} &= \mathbb{E}_\theta\big[v_T(X_T^*,\mathbb{E}_\theta[X_T^*])\big]-v_0(x_0,x_0)\\
&= -\lambda\mathrm{Var}_\theta(X_T^*)+\mathbb{E}_\theta[X_T^*]-x_0-\frac{1}{4\lambda}(e^{R(\theta^*)T}-1)\\
&= -\frac{e^{2R(\theta^*)T}}{4\lambda}\mathrm{Var}_\theta(N_T^*)+\frac{e^{R(\theta^*)T}}{2\lambda}(1-\mathbb{E}_\theta[N_T^*])-\frac{1}{4\lambda}(e^{R(\theta^*)T}-1)\\
&= \frac{e^{R(\theta^*)T}}{4\lambda}\Big(1-e^{R(\theta^*)T}\mathbb{E}_\theta[|N_T^*|^2]\Big)+\frac{1}{4\lambda}\Big(e^{R(\theta^*)T}\mathbb{E}_\theta[N_T^*]-1\Big)^2\\
&\ge \frac{e^{R(\theta^*)T}}{4\lambda}\Big(1-e^{R(\theta^*)T}\mathbb{E}_\theta[|N_T^*|^2]\Big) \;=:\; \frac{e^{R(\theta^*)T}}{4\lambda}\Delta_T^*(\theta).
\end{aligned} \tag{3.15}$$

Noting that N^* is rewritten in terms of H introduced in Lemma 3.2 as

$$N_t^* \;=\; \exp\Big(-\int_0^t \big(H(b_s,\rho^*) + \frac{1}{2}H(b^*,\rho_s)\big)ds - \int_0^t (b^*)^\intercal \Sigma(\rho^*)^{-1}\sigma(\rho_s)dW_s^\theta\Big), 0\le t\le T,\;\; \mathbb{P}^\theta - a.s.$$

and observing that $|(b^*)^\intercal\Sigma(\rho^*)^{-1}\sigma(\rho_s)|^2 = H(b^*,\rho_s)$, we see that

$$|N_t^*|^2 \;=\; \exp\Big(-\int_0^t \big(2H(b_s,\rho^*) - H(b^*,\rho_s)\big)ds\Big)M_t^*,$$

where

$$M_t^* \;=\; \exp\Big(-2\int_0^t |(b^*)^\intercal\Sigma(\rho^*)^{-1}\sigma(\rho_s)|^2 ds - 2\int_0^t (b^*)^\intercal\Sigma(\rho^*)^{-1}\sigma(\rho_s)dW_s^\theta\Big),\; 0\le t\le T,\;\; \mathbb{P}^\theta - a.s.$$

is an exponential Doléans-Dade local martingale under any $\mathbb{P}^\theta$, $\theta\in\mathcal{V}_\Theta$. Actually, the Novikov criterion

$$\begin{array}{rcl}\mathbb{E}_\theta\Big[\exp\Big(\frac{1}{2}\displaystyle\int_0^T |2(b^*)^\intercal\Sigma(\rho^*)^{-1}\sigma(\rho_t)|^2dt\Big)\Big] &=& \mathbb{E}_\theta\Big[\exp\Big(2\displaystyle\int_0^T H(b^*,\rho_t)dt\Big)\Big]\\ &\le& \exp\big(2R(\theta^*)T\big) \;<\; \infty,\end{array}$$

is satisfied by (3.5), and then $(M_t^*)_{0\le t\le T}$ is a martingale under any $\mathbb{P}^\theta$, $\theta\in\mathcal{V}_\Theta$. Consequently, we have

$$\begin{array}{rcl}\Delta_T^*(\theta) &=& 1-\mathbb{E}_\theta\Big[\exp\Big(\displaystyle\int_0^t \big(R(\theta^*) - 2H(b_s,\rho^*) + H(b^*,\rho_s)\big)ds\Big)M_T^*\Big]\\ &\ge& 1-\mathbb{E}_\theta[M_T^*] \;=\; 1-M_0^* \;=\; 0,\end{array}$$

where we used (3.5) in the above inequality. From (3.15), this proves condition (iii) of Lemma (3.1), and finally concludes the proof of Theorem 3.1. □

Remark 3.2 The optimal strategy α^* given in (3.1) can be expressed in feedback form as

$$\alpha_t^* \;=\; \hat{a}_t(X_t^*, \mathbb{E}_{\theta^*}[X_t^*]),\quad 0\le t\le T,\; \mathcal{P}^\Theta - q.s. \tag{3.16}$$

where $\hat{a}_t$ is defined in (3.13). Indeed, denoting by $\hat{\alpha}\in\mathcal{A}$ the process defined by $\hat{\alpha}_t = \hat{a}_t(\hat{X}_t, \mathbb{E}_{\theta^*}[\hat{X}_t])$, $0\le t\le T$, $\mathcal{P}^\Theta - q.s$, where $\hat{X}$ is the wealth process associated to $\hat{\alpha}$, we see from (2.3) that $\hat{X}$ satisfies the dynamics under $\mathbb{P}^{\theta^*}$:

$$d\hat{X}_t \;=\; -\Big[\hat{X}_t - \mathbb{E}_{\theta^*}[\hat{X}_t] - \frac{1}{2\lambda}e^{R(\theta^*)(T-t)}\Big](b^*)^\intercal\Sigma(\rho^*)^{-1}\big[b^*dt + \sigma(\rho^*)dW_t^{\theta^*}\big].$$

By taking expectation under $\mathbb{P}^{\theta^*}$, we get: $d\mathbb{E}_{\theta^*}[\hat{X}_t] = \frac{1}{2\lambda}e^{R(\theta^*)(T-t)}R(\theta^*)dt$, and thus

$$\begin{array}{rcl}\mathbb{E}_{\theta^*}[\hat{X}_t] &=& x_0 + \dfrac{e^{R(\theta^*)T}}{2\lambda}\big[1-e^{-R(\theta^*)t}\big],\\ \hat{\alpha}_t &=& \Lambda_{\theta^*}(\hat{X}_t)\Sigma(\rho^*)^{-1}b^*,\quad 0\le t\le T,\; \mathcal{P}^\Theta - q.s.\end{array}$$

This implies that $\hat{X}$ and X^* satisfy the same linear SDE under $\mathbb{P}^\theta$, for any $\theta\in\mathcal{V}_\Theta$, and so $\hat{X}_t = X_t^*$, $0\le t\le T$, $\mathcal{P}^\Theta$-q.s. This proves that $\alpha^* = \hat{\alpha}$, equal to (3.16). ◊

Remark 3.3 By similar derivation as in (3.9), and using (3.10), (3.16), we have that for all $\theta = (\theta_t)_t = (b_t, \rho_t)_t \in \mathcal{V}_\Theta$, $t \in [0,T]$,

$$\begin{aligned}\frac{d\mathbb{E}_\theta[V_t^{\alpha^*,\theta}]}{dt} &= K_t\big(R(\theta^*) - R(\theta_t)\big)\mathrm{Var}_\theta(X_t^*) + \frac{1}{4K_t}\big(R(\theta^*) - R(\theta_t)\big) \qquad (3.17)\\ &\quad + K_t\mathbb{E}_\theta\Big[\big(\hat{a}_t(X_t^*, \mathbb{E}_{\theta^*}[X_t^*]) - \hat{a}_t(X_t^*, \mathbb{E}_\theta[X_t^*])\big)^\intercal \Sigma(\rho_t)\\ &\qquad \big(\hat{a}_t(X_t^*, \mathbb{E}_{\theta^*}[X_t^*]) - \hat{a}_t(X_t^*, \mathbb{E}_\theta[X_t^*])\big)\Big]\\ &\geq K_t\mathbb{E}_\theta\Big[\big(\hat{a}_t(X_t^*, \mathbb{E}_{\theta^*}[X_t^*]) - \hat{a}_t(X_t^*, \mathbb{E}_\theta[X_t^*])\big)^\intercal \Sigma(\rho_t)\\ &\qquad \big(\hat{a}_t(X_t^*, \mathbb{E}_{\theta^*}[X_t^*]) - \hat{a}_t(X_t^*, \mathbb{E}_\theta[X_t^*])\big)\Big] \qquad (3.18)\end{aligned}$$

by definition of $\theta^* \in \arg\min_{\theta\in\Theta} R(\theta)$, and as $K_t < 0$. In the case when there is no uncertainty on the drift, i.e., for any $\theta = (b,\rho) \in \mathcal{V}_\Theta$, b is a constant equal to b^o, the dynamics of X^* under any $\mathbb{P}^\theta$, $\theta \in \mathcal{V}_\Theta$, is given by

$$dX_t^* = \big[x_0 + \frac{e^{R(\theta^*)T}}{2\lambda} - X_t^*\big](b^o)^\intercal\Sigma(\rho^*)^{-1}\big[b^o dt + \sigma(\rho_t)dW_t^\theta\big],$$

from which, we deduce by taking expectation under $\mathbb{P}^\theta$:

$$\mathbb{E}_\theta[X_t^*] = x_0 + \frac{e^{R(\theta^*)T}}{2\lambda}\big[1 - e^{-R(\theta^*)t}\big].$$

This means that the expectation under $\mathbb{P}^\theta$ of the optimal wealth process X^* does not depend on $\theta \in \mathcal{V}_\Theta$, and the r.h.s. of (3.18) is then equal to zero. Therefore, the function $t \mapsto \mathbb{E}_\theta[V_t^{\alpha^*,\theta}]$ is nondecreasing for all $\theta \in \mathcal{V}_\Theta$, which implies in particular condition (iii) of Lemma 3.1.

However, in the case of drift uncertainty, we cannot conclude as above, and actually this nondecreasing property does not always hold true. Indeed, consider for example the case where there is only drift uncertainty in a single asset model $d = 1$, with $\Theta = \{\theta \in [\underline{b}, \bar{b}]\}$, $0 \leq \underline{b} < \bar{b}$, and known variance Σ^o normalized to one. Notice that $R(\theta) = \theta^2$, and $\theta^* = \arg\min_{\theta\in\Theta} R(\theta) = \underline{b}$. For any $\theta \in \Theta$, we can compute explicitly from (3.14) the expectation and variance of X^* under $\mathbb{P}^\theta$:

$$\begin{aligned}\mathbb{E}_\theta[X_t^*] &= \frac{1}{2\lambda}e^{R(\theta^*)T}\big[1 - e^{-\theta\theta^* t}\big],\\ \mathrm{Var}_\theta(X_t^*) &= \frac{1}{4\lambda^2}e^{2R(\theta^*)T}\big[e^{(R(\theta^*)-2\theta\theta^*)t} - e^{-2\theta\theta^* t}\big].\end{aligned}$$

Plugging into (3.17), and using also the expression of K, $\hat{a}$ in (3.11), (3.13), we have for all $\theta \in \Theta$, $t \in [0,T]$, after some straightforward rearrangement:

$$\begin{aligned}\frac{d\mathbb{E}_\theta[V_t^{\alpha^*,\theta}]}{dt} &= \frac{1}{2\lambda}e^{R(\theta^*)T}\Big[ce^{-2ct} - e^{-R(\theta^*)t}(1 - e^{-ct})\Big(\frac{R(\theta^*)}{2} - \big(\frac{R(\theta^*)}{2} + c\big)e^{-ct}\Big)\Big]\\ &=: f(t,c),\end{aligned}$$

where we set $c = (\theta - \theta^*)\theta^* \geq 0$. Now, we easily see that for all $t \in [0,T]$, $f(t,c)$ converges to $-\frac{R(\theta^*)}{4\lambda}e^{R(\theta^*)(T-t)} < 0$, as c goes to infinity. Then, by continuity of f with respect to c, we deduce that for θ large enough (hence for c large enough), $\frac{d\mathbb{E}_\theta[V_t^{\alpha^*,\theta}]}{dt}$ is negative, which means that the function $t \mapsto \mathbb{E}_\theta[V_t^{\alpha^*,\theta}]$ is not nondecreasing for all $\theta \in \Theta$. Actually, we have proved in Theorem 3.1 the weaker condition (iii) of Lemma 3.1 that $V_0^{\alpha^*,\theta} \leq \mathbb{E}_\theta[V_T^{\alpha^*,\theta}]$, for all $\theta \in \mathcal{V}_\Theta$. $\diamondsuit$

4 Applications and examples

We provide in this section several examples for the determination of the minimal risk premium arising from the separation principle in Theorem 3.1, and the implications for the optimal robust portfolio strategy and the portfolio diversification. We shall focus in this section on the case of ellipsoidal uncertainty set Θ as in **(HΘ)**(ii), i.e., in the form:

$$\Theta \;=\; \{(b,\rho) \in \mathbb{R}^d \times \Gamma : \|\sigma(\rho)^{-1}(b-\hat{b})\|_2 \leq \delta\}, \tag{4.1}$$

and refer to the thesis manuscript [36] for the case of rectangular uncertainty set.

4.1 Minimal risk premium and worst-case scenario

We compute the infimum of the prior risk premium function $\theta \in \Theta \mapsto R(\theta)$ as defined in (2.1), and (when it exists) the element $\theta^* \in \Theta$ which achieves this minimum, i.e., the worst-case scenario for uncertain parameters.

Given Θ as in (4.1), we denote by $\hat{\beta}_i := \frac{\hat{b}_i}{\sigma_i}$ the Sharpe ratio of the i-th asset associated with *a priori* expected rate of return $\hat{b}_i$, and marginal volatility $\sigma_i > 0$, $i = 1, \ldots, d$. In what follows, we always assume w.l.o.g that $|\hat{\beta}_i|$ is in descending order: $|\hat{\beta}_1| \geq |\hat{\beta}_2| \geq \ldots |\hat{\beta}_d|$ and $\hat{\beta}_1 \neq 0$, i.e., $\hat{b} \neq 0$ (in the trivial case $\hat{b} = 0$, the optimal portfolio strategy is to never trade, i.e. $\alpha_t^* \equiv 0$), and define the Sharpe ratio "proximity" between asset i and asset j by

$$\hat{\varrho}_{ij} \;=\; \frac{\hat{\beta}_j}{\hat{\beta}_i} \in [-1,1], \quad 1 \leq i < j \leq d, \tag{4.2}$$

with the convention that $\hat{\varrho}_{ij} = 0$ when $\hat{\beta}_i = 0$.

Lemma 4.1 *Let Θ be an ellipsoidal set as in (4.1), and assume that there exists $\rho^* \in \arg\min_{\rho\in\Gamma} R(\hat{b},\rho) = \arg\min_{\rho\in\Gamma} \left\|\sigma(\rho)^{-1}\hat{b}\right\|_2$. Then $\theta^* = (b^*,\rho^*)$ with*

$$\begin{aligned} b^* \;&=\; \Big(1 - \frac{\delta}{\|\sigma(\rho^*)^{-1}\hat{b}\|_2}\Big) 1_{\{\|\sigma(\rho^*)^{-1}\hat{b}\|_2 > \delta\}}\, \hat{b}, \\ \textit{and} \quad R(\theta^*) \;&=\; \big(\|\sigma(\rho^*)^{-1}\hat{b}\|_2 - \delta\big)^2 1_{\{\|\sigma(\rho^*)^{-1}\hat{b}\|_2 > \delta\}}. \end{aligned}$$

Proof. See proof of Lemma A.3 in Appendix. □

Remark 4.1 The existence of ρ^* is guaranteed when Γ is a compact set of $\mathbb{C}^d_{>+}$ by continuity of the function $\rho \mapsto \left\|\sigma(\rho)^{-1}\hat{b}\right\|_2$. We also show in Proposition 4.1 its existence when $\Gamma = \mathbb{C}^d_{>+}$, and under the condition that there exists a highest *a priori* Sharpe ratio. ◊

In the particular case when there is full ambiguity about the correlation, i.e., $\Gamma = \mathbb{C}^d_{>+}$, and there is an asset with *a priori* highest (absolute value) Sharpe ratio, one can compute explicitly the worst-case scenario $\rho^* \in \Theta$ for correlation.

Proposition 4.1 (Full ambiguity correlation) *Let Θ as in (4.1), with $\Gamma = \mathbb{C}^d_{>+}$, and assume that $|\hat{\beta}_1| > |\hat{\beta}_2| = \max_{i\neq 1} |\hat{\beta}_i|$. Then, we have $\arg\min_{\Theta} R(b,\rho) \neq \varnothing$, attained at b^*, $\rho^* = (\rho^*_{ij})_{1\leq i<j\leq d} \in \mathbb{C}^d_{>+}$ with*

$$\rho^*_{1j} \;=\; \hat{\varrho}_{1j}, \qquad b^* \;=\; (1 - \frac{\delta}{|\hat{\beta}_1|})1_{\{|\hat{\beta}_1|>\delta\}}\hat{b}, \quad 1 < j \leq d,$$

and

$$R(\theta^*) \;=\; (|\hat{\beta}_1| - \delta)^2 1_{\{|\hat{\beta}_1|>\delta\}}.$$

In particular, when $|\hat{\beta}_1| > |\hat{\beta}_2| > \ldots > |\hat{\beta}_d|$, then $\rho^*_{ij} = \hat{\varrho}_{ij}$, $1 \leq i < j \leq d$, and the associated correlation matrix $C(\rho^*)$ is positive definite.

Proof. See A.2 in Appendix. □

Remark 4.2 The proof of proposition 4.1 states that $R(\hat{b},\rho)$ has minimum value over $\mathbb{C}^d_{>+}$ if and only if $|\hat{\beta}_1| > \max_{j\neq 1} |\hat{\beta}_j|$. Indeed, if there are more than one greatest (absolute) Sharpe ratio, $R(\hat{b},\rho)$ does not have minimum value. For example, if $d = 2$, $\hat{\beta}_1 = \hat{\beta}_2$, then $\mathbb{C}^2_{>+} = (-1,1)$, and $R(\hat{b},\rho) = \frac{2|\hat{\beta}_1|^2}{1+\rho}$. We thus obtain $\inf_{\mathbb{C}^2_{>+}} R(\hat{b},\rho) = |\hat{\beta}_1|^2$, and see that this infimum cannot be attained over $\mathbb{C}^2_{>+}$. ◊

We now consider a model for two-risky assets, i.e., with $d = 2$, mixing partial ambiguity about correlation and drift uncertainty. In this case, the following result provides the explicit expression of the worst-case scenario achieving the minimal risk premium.

Proposition 4.2 (Ambiguous drift and correlation in the two-assets case) *Let $\Theta = \{(b,\rho) \in \mathbb{R}^2 \times [\underline{\rho},\bar{\rho}] : \|\sigma(\rho)^{-1}(b - \hat{b})\|_2 \leq \delta\}$, with $-1 < \underline{\rho} \leq \bar{\rho} < 1$, and assume w.l.o.g that $|\hat{\beta}_1| \geq |\hat{\beta}_2|$, $\hat{\beta}_1 \neq 0$. Recall that $\hat{\rho}_{12} := \frac{\hat{\beta}_2}{\hat{\beta}_1}$. Then,*

1. *If $\hat{\varrho}_{12} \in [\underline{\rho},\bar{\rho}]$, then $\rho^* = \hat{\varrho}_{12}$, and $b^* = \hat{b}(1 - \frac{\delta}{|\hat{\beta}_1|})1_{\{|\hat{\beta}_1|>\delta\}}$.*
2. *If $\bar{\rho} < \hat{\varrho}_{12}$, then $\rho^* = \bar{\rho}$, and $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma(\bar{\rho})^{-1}\hat{b}\|_2})1_{\{\|\sigma(\bar{\rho})^{-1}\hat{b}\|_2>\delta\}}$.*
3. *If $\underline{\rho} > \hat{\varrho}_{12}$, then $\rho^* = \underline{\rho}$, and $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma(\underline{\rho})^{-1}\hat{b}\|_2})1_{\{\|\sigma(\underline{\rho})^{-1}\hat{b}\|_2>\delta\}}$.*

Proof. See A.3 in Appendix. □

Remark 4.3 The computation of the worst-case correlation ρ^* is determined according to three cases depending on the relation between the Sharpe ratio proximity $\hat{\rho}_{12}$ and the two correlation bounds $\underline{\rho}$ and $\bar{\rho}$.

In the first case when the range of correlation ambiguity is large enough so that $\hat{\rho}_{12} \in [\underline{\rho},\bar{\rho}]$, or in other words, no stock is clearly dominating the other one in terms of Sharpe ratio, then the worst-case correlation is attained at the point $\hat{\rho}_{12}$ inside the interval $[\underline{\rho},\bar{\rho}]$.

In the second case, when $\bar{\rho} < \hat{\varrho}_{12}$, meaning that both assets have close Sharpe ratios with a correlation upper bound not too large, then the worst-case correlation is attained at the prior highest correlation $\bar{\rho}$.

In the third case, when $\underline{\rho} > \hat{\varrho}_{12}$, meaning that Sharpe ratios of the two assets are rather distinctive with respect to the correlation lower bound, then the worst-case correlation is given by the prior lowest correlation $\underline{\rho}$. $\Diamond$

We finally consider a model for three-risky assets ($d = 3$) mixing partial ambiguity about correlation and drift uncertainty, hence with Θ in the form $\Theta = \{(b, \rho) \in \mathbb{R}^3 \times \Gamma : \|\sigma(\rho)^{-1}(b - \hat{b})\|_2 \leq \delta\}$ with $\Gamma = [\underline{\rho}_{12}, \bar{\rho}_{12}] \times [\underline{\rho}_{13}, \bar{\rho}_{13}] \times [\underline{\rho}_{23}, \bar{\rho}_{23}]$, a subset of $\mathbb{C}^3_{>+}$.

Recall that

$$\hat{\varrho}_{12} := \frac{\hat{\beta}_2}{\hat{\beta}_1} \in [-1, 1], \quad \hat{\varrho}_{13} := \frac{\hat{\beta}_3}{\hat{\beta}_1} \in [-1, 1], \quad \hat{\varrho}_{23} := \frac{\hat{\beta}_3}{\hat{\beta}_2} \in [-1, 1]. \tag{4.3}$$

We introduce the so-called variance risk ratio $\hat{\kappa}(\rho)$,

$$\Sigma(\rho)^{-1}\hat{b} =: \hat{\kappa}(\rho) = (\hat{\kappa}^1(\rho), \hat{\kappa}^2(\rho), \hat{\kappa}^3(\rho))^\intercal, \tag{4.4}$$

which represents (up to a scalar term) the vector of allocation in the assets when the drift is $\hat{b}$ and the correlation is ρ.

We denote by $\hat{b}_{-i}$ the *a priori* expected rate of return $\hat{b}$ with the i-th component $\hat{b}_i$ removed, and by $\Sigma_{-i}(\rho)$ the covariance matrix $\Sigma(\rho)$ with i-th row and i-th column removed, and $\sigma_{-i}(\rho) = \Sigma_{-i}(\rho)^{\frac{1}{2}}$. Notice that $\Sigma_{-1}(\rho)$ depends only on ρ_{23}. We will write $\Sigma_{-1}(\rho)$ as $\Sigma_{-1}(\rho_{23})$, similarly, $\Sigma_{-2}(\rho_{13})$, $\Sigma_{-3}(\rho_{12})$.

In this case, the following result provides the explicit expression of the worst-case scenario achieving the minimal risk premium.

Proposition 4.3 (Ambiguous drift and correlation in the three-asset case) *Let* $\Theta = \{(b, \rho) \in \mathbb{R}^3 \times \Gamma : \|\sigma(\rho)^{-1}(b - \hat{b})\|_2 \leq \delta\}$ *with* $\Gamma = [\underline{\rho}_{12}, \bar{\rho}_{12}] \times [\underline{\rho}_{13}, \bar{\rho}_{13}] \times [\underline{\rho}_{23}, \bar{\rho}_{23}] \subset \mathbb{C}^3_{>+}$, *and assume w.l.o.g that* $|\hat{\beta}_1| \geq |\hat{\beta}_2| \geq |\hat{\beta}_3|$, $\hat{\beta}_1 \neq 0$. *Then, we have the following possible exclusive cases:*

1. *(High-level correlation ambiguity for the second and third assets)*

   If $\hat{\varrho}_{12} \in [\underline{\rho}_{12}, \bar{\rho}_{12}]$, $\hat{\varrho}_{13} \in [\underline{\rho}_{13}, \bar{\rho}_{13}]$, *then* $\rho^* = (\hat{\varrho}_{12}, \hat{\varrho}_{13}, \rho^*_{23})$ *for any* $\rho^*_{23} \in [\underline{\rho}_{23}, \bar{\rho}_{23}]$, *and* $b^* = \hat{b}(1 - \frac{\delta}{|\hat{\beta}_1|})1_{\{|\hat{\beta}_1| > \delta\}}$.

2. *(High-level correlation ambiguity for the third asset)*

   (i) If $\bar{\rho}_{12} < \hat{\varrho}_{12}$, $\hat{\kappa}^3(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23})\hat{\kappa}^3(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23}) \leq 0$, *then* $\rho^* = (\bar{\rho}_{12}, \rho^*_{13}, \rho^*_{23})$ *satisfying* $\hat{\kappa}^3(\bar{\rho}_{12}, \rho^*_{13}, \rho^*_{23}) = 0$, *and* $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma_{-3}(\bar{\rho}_{12})^{-1}\hat{b}_{-3}\|_2})1_{\{\|\sigma_{-3}(\bar{\rho}_{12})^{-1}\hat{b}_{-3}\|_2 > \delta\}}$.

   (ii) If $\underline{\rho}_{12} > \hat{\varrho}_{12}$, $\hat{\kappa}^3(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23})\hat{\kappa}^3(\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23}) \leq 0$, *then* $\rho^* = (\underline{\rho}_{12}, \rho^*_{13}, \rho^*_{23})$ *satisfying* $\hat{\kappa}^3(\underline{\rho}_{12}, \rho^*_{13}, \rho^*_{23}) = 0$, *and* $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma_{-3}(\underline{\rho}_{12})^{-1}\hat{b}_{-3}\|_2})1_{\{\|\sigma_{-3}(\underline{\rho}_{12})^{-1}\hat{b}_{-3}\|_2 > \delta\}}$.

3. *(High-level correlation ambiguity for the second asset)*

   (i) If $\bar{\rho}_{13} < \hat{\varrho}_{13}$, $\hat{\kappa}^2(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23})\hat{\kappa}^2(\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23}) \leq 0$, *then* $\rho^* = (\rho^*_{12}, \bar{\rho}_{13}, \rho^*_{23})$ *satisfying* $\hat{\kappa}^2(\rho^*_{12}, \bar{\rho}_{13}, \rho^*_{23}) = 0$, *and* $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma_{-2}(\bar{\rho}_{13})^{-1}\hat{b}_{-2}\|_2})1_{\{\|\sigma_{-2}(\bar{\rho}_{13})^{-1}\hat{b}_{-2}\|_2 > \delta\}}$.

   (ii) If $\underline{\rho}_{13} > \hat{\varrho}_{13}$, $\hat{\kappa}^2(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23})\hat{\kappa}^2(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23}) \leq 0$, *then* $\rho^* = (\rho^*_{12}, \underline{\rho}_{13}, \rho^*_{23})$ *satisfying* $\hat{\kappa}^2(\rho^*_{12}, \underline{\rho}_{13}, \rho^*_{23}) = 0$, *and* $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma_{-2}(\underline{\rho}_{13})^{-1}\hat{b}_{-2}\|_2})1_{\{\|\sigma_{-2}(\underline{\rho}_{13})^{-1}\hat{b}_{-2}\|_2 > \delta\}}$.

4. *(High-level correlation ambiguity for the first asset)*

(i) If $\bar{\rho}_{23} < \hat{\varrho}_{23}$, $\hat{\kappa}^1(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23})\hat{\kappa}^1(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23}) \leq 0$, *then* $\rho^* = (\rho^*_{12}, \rho^*_{13}, \bar{\rho}_{23})$ *satisfying* $\hat{\kappa}^1(\rho^*_{12}, \rho^*_{13}, \bar{\rho}_{23}) = 0$, *and* $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma_{-1}(\bar{\rho}_{23})^{-1}\hat{b}_{-1}\|_2})1_{\{\|\sigma_{-1}(\bar{\rho}_{23})^{-1}\hat{b}_{-1}\|_2 > \delta\}}$.

(ii) If $\underline{\rho}_{23} > \hat{\varrho}_{23}$, $\hat{\kappa}^1(\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23})\hat{\kappa}^1(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23}) \leq 0$, *then* $\rho^* = (\rho^*_{12}, \rho^*_{13}, \underline{\rho}_{23})$ *satisfying* $\hat{\kappa}^1(\rho^*_{12}, \rho^*_{13}, \underline{\rho}_{23}) = 0$, *and* $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma_{-1}(\underline{\rho}_{23})^{-1}\hat{b}_{-1}\|_2})1_{\{\|\sigma_{-1}(\underline{\rho}_{23})^{-1}\hat{b}_{-1}\|_2 > \delta\}}$.

5. *(Small ambiguity about correlation)*

(i) If $\hat{\kappa}^1\hat{\kappa}^2(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23}) > 0$, $\hat{\kappa}^1\hat{\kappa}^3(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23}) > 0$, *then* $\rho^* = (\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23})$, *and* $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23})^{-1}\hat{b}\|_2})1_{\{\|\sigma(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23})^{-1}\hat{b}\|_2 > \delta\}}$.

(ii) If $\hat{\kappa}^1\hat{\kappa}^2(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23}) < 0$, $\hat{\kappa}^1\hat{\kappa}^3(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23}) < 0$, *then* $\rho^* = (\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23})$, *and* $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23})^{-1}\hat{b}\|_2})1_{\{\|\sigma(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23})^{-1}\hat{b}\|_2 > \delta\}}$.

(iii) If $\hat{\kappa}^1\hat{\kappa}^2(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23}) > 0$, $\hat{\kappa}^1\hat{\kappa}^3(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23}) < 0$, *then* $\rho^* = (\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23})$, *and* $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23})^{-1}\hat{b}\|_2})1_{\{\|\sigma(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23})^{-1}\hat{b}\|_2 > \delta\}}$.

(iv) If $\hat{\kappa}^1\hat{\kappa}^2(\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23}) < 0$, $\hat{\kappa}^1\hat{\kappa}^3(\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23}) > 0$, *then* $\rho^* = (\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23})$, *and* $b^* = \hat{b}(1 - \frac{\delta}{\|\sigma(\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23})^{-1}\hat{b}\|_2})1_{\{\|\sigma(\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23})^{-1}\hat{b}\|_2 > \delta\}}$.

Proof. See A.4 in Appendix. □

Remark 4.4 The different cases in the above Proposition depend on the relation between the Sharpe ratio proximities and the correlation intervals bounds, and can be roughly divided into 5 cases with subcases with the following interpretation:

In case **1** where the range of correlation ambiguity for the second and third asset is large enough, in the sense that the intervals $[\underline{\rho}_{12}, \bar{\rho}_{12}]$ and $[\underline{\rho}_{13}, \bar{\rho}_{13}]$ contain respectively $\hat{\rho}_{12}$ and $\hat{\rho}_{13}$, then the worst-case correlation is attained at the Sharpe ratio proximity value $\rho^* = (\hat{\varrho}_{12}, \hat{\varrho}_{13}, \rho^*_{23})$.

Let us now discuss case **2**, and more specifically (i). This corresponds to the situation where the assets 1 and 2 have close Sharpe ratios with a correlation upper bound between these assets not too large, while the correlation ambiguity for the third asset is high, which is quantified by the fact that the function $(\rho_{13}, \rho_{23}) \mapsto \hat{\kappa}(\bar{\rho}_{12}, \rho_{13}, \rho_{23})$ evaluated at the prior lower bounds $(\underline{\rho}_{13}, \underline{\rho}_{23})$ and the prior upper bounds $(\bar{\rho}_{13}, \bar{\rho}_{23})$ have opposite signs. In this case, the worst-case correlation is achieved at the prior highest correlation $\bar{\rho}_{12}$ for ρ_{12}, and at the point $(\rho^*_{13}, \rho^*_{23})$ cancelling the term $\hat{\kappa}(\bar{\rho}_{12}, \rho^*_{13}, \rho^*_{23})$. Similar interpretations hold for cases **3** and **4**.

Let us finally discuss case **5**, which involves explicitly the signs of $\hat{\kappa}^1\hat{\kappa}^2$ and $\hat{\kappa}^1\hat{\kappa}^3$ at the prior correlation bounds. Assuming that these functions $\hat{\kappa}^1\hat{\kappa}^2$ and $\hat{\kappa}^1\hat{\kappa}^3$ do not vanish at some point $\rho \in [\underline{\rho}_{12}, \bar{\rho}_{12}] \times [\underline{\rho}_{13}, \bar{\rho}_{13}] \times [\underline{\rho}_{23}, \bar{\rho}_{23}]$, then by continuity, and provided that the range of these correlation bounds are small enough, we see that one should fall into one of the 4 subcases **5.**(i), (ii), (iii), (iv), and for which the worst-case correlation is obtained on the prior upper or lower correlation bounds. ◊

4.2 Optimal robust strategy and portfolio diversification

We first provide the general explicit expression of the robust optimal strategy in the case of ellipsoidal ambiguity set. This follows directly from Theorem 3.1 and Lemma 4.1.

Proposition 4.4 *Let Θ be an ellipsoidal set as in* (4.1)*, and assume that there exists $\rho^* \in \arg\min_{\rho\in\Gamma} \big\|\sigma(\rho)^{-1}\hat{b}\big\|_2$. Then, an optimal portfolio strategy for* (2.4) *is given by*

$$\alpha_t^* = \big[x_0 + \frac{1}{2\lambda} e^{(\|\sigma(\rho^*)^{-1}\hat{b}\|_2 - \delta)^2 T} - X_t^*\big]\Big(1 - \frac{\delta}{\|\sigma(\rho^*)^{-1}\hat{b}\|_2}\Big) 1_{\{\|\sigma(\rho^*)^{-1}\hat{b}\|_2 > \delta\}} \Sigma(\rho^*)^{-1}\hat{b}. \tag{4.5}$$

Remark 4.5 We have seen in the previous section that ρ^* exists when Γ is compact (in particular when it is a singleton, i.e., there is no ambiguity on correlation) or when $\Gamma = \mathbb{C}^d_{>+}$, i.e., there is full ambiguity on correlation. From (4.5), we observe notably that whenever $\delta \geq \|\sigma(\rho^*)^{-1}\hat{b}\|_2$, then $\alpha^* \equiv 0$. In other words, when the level of ambiguity about the expected rate of return is high (or when the investor is poorly confident about her estimation $\hat{b}$ on the expected rate of return), then she does not make risky investment at all. $\Diamond$

4.2.1 Full ambiguity correlation and anti-diversification

In this paragraph, we consider the case of full ambiguity on correlation, i.e., $\Gamma = \mathbb{C}^d_{>+}$, and investigate the impact on optimal robust portfolio strategy.

Theorem 4.1 (Full ambiguity correlation) *Let Θ be an ellipsoidal set as in* (4.1)*, with $\Gamma = \mathbb{C}^d_{>+}$, and assume that $|\hat{\beta}_1| > |\hat{\beta}_2| = \max_{j\neq 1} |\hat{\beta}_j|$, $\beta_1 \neq 0$. Then an optimal portfolio strategy for the robust mean-variance problem* (2.4) *is explicitly given by*

$$\alpha_t^* = \big[x_0 + \frac{1}{2\lambda} e^{(|\hat{\beta}_1| - \delta)^2 T} - X_t^*\big]\Big(1 - \frac{\delta}{|\hat{\beta}_1|}\Big) 1_{|\hat{\beta}_1|>\delta} \big(\frac{\hat{b}_1}{\sigma_1^2}, 0, \ldots, 0\big)^\intercal, \quad 0 \leq t \leq T,\ \mathcal{P}^\Theta - q.s.$$

Proof. From the formula (4.5) of the optimal portfolio strategy in Proposition 4.4, we only have to compute the vector $\hat{\kappa}(\rho^*) = \Sigma(\rho^*)^{-1}\hat{b}$, and $R(\hat{b}, \rho^*)$, which have been already given in (A.28), (A.29) in Section A.2. $\square$

Remark 4.6 (Financial interpretation: anti-diversification) If the investor is poorly confident on the drift estimate, i.e., whenever δ is large enough, then she does not make risky investments at all, i.e. $\alpha_t^* \equiv 0$. When the investor has good knowledge on drift estimates but is poorly confident on correlation estimates, she only invests in one asset, namely the one with the highest *a priori* Sharpe ratio. This anti-diversification result under full ambiguity about correlation has been also observed in [24] for a single-period mean-variance problem without drift uncertainty, and is extended here in a continuous time framework. $\Diamond$

4.2.2 Partial diversification

• **Two-asset model:** $d = 2$

We provide a complete picture of the optimal robust portfolio strategy in a two-asset model with ambiguous drift and correlation.

Theorem 4.2 (Ambiguous drift and correlation in the two-assets case) *Let $\Theta = \{(b, \rho) \in \mathbb{R}^2 \times [\underline{\rho}, \bar{\rho}] : \|\sigma(\rho)^{-1}(b - \hat{b})\|_2 \leq \delta\}$, with $-1 < \underline{\rho} \leq \bar{\rho} < 1$, and assume w.l.o.g that $|\hat{\beta}_1| \geq |\hat{\beta}_2|$, $\hat{\beta}_1 \neq 0$. We have the following possible cases:*

(1) *If* $\hat{\varrho}_{12} \in [\underline{\rho}, \bar{\rho}]$*, then an optimal portfolio strategy is explicitly given by*

$$\alpha_t^* = \big[x_0 + \frac{1}{2\lambda} e^{(|\hat{\beta}_1|-\delta)^2 T} - X_t^*\big]\big(1 - \frac{\delta}{|\hat{\beta}_1|}\big) 1_{\{|\hat{\beta}_1|>\delta\}} \begin{pmatrix} \frac{\hat{b}_1}{\sigma_1^2} \\ 0 \end{pmatrix}, \ \ 0 \leq t \leq T, \ \mathcal{P}^\Theta - q.s.$$

(2) *If* $\bar{\rho} < \hat{\varrho}_{12}$*, then an optimal portfolio strategy is explicitly given by*

$$\alpha_t^* = \big[x_0 + \frac{1}{2\lambda} e^{(\|\sigma(\bar{\rho})^{-1}\hat{b}\|_2-\delta)^2 T} - X_t^*\big]\Big(1 - \frac{\delta}{\|\sigma(\bar{\rho})^{-1}\hat{b}\|_2}\Big) 1_{\{\|\sigma(\bar{\rho})^{-1}\hat{b}\|_2>\delta\}} \Sigma(\bar{\rho})^{-1}\hat{b},$$

and if $\|\sigma(\bar{\rho})^{-1}\hat{b}\|_2 > \delta$*, then* $\alpha_t^{1,*}\alpha_t^{2,*} > 0$*.*

(3) *If* $\underline{\rho} > \hat{\varrho}_{12}$*, then an optimal portfolio strategy is explicitly given by*

$$\alpha_t^* = \big[x_0 + \frac{1}{2\lambda} e^{(\|\sigma(\underline{\rho})^{-1}\hat{b}\|_2-\delta)^2 T} - X_t^*\big]\Big(1 - \frac{\delta}{\|\sigma(\underline{\rho})^{-1}\hat{b}\|_2}\Big) 1_{\{\|\sigma(\underline{\rho})^{-1}\hat{b}\|_2>\delta\}} \Sigma(\underline{\rho})^{-1}\hat{b},$$

and if $\|\sigma(\underline{\rho})^{-1}\hat{b}\|_2 > \delta$*, then* $\alpha_t^{1,*}\alpha_t^{2,*} < 0$*.*

Proof. In light of formula (4.5) of the optimal portfolio strategy in Proposition 4.4, we only need to compute $\|\sigma(\rho^*)^{-1}\hat{b}\|_2$, i.e., $R(\hat{b}, \rho^*)$, vector $\hat{\kappa}(\rho^*) = \Sigma(\rho^*)^{-1}\hat{b}$, explicitly given in the proof of Proposition 4.2 (see A.3 in Appendix) when computing ρ^*, which leads to the three cases of Theorem 4.2. □

Remark 4.7 When there is no ambiguity on the drift, which corresponds to $\delta = 0$, we retrieve the results obtained in [17] for the correlation ambiguity between two assets (see their Theorem 4.2). Our Theorem includes in addition the case when there is uncertainty on the expected rate of return. ♢

Remark 4.8 (Financial interpretation) In the first case when the range of correlation ambiguity is large enough so that $\hat{\rho}_{12} \in (\underline{\rho}, \bar{\rho})$, and thus no stock is clearly dominating the other one in terms of Sharpe ratio, it is optimal to invest only in one asset, namely the one with the highest estimated Sharpe ratio.

In the second case when $\bar{\rho} < \hat{\rho}_{12}$, this means that no stock is "dominating" the other one in terms of Sharpe ratio, and it is optimal to invest in both assets with a directional trading, that is buying or selling simultaneously, and the worst-case correlation refers to the highest prior correlation $\bar{\rho}$ (recall Remark 4.3) where the diversification effect is minimal.

Finally, when $\underline{\rho} > \hat{\rho}_{12}$, this means that one asset is clearly dominating the other one, and it is optimal to invest in both assets with a spread trading, that is buying one and selling another, and the worst-case correlation corresponds to the lowest prior correlation $\underline{\rho}$ where the profit from the spread trading is minimal.

This diversification result has been also observed in [24] for a single-period mean-variance problem with only correlation uncertainty, and is extended here in a continuous time framework with both drift and correlation uncertainty. ♢

• **Three-asset model:** $d = 3$

We finally provide an explicit description of the optimal robust strategy in a three-asset model under drift uncertainty and ambiguous correlation.

Theorem 4.3 *Let* $\Theta = \{(b,\rho) \in \mathbb{R}^3 \times \Gamma : \|\sigma(\rho)^{-1}(b-\hat{b})\|_2 \leq \delta\}$, *with* $\Gamma = [\underline{\rho}_{12}, \bar{\rho}_{12}] \times [\underline{\rho}_{13}, \bar{\rho}_{13}] \times [\underline{\rho}_{23}, \bar{\rho}_{23}] \subset \mathbb{C}^3_{>+}$, *and assume w.l.o.g that* $|\hat{\beta}_1| \geq |\hat{\beta}_2| \geq |\hat{\beta}_3|$ *and* $\hat{\beta}_1 \neq 0$. *Then, we have the following possible exclusive cases:*

1. *(Anti-diversification) If* $\hat{\varrho}_{12} \in [\underline{\rho}_{12}, \bar{\rho}_{12}]$, *and* $\hat{\varrho}_{13} \in [\underline{\rho}_{13}, \bar{\rho}_{13}]$, *then an optimal portfolio strategy is explicitly given by*

$$\alpha^*_t = \big[x_0 + \frac{1}{2\lambda} e^{(|\hat{\beta}_1|-\delta)^2 T} - X^*_t\big]\Big(1 - \frac{\delta}{|\hat{\beta}_1|}\Big) 1_{\{|\hat{\beta}_1|>\delta\}} \begin{pmatrix} \frac{\hat{b}_1}{\sigma_1^2} \\ 0 \\ 0 \end{pmatrix}, \ 0 \leq t \leq T, \ \mathcal{P}^\Theta - q.s..$$

2. *(Under-diversification: no investment in the third asset)*

(i) If $\bar{\rho}_{12} < \hat{\varrho}_{12}$, *and* $\hat{\kappa}^3(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23})\hat{\kappa}^3(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23}) \leq 0$, *then an optimal portfolio strategy is*

$$\begin{aligned} \begin{pmatrix} \alpha^{1,*}_t \\ \alpha^{2,*}_t \end{pmatrix} &= \big[x_0 + \frac{1}{2\lambda} e^{(\|\sigma_{-3}(\bar{\rho}_{12})^{-1}\hat{b}_{-3}\|_2 - \delta)^2 T} - X^*_t\big]\Big(1 - \frac{\delta}{\|\sigma_{-3}(\bar{\rho}_{12})^{-1}\hat{b}_{-3}\|_2}\Big) \\ &\qquad 1_{\{\|\sigma_{-3}(\bar{\rho}_{12})^{-1}\hat{b}_{-3}\|_2>\delta\}} \Sigma_{-3}(\bar{\rho}_{12})^{-1}\hat{b}_{-3} \\ \alpha^{3,*}_t &\equiv 0, \end{aligned}$$

and if $\|\sigma_{-3}(\bar{\rho}_{12})^{-1}\hat{b}_{-3}\|_2 > \delta$, *then* $\alpha^{1,*}_t \alpha^{2,*}_t > 0$.

(ii) If $\underline{\rho}_{12} > \hat{\varrho}_{12}$, *and* $\hat{\kappa}^3(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23})\hat{\kappa}^3(\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23}) \leq 0$, *then an optimal portfolio strategy is*

$$\begin{aligned} \begin{pmatrix} \alpha^{1,*}_t \\ \alpha^{2,*}_t \end{pmatrix} &= \big[x_0 + \frac{1}{2\lambda} e^{(\|\sigma_{-3}(\underline{\rho}_{12})^{-1}\hat{b}_{-3}\|_2) - \delta)^2 T} - X^*_t\big]\Big(1 - \frac{\delta}{\|\sigma_{-3}(\underline{\rho}_{12})^{-1}\hat{b}_{-3}\|_2}\Big) \\ &\qquad 1_{\{\|\sigma_{-3}(\underline{\rho}_{12})^{-1}\hat{b}_{-3}\|_2>\delta\}} \Sigma_{-3}(\bar{\rho}_{12})^{-1}\hat{b}_{-3} \\ \alpha^{3,*}_t &\equiv 0, \end{aligned}$$

and if $\|\sigma_{-3}(\underline{\rho}_{12})^{-1}\hat{b}_{-3}\|_2 > \delta$, *then* $\alpha^{1,*}_t \alpha^{2,*}_t < 0$.

3. *(Under-diversification: no investment in the second asset)*

(i) If $\bar{\rho}_{13} < \hat{\varrho}_{13}$, *and* $\hat{\kappa}^2(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23})\hat{\kappa}^2(\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23}) \leq 0$, *then an optimal portfolio strategy is*

$$\begin{aligned} \begin{pmatrix} \alpha^{1,*}_t \\ \alpha^{3,*}_t \end{pmatrix} &= \big[x_0 + \frac{1}{2\lambda} e^{(\|\sigma_{-2}(\bar{\rho}_{13})^{-1}\hat{b}_{-2}\|_2 - \delta)^2 T} - X^*_t\big]\Big(1 - \frac{\delta}{\|\sigma_{-2}(\bar{\rho}_{13})^{-1}\hat{b}_{-2}\|_2}\Big) \\ &\qquad 1_{\{\|\sigma_{-2}(\bar{\rho}_{13})^{-1}\hat{b}_{-2}\|_2>\delta\}} \Sigma_{-2}(\bar{\rho}_{13})^{-1}\hat{b}_{-2} \\ \alpha^{2,*}_t &\equiv 0, \end{aligned}$$

and if $\|\sigma_{-2}(\bar{\rho}_{13})^{-1}\hat{b}_{-2}\|_2 > \delta$, *then* $\alpha^{1,*}_t \alpha^{3,*}_t > 0$.

(ii) If $\underline{\rho}_{13} > \hat{\varrho}_{13}$, *and* $\hat{\kappa}^2(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23})\hat{\kappa}^2(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23}) \leq 0$, *then an optimal portfolio strategy is given by*

$$\begin{aligned} \begin{pmatrix} \alpha^{1,*}_t \\ \alpha^{3,*}_t \end{pmatrix} &= \big[x_0 + \frac{1}{2\lambda} e^{(\|\sigma_{-2}(\underline{\rho}_{13})^{-1}\hat{b}_{-2}\|_2 - \delta)^2 T} - X^*_t\big]\Big(1 - \frac{\delta}{\|\sigma_{-2}(\underline{\rho}_{13})^{-1}\hat{b}_{-2}\|_2}\Big) \\ &\qquad 1_{\{\|\sigma_{-2}(\underline{\rho}_{13})^{-1}\hat{b}_{-2}\|_2>\delta\}} \Sigma_{-2}(\underline{\rho}_{13})^{-1}\hat{b}_{-2} \\ \alpha^{2,*}_t &\equiv 0, \end{aligned}$$

and if $\|\sigma_{-2}(\underline{\rho}_{13})^{-1}\hat{b}_{-2}\|_2 > \delta$, *then* $\alpha_t^{1,*}\alpha_t^{3,*} < 0$.

4. *(Under-diversification: no investment in the first asset)*

(i) If $\bar{\rho}_{23} < \hat{\underline{\rho}}_{23}$, *and* $\hat{\kappa}^1(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23})\hat{\kappa}^1(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23}) \leq 0$, *then an optimal portfolio strategy is*

$$\begin{aligned}\begin{pmatrix}\alpha_t^{2,*}\\ \alpha_t^{3,*}\end{pmatrix} &= \big[x_0 + \frac{1}{2\lambda}e^{(\|\sigma_{-1}(\bar{\rho}_{23})^{-1}\hat{b}_{-1}\|_2-\delta)^2 T} - X_t^*\big]\Big(1 - \frac{\delta}{\|\sigma_{-1}(\bar{\rho}_{23})^{-1}\hat{b}_{-1}\|_2}\Big)\\ &\qquad 1_{\{\|\sigma_{-1}(\bar{\rho}_{23})^{-1}\hat{b}_{-1}\|_2>\delta\}}\Sigma_{-1}(\bar{\rho}_{23})^{-1}\hat{b}_{-1}\\ \alpha_t^{1,*} &\equiv 0,\end{aligned}$$

and if $\|\sigma_{-1}(\bar{\rho}_{23})^{-1}\hat{b}_{-1}\|_2 > \delta$, *then* $\alpha_t^{2,*}\alpha_t^{3,*} > 0$.

(ii) If $\underline{\rho}_{23} > \hat{\underline{\rho}}_{23}$, *and* $\hat{\kappa}^1(\underline{\rho}_{12}, \bar{\rho}_{13}, \underline{\rho}_{23})\hat{\kappa}^1(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23}) \leq 0$, *then an optimal portfolio strategy is*

$$\begin{aligned}\begin{pmatrix}\alpha_t^{2,*}\\ \alpha_t^{3,*}\end{pmatrix} &= \big[x_0 + \frac{1}{2\lambda}e^{(\|\sigma_{-1}(\underline{\rho}_{23})^{-1}\hat{b}_{-1}\|_2-\delta)^2 T} - X_t^*\big]\Big(1 - \frac{\delta}{\|\sigma_{-1}(\underline{\rho}_{23})^{-1}\hat{b}_{-1}\|_2}\Big)\\ &\qquad 1_{\{\|\sigma_{-1}(\underline{\rho}_{23})^{-1}\hat{b}_{-1}\|_2>\delta\}}\Sigma_{-1}(\underline{\rho}_{23})^{-1}\hat{b}_{-1}\\ \alpha_t^{1,*} &\equiv 0,\end{aligned}$$

and if $\|\sigma_{-1}(\underline{\rho}_{23})^{-1}\hat{b}_{-1}\|_2 > \delta$, *then* $\alpha_t^{2,*}\ \alpha_t^{3,*} < 0$.

5. *(Well-diversification)*

(i) If $\hat{\kappa}^1\hat{\kappa}^2(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23}) > 0$, *and* $\hat{\kappa}^1\hat{\kappa}^3(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23}) > 0$, *then an optimal portfolio strategy is given by*

$$\begin{aligned}\alpha_t^* &= \big[x_0 + \frac{1}{2\lambda}e^{(\|\sigma(\bar{\rho}_{12},\bar{\rho}_{13},\bar{\rho}_{23})^{-1}\hat{b}\|_2-\delta)^2 T} - X_t^*\big]\Big(1 - \frac{\delta}{\|\sigma(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23})^{-1}\hat{b}\|_2}\Big)\\ &\qquad 1_{\{\|\sigma(\bar{\rho}_{12},\bar{\rho}_{13},\bar{\rho}_{23})^{-1}\hat{b}\|_2>\delta\}}\Sigma(\bar{\rho}_{12}, \bar{\rho}_{13}, \bar{\rho}_{23})^{-1}\hat{b}.\end{aligned}$$

(ii) If $\hat{\kappa}^1\hat{\kappa}^2(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23}) < 0$, *and* $\hat{\kappa}^1\hat{\kappa}^3(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23}) < 0$, *then an optimal portfolio strategy is given by*

$$\begin{aligned}\alpha_t^* &= \big[x_0 + \frac{1}{2\lambda}e^{(\|\sigma(\underline{\rho}_{12},\underline{\rho}_{13},\bar{\rho}_{23})^{-1}\hat{b}\|_2-\delta)^2 T} - X_t^*\big]\Big(1 - \frac{\delta}{\|\sigma(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23})^{-1}\hat{b}\|_2}\Big)\\ &\qquad 1_{\{\|\sigma(\underline{\rho}_{12},\underline{\rho}_{13},\bar{\rho}_{23})^{-1}\hat{b}\|_2>\delta\}}\Sigma(\underline{\rho}_{12}, \underline{\rho}_{13}, \bar{\rho}_{23})^{-1}\hat{b}.\end{aligned}$$

(iii) If $\hat{\kappa}^1\hat{\kappa}^2(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23}) > 0$, *and* $\hat{\kappa}^1\hat{\kappa}^3(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23}) < 0$, *then an optimal portfolio strategy is given by*

$$\begin{aligned}\alpha_t^* &= \big[x_0 + \frac{1}{2\lambda}e^{(\|\sigma(\bar{\rho}_{12},\underline{\rho}_{13},\underline{\rho}_{23})^{-1}\hat{b}\|_2-\delta)^2 T} - X_t^*\big]\Big(1 - \frac{\delta}{\|\sigma(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23})^{-1}\hat{b}\|_2}\Big)\\ &\qquad 1_{\{\|\sigma(\bar{\rho}_{12},\underline{\rho}_{13},\underline{\rho}_{23})^{-1}\hat{b}\|_2>\delta\}}\Sigma(\bar{\rho}_{12}, \underline{\rho}_{13}, \underline{\rho}_{23})^{-1}\hat{b}.\end{aligned}$$

(iv) If $\hat{\kappa}^1\hat{\kappa}^2(\underline{\rho}_{12},\bar{\rho}_{13},\underline{\rho}_{23}) < 0$, *and* $\hat{\kappa}^1\hat{\kappa}^3(\underline{\rho}_{12},\bar{\rho}_{13},\underline{\rho}_{23}) > 0$, *then an optimal portfolio strategy is given by*

$$\alpha_t^* = \big[x_0 + \frac{1}{2\lambda}e^{\|(\sigma(\underline{\rho}_{12},\bar{\rho}_{13},\underline{\rho}_{23})^{-1}\hat{b}\|_2-\delta)^2 T} - X_t^*\big]\Big(1 - \frac{\delta}{\|\sigma(\underline{\rho}_{12},\bar{\rho}_{13},\underline{\rho}_{23})^{-1}\hat{b}\|_2}\Big) 1_{\{\|\sigma(\underline{\rho}_{12},\bar{\rho}_{13},\underline{\rho}_{23})^{-1}\hat{b}\|_2>\delta\}}\Sigma(\underline{\rho}_{12},\bar{\rho}_{13},\underline{\rho}_{23})^{-1}\hat{b}.$$

Proof. In view of formula (4.5) of the optimal portfolio strategy in Proposition 4.4, we only need to compute $\hat{\kappa}(\rho^*) = \Sigma(\rho^*)^{-1}\hat{b}$, and $\|\sigma(\rho^*)^{-1}\hat{b}\|_2$, i.e., $R(\hat{b},\rho^*)$, which have been given explicitly in the proof of Proposition 4.3 (see A.4 in Appendix) when computing ρ^*. In the case **1.**, we obtained (see (A.33)) $\hat{\kappa}(\rho^*) = (\frac{\hat{b}_1}{\sigma_1^2},0,0)^\intercal$, and $R(\hat{b},\rho^*) = \hat{b}^\intercal\hat{\kappa}(\rho^*) = \hat{\beta}_1^2$. In the case **2.**, let us focus on subcase (i) as the other subcase (ii) is dealt with similarly: we have $\rho_{12}^* = \bar{\rho}_{12}$, $(\hat{\kappa}^1(\bar{\rho}_{12}),\hat{\kappa}^2(\bar{\rho}_{12}))^\intercal = \Sigma_{-3}(\bar{\rho}_{12})\hat{b}_{-3}$, $\hat{\kappa}^3(\bar{\rho}_{12},\rho_{13}^*,\rho_{23}^*) = 0$, and $R(\hat{b},\rho^*) = \hat{b}_{-3}^\intercal\Sigma_{-3}(\bar{\rho}_{12})\hat{b}_{-3}$, by (A.40), (A.41). The other cases are computed in the same way and omitted here. □

Remark 4.9 (Financial interpretation) In case **1**, corresponding to large correlation ambiguity for the second and third asset (recall Remark 4.4), it is optimal to invest only in the first asset, namely the one with the highest *a priori* Sharpe ratio, which is consistent with the anti-diversification result obtained in Theorem 4.1 (see also Remark 4.6).

In case **2**, corresponding to a large correlation ambiguity for the third asset (see Remark 4.4), the investor does not invest in the third asset, but only in the first and second assets. Moreover, depending on whether the assets 1 and 2 have close Sharpe ratios with a correlation upper bound between these assets not too large (subcase (i)), or the asset 1 dominates the asset 2 in terms of Sharpe ratio (subcase (ii)), the investment in assets 1 or 2 follows a directional trading or a spread trading.

We have similar under-diversification effect in cases **3** and **4**, and notice that it may happen that one does not invest in the first asset even though it has the highest *a priori* Sharpe ratio. The result in case **4** is quite interesting and is *a priori* unexpected. Intuitively, an investor should always invest in the the asset with the greatest absolute Sharpe ratio. For example, this is the case when anti-diversification occurs and also in cases **1**, **2**, **3**, **5**. However, the case **4** means that the asset with the greatest absolute Sharpe ratio (asset 1) may not be traded in the optimal portfolio while the one with the smallest absolute Sharpe ratio (asset 3) may be traded. The idea is that depending on the drift and correlation ambiguity levels, investing in the two other assets may achieve a higher Sharpe ratio than investing in asset 1. In [24], the authors constructed a simple example (with no ambiguity in the correlation between asset 2 and asset 3) for this scenario.

Finally, in the case **5**, corresponding to a small correlation ambiguity (see Remark 4.4), the investor has interest to well-diversify her portfolio among the three assets. ♢

A Appendix

Notations, differentiation and characterization of convex function

Let us introduce some notations and state some results which will be used frequently in the proof of some Lemmas and Propositions.

1. We introduce the so-called variance risk ratios

$$\begin{array}{rcl} \hat{\kappa}(\rho) & := & \Sigma(\rho)^{-1}\hat{b} \;=\; (\hat{\kappa}^1(\rho),\ldots,\hat{\kappa}^d(\rho))^\intercal, \\ \kappa(b,\rho) & := & \Sigma(\rho)^{-1}b \;=\; (\kappa^1(b,\rho),\ldots,\kappa^d(b,\rho))^\intercal. \end{array}$$

2. From some matrix calculations (see e.g. corollary 95 and corollary 105 in [9]), we obtain the explicit expressions of the first partial derivatives of $R(b,\rho) = b^\intercal\kappa(b,\rho)$ with respect to b_i, ρ_{ij} denoted by $\frac{\partial R(b,\rho)}{\partial b_i}$ and $\frac{\partial R(b,\rho)}{\partial \rho_{ij}}$, $1 \le i < j \le d$,

$$\frac{\partial R(b,\rho)}{\partial b_i} \;=\; 2\kappa^i(b,\rho), \qquad \frac{\partial R(b,\rho)}{\partial \rho_{ij}} \;=\; -\sigma_i\sigma_j\kappa^i(b,\rho)\kappa^j(b,\rho). \tag{A.6}$$

We also denote by $\nabla_b R(b,\rho)$ and $\nabla_\rho R(b,\rho)$ the gradients of $R(b,\rho)$ with respect to b and ρ respectively,

$$\left\{ \begin{array}{rcl} \nabla_b R(b,\rho) & = & (\frac{\partial R(b,\rho)}{\partial b_1},\ldots,\frac{\partial R(b,\rho)}{\partial b_d})^\intercal \\ \nabla_\rho R(b,\rho) & = & (\frac{\partial R(b,\rho)}{\partial \rho_{12}},\ldots,\frac{\partial R(b,\rho)}{\partial \rho_{1d}},\ldots,\frac{\partial R(b,\rho)}{\partial \rho_{(d-1)d}})^\intercal \end{array} \right. \tag{A.7}$$

3. (*Sufficient and necessary optimality condition*). It is known (see e.g. Lemma 2.2 in [10]) that $R(b,\rho)$ is jointly convex in b and ρ. Similarly, $R(\hat{b},\rho)$ is convex in ρ. Then ρ^* is a global minimum of $\rho \mapsto R(\hat{b},\rho)$ over Γ convex set of $\mathbb{C}_+^d$ if and only if, for any $\rho \in \Gamma$ (see e.g. section 4.2.3 in [4]),

$$(\rho-\rho^*)^\intercal \nabla_\rho R(\hat{b},\rho^*) \;=\; \sum_{j=1}^{d}\sum_{i=1}^{j-1} \frac{\partial R(\hat{b},\rho^*)}{\partial \rho_{ij}}(\rho_{ij}-\rho^*_{ij}) \;\ge\; 0,$$

which is written from (A.6) as,

$$\sum_{j=1}^{d}\sum_{i=1}^{j-1} \sigma_i\sigma_j\hat{\kappa}^i\hat{\kappa}^j(\rho^*)(\rho_{ij}-\rho^*_{ij}) \;\le\; 0. \tag{A.8}$$

A.1 Proof of Lemma 3.2

The statement of Lemma 3.2 is minimax type theorem, as it implies obviously in the case where $\Theta = \Delta \times \Gamma$ is a rectangular set that the function H in (3.4) satisfies

$$\min_{b\in\Delta}\max_{\rho\in\Gamma} H(b,\rho) \;=\; \max_{\rho\in\Gamma}\min_{b\in\Delta} H(b,\rho).$$

However, its proof cannot be deduced directly from standard minimax theorem (see e.g. Theorem 45.8 in [32]), as it does not fulfill totally their conditions: the function H is linear (hence convex) in b, linear (hence concave) in ρ, but we do not assume that Γ is a compact set, and we also consider the case where Θ is an ellipsoidal set.

We distinguish the two cases in **(HΘ)** whether Θ is a rectangular or ellipsoidal set.

Lemma A.2 *Suppose that $\Theta = \Delta \times \Gamma$ is in product set as in* **(HΘ)***(i), and assume that there exists $\theta^* \in \arg\min_{\Theta} R(\theta)$. Then, we have for all $b \in \Delta$, $\rho \in \Gamma$:*

$$H(b^*, \rho) \;\le\; H(\theta^*) \;=\; R(\theta^*) \;\le\; H(b, \rho^*).$$

Moreover, (b^*, ρ^*) is a saddle point, namely,

$$\inf_{b\in\Delta}\sup_{\rho\in\Gamma} H(b,\rho) \;=\; \sup_{\rho\in\Gamma}\inf_{b\in\Delta} H(b,\rho) \;=\; H(b^*,\rho^*).$$

Proof. Note that if there exists $(b^*, \rho^*) \in \arg\min_{\Theta} R(\theta)$, the first-order condition implies that for any $(b, \rho) \in \Theta$,

$$(b - b^*)^\intercal \nabla_b R(\theta^*) + (\rho - \rho^*)^\intercal \nabla_\rho R(\theta^*) \;\ge\; 0. \tag{A.9}$$

Recalling $H(b, \rho)$ in (3.4) and explicit expressions (A.6), (A.7) of $\nabla_b R(\theta^*)$ and $\nabla_\rho R(\theta^*)$, and fixing $b = b^*$ or $\rho = \rho^*$ in (A.9) respectively, we get for any $b \in \Delta$, $\rho \in \Gamma$,

$$\begin{aligned}
H(b^*, \rho) - H(b^*, \rho^*) &= \sum_{j=1}^{d}\sum_{i=1}^{j-1} \kappa^i(b^*,\rho^*)\kappa^j(b^*,\rho^*)\sigma_i\sigma_j(\rho_{ij} - \rho^*_{ij}) \\
&= (\rho^* - \rho)^\intercal \nabla_\rho R(\theta^*) \;\le\; 0, \\
H(b, \rho^*) - H(b^*, \rho^*) &= \sum_{i=1}^{d} (b_i - b^*_i)\kappa^i(b^*,\rho^*) \\
&= \frac{1}{2}(b - b^*)^\intercal \nabla_b R(\theta^*) \;\ge\; 0,
\end{aligned}$$

where by convention, we set: $\sum_{i=1}^{0} \cdot = 0$. It follows that

$$\inf_{b\in\Delta}\sup_{\rho\in\Gamma} H(b,\rho) \;\le\; \sup_{\rho\in\Gamma} H(b^*,\rho) \;=\; H(b^*,\rho^*) \;=\; \inf_{b\in\Delta} H(b,\rho^*) \;\le\; \sup_{\rho\in\Gamma}\inf_{b\in\Delta} H(b,\rho). \tag{A.10}$$

Since we always have $\inf_{b\in\Delta}\sup_{\rho\in\Gamma} H(b,\rho) \ge \sup_{\rho\in\Gamma}\inf_{b\in\Delta} H(b,\rho)$, the above inequality is indeed an equality, and this proves the required result. □

Lemma A.3 *Suppose that $\Theta = \{(b, \rho) \in \mathbb{R}^d \times \Gamma : \|\sigma(\rho)^{-1}(b - \hat{b})\|_2 \le \delta\}$ is an ellipsoidal set as in* **(HΘ)***(ii), and assume that $\rho^* = \arg\min_{\Gamma} \|\sigma(\rho)^{-1}\hat{b}\|_2$ exists. Then there exists $\theta^* \in \arg\min_{\Theta} R(\theta)$ with $\theta^* = (b^*, \rho^*)$,*

$$b^* \;=\; \hat{b}\Big(1 - \frac{\delta}{\|\sigma(\rho^*)^{-1}\hat{b}\|_2}\Big)1_{\{\|\sigma(\rho^*)^{-1}\hat{b}\|_2 > \delta\}}, \tag{A.11}$$

and

$$R(\theta^*) \;=\; \big(\|\sigma(\rho^*)^{-1}\hat{b}\|_2 - \delta\big)^2 1_{\{\|\sigma(\rho^*)^{-1}\hat{b}\|_2 > \delta\}}. \tag{A.12}$$

Moreover, we have for all $b \in \Theta_b$, $\rho \in \Gamma$:

$$H(b^*, \rho) \;\le\; H(\theta^*) \;=\; R(\theta^*) \;\le\; H(b, \rho^*). \tag{A.13}$$

where Θ_b denotes the projection of the set Θ onto b-plane, i.e., $\Theta_b := \{b \in \mathbb{R}^d : b \in \Delta_\rho, \rho \in \Gamma\}$ with $\Delta_\rho := \{b \in \mathbb{R}^d : \|\sigma(\rho)^{-1}(b - \hat{b})\|_2 \le \delta\}$.

Proof. Due to the dependence of b on ρ in the ellipsoidal set Θ written as $\Theta = \{(b,\rho) \in \mathbb{R}^d \times \Gamma : b \in \Delta_\rho\}$ where Δ_ρ is defined in Lemma A.3. We use a Lagrangian approach here. *Step 1.* For fixed $\rho \in \Gamma$, let us first focus on the inner minimization

$$\min_{b\in\Delta_\rho} R(b,\rho). \tag{A.14}$$

The Lagrangian with nonnegative multiplier μ associated to this constrained minimization problem is

$$L_1(b,\mu) \;=\; R(b,\rho) - \mu\Big(\delta^2 - (b-\hat{b})^\intercal \Sigma(\rho)^{-1}(b-\hat{b})\Big), \tag{A.15}$$

and the first-order condition gives

$$\begin{array}{rcl} \dfrac{\partial L_1(b,\mu)}{\partial b} &=& 2\Sigma(\rho)^{-1}b + 2\mu\Sigma(\rho)^{-1}(b-\hat{b}) \;=\; 0 \\ \dfrac{\partial L_1(b,\mu)}{\partial \mu} &=& \delta^2 - (b-\hat{b})^\intercal\Sigma(\rho)^{-1}(b-\hat{b}) \;=\; 0. \end{array}$$

Solving these two equations for fixed ρ, and recalling that the Lagrange multiplier is nonnegative, yield

$$\left\{\begin{array}{rcl} \mu^*(\rho) &=& (\frac{\|\sigma(\rho)^{-1}\hat{b}\|_2}{\delta} - 1)1_{\{\|\sigma(\rho)^{-1}\hat{b}\|_2>\delta\}}, \\ b^*(\rho) &=& \hat{b}(1 - \frac{\delta}{\|\sigma(\rho)^{-1}\hat{b}\|_2})1_{\{\|\sigma(\rho)^{-1}\hat{b}\|_2>\delta\}}. \end{array}\right. \tag{A.16}$$

Substituting these expressions into the Lagrangian (A.15), we get

$$L_1(b^*(\rho),\rho) \;=\; R(b^*(\rho),\rho) \;=\; \big(\|\sigma(\rho)^{-1}\hat{b}\|_2 - \delta\big)^2 1_{\{\|\sigma(\rho)^{-1}\hat{b}\|_2>\delta\}}.$$

and thus, the original problem $\inf_\Theta R(\theta)$ is reduced to

$$\begin{array}{rcl} \inf\limits_{\theta=(b,\rho)\in\Theta} R(\theta) &=& \inf\limits_{\rho\in\Gamma}\inf\limits_{b\in\Delta_\rho} R(b,\rho) \;=\; \inf\limits_{\rho\in\Gamma} R(b^*(\rho),\rho) \\ &=& \inf\limits_{\rho\in\Gamma}\Big\{\big(\|\sigma(\rho)^{-1}\hat{b}\|_2 - \delta\big)^2 1_{\{\|\sigma(\rho)^{-1}\hat{b}\|_2>\delta\}}\Big\} \\ &=& \Big(\inf\limits_{\rho\in\Gamma}\|\sigma(\rho)^{-1}\hat{b}\|_2 - \delta\Big)^2 1_{\big\{\inf\limits_{\rho\in\Gamma}\|\sigma(\rho)^{-1}\hat{b}\|_2 \,>\, \delta\big\}}. \end{array} \tag{A.17}$$

Therefore, whenever $\rho^* \in \arg\min_\Gamma \|\sigma(\rho)^{-1}\hat{b}\|_2$ exists, we see from (A.17) that R attains its infimum at $\theta^* = (b^*,\rho^*)$ with $b^* = b^*(\rho^*)$ as in (A.16) with $\rho = \rho^*$, which leads to the expressions as described in (A.11) and (A.12) of Lemma A.3.

Step 2. Suppose that there exists $\rho^* \in \arg\min_\Gamma \|\sigma(\rho)^{-1}\hat{b}\|_2$. From *Step 1*, there exists $\theta^* = (b^*,\rho^*) \in \arg\min_\Theta R(\theta)$. Let us now prove that $H(b^*,\rho) \le R(\theta^*)$ for any $\rho \in \Gamma$. Substituting the expression (A.11) of b^* in $H(b^*,\rho)$, we rewrite $H(b^*,\rho)$ as

$$H(b^*,\rho) \;=\; \Big(1 - \frac{\delta}{\|\sigma(\rho^*)^{-1}\hat{b}\|_2}\Big)^2 \hat{b}^\intercal\Sigma(\rho^*)^{-1}\Sigma(\rho)\Sigma(\rho^*)^{-1}\hat{b}1_{\{\|\sigma(\rho^*)^{-1}\hat{b}\|_2>\delta\}}.$$

As $\rho^* \in \arg\min_\Gamma \hat{b}^\intercal\Sigma(\rho)^{-1}\hat{b}$, we use Lemma A.2 by setting $\Delta = \{\hat{b}\}$, and immediately obtain

$$\sup_{\rho\in\Gamma} \hat{b}^\intercal\Sigma(\rho^*)^{-1}\Sigma(\rho)\Sigma(\rho^*)^{-1}\hat{b} \;=\; \hat{b}^\intercal\Sigma(\rho^*)^{-1}\hat{b}. \tag{A.18}$$

By multiplying both sides of the above equality with the constant $\left(1-\frac{\delta}{\|\sigma(\rho^*)^{-1}\hat{b}\|_2}\right)^2 1_{\{\|\sigma(\rho^*)^{-1}\hat{b}\|_2>\delta\}}$, we get

$$\sup_{\rho\in\Gamma} H(b^*,\rho) \;=\; H(b^*,\rho^*) \;=\; R(\theta^*), \tag{A.19}$$

which shows that

$$H(b^*,\rho) \;\leq\; R(\theta^*), \;\text{ for all } \rho \;\in\; \Gamma.$$

Step 3. Let us finally prove that $H(b,\rho^*) \geq R(\theta^*)$ for any $b \in \Theta_b$. Again, we use a Lagrangian approach. For fixed $\rho \in \Gamma$, we focus on the inner minimization

$$\inf_{b\in\Delta_\rho} H(b,\rho^*),$$

and consider the associated Lagrangian function with nonnegative multiplier μ

$$L_2(b,\mu) \;=\; H(b,\rho^*) - \mu\big(\delta^2 - (b-\hat{b})^\intercal\Sigma(\rho)^{-1}(b-\hat{b})\big). \tag{A.20}$$

The first-order condition gives

$$\begin{aligned}
\frac{\partial L_2(b,\mu)}{\partial b} &= \Sigma(\rho^*)^{-1}b^* + 2\mu\Sigma(\rho)^{-1}(b-\hat{b}) \;=\; 0,\\
\frac{\partial L_2(b,\mu)}{\partial \mu} &= \delta^2 - (b-\hat{b})^\intercal\Sigma(\rho)^{-1}(b-\hat{b}) \;=\; 0,
\end{aligned}$$

and by solving these two equations for fixed ρ (recalling also that the Lagrangian multiplier is nonnegative), we get

$$\left\{\begin{array}{rcl}
\mu^{**}(\rho) &=& \frac{\sqrt{H(b^*,\rho)}}{2\delta} \;\geq\; 0\\
b^{**}(\rho) &=& \hat{b} - \frac{\delta}{\sqrt{H(b^*,\rho)}}\Sigma(\rho)\Sigma(\rho^*)^{-1}b^*.
\end{array}\right. \tag{A.21}$$

Substituting these expressions into the Lagrangian (A.20), we get

$$L_2(b^{**}(\rho),\rho) \;=\; H(b^{**}(\rho),\rho^*) \;=\; \hat{b}^\intercal\Sigma(\rho^*)^{-1}b^* - \delta\sqrt{H(b^*,\rho)}.$$

The outer minimization over Γ then yields

$$\begin{aligned}
\inf_{b\in\Theta_b} H(b,\rho^*) &= \inf_{\rho\in\Gamma}\inf_{b\in\Delta_\rho} H(b,\rho^*) \;=\; \inf_{\rho\in\Gamma} H(b^{**}(\rho),\rho^*)\\
&= \inf_{\rho\in\Gamma}\left\{\hat{b}^\intercal\Sigma(\rho^*)^{-1}b^* - \delta\sqrt{H(b^*,\rho)}\right\}\\
&= \hat{b}^\intercal\Sigma(\rho^*)^{-1}b^* - \delta\sup_{\rho\in\Gamma}\sqrt{H(b^*,\rho)}\\
&= \hat{b}^\intercal\Sigma(\rho^*)^{-1}b^* - \delta\sqrt{R(\theta^*)}\\
&= R(\theta^*),
\end{aligned}$$

where we used (A.19) in the last second equality, and last equality comes from (A.11). This shows that the infimum of $H(b,\rho^*)$ over $b \in \Theta_b$ is attained at $b^{**}(\rho^*) = b^*$ as in (A.21) with $\rho = \rho^*$. We conclude that for any $b \in \Theta_b$,

$$H(b,\rho^*) \;\geq\; R(\theta^*), \tag{A.22}$$

which completes the proof. □

A.2 Proof of Proposition 4.1

Let us prove that under the condition $|\hat{\beta}_1| > |\hat{\beta}_2| = \max_{i\neq 1} |\hat{\beta}_i|$, the function $\rho \mapsto R(\hat{b}, \rho)$ attains its infimum over $\mathbb{C}^d_{>+} \subset (-1,1)^{d(d-1)/2}$, and this infimum ρ^* can be computed explicitly. By convexity and differentiability of $\rho \mapsto R(\hat{b}, \rho)$ over the convex open set $\Gamma = \mathbb{C}^d_{>+}$, the existence of such minimum is equivalent to the existence of critical points to $R(\hat{b}, .)$, i.e.,

$$\frac{\partial R(\hat{b}, \rho^*)}{\partial \rho_{ij}} \;=\; 0, \;\; 1 \leq i < j \leq d. \tag{A.23}$$

Recalling that $\sigma_i > 0$ $i = 1, \ldots, d$, this is written from (A.6) as the system of $d(d-1)/2$ equations:

$$\hat{\kappa}^i(\rho^*)\hat{\kappa}^j(\rho^*) \;=\; 0, \quad 1 \leq i < j \leq d, \tag{A.24}$$

which indicates that at most one component of $\hat{\kappa}(\rho^*)$ is not zero. Notice that due to the assumption that $\hat{b} \neq 0$, $\hat{\kappa}(\rho^*) = \Sigma(\rho^*)^{-1}\hat{b}$ is never zero, i.e. at least one component of $\hat{\kappa}(\rho^*)$ is not zero. Therefore, exactly one component of $\hat{\kappa}(\rho^*)$ is not zero. Then (A.24) is equivalent to $\hat{\kappa}^{i_1}(\rho^*) \neq 0$, $\hat{\kappa}^j(\rho^*) = 0$, $j \neq i_1$, for some $i_1 = 1, \ldots, d$. In other words, we have

$$(0, \ldots, 0, \hat{\kappa}^{i_1}(\rho^*), 0, \ldots, 0)^\intercal \;=\; \Sigma(\rho^*)^{-1}\hat{b}, \qquad \text{for some } i_1 = 1, \ldots, d. \tag{A.25}$$

Pre-multiplying $\Sigma(\rho^*)$ on both sides of (A.25) and then writing out l.h.s, we obtain

$$\left\{ \begin{array}{rcl} \sigma_{i_1}^2 \hat{\kappa}^{i_1}(\rho^*) & = & \hat{b}_{i_1} \\ \sigma_{i_1}\sigma_i \rho^*_{i_1 i}\hat{\kappa}^{i_1}(\rho^*) & = & \hat{b}_i \quad 1 \leq i \leq d, \; i \neq i_1, \end{array} \right. \tag{A.26}$$

which yields the explicit form:

$$\left\{ \begin{array}{rcl} \hat{\kappa}^{i_1}(\rho^*) & = & \frac{\hat{b}_{i_1}}{\sigma^2_{i_1}} \\ \rho^*_{i_1 i} & = & \frac{\hat{\beta}_i}{\hat{\beta}_{i_1}} \quad i \neq i_1, \; 1 \leq i \leq d. \end{array} \right. \tag{A.27}$$

As $|\rho^*_{i_1 i}| < 1$ in (A.27), together with condition $|\hat{\beta}_1| > \max_{i\neq 1} |\hat{\beta}_i|$, we thus have $i_1 = 1$ and

$$\Sigma(\rho^*)^{-1}\hat{b} \;=\; \hat{\kappa}(\rho^*) \;=\; (\frac{\hat{b}_1}{\sigma_1^2}, 0, \ldots, 0)^\intercal, \quad \rho^*_{1i} \;=\; \hat{\varrho}_{1i}, \quad 2 \leq i \leq d. \tag{A.28}$$

Once $\{\rho^*_{1i}\}_{2\leq i\leq d}$ is given as in (A.28), we can complete the other values of $\rho^*_{ij} \in (-1,1)$ such that ρ^* belongs to $\mathbb{C}^d_{>+}$. For instance, by choosing as in Corollary 2 in [24], $\rho^*_{ij} = \rho^*_{1i}\rho^*_{1j} = \hat{\varrho}_{1i}\hat{\varrho}_{1j}$, $2 \leq i < j \leq d$, we check that $C(\rho^*) \in \mathbb{S}^d_{>+}$. Indeed, in this case we have

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ -\hat{\varrho}_{12} & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\hat{\varrho}_{1d} & 0 & \ldots & 1 \end{pmatrix} \begin{pmatrix} 1 & \hat{\varrho}_{12} & \hat{\varrho}_{13} & \ldots & \hat{\varrho}_{1d} \\ \hat{\varrho}_{12} & 1 & \hat{\varrho}_{12}\hat{\varrho}_{13} & \ldots & \hat{\varrho}_{12}\hat{\varrho}_{1d} \\ \hat{\varrho}_{13} & \hat{\varrho}_{12}\hat{\varrho}_{13} & 1 & \ldots & \hat{\varrho}_{13}\hat{\varrho}_{1d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\varrho}_{1d} & \hat{\varrho}_{12}\hat{\varrho}_{1d} & \hat{\varrho}_{13}\hat{\varrho}_{1d} & \ldots & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \ldots & 0 \\ -\hat{\varrho}_{12} & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\hat{\varrho}_{1d} & 0 & \ldots & 1 \end{pmatrix}^\intercal$$
$$= \mathrm{diag}\{1, 1-|\hat{\varrho}_{12}|^2, 1-|\hat{\varrho}_{13}|^2, \ldots, 1-|\hat{\varrho}_{1d}|^2\},$$

which is positive definite since $1-|\hat{\varrho}_{1i}|^2>0$, $i=2,\ldots,d$.

In particular when $|\hat{\beta}_i|$ is in strictly descending order, i.e. $|\hat{\beta}_1|>|\hat{\beta}_2|>\ldots>|\hat{\beta}_d|$, $\rho^*=(\rho^*_{ij})_{1\le i<j\le d}=(\hat{\varrho}_{ij})_{1\le i\le j\le d}=:\hat{\varrho}$ also belongs to $\mathbb{C}^d_{>+}$. Indeed, in this case, observe that

$$\begin{pmatrix} 1 & 0 & \ldots & 0\\ -\hat{\varrho}_{12} & 1 & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ -\hat{\varrho}_{1d} & 0 & \ldots & 1\end{pmatrix} C(\hat{\varrho}) \begin{pmatrix} 1 & 0 & \ldots & 0\\ -\hat{\varrho}_{12} & 1 & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ -\hat{\varrho}_{1d} & 0 & \ldots & 1\end{pmatrix}^\intercal = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0\\ 0 & 1-\hat{\varrho}^2_{12} & \hat{\varrho}_{23}-\hat{\varrho}_{12}\hat{\varrho}_{13} & \ldots & \hat{\varrho}_{2d}-\hat{\varrho}_{12}\hat{\varrho}_{1d}\\ 0 & \hat{\varrho}_{23}-\hat{\varrho}_{12}\hat{\varrho}_{13} & 1-\hat{\varrho}^2_{13} & \ldots & \hat{\varrho}_{3d}-\hat{\varrho}_{13}\hat{\varrho}_{1d}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & \hat{\varrho}_{2d}-\hat{\varrho}_{12}\hat{\varrho}_{1d} & \hat{\varrho}_{3d}-\hat{\varrho}_{13}\hat{\varrho}_{1d} & \ldots & 1-\hat{\varrho}^2_{1d}\end{pmatrix}.$$

Denote by $C_1(\hat{\varrho})$ the matrix in the r.h.s of the above equality and note that $\hat{\varrho}_{1i}=\hat{\varrho}_{12}\hat{\varrho}_{2i}$, $i=3,\ldots,d$. Then we have

$$\begin{aligned} &\begin{pmatrix} 1 & 0 & 0 & \ldots & 0\\ 0 & 1 & 0 & \ldots & 0\\ 0 & -\hat{\varrho}_{23} & 1 & \ldots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & -\hat{\varrho}_{2d} & 0 & \ldots & 1\end{pmatrix} C_1(\hat{\varrho}) \begin{pmatrix} 1 & 0 & 0 & \ldots & 0\\ 0 & 1 & 0 & \ldots & 0\\ 0 & -\hat{\varrho}_{23} & 1 & \ldots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & -\hat{\varrho}_{2d} & 0 & \ldots & 1\end{pmatrix}^\intercal\\ =\ &\begin{pmatrix} 1 & 0 & 0 & 0 & \ldots & 0\\ 0 & 1-\hat{\varrho}^2_{12} & 0 & 0 & \ldots & 0\\ 0 & 0 & 1-\hat{\varrho}^2_{23} & \hat{\rho}_{34}-\hat{\rho}_{23}\hat{\rho}_{24} & \ldots & \hat{\rho}_{3d}-\hat{\varrho}_{23}\hat{\varrho}_{2d}\\ 0 & 0 & \hat{\varrho}_{34}-\hat{\varrho}_{23}\hat{\varrho}_{24} & 1-\hat{\varrho}^2_{24} & \ldots & \hat{\varrho}_{4d}-\hat{\varrho}_{24}\hat{\varrho}_{2d}\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \hat{\varrho}_{3d}-\hat{\varrho}_{23}\hat{\varrho}_{2d} & \hat{\varrho}_{4d}-\hat{\varrho}_{24}\hat{\varrho}_{2d} & \ldots & 1-\hat{\varrho}^2_{2d}\end{pmatrix}. \end{aligned}$$

Denote by $C_2(\hat{\varrho})$ the matrix in the r.h.s of the above equality and again note that $\hat{\varrho}_{2i}=\hat{\varrho}_{23}\,\hat{\varrho}_{3i}$, $i=4,\ldots,d$. Then we can do the similar matrix congruence with $C_2(\hat{\varrho})$ as with $C_1(\hat{\varrho})$. And so on. After $d-1$ steps of matrix congruence, we arrive at the diagonalization of the matrix $C(\hat{\varrho})$

$$TC(\hat{\varrho})T^\intercal \;=\; \mathrm{diag}\{1,1-|\hat{\varrho}_{12}|^2,1-|\hat{\varrho}_{23}|^2,\ldots,1-|\hat{\varrho}_{d-1d}|^2\},$$

where $T=T_d\cdots T_1$ with T_i being invertible matrix with diagonal terms 1, (j,i)-th term $-\hat{\varrho}_{ij}$, $j>i$, and other terms 0.

We deduce that the system of equations (A.24) has solutions in $\mathbb{C}^d_{>+}$ given by (A.28). Moreover, we have from (A.28)

$$\min_{\rho\in\mathbb{C}^d_{>+}} R(\hat{b},\rho) \;=\; R(\hat{b},\rho^*) \;=\; \hat{b}^\intercal\hat{\kappa}(\rho^*) \;=\; \hat{b}_1\hat{\kappa}^1(\rho^*) \;=\; \hat{\beta}_1^2. \tag{A.29}$$

Combining this with Lemma A.3, we obtain b^* described in 4.1. □

A.3 Proof of Proposition 4.2

As $\Gamma=[\underline{\rho},\bar{\rho}]$ is compact, we already know that $\rho^*=\arg\min_{\rho\in\Gamma}R(\hat{b},\rho)$ exists, and from Lemma A.3, we only need to compute the minimum of the function $\rho\mapsto R(\hat{b},\rho)$ over Γ. From (A.8) with $d=2$, we obtain the sufficient and necessary condition of ρ^* for being global minimum of $R(\hat{b},\rho)$ over Γ:

$$\sigma_1\sigma_2\hat{\kappa}^1(\rho^*)\hat{\kappa}^2(\rho^*)(\rho-\rho^*)\le 0,\quad \forall\rho\in[\underline{\rho},\bar{\rho}], \tag{A.30}$$

where $\hat{\kappa}(\rho)$ is explicitly written as

$$\hat{\kappa}(\rho) \;=\; \frac{1}{1-\rho^2}\begin{pmatrix} \frac{\hat{b}_1}{\sigma_1^2} - \frac{\hat{b}_2}{\sigma_1\sigma_2}\rho \\ \frac{\hat{b}_2}{\sigma_2^2} - \frac{\hat{b}_1}{\sigma_1\sigma_2}\rho \end{pmatrix} \;=\; \frac{1}{1-\rho^2}\begin{pmatrix} \frac{\hat{\beta}_1 - \hat{\beta}_2\rho}{\sigma_1} \\ \frac{\hat{\beta}_2 - \hat{\beta}_1\rho}{\sigma_2} \end{pmatrix}. \tag{A.31}$$

From (A.30), we have three possible cases:

- $\hat{\kappa}^1(\rho^*)\hat{\kappa}^2(\rho^*) = 0$. From the explicit expression (A.31) of $\hat{\kappa}(\rho^*)$, and as ρ^* has to belong to $[\underline{\rho}, \bar{\rho}] \subset (-1, 1)$, we obtain $\hat{\kappa}^2(\rho^*) = 0$, i.e., $\rho^* = \hat{\varrho}_{12}$, and so $R(\hat{b}, \rho^*) = \hat{\beta}_1^2$.
- $\hat{\kappa}^1(\rho^*)\hat{\kappa}^2(\rho^*) > 0$. Then (A.30) is satisfied iff $\rho^* = \bar{\rho}$. Moreover, from the above explicit expression of $\hat{\kappa}(\rho^*)$, we obtain $\bar{\rho} < \hat{\varrho}_{12}$.
- $\hat{\kappa}^1(\rho^*)\hat{\kappa}^2(\rho^*) < 0$. Then (A.30) is satisfied iff $\rho^* = \underline{\rho}$. Moreover, from the explicit expression of $\hat{\kappa}(\rho^*)$, we obtain $\underline{\rho} > \hat{\varrho}_{12}$.

By combining this with Lemma A.3, we obtain b^* described as in Proposition 4.2. □

A.4 Proof of Proposition 4.3

As $\Gamma = \prod_{j=1}^{3}\prod_{i=1}^{j-1}[\underline{\rho}_{ij}, \bar{\rho}_{ij}]$ is compact, we already know that $\rho^* = \arg\min_{\rho\in\Gamma} R(\hat{b}, \rho)$ exists. From Lemma A.3, we only need to compute the minimum of the function $\rho \mapsto R(\hat{b}, \rho)$ over Γ by applying the optimality principle (A.8) when $d = 3$,

$$\sum_{j=1}^{3}\sum_{i=1}^{j-1}\sigma_i\sigma_j\hat{\kappa}^i(\rho^*)\hat{\kappa}^j(\rho^*)(\rho_{ij} - \rho^*_{ij}) \le 0 \quad \text{for any } \rho \in \Gamma. \tag{A.32}$$

We observe from (A.32) that similar as Proposition 4.2, each ρ^*_{ij}, $1 \le i < j \le 3$ may be lower bound $\underline{\rho}_{ij}$, upper bound $\bar{\rho}_{ij}$, or an interior point in $(\underline{\rho}_{ij}, \bar{\rho}_{ij})$, which corresponds to $\kappa^i\kappa^j(\rho^*) > 0$, $\kappa^i\kappa^j(\rho^*) < 0$, or $\kappa^i\kappa^j(\rho^*) = 0$ respectively. Therefore, we consider the following possible exclusive cases depending on the number of zero components in $\hat{\kappa}(\rho^*)$:

1. $\hat{\kappa}^1\hat{\kappa}^2(\rho^*) = 0$, $\hat{\kappa}^1\hat{\kappa}^3(\rho^*) = 0$, $\hat{\kappa}^2\hat{\kappa}^3(\rho^*) = 0$.
In this case, (A.32) is immediately satisfied. As we assume that $\hat{b} \neq 0$, $\hat{\kappa}(\rho^*)$ is not zero, i.e. at least one component of $\hat{\kappa}(\rho^*)$ is nonzero. Then, two components of $\hat{\kappa}(\rho^*)$ are zero. Under the assumption that $|\hat{\beta}_1| \ge |\hat{\beta}_2| \ge |\hat{\beta}_3|$, (A.28) and (A.29) in Section A.2 yield the explicit expressions of ρ^*, $\hat{\kappa}(\rho^*)$ and $R(\hat{b}, \rho^*)$

$$\rho^*_{12} \;=\; \hat{\varrho}_{12} \in [\underline{\rho}_{12}, \bar{\rho}_{12}],\; \rho^*_{13} \;=\; \hat{\varrho}_{13} \in [\underline{\rho}_{13}, \bar{\rho}_{13}],\; \text{any } \rho^*_{23} \in [\underline{\rho}_{23}, \bar{\rho}_{23}]$$

and

$$\hat{\kappa}(\rho^*) \;=\; (\frac{\hat{b}_1}{\sigma_1^2}, 0, 0)^\intercal, \quad R(\hat{b}, \rho^*) \;=\; \hat{\beta}_1^2. \tag{A.33}$$

Let us show that $\hat{\beta}_1^2$ in (A.33) is strict minimum value in the sense that $R(\hat{b}, \rho^*) = \hat{\beta}_1^2$ iff $\rho^*_{12} = \hat{\varrho}_{12} \in [\underline{\rho}_{12}, \bar{\rho}_{12}]$, $\rho^*_{13} = \hat{\varrho}_{13} \in [\underline{\rho}_{13}, \bar{\rho}_{13}]$ and any $\rho^*_{23} \in [\underline{\rho}_{23}, \bar{\rho}_{23}]$. We express $\Sigma(\rho)$ as the following block matrix

$$\Sigma(\rho) = \begin{pmatrix} \sigma_1^2 & C_1^\intercal \\ C_1 & \Sigma_{-1}(\rho_{23}) \end{pmatrix},$$

where the vector $C_1 = (\sigma_1\sigma_2\rho_{12}, \sigma_1\sigma_3\rho_{13})^\intercal$.

$$\begin{pmatrix} 1 & \mathbf{0}_{1\times 2} \\ -\frac{C_1}{\sigma_1^2} & I_{2\times 2} \end{pmatrix} \begin{pmatrix} \sigma_1^2 & C_1^\intercal \\ C_1 & \Sigma_{-1}(\rho_{23}) \end{pmatrix} \begin{pmatrix} 1 & -\frac{C_1^\intercal}{\sigma_1^2} \\ \mathbf{0}_{2\times 1} & I_{2\times 2} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \mathbf{0}_{1\times 2} \\ \mathbf{0}_{2\times 1} & A \end{pmatrix}, \quad \text{(A.34)}$$

where $I_{2\times 2}$ denotes 2×2 identity matrix and $A = \Sigma_{-1}(\rho_{23}) - \frac{C_1 C_1^\intercal}{\sigma_1^4}$ is 2×2 positive definite matrix.
Inverting on both sides of (A.34), we get

$$\Sigma^{-1}(\rho) = \begin{pmatrix} 1 & -\frac{C_1^\intercal}{\sigma_1^2} \\ \mathbf{0}_{2\times 1} & I_{2\times 2} \end{pmatrix} \begin{pmatrix} \sigma_1^{-2} & \mathbf{0}_{1\times 2} \\ \mathbf{0}_{2\times 1} & A^{-1} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}_{1\times 2} \\ -\frac{C_1}{\sigma_1^2} & I_{2\times 2} \end{pmatrix}. \quad \text{(A.35)}$$

We express $\hat{b}$ as $(\hat{b}_1, \hat{b}_{-1}^\intercal)^\intercal$ and then write $R(\hat{b}, \rho)$ as two nonnegative decompositions from (A.35),

$$\begin{aligned} R(\hat{b}, \rho) &= \hat{\beta}_1^2 + (\hat{b}_{-1} - \frac{\hat{b}_1}{\sigma_1^2} C_1)^\intercal A^{-1} (\hat{b}_{-1} - \frac{\hat{b}_1}{\sigma_1^2} C_1) \\ &\geq \hat{\beta}_1^2, \end{aligned}$$

where in the last inequality, '=' holds if and only if $\hat{b}_{-1} - \frac{\hat{b}_1}{\sigma_1^2} C_1 = 0$, i.e. $\rho_{12}^* = \hat{\varrho}_{12}$, $\rho_{13}^* = \hat{\varrho}_{13}$. This corresponds to case **1.** of Proposition 4.3.

2. $\hat{\kappa}^1\hat{\kappa}^2(\rho^*) \neq 0$, $\hat{\kappa}^1\hat{\kappa}^3(\rho^*) = 0$, $\hat{\kappa}^2\hat{\kappa}^3(\rho^*) = 0$.
In this case, we express $\Sigma(\rho)$ as the following block-matrix form for convenience,

$$\Sigma(\rho) = \begin{pmatrix} \Sigma_{-3}(\rho_{12}) & C_3 \\ C_3^\intercal & \sigma_3^2 \end{pmatrix},$$

where the vector $C_3 = (\sigma_1\sigma_3\rho_{13}, \sigma_2\sigma_3\rho_{23})^\intercal$.
By first transforming $\Sigma(\rho)$ to block diagonal matrix as (A.35) and then taking inverse, we obtain

$$\begin{aligned} \Sigma(\rho)^{-1} = &\begin{pmatrix} I_{2\times 2} & -\Sigma_{-3}(\rho_{12})^{-1} C_3 \\ \mathbf{0}_{1\times 2} & 1 \end{pmatrix} \begin{pmatrix} \Sigma_{-3}(\rho_{12})^{-1} & \mathbf{0}_{2\times 1} \\ \mathbf{0}_{1\times 2} & a(\rho)^{-1} \end{pmatrix} \\ &\begin{pmatrix} I_{2\times 2} & \mathbf{0}_{2\times 1} \\ -C_3^\intercal \Sigma_{-3}(\rho_{12})^{-1} & 1 \end{pmatrix}, \end{aligned} \quad \text{(A.36)}$$

where $a(\rho) = \sigma_3^2 - C_3^\intercal \Sigma_{-3}(\rho_{12})^{-1} C_3$ is positive.

Recalling the definition of $\kappa(\hat{b}, \rho)$ and $R(\hat{b}, \rho)$, we obtain from (A.36)

$$\begin{cases} \begin{pmatrix} \hat{\kappa}^1(\rho) \\ \hat{\kappa}^2(\rho) \end{pmatrix} = \Sigma_{-3}(\rho_{12})^{-1} \hat{b}_{-3} - \hat{\kappa}^3(\rho) \Sigma_{-3}(\rho_{12})^{-1} C_3 \\ \hat{\kappa}^3(\rho) = \frac{1}{a(\rho)} (\hat{b}_3 - C_3^\intercal \Sigma_{-3}(\rho_{12})^{-1} \hat{b}_{-3}) \end{cases} \quad \text{(A.37)}$$

and

$$R(\hat{b}, \rho) = \hat{b}_{-3}^\intercal \Sigma_{-3}(\rho_{12})^{-1} \hat{b}_{-3} + a(\rho) \big(\hat{\kappa}^3(\rho)\big)^2. \quad \text{(A.38)}$$

In the following, we write $\hat{b}_{-3}^{\intercal}\Sigma_{-3}(\rho_{12})^{-1}\hat{b}_{-3}$ as $R(\hat{b}_{-3},\rho_{12})$.
As $\hat{\kappa}^3(\rho^*)=0$, we obtain from (A.32) that

$$\sigma_1\sigma_2\hat{\kappa}^1\hat{\kappa}^2(\rho_{12}^*)(\rho_{12}-\rho_{12}^*) \;\leq\; 0 \quad \text{for all } \rho_{12}\in[\underline{\rho}_{12},\bar{\rho}_{12}] \tag{A.39}$$

and from (A.37) and (A.38) that

$$\begin{cases} \begin{pmatrix}\hat{\kappa}^1(\rho_{12}^*)\\ \hat{\kappa}^2(\rho_{12}^*)\end{pmatrix} &= \Sigma_{-3}(\rho_{12}^*)^{-1}\hat{b}_{-3} \\ R(\hat{b},\rho^*) &= R(\hat{b}_{-3},\rho_{12}^*). \end{cases} \tag{A.40}$$

This is the case of ambiguous correlation in the two-risky assets: risky asset 1 and risky asset 2 with ambiguous correlation ρ_{12} in $[\underline{\rho}_{12},\bar{\rho}_{12}]$. In this case, $\hat{\kappa}^1(\rho^*)$ and $\hat{\kappa}^2(\rho^*)$ are not zero, therefore we have that from Proposition 4.2

$$\rho_{12}^* \;=\; \bar{\rho}_{12}1_{\{\bar{\rho}_{12}<\hat{\varrho}_{12}\}} \,+\, \underline{\rho}_{12}1_{\{\underline{\rho}_{12}>\hat{\varrho}_{12}\}}. \tag{A.41}$$

By setting $g(\rho_{12}^*,\rho_{13},\rho_{23}) := a(\rho_{12}^*,\rho_{13},\rho_{23})\kappa^3(\rho_{12}^*,\rho_{13},\rho_{23})$ for fixed ρ_{12}^* in (A.41), we deduce from (A.37) that the function

$$(\rho_{13},\rho_{23})\mapsto g(\rho_{12}^*,\rho_{13},\rho_{23}) \;=\; \hat{b}_3-\sigma_1\sigma_3\hat{\kappa}^1(\rho_{12}^*)\rho_{13}-\sigma_2\sigma_3\hat{\kappa}^2(\rho_{12}^*)\rho_{23},$$

is linear in $(\rho_{13},\rho_{23})\in[\underline{\rho}_{13},\bar{\rho}_{13}]\times[\underline{\rho}_{23},\bar{\rho}_{23}]$, and has the same sign as $\hat{\kappa}^3(\rho_{12}^*,\rho_{13},\rho_{23})$ due to the positiveness of $a(\rho_{12}^*,\rho_{13},\rho_{23})$. To study the condition of $\kappa^3(\rho^*)=0$, we discuss it in the following two cases:

(i) if $\bar{\rho}_{12}<\hat{\varrho}_{12}$, then $\hat{\kappa}^1\hat{\kappa}^2(\bar{\rho}_{12})>0$, the function $(\rho_{13},\rho_{23})\mapsto g(\bar{\rho}_{12},\rho_{13},\rho_{23})$ has the same monotonicity with respect to ρ_{13}, ρ_{23}. Therefore, to ensure that the function $g(\bar{\rho}_{12},\rho_{13},\rho_{23})$ has a root in $[\underline{\rho}_{13},\bar{\rho}_{13}]\times[\underline{\rho}_{23},\bar{\rho}_{23}]$, we need $g(\bar{\rho}_{12},\underline{\rho}_{13},\underline{\rho}_{23})$ $g(\bar{\rho}_{12},\bar{\rho}_{13},\bar{\rho}_{23})\leq 0$, or equivalently $\hat{\kappa}^3(\bar{\rho}_{12},\underline{\rho}_{13},\underline{\rho}_{23})$ $\hat{\kappa}^3(\bar{\rho}_{12},\bar{\rho}_{13},\bar{\rho}_{23})\leq 0$.

(ii) if $\underline{\rho}_{12}>\hat{\varrho}_{12}$, then $\hat{\kappa}^1\hat{\kappa}^2(\underline{\rho}_{12})<0$, the function $(\rho_{13},\rho_{23})\mapsto g(\underline{\rho}_{12},\rho_{13},\rho_{23})$ has the opposite monotonicity with respect to ρ_{13}, ρ_{23}. Therefore, when $g(\underline{\rho}_{12},\bar{\rho}_{13},\underline{\rho}_{23})$ $g(\underline{\rho}_{12},\underline{\rho}_{13},\bar{\rho}_{23})\leq 0$, or equivalently $\hat{\kappa}^3(\underline{\rho}_{12},\bar{\rho}_{13},\underline{\rho}_{23})\hat{\kappa}^3(\underline{\rho}_{12},\underline{\rho}_{13},\bar{\rho}_{13})\leq 0$, the function $g(\underline{\rho}_{12},\rho_{13},\rho_{23})$ has a root in $[\underline{\rho}_{13},\bar{\rho}_{13}]\times[\underline{\rho}_{23},\bar{\rho}_{23}]$.

Therefore, we deduce that $R(\hat{b},\rho)\geq R(\hat{b}_{-3},\rho_{12})\geq R(\hat{b}_{-3},\bar{\rho}_{12}1_{\{\bar{\rho}_{12}<\hat{\varrho}_{12}\}}+\underline{\rho}_{12}1_{\{\underline{\rho}_{12}>\hat{\varrho}_{12}\}})$ and that '=' holds if and only if $\rho_{12}^*=\bar{\rho}_{12}1_{\{\bar{\rho}_{12}<\hat{\varrho}_{12}\}}+\underline{\rho}_{12}1_{\{\underline{\rho}_{12}>\hat{\varrho}_{12}\}}$ and ρ_{13}^*, ρ_{23}^* satisfies $\hat{\kappa}^3(\rho_{12}^*,\rho_{13}^*,\rho_{23}^*)=0$.This corresponds to subcases **2.**(i) and **2.**(ii) of Proposition 4.3.

3. $\hat{\kappa}^1\hat{\kappa}^2(\rho^*)=0$, $\hat{\kappa}^1\hat{\kappa}^3(\rho^*)\neq 0$, $\hat{\kappa}^2(\rho^*)\hat{\kappa}^3(\rho^*)=0$.
In this case, we make permutations as follows,

$$\begin{pmatrix}\hat{\kappa}^{-2}(\rho)\\ \hat{\kappa}^2(\rho)\end{pmatrix} \;=\; \begin{pmatrix}\Sigma_{-2}(\rho_{13}) & C_2\\ C_2^{\intercal} & \sigma_2^2\end{pmatrix}^{-1}\begin{pmatrix}\hat{b}_{-2}\\ \hat{b}_2\end{pmatrix}, \tag{A.42}$$

where $\hat{\kappa}^{-2}(\rho)=(\hat{\kappa}^1(\rho),\hat{\kappa}^3(\rho))^{\intercal}$ and $C_2=(\sigma_1\sigma_2\rho_{12},\sigma_2\sigma_3\rho_{23})^{\intercal}$. Using (A.42) and proceeding with the same arguments as in the case **2.**, we obtain the result of $\hat{\kappa}^2(\rho^*)=0$, $\hat{\kappa}^1(\rho^*)\hat{\kappa}^3(\rho^*)\neq 0$ as described in the subcases **3.**(i) and **3.**(ii) of Proposition 4.3.

4. $\hat{\kappa}^1\hat{\kappa}^2(\rho^*) = 0$, $\hat{\kappa}^1\hat{\kappa}^3(\rho^*) = 0$, $\hat{\kappa}^2(\rho^*)\hat{\kappa}^3(\rho^*) \neq 0$.
Note that

$$\begin{pmatrix} \hat{\kappa}^{-1}(\rho) \\ \hat{\kappa}^1(\rho) \end{pmatrix} = \begin{pmatrix} \Sigma_{-1}(\rho_{23}) & C_1 \\ C_1^\intercal & \sigma_1^2 \end{pmatrix}^{-1} \begin{pmatrix} \hat{b}_{-1} \\ \hat{b}_1 \end{pmatrix}, \tag{A.43}$$

where $\hat{\kappa}^{-1}(\rho) = (\hat{\kappa}^2(\rho), \hat{\kappa}^3(\rho))^\intercal$ and $C_1 = (\sigma_1\sigma_2\rho_{12}, \sigma_1\sigma_3\rho_{13})^\intercal$. Using (A.43) and proceeding with the same arguments as in the case **2.**, we obtain the result of $\hat{\kappa}^1(\rho^*) = 0$, $\hat{\kappa}^2(\rho^*)\hat{\kappa}^3(\rho^*) \neq 0$ as described in subcases **4.**(i) and **4.**(ii) of Proposition 4.3.

5. $\hat{\kappa}^1\hat{\kappa}^2(\rho^*) \neq 0$, $\hat{\kappa}^1\hat{\kappa}^3(\rho^*) \neq 0$, $\hat{\kappa}^2(\rho^*)\hat{\kappa}^3(\rho^*) \neq 0$.
In this case, we see from (A.32) that each ρ_{ij}^* takes value in $\{\underline{\rho}_{ij}, \bar{\rho}_{ij}\}$ relying on the sign of $\hat{\kappa}^i\hat{\kappa}^j(\rho^*)$. Note that once the signs of $\hat{\kappa}^1\hat{\kappa}^2(\rho^*)$ and $\hat{\kappa}^1\hat{\kappa}^3(\rho^*)$ are known, the sign of $\hat{\kappa}^2(\rho^*)\hat{\kappa}^3(\rho^*)$ is determined. Therefore, by combination, there are 4 possible sub-cases as described in the case **5.** of Proposition 4.3.
As $\hat{\kappa}^i(\rho^*)\hat{\kappa}^j(\rho^*) \neq 0$ in each subcase, left right hand of (A.32) is strictly negative for any $\rho \in \Gamma \setminus \{\rho^*\}$. From the first-order characterization for convexity of $R(\hat{b}, \rho)$ (see e.g. Section 3.1.3 in [4]) and (A.6), we obtain for any $\rho \in \Gamma \setminus \{\rho^*\}$,

$$\begin{aligned} R(\hat{b}, \rho) &\geq R(\hat{b}, \rho^*) + (\rho - \rho^*)^\intercal \nabla_\rho R(\hat{b}, \rho^*) \\ &= R(\hat{b}, \rho^*) - \sum_{j=1}^{3} \sum_{i=1}^{j-1} \sigma_i\sigma_j\hat{\kappa}^i(\rho^*)\hat{\kappa}^j(\rho^*)(\rho_{ij} - \rho_{ij}^*) \\ &> R(\hat{b}, \rho^*), \end{aligned}$$

which indicates that ρ^* in each sub-case of case **5.** in Proposition 4.3 is a strict minimum of $R(\hat{b}, \rho)$.

As $R(\hat{b}, \rho^*)$ in this subcase is strict minimum value, we conclude that each subcase in Proposition A.3 is exclusive.
By combining this with Lemma A.3, we obtain b^* described as in Proposition 4.3. □

References

[1] Becherer D. and K. Kentia (2017): "Good deal hedging and valuation under combined uncertainty about drift and volatility". *Probability, Uncertainty and Quantitative Risk*, **2**(1), 13.

[2] Ben-Tal A., Margelit T. and A. Nemirovski (2000): "Robust modeling of multi-stage portfolio problems", In: Frenk H, Roos K, Terlaky T, Zhang S, editors. High-performance optimization. Dordrecht: Kluwer Academic Publishers, 303-328.

[3] Biagini S. and M. C. Pinar (2017): "The robust Merton problem of an ambiguity averse investor", *Mathematics and Financial Economics*, **11**(1), 1-24.

[4] Boyd, S. and L. Vandenberghe(2004): "Convex optimization", *Cambridge university press*.

[5] Boyle P., Garlappi L., Uppal R. and T. Wang (2012): "Keynes meets Markowitz: The trade-off between familiarity and diversification", *Management Science*, **58**(2), 253-272.

[6] Calvet L. E., Campbell J. Y. and P. Sodini (2007): "Down or out: Assessing the welfare costs of household investment mistakes", *Journal of Political Economy*, **115**(5), 707-747.

[7] Cooper I. and E. Kaplanis (1994): "Home bias in equity portfolios, inflation hedging, and international capital market equilibrium", *The Review of Financial Studies*, **7**(1), 45-60.

[8] Denis L. and M. Kervarec (2013): "Optimal investment under model uncertainty in nondominated models", *SIAM Journal on Control and Optimization*, **51**(3), 1803-1822.

[9] Dhrymes P. J.(1978): "Mathematics for econometrics" (No. 04; QA184, D4.). *New York: Springer-Verlag.*

[10] Ding K.W., Chen Z.Y. and N.J. Huang (2018): "Robust mean variance optimization problem under Rényi divergence information", *Optimization*, **67**(2), 287-307.

[11] Fouque J. P., Pun C. S. and H. Y. Wong (2016): "Portfolio optimization with ambiguous correlation and stochastic volatilities", *SIAM Journal on Control and Optimization*, **54**(5), 2309-2338.

[12] French K.R. and J.M. Poterba (1991): "Investor diversification and international equity markets", *National Bureau of Economic Research*, No. w3609.

[13] Garlappi L., Uppal, R. and T. Wang (2006): "Portfolio selection with parameter and model uncertainty: A multi-prior approach", *The Review of Financial Studies*, **20**(1), 41-81.

[14] Glasserman P. and X. Xu (2013): "Robust Portfolio Control with Stochastic Factor Dynamics", *Operations Research*, **61**(4), 874-893.

[15] Guidolin, M. and H. Liu (2016): "Ambiguity aversion and underdiversification", *Journal of Financial and Quantitative Analysis*, **51**(4), 1297-1323.

[16] Huberman G. (2001): "Familiarity breeds investment", *The Review of Financial Studies*, **14**(3), 659-680.

[17] Ismail A. and H. Pham (2017): "Robust Markowitz mean-variance portfolio selection under ambiguous covariance matrix", arXiv:1610.06805, to appear in *Mathematical Finance*.

[18] Jagannathan, R. and T. Ma (2003): "Risk reduction in large portfolios: Why imposing the wrong constraints helps", *The Journal of Finance*, **58**(4), 1651-1683.

[19] Jin H. and X. Y. Zhou (2015): "Continuous-time portfolio selection under ambiguity", *Mathematical Control and Related Fields*, **5**, 475-488.

[20] Ledoit, O. and M. Wolf. (2004): " A well-conditioned estimator for large-dimensional covariance matrices", *Journal of Multivariate Analysis*, **88**(2), 365-411.

[21] Li D. and W.L. Ng (2000): "Optimal dynamic portfolio selection: Multiperiod mean variance formulation", *Mathematical Finance*, **10**(3), 387-406.

[22] Li D. and X.Y. Zhou (2000): "Continuous-time mean-variance portfolio selection: A stochastic LQ framework", *Applied Mathematics and Optimization*, **42**(1), 19-33.

[23] Lin, Q. and F. Riedel (2014): "Optimal consumption and portfolio choice with ambiguity", arXiv preprint arXiv:1401.1639.

[24] Liu J. and X.D. Zeng (2017): "Correlation Ambiguity and Under-Diversification", preprint, available at SSRN: https://ssrn.com/abstract=2692692 or http://dx.doi.org/10.2139/ssrn.2692692.

[25] Markowitz H. (1952): "Portfolio Selection", *Journal of Finance*, **7**, 77-91.

[26] Matoussi A., Possamai D. and C. Zhou (2015): "Robust utility maximization in nondominated models with 2BSDE: the uncertain volatility model", *Mathematical Finance*, **25**(2), 258-287.

[27] Mitton T. and K. Vorkink (2007): "Equilibrium underdiversification and the preference for skewness", *The Review of Financial Studies*, **20**(4), 1255-1288.

[28] Neufeld A. and M. Nutz (2018): "Robust Utility Maximization with Lévy Processes", *Mathematical Finance*, **28**(1), 82-105.

[29] Rapisarda F., Brigo D. and F. Mercurio (2007): "Parameterizing correlations: a geometric interpretation", *IMA Journal of Management Mathematics*, **18**(1), 55-73.

[30] Schied A. (2007): "Optimal investments for risk-and ambiguity-averse preferences: a duality approach", *Finance and Stochastics*, **11**(1), 107-129.

[31] Schultz E. (1996): "Workers put too much in their employer's stock", *Wall Street Journal*, **1**.

[32] Strasser H. (1985): Mathematical theory of Statistics, de Gruyter studies in Mathematics.

[33] Tevzadze R., Toronjadze T. and T. Uzunashvili (2013): "Robust utility maximization for a diffusion market model with misspecified coefficients", *Finance and Stochastics*, **17**(3), 535-563.

[34] Uppal R. and T. Wang (2003): "Model Misspecification and Underdiversification", *The Journal of Finance*, **58**(6), 2465-2486.

[35] Van Nieuwerburgh S. and L. Veldkamp (2010): "Information acquisition and underdiversification", *The Review of Economic Studies*, **77**(2), 779-805.

[36] Wei X. (2018): "McKean-Vlasov control problems and applications", PhD thesis, University Paris Diderot.