Robust Markowitz mean-variance portfolio selection under ambiguous covariance matrix *

Amine Ismail, Huyên Pham

Robust Markowitz mean-variance portfolio selection under ambiguous covariance matrix *

Amine ISMAIL † Huyên PHAM‡

March 11, 2017

Abstract

This paper studies a robust continuous-time Markowitz portfolio selection problem where the model uncertainty carries on the covariance matrix of multiple risky assets. This problem is formulated into a min-max mean-variance problem over a set of non-dominated probability measures that is solved by a McKean-Vlasov dynamic programming approach, which allows us to characterize the solution in terms of a Bellman-Isaacs equation in the Wasserstein space of probability measures. We provide explicit solutions for the optimal robust portfolio strategies and illustrate our results in the case of uncertain volatilities and ambiguous correlation between two risky assets. We then derive the robust efficient frontier in closed-form, and obtain a lower bound for the Sharpe ratio of any robust efficient portfolio strategy. Finally, we compare the performance of Sharpe ratios for a robust investor and for an investor with a misspecified model.

MSC Classification: 91G10, 91G80, 60H30

Key words: Continuous-time Markowitz problem, covariance matrix uncertainty, ambiguous correlation, McKean-Vlasov, dynamic programming, Wasserstein space.

1 Introduction

The Markowitz mean-variance portfolio selection problem [25], initially considered in a single period model, is the cornerstone of modern portfolio allocation theory. Investment decisions rules are made according to the objective of maximizing the expected return for a given financial risk quantified by the variance of the portfolio, and lead to the concept of efficient frontier, which proposes a simple illustration of the trade-off between return and risk. The use of Markowitz efficient portfolio strategies in the financial industry has become quite popular mainly due to its natural and intuitive formulation.

*This work is issued from a CIFRE collaboration between NATIXIS and LPMA. We would like to thank Carmine De Franco, Johan Nicolle and Nizar Touzi for helpful discussions. We are grateful to both referees and the AE for numerous comments, which help to improve the first version [17] of this paper.

†Natixis, Equity Markets, and LPMA,Université Paris-Diderot, ami.ismael@gmail.com

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

In a continuous-time dynamic setting, the mean-variance criterion involves in a nonlinear way the expected terminal wealth due to the variance term, and induces the so-called time inconsistency. This nonstandard feature in stochastic control problem has generated various resolution approaches. A first approach in [38] consists in embedding the mean-variance problem into an auxiliary standard control problem that can be solved by using stochastic linear quadratic theory. A second approach relies on the observation that the dynamic mean-variance problem can be reformulated as a control problem of McKean-Vlasov type, where the cost functional may depend nonlinearly on the law of the wealth state process. It has then been solved in [2] where the authors have derived a version of the Pontryagin maximum principle. More recently, the paper [29] has developed a general dynamic programming approach for the control of McKean-Vlasov dynamics and applied their method for the resolution of the mean-variance portfolio selection problem. We also mention the recent paper [13], where the mean-variance problem is viewed as the McKean-Vlasov limit of a family of controlled many-component weakly interacting systems. These prelimit problems are solved by standard dynamic programming, and the solution to the original problem is obtained by passage to the limit.

In the above cited papers, the continuous-time Markowitz problem was essentially studied in the framework of a Black-Scholes model, and abundant research has been conducted to extend this setup by including models with random parameters. Among this large literature, we cite the recent paper [8] which uses a stochastic correlation model for taking into account the correlation risk between risky assets. In all these works, it is assumed that investors have a perfect knowledge of the stochastic dynamics governing the price process, that is a "correct" model has to be first specified, and then the parameters have to be accurately estimated or calibrated. However, in finance, a model is clearly an approximation of the reality, and moreover within a model, the estimation problem is a difficult issue. For example, it is known that the estimation of correlation between assets may be extremely inaccurate due to asynchronous data and lead-lag effect, especially when the number of assets is large, and the correlation estimate converges to its true value less rapidly than the estimates of volatilities that are based on the full sets of marginal observations, see e.g. [18], [16] and [1]. On the other hand, optimal portfolios are typically sensitive to the model and the parameters, and may perform badly when the parameters are not sufficiently accurate. Therefore, the impact of model misspecification, due to erroneous models and measurements, is an important issue in the practical implementation of trading strategies, and is usually refereed to as model risk.

In order to address the model risk related to uncertainty or ambiguous model parameters, the robust approach, which consists in taking decisions under the worst-case scenario over all conceivable models, is a notable research direction in mathematical finance. A common robust modeling is to consider a family of probability measures representing all the prior beliefs of the investor on the model parameters. For example, drift uncertainty is modeled via Girsanov's theorem by a set of dominated probability measures, and has been first considered in the context of portfolio selection in [15], and then largely studied in the literature, see the recent paper [20] and the references therein.

We focus here on uncertainty or ambiguity on the covariance matrix of the risky assets, assuming that the instantaneous return (drift) is known (or by considering that we have a strong belief on its value). Uncertain volatility models have been considered in [3], [24], or [10] in the context of option pricing, and in [26], [22] for robust portfolio optimization with expected utility criterion. As in [14], we are also interested in a setting with ambiguous correlation between two risky assets since, as already mentioned above, the correlation parameter is hard in practice to infer with accuracy from market information.

In this paper, we investigate the robust Markowitz mean-variance portfolio selection under uncertainty on the volatilities and correlation of multi risky assets. Robust mean-variance problems have been considered in the economic and engineering literature, mostly on single period or multiperiod models, see e.g. [12], [30] and [23]. Here, in our continuous-time modeling, we adopt the probabilistic framework in [11], related to the theory of G-expectation [28] (see also [35]), in order to capture model uncertainty and ambiguity on the covariance matrix, which leads to a set of non-dominated probability measures for the prior probabilities. We also make some concavity assumption on the set of prior covariance matrix. From a mathematical viewpoint, and compared to robust problem with expected utility, we face two additional difficulties: (i) it cannot be tackled a priori by classical stochastic differential game approach due to the nonlinear variance term, (ii) moreover, since the worst-case scenario is not the same for the mean and the variance, it is not straightforward that it can be put into a min-max problem. We then use the following methodology. We consider a robust mean-variance criterion, which is actually formulated as a min-max problem, and show a posteriori how it is connected to the robust Markowitz problem. We tackle the former problem by a McKean-Vlasov dynamic programming approach: we first reformulate the robust mean-variance problem into a deterministic differential game problem with the law of the wealth process under a prior probability measure as state variable. Then, adapting optimality arguments from dynamic programming principle, and using recent chain rule for flow of probability measures derived in [5] and [7], we state a verification theorem which gives the optimal strategy and performance in terms of a Bellman-Isaacs equation in the Wasserstein space of probability measures. We next apply this analytic partial differential equation characterization of the solution to the robust mean-variance problem, and show that the problem can be reduced into two steps: first, we determine the worst-case scenario, and the remarkable point is that it corresponds to a constant variance/covariance matrix obtained by the minimization of the risk premium, which is a direct input of the model. Secondly, we obtain the optimal mean-variance strategy as in the Black-Scholes model with the known instantaneous return and the worst-case constant covariance matrix. We illustrate our results with closed-form expressions for the optimal portfolio strategies in two examples: uncertain volatilities and ambiguous correlation between two risky assets. Moreover, we are able to derive explicitly the corresponding robust efficient frontier of the robust Markowitz problem. In particular, we obtain a lower bound for the Sharpe ratio of any robust efficient portfolio strategy, which is independent of any modelling on the covariance matrix.

How can robust mean-variance portfolio strategies help to improve performance of investors? We address this question by using simulations to evaluate and compare the Sharpe ratio of a robust investor and a simple investor who implements mean-variance strategies with a misspecified model in two examples: (i) in the first example, the true dynamics of the stock price is assumed to be governed by a Heston type stochastic volatility model that makes the volatility bounded, and the simple investor considers that the risky asset is governed by a Black-Scholes model with constant volatility, (ii) in the second example, the two-assets price is given in reality by a stochastic correlation model, but the simple investor considers a constant correlation between the risky assets. Our results show that the robust Sharpe ratio can perform noticeably better than the misspecified Sharpe ratio for some choice of the parameters describing the true dynamics.

The rest of the paper is organized as follows. Section 2 formulates the probabilistic framework for the robust Markowitz mean-variance problem. We present in Section 3 the McKean-Vlasov dynamic programming approach for solving our problem. In Section 4, we derive explicit solutions in the context of ambiguous covariance matrix including uncertain volatilities and ambiguous correlation. Section 5 is devoted to the derivation of the robust efficient frontier in closed form, and the last Section 6 discusses the benefit of a robust investor compared to a misspecified investor.

2 Problem formulation

We consider a financial market with one risk-free asset, assumed to be constant equal to one (zero interest rate), and d risky stocks on a finite investment horizon $[0, T]$. We model the uncertainty about the volatility matrix of the risky assets by using the probabilistic setup as in [10], [28] or [35]. We define the canonical state space by $\Omega = \{\omega = (\omega(t))_{t\in[0,T]} \in C([0,T];\mathbb{R}^n) : \omega(0) = 0\}$ representing the continuous paths driving d risky assets, and possibly m (non tradable) factor processes ($n = d + m$), by $\mathcal{F}$ its Borel σ-field, and denote by $\bar{B} = (\bar{B}_t)_{t\in[0,T]}$ the canonical process, i.e. $\bar{B}_t(\omega) = \omega(t)$, by $\mathbb{P}_0$ the Wiener measure, i.e. making $\bar{B}$ a n-dimensional Brownian motion under $\mathbb{P}_0$, and by $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ the canonical filtration, i.e. the natural filtration generated by $\bar{B}$. We distinguish the d-dimensional components of $\bar{B}$, denoted by B, and representing the continuous paths of the risky assets, and the other $(n-d)$-dimensional components are denoted by $\check{B}$.

The investor knows (or has estimated) the constant drift $b = (b_1, \dots, b_d) \in \mathbb{R}^d$ of the assets, but is uncertain about the volatility matrix (possibly random) of the d risky assets. We adopt the concept of ambiguous volatility as defined in [11], which means that the investor only knows that the covariance matrix belongs to some prior compact set Γ of $\mathbb{S}^d_{>+}$, the set of strictly positive definite matrices in $\mathbb{R}^{d\times d}$. We assume that $\Gamma = \Gamma(\Theta)$ is parametrized by a prior convex set Θ of $\mathbb{R}^q$, that is there exists some measurable function $\gamma : \mathbb{R}^q \to \mathbb{S}^d_{>+}$ s.t. any $\Sigma \in \Gamma$ is in the form $\Sigma = \gamma(\theta)$ for some $\theta \in \Theta$. For any $\Sigma \in \Gamma$, we denote by $\sigma = \Sigma^{\frac{1}{2}}$ its square-root matrix, and we shall often identify a covariance matrix with its square-root matrix called volatility matrix. Here are some examples of this modeling:

Example 1 (uncertain volatilities). In dimension $d = 1$, this is modelled through $\Gamma = \Theta = [\underline{\sigma}^2, \bar{\sigma}^2]$ with positive constants $0 < \underline{\sigma} \leq \bar{\sigma} < \infty$, see [3], [24]. The extension to the multivariate assets case with zero correlation is modelled through $\Theta = \prod_{i=1}^{d}[\underline{\sigma_i}^2, \bar{\sigma}_i^2]$ with $0 < \underline{\sigma_i} \leq \bar{\sigma}_i < \infty$, $i = 1, \dots, d$, and

$$\gamma(\theta) \;=\; \begin{pmatrix} \sigma_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_d^2 \end{pmatrix}, \quad \text{for } \; \theta = (\sigma_1^2, \dots, \sigma_d^2).$$

Example 2 (ambiguous correlation). The uncertainty about the correlation between risky assets in dimension $d = 2$ has been recently considered in [14], and can be formalized here with $\Theta = [\underline{\varrho}, \bar{\varrho}] \subset (-1, 1)$, and

$$\gamma(\theta) \;=\; \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\theta \\ \sigma_1\sigma_2\theta & \sigma_2^2 \end{pmatrix},$$

for some known positive constants σ_1 and σ_2 representing the marginal volatilities of the assets, and where θ represents the unknown correlation parameter varying between $\underline{\varrho}$ and $\bar{\varrho}$. The extension to multivariate assets for $d \geq 2$ can also be done within our framework with a parametric form for the correlation matrix using for instance $d(d-1)/2$ angular coordinates as in [31].

We denote by $\mathcal{V}_\Theta$ the set of $\mathbb{F}$-progressively measurable processes $\Sigma = (\Sigma_t)$ valued in $\Gamma = \Gamma(\Theta)$, and introduce the set of prior probability measures $\mathcal{P}^\Theta$:

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

where $\mathbb{P}^\sigma$ is the probability measure on $(\Omega, \mathcal{F}_T)$ induced by $\mathbb{P}_0$ via:

$$\mathbb{P}^\sigma \quad := \quad \mathbb{P}_0 \circ (\bar{B}^\sigma)^{-1}, \quad \text{ with } \sigma_t := \Sigma_t^{\frac{1}{2}},\ B_t^\sigma := \int_0^t \sigma_s dB_s,\ \ 0 \leq t \leq T,\ \mathbb{P}_0 - a.s.,$$

and $\bar{B}^\sigma$ is the $\mathbb{R}^n$-valued process on Ω defined by $\bar{B}^\sigma := (B^\sigma\ \check{B})$.

Under any $\mathbb{P}^\sigma$, $\Sigma \in \mathcal{V}_\Theta$, the process B is a martingale, hence admits from [21], a quadratic variation, which is given by:

$$d < B >_t \quad = \quad \Sigma_t dt.$$

Remark 2.1 Ambiguity in volatility leads to a set of prior probabilities in $\mathcal{P}^\Theta$, which are non-equivalent, actually mutually singular. Such a specification for the set of prior probabilities $\mathbb{P}^\sigma$ is closely connected to the theory of G-Brownian motion introduced in [28], and requires tools from quasi-sure analysis as pointed out in [10], and further studied in [35]. In particular, we say that a property holds $\mathcal{P}^\Theta$-quasi surely ($\mathcal{P}^\Theta - q.s.$ in short), if it holds $\mathbb{P}^\sigma - a.s.$ for all $\mathbb{P}^\sigma \in \mathcal{P}^\Theta$. □

The price process S of the d risky assets is given by

$$dS_t \quad = \quad \text{diag}(S_t)\big(bdt + dB_t\big), \quad 0 \leq t \leq T,\ \ \mathcal{P}^\Theta - q.s.$$

Remark 2.2 Under each $\mathbb{P}^\sigma \in \mathcal{P}^\Theta$, for $\Sigma \in \mathcal{V}_\Theta$, we have $dB_t = \sigma_t dW_t^\sigma$ where W^σ is a Brownian motion under $\mathbb{P}^\sigma$, and so the price process is governed under $\mathbb{P}^\sigma$ by

$$dS_t \quad = \quad \text{diag}(S_t)\big(bdt + \sigma_t dW_t^\sigma\big), \quad 0 \leq t \leq T,\ \ \mathbb{P}^\sigma - a.s.$$

□

A portfolio strategy $\alpha = (\alpha_t)_{0 \leq t \leq T}$, representing the amount invested in the d risky assets, is a d-dimensional $\mathbb{F}$-progressively measurable process, valued in some closed convex set A of $\mathbb{R}^d$, satisfying the integrability condition

$$\sup_{\mathbb{P}^\sigma \in \mathcal{P}^\Theta} \mathbb{E}_\sigma\Big[\int_0^T \alpha_t^\intercal \Sigma_t \alpha_t dt\Big] \quad < \quad \infty, \tag{2.1}$$

and denoted by $\alpha \in \mathcal{A}$. Here $^\intercal$ denotes the transpose of a matrix, and $\mathbb{E}_\sigma$ denotes the expectation under $\mathbb{P}^\sigma$. Given a portfolio strategy $\alpha \in \mathcal{A}$, and an initial capital $x_0 \in \mathbb{R}$, the evolution of the self-financing wealth process X^α is given by

$$dX_t^\alpha \quad = \quad \alpha_t^\intercal \text{diag}(S_t)^{-1} dS_t \ = \ \alpha_t^\intercal\big(bdt + dB_t\big), \quad 0 \leq t \leq T,\ X_0^\alpha = x_0,\ \mathcal{P}^\Theta - q.s. \tag{2.2}$$

Remark 2.3 Given $\alpha \in \mathcal{A}$, the existence of a $\mathcal{P}^{\Theta}$-quasi surely aggregated solution to (2.2) is ensured by Theorem 2.2 in [27] under the Zermelo Fraenkel set theory with axiom of choice (ZFC) plus the Continuum Hypothesis. Moreover, for $\alpha \in \mathcal{A}$, and from Remark 2.2, we see that the evolution of X^α under any $\mathbb{P}^\sigma \in \mathcal{P}^\Theta$, $\Sigma \in \mathcal{V}_\theta$, is given by

$$dX_t^\alpha \;=\; \alpha_t^\intercal(bdt + \sigma_t dW_t^\sigma), \quad 0 \leq t \leq T,\; X_0^\alpha = x_0,\; \mathbb{P}^\sigma - a.s. \tag{2.3}$$

where W^σ is a Brownian motion under $\mathbb{P}^\sigma$, and we have

$$\sup_{\mathbb{P}^\sigma \in \mathcal{P}^\Theta} \mathbb{E}_\sigma\big[\sup_{0\leq t\leq T} |X_t^\alpha|^2\big] \;<\; \infty.$$

$\square$

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

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

where $\mathrm{Var}_\sigma(X)$ denotes the variance of X under $\mathbb{P}^\sigma$, and the robust mean-variance portfolio selection problem is then formulated as

$$V_0 \;=\; \inf_{\alpha\in\mathcal{A}} J_{wc}(\alpha) \;=\; \inf_{\alpha\in\mathcal{A}} \sup_{\mathbb{P}^\sigma \in \mathcal{P}^\Theta} \Big(\lambda \mathrm{Var}_\sigma(X_T^\alpha) - \mathbb{E}_\sigma[X_T^\alpha]\Big). \tag{2.4}$$

A related problem to the robust mean-variance portfolio selection problem is the robust Markowitz problem, which is formulated as follows: given a variance risk $\vartheta > 0$,

$$\begin{cases} \text{maximize over } \alpha \in \mathcal{A}, & \mathcal{E}(\alpha) \;:=\; \inf_{\mathbb{P}^\sigma \in \mathcal{P}^\Theta} \mathbb{E}_\sigma[X_T^\alpha] \\ \qquad\qquad \text{subject to} & \mathcal{R}(\alpha) \;:=\; \sup_{\mathbb{P}^\sigma \in \mathcal{P}^\Theta} \mathrm{Var}_\sigma(X_T^\alpha) \;\leq\; \vartheta. \end{cases} \tag{2.5}$$

A solution $\hat\alpha^\vartheta$ to (2.5), when it exists, is called *robust efficient portfolio strategy* with respect to ϑ. In other words, a robust efficient portfolio strategy maximizes the worst case expected terminal wealth given a financial risk measured by the worst case variance of the terminal wealth. The pair $(\mathcal{R}(\hat\alpha^\vartheta), \mathcal{E}(\hat\alpha^\vartheta))$ is called a robust efficient point, and the set of all robust efficient points, when varying ϑ, is called *robust efficient frontier*. By standard convex optimization theory, the constrained optimization problem (2.5) is connected by duality to the Lagrangian optimization problem, which is defined as

$$\inf_{\alpha\in\mathcal{A}} \big[\lambda\mathcal{R}(\alpha) - \mathcal{E}(\alpha)\big] \;=\; \inf_{\alpha\in\mathcal{A}} \Big\{\lambda \sup_{\mathbb{P}^\sigma \in \mathcal{P}^\Theta} \mathrm{Var}_\sigma(X_T^\alpha) - \inf_{\mathbb{P}^\sigma \in \mathcal{P}^\Theta} \mathbb{E}_\sigma[X_T^\alpha]\Big\}.$$

Notice that this Lagrangian optimization problem is equal to problem (2.4) when $\mathcal{P}^\Theta$ is a singleton, but differs a priori from (2.4). We shall solve in the two next sections the robust mean-variance portfolio selection problem (2.4), and show in the last section that it is actually equal by duality to the Lagrangian optimization problem, and so leads to the solution of the robust Markowitz problem (2.5) and the construction of the robust efficient frontier.

3 McKean-Vlasov approach

Problem (2.4) can be viewed as a zero-sum stochastic differential game problem with gain/cost functional

$$J(\alpha,\sigma) \;=\; \lambda \mathrm{Var}_\sigma(X_T^\alpha) - \mathbb{E}_\sigma[X_T^\alpha], \quad \alpha \in \mathcal{A}, \Sigma \in \mathcal{V}_\Theta, \tag{3.1}$$

so that $V_0 = \inf_{\alpha\in\mathcal{A}} \sup_{\Sigma\in\mathcal{V}_\Theta} J(\alpha,\sigma)$. The peculiarity of this differential game problem is the nonlinear dependence of the law of the state process via the variance term, making the problem a priori time inconsistent. Following the idea in [4] and [29] for control problem, we first reformulate our problem into a deterministic differential game problem, taking into account the uncertainty about the probability law governing the risky asset. For any $\alpha \in \mathcal{A}$, and $t \in [0,T]$, let us denote by $\rho_t^{\alpha,\sigma} = \mathbb{P}^\sigma_{X_t^\alpha}$ the law of X_t^α under $\mathbb{P}^\sigma$, $\Sigma \in \mathcal{V}_\Theta$, which defines a deterministic process valued in the Wasserstein space $\mathcal{P}_2(\mathbb{R})$ of square-integrable probability measures on $\mathbb{R}$, which is a metric space when equipped with the Wasserstein distance $\mathcal{W}_2$:

$$\mathcal{W}_2(\mu,\mu') \;=\; \inf\Big\{\Big(\int_{\mathbb{R}\times\mathbb{R}} |x-y|^2 \pi(dx,dy)\Big)^{\frac{1}{2}} : \pi \in \mathcal{P}_2(\mathbb{R}\times\mathbb{R}) \text{ with marginals } \mu \text{ and } \mu'\Big\}$$

We also set $\|\mu\|_2 := \mathcal{W}_2(\mu,\delta_0) = \big(\int |x|^2\mu(dx)\big)^{\frac{1}{2}}$.

We also introduce the following convenient notations: for any $\mu \in \mathcal{P}_2(\mathbb{R})$, we denote by

$$\bar\mu \;:=\; \int_{\mathbb{R}} x\mu(dx), \qquad \mathrm{Var}(\mu) \;:=\; \int_{\mathbb{R}} (x-\bar\mu)^2\mu(dx).$$

We can then rewrite the functional in (3.1) and the worst-case mean-variance functional as

$$J_{wc}(\alpha) \;=\; \sup_{\Sigma\in\mathcal{V}_\Theta} J(\alpha,\sigma) \;=\; \sup_{\Sigma\in\mathcal{V}_\Theta} \big[\lambda \mathrm{Var}(\rho_T^{\alpha,\sigma}) - \overline{\rho_T^{\alpha,\sigma}}\big], \quad \alpha \in \mathcal{A}. \tag{3.2}$$

The robust mean-variance portfolio selection problem is therefore reformulated as a deterministic differential game problem with controlled state variable $\rho^{\alpha,\sigma}$ valued in the infinite-dimensional space $\mathcal{P}_2(\mathbb{R})$. To solve this problem, we use general dynamic programming optimality principle, which takes the following formulation in our context:

Optimality principle

Let $\{V^{\alpha,\sigma}, \alpha \in \mathcal{A}, \Sigma \in \mathcal{V}_\Theta\}$ be a family of deterministic processes in the form $V_t^{\alpha,\sigma} = v(t,\rho_t^{\alpha,\sigma})$ for some real-valued measurable function v on $[0,T]\times\mathcal{P}_2(\mathbb{R})$ satisfying

(i) $v(T,\mu) = \lambda \mathrm{Var}(\mu) - \bar\mu$, for any $\mu \in \mathcal{P}_2(\mathbb{R})$

(ii) $t \in [0,T] \longmapsto \sup_{\Sigma\in\mathcal{V}_\Theta} V_t^{\alpha,\sigma}$ is nondecreasing for all $\alpha \in \mathcal{A}$

(iii) $t \in [0,T] \longmapsto \sup_{\Sigma\in\mathcal{V}_\Theta} V_t^{\alpha^*,\sigma}$ is nonincreasing (hence constant) for some $\alpha^* \in \mathcal{A}$.

Then, α^* is an optimal control for the robust mean-variance problem (2.4) with optimal value

$$V_0 \;=\; v(0,\delta_{x_0}) \;=\; J_{wc}(\alpha^*). \tag{3.3}$$

Indeed, observe that at time $t = 0$, $\rho_0^{\alpha,\sigma} = \delta_{x_0}$ for any $\alpha \in \mathcal{A}, \Sigma \in \mathcal{V}_\Theta$, since X_0^α is equal to the constant x_0, which implies that $V_0^{\alpha,\sigma} = v(0, \delta_{x_0})$ does not depend on $\alpha \in \mathcal{A}, \Sigma \in \mathcal{V}_\Theta$. From properties (i) and (ii), we then have for all $\alpha \in \mathcal{A}$,

$$v(0, \delta_{x_0}) \;=\; \sup_{\Sigma \in \mathcal{V}_\Theta} V_0^{\alpha,\sigma} \;\leq\; \sup_{\Sigma \in \mathcal{V}_\Theta} V_T^{\alpha,\sigma} \;=\; \sup_{\Sigma \in \mathcal{V}_\Theta} v(T, \rho_T^{\alpha,\sigma}) \;=\; \sup_{\Sigma \in \mathcal{V}_\Theta} J(\alpha, \sigma) \;=\; J_{wc}(\alpha),$$

by (3.2). Since α is arbitrary in $\mathcal{A}$, this gives: $v(0, \delta_{x_0}) \leq \inf_{\alpha \in \mathcal{A}} J_{wc}(\alpha) = V_0$. Similarly, from properties (i) and (iii), we obtain $v(0, \delta_{x_0}) = \sup_{\Sigma \in \mathcal{V}_\Theta} J(\alpha^*, \sigma) = J_{wc}(\alpha^*) \geq V_0$, which proves (3.3).

In order to construct a process $V_t^{\alpha,\sigma} = v(t, \rho_t^{\alpha,\sigma})$ satisfying the above conditions (i), (ii), (iii) for the optimality principle, we shall rely on the recent notion of derivatives in the Wasserstein space introduced by P.L. Lions, and the corresponding chain rule (Itô's formula) for flow of probability measures, that we recall in the appendix. The derivative (when it exists) of a function $\varphi(\mu)$ on $\mathcal{P}_2(\mathbb{R})$ is denoted by $\partial_\mu \varphi(\mu)$, and is a function from $\mathbb{R}$ into $\mathbb{R}$, which is in $L^2(\mu)$, and when a version of the function $x \in \mathbb{R} \mapsto \partial_\mu \varphi(\mu)(x)$ is differentiable, we denote by $\partial_x \partial_\mu \varphi(\mu)(x)$ its derivative. Assuming that $v(t, \mu)$ is smooth on $[0, T] \times \mathcal{P}_2(\mathbb{R})$, i.e. continuously differentiable w.r.t. to t, and partially $\mathcal{C}^2$ w.r.t. μ (see Appendix B), we have by Itô's formula (A.2) (recalling (2.3)):

$$dV_t^{\alpha,\sigma} \;=\; dv(t, \rho_t^{\alpha,\sigma}) \;=\; D_t^{\alpha,\sigma} dt, \tag{3.4}$$

where

$$D_t^{\alpha,\sigma} \;=\; \partial_t v(t, \rho_t^{\alpha,\sigma}) + \mathbb{E}_\sigma\big[H(\partial_\mu v(t, \rho_t^{\alpha,\sigma})(X_t^\alpha), \partial_x \partial_\mu v(t, \rho_t^{\alpha,\sigma})(X_t^\alpha), \alpha_t, \Sigma_t)\big], \tag{3.5}$$

with H the function defined on $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}^d_{>+}$ by

$$H(p, M, a, \Sigma) \;=\; p a^\intercal b + \frac{1}{2} M a^\intercal \Sigma a. \tag{3.6}$$

We state some easy properties for the function H, which allows us to introduce some useful notations.

Lemma 3.1 For all $(p, M) \in \mathbb{R} \times (0, \infty)$, $a \in A$, we have

$$\sup_{\Sigma \in \Gamma} H(p, M, a, \Sigma) \;=\; H(p, M, a, \hat{\Sigma}(a)) \;<\; \infty, \qquad \text{with } \hat{\Sigma}(a) \in \arg\max_{\Sigma \in \Gamma} a^\intercal \Sigma a.$$

There exists a measurable function $(p, M) \in \mathbb{R} \times (0, \infty) \mapsto a^*(p, M) \in A$ such that

$$H^*(p, M) \;:=\; \inf_{a \in A} \sup_{\Sigma \in \Gamma} H(p, M, a, \Sigma) \;=\; \sup_{\Sigma \in \Gamma} H(p, M, a^*(p, M), \Sigma). \tag{3.7}$$

Proof. For fixed $(p, M) \in \mathbb{R} \times (0, \infty)$, $a \in A$, it is clear that the continuous function $\Sigma \mapsto H(p, M, a, \Sigma)$ attains its maximum on the compact set Γ at some point $\hat{\Sigma}(a)$ given by $\hat{\Sigma}(a) \in \arg\max_{\Sigma \in \Gamma} a^\intercal \Sigma a$, from the expression of H, hence not depending on (p, M). By convexity of the function $a \mapsto |a|^2$, it is clear that the function $a \in A \mapsto \bar{H}(p, M, a) := \sup_{\Sigma \in \Gamma} H(p, M, a, \Sigma)$ is also convex. Moreover, since $\bar{H}(p, M, a) \geq p a^\intercal b + \frac{1}{2} M a^\intercal \underline{\Sigma} a$, with $\underline{\Sigma}$ positive definite, we see that $\bar{H}(p, M, a)$ goes to infinity when $|a|$ goes to infinity. It follows that $a \mapsto \bar{H}(p, M, a)$ attains its infimum on the closed convex set A at some $a^*(p, M)$ which can be chosen measurable by continuity of H and Carathéodory-type measurable selection theorem, see e.g. [37]. $\square$

We can now state an analytic verification theorem for the robust mean-variance portfolio selection problem, which provides a characterization of the optimal portfolio strategy.

Theorem 3.1 (Verification theorem)
Let v be a smooth function on $[0,T]\times\mathcal{P}_2(\mathbb{R})$ satisfying $\partial_x\partial_\mu v(t,\mu)(x) > 0$ for all $(t,x,\mu)\in[0,T)\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R})$, and suppose that v is solution to the Bellman-Isaacs partial differential equation (PDE):

$$\begin{cases} \partial_t v(t,\mu) + \displaystyle\int_{\mathbb{R}} H^*(\partial_\mu v(t,\mu)(x), \partial_x\partial_\mu v(t,\mu)(x))\mu(dx) &= 0, \quad (t,\mu)\in[0,T)\times\mathcal{P}_2(\mathbb{R}) \\ v(T,\mu) &= \lambda \mathrm{Var}(\mu) - \bar{\mu}, \quad \mu\in\mathcal{P}_2(\mathbb{R}), \end{cases} \tag{3.8}$$

s.t. the function $(x,\mu)\in\mathbb{R}\times\mathcal{P}_2(\mathbb{R})\mapsto \hat{a}(t,x,\mu) := a^*(\partial_\mu v(t,\mu)(x), \partial_x\partial_\mu v(t,\mu)(x))$ is Lipschitz, for any $t\in[0,T]$, and $\int_0^T|\hat{a}(t,0,\delta_0)|^2dt < \infty$. For any $\Sigma\in\mathcal{V}_\Theta$, denote by $X^{\mathbb{P}^\sigma}$ the solution to the McKean-Vlasov SDE under $\mathbb{P}^\sigma$:

$$dX_t \;=\; \hat{a}(t,X_t,\mathbb{P}^\sigma_{X_t})[bdt + \sigma_t dW_t^\sigma], \quad 0\le t\le T,\ X_0 = x_0,\ \mathbb{P}^\sigma-\text{p.s.} \tag{3.9}$$

and suppose that the family of processes $\{X^{\mathbb{P}^\sigma}, \Sigma\in\mathcal{V}_\Theta\}$ can be aggregated into a $\mathcal{P}^\Theta$-quasi surely aggregated solution, i.e. there exists X^* s.t.

$$X_t^* \;=\; X_t^{\mathbb{P}^\sigma}, \quad 0\le t\le T,\ \mathbb{P}^\sigma-\text{p.s.},\ \forall\Sigma\in\mathcal{V}_\Theta.$$

Then, the family of processes $\{\hat{a}(t,X_t^{\mathbb{P}^\sigma},\mathbb{P}^\sigma_{X_t^{\mathbb{P}^\sigma}}), 0\le t\le T, \Sigma\in\mathcal{V}_\Theta\}$ can also be aggregated, i.e. there exists a process α^* s.t.

$$\alpha_t^* \;=\; \hat{a}(t,X_t^{\mathbb{P}^\sigma},\mathbb{P}^\sigma_{X_t^{\mathbb{P}^\sigma}}) \quad 0\le t\le T,\ \mathbb{P}^\sigma-\text{p.s.},\ \forall\Sigma\in\mathcal{V}_\Theta, \tag{3.10}$$

and the process α^* defines a portfolio strategy in $\mathcal{A}$, which is optimal for (2.4), i.e. $V_0 = J_{wc}(\alpha^*)$, and we have $V_0 = v(0,\delta_{x_0})$.

Remark 3.1 **1.** In standard stochastic control problem where the criterion involves linear functional of the law of the state process, we look for a value function $v(t,\mu)$, which is also linear in μ, hence of the form $v(t,\mu) = \int w(t,x)\mu(dx)$ for some smooth function w on $[0,T]\times\mathbb{R}$ solution to the standard Hamilton-Jacobi-Bellman-Isaacs equation. In this case, $\partial_x\partial_\mu v(t,\mu)(x) = D_x^2 w(t,x)$, and the above condition in the verification theorem: $\partial_x\partial_\mu v(t,\mu)(x) > 0$ for all (t,x,μ), simply means that we look for a convex function w, which usually follows from the convexity of the terminal cost and the linear dynamics of the wealth process. Here for the mean-variance criterion, the condition $\partial_x\partial_\mu v(t,\mu)(x) > 0$ is related to the positivity of the variance penalization parameter λ, see (4.15).
2. For fixed $\Sigma\in\mathcal{V}_\Theta$, the existence and uniqueness of a $\mathbb{P}^\sigma$-solution $X^{\mathbb{P}^\sigma}$ to the McKean-Vlasov SDE (3.9) under the Lipschitz condition on $\hat{a}$ and the square-integrability condition of $\hat{a}(.,0,\delta_0)$ follows from standard arguments (recall that $\Sigma\in\mathcal{V}_\Theta$ is bounded) as in [33] or [19], and we have the estimate:

$$\mathbb{E}_\sigma\big[\sup_{0\le t\le T}|X_t^{\mathbb{P}^\sigma}|^2\big] \;\le\; C\Big(1+\int_0^T|\hat{a}(t,0,\rho_t^*)|^2dt\Big) \;<\; \infty, \tag{3.11}$$

for some positive constant C depending on the Lipschitz condition on the function $x\mapsto\hat{a}(t,x,\rho_t^*)$, and independent of Σ. The key assumption in the above verification theorem is the fact one can aggregate the family of processes $\{X^{\mathbb{P}^\sigma},\Sigma\in\mathcal{V}_\Theta\}$ in order to define a universal process X^* defined $\mathcal{P}^\Theta$-quasi surely. This point is discussed more precisely in the next section, where it is shown that the aggregation condition is satisfied when prior probability measures are related to uncertainty on covariance matrix (see Theorem 4.1), but not in general on drift uncertainty (see Remark 4.3). Once this aggregation condition is satisfied, we notice that the i-th component of the $\mathbb{R}^d$-valued process $\{\hat{a}(t, X_t^{\mathbb{P}^\sigma}, \mathbb{P}^\sigma_{X_t^{\mathbb{P}^\sigma}}), 0 \leq t \leq T\}$ is obtained as the Radon-Nikodym derivative

$$\hat{a}^i(t, X_t^{\mathbb{P}^\sigma}, \mathbb{P}^\sigma_{X_t^{\mathbb{P}^\sigma}}) \;=\; \frac{d < X^*, B^i >_t}{d < B^i >_t}, \quad 0 \leq t \leq T, \ \mathbb{P}^\sigma - \text{p.s.}, \ \forall \Sigma \in \mathcal{V}_\Theta,$$

where $< X^*, B^i >$ is the quadratic covariation (covariance) process associated to X^* and B^i, which is defined $\mathcal{P}^\Theta$-quasi surely. Therefore, the family of processes $\{\hat{a}(t, X_t^{\mathbb{P}^\sigma}, \mathbb{P}^\sigma_{X_t^{\mathbb{P}^\sigma}}), 0 \leq t \leq T, \Sigma \in \mathcal{V}_\Theta\}$ can be aggregated into α^* as in (3.10), and we easily see from (3.11) that α^* satisfies the integrability condition (2.1), hence lies in $\mathcal{A}$. By construction, we then see that $X^* = X^{\alpha^*}$ the associated (self-financing) wealth process, and the remaining point in the verification theorem is to check that α^* is optimal, as proved below. □

Proof of Theorem 3.1. It suffices to check that the family of (deterministic) processes $V_t^{\alpha,\sigma} = v(t, \rho_t^{\alpha,\sigma})$, $0 \leq t \leq T$, with v solution to the PDE (3.8), satisfies the conditions of the optimality principle with α^*. Condition (i) is already satisfied and in view of (3.4), it suffices to check that (ii) for all $\alpha \in \mathcal{A}$, there exists $\bar{\Sigma}$ depending on $\alpha \in \mathcal{V}_\theta$ s.t. $D_t^{\alpha,\bar{\sigma}} \geq 0$, $0 \leq t \leq T$, and (iii) $D_t^{\alpha^*,\sigma} \leq 0$, $0 \leq t \leq T$, for all $\Sigma \in \mathcal{V}_\Theta$, hold true. Given $\alpha \in \mathcal{A}$, consider the process $\bar{\Sigma} \in \mathcal{V}_\Theta$ defined by $\bar{\Sigma}_t = \hat{\Sigma}(\alpha_t)$, $0 \leq t \leq T$, where $\hat{\Sigma}(.)$ is defined in Lemma 3.1. Recalling the expression of $D^{\alpha,\bar{\sigma}}$ in (3.5), we have for all $t \in [0,T]$,

$$\begin{aligned} D_t^{\alpha,\bar{\sigma}} &= \mathbb{E}_{\bar{\sigma}}\big[\partial_t v(t, \rho_t^{\alpha,\bar{\sigma}}) + H(\partial_\mu v(t, \rho_t^{\alpha,\bar{\sigma}})(X_t^\alpha), \partial_x\partial_\mu v(t, \rho_t^\alpha)(X_t^\alpha), \alpha_t, \hat{\Sigma}(\alpha_t))\big] \\ &= \mathbb{E}_{\bar{\sigma}}\big[\partial_t v(t, \rho_t^{\alpha,\bar{\sigma}}) + \sup_{\gamma\in\Gamma} H(\partial_\mu v(t, \rho_t^{\alpha,\bar{\sigma}})(X_t^\alpha), \partial_x\partial_\mu v(t, \rho_t^{\alpha,\bar{\sigma}})(X_t^\alpha), \alpha_t, \gamma)\big] \\ &\geq \mathbb{E}_{\bar{\sigma}}\big[\partial_t v(t, \rho_t^{\alpha,\bar{\sigma}}) + H^*(\partial_\mu v(t, \rho_t^{\alpha,\bar{\sigma}})(X_t^\alpha), \partial_x\partial_\mu v(t, \rho_t^{\alpha,\bar{\sigma}})(X_t^\alpha))\big] \;=\; 0, \end{aligned}$$

where the second equality comes from the definition of $\bar{\Sigma}_t = \hat{\Sigma}(\alpha_t)$, the inequality $\geq$ from the fact that $H^*(p, M) \leq \sup_{\gamma\in\Gamma} H(p, M, a, \gamma)$ for all $a \in A$, and the last equality $= 0$ from the PDE (3.8) satisfied by v at point $(t, \rho_t^{\alpha,\bar{\sigma}})$ and recalling that $\rho^{\alpha,\bar{\sigma}}$ is the law of X_t^α under $\mathbb{P}^{\bar{\sigma}}$. This proves the condition (ii). On the other hand, let us consider the universal process $\alpha^* \in \mathcal{A}$ defined in (3.10). We then have for all $\Sigma \in \mathcal{V}_\Theta$, and $t \in [0,T]$,

$$\begin{aligned} D_t^{\alpha^*,\sigma} &= \mathbb{E}_\sigma\big[\partial_t v(t, \rho_t^{\alpha^*,\sigma}) + H(\partial_\mu v(t, \rho_t^{\alpha^*,\sigma})(X_t^*), \partial_x\partial_\mu v(t, \rho_t^{\alpha^*,\sigma})(X_t^*), \alpha_t^*, \Sigma_t)\big] \\ &\leq \mathbb{E}_\sigma\big[\partial_t v(t, \rho_t^{\alpha^*,\sigma}) + \sup_{\gamma\in\Gamma} H(\partial_\mu v(t, \rho_t^{\alpha^*,\sigma})(X_t^*), \partial_x\partial_\mu v(t, \rho_t^{\alpha^*,\sigma})(X_t^*), \alpha_t^*, \gamma)\big] \\ &= \mathbb{E}_\sigma\big[\partial_t v(t, \rho_t^{\alpha^*,\sigma}) + H^*(\partial_\mu v(t, \rho_t^{\alpha^*,\sigma})(X_t^*), \partial_x\partial_\mu v(t, \rho_t^{\alpha^*,\sigma})(X_t^*))\big] \;=\; 0, \end{aligned}$$

where the second equality follows from the definition of α^* and relation (3.7). This proves condition (iii), and ends the proof of this theorem. □

4 Explicit solutions

We provide in this section explicit solutions to the Bellman-Isaacs PDE (3.8), hence to the robust mean-variance portfolio selection problem (2.4), when $A = \mathbb{R}^d$, and for a class of prior models Γ on the covariance matrix satisfying a concavity assumption. Recall our parametrization of the covariance matrix: there is some convex set $\Theta \subset \mathbb{R}^q$, and a measurable function $\gamma : \mathbb{R}^q \to \mathbb{S}^d_{>+}$ s.t. any $\Sigma \in \Gamma = \Gamma(\Theta)$ is in the form $\Sigma = \gamma(\theta)$ for some θ in Θ. We shall assume that

(IC) $A = \mathbb{R}^d$ and $\gamma : \mathbb{R}^q \to \mathbb{S}^d_{>+}$ is concave[1] on Θ, i.e. for all θ_1, $\theta_2 \in \Theta$,

$$\frac{1}{2}\big(\gamma(\theta_1) + \gamma(\theta_2)\big) \quad \preceq \quad \gamma\big(\frac{1}{2}(\theta_1 + \theta_2)\big). \tag{4.1}$$

Notice that this assumption is trivially satisfied in the Examples 1 and 2 of uncertain volatilities and ambiguous correlation detailed in Section 2 where we have actually equality in (4.1).

Let us denote by R the (square) risk premium function:

$$R(\theta) \quad := \quad b^\intercal \gamma(\theta)^{-1} b, \quad \theta \in \Theta. \tag{4.2}$$

The next Lemma provides a key result on the Hamiltonian function H in (3.6), which will be useful for the elucidation of our problem.

Lemma 4.1 Let condition **(IC)** hold. Then, for all $p \in \mathbb{R}$, $M > 0$, we have

$$\begin{aligned} H^*(p, M) \quad &= \quad -\frac{1}{2}\frac{p^2}{M} b^\intercal (\Sigma^*)^{-1} b \\ &= \quad H(p, M, a^*(p, M), \Sigma^*) \end{aligned} \tag{4.3}$$

where $\Sigma^* = \gamma(\theta^*)$ is a constant in $\Gamma = \Gamma(\Theta)$ defined by

$$\Sigma^* \quad \in \quad \arg\min_{\Sigma \in \Gamma} \big[b^\intercal \Sigma^{-1} b\big], \quad \text{i.e.} \quad \theta^* \ \in \ \arg\min_{\theta \in \Theta} R(\theta). \tag{4.4}$$

Moreover, the pair (a^*, Σ^*) is a saddle-point for H i.e. for all $p \in \mathbb{R}$, $M > 0$,

$$\begin{cases} H(p, M, a^*(p, M), \Sigma) \ \leq \ H(p, M, a^*(p, M), \Sigma^*) \ = \ H^*(p, M), \quad \forall \Sigma \in \Gamma, \\ \qquad H(p, M, a, \Sigma^*) \ \geq \ H(p, M, a^*(p, M), \Sigma^*) \ = \ H^*(p, M), \quad \forall a \in \mathbb{R}^d, \end{cases} \tag{4.5}$$

and a^* is explicitly given by:

$$a^*(p, M) \quad = \quad -\frac{p}{M} (\Sigma^*)^{-1} b. \tag{4.6}$$

Proof. Denote by $\tilde{H}$ the function defined on $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^d \times \Theta$ by

$$\tilde{H}(p, M, a, \theta) \quad := \quad H(p, M, a, \gamma(\theta)) \ = \ p a^\intercal b + \frac{1}{2} M a^\intercal \gamma(\theta) a.$$

Under the concavity assumption of γ in **(IC)**, we clearly see that for fixed $(p, M) \in \mathbb{R} \times (0, \infty)$, the function $\tilde{H}(p, M, ., .)$ is convex in $a \in \mathbb{R}^d$, and concave in θ lying in the convex-compact set Θ. By the min-max theorem (see e.g. Theorem 45.8 in [36]), we then get the so-called Isaacs relation:

$$\begin{aligned} \inf_{a \in \mathbb{R}^d} \sup_{\theta \in \Theta} \tilde{H}(p, M, a, \theta) \quad &= \quad \sup_{\theta \in \Theta} \inf_{a \in \mathbb{R}^d} \tilde{H}(p, M, a, \theta), \\ \text{i.e.} \quad \inf_{a \in \mathbb{R}^d} \sup_{\Sigma \in \Gamma} H(p, M, a, \Sigma) \quad &= \quad \sup_{\Sigma \in \Gamma} \inf_{a \in \mathbb{R}^d} H(p, M, a, \Sigma). \end{aligned}$$

[1] We use the partial ordering $\preceq$ on the set of $d \times d$-symmetric matrices: $M \preceq N \Leftrightarrow N - M$ is positive semi-definite $\Leftrightarrow a^\intercal (N - M) a \geq 0$ for all $a \in \mathbb{R}^d$.

By square completion, we can rewrite the function H as:

$$H(p,M,a,\Sigma) \;=\; \frac{M}{2}(a+\frac{p}{M}\Sigma^{-1}b)^\intercal\Sigma(a+\frac{p}{M}\Sigma^{-1}b)-\frac{1}{2}\frac{p^2}{M}b^\intercal\Sigma^{-1}b, \tag{4.7}$$

from which we get

$$\inf_{a\in\mathbb{R}^d} H(p,M,a,\Sigma) \;=\; H(p,M,\bar{a}(p,M,\Sigma),\Sigma) \;=\; -\frac{1}{2}\frac{p^2}{M}b^\intercal\Sigma^{-1}b, \tag{4.8}$$

where we set: $\bar{a}(p,M,\Sigma) := -\frac{p}{M}\Sigma^{-1}b$, and then the explicit expression of $H^*(p,M)$

$$H^*(p,M) \;=\; \sup_{\Sigma\in\Gamma}\inf_{a\in\mathbb{R}^d} H(p,M,a,\Sigma) \;=\; -\frac{1}{2}\frac{p^2}{M}\inf_{\Sigma\in\Gamma} b^\intercal\Sigma^{-1}b \;=\; -\frac{1}{2}\frac{p^2}{M}b^\intercal(\Sigma^*)^{-1}b.$$

Let us now check the saddle-point property of (a^*,Σ^*). By definition of $a^*(p,M)$, we have

$$\begin{aligned}
\sup_{\Sigma\in\Gamma} H(p,M,a^*(p,M),\Sigma) &= \inf_{a\in\mathbb{R}^d}\sup_{\Sigma\in\Gamma} H(p,M,a,\Sigma) \\
&= \sup_{\Sigma\in\Gamma}\inf_{a\in\mathbb{R}^d} H(p,M,a,\Sigma) \;=\; H^*(p,M) \\
&= \inf_{a\in\mathbb{R}^d} H(p,M,a,\Sigma^*) \;\le\; H(p,M,a,\Sigma^*), \quad \forall a\in\mathbb{R}^d,
\end{aligned}$$

where we used in the second equality Isaacs condition, and noticed in the last equality that Σ^* attains the supremum of $\Sigma\mapsto\inf_{a\in\mathbb{R}^d}H(p,M,a,\Sigma)$ by (4.8). We then deduce

$$\begin{aligned}
H(p,M,a^*(p,M),\Sigma^*) \;\le\; \sup_{\Sigma\in\Gamma} H(p,M,a^*(p,M),\Sigma) &= H^*(p,M) \\
&\le H(p,M,a,\Sigma^*), \quad \forall a\in\mathbb{R}^d,
\end{aligned}$$

which shows the second inequality in (4.5). Similarly, we have

$$\begin{aligned}
\inf_{a\in A} H(p,M,a,\Sigma^*) &= \sup_{\Sigma\in\Gamma}\inf_{a\in\mathbb{R}^d} H(p,M,a,\Sigma) \\
&= \inf_{a\in\mathbb{R}^d}\sup_{\Sigma\in\Gamma} H(p,M,a,\Sigma) \;=\; H^*(p,M) \\
&= \sup_{\Sigma\in\Gamma} H(p,M,a^*(p,M),\Sigma) \;\ge\; H(p,M,a^*(p,M),\Sigma), \quad \forall\Sigma\in\Gamma,
\end{aligned}$$

which implies that

$$\begin{aligned}
H(p,M,a^*(p,M),\Sigma^*) \;\ge\; \inf_{a\in\mathbb{R}^d} H(p,M,a,\Sigma^*) &= H^*(p,M) \\
&\ge H(p,M,a^*(p,M),\Sigma), \quad \forall\Sigma\in\Gamma.
\end{aligned}$$

This proves the first inequality in (4.5), hence the saddle-point property, and also that $H^*(p,M) = H(p,M,a^*(p,M),\Sigma^*)$.

On the other hand, by applying relation (4.8) to $\Sigma=\Sigma^*$, we have

$$H^*(p,M) \;=\; H(p,M,\bar{a}(p,M,\Sigma^*),\Sigma^*),$$

which combined with the saddle-point property of (a^*,Σ) shows that: $H(p,M,a^*(p,M),\Sigma^*)$ $= H(p,M,\bar{a}(p,M,\Sigma^*),\Sigma^*) = H^*(p,M)$, and then from the expression (4.7) of H:

$$\frac{M}{2}(a^*(p,M)-\bar{a}(p,M,\Sigma^*))^\intercal\Sigma^*(a-\bar{a}(p,M,\Sigma^*))+H^*(p,M) \;=\; H^*(p,M).$$

This proves that $a^*(p,M)=\bar{a}(p,M,\Sigma^*)$, i.e. the expression (4.6). $\square$

Remark 4.1 Under condition **(IC)**, and if the conditions of the verification theorem 3.1 are satisfied with a solution v to the Bellman-Isaacs PDE (3.8), and an optimal feedback control α^*, then we see from the saddle-point relation (4.5), that the drift $D_t^{\alpha,\sigma}$ of the deterministic process $V_t^{\alpha,\sigma} = v(t, \rho_t^{\alpha,\sigma})$ satisfies for all $\alpha \in \mathcal{A}$, $\Sigma \in \mathcal{V}_\Theta$,

$$D_t^{\alpha,\sigma^*} \;\geq\; D_t^{\alpha^*,\sigma^*} \;=\; 0 \;\geq\; D_t^{\alpha^*,\sigma}, \quad 0 \leq t \leq T, \; a.s.,$$

where $\sigma^* = (\Sigma^*)^{\frac{1}{2}}$. This means that the process (i) V_t^{α,σ^*} is nondecreasing for all $\alpha \in \mathcal{A}$, (ii) the process $V_t^{\alpha^*,\sigma}$ is nonincreasing for all $\Sigma \in \mathcal{V}_\Theta$, from which we easily deduce the min-max property:

$$V_0 \;=\; v(0,\delta_{x_0}) \;=\; \inf_{\alpha\in\mathcal{A}} \sup_{\Sigma\in\mathcal{V}_\Theta} J(\alpha,\sigma) \;=\; \sup_{\Sigma\in\mathcal{V}_\Theta} \inf_{\alpha\in\mathcal{A}} J(\alpha,\sigma) \;=\; J(\alpha^*,\sigma^*).$$

This shows in particular that σ^*, which is a constant explicitly computed from (4.4), i.e. minimizing the risk premium, is an optimal worst-case volatility for the robust mean-variance problem. $\Box$

Proposition 4.1 Assume that **(IC)** holds. Then, the function defined on $[0,T] \times \mathcal{P}_2(\mathbb{R})$ by

$$v(t,\mu) \;=\; K(t)\mathrm{Var}(\mu) - \bar{\mu} + \chi(t), \tag{4.9}$$

with

$$\begin{cases} K(t) & = & \lambda \exp\big(-R^*(T-t)\big) \\ \chi(t) & = & -\frac{1}{4\lambda}\Big[\exp\big(R^*(T-t)\big) - 1\Big], \quad 0 \leq t \leq T, \\ R^* & = & b^\intercal(\Sigma^*)^{-1}b, \end{cases} \tag{4.10}$$

is solution to the Bellman-Isaacs PDE (3.8).

Proof. We look for a function solution to (3.8) in the form:

$$v(t,\mu) \;=\; K(t)\mathrm{Var}(\mu) + Y(t)\bar{\mu} + \chi(t), \tag{4.11}$$

for some continuously differentiable functions $K > 0$, Y and χ on $[0,T]$. Such function is smooth and we have

$$\partial_\mu v(t,\mu)(x) \;=\; 2K(t)(x-\bar{\mu}) + Y(t), \qquad \partial_x\partial_\mu v(t,\mu)(x) \;=\; 2K(t) > 0,$$

From the expression of H^* in (4.3), we then get

$$\begin{aligned} & \partial_t v(t,\mu) + \int_{\mathbb{R}} H^*\big(\partial_\mu v(t,\mu)(x), \partial_x\partial_\mu v(t,\mu)(x)\big)\mu(dx) \\ = \; & \dot{K}(t)\mathrm{Var}(\mu) + \dot{Y}(t)\bar{\mu} + \dot{\chi}(t) \\ & \quad - \frac{1}{2} b^\intercal(\Sigma^*)^{-1}b \int \frac{4K(t)^2(x-\bar{\mu})^2 + Y(t)^2 + 4K(t)Y(t)(x-\bar{\mu})}{2K(t)}\mu(dx) \\ = \; & \big[\dot{K}(t) - b^\intercal(\Sigma^*)^{-1}bK(t)\big]\mathrm{Var}(\mu) + \dot{Y}(t)\bar{\mu} + \dot{\chi}(t) - \frac{1}{4} b^\intercal(\Sigma^*)^{-1}b\frac{Y(t)^2}{K(t)}. \end{aligned}$$

It follows that v in (4.11) satisfies the Bellman-Isaacs PDE (3.8) iff K, Y and χ satisfy the system of ordinary differential equations:

$$\begin{aligned}
\dot{K}(t) - b^\intercal(\Sigma^*)^{-1}bK(t) &= 0, \quad K(T) \;=\; \lambda \\
\dot{Y}(t) &= 0, \quad Y(T) \;=\; -1 \\
\dot{\chi}(t) - \frac{1}{4}b^\intercal(\Sigma^*)^{-1}b\frac{Y(t)^2}{K(t)} &= 0, \quad \chi(T) \;=\; 0,
\end{aligned}$$

which leads to the explicit solution $Y = -1$, K, χ as in (4.10). □

We can now provide a complete and explicit resolution of the robust mean-variance problem for a general class of covariance matrix uncertainty model satisfying **(IC)**.

Theorem 4.1 Let condition **(IC)** hold. There exists an optimal robust mean-variance strategy solution to (2.4), and given explicitly by

$$\alpha_t^* \;=\; \Big[x_0 + \frac{1}{2\lambda}\exp\big(R^*T\big) - X_t^*\Big](\Sigma^*)^{-1}b, \quad 0 \le t \le T, \;\; \mathcal{P}^\Theta - q.s. \tag{4.12}$$

where $R^* = b^\intercal(\Sigma^*)^{-1}b$ is the minimal risk premium corresponding to the worst-case covariance matrix parameter Σ^*, and with an optimal corresponding wealth process X^*, whose terminal return under any $\mathbb{P}^\sigma$, $\Sigma \in \mathcal{V}_\Theta$ is given by:

$$\mathbb{E}_\sigma[X_T^*] \;=\; x_0 + \frac{1}{2\lambda}\Big[\exp\big(R^*T\big) - 1\Big]. \tag{4.13}$$

Moreover, the optimal cost is given by

$$V_0 \;=\; v(0, \delta_{x_0}) \;=\; -\frac{1}{4\lambda}\big[\exp\big(R^*T\big) - 1\big] - x_0. \tag{4.14}$$

Proof. Let us consider the function $v(t,\mu)$ in (4.9), which satisfies the Bellman-Isaacs PDE (3.8). For this smooth function on $[0,T] \times \mathcal{P}_2(\mathbb{R})$, we have

$$\partial_\mu v(t,\mu)(x) \;=\; 2K(t)(x - \bar{\mu}) - 1, \qquad \partial_x\partial_\mu v(t,\mu)(x) \;=\; 2K(t) > 0, \tag{4.15}$$

with K as in (4.10). From the expression of a^* in (4.6), the candidate $\hat{a}(t,x,\mu)$ for the optimal feedback control in the verification Theorem 3.1 is then equal to:

$$\begin{aligned}
\hat{a}(t,x,\mu) &:= a^*(\partial_\mu v(t,\mu)(x), \partial_x\partial_\mu v(t,\mu)(x)) \\
&= -\Big[x - \bar{\mu} - \frac{1}{2K(t)}\Big](\Sigma^*)^{-1}b,
\end{aligned}$$

which is clearly Lipschitz in (x,μ). The solution $X^{\mathbb{P}^\sigma}$ to the McKean-Vlasov SDE (3.9) under $\mathbb{P}^\sigma \in \mathcal{P}^\Theta$, is thus governed by

$$dX_t^{\mathbb{P}^\sigma} \;=\; -\Big[X_t^{\mathbb{P}^\sigma} - \mathbb{E}_\sigma[X_t^{\mathbb{P}^\sigma}] - \frac{1}{2K(t)}\Big]b^\intercal(\Sigma^*)^{-1}[bdt + \sigma_t dW_t^\sigma], \;\; 0 \le t \le T, \; \mathbb{P}^\sigma - \text{p.s.},$$

which yields by taking expectation under $\mathbb{P}^\sigma$:

$$d\mathbb{E}_\sigma[X_t^{\mathbb{P}^\sigma}] \;=\; \frac{R^*}{2K(t)}dt,$$

and thus

$$\mathbb{E}_\sigma[X_t^{\mathbb{P}^\sigma}] = x_0 + \int_0^t \frac{R^*}{2K(s)} ds, \quad 0 \le t \le T.$$

The crucial observation is that this expectation does not depend on $\Sigma \in \mathcal{V}_\Theta$. Plugging into the SDE (3.9), this can be rewritten as (we now simply write $X_t = X_t^{\mathbb{P}^\sigma}$ to alleviate notations):

$$dX_t = -\Big[X_t - x_0 - \int_0^t \frac{R^*}{2K(s)} ds - \frac{1}{2K(t)}\Big] b^\intercal (\Sigma^*)^{-1}[b dt + dB_t], \quad 0 \le t \le T, \ \mathcal{P}^\Theta - \text{q.s.}.$$

This is now a standard SDE under $\mathcal{P}^\Theta$, and we know from Proposition 6.10 in [34] that there exists a $\mathcal{P}^\Theta$ quasi surely aggregated solution X^*, i.e. $X^* = X^{\mathbb{P}^\sigma}$, $\mathbb{P}^\sigma$-p.s, for all $\Sigma \in \mathcal{V}_\Theta$. Consequently, the family of processes $\{\hat a(t, X_t^{\mathbb{P}^\sigma}, \mathbb{P}^\sigma_{X_t^{\mathbb{P}^\sigma}}), 0 \le t \le T, \Sigma \in \mathcal{V}_\Theta\}$ can be aggregated into a $\mathcal{P}^\Theta$-q.s. defined by

$$\begin{aligned} \alpha_t^* &= \hat a(t, X_t^{\mathbb{P}^\sigma}, \mathbb{P}^\sigma_{X_t^{\mathbb{P}^\sigma}}), \quad 0 \le t \le T, \ \mathbb{P}^\sigma - \text{p.s.}, \ \forall \Sigma \in \mathcal{V}_\Theta \\ &= -\Big[X_t^* - x_0 - \int_0^t \frac{R^*}{2K(s)} ds - \frac{1}{2K(t)}\Big](\Sigma^*)^{-1} b, \quad 0 \le t \le T, \ \mathcal{P}^\Theta - \text{q.s.}, \end{aligned}$$

which gives from (4.10) the expression in (4.12). We conclude from the verification Theorem 3.1 that α^* is an optimal solution to (2.4), and the optimal cost is equal to $V_0 = v(0, \delta_{x_0})$, hence given by (4.14) from the explicit form of v in (4.9). □

Remark 4.2 Although the original robust mean-variance problem is *a priori* a complex and non standard stochastic differential game problem, the message of the main result in Theorem 4.1 is quite simple with an intuitive interpretation. It says that the resolution of this problem can be reduced into two steps: first, we determine the worst-case scenario, and the remarkable point is that it corresponds to a constant covariance matrix Σ^* obtained by the minimization (4.4) of the risk premium. This constant is directly computed from the inputs of the model: the instantaneous return b (assumed to be known), and the function γ parametrizing the uncertainty on the covariance matrix of the assets (we shall give in the sequel some examples for explicit computations of Σ^*). Secondly, we obtain the optimal mean-variance strategy as in the Black-Scholes model with instantaneous return b and covariance matrix Σ^*, whose expression has been derived in [38], and that we recover here by a different approach as a particular case when there is no uncertainty on the model.
□

Remark 4.3 (About drift uncertainty)
Let us discuss the case when there is ambiguity on the drift of the d risky assets (but with known covariance matrix Σ for simplicity). This is modeled by considering that the drift process $b = (b_t)_t \in \mathcal{V}_\Theta$ is an unobservable process, which is only known to be valued in a given convex set Θ of $\mathbb{R}^d$. The Hamiltonian function for the corresponding robust optimization problem is then given by (by abuse of notation, we keep the same notation H as in the case of uncertain covariance matrix):

$$H(p, M, a, \theta) = p a^\intercal \theta + \frac{1}{2} M a^\intercal \Sigma a, \quad (p, M, a, \theta) \in \mathbb{R} \times (0, \infty) \times \mathbb{R}^d \times \Theta.$$

By similar arguments as in Lemma 4.1, for fixed $(p, M) \in \mathbb{R} \times (0, \infty)$, there is a saddle-point $(a^*(p, M), \theta^*)$ for $H(p, M, ., .)$ given by

$$a^*(p, M) \;=\; -\frac{p}{M}\Sigma^{-1}\theta^*, \qquad \theta^* \;\in\; \arg\min_{\theta \in \Theta}\big[\theta^\intercal \Sigma^{-1}\theta\big].$$

Then, similarly as in Proposition 4.1, we find that v given in (4.9)-(4.10) is solution to the associated Bellman-Isaacs PDE for the robust mean-variance problem, where the "worst-case" risk premium R^* is now given by

$$R^* \;=\; (\theta^*)^\intercal \Sigma^{-1}\theta^*.$$

Following arguments as in the verification Theorem 3.1, this leads to a candidate for the optimal feedback control in the form

$$\begin{aligned}\hat{a}(t, x, \mu) \;&:=\; a^*(\partial_\mu v(t, \mu)(x), \partial_x \partial_\mu v(t, \mu)(x)) \\ &=\; -\Big[x - \bar{\mu} - \frac{1}{2K(t)}\Big]\Sigma^{-1}\theta^*,\end{aligned}$$

and we then has to consider the solution $X^{\mathbb{P}^b}$ to the McKean-Vlasov SDE (3.9) under any (equivalent) prior probability measure $\mathbb{P}^b$, $b \in \mathcal{V}_\Theta$, governed by

$$dX_t^{\mathbb{P}^b} \;=\; -\Big[X_t^{\mathbb{P}^b} - \mathbb{E}_b[X_t^{\mathbb{P}^b}] - \frac{1}{2K(t)}\Big](\theta^*)^\intercal \Sigma^{-1}[b_t dt + \sigma dW_t^b], \;\; 0 \le t \le T, \; \mathbb{P}^b - \text{p.s.}.$$

By taking expectation under $\mathbb{P}^b$, we get

$$\mathbb{E}_b[X_t^{\mathbb{P}^b}] \;=\; x_0 + \int_0^t \frac{1}{2K(s)}(\theta^*)^\intercal \Sigma^{-1}\mathbb{E}_b[b_s]ds,$$

and see that, in contrast with the case of covariance matrix uncertainty, this expectation depends on the prior probability measure $\mathbb{P}^b$. Consequently, the family of processes $\{\hat{a}(t, X_t^{\mathbb{P}^b}, \mathbb{P}^b_{X_t^{\mathbb{P}^b}}), 0 \le t \le T, \Sigma \in \mathcal{V}_\Theta\}$ cannot be aggregated into a universal process α^*, which would allow us to conclude that α^* is an optimal strategy. The main issue in the mean-variance framework, compared to classical (robust) expected utility maximization where the optimal strategy depends in feedback form only on the wealth state process, arises from the feedback form dependence of the optimal wealth process not only upon the wealth process, but also on the expected wealth process, which depends on the prior probability measure when considering drift uncertainty. The robust mean-variance and Markowitz problem is then a challenging problem that could not be directly tackled by our approach and that we postpone for future research. □

4.1 Example 1: uncertain volatility

We consider the uncertain volatility model in the multivariate case with zero correlation as presented in Example 1: $\Theta = \prod_{i=1}^d [\underline{\sigma_i}^2, \bar{\sigma}_i^2]$ with $0 < \underline{\sigma_i} \le \bar{\sigma}_i < \infty$, $i = 1, \ldots, d$, and

$$\gamma(\theta) \;=\; \begin{pmatrix} \sigma_1^2 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \sigma_d^2 \end{pmatrix}, \quad \text{for } \; \theta = (\sigma_1^2, \ldots, \sigma_d^2).$$

In this case, the risk premium function is simply given by

$$R(\theta) \ := \ b^\intercal \gamma(\theta)^{-1} b \ = \ \sum_{i=1}^{d} \frac{b_i^2}{\sigma_i^2}, \quad \text{for } \theta = (\sigma_1^2, \ldots, \sigma_d^2),$$

and its is clear that the worst-case scenario corresponds to the covariance matrix $\Sigma^* = \bar{\Sigma}$ $:= \gamma(\bar{\theta})$ with $\bar{\theta} = (\bar{\sigma}_1^2, \ldots, \bar{\sigma}_d^2)$, i.e. for the highest marginal volatilities.

From Theorem 4.1, we obtain an explicit optimal portfolio strategy for the robust mean-variance problem under uncertain volatility:

$$\begin{aligned} \alpha_t^* \ &= \ \Big[x_0 + \frac{1}{2\lambda}\exp\big(\bar{R}\,T\big) - X_t^*\Big]\bar{\Sigma}^{-1}b, \quad 0 \leq t \leq T, \ \ \mathcal{P}^\Theta q.s., \\ \text{with} \ \ \bar{R} \ &:= \ R(\bar{\theta}) \ = \ b^\intercal \bar{\Sigma}^{-1} b. \end{aligned}$$

This corresponds to the optimal mean-variance portfolio strategy in a multidimensional Black-Scholes model with uncorrelated assets of drift b and covariance matrix $\bar{\Sigma}$, as derived in [38] and [13]. The financial interpretation is natural: the worst-case scenario corresponds to the highest variance $\bar{\Sigma}$, and the risk-averse investor makes her/his portfolio decision by referring to this case.

4.2 Example 2: ambiguous correlation

We consider the model for a two-risky assets model with ambiguous correlation, i.e. $\Theta = [\underline{\varrho}, \bar{\varrho}] \subset (-1,1)$, and

$$\gamma(\theta) \ = \ \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\theta \\ \sigma_1\sigma_2\theta & \sigma_2^2 \end{pmatrix}, \quad \theta \in \Theta,$$

for some known positive constants $\sigma_1 > 0$ and $\sigma_2 > 0$.

In this case, the risk premium function is given by

$$R(\theta) \ := \ b^\intercal \gamma(\theta)^{-1} b \ = \ \frac{1}{1-\theta^2}\big(\beta_1^2 + \beta_2^2 - 2\beta_1\beta_2\theta\big), \tag{4.16}$$

where we denote by $\beta_i = \frac{b_i}{\sigma_i}$, $i = 1,2$, the instantaneous Sharpe ratio of each risky asset. When the asset S^i is a stock, its sharpe ratio is usually positive (otherwise it would perform less than the riskless bond). We may also want to consider the case when β_i is nonpositive, which would correspond typically to the case when the asset S^i is a spread between two stocks. In the sequel, we shall assume w.l.o.g. that $(\beta_1, \beta_2) \neq (0,0)$ (in this trivial case, the optimal portfolio strategy is clearly to never trade, i.e. $\alpha^* \equiv 0$), and we set:

$$\varrho_0^+ \ := \ \frac{\min(|\beta_1|, |\beta_2|)}{\max(|\beta_1|, |\beta_2|)} \ \in \ [0,1], \qquad \varrho_0^- \ := \ -\varrho_0^+. \tag{4.17}$$

Let us also introduce the extremal covariance matrices

$$\bar{\Sigma} \ := \ \gamma(\bar{\varrho}) \ = \ \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\bar{\varrho} \\ \sigma_1\sigma_2\bar{\varrho} & \sigma_2^2 \end{pmatrix}, \qquad \underline{\Sigma} \ := \ \gamma(\underline{\varrho}) \ = \ \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\underline{\varrho} \\ \sigma_1\sigma_2\underline{\varrho} & \sigma_2^2 \end{pmatrix},$$

and their corresponding variance risk ratios:

$$\bar{\Sigma}^{-1}b \ = \ \frac{1}{1-\bar{\varrho}^2}\begin{pmatrix} \frac{b_1}{\sigma_1^2} - \frac{b_2\bar{\varrho}}{\sigma_1\sigma_2} \\ \frac{b_2}{\sigma_2^2} - \frac{b_1\bar{\varrho}}{\sigma_1\sigma_2} \end{pmatrix} \ =: \ \begin{pmatrix} \bar{\kappa}_1 \\ \bar{\kappa}_2 \end{pmatrix}, \qquad \underline{\Sigma}^{-1}b \ = \ \frac{1}{1-\underline{\varrho}^2}\begin{pmatrix} \frac{b_1}{\sigma_1^2} - \frac{b_2\underline{\varrho}}{\sigma_1\sigma_2} \\ \frac{b_2}{\sigma_2^2} - \frac{b_1\underline{\varrho}}{\sigma_1\sigma_2} \end{pmatrix} \ =: \ \begin{pmatrix} \underline{\kappa}_1 \\ \underline{\kappa}_2 \end{pmatrix}.$$

The following result provides the explicit determination of the correlation θ^* achieving the minimal risk premium.

Lemma 4.2 We distinguish two cases depending on the sign of $\beta_1\beta_2$.
I. For $\beta_1\beta_2 > 0$, we have:

1. if $\bar{\varrho} < \varrho_0^+$, then $\theta^* = \bar{\varrho}$. Moreover, $\bar{\kappa}_1\bar{\kappa}_2 > 0$ and $\underline{\kappa}_1\underline{\kappa}_2 > 0$.

2. if $\underline{\varrho} > \varrho_0^+$, then $\theta^* = \underline{\varrho}$. Moreover, $\underline{\kappa}_1\underline{\kappa}_2 < 0$ and $\bar{\kappa}_1\bar{\kappa}_2 < 0$.

3. if $\varrho_0^+ \in \Theta = [\underline{\varrho}, \bar{\varrho}]$, then $\theta^* = \varrho_0^+$. Moreover, $\underline{\kappa}_1\underline{\kappa}_2 \geq 0$ and $\bar{\kappa}_1\bar{\kappa}_2 \leq 0$.

I'. For $\beta_1\beta_2 \leq 0$, we have:

1'. if $\bar{\varrho} < \varrho_0^-$, then $\theta^* = \bar{\varrho}$. Moreover, $\bar{\kappa}_1\bar{\kappa}_2 > 0$ and $\underline{\kappa}_1\underline{\kappa}_2 > 0$.

2'. if $\underline{\varrho} > \varrho_0^-$, then $\theta^* = \underline{\varrho}$. Moreover, $\underline{\kappa}_1\underline{\kappa}_2 < 0$ and $\bar{\kappa}_1\bar{\kappa}_2 < 0$.

3'. if $\varrho_0^- \in \Theta = [\underline{\varrho}, \bar{\varrho}]$, then $\theta^* = \varrho_0^-$. Moreover, $\underline{\kappa}_1\underline{\kappa}_2 \geq 0$ and $\bar{\kappa}_1\bar{\kappa}_2 \leq 0$.

Proof. The risk premium function R is differentiable on $\Theta = [\underline{\varrho}, \bar{\varrho}]$, with a derivative given by:

$$R'(\theta) \;=\; -\frac{2}{(1-\theta^2)^2} f(\theta), \quad \text{with} \quad f(\theta) \;=\; \beta_1\beta_2(1+\theta^2) - (\beta_1^2+\beta_2^2)\theta.$$

For any $\theta \in \Theta$, let us also denote by $\kappa_1(\theta)$, $\kappa_2(\theta)$ the components of the variance risk ratio $\Sigma(\theta)^{-1}b$, i.e.

$$\kappa_1(\theta) \;=\; \frac{1}{1-\theta^2}\Big(\frac{b_1}{\sigma_1^2} - \frac{b_2\theta}{\sigma_1\sigma_2}\Big), \qquad \kappa_2(\theta) \;=\; \frac{1}{1-\theta^2}\Big(\frac{b_2}{\sigma_2^2} - \frac{b_1\theta}{\sigma_1\sigma_2}\Big),$$

so that $\bar{\kappa}_i = \kappa_i(\bar{\varrho})$, and $\underline{\kappa}_i = \kappa_i(\underline{\varrho})$, $i = 1, 2$, and notice that

$$\kappa_1(\theta)\kappa_2(\theta) \;=\; \frac{1}{\sigma_1\sigma_2(1-\theta^2)^2} f(\theta). \tag{4.18}$$

I. We first consider the case when $\beta_1\beta_2 > 0$. In this case, the function f is a strictly convex parabolic function attaining its infimum on $\mathbb{R}$ at $\bar{\theta} = \frac{\beta_1^2+\beta_2^2}{2\beta_1\beta_2} \geq 1$, which implies that f is strictly decreasing on $(-\infty, \bar{\theta}]$ hence on Θ. Since $f(0) = \beta_1\beta_2 > 0$ and $f(1) = -(\beta_1-\beta_2)^2 \leq 0$, there exists a unique $\varrho_0^+ \in (0, 1]$ s.t. $f(\varrho_0^+) = 0$, which is exactly given by the expression in (4.17). We are then led to distinguish the following cases:
1. $\bar{\varrho} < \varrho_0^+$.
In this case, recalling that f is strictly decreasing on $\Theta = [\underline{\varrho}, \bar{\varrho}]$, we see that for all $\theta \in \Theta$, $f(\theta) > f(\varrho_0^+) = 0$, i.e. $R'(\theta) < 0$ on Θ, i.e. R is strictly decreasing on Θ, and thus: $\theta^* = \arg\min_{\theta\in\Theta} R(\theta) = \bar{\varrho}$. Moreover, by (4.18), we have $\kappa_1(\theta)\kappa_2(\theta) > 0$ for all $\theta \in \Theta$, and thus: $\underline{\kappa}_1\underline{\kappa}_2 > 0$ and $\bar{\kappa}_1\bar{\kappa}_2 > 0$.
2. $\underline{\varrho} > \varrho_0^+$. In this case, $f(\bar{\varrho}) \leq f(\underline{\varrho}) < f(\varrho_0^+) = 0$, and thus for all $\theta \in \Theta$, $f(\theta) < 0$, $\kappa_1(\theta)\kappa_2(\theta) < 0$, $R'(\theta) > 0$, i.e. R is strictly increasing on Θ. This implies that $\theta^* = \arg\min_{\theta\in\Theta} R(\theta) = \underline{\varrho}$, and also $\underline{\kappa}_1\underline{\kappa}_2 < 0$, $\bar{\kappa}_1\bar{\kappa}_2 < 0$.
3. $\underline{\varrho} \leq \varrho_0^+ \leq \bar{\varrho}$, i.e. $\varrho_0^+ \in \Theta$. Notice that in this case, ϱ_0^+ is strictly smaller than 1 (recall that $\bar{\varrho} < 1$), and thus $\beta_1 \neq \beta_2$. Again, since f is decreasing, we have $f(\theta) \geq f(\varrho_0^+) = 0$ for

$\theta \in [\underline{\varrho}, \varrho_0^+]$, and $f(\theta) \leq f(\varrho_0^+) = 0$ for $\theta \in [\varrho_0^+, \bar{\varrho}]$. Therefore, $\kappa_1(\theta)\kappa_2(\theta) \geq 0$, $R'(\theta) \leq 0$ for $\theta \in [\underline{\varrho}, \varrho_0^+]$, i.e. R is decreasing on $[\underline{\varrho}, \varrho_0^+]$, and $\kappa_1(\theta)\kappa_2(\theta) \leq 0$, $R'(\theta) \geq 0$ for $\theta \in [\varrho_0^+, \bar{\varrho}]$, i.e. R is increasing on $[\varrho_0^+, \bar{\varrho}]$. Therefore, $\theta^* = \arg\min_{\theta\in\Theta} R(\theta) = \varrho_0^+$, and we also have $\underline{\kappa}_1\underline{\kappa}_2 \geq 0$ and $\bar{\kappa}_1\bar{\kappa}_2 \leq 0$.

I'. We finally consider the case when $\beta_1\beta_2 \leq 0$. When $\beta_1\beta_2 < 0$, the function f is a strictly concave parabolic function attaining its infimum on $\mathbb{R}$ at $\bar{\theta} = \frac{\beta_1^2+\beta_2^2}{2\beta_1\beta_2} \leq -1$, and when $\beta_1\beta_2 = 0$, f is a linear function with strictly negative slope. In any case, the function f is strictly decreasing on $[\bar{\theta}, \infty)$ hence on Θ. Since $f(0) = \beta_1\beta_2 \leq 0$, $f(-1) = (\beta_1 + \beta_2)^2 \geq 0$, there exists a unique $\varrho_0^- \in [-1, 0]$ s.t. $f(\varrho_0^-) = 0$, which is exactly given by the expression in (4.17), i.e. $\varrho_0^- = -\varrho_0^+$. Then, by distinguishing the cases when $\varrho_0^- > \underline{\varrho}$, $\varrho_0^- < \underline{\varrho}$ and $\varrho_0^- \in \Theta$, and proceeding by the same arguments as in Case **I**, we obtain the results described in **1'**, **2'** and **3'**. $\square$

By applying Theorem 4.1, we can now provide an explicit description of the optimal strategy under ambiguous correlation.

Theorem 4.2 The solution to problem (2.4) is explicitly described through the following cases[2] :

I. If $\beta_1\beta_2 > 0$, and

1. $\bar{\varrho} < \varrho_0^+$, then an optimal portfolio strategy is explicitly given by

$$\alpha_t^* = \Big[x_0 + \frac{1}{2\lambda}\exp\big(\bar{R}T\big) - X_t^*\Big]\bar{\Sigma}^{-1}b, \quad 0 \leq t \leq T, \ \ \mathcal{P}^{\Theta} q.s.,$$

with $\bar{R} = b^{\intercal}\bar{\Sigma}^{-1}b$, and the optimal cost is

$$V_0 = -\frac{1}{4\lambda}\big[\exp\big(\bar{R}T\big) - 1\big] - x_0.$$

2. $\underline{\varrho} > \varrho_0^+$, then an optimal portfolio strategy is explicitly given by

$$\alpha_t^* = \Big[x_0 + \frac{1}{2\lambda}\exp\big(\underline{R}T\big) - X_t^*\Big]\underline{\Sigma}^{-1}b, \quad 0 \leq t \leq T, \ \ \mathcal{P}^{\Theta} q.s.,$$

with $\underline{R} = b^{\intercal}\underline{\Sigma}^{-1}b$, and the optimal cost is

$$V_0 = -\frac{1}{4\lambda}\big[\exp\big(\underline{R}T\big) - 1\big] - x_0.$$

3. $\underline{\varrho} \leq \varrho_0^+ \leq \bar{\varrho}$, then an optimal portfolio strategy is explicitly given by

$$\alpha_t^* = \begin{cases} \begin{pmatrix} \Big[x_0 + \frac{1}{2\lambda}\exp\big(\beta_1^2 T\big) - X_t^*\Big]\frac{b_1}{\sigma_1^2} \\ 0 \end{pmatrix}, & 0 \leq t \leq T, \ \ \mathcal{P}^{\Theta} q.s., \quad \text{when } \beta_1^2 > \beta_2^2, \\[2ex] \begin{pmatrix} 0 \\ \Big[x_0 + \frac{1}{2\lambda}\exp\big(\beta_2^2 T\big) - X_t^*\Big]\frac{b_2}{\sigma_2^2} \end{pmatrix}, & 0 \leq t \leq T, \ \ \mathcal{P}^{\Theta} q.s., \quad \text{when } \beta_2^2 > \beta_1^2. \end{cases}$$

and the optimal cost

$$V_0 = -\frac{1}{4\lambda}\big[\exp\big(\max(\beta_1^2, \beta_2^2)T\big) - 1\big] - x_0.$$

[2]By misuse of notation, we write indifferently $a = (a_1, a_2)$ or $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ for an element in $\mathbb{R}^2$.

I'. If $\beta_1\beta_2 \leq 0$, and

1'. $\bar{\varrho} < \varrho_0^-$, then an optimal portfolio strategy is explicitly given by

$$\alpha_t^* = \Big[x_0 + \frac{1}{2\lambda}\exp\big(\bar{R}T\big) - X_t^*\Big]\bar{\Sigma}^{-1}b, \quad 0 \leq t \leq T, \ \ \mathcal{P}^\Theta q.s.,$$

and the optimal cost is

$$V_0 = -\frac{1}{4\lambda}\big[\exp\big(\bar{R}T\big) - 1\big] - x_0.$$

Moreover, $\bar{\kappa}_1\bar{\kappa}_2 > 0$ and $\underline{\kappa}_1\underline{\kappa}_2 > 0$.

2'. $\underline{\varrho} > \varrho_0^-$, then an optimal portfolio strategy is explicitly given by

$$\alpha_t^* = \Big[x_0 + \frac{1}{2\lambda}\exp\big(\underline{R}T\big) - X_t^*\Big]\underline{\Sigma}^{-1}b, \quad 0 \leq t \leq T, \ \ \mathcal{P}^\Theta q.s.,$$

and the optimal cost

$$V_0 = -\frac{1}{4\lambda}\big[\exp\big(\underline{R}T\big) - 1\big] - x_0.$$

Moreover, $\underline{\kappa}_1\underline{\kappa}_2 < 0$ and $\bar{\kappa}_1\bar{\kappa}_2 < 0$.

3'. $\underline{\varrho} \leq \varrho_0^- \leq \bar{\varrho}$, then an optimal portfolio strategy is explicitly given by

$$\alpha_t^* = \begin{cases} \begin{pmatrix} \Big[x_0 + \frac{1}{2\lambda}\exp\big(\beta_1^2 T\big) - X_t^*\Big]\frac{b_1}{\sigma_1^2} \\ 0 \end{pmatrix}, & 0 \leq t \leq T, \ \ \mathcal{P}^\Theta q.s., \quad \text{when } \beta_1^2 > \beta_2^2, \\ \\ \begin{pmatrix} 0 \\ \Big[x_0 + \frac{1}{2\lambda}\exp\big(\beta_2^2 T\big) - X_t^*\Big]\frac{b_2}{\sigma_2^2} \end{pmatrix}, & 0 \leq t \leq T, \ \ \mathcal{P}^\Theta q.s., \quad \text{when } \beta_2^2 > \beta_1^2. \end{cases}$$

and the optimal cost

$$V_0 = -\frac{1}{4\lambda}\big[\exp\big(\max(\beta_1^2, \beta_2^2)T\big) - 1\big] - x_0.$$

Moreover, $\underline{\kappa}_1\underline{\kappa}_2 \geq 0$ and $\bar{\kappa}_1\bar{\kappa}_2 \leq 0$.

Proof. In view of the formulae (4.12) and (4.14) of the optimal portfolio strategy and optimal cost in Theorem 4.1, we only has to compute the minimal risk premium $R^* = R(\theta^*)$, and the vector $(\Sigma^*)^{-1}b$ with $\Sigma^* = \gamma(\theta^*)$, and θ^* explicitly given in Lemma 4.2. We only consider the case **I** when $\beta_1\beta_2 > 0$ since the other case **I'** is dealt with similarly. The subcases **1** and **2** are immediate, and we only focus on the third case **3** when $\varrho_0^+ \in [\underline{\varrho}, \bar{\varrho}]$. In this case $\theta^* = \varrho_0^+$, and a simple computation from the expressions of R in (4.16) and ϱ_0^+ in (4.17) gives: $R^* = R(\varrho_0^+) = \max(\beta_1^2, \beta_2^2)$. Moreover, a straightforward calculation shows that

$$(\Sigma^*)^{-1}b = \gamma(\varrho_0^+)^{-1}b = \begin{cases} \begin{pmatrix} \frac{b_1}{\sigma_1^2} \\ 0 \end{pmatrix}, & \text{when } \beta_1^2 > \beta_2^2, \\ \\ \begin{pmatrix} 0 \\ \frac{b_2}{\sigma_2^2} \end{pmatrix}, & \text{when } \beta_2^2 > \beta_1^2, \end{cases}$$

which leads to the expression of the optimal portfolio strategy in the assertion of the Theorem. $\square$

Remark 4.4 (Financial interpretation)
To fix the idea, we focus on the usual case of two stocks when $\beta_1 > 0, \beta_2 > 0$. The coefficient ϱ_0^+ can be viewed as a measure for the "proximity" between the two stocks: a small ϱ_0^+ (close to zero) means that one stock is much better than the other one in the sense that it has a much larger instantaneous Sharpe ratio, while large ϱ_0^+ (close to one) means that the two stocks are similar in terms of instantaneous Sharpe ratio.

When $\bar{\varrho} < \varrho_0^+$, this means that no stock is "dominating" the other one, and it is optimal to invest in both assets with a directional trading, that is buying or selling simultaneously (recall from Lemma 4.2 that in this case $\bar{\kappa}_1\bar{\kappa}_2 > 0$), and the worst-case scenario refers to the highest correlation $\bar{\rho}$ where the diversification effect is minimal. The optimal strategy corresponds to the optimal mean-variance portfolio strategy in a market with constant covariance matrix $\bar{\Sigma}$.

When $\underline{\varrho} > \varrho_0^+$, 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 (recall that in this case $\underline{\kappa}_1\underline{\kappa}_2 < 0$), and the worst-case scenario corresponds to the lowest correlation where the profit from the spread trading is minimal.

When $\underline{\varrho} \leq \varrho_0^+ \leq \bar{\varrho}$, it is optimal to invest in either one of the stocks, but not both, since the directional trading is not optimal for high correlation and the spread trading is not optimal for low correlation. The selection for the risky asset is then naturally made on the one with the highest instantaneous Sharpe ratio.

We notice that a similar interpretation was derived in [14] for robust portfolio optimization with utility function, but in this cited paper, the authors derived the worst-case scenario by distinguish four cases (see their Theorem 2.2): (1) $\bar{\kappa}_1\bar{\kappa}_2 > 0$ and $\underline{\kappa}_1\underline{\kappa}_2 \geq 0$, (2) $\bar{\kappa}_1\bar{\kappa}_2 \leq 0$ and $\underline{\kappa}_1\underline{\kappa}_2 < 0$, (3) $\bar{\kappa}_1\bar{\kappa}_2 \leq 0$ and $\underline{\kappa}_1\underline{\kappa}_2 \geq 0$, and (4) $\bar{\kappa}_1\bar{\kappa}_2 > 0$ and $\underline{\kappa}_1\underline{\kappa}_2 < 0$. Compared to [14], we push further the calculations and reduce the different cases on the variance risk ratios $\bar{\kappa}_1$ and $\bar{\kappa}_2$ to an explicit description with three cases in terms of the correlations $\bar{\varrho}$, $\underline{\varrho}$, and ϱ_0^+. Actually, as shown in Lemma 4.2, our cases **1**, resp. **2**, resp. **3** in Theorem 4.2 are equivalent to their cases (1), resp. (2), resp. (3), and it appears that their last case (4) can never happen. Let us also mention that a similar description with three cases in terms of the correlation was done in [23] (see their Proposition 2) for a single-period mean-variance problem under correlation ambiguity. □

5 Robust efficient frontier

Let us denote by $U_0(\vartheta)$ the optimal worst-case expected terminal wealth given a worst-case variance risk $\vartheta > 0$, i.e.,

$$U_0(\vartheta) \quad = \quad \sup\Big\{\mathcal{E}(\alpha) : \alpha \in \mathcal{A}, \mathcal{R}(\alpha) \leq \vartheta\Big\},$$

where we recall the notations from the robust Markowitz problem (2.5):

$$\mathcal{E}(\alpha) \ := \inf_{\mathbb{P}^\sigma \in \mathcal{P}^\Theta} \mathbb{E}_\sigma[X_T^\alpha], \qquad \mathcal{R}(\alpha) \ := \sup_{\mathbb{P}^\sigma \in \mathcal{P}^\Theta} \mathrm{Var}_\sigma(X_T^\alpha).$$

By the linearity of X^α w.r.t. α lying in the convex set $\mathcal{A}$, the convexity (resp. the linearity) of $X \in L^2(\mathcal{F}_T, \mathbb{P}^\sigma) \mapsto \mathrm{Var}_\sigma(X)$ (resp. $\mathbb{E}_\sigma[X]$), it is easily seen that the function U_0 is concave w.r.t. $\vartheta \in (0, \infty)$.

We consider the general framework of Section 4 under condition **(IC)**, and emphasize the dependence of $V_0 = V_0(\lambda)$, and $\alpha^* = \alpha^{*,\lambda}$, for the optimal cost and optimal portfolio strategy to the robust mean-variance portfolio selection problem (2.4) with risk-aversion parameter λ:

$$\begin{aligned} V_0(\lambda) &= \inf_{\alpha\in\mathcal{A}} \sup_{\mathbb{P}^\sigma\in\mathcal{P}^\Theta} \Big(\lambda \mathrm{Var}_\sigma(X_T^\alpha) - \mathbb{E}_\sigma[X_T^\alpha]\Big) \\ &= \sup_{\mathbb{P}^\sigma\in\mathcal{P}^\Theta} \Big(\lambda \mathrm{Var}_\sigma(X_T^{\alpha^{*,\lambda}}) - \mathbb{E}_\sigma[X_T^{\alpha^{*,\lambda}}]\Big). \end{aligned}$$

From (4.14) in Theorem 4.1, we recall that

$$V_0(\lambda) = -\frac{1}{4\lambda}\Big[\exp\big(R(\theta^*)T\big) - 1\Big] - x_0, \tag{5.1}$$

where $R(\theta) = b^\intercal \gamma(\theta)^{-1} b$, and $\theta^* = \arg\min_{\theta\in\Theta} R(\theta)$. Moreover, a crucial observation (see (4.13)) is that the expected optimal terminal wealth $\mathbb{E}_\sigma[X_T^{\alpha^{*,\lambda}}]$ under any prior probability measure $\mathbb{P}^\sigma$ does not depend actually on $\Sigma \in \mathcal{V}_\Theta$, and thus:

$$\mathcal{E}(\alpha^{*,\lambda}) = \mathbb{E}_\sigma[X_T^{\alpha^{*,\lambda}}] =: \bar\rho_T^{*,\lambda}, \quad \forall \Sigma \in \mathcal{V}_\Theta, \tag{5.2}$$

with

$$\bar\rho_T^{*,\lambda} = x_0 + \frac{1}{2\lambda}\Big[\exp\big(R(\theta^*)T\big) - 1\Big]. \tag{5.3}$$

By adapting standard arguments from convex optimization theory, we show the duality relation between the robust mean-variance problem and the robust Markowitz problem, namely:

$$\begin{aligned} V_0(\lambda) &= \inf_{\vartheta>0}\big[\lambda\vartheta - U_0(\vartheta)\big], \quad \forall \lambda > 0, \\ U_0(\vartheta) &= \inf_{\lambda>0}\big[\lambda\vartheta - V_0(\lambda)\big], \quad \forall \vartheta > 0. \end{aligned} \tag{5.4}$$

Indeed, for fixed $\vartheta > 0$, and for any $\varepsilon > 0$, there exists an ε-optimal control for $U_0(\vartheta)$, that is a control $\tilde\alpha^\varepsilon \in \mathcal{A}$ s.t. $U_0(\vartheta) \leq \mathcal{E}(\tilde\alpha^\varepsilon) + \varepsilon$, and $\mathcal{R}(\tilde\alpha^\varepsilon) \leq \vartheta$. It follows that for all $\lambda > 0$,

$$\begin{aligned} V_0(\lambda) &\leq \sup_{\mathbb{P}^\sigma\in\mathcal{P}^\Theta} \Big(\lambda \mathrm{Var}_\sigma(X_T^{\tilde\alpha^\varepsilon}) - \mathbb{E}_\sigma[X_T^{\tilde\alpha^\varepsilon}]\Big) \\ &\leq \lambda \sup_{\mathbb{P}^\sigma\in\mathcal{P}^\Theta} \mathrm{Var}_\sigma(X_T^{\tilde\alpha^\varepsilon}) - \inf_{\mathbb{P}^\sigma\in\mathcal{P}^\Theta} \mathbb{E}_\sigma[X_T^{\tilde\alpha^\varepsilon}] = \lambda\mathcal{R}(\tilde\alpha^\varepsilon) - \mathcal{E}(\tilde\alpha^\varepsilon) \\ &\leq \lambda\vartheta - U_0(\vartheta) + \varepsilon. \end{aligned}$$

Since ε is arbitrary, and the above relation holds for any fixed $\vartheta > 0$, this shows that

$$V_0(\lambda) \leq \inf_{\vartheta>0}\big[\lambda\vartheta - U_0(\vartheta)\big], \quad \forall \lambda > 0. \tag{5.5}$$

Conversely, for fixed $\lambda > 0$, let us consider the optimal control $\alpha^{*,\lambda} \in \mathcal{A}$ for $V_0(\lambda)$, and set $\vartheta_\lambda := \mathcal{R}(\alpha^{*,\lambda})$ which is strictly positive since the terminal wealth $X_T^{\alpha^{*,\lambda}}$ is not constant. Then, by definition of $U_0(\vartheta_\lambda)$, we have $\mathcal{E}(\alpha^{*,\lambda}) \leq U_0(\vartheta_\lambda)$, and so by (5.2)

$$\begin{aligned} V_0(\lambda) &= \sup_{\mathbb{P}^\sigma\in\mathcal{P}^\Theta} \Big(\lambda \mathrm{Var}_\sigma(X_T^{\alpha^{*,\lambda}}) - \mathbb{E}_\sigma[X_T^{\alpha^{*,\lambda}}]\Big) = \lambda\mathcal{R}(\alpha^{*,\lambda}) - \mathcal{E}(\alpha^{*,\lambda}) \\ &\geq \lambda\vartheta_\lambda - U_0(\vartheta_\lambda). \end{aligned} \tag{5.6}$$

Together with (5.5), this shows the first duality relation in (5.4), i.e., V_0 is the Fenchel-Legendre transform of U_0, and ϑ_λ attains the infimum in this transform:

$$V_0(\lambda) = \inf_{\vartheta>0}\big[\lambda\vartheta - U_0(\vartheta)\big] = \lambda\vartheta_\lambda - U_0(\vartheta_\lambda). \tag{5.7}$$

By concavity of U_0, we deduce (see e.g. [32]) the second duality relation in (5.4), i.e., U_0 is the Fenchel-Legendre transform of V_0.

Next, observe from the explicit expression of V_0 in (5.1), that V_0 is a strictly concave C^1 function on $(0, \infty)$, with $V_0'(0^+) = \infty$, $V_0'(\infty) = 0$. Then, for any fixed $\vartheta > 0$, there exists a unique $\lambda_\vartheta > 0$ that attains the infimum of $\lambda \in (0, \infty) \mapsto \lambda\vartheta - V_0(\lambda)$, characterized by $V_0'(\lambda_\vartheta) = \vartheta$, and explicitly given by

$$\lambda_\vartheta \quad = \quad \sqrt{\frac{\exp\left(R(\theta^*)T\right) - 1}{4\vartheta}}. \tag{5.8}$$

Relation (5.8) gives the explicit link between the variance risk in the robust Markowitz problem and the Lagrange multiplier in the robust mean-variance problem. This Lagrange multiplier λ is then interpreted as a risk-aversion parameter: the larger is λ_ϑ, the lower is the variance risk ϑ. From the duality relation (5.4), we then have:

$$V_0(\lambda_\vartheta) \quad = \quad \lambda_\vartheta\vartheta - U_0(\vartheta) \;=\; \inf_{\vartheta'>0}[\lambda_\vartheta\vartheta' - U_0(\vartheta')],$$

which means that ϑ attains the infimum of $\vartheta' \in (0, \infty) \mapsto \lambda_\vartheta\vartheta' - U_0(\vartheta')$. Since V_0 is strictly concave, its Fenchel-Legendre transform U_0 is also strictly concave (see e.g. [32]), and thus this infimum is unique. Recalling (5.7), this shows that $\vartheta = \vartheta_{\lambda_\vartheta} = \mathcal{R}(\alpha^{*,\lambda_\vartheta})$. Together with (5.6), we then obtain:

$$\begin{aligned} U_0(\vartheta) &= \lambda_\vartheta\vartheta - V_0(\lambda_\vartheta) \\ &= \lambda_\vartheta\mathcal{R}(\alpha^{*,\lambda_\vartheta}) - \left[\lambda_\vartheta\mathcal{R}(\alpha^{*,\lambda_\vartheta}) - \mathcal{E}(\alpha^{*,\lambda_\vartheta})\right] \;=\; \mathcal{E}(\alpha^{*,\lambda_\vartheta}), \end{aligned}$$

which proves that $\hat{\alpha}^\vartheta = \alpha^{*,\lambda_\vartheta}$ is a solution to the robust Markowitz problem $U_0(\vartheta)$, i.e., a robust efficient portfolio strategy given a worst-case variance risk $\vartheta > 0$. From (5.2), (5.3) and (5.8), we get the explicit form of the robust efficient frontier:

$$\begin{aligned} U_0(\vartheta) \;=\; \mathcal{E}(\hat{\alpha}^\vartheta) &= \bar{\rho}_T^{*,\lambda_\vartheta} \\ &= x_0 + \sqrt{\vartheta}\sqrt{\exp\left(R(\theta^*)T\right) - 1}, \quad \vartheta > 0 \\ &= x_0 + \sqrt{\mathcal{R}(\hat{\alpha}^\vartheta)}\sqrt{\exp\left(R(\theta^*)T\right) - 1}. \end{aligned} \tag{5.9}$$

To summarize the above discussion, we have the following result:

Theorem 5.1 Under **(IC)**, the efficient frontier of the robust Markowitz problem (2.5) is explicitly given by the relation (5.9).

The relation (5.9) determines explicitly the tradeoff between the worst-case mean (return) and worst-case variance (risk), and can be inverted: given an expected return level $m > x_0$, the risk that the robust investor can take is:

$$\hat{\vartheta}(m) \;=\; U_0^{-1}(m) \quad = \quad \frac{(m - x_0)^2}{\exp\left(R(\theta^*)T\right) - 1}, \quad m > x_0.$$

Notice that the robust efficient frontier (5.9) involves a square-root shape as in the classical efficient frontier in Markowitz problem, see e.g. [38].

Let us consider the Sharpe ratio for a portfolio strategy $\alpha \in \mathcal{A}$, defined by

$$\mathcal{S}(\alpha) \quad = \quad \frac{\mathbb{E}[X_T^\alpha] - x_0}{\sqrt{\mathrm{Var}(X_T^\alpha)}},$$

that is the excess of the expected return per unit of the standard deviation, evaluated under the true historical probability measure. By definition of the robust Markowitz problem, and from the relation (5.9), we have a lower bound for the Sharpe ratio of any robust efficient portfolio strategy $\hat{\alpha}^\vartheta$:

$$\mathcal{S}(\hat{\alpha}^\vartheta) \quad \geq \quad \frac{\mathcal{E}(\hat{\alpha}^\vartheta) - x_0}{\sqrt{\mathcal{R}(\hat{\alpha}^\vartheta)}} \quad = \quad \sqrt{\exp\big(R(\theta^*)T\big) - 1} \ =: \ \underline{\mathcal{S}}.$$

In other words, a robust investor can achieve a Sharpe ratio at least greater than $\underline{\mathcal{S}} > 0$, and this lower bound is robust to any model misspecification on the covariance matrix.

6 Robust Sharpe ratio vs model misspecification

In this section, we illustrate through two examples how robust mean-variance portfolio strategies may help to protect the investor from model misspecification, and can sometimes increase the Sharpe ratio for a specific choice of parameters.

6.1 A Heston-type stochastic volatility model

We consider a market with one risky asset, and assume that the true dynamics of the stock price is given by a Heston-type stochastic volatility model

$$\begin{cases} dS_t &= S_t(bdt + \sigma_t dW_t) \\ d\sigma_t^2 &= \kappa(\sigma_\infty^2 - \sigma_t^2)dt + \eta\sqrt{(\sigma_t^2 - \underline{\sigma}^2)(\bar{\sigma}^2 - \sigma_t^2)}d\tilde{W}_t \end{cases} \tag{6.1}$$

where $W, \tilde{W}$ are two Brownian motions under the real probability measure $\mathbb{P}$, with negative correlation ϱ representing the leverage effect, $\kappa > 0$, $\sigma_\infty \in [\underline{\sigma}, \bar{\sigma}]$, $0 < \underline{\sigma} \leq \bar{\sigma} < \infty$. Compared to the original Heston stochastic volatility model where the variance σ_t^2 follows a Cox-Ingersoll-Ross process, and is thus valued in $(0, \infty)$, we consider here a variation where the variance follows a Wright-Fisher dynamics, and is bounded, valued in $[\underline{\sigma}^2, \bar{\sigma}^2]$.

We now consider a simple investor who knows the drift b but specifies incorrectly the volatility by considering that it is equal to a constant $\tilde{\sigma}_0$. In other words, she/he believes that the stock price is governed by a Black-Scholes model of parameters $(b, \tilde{\sigma}_0)$. Therefore, from the result in [38] or as a particular case of our paragraph 4.1 when Θ is reduced to the singleton $\{\tilde{\sigma}_0^2\}$, the optimal mean-variance portfolio strategy of this "misspecified" investor with risk-aversion parameter $\lambda > 0$, and initial capital x_0 is given by:

$$\tilde{\alpha}_t \quad = \quad \frac{b}{\tilde{\sigma}_0^2}\Big[x_0 + \frac{1}{2\lambda}\exp\big(\frac{b^2}{\tilde{\sigma}_0^2}T\big) - \tilde{X}_t\Big], \quad 0 \leq t \leq T, \tag{6.2}$$

where $\tilde{X}_t$ is the wealth process with feedback strategy $\tilde{\alpha}$. Notice that the evolution of the wealth process $\tilde{X}$ under the real probability measure $\mathbb{P}$ is

$$d\tilde{X}_t \quad = \quad \tilde{\alpha}_t\frac{dS_t}{S_t} \quad = \quad \tilde{\alpha}_t b dt + \tilde{\alpha}_t\sigma_t dW_t,$$

which implies that its expected return under $\mathbb{P}$ is governed by

$$d\mathbb{E}[\tilde{X}_t] \quad = \quad b\mathbb{E}[\tilde{\alpha}_t]dt \ = \ \frac{b^2}{\tilde{\sigma}_0^2}\Big[x_0 + \frac{1}{2\lambda}\exp\big(\frac{b^2}{\tilde{\sigma}_0^2}T\big) - \mathbb{E}[\tilde{X}_t]\Big]dt.$$

where we used (6.2). Therefore, the excess expected return under $\mathbb{P}$ is explicitly given by:

$$\mathbb{E}[\tilde{X}_T] - x_0 \quad = \quad \frac{1}{2\lambda}\Big[\exp\big(\frac{b^2}{\tilde{\sigma}_0^2}T\big) - 1\Big].$$

The variance risk of $\tilde{X}_T$ under $\mathbb{P}$ is not explicit, but can be approximated by N Monte-Carlo simulations $(\tilde{X}^i)_{i=1,\ldots,N}$ of $\tilde{X}$ under $\mathbb{P}$ via:

$$\mathrm{Var}(\tilde{X}_T) \quad \simeq \quad \frac{1}{N-1}\sum_{i=1}^{N}\big(\tilde{X}_T^i - \mathbb{E}[\tilde{X}_T]\big)^2.$$

We can then compute the Sharpe ratio $\mathcal{S}(\tilde{\alpha}) = \frac{\mathbb{E}[\tilde{X}_T]-x_0}{\sqrt{\mathrm{Var}(\tilde{X}_T)}}$ for the “misspecified" investor.

The model parameters used in the simulations for the bounded Heston stochastic volatility model (6.1) are given in Table 1. We used the simulation method as in [9] for dealing with the discretization of the CIR process for the volatility, which means that when a volatility trajectory breachs the bounds $\underline{\sigma}$ or $\bar{\sigma}$, we project its value according to its closest neighbor on $[\underline{\sigma}, \bar{\sigma}]$.

We fix an investment horizon $T = 1$ year, a risk-aversion parameter $\lambda = 5$, and use $N = 500000$ simulations for each set of parameters.

On the other hand, let us consider a robust investor with risk-aversion parameter λ, initial capital x_0, who knows only the bounds $\underline{\sigma}$, $\bar{\sigma}$ of the volatility, and then follows a robust efficient portfolio strategy $\alpha^* = \alpha^{*,\lambda}$ given by

$$\alpha_t^* \quad = \quad \frac{b}{\bar{\sigma}^2}\Big[x_0 + \frac{1}{2\lambda}\exp\big(\frac{b^2}{\bar{\sigma}^2}T\big) - X_t^*\Big], \quad 0 \leq t \leq T.$$

Her/his excess expected return under $\mathbb{P}$ is then explicitly given by

$$\mathbb{E}[X_T^*] - x_0 \quad = \quad \frac{1}{2\lambda}\Big[\exp\big(\frac{b^2}{\bar{\sigma}^2}T\big) - 1\Big].$$

The variance risk of X_T^* under $\mathbb{P}$ is approximated by Monte-Carlo simulations of X^* under $\mathbb{P}$, and we then compute the Sharpe ratio $\mathcal{S}(\alpha^*) = \frac{\mathbb{E}[X_T^*]-x_0}{\sqrt{\mathrm{Var}(X_T^*)}}$ for the robust investor, which is known a priori to be larger than $\underline{\mathcal{S}} = \sqrt{\exp\big(\frac{b^2}{\bar{\sigma}^2}T\big) - 1}$. Notice that the optimal strategy of the robust investor corresponds to the optimal strategy of a simple investor with misspecified volatility $\bar{\sigma}$.

Table 2 and Figure 1 show the Sharpe ratios of the robust investor and of the simple investor when varying the misspecified volatility $\tilde{\sigma}_0$. Since the Sharpe ratios are computed by Monte-Carlo simulations, we also put in Table 2 a confidence interval. We see that the Sharpe ratio of the robust investor can perform noticeably better than the one of the simple investor who uses a misspecified volatility: this gap is all the more important as the misspecified volatility is far from the stationary value σ_∞ of the true volatility, for example when $\tilde{\sigma}_0 = \underline{\sigma}$. On the other hand, we notice that the Sharpe ratio of the simple investor is obviously equal to the one of the robust investor when the misspecified volatility $\tilde{\sigma}_0$ is

b	κ	η	σ_0	$\underline{\sigma}$	σ_∞	$\bar{\sigma}$	ρ
20%	2	1	30%	15%	30%	45%	-0.7

Table 1: Parameter values used in the bounded Heston stochastic volatility model.

$\tilde{\sigma}_0$	$\underline{\sigma}$	20%	σ_∞	$\bar{\sigma}$	50%
$\underline{\mathcal{S}}$	0.4673	0.4673	0.4673	0.4673	0.4673
$\mathcal{S}(\alpha^*)$	0.6831	0.6831	0.6831	0.6831	0.6831
95% confidence interval for $\mathcal{S}(\alpha^*)$	[0.6817,0.6844]	[0.6817,0.6844]	[0.6817,0.6844]	[[0.6817,0.6844]	[0.6817,0.6844]
$\mathcal{S}(\tilde{\alpha})$	0.1666	0.1839	0.64	0.6831	0.6809
95% confidence interval for $\mathcal{S}(\tilde{\alpha})$	[0.1662,0.1669]	[0.1835,0.1842]	[0.6387,0.6412]	[0.6817,0.6844]	[0.6795,0.6822]

Table 2:
Sharpe ratios $\mathcal{S}(\alpha^*)$ of the robust investor and $\mathcal{S}(\tilde{\alpha})$ of the investor for different misspecified values of $\tilde{\sigma}_0$ and parameter values as in Table 1.

equal to the worst-case scenario of volatility $\bar{\sigma}$. Let us mention that the outperformance of the robust strategies with respect to the misspecified Black-Scholes strategies is illustrated in our example for a specific choice of parameters. However, it may happen that when the vol-of-vol η is low, and/or the speed of mean-reversion κ is high, then the Black-Scholes investor using a misspecified volatility closed to the long-run volatility σ_∞ will perform better than the robust investor, as shown through Table 3 and Figure 2, where we have used $\eta = 0.25$ and $\kappa = 5$ while keeping the other parameters as specified in Table 1.

6.2 A stochastic correlation model

We consider a market with two risky assets, and motivated by the model in [8], assume that the true dynamics of the stock price $S = (S^1, S^2)$ is governed by

$$\begin{aligned} dS_t &= \mathrm{diag}(S_t)\big[bdt + \varsigma(\varrho_t)dW_t\big] \\ &= \begin{pmatrix} S_t^1\big[b_1dt + \sigma_1\sqrt{1-\varrho_t^2}dW_t^1 + \sigma_1\varrho_t dW_t^2\big] \\ S_t^2\big[b_2dt + \sigma_2 dW_t^2\big] \end{pmatrix}, \end{aligned} \tag{6.3}$$

where $b = (b_1, b_2)$, $\sigma_1 > 0$, $\sigma_2 > 0$ are known constants, and (ϱ_t) is a stochastic correlation process valued in $[0, \bar{\varrho}]$, with a known positive constant $\bar{\varrho} < 1$, and governed by a Wright-Fisher dynamics

$$d\varrho_t = \kappa(\varrho_\infty - \varrho_t)dt + \eta\sqrt{\varrho_t(\bar{\varrho} - \varrho_t)}d\tilde{W}_t, \tag{6.4}$$

$\tilde{\sigma}_0$	$\underline{\sigma}$	20%	σ_∞	$\bar{\sigma}$	50%
$\underline{\mathcal{S}}$	0.4673	0.4673	0.4673	0.4673	0.4673
$\mathcal{S}(\alpha^*)$	0.7135	0.7135	0.7135	0.7135	0.7135
95% confidence interval for $\mathcal{S}(\alpha^*)$	[0.7108,0.7136]	[0.7108,0.7136]	[0.7108,0.7136]	[0.7108,0.7136]	[0.7108,0.7136]
$\mathcal{S}(\tilde{\alpha})$	0.1581	0.5515	0.7282	0.7135	0.7069
95% confidence interval for $\mathcal{S}(\tilde{\alpha})$	[0.1578,0.1584]	[0.5503,0.5525]	[0.7273,0.7301]	[0.7108,0.7136]	[0.7075,0.7102]

Table 3:
Sharpe ratios $\mathcal{S}(\alpha^*)$ of the robust investor and $\mathcal{S}(\tilde{\alpha})$ of the investor for different misspecified values of $\tilde{\sigma}_0$ and parameter values as in Table 1 but with $\kappa = 5$, $\eta = 0.25$.

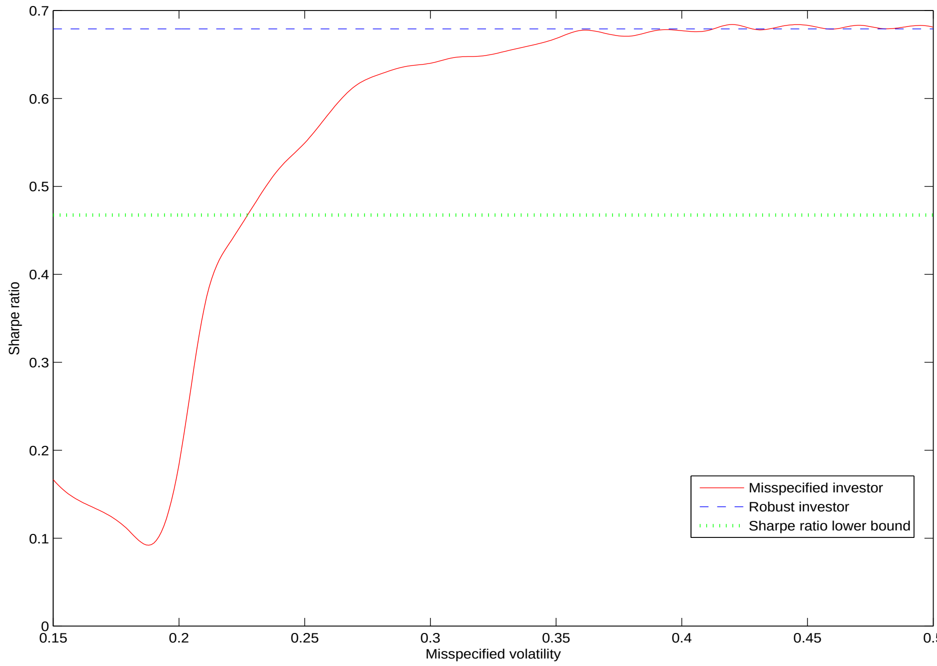

Figure 1: Sharpe ratio $\mathcal{S}(\tilde{\alpha})$ for different values of $\tilde{\sigma}_0$ with parameter values as in Table 1.

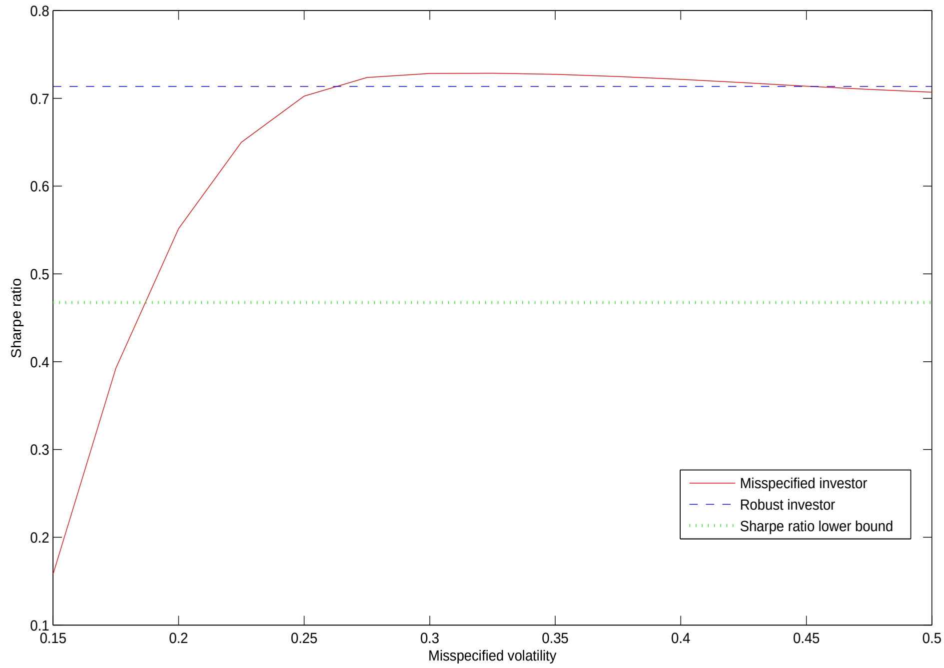

Figure 2: Sharpe ratio $\mathcal{S}(\tilde{\alpha})$ for different values of $\tilde{\sigma}_0$ with parameter values as in Table 1 but with $\kappa = 5$, $\eta = 0.25$.

where $\kappa \geq 0$, $\varrho_\infty \in [0, \bar{\varrho}]$, $\eta > 0$, and $\tilde{W}$ is a Brownian motion, assumed here for simplicity, to be independent of the two dimensional Brownian motion $W = (W^1, W^2)$ under the real probability measure $\mathbb{P}$.

We now consider a simple investor who knows the drifts b_i, the volatilities σ_i, hence the corresponding instantaneous Sharpe ratios $\beta_i = b_i/\sigma_i$, of the two assets $i = 1, 2$, but specifies incorrectly the correlation by considering that it is equal to a constant $\tilde{\varrho}_0 \in (-1, 1)$. Therefore, from the result in [38] or as a particular case of our paragraph 4.2 when Θ is reduced to the singleton $\tilde{\varrho}_0$, the optimal mean-variance portfolio strategy of this "misspecified" investor with risk-aversion parameter $\lambda > 0$, and initial capital x_0 is given by:

$$\tilde{\alpha}_t \;=\; \Big[x_0 + \frac{1}{2\lambda}\exp\big(\tilde{R}_0 T\big) - \tilde{X}_t\Big]\tilde{\Sigma}_0^{-1} b, \quad 0 \leq t \leq T,$$

where

$$\tilde{\Sigma}_0 \;:=\; \gamma(\tilde{\varrho}_0) \;=\; \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\tilde{\varrho}_0 \\ \sigma_1\sigma_2\tilde{\varrho}_0 & \sigma_2^2 \end{pmatrix}, \qquad \tilde{\Sigma}_0^{-1} b \;=\; \frac{1}{1-\tilde{\varrho}_0^2}\begin{pmatrix} \frac{\beta_1 - \beta_2\tilde{\varrho}_0}{\sigma_1} \\ \frac{\beta_2 - \beta_1\tilde{\varrho}_0}{\sigma_2} \end{pmatrix}$$

$$\tilde{R}_0 \;:=\; b^\intercal \tilde{\Sigma}_0^{-1} b \;=\; \frac{1}{1-\tilde{\varrho}_0^2}\big(\beta_1^2 + \beta_2^2 - 2\beta_1\beta_2\tilde{\varrho}_0\big),$$

and $\tilde{X}$ is the wealth process with feedback strategy $\tilde{\alpha}$, governed under the real probability measure $\mathbb{P}$ by

$$d\tilde{X}_t \;=\; \tilde{\alpha}_t^\intercal b dt + \tilde{\alpha}_t^\intercal \varsigma(\varrho_t) dW_t.$$

Its expected return under $\mathbb{P}$ is then governed by

$$d\mathbb{E}[\tilde{X}_t] \;=\; b^\intercal \mathbb{E}[\tilde{\alpha}_t] dt \;=\; \tilde{R}_0\Big[x_0 + \frac{1}{2\lambda}\exp\big(\tilde{R}_0 T\big) - \mathbb{E}[\tilde{X}_t]\Big] dt,$$

which gives the excess of expected return at T:

$$\mathbb{E}[\tilde{X}_T] - x_0 \;=\; \frac{1}{2\lambda}\big[\exp\big(\tilde{R}_0 T\big) - 1\big].$$

The variance risk of $\tilde{X}_T$ under $\mathbb{P}$ is not explicit, but can be approximated by N Monte-Carlo simulations $(\tilde{X}^i)_{i=1,\ldots,N}$ of $\tilde{X}$ under $\mathbb{P}$ via:

$$\mathrm{Var}(\tilde{X}_T) \;\simeq\; \frac{1}{N}\sum_{i=1}^{N}\big(\tilde{X}_T^i - \mathbb{E}[\tilde{X}_T]\big)^2.$$

We can then compute the Sharpe ratio $\mathcal{S}(\tilde{\alpha}) = \frac{\mathbb{E}[\tilde{X}_T] - x_0}{\sqrt{\mathrm{Var}(\tilde{X}_T)}}$ for the "misspecified" investor.

The model parameters used in the simulations of $\tilde{X}$ in the stochastic correlation model (6.3)-(6.4) are given in Table 4. Again, we used the simulation method as in [9] for dealing with the discretization of the Wright-Fisher process for the correlation, which means that when a correlation trajectory breachs the bounds 0 or $\bar{\varrho}$, we project its value according to its closest neighbor on $[0, \bar{\varrho}]$. We fix an investment horizon $T = 1$ year, a risk-aversion parameter $\lambda = 5$, and use $N = 500000$ simulations for each set of parameters.

On the other hand, let us consider a robust investor with risk-aversion parameter λ, initial capital x_0. By taking the parameters in Table 4, we notice that $\varrho_0^+ = \beta_2/\beta_1 \in [0, \bar{\varrho}]$, and thus from the result in Theorem 4.2, her/his robust efficient portfolio strategy $\alpha^* = \alpha^{*,\lambda}$ is given by

$$\alpha_t^* = \begin{pmatrix} \Big[x_0 + \frac{1}{2\lambda}\exp\big(\beta_1^2 T\big) - X_t^*\Big]\frac{b_1}{\sigma_1^2} \\ 0 \end{pmatrix}, \quad 0 \le t \le T,$$

and her/his wealth process X^* is governed under the real probability measure $\mathbb{P}$ by

$$\begin{aligned} dX_t^* &= (\alpha_t^*)^\intercal b dt + (\alpha_t^*)^\intercal \varsigma(\varrho_t) dW_t \\ &= \Big[x_0 + \frac{1}{2\lambda}\exp\big(\beta_1^2 T\big) - X_t^*\Big]\beta_1^2 dt + \Big[x_0 + \frac{1}{2\lambda}\exp\big(\beta_1^2 T\big) - X_t^*\Big]\beta_1\sqrt{1-\rho_t^2}dW_t^1 \\ &\quad + \Big[x_0 + \frac{1}{2\lambda}\exp\big(\beta_1^2 T\big) - X_t^*\Big]\beta_1\rho_t dW_t^2. \end{aligned} \tag{6.5}$$

The excess of expected return under $\mathbb{P}$ is explicitly given by

$$\mathbb{E}[X_t^*] - x_0 = \frac{1}{2\lambda}\big[\exp\big(\beta_1^2 T\big) - \exp\big(\beta_1^2(T-t)\big)\big], \quad 0 \le t \le T. \tag{6.6}$$

and we can actually compute explicitly in this case the variance risk of X_t^* under the real probability measure. Indeed, denoting by $Y_t^* = X_t^* - \mathbb{E}[X_t^*]$, we see from (6.5)-(6.6) that

$$dY_t^* = -\beta_1^2 Y_t^* dt + \big(\frac{1}{2\lambda}e^{\beta_1^2(T-t)} - Y_t^*\big)\big[\beta_1\sqrt{1-\rho_t^2}dW_t^1 + \beta_1\rho_t dW_t^2\big],$$

so that by Itô's formula, and taking expectation under $\mathbb{P}$:

$$d\mathbb{E}|Y_t^*|^2 = \big(-\beta_1^2\mathbb{E}|Y_t^*|^2 + \frac{\beta_1^2}{4\lambda^2}e^{2\beta_1^2(T-t)}\big)dt.$$

It follows that

$$\mathrm{Var}(X_t^*) = \mathbb{E}|Y_t^*|^2 = \frac{e^{2\beta_1^2(T-t)}}{4\lambda^2}\big(e^{\beta_1^2 t} - 1\big), \quad 0 \le t \le T.$$

In particular, we deduce the Sharpe ratio of the robust investor:

$$\mathcal{S}(\alpha^*) = \frac{\mathbb{E}[X_T^*] - x_0}{\sqrt{\mathrm{Var}(X_T^*)}} = \sqrt{\exp\big(\beta_1^2 T\big) - 1} = \underline{\mathcal{S}},$$

which means that in the case when $\varrho_0^+ \in [0, \bar{\varrho}]$, the Sharpe ratio attains its lower bound $\underline{\mathcal{S}}$. Notice that the optimal strategy of the robust investor is equal to the optimal strategy of a simple investor with misspecified correlation $\tilde{\varrho}_0 = \varrho_0^+$.

Table 5 and Figure 3 show the Sharpe ratios of the robust investor and of the simple investor when varying the misspecified correlation $\tilde{\varrho}_0$ (since the Sharpe ratio of the simple investor is computed by Monte-Carlo simulations, we also put in Table 5 its confidence interval at level 95%). They obviously coincide by definition when the misspecified correlation $\tilde{\varrho}_0$ is equal to ϱ_0^+ (here equal to $\beta_2/\beta_1 = 1/3$). On the other hand, we see that the Sharpe ratio of the robust investor may perform worse than the one of the simple investor, especially when the misspecified correlation $\tilde{\varrho}_0$ is close from the true stationary correlation ϱ_∞ (and when the vol-of-correl η is low, and/or the speed of mean-reversion κ is high), but performs better when $\tilde{\varrho}_0$ is smaller than ϱ_0^+.

β_1	β_2	ϱ_0	$\bar{\varrho}$	κ	ϱ_∞	η
1.5	0.5	0.7	0.95	5	0.7	20%

Table 4: Parameter values used in the stochastic correlation model

$\tilde{\varrho}_0$	0.1	$\varrho_0^+ = 1/3$	ϱ_∞	0.8
$\mathcal{S}(\alpha^*) = \underline{\mathcal{S}}$	2.9134	2.9134	2.9134	2.9134
$\mathcal{S}(\tilde{\alpha})$	2.1085	2.9134	4.2008	5.6798
95% confidence interval for $\mathcal{S}(\tilde{\alpha})$	[2.1043,2.1126]	[2.9076,2.9191]	[4.1925,4.2090]	[5.6686,5.6909]

Table 5:
Sharpe ratios $\mathcal{S}(\alpha^*)$ of the robust investor and $\mathcal{S}(\tilde{\alpha})$ of the investor for different misspecified values of $\tilde{\varrho}_0$.

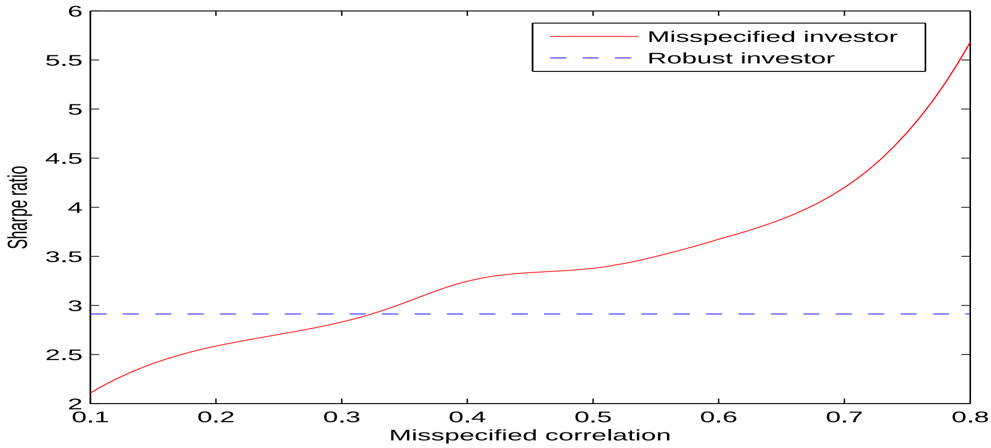

Figure 3: Sharpe ratio $\mathcal{S}(\tilde{\alpha})$ for different values of $\tilde{\varrho}_0$

A Appendix: Differentiability on Wasserstein space and Itô's formula

We first recall the notion of derivative with respect to a probability measure, as introduced by P.L. Lions in his course at Collège de France, and detailed in the lecture notes [6].

This notion is based on the lifting of functions $u : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ into functions U defined on $L^2(\mathcal{G};\mathbb{R}) = L^2(\Omega,\mathcal{G},\mathbb{P};\mathbb{R})$ (the set of square-integrable random variables on some probability space $(\Omega,\mathcal{G},\mathbb{P})$) by $U(X) = u(\mathcal{L}(X))$, where $\mathcal{L}(X)$ is the law of X on $(\Omega,\mathcal{G},\mathbb{P})$. We say that u is differentiable (resp. $\mathcal{C}^1$) on $\mathcal{P}_2(\mathbb{R})$ if the lift U is Fréchet differentiable (resp. Fréchet differentiable with continuous derivatives) on $L^2(\mathcal{G};\mathbb{R})$. In this case, the Fréchet derivative $[DU](X)$, which is identified as an element $DU(X)$ of $L^2(\mathcal{G};\mathbb{R})$ by Riesz' theorem through the relation: $[DU](X)(Y) = \mathbb{E}[DU(X)Y]$, can be represented as

$$DU(X) \quad = \quad \partial_\mu u(\mathcal{L}(X))(X), \tag{A.1}$$

for some function $\partial_\mu u(\mathcal{L}(X)) : \mathbb{R} \to \mathbb{R}$, which is called derivative of u at $\mu = \mathcal{L}(X)$. Moreover, $\partial_\mu u(\mu) \in L^2(\mu)$ for $\mu \in \mathcal{P}_2(\mathbb{R}) = \{\mathcal{L}(X), X \in L^2(\mathcal{G};\mathbb{R})\}$. We say that u is partially $\mathcal{C}^2$ if it is $\mathcal{C}^1$, and one can find, for any $\mu \in \mathcal{P}_2(\mathbb{R})$, a continuous version of the mapping $x \in \mathbb{R} \mapsto \partial_\mu u(\mu)(x)$, such that the mapping $(\mu,x) \in \mathcal{P}_2(\mathbb{R}) \times \mathbb{R} \mapsto \partial_\mu u(\mu)(x)$ is continuous at any point (μ,x) such that $x \in \mathrm{Supp}(\mu)$, and if for any $\mu \in \mathcal{P}_2(\mathbb{R})$, the mapping $x \in \mathbb{R} \mapsto \partial_\mu u(\mu)(x)$ is differentiable, its derivative being jointly continuous at any point (μ,x) such that $x \in \mathrm{Supp}(\mu)$. The gradient is then denoted by $\partial_x\partial_\mu u(\mu)(x)$.

For example, consider a linear function: $u(\mu) = \int \varphi(x)\mu(dx)$. Its lifted function is $U(X) = \mathbb{E}[\varphi(X)]$, whose Fréchet derivative is given by: $[DU](X)(Y) = \mathbb{E}[D_x\varphi(X).Y]$, from which we see that $\partial_\mu u(\mu) = D_x\varphi$, and thus $\partial_x\partial_\mu u(\mu) = D_x^2\varphi$. In particular, when $\varphi(x) = x$, i.e., $u(\mu) = \bar\mu := \int x\mu(dx)$, then $\partial_\mu u(\mu) = 1$. Another example used in this paper is a function $u(\mu) = \mathrm{Var}(\mu) := \int (x-\bar\mu)^2\mu(dx)$. In this case, its lifted function is $U(X) = \mathrm{Var}(X)$, from which we see that $DU(X) = 2(X - \mathbb{E}[X])$, and thus $\partial_\mu u(\mu)(x) = 2(x-\bar\mu)$, $\partial_x\partial_\mu u(\mu)(x) = 2$.

We next recall a chain rule (or Itô's formula) for functions defined on $\mathcal{P}_2(\mathbb{R})$, proved independently in [5] and [7]. Let us consider a real-valued Itô process

$$dX_t \quad = \quad b_t dt + \sigma_t dW_t, \quad X_0 \in L^2(\mathcal{F}_0;\mathbb{R}),$$

where (b_t) and (σ_t) are progressively measurable processes with respect to the filtration generated by the d-dimensional Brownian motion W, valued respectively in $\mathbb{R}$ and $\mathbb{R}^{1\times d}$, and satisfying the integrability condition: $\mathbb{E}\Big[\int_0^T |b_t|^2 + |\sigma_t|^2 dt\Big] < \infty$. Let u be a partially $\mathcal{C}^2$ function on $\mathcal{P}_2(\mathbb{R})$. Then, for all $t \in [0,T]$,

$$\begin{aligned} u(\mathcal{L}(X_t)) \quad = \quad u(\mathcal{L}(X_0)) \; + \; & \int_0^t \mathbb{E}\big[\partial_\mu u(\mathcal{L}(X_s))(X_s) b_s \\ & + \frac{1}{2}\partial_x\partial_\mu u(\mathcal{L}(X_s))(X_s)|\sigma_s|^2\big] ds. \end{aligned} \tag{A.2}$$

References

[1] Ait-Sahalia Y., Fan J. and D. Xiu (2010): "High Frequency Covariance Estimates with Noisy and Asynchronous Financial Data", *Journal of American Statistical Association*, **105**, 1504-1517.

[2] Andersson D. and B. Djehiche (2011): "A maximum principle for SDEs of mean-field type", *Applied Mathematics and Optimization*, **63**(3), 341-356.

[3] Avellaneda M., Levy A. and A. Paras (1995): "Pricing and Hedging Derivative Securities in Markets with Uncertain Volatilities", *Applied Mathematical Finance*, **2**, 73-88.

[4] Bensoussan A., Frehse J. and P. Yam (2015): “The Master equation in mean-field theory", *Journal de Mathématiques Pures et Appliquées*, **103**(6), 1441-1474.

[5] Buckdahn R., Li J., Peng S. and C. Rainer (2014): “Mean-field stochastic differential equations and associated PDEs", http://arxiv.org/abs/1407.1215, to appear in the *Annals of Probability.*

[6] Cardaliaguet P. (2012): “Notes on mean field games", Notes from P.L. Lions lectures at Collège de France, https://www.ceremade.dauphine.fr/cardalia/MFG100629.pdf

[7] Chassagneux J.F., Crisan D. and F. Delarue (2015): “A probabilistic approach to classical solutions of the master equation for large population equilibria", http://arxiv.org/pdf/1411.3009.pdf

[8] Chiu M.C. and H.Y. Wong (2014): “Mean-variance portfolio selection with correlation risk", *Journal of Computational and Applied Mathematics*, **263**, 432-444.

[9] Deelstra G. and F. Delbaen (1998): “Convergence of Discretized Stochastic (Interest Rate) Processes with Stochastic Drift Term", *Appl. Stochastic Models Data Anal.*, **14**, 77-84.

[10] Denis L., and C. Martini (2006): "A Theoretical Framework for the Pricing of Contingent Claims in the Presence of Model Uncertainty", *Annals of Applied Probability*, **16**, 827-852.

[11] Epstein L.G. and S. Ji (2013): “Ambiguous Volatility and Asset Pricing in Continuous Time", *Review of Financial Studies*, **26**, 1740-1786.

[12] Fabozzi F., Huang D. and G. Zhou (2010): “Robust portfolios: contributions from operations research and finance", *Annals of Operation Research*, **176**, 191-220.

[13] Fisher M. and G. Livieri (2015): “Continuous time mean-variance portfolio optimization through the mean-field approach", to appear in *ESAIM Probability and Statistics.*

[14] Fouque J.P., Pun C.S. and H.Y. Wong (2015): “Portfolio optimization with ambiguous correlation and stochastic volatilities", to appear in *SIAM Journal on Control and Optimization.*

[15] Hansen L.P. and T. J. Sargent (2001): “Robust control and model uncertainty", *American Economic Review*, **91**, 60-66.

[16] Hayashi T. and N. Yoshida (2005): “On covariance estimation of non-synchronously observed diffusion processes", *Bernoulli*, **11**, 359-379.

[17] Ismail A. and H. Pham (2016): “Robust Markowitz mean-variance portfolio selection under ambiguous volatility and correlation", arXiv:1610.06805v1

[18] Jagannathan R. and T. Ma (2003): “Risk reduction in large portfolios: Why imposing the wrong constraints helps", *Journal of Finance*, **58**, 1651-1684.

[19] Jourdain B., Méléard S. and W.A. Woyczynski (2008): “Nonlinear SDEs driven by Lévy processes and related PDEs", *ALEA Lat. Am. J. Probab. Math. Stat.*, **4**, 1-29.

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

[21] Karandikar R. (1995): "On Pathwise Stochastic Integration", *Stochastic Processes and their Applications*, **57**, 11-18.

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

[23] Liu J. and X. Zeng (2016): “Correlation ambiguity", available at https://ssrn.com/abstract=2692692

[24] Lyons, T. (1995): "Uncertain Volatility and the Risk Free Synthesis of Derivatives", *Applied Mathematical Finance*, **2**, 117-133.

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

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

[27] Nutz M. (2012): "Pathwise Construction of Stochastic Integrals", *Electronic Communications in Probability*, **17**, 1-7.

[28] Peng S. (2006): "G-Expectation, G-Brownian Motion and Related Stochastic Calculus of Itô type", in The Abel Symposium 2005, Abel Symposia 2, ed. by Benth et. al., Springer-Verlag, 541-567.

[29] Pham H. and X. Wei (2016): "Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics", arXiv:1604.04057, to appear in *SIAM Journal on Control and Optimization.*

[30] Pinar M. (2016): "On robust mean-variance portfolios", *Optimization: Journal of Mathematical Programming and Operations Research*, **65**, 1039-1048.

[31] Rebonato R. and P. Jäckel (1999): "The most general methodology to create a valid correlation matrix for risk management and option pricing purposes", *Applied Economics Letters*, **19**, 1767-1768.

[32] Rockafellar T. (1970: Convex analysis, Princeton University Press.

[33] Sznitman A.S. (1989): Topics in propagation of chaos, in *Lecture Notes in Mathematics*, Springer, **1464**, 165-251.

[34] Soner, H.M., Touzi, N., Zhang, J. (2011): "Quasi-sure stochastic analysis through aggregation", *Electronic Journal of Probability*, **16**, 1844-1879.

[35] Soner H.M., Touzi N. and J. Zhang (2012): "Wellposedness of second order BSDEs", *Probability Theory and Related Fields*, **153**,(1), 149-190.

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

[37] Wagner D. (1980): "Survey of measurable selection theorem: an update", Lect. Notes in Math., 794, Springer Verlag.

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