

Bayesian learning for the Markowitz portfolio selection problem

Carmine de Franco, Johann Nicolle, Huyên Pham

▶ **To cite this version:**

Carmine de Franco, Johann Nicolle, Huyên Pham. Bayesian learning for the Markowitz portfolio selection problem. 2018. hal-01923917

HAL Id: hal-01923917
https://hal.science/hal-01923917

Preprint submitted on 15 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bayesian learning for the Markowitz portfolio selection problem *

Carmine De Franco† Johann Nicolle‡ Huyên Pham §

November 15, 2018

Abstract

We study the Markowitz portfolio selection problem with unknown drift vector in the multidimensional framework. The prior belief on the uncertain expected rate of return is modeled by an arbitrary probability law, and a Bayesian approach from filtering theory is used to learn the posterior distribution about the drift given the observed market data of the assets. The Bayesian Markowitz problem is then embedded into an auxiliary standard control problem that we characterize by a dynamic programming method and prove the existence and uniqueness of a smooth solution to the related semi-linear partial differential equation (PDE). The optimal Markowitz portfolio strategy is explicitly computed in the case of a Gaussian prior distribution. Finally, we measure the quantitative impact of learning, updating the strategy from observed data, compared to non-learning, using a constant drift in an uncertain context, and analyze the sensitivity of the value of information w.r.t. various relevant parameters of our model.

Key Words: Bayesian learning, optimal portfolio, Markowitz problem, portfolio selection.

1 Introduction

Portfolio selection is a core problem in mathematical finance and investment management. Its purpose is to choose the best portfolio according to a criterion of optimality. The mean-variance optimization provides a criterion of optimality considering the best portfolio as the one that maximizes an expected level of return given a certain level of risk, the one that the investor can bear, or conversely, minimizes the risk of a portfolio given an expected level of return. Markowitz (1952) pioneered modern portfolio theory by settling the basic concepts. This conceptual framework, still widely used in the industry, leads to the efficient frontier principle which exhibits an intuitive relationship between risk and return.

Later, portfolio selection theory was extended several times to encompass multi-period problems, in discrete time by Samuelson (1969) and in continuous time by Merton (1969). Karatzas et al. (1987) made a decisive step forward when they solved Merton's problem for a large class of smooth utility functions using dual martingale methods under the no-bankruptcy constraint.

Originally, the literature on portfolio selection assumed that the parameters of the model, drifts and volatilities, are known and constant. This assumption raises the question

*This work is issued from a CIFRE collaboration between OSSIAM and LPSM.

†OSSIAM, E-mail: carmine.de-franco@ossiam.com

‡OSSIAM and LPSM, Université Paris Diderot, E-mail: johann.nicolle@ossiam.com

§LPSM, Université Paris Diderot, E-mail: pham@lpsm.paris

of estimating these parameters. Typically, the parameters are estimated from past data and are fixed once and for all. This convenient solution does not look realistic in practice since it does not adapt to changing market conditions. Moreover, as Merton (1980) among others showed, estimates of variances and covariances of the assets are more accurate than the estimates of the means. Indeed, he demonstrated the slow convergence of his estimators of the instantaneous expected return in a log-normal diffusion price model. Later, Best and Grauer (1991) argued that mean-variance optimal portfolios are very sensitive to the level of expected returns. As a consequence, estimating the expected return is not only more complicated than estimating the variance/covariance, but also a wrong estimation of the expected return can result in a very suboptimal portfolio a posteriori.

To circumvent these issues, an extensive literature incorporating parameters uncertainty in portfolio analysis has emerged, see for example Barry (1974) and Klein and Bawa (1976) and methods using Bayesian statistics have been developed, see Frost and Savarino (1986), Aguilar and West (2000), Avramov and Zhou (2010) and Bodnar et al. (2017). In particular, the Black and Litterman (1992) model, based on economic arguments and equilibrium relations, provides more stable and diversified portfolios than simple mean-variance optimization. Based upon the Markowitz problem and the capital asset pricing model (CAPM), this model remains static and cannot benefit from the flow of information fed by the market prices of the assets. This loss of information is detrimental in the optimization process, since it does not allow for an update of the model given the most recent available information.

Consequently, research then focused on taking advantage of the latest information conveyed by the prices. It explains the subsequent growing literature on filtering and learning techniques in a partial information framework, see Lakner (1995), Lakner (1998), Rogers (2001) and Cvitanić et al. (2006). The most noticeable research involving optimization and Bayesian learning techniques is by Karatzas and Zhao (2001). Using martingale methods, they computed the optimal portfolio allocation for a large class of utility functions in the case of an unknown drift and Gaussian asset returns. Guéant and Pu (2017) extend the previous results to take into account both the liquidity and the expected returns of the assets, coupling Bayesian learning and dynamic programming techniques while addressing optimal portfolio liquidation and transition problems. Recently, Bauder et al. (2018) suggest to deal with the Markowitz problem using the posterior predictive distribution. A Bayesian efficient frontier is also derived and proved to outperform the overoptimistic sample efficient frontier.

In this paper, we consider an investor who is willing to invest in an optimal portfolio in the Markowitz sense for a finite given time horizon. The investment universe consists of assets for which we assume the covariance matrix known and constant. The drift vector is uncertain and assumed to have a prior probability distribution.

We contribute to the literature of optimal portfolio selection by formulating and solving, in the multidimensional case, the Markowitz problem in the case of an uncertain drift modeled by a prior distribution. This a priori time inconsistent problem cannot be tackled by convex duality and martingale method as in Karatzas and Zhao (2001). Instead, we adapt the methodology of Zhou and Li (2000) to our Bayesian learning framework, and show how the Bayesian-Markowitz problem can be embedded into an auxiliary standard control problem that we study by the dynamic programming approach. Optimal strategies are then characterized in terms of a smooth solution to a semi-linear partial differential equation, and we highlight the effect of Bayesian learning compared to classical strategies based on constant drift. Although prior conjugates are widely used in the literature, here we extend to the multidimensional case, a recent result by Ekstrom and Vaicenavicius

(2016) which enables us to use any prior distribution with invertible covariance matrix to model the uncertain drift allowing for a wide range of investors' strategies. In the particular case of a multidimensional Gaussian prior conjugate, we are able to exhibit a closed-form analytical formula.

Next, we measure the benefit of learning on the Markowitz strategy. To do so, we compare the Bayesian learning strategy to the subsequently called non-learning strategy which considers the coefficients of the model, especially the drift, constant. The Bayesian learning strategy is characterized by the fact that it uses, as information, the updated market prices of the assets to adjust the investment strategy and reaches optimality in the Markowitz sense. On the other hand, the non-learning strategy is a typical constant parameter strategy that estimates once and for all its unknown parameters with past data at time 0, leading to sub-optimal solutions to the Markowitz problem because it misses the most recent information. The benefit of learning, measured as the difference between the Sharpe ratios of the portfolios based on the Bayesian learning and the non-learning strategies, is called the value of information in the sequel. We then analyze in a one-dimensional model the sensitivities of the value of information w.r.t. four essential variables: drift volatility, Sharpe ratio of the asset, time and investment horizon.

The paper is organized as follows: Section 2 depicts the Markowitz problem in our framework of drift uncertainty modeled by a prior distribution and the steps to turn it into a solvable problem. Section 3 introduces the Bayesian learning framework and the methodology to consider any prior with positive definite covariance matrices. Section 4 provides the main results and some specific examples where some computations are possible analytically. Section 5 concerns the value of information and its sensitivity to different parameters of an unidimensional model: the volatility of the drift, the sharpe ratio of the asset, the time, and the investment horizon. Section 6 concludes this paper.

2 Markowitz problem with prior law on the uncertain drift

We consider a financial market on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$ that satisfies the usual conditions, and on which is defined a standard n-dimensional Brownian motion W. The market consists of n risky assets that are continuously traded. The assets prices process S is defined as:

$$\left\{\begin{array}{rcl} dS_t &=& \mathrm{diag}(S_t)\big(Bdt+\sigma dW_t\big), \quad t\in[0,T], \\ S_0 &=& s_0\in\mathbb{R}^n, \end{array}\right. \tag{2.1}$$

where T is the investment horizon, B is a random vector in $\mathbb{R}^n$, independent of W, with distribution μ and such that $\mathbb{E}[|B|^2]<\infty$. The prior law of B represents the subjective beliefs of the investor about the likelihood of the different values that B might take.

In the sequel, we consider the following assumptions satisfied:

Assumption 2.1. *The covariance matrix* $\mathrm{Cov}\,(B)$ *of the random vector* B *is positive definite,*

Assumption 2.2. *The* $n\times n$ *volatility matrix* σ *is known and invertible, and we denote* $\Sigma=\sigma\sigma^\intercal$.

We write $\mathbb{F}^{\mathbb{S}}=\{\mathcal{F}_t^S\}_{t\geq 0}$ the filtration generated by the price process S augmented by the null sets of $\mathcal{F}$. The filtration $\mathbb{F}^S$ represents the only available source of information: the investor observes the price process S but not the drift.

We denote by $\alpha = (\alpha_t)_{0\leq t\leq T}$ an admissible investment strategy representing the amount invested in each of the n risky assets. It is a n-dimensional $\mathbb{F}^{\mathbb{S}}$-progressively measurable process, valued in $\mathbb{R}^n$, and satisfying the following integrability condition

$$\mathbb{E}\Big[\int_0^{\intercal} |\alpha_t|^2 dt\Big] < \infty.$$

We define $\mathcal{A}$ as the set of all admissible investment strategies.

The evolution of the self-financing wealth process X^α, given $\alpha \in \mathcal{A}$ and an initial wealth $x_0 \in \mathbb{R}$, is given by:

$$dX_t^\alpha = \alpha_t^\intercal \mathrm{diag}(S_t)^{-1} dS_t = \alpha_t^\intercal (Bdt + \sigma dW_t), \quad 0 \leq t \leq T, \quad X_0^\alpha = x_0. \tag{2.2}$$

The Markowitz problem is then formulated as :

$$U_0(\vartheta) \;:=\; \sup_{\alpha\in\mathcal{A}} \Big\{\mathbb{E}[X_T^\alpha] : \mathrm{Var}(X_T^\alpha) \leq \vartheta\Big\}, \tag{2.3}$$

where $\vartheta > 0$ is the variance budget the investor can allow on her portfolio. The value function U_0 is then the expected optimal terminal wealth value the investor can achieve given her variance constraint. It is very standard to transform the initial problem (2.3) into a so called mean-variance problem, where the constraint becomes part of the objective function:

$$V_0(\lambda) \;:=\; \inf_{\alpha\in\mathcal{A}} \big[\lambda \mathrm{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha]\big], \quad \lambda > 0. \tag{2.4}$$

The well-known connection between the Markowitz and mean-variance problem is stated in the following lemma.

Lemma 2.3. *We have the duality relation*

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

Furthermore, if $\alpha^{*,\lambda}$ is an optimal control for problem (2.4), then $\hat{\alpha}^\vartheta := \alpha^{*,\lambda(\vartheta)}$ is an optimal control for problem $U_0(\vartheta)$ in (2.3), with $\lambda(\vartheta) := \arg\min_{\lambda>0} \big[\lambda\vartheta - V_0(\lambda)\big]$, and $\mathrm{Var}(X_T^{\hat{\alpha}^\vartheta}) = \vartheta$.

The proof of this result is standard and detailed in A.

Because of the variance term, mean-variance problem (2.4) does not fit into a classical time consistent control problem. To circumvent this issue, we adopt a similar approach as in Zhou and Li (2000) in order to embed it into an auxiliary control problem, for which we can use the dynamic programming method.

Lemma 2.4. *The mean-variance optimization problem (2.4) can be written as follows:*

$$V_0(\lambda) \;=\; \inf_{\gamma\in\mathbb{R}} \big[\tilde{V}_0(\lambda,\gamma) + \lambda\gamma^2\big] \tag{2.5}$$

where

$$\tilde{V}_0(\lambda,\gamma) \;:=\; \inf_{\alpha\in\mathcal{A}} \big[\lambda\mathbb{E}[(X_T^\alpha)^2] - (1+2\lambda\gamma)\mathbb{E}[X_T^\alpha]\big]. \tag{2.6}$$

Moreover, if $\tilde{\alpha}^{\lambda,\gamma}$ *is an optimal control for* (2.6) *and* $\gamma^*(\lambda) := \text{argmin}_{\gamma\in\mathbb{R}}\big[\tilde{V}_0(\lambda,\gamma) + \lambda\gamma^2\big]$ *then* $\alpha^{*,\lambda} := \tilde{\alpha}^{\lambda,\gamma^*(\lambda)}$ *is an optimal control for* $V_0(\lambda)$ *in* (2.4).

Furthermore, the following equality holds true

$$\gamma^*(\lambda) \;=\; \mathbb{E}[X_T^{\alpha^{*,\lambda}}]. \tag{2.7}$$

The proof of this lemma follows Zhou and Li (2000) and is provided in B. As we shall see in Section 4, problem (2.6) is well suited for dynamic programming.

We end this section by recalling the solution to the Markowitz problem in the case of constant known drifts and volatilities, which will serve as a benchmark to refer to.

Remark 2.5 (Case of constant known drift)**.** In the case where the drift vector b_0 and the volatility matrix σ are known and constant, we have the following results :

- The value function of the mean-variance problem with initial wealth x_0 is given by

$$V_0(\lambda) \;=\; v^\lambda(0,x_0) = -\frac{1}{4\lambda}\Big(e^{|\sigma^{-1}b_0|^2T} - 1\Big) - x_0$$

with

$$v^\lambda(t,x) \;=\; \lambda e^{-|\sigma^{-1}b_0|^2(T-t)}\,(x-x_0)^2 - \frac{1}{4\lambda}\Big(e^{|\sigma^{-1}b_0|^2(T-t)} - 1\Big) - x.$$

The corresponding optimal portfolio strategy for $V_0(\lambda)$ is given in feedback form by

$$\alpha_t^{*,\lambda} \;=\; a_0^\lambda(t, X_t^*),$$

where

$$a_0^\lambda(t,x) \;:=\; \Big(x_0 - x + \frac{e^{|\sigma^{-1}b_0|^2T}}{2\lambda}\Big)\Sigma^{-1}b_0$$

and X_t^* is the optimal wealth obtained with the optimal feedback strategy.

- The optimal portfolio strategy of the Markowitz problem $U_0(\vartheta)$ is then given by

$$\hat{\alpha}_t^\vartheta \;=\; \alpha_t^{*,\lambda_0(\vartheta)}$$

where

$$\lambda_0(\vartheta) \;:=\; \sqrt{\frac{e^{|\sigma^{-1}b_0|^2T} - 1}{4\vartheta}}.$$

Moreover, the solution to the Markowitz problem is

$$U_0(\vartheta) \;=\; x_0 + \sqrt{\vartheta\big(e^{|\sigma^{-1}b_0|^2T} - 1\big)}.$$

See for instance Zhou and Li (2000) for more details. ◊

3 Bayesian learning

Since the investor does not observe the assets drift vector B, she needs to have a subjective belief on its potential value. It is represented as a prior probability distribution μ on $\mathbb{R}^n$, assumed to satisfy for some $a > 0$,

$$\int_{\mathbb{R}^n} e^{a|b|^2} \mu(db) < \infty.$$

The prior probability distribution will then learn and infer the value of the drift from observable samples of the assets prices. Using filtering theory, and following in particular the recent work by Ekstrom and Vaicenavicius (2016) that we extend to the multi-dimensional case, we compute the posterior probability distribution of B given the observed assets prices.

Let us first introduce the $\mathbb{R}^n$-valued process $Y_t = \sigma^{-1}Bt + W_t$, which clearly generates the same filtrations as S, and write the observation process

$$S_t^i \;=\; S_0^i \exp\Big(\sigma^i Y_t - \frac{|\sigma^i|^2}{2}t\Big), \quad i = 1, \ldots, n, \;\; 0 \le t \le T,$$

where σ^i denotes the i-th line of the matrix σ.

Equivalently, we can express Y in terms of S as: $Y_t = h(t, S_t)$ where $h : [0,T] \times (0,\infty)^n \to \mathbb{R}^n$ is defined by

$$h(t,s) \;:=\; \sigma^{-1} \begin{pmatrix} \ln\left(\frac{s^1}{S_0^i}\right) + \frac{|\sigma^1|^2}{2}t \\ \vdots \\ \ln\left(\frac{s^n}{S_0^i}\right) + \frac{|\sigma^n|^2}{2}t \end{pmatrix}, \quad s = (s^1, \ldots, s^n) \in (0,\infty)^n.$$

The following result gives the conditional distribution of B given observations of the assets market prices in terms of the current value of Y. We refer to Proposition 3.16 in Bain and Crisan (2009) as well as Karatzas and Zhao (2001) and Pham (2008) for a proof.

Lemma 3.1. *Let $g : \mathbb{R}^n \to \mathbb{R}^n$ satisfying $\int_{\mathbb{R}^n} |g(b)|\mu(db) < \infty$. Then*

$$\mathbb{E}\big[g(B)|\mathcal{F}_t^S\big] \;=\; \mathbb{E}[g(B)|Y_t] \;=\; \frac{\int_{\mathbb{R}^n} g(b) e^{<\sigma^{-1}b, Y_t> - \frac{|\sigma^{-1}b|^2}{2}t} \mu(db)}{\int_{\mathbb{R}^n} e^{<\sigma^{-1}b, Y_t> - \frac{|\sigma^{-1}b|^2}{2}t} \mu(db)}.$$

From Lemma 3.1, the conditional distribution $\mu_{t,y}$ of B given $Y_t = y$ is given by

$$\mu_{t,y}(db) \;=\; \frac{e^{<\sigma^{-1}b, y> - \frac{|\sigma^{-1}b|^2}{2}t} \mu(db)}{\int_{\mathbb{R}^n} e^{<\sigma^{-1}b, y> - \frac{|\sigma^{-1}b|^2}{2}t} \mu(db)}, \tag{3.1}$$

and the posterior predictive mean of the drift is

$$\hat{B}_t \;:=\; \mathbb{E}[B|\mathcal{F}_t^S] \;=\; \mathbb{E}[B|Y_t] \;=\; f(t, Y_t),$$

where

$$f(t,y) \;:=\; \int_{\mathbb{R}^n} b\mu_{t,y}(db) \;=\; \frac{\int_{\mathbb{R}^n} b e^{<\sigma^{-1}b, y> - \frac{|\sigma^{-1}b|^2}{2}t} \mu(db)}{\int_{\mathbb{R}^n} e^{<\sigma^{-1}b, y> - \frac{|\sigma^{-1}b|^2}{2}t} \mu(db)}. \tag{3.2}$$

The following result shows some useful properties regarding the function f.

Lemma 3.2. *For each $t \in [0,T]$, the function $f_t : \mathbb{R}^n \to \mathbb{R}^n$ defined as $f_t(y) := f(t,y)$, where f is given in (3.2) is invertible when restricted to its image $\mathcal{B}_t := f_t(\mathbb{R}^n)$, and we have*

$$\nabla_y f(t,y) \;=\; \mathrm{Cov}\left(B\,|Y_t = y\right)\left(\sigma^{-1}\right)^{\intercal}, \quad y \in \mathbb{R}^n.$$

and the conditional covariance matrix of B given $Y_t = y$ is positive definite.

Proof. The vector-valued function $y \mapsto f_t(y)$ is differentiable, hence we can compute the matrix function $\nabla_y f$ element-wise,

$$\begin{aligned}
\left[\partial_{y_j} f_t(y)\right]_i &= \int_{\mathbb{R}^n} [\sigma^{-1}b]_j b_i \mu_{t,y}(db) - \int_{\mathbb{R}^n} b_i \mu_{t,y}(db) \int_{\mathbb{R}^n} [\sigma^{-1}b]_j \mu_{t,y}(db) \\
&= \sum_{k=1}^{n} \left(\sigma^{-1}_{j,k} \int_{\mathbb{R}^n} b_k b_i \mu_{t,y}(db) - \int_{\mathbb{R}^n} b_i \mu_{t,y}(db) \int_{\mathbb{R}^n} b_k \mu_{t,y}(db) \right) \\
&= \sum_{k=1}^{n} \sigma^{-1}_{j,k} \mathrm{cov}(B_i, B_k | Y_t = y) \\
&= \left(\sigma^{-1} \mathrm{cov}\left(B \mid Y_t = y\right)\right)_{j,i} \\
&= \left(\mathrm{cov}\left(B \mid Y_t = y\right)(\sigma^{-1})^{\intercal}\right)_{i,j}.
\end{aligned}$$

If $\mathrm{cov}\left(B \mid Y_t\right)$ is not positive definite, then for some linear combination of B, we would have $\mathrm{Var}\left(\sum_j q_j B_j \mid Y_t\right) = 0$ for some $q \in \mathbb{R}^n$, meaning that $\sum_j q_j B_j = C$, $C \in \mathbb{R}$, $\mu_{t,y}$-a.e.. But from (3.1), this would imply $\sum_j q_j B_j = C$, $C \in \mathbb{R}$, μ-a.e which contradicts Assumption 2.1.

Fix $t \geq 0$, the function $f_t : \mathbb{R}^n \to \mathbb{R}^n$ is obviously surjective on its image $\mathcal{B}_t$. To show that it is injective, we define $\tilde{f} = f_t \sigma^{\intercal}$ and take two $\mathbb{R}^n$-vectors $y_1 \neq y_2$. If we compute the Taylor expansion of $\tilde{f}$,

$$\tilde{f}(y_1) - \tilde{f}(y_2) = \nabla_y \tilde{f}(\eta)(y_1 - y_2),$$

where η is a convex combination of y_1 and y_2. Since $\mathrm{cov}(B|Y_t)$ is positive definite $\nabla_y \tilde{f}$ is positive definite so $\tilde{f}(y_1) \neq \tilde{f}(y_2)$ and f_t is injective.

□

Let us now introduce the so-called innovation process

$$\hat{W}_t \;:=\; \sigma^{-1} \int_0^t (B - \hat{B}_s) ds + W_t, \quad 0 \leq t \leq T,$$

which is a $(\mathbb{P}, \mathbb{F}^S)$-Brownian motion, see Proposition 2.30 in Bain and Crisan (2009). The observation process Y is written in terms of the innovation process as

$$Y_t \;=\; \sigma^{-1} \int_0^t \hat{B}_s ds + \hat{W}_t. \tag{3.3}$$

Applying Itô's formula to $\hat{B}_t = f(t, Y_t)$ with (3.3), and recalling that $\hat{B}$ is a $(\mathbb{P}, \mathbb{F}^S)$-martingale, we see that $d\hat{B}_t = \nabla_y f(t, Y_t) d\hat{W}_t$. By Lemma 3.2, and defining the matrix-valued function ψ by

$$\psi(t,b) \;:=\; \nabla_y f(t, f_t^{-1}(b)), \quad t \in [0,T],\; b \in \mathcal{B}_t, \tag{3.4}$$

where $y = f_t^{-1}(b)$ is the unique value of the observation process Y_t that yields $\hat{B}_t = b$, we have

$$d\hat{B}_t \;=\; \psi(t,\hat{B}_t)d\hat{W}_t. \tag{3.5}$$

Remark 3.3 (Dirac case)**.** *The Dirac prior distribution $\mu = \mathbb{1}_{b_0}$ does not verify Assumption* (2.1)*. Indeed in this case, $\forall\ (t,b)\ f(t,b) := b_0$, consequently we cannot properly define the function f_t^{-1} and the matrix-valued function ψ. A natural way to extend our framework to the Dirac case is by setting $\forall (t,b)$, $\psi(t,b) := 0$. Indeed when B is known then $cov(B) = 0$ and thus $\psi = 0$.* $\Diamond$

We can rewrite Model (2.1) in terms of observable variables:

$$\begin{cases} dS_t &= \mathrm{diag}(S_t)\left(\hat{B}_t dt + \sigma d\hat{W}_t\right), \quad t\in[0,T], \qquad S_0 \;=\; s_0, \\ d\hat{B}_t &= \psi(t,\hat{B}_t)d\hat{W}_t, \qquad\qquad\quad t\in[0,T], \quad \hat{B}_0 \;=\; b_0 \;= \mathbb{E}[B]. \end{cases}$$

Notice that the dynamics of the observable process $\hat{B}_t = f(t,Y_t) = f(t,h(t,S_t))$ is fully determined by the function ψ given in analytical form by (3.4) (explicit computations will be given later in the next sections). The dynamics of the wealth process in (2.2) becomes

$$dX_t^\alpha \;=\; \alpha_t^\intercal \mathrm{diag}(S_t)^{-1} dS_t = \alpha_t^\intercal(\hat{B}_t dt + \sigma d\hat{W}_t), \quad 0\le t\le T, \quad X_0^\alpha = x_0. \tag{3.6}$$

4 Solution to the Bayesian-Markowitz problem

4.1 Main result

In order to solve the initial Markovitz problem in (2.3), we first solve problem (2.6), then use Lemma 2.4 to find the optimal γ and obtain the dual function $V_0(\lambda)$ stated in (2.5), and finally apply Lemma 2.3 to find the optimal Lagrange multiplier λ which gives us the original value function $U_0(\vartheta)$ and the associated optimal strategy.

Let us define the dynamic value function associated to problem (2.3):

$$v^{\lambda,\gamma}(t,x,b) := \inf_{\alpha\in\mathcal{A}} J^{\lambda,\gamma}(t,x,b,\alpha), \qquad t\in[0,T],\ x\in\mathbb{R},\ b\in\mathcal{B}_t, \tag{4.1}$$

with

$$J^{\lambda,\gamma}(t,x,b,\alpha) := \mathbb{E}\Big[\lambda(X_T^{t,x,b,\alpha})^2 - (1+2\lambda\gamma)X_T^{t,x,b,\alpha}\Big],$$

where $X^{t,x,b,\alpha}$ is the solution to (3.6) on $[t,T]$, starting at $X_t^{t,x,b,\alpha} = x$ and $\hat{B}_t = b$ at time $t\in[0,T]$, controlled by $\alpha\in\mathcal{A}$, so that $V_0(\lambda,\gamma) = v^{\lambda,\gamma}(0,x_0,b_0)$, with $b_0 = \mathbb{E}[B]$.

Problem (4.1) is a standard stochastic control problem that can be characterized by the dynamic programming Hamilton-Jacobi-Bellman (HJB) equation, which is a fully nonlinear PDE. Actually, by exploiting the quadratic structure of this control problem (as detailed below), the HJB equation can be reduced to the following semi-linear PDE on $\mathcal{R} = \{(t,b) : t\in[0,T], b\in\mathcal{B}_t\}$:

$$-\partial_t R - \frac{1}{2}\mathrm{tr}\big(\psi\psi^\intercal \mathcal{D}_b^2 R\big) + \tilde{F}(t,b,\nabla_b R) = 0, \tag{4.2}$$

with

$$\tilde{F}(t,b,p) := 2\left(\psi\sigma^{-1}b\right)^\intercal p - \frac{1}{2}|\psi^\intercal p|^2 - |\sigma^{-1}b|^2, \tag{4.3}$$

and terminal condition

$$R(T,b) = 0, \quad b \in \mathcal{B}_t, \tag{4.4}$$

We can then state the main result of this paper, which provides the analytic solution to the Bayesian Markowitz problem.

Theorem 4.1. *Suppose there exists a solution $R \in C^{1,2}([0,T) \times \mathcal{B}_t; \mathbb{R}_+) \cap C^0(\mathcal{R}; \mathbb{R}_+)$ to the semi-linear Eq. (4.2) with terminal condition (4.4). Then, for any $\lambda > 0$,*

$$V_0(\lambda) \;=\; -\frac{1}{4\lambda}\left(e^{R(0,b_0)} - 1\right) - x_0$$

with the associated optimal mean-variance control given in feedback form by

$$\alpha_t^{*,\lambda} \;=\; a_0^{Bayes,\lambda}(t, X_t^{\alpha^{*,\lambda}}, \hat{B}_t) \;=\; a_0^{Bayes,\lambda}\big(t, X_t^{\alpha^{*,\lambda}}, f(t,h(t,S_t))\big) \tag{4.5}$$

where

$$a_0^{Bayes,\lambda}(t,x,b) \;:=\; \left(x_0 - x + \frac{e^{R(0,b_0)}}{2\lambda}\right)\left(\Sigma^{-1}b - \left(\psi\sigma^{-1}\right)^\intercal \nabla_b R(t,b)\right), \tag{4.6}$$

and the corresponding optimal terminal wealth is equal to

$$\mathbb{E}\big[X_T^{\alpha^{*,\lambda}}\big] \;=\; x_0 + \frac{1}{2\lambda}\big(e^{R(0,b_0)} - 1\big). \tag{4.7}$$

Moreover, for any $\vartheta > 0$,

$$U_0(\vartheta) \;=\; x_0 + \sqrt{\vartheta\left(e^{R(0,b_0)} - 1\right)}$$

with the associated optimal Bayesian-Markowitz strategy given by

$$\hat{\alpha}^\vartheta \;=\; \alpha^{*,\lambda(\vartheta)},$$

with

$$\lambda(\vartheta) \;=\; \sqrt{\frac{e^{R(0,b_0)} - 1}{4\vartheta}}. \tag{4.8}$$

Proof. The detailed proof is postponed in C, and we only sketch here the main arguments.

For fixed (λ, γ), the HJB equation associated to the standard stochastic control problem (4.1) is written as

$$\partial_t v^{\lambda,\gamma} + \frac{1}{2}\mathrm{tr}\big(\psi\psi^\intercal \mathcal{D}_b^2 v^{\lambda,\gamma}\big) + \inf_{\alpha\in\mathcal{A}}\left[\alpha^\intercal b \partial_x v^{\lambda,\gamma} + \frac{1}{2}|\sigma^\intercal\alpha|^2\partial_{xx}^2 v^{\lambda,\gamma} + \alpha^\intercal\sigma\psi^\intercal\partial_{xb}^2 v^{\lambda,\gamma}\right] \;=\; 0, \tag{4.9}$$

with terminal condition

$$v^{\lambda,\gamma}(T,x,b) \;=\; \lambda x^2 - (1+2\lambda\gamma)x.$$

We look for a solution in the ansatz form: $v^{\lambda,\gamma}(t,x,b) = K(t,b)x^2 + \Gamma(t,b)x + \chi(t,b)$. By plugging into the above HJB equation, and identifying the terms in x^2, x, we find after some straightforward calculations that

$$v^{\lambda,\gamma}(t,x,b) \;=\; e^{-R(t,b)}\left[\lambda x^2 - (1+2\lambda\gamma)x + \frac{(1+2\lambda\gamma)^2}{4\lambda}\right] - \frac{(1+2\lambda\gamma)^2}{4\lambda},$$

where R has to satisfy the semi-linear PDE (4.2) (which does not depend on λ, γ). Since R is assumed to exist smooth, the optimal feedback control achieving the argmin in the HJB equation (4.9) is given by

$$\tilde{\alpha}_t^{\lambda,\gamma} = \tilde{a}^{\lambda,\gamma}(t, X_t^{\tilde{\alpha}^{\lambda,\gamma}}, \hat{B}_t), \tag{4.10}$$

with

$$\tilde{a}^{\lambda,\gamma}(t,x,b) \ := \ \Big(\frac{1}{2\lambda} + \gamma - x\Big)\big(\Sigma^{-1}b - (\psi\sigma^{-1})^\intercal \nabla_b R(t,b)\big), \tag{4.11}$$

while $\gamma^*(\lambda) := \arg\min_{\gamma\in\mathbb{R}} \big[\tilde{V}_0(\lambda,\gamma) + \lambda\gamma^2\big] = \arg\min_{\gamma\in\mathbb{R}} \big[\tilde{v}^{\lambda,\gamma}(0,x_0,b_0) + \lambda\gamma^2\big]$ is given by

$$\gamma^*(\lambda) \ = \ x_0 + \frac{1}{2\lambda}\big(e^{R(0,b_0)} - 1\big), \tag{4.12}$$

which leads to the expression (4.7) from (2.7). We deduce that

$$\begin{aligned} V_0(\lambda) &= \tilde{V}_0(\lambda, \gamma^*(\lambda)) + \lambda(\gamma^*(\lambda))^2 \\ &= v^{\lambda,\gamma^*(\lambda)}(0,x_0,b_0) + \lambda(\gamma^*(\lambda))^2 \\ &= -\frac{1}{4\lambda}\Big(e^{R(0,b_0)} - 1\Big) - x_0. \end{aligned} \tag{4.13}$$

From Lemma 2.4, we know $\alpha_t^{*,\lambda} = \tilde{\alpha}_t^{\lambda,\gamma^*(\lambda)}$ and the optimal control for $V_0(\lambda)$ is thus

$$\alpha_t^{*,\lambda} = \tilde{a}^{\lambda,\gamma^*(\lambda)}(t, X_t^{\tilde{a}^{\lambda,\gamma^*(\lambda)}}, \hat{B}_t) = a_0^{Bayes,\lambda}(t, X_t^{\alpha^{*,\lambda}}, \hat{B}_t), \quad 0 \le t \le T,$$

with the function $a_0^{Bayes,\lambda}$ as in (4.6). This leads to the expression in (4.5) from (4.10), (4.11) and (4.12). Finally, the Lagrange multiplier $\lambda(\vartheta) = \arg\min_{\lambda>0}\big[\lambda\vartheta - V_0(\lambda)\big]$ is explicitly computed from (4.13), and is equal to the expression in (4.8). The optimal performance of the Markowitz problem is then equal to

$$U_0(\vartheta) \ = \ \lambda(\vartheta)\vartheta - V_0\big(\lambda(\vartheta)\big) \ = \ x_0 + \sqrt{\vartheta\,\big(e^{R(0,b_0)} - 1\big)},$$

by (4.8) and (4.13). The optimal Bayesian-Markowitz strategy is then given by $\hat{\alpha}^\vartheta = \alpha^{*,\lambda(\vartheta)}$ according to Lemma 2.3. □

Remark 4.2 (Financial interpretation)**.** When the drift $B = b_0$ is known, the function R is simply given by $R(t) = |\sigma^{-1}b_0|^2(T-t)$, and we retrieve the results recalled in Remark 2.5. Under a prior probability distribution on the drift B, we see that the optimal performance of the Bayesian-Markowitz problem has a similar form as in the constant drift case, when substituting $|\sigma^{-1}b_0|^2T$ by $R(0,b_0)$. For the optimal Bayesian-Markowitz strategy, we have to substitute the vector term $\Sigma^{-1}b_0$ by $\Sigma^{-1}\hat{B}_t$, i.e. replacing b_0 by the posterior predictive mean $\hat{B}_t$, and correcting with the additional term $(\psi\sigma^{-1})^\intercal\nabla_b R(t,\hat{B}_t)$. We call R the Bayesian risk premium function. ♢

Our main result in Theorem 4.1 is stated under the condition that R exists sufficiently smooth. In the next paragraph, we give some sufficient conditions ensuring this regularity, thus the existence of R, and provide some examples.

4.2 On existence and smoothness of the Bayesian risk premium

In this section we provide sufficient conditions ensuring the existence of the Bayesian risk premium function R, solution to the semi-linear PDE (4.2). Note that the standard assumptions of existence and uniqueness that we can find in Ladyzhenskaia et al. (1968) do not apply here since the function $\tilde{F}$ defined in (4.3) is not globally Lipschitz in p. The difficulty of our framework comes from the unboundedness of the domain and the quadratic growth in p of the function $\tilde{F}$ which entails that the solution may grow quadratically.

Theorem 4.3. *Suppose the following conditions hold true:*

- $\forall (t,b) \in \mathcal{R}$, $\psi(t,b)$ *and* $\psi^{-1}(t,b)$ *are bounded,*
- *The matrix* $\nabla_b \psi$ *exists and is bounded,*

then there exists a classical solution $R \in \mathcal{C}^{1,2}(\mathcal{R})$ *to PDE* (4.2)*. In addition,* $R(t,b)$ *is at most quadrically growing in* b *and* $\nabla_b R(t,b)$ *is at most linearly growing in* b*.*

Proof. We provide a sketch of the proof and details in D. We follow a similar approach as Pham (2002) and Benth and Karlsen (2005).Without loss of generality we consider the function $\tilde{R} := -R$. We rewrite the PDE characterizing $\tilde{R}$ as

$$-\partial_t \tilde{R} - \frac{1}{2}\mathrm{tr}(\psi\psi^\intercal \mathcal{D}_b^2 \tilde{R}) + F(t,b,\nabla_b \tilde{R}) \;=\; 0$$

with terminal condition $\tilde{R}(T,.) = 0$, where the function $F : \mathcal{R} \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$F(t,b,p) \;=\; \frac{1}{2}|\psi^\intercal p|^2 + 2\big(\psi\sigma^{-1}b\big)^\intercal p + |\sigma^{-1}b|^2,$$

is quadratic (hence convex in the gradient argument p). Let us introduce the Fenchel-Legendre transform of $F(t,b,.)$ as

$$L_t(b,q) \;:=\; \max_{p\in\mathbb{R}^n}\big[-q^\intercal\psi^\intercal p - F(t,b,p)\big], \quad b \in \mathcal{B}_t,\; q \in \mathbb{R}^n,$$

which explicitly yields

$$L_t(b,q) \;=\; \frac{1}{2}|q|^2 + 2(\sigma^{-1}b)^\intercal q + |\sigma^{-1}b|^2. \tag{4.14}$$

The explicit form shows that L_t only depends on t through the domain $\mathcal{B}_t$ of b. It is also known that the following duality relationship holds:

$$F(t,b,p) \;=\; \max_{q\in\mathbb{R}^n}\big[-q^\intercal\psi^\intercal p - L_t(b,q)\big].$$

From (4.14) we deduce the following estimates on L and its gradient w.r.t. b:

$$\left\{\begin{array}{rcl} |L_t(b,q)| & \leq & C(|q|^2+|b|^2), \\ |\nabla_b L_t(b,q)| & \leq & C(|q|+|b|), \quad \forall b \in \mathcal{B}_t,\; q \in \mathbb{R}^n, \end{array}\right. \tag{4.15}$$

for some independent of t positive constant C. Let us now consider the truncated auxiliary function

$$F^k(t,b,p) \;=\; \max_{|q|\leq k}\big[-q^\intercal\psi^\intercal p - L_t(b,q)\big], \quad b \in \mathcal{B}_t,\; p \in \mathbb{R}^n,$$

for $k \in \mathbb{N}$. By (4.15) and Theorem 4.3 p. 163 in Fleming and Soner (2006), there exists a unique smooth solution $\tilde{R}^k$ with quadratic growth condition to the truncated semi-linear PDE

$$-\partial_t \tilde{R}^k - \frac{1}{2}\mathrm{tr}(\psi\psi^\intercal \mathcal{D}_b^2 \tilde{R}^k) + F^k(t, b, \nabla_b \tilde{R}^k) \;=\; 0,$$

with terminal condition $\tilde{R}^k(T, .) = 0$.

The next step is to obtain estimates on $\tilde{R}^k$ and its gradient w.r.t. b, uniformly in k:

$$\left\{ \begin{array}{rcl} |\tilde{R}^k(t,b)| & \leq & C(1+|b|^2), \\ |\nabla_b \tilde{R}^k(t,b)| & \leq & C(1+|b|), \quad \forall (t,b) \in \mathcal{R}, \end{array} \right.$$

for some constant C independent of k. Then, following similar arguments as in Pham (2002) and Benth and Karlsen (2005), we show that for k large enough, $R^k = -\tilde{R}^k$ solves PDE (4.2). This proves the existence of a smooth solution to this semi-linear PDE. □

4.3 Examples

We illustrate with some relevant examples the explicit computation of the diffusion coefficient ψ appearing in the dynamic of the posterior mean of the drift described in (3.5), as well as the computation of the Bayesian risk premium function R.

4.3.1 Prior discrete law

We consider the case when the drift vector B has a prior discrete distribution

$$\mu(db) \;=\; \sum_{i=1}^N \pi_i \delta_{V_i}(db), \tag{4.16}$$

where for $i = 1, \ldots, N$, V_i are vectors in $\mathbb{R}^n$, $\pi_i \in (0,1)$ and $\sum_{i=1}^N \pi_i = 1$. We denote by V the $n \times N$-matrix $V = (V^1 \ldots V^N)$, with rank(V) $= n < N$ and we assume that the vectors V_i are chosen such that

$$\mathrm{Cov}(B) = \sum_{i=1}^N \pi_i V_i V_i^\intercal - b_0 b_0^\intercal > 0.$$

From (4.16) and (3.1), we easily compute the conditional distribution of B w.r.t. $Y_t = y \in \mathbb{R}^n$, which is given by

$$\mu_{t,y}(db) \;=\; \sum_{i=1}^N p^i_{t,y} \delta_{V_i}(db) \quad t \in [0,T],$$

where $p_{t,y} = (p^1_{t,y} \ldots p^N_{t,y}) \in [0,1]^N$ is determined by

$$p^i_{t,y} \;=\; \frac{\pi_i e^{y^\intercal \sigma^{-1} V_i - \frac{1}{2} V_i^\intercal \Sigma^{-1} V_i t}}{\sum_{j=1}^N \pi_j e^{y^\intercal \sigma^{-1} V_j - \frac{1}{2} V_j^\intercal \Sigma^{-1} V_j t}}, \quad i = 1, \ldots, N.$$

It follows that the function f in (3.2) is equal to

$$f(t,y) \;=\; \int_{\mathbb{R}^n} b \mu_{t,y}(db) \;=\; \sum_{i=1}^N p^i_{t,y} V_i \;=\; V p_{t,y},$$

from which we calculate its gradient:

$$\nabla_y f(t,y) \quad = \quad \left(\sum_{i=1}^{N} p^i_{t,y} V_i V_i^\intercal - V p_{t,y} (V p_{t,y})^\intercal\right) (\sigma^{-1})^\intercal.$$

In this case the domain $\mathcal{B}_t$ of b is the convex hull of the vectors V_i, $i = 1, \dots, N$ mathematically defined as

$$\mathcal{B}_t = \left\{ b : V p_{t,y^*} = b \text{ and } \forall j \in \{1,\dots,N\},\ \ p^j_{t,y^*} \in (0,1) \ \text{ and } \sum_{j=1}^{N} p^j_{t,y^*} = 1 \right\},$$

where for fixed $(t,b) \in \mathcal{R}$, $y^* = f_t^{-1}(b)$. More precisely, y^* is the solution to the equation $f(t,y^*) = V p_{t,y^*} = b$. We note $\overline{\mathcal{B}_t}$ the closure of $\mathcal{B}_t$.
The function ψ defined in (3.4) is therefore equal to

$$\begin{aligned} \psi(t,b) &= \nabla_y f(t,y^*) \\ &= \left(\sum_{i=1}^{N} p^i_{t,y^*} V_i V_i^\intercal - b b^\intercal\right) (\sigma^{-1})^\intercal. \end{aligned} \tag{4.17}$$

Remark 4.4. In the discrete case we cannot apply Theorem 4.3 for two main reasons: $\mathcal{B}_t$ is a bounded non-smooth domain, and ψ can vanish on the boundary of $\mathcal{B}_t$, leading to ψ^{-1} being ill defined on $\overline{\mathcal{B}_t}$. It is, in particular, easy to see that ψ vanishes at the points V_i. Indeed, b equals V_i implies that $\forall\ j = 1, \dots, N$, $p^j_{t,y^*} = 0$ for $j \neq i$ and $p^i_{t,y^*} = 1$ leading to Eq. (4.17) being zero. The existence theorem of a solution to problem (4.2)-(4.4), in our context, can be found in Krylov (1987), especially chapter 6, section 3. $\lozenge$

4.3.2 The Gaussian case

We consider the case when the drift vector B follows a n-dimensional normal distribution with mean vector b_0 and covariance matrix Σ_0:

$$\mu \quad \sim \quad \mathcal{N}(b_0, \Sigma_0). \tag{4.18}$$

In this conjugate case, it is well-known (see Proposition 10 in Guéant and Pu (2017)) that the posterior distribution of B, i.e., the conditional distribution of B given $Y_t = y$ is also Gaussian with mean

$$f(t,y) \quad = \quad \left(\Sigma_0^{-1} + \Sigma^{-1} t\right)^{-1} \left(\Sigma_0^{-1} b_0 + (\sigma^\intercal)^{-1} y\right), \tag{4.19}$$

and covariance

$$\Sigma(t,y) \quad = \quad \left(\Sigma_0^{-1} + \Sigma^{-1} t\right)^{-1}.$$

For the sake of completeness, we show this result. From (3.1) and (4.18), the conditional distribution of B w.r.t. $Y_t = y$ is given by

$$\begin{aligned} \mu_{t,y}(db) &= \frac{e^{y^\intercal \sigma^{-1} b - \frac{1}{2} b^\intercal \Sigma^{-1} b t - \frac{1}{2}(b-b_0)^\intercal \Sigma_0^{-1}(b-b_0)}}{\int_{\mathbb{R}^n} e^{y^\intercal \sigma^{-1} b - \frac{1}{2} b^\intercal \Sigma^{-1} b t - \frac{1}{2}(b-b_0)^\intercal \Sigma_0^{-1}(b-b_0)} db} db \\ &= \frac{e^{-\frac{1}{2}\left(b^\intercal\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)b - 2b^\intercal\left(\Sigma_0^{-1}b_0 + (\sigma^{-1})^\intercal y\right)\right)}}{\int_{\mathbb{R}^n} e^{-\frac{1}{2}\left(b^\intercal\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)b - 2b^\intercal\left(\Sigma_0^{-1}b_0 + (\sigma^{-1})^\intercal y\right)\right)} db} db \\ &= \frac{e^{-\frac{1}{2}\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}\left(\Sigma_0^{-1}b_0+(\sigma^{-1})^\intercal y\right)\right)^\intercal \left(\Sigma_0^{-1}+\Sigma^{-1}t\right)\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}\left(\Sigma_0^{-1}b_0+(\sigma^{-1})^\intercal y\right)\right)}}{\int_{\mathbb{R}^n} e^{-\frac{1}{2}\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}\left(\Sigma_0^{-1}b_0+(\sigma^{-1})^\intercal y\right)\right)^\intercal \left(\Sigma_0^{-1}+\Sigma^{-1}t\right)\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}\left(\Sigma_0^{-1}b_0+(\sigma^{-1})^\intercal y\right)\right)} db} db. \end{aligned}$$

Recognising the form of a Gaussian distribution, we find

$$\mu_{t,y}(db) \;=\; \frac{e^{-\frac{1}{2}\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}\left(\Sigma_0^{-1}b_0+(\sigma^{-1})^\intercal y\right)\right)^\intercal\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}\left(\Sigma_0^{-1}b_0+(\sigma^{-1})^\intercal y\right)\right)}}{(2\pi)^{\frac{n}{2}}\left|\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}\right|^{\frac{1}{2}}}db,$$

which shows that

$$\mu_{t,y} \;\sim\; \mathcal{N}\left(f(t,y),\Sigma(t,y)\right).$$

Next, from (4.19), we have

$$\nabla_y f(t,y) \;=\; \left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}(\sigma^{-1})^\intercal,$$

Left multiplying by $\Sigma_0\Sigma_0^{-1}$ and right multiplying by $\sigma^\intercal\Sigma^{-1}\sigma$ the previous equation, we obtain the multi-dimensional independent from b familiar expression of the diffusion coefficient ψ stated in Ekstrom and Vaicenavicius (2016) in dimension one:

$$\psi(t) \;=\; \Sigma_0\left(\Sigma+\Sigma_0 t\right)^{-1}\sigma.$$

For the computation of the Bayesian risk premium function R solution to (4.2), we look for a solution of the form:

$$R(t,b) \;=\; b^\intercal M(t)b+U(t), \tag{4.20}$$

for some $\mathbb{R}^{n\times n}$-valued function M, and $\mathbb{R}^n$-valued function U defined on $[0,T]$, to be determined. Plugging into (4.2), we see that M and U should satisfy the first order ODE system:

$$\left\{\begin{array}{rcl} -M'(t)-2M(t)^\intercal G(t)^\intercal\Sigma G(t)M(t)+4G(t)M(t)-\Sigma^{-1} &=& 0\\ -U'(t)-\mathrm{tr}\left[G(t)^\intercal\Sigma G(t)M(t)\right] &=& 0,\end{array}\right. \tag{4.21}$$

with $G(t)=(\Sigma+\Sigma_0 t)^{-1}\Sigma_0$ and terminal conditions:

$$M(T) \;=\; 0, \qquad U(T) \;=\; 0. \tag{4.22}$$

The solution to (4.21)-(4.22) is given by the following Lemma, whose proof is detailed in E.

Lemma 4.5. *The solution to the ODE system* (4.21) *with terminal condition* (4.22) *is:*

$$\begin{aligned} M(t) &= \Sigma_0^{-1}+\Sigma^{-1}t-\left[\left(\Sigma_0^{-1}+\Sigma^{-1}T\right)^{-1}+2\int_t^T\Sigma_0\left(\Sigma+\Sigma_0 s\right)^{-1}\Sigma\left(\Sigma+\Sigma_0 s\right)^{-1}\Sigma_0 ds\right]^{-1}\\ U(t) &= \int_t^T\mathrm{tr}\left(\Sigma_0\left(\Sigma+\Sigma_0 s\right)^{-1}-\left[\left(\left(\Sigma_0^{-1}+\Sigma^{-1}T\right)^{-1}+2\int_s^T G(u)^\intercal\Sigma G(u)du\right)\left(G(s)^\intercal\Sigma G(s)\right)^{-1}\right]^{-1}\right)ds \end{aligned}$$

We conclude that the Bayesian risk premium function R is explicitly given by (4.20) with M and U as in Lemma 4.5. The optimal Bayesian mean-variance strategy (4.5) is explicitly written as

$$\alpha_t^{*,\lambda} \;=\; \left(x_0-X_t^{\alpha^{*,\lambda}}+\frac{e^{R(0,b_0)}}{2\lambda}\right)\left[\Sigma^{-1}-\left(\Sigma+\Sigma_0 t\right)^{-1}\Sigma_0 M(t)\right]\hat{B}_t$$

where $\hat{B}_t = f(t, Y_t) = f(t, h(t, S_t))$.

In the one-dimensional case $n = 1$, hence with $\Sigma = \sigma^2$, $\Sigma_0 = \sigma_0^2$, the expressions for M and U are simplified into

$$\begin{aligned} M(t) &= \frac{\sigma^2 + \sigma_0^2 t}{\sigma^2\big(\sigma^2 + \sigma_0^2(2T - t)\big)}(T - t), \\ U(t) &= \ln\left(\frac{\sigma^2 + \sigma_0^2 T}{\sqrt{(\sigma^2 + \sigma_0^2 t)(\sigma^2 + \sigma_0^2(2T - t))}}\right). \end{aligned}$$

5 Impact of learning on the Markowitz strategy

This section shows the benefit of using a Bayesian learning approach to solve the Markowitz problem. In contrast to classical approaches that estimate unknown parameters in a second step, leading to sub-optimal solutions, the Bayesian learning approach uses the updated data from the most recent prices of the assets to adjust the controls of the investment strategy and reaches optimality in the Markowitz sense.

In a framework where the drift B is unknown, following a prior probability distribution μ, we will compare the performance, in terms of Sharpe ratio, of the Bayesian learning strategy to the non-learning strategy. Recall, the non-learning strategy considers the drift constant set at $b_0 = \mathbb{E}[B] = \int b\mu(db)$. This allows us to exhibit the benefit of integrating the most recent available information into the strategy.

5.1 Computation of the Sharpe ratios

The Sharpe ratio associated to a portfolio strategy α is defined by

$$Sh_T^\alpha := \frac{\mathbb{E}[X_T^\alpha] - x_0}{\sqrt{\mathrm{Var}(X_T^\alpha)}}.$$

We denote by $\hat{\alpha}^{\vartheta,L}$ the optimal Bayesian-Markowitz strategy obtained in Theorem 4.1, $X^L = X^{\hat{\alpha}^{\vartheta,L}}$ the associated wealth process and Sh_T^L the corresponding Sharpe ratio. We also write $\hat{\alpha}^{\vartheta,NL}$ the non-learning Markowitz strategy based on a constant drift parameter b_0 as described in Remark 2.5, $X^{NL} = X^{\hat{\alpha}^{\vartheta,NL}}$ the associated wealth process and Sh_T^{NL} the corresponding Sharpe ratio.

Proposition 5.1. *The Sharpe ratios of the learning and non-learning Markowitz strategies are explicitly given by*

$$Sh_T^L = \sqrt{e^{R(0,b_0)} - 1}, \tag{5.1}$$

$$Sh_T^{NL} = \frac{1 - \int_{\mathbb{R}^n} e^{-b_0^\intercal \Sigma^{-1} bT}\mu(db)}{\sqrt{\int_{\mathbb{R}^n} e^{-b_0^\intercal \Sigma^{-1}(2b - b_0)T}\mu(db) - \left(\int_{\mathbb{R}^n} e^{-b_0^\intercal \Sigma^{-1} bT}\mu(db)\right)^2}}. \tag{5.2}$$

Proof.

1. We first focus on the Bayesian learning strategy. From (4.7), we have $\mathbb{E}[X^L] = x_0 + \frac{1}{2\lambda(\vartheta)}\left(e^{R(0,b_0)} - 1\right)$ with $\lambda(\vartheta) = \sqrt{\frac{e^{R(0,b_0)} - 1}{4\vartheta}}$ binding the variance of the optimal terminal wealth to ϑ. We thus obtain

$$Sh_T^L = \frac{\mathbb{E}[X^L] - x_0}{\sqrt{\vartheta}} = \frac{\sqrt{\vartheta\left(e^{R(0,b_0)} - 1\right)}}{\sqrt{\vartheta}} = \sqrt{e^{R(0,b_0)} - 1}.$$

2. Let us now consider the non-learning strategy $\hat{\alpha}^{\vartheta,NL}$ for which we need to compute its expectation and variance. From the expression of $\hat{\alpha}^{\vartheta,NL}$ given in Remark 2.5, and the self-financed equation of the wealth process (2.2), we deduce the dynamic of the conditional expectation of X^{NL} given B:

$$\begin{aligned} d\mathbb{E}[X_t^{NL}|B] &= \mathbb{E}\Big[\big(x_0 - X_t^{NL} + C_1\big)C_2^\intercal B|B\Big]dt \\ &= \Big[\,(x_0 + C_1)\,C_2^\intercal B - \mathbb{E}\big[X_t^{NL}|B\big]C_2^\intercal B\Big]dt, \end{aligned}$$

where we set $C_1 = \frac{e^{|\sigma^{-1}b_0|^2T}}{2\lambda_0(\vartheta)}$, $C_2 = \Sigma^{-1}b_0$ to alleviate notations. We thus obtain an ordinary differential equation for $t \mapsto \mathbb{E}[X_t^{NL}|B]$ with initial condition $\mathbb{E}[X_0^{NL}|B] = x_0$, whose solution is explicitly given by

$$\mathbb{E}[X_t^{NL}|B] = x_0 + C_1\big(1 - e^{-C_2^\intercal Bt}\big), \quad 0 \le t \le T. \tag{5.3}$$

Integrating w.r.t. the law of B, we obtain the expression of the expectation:

$$\mathbb{E}[X_t^{NL}] = x_0 + \frac{e^{||\sigma^{-1}b_0||^2T}}{2\lambda_0(\vartheta)}\left(1 - \int_{\mathbb{R}^n} e^{-b_0^\intercal\Sigma^{-1}bt}\mu(db)\right). \tag{5.4}$$

Next, let us compute the variance of X_T^{NL}. Applying Itô's formula to $|X^{NL}|^2$, and taking conditional expectation w.r.t. B, we have

$$\begin{aligned} d\mathbb{E}[|X_t^{NL}|^2|B] &= \Big[C_2^\intercal\,(\Sigma C_2 - 2B)\,\mathbb{E}[|X_t^{NL}|^2|B] - 2C_2^\intercal\,(x_0 + C_1)\,(\Sigma C_2 - B)\,\mathbb{E}[X_t^{NL}|B] \\ &\qquad + C_2^\intercal\Sigma C_2\,(x_0 + C_1)^2\,\Big]dt. \end{aligned}$$

Replacing $\mathbb{E}[X_t^{NL}|B]$ by its value calculated in (5.3), we obtain an ODE in $\mathbb{E}[|X_t^{NL}|^2|B]$ that we explicitly solve:

$$\mathbb{E}[|X_t^{NL}|^2|B] = (x_0 + C_1)^2 - 2C_1\,(x_0 + C_1)\,e^{C_2^\intercal Bt} + C_1^2 e^{C_2^\intercal(\Sigma C_2 - 2B)t}.$$

By integrating over B, we obtain

$$\mathbb{E}[|X_t^{NL}|^2] = (x_0 + C_1)^2 - 2C_1\,(x_0 + C_1)\int_{\mathbb{R}^n} e^{C_2^\intercal bt}\mu(db) + C_1^2 e^{C_2^\intercal\Sigma C_2 t}\int_{\mathbb{R}^n} e^{-2C_2^\intercal bt}\mu(db).$$

From $\mathrm{Var}(X_t^{NL}) = \mathbb{E}[|X_t^{NL}|^2] - \big(\mathbb{E}[X_t^{NL}]\big)^2$, we obtain the expression of the variance

$$\mathrm{Var}(X_t^{NL}) = \frac{e^{2|\sigma^{-1}b_0|^2T}}{4\lambda_0(\vartheta)^2}\left[\int_{\mathbb{R}^n} e^{-b_0^\intercal\Sigma^{-1}(2b-b_0)t}\mu(db) - \left(\int_{\mathbb{R}^n} e^{-b_0^\intercal\Sigma^{-1}bt}\mu(db)\right)^2\right]. \tag{5.5}$$

Finally, coupling (5.4) and (5.5) we infer the value of Sh_T^{NL} in (5.2). □

Next we prove that $Sh_T^{NL} \le Sh_T^L$. From (2.2) we know that X^{NL} is linear in its control $\hat{\alpha}^{\vartheta,NL}$, so we can always find a leveraged strategy $\tilde{\alpha}^{\vartheta,NL} = \delta\hat{\alpha}^{\vartheta,NL}$ with $\delta > 0$ such that $\mathrm{Var}(X_T^{\tilde{\alpha}^{\vartheta,NL}}) = \vartheta$, simply by taking $\delta = \sqrt{\frac{\vartheta}{\mathrm{Var}(X_T^{NL})}}$. Thus, by invariance of the Sharpe ratio w.r.t. the leverage we obtain,

$$\begin{aligned} Sh_T^L - Sh_T^{NL} &= \frac{\mathbb{E}[X_T^{\hat{\alpha}^{\vartheta,L}}] - x_0}{\sqrt{\vartheta}} - \frac{\mathbb{E}[X_T^{\hat{\alpha}^{\vartheta,NL}}] - x_0}{\sqrt{\mathrm{Var}(X_T^{\hat{\alpha}^{\vartheta,NL}})}} \\ &= \frac{\mathbb{E}[X_T^{\hat{\alpha}^{\vartheta,L}}] - x_0}{\sqrt{\vartheta}} - \frac{\mathbb{E}[X_T^{\tilde{\alpha}^{\vartheta,NL}}] - x_0}{\sqrt{\vartheta}} \\ &= \mathbb{E}[X_T^{\hat{\alpha}^{\vartheta,L}}] - \mathbb{E}[X_T^{\tilde{\alpha}^{\vartheta,NL}}] \\ &\ge 0 \end{aligned}$$

The last inequality comes from $\hat{\alpha}^{\vartheta,L}$ being the optimal control.

Remark 5.2. We have an upper-bound for the Sharpe ratio of the non-learning strategy:

$$Sh_T^{NL} \leq \sqrt{e^{|\sigma^{-1}b_0|^2T} - 1}.$$

Indeed, by Jensen's inequality, we have

$$\begin{aligned}
\left(\int_{\mathbb{R}^n} e^{-b_0\Sigma^{-1}bT}\mu(db)\right)^2 &\leq \int_{\mathbb{R}^n} e^{-2b_0\Sigma^{-1}bT}\mu(db), \\
\int_{\mathbb{R}^n} e^{-b_0\Sigma^{-1}bT}\mu(db) &\geq e^{-T\int_{\mathbb{R}^n} b_0\Sigma^{-1}b\mu(db)} = e^{-|\sigma^{-1}b_0|^2T}, \\
\int_{\mathbb{R}^n} e^{-2b_0\Sigma^{-1}bT}\mu(db) &\geq e^{-2|\sigma^{-1}b_0|^2T}.
\end{aligned}$$

From (5.2), we thus deduce

$$\begin{aligned}
Sh_T^{NL} &\leq \frac{1 - \int_{\mathbb{R}^n} e^{-b_0\Sigma^{-1}bT}\mu(db)}{\sqrt{\int_{\mathbb{R}^n} e^{-b_0\Sigma^{-1}(2b-b_0)T}\mu(db) - \int_{\mathbb{R}^n} e^{-2b_0\Sigma^{-1}bT}\mu(db)}} \\
&= \frac{1 - \int_{\mathbb{R}^n} e^{-b_0\Sigma^{-1}bT}\mu(db)}{\sqrt{\left(e^{|\sigma^{-1}b_0|^2T} - 1\right)\int_{\mathbb{R}^n} e^{-2b_0\Sigma^{-1}bT}\mu(db)}} \\
&\leq \frac{1 - e^{-|\sigma^{-1}b_0|^2T}}{\sqrt{\left(e^{|\sigma^{-1}b_0|^2T} - 1\right)e^{-2|\sigma^{-1}b_0|^2T}}} = \sqrt{e^{|\sigma^{-1}b_0|^2T} - 1}
\end{aligned}$$

We notice that when $\mu = \mathbb{1}_{b_0}$ the Sharpe ratio of the non-learning strategy coincides with $\sqrt{e^{|\sigma^{-1}b_0|^2T} - 1}$. ◊

5.2 Information Value

We illustrate our results in the case where B follows a prior one-dimensional Gaussian distribution $\mathcal{N}(b_0, \sigma_0^2)$. To assess the value of information (VI), we consider the Sharpe ratio Sh_T^L and Sh_T^{NL} of the learning and non-learning strategies as computed in Proposition 5.1. We define the value of information as the difference $Sh_T^L - Sh_T^{NL}$, and measure its sensitivity w.r.t. various parameters. So, in the sequel, we denote VI(Asset i) the value of information obtained using the set of parameters corresponding to Asset i.

The following explicit formulas for the Sharpe ratios learning and non-learning, used to create the graphs of this section, are computed from (5.1) and (5.2) with $\mu \sim \mathcal{N}(b_0, \sigma_0^2)$.

$$Gaussian_Sh_T^L = \sqrt{\frac{\sigma^2 + \sigma_0^2 T}{\sigma\sqrt{\sigma^2 + 2\sigma_0^2 T}} e^{\frac{b_0^2 T}{\sigma^2 + 2\sigma_0^2 T}} - 1} \tag{5.6}$$

and

$$Gaussian_Sh_T^{NL} = \frac{e^{\frac{b_0^2}{\sigma^2}T\left(1 - \frac{\sigma_0^2}{2\sigma^2}T\right)} - 1}{\sqrt{e^{\frac{b_0^2}{\sigma^2}T\left(1 + \frac{\sigma_0^2}{\sigma^2}T\right)} - 1}} \tag{5.7}$$

5.2.1 Volatility of the drift

Intuitively, the volatility of the drift σ_0 measures the confidence the investor puts in her estimate of b_0. The higher σ_0, the more confident the investor is about her estimate. It is an important parameter since it legitimates the use of a learning strategy.

To estimate the sensitivity of the value of information w.r.t. the volatility of the drift, we choose an investment horizon of $T = 1$ and a sample of three different assets with the same mean $b_0 = 5\%$ and different volatilities, $\sigma = \{5\%, 10\%, 20\%\}$ resulting in realistic Sharpe ratios of 1, 0.5 and 0.25 respectively. From Formulas (5.6) and (5.7) we compute the value of information according to the volatility of the drift ranging from 0 to 100%. Figure 1 shows the value of information as a function of the volatility of the drift for the three assets described previously, which parameters are summed up in Table 1. As we can see on the graph, the value of information is a monotone increasing function of the volatility of the drift. When the volatility of the drift is zero, as in the Dirac case, the Sharpe ratios of the learning and non-learning strategies are equal. It is clear since a volatility of the drift equal to zero means the drift is simply constant, so updating it does not bring any additional value to the learning strategy. For all three assets, we notice that as σ_0 increases the value of information becomes more and more valuable depending on the level of the Sharpe ratio of each asset. The higher the Sharpe ratio, the bigger the value of information and its rate of increase. For instance, the value of information of the highest Sharpe ratio asset, Asset 1, increases rapidly to 1 when $\sigma_0 = 10\%$ and reaches rapidly 3.5 when $\sigma_0 = 100\%$ whereas the lowest Sharpe ratio asset, Asset 3, equals roughly 0.1 when $\sigma_0 = 0.1$ and nearly reaches 2 when $\sigma_0 = 100\%$ at a relatively slower pace.

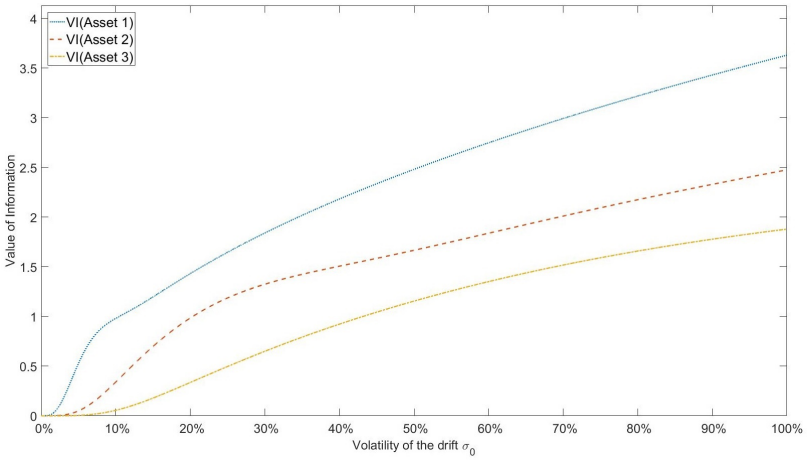


Figure 1: Value of information as a function of σ_0 with parameters as in Table 1.

Table 1: Parameter values used in Figure 1.

	$\mathbf{b_0}$	$\boldsymbol{\sigma}$	**Sharpe ratio**	**T**	$\boldsymbol{\sigma_0}$
Asset 1	5%	5%	1	1	[0-100%]
Asset 2	5%	10%	0.5	1	[0-100%]
Asset 3	5%	20%	0.25	1	[0-100%]

An interesting fact happens in the case of an asset with a high Sharpe ratio. Figure 2 shows the value of information, the curves of the Sharpe ratio of the learning and the

non-learning strategy based on Asset 4. This asset has a high Sharpe ratio of 2, a set of parameters as in Table 2 and a volatility of the drift ranging from 0 to 30%.

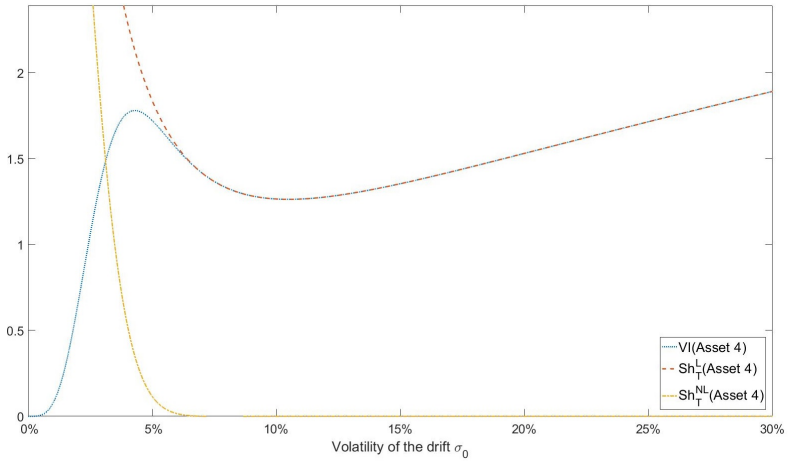


Figure 2: Value of information, Sh_T^L and Sh_T^{NL} of Asset 4 as a function of σ_0 with parameters as in Table 2.

Table 2: Parameter values used in Figure 2.

	$\mathbf{b_0}$	$\boldsymbol{\sigma}$	**Sharpe ratio**	**T**	$\boldsymbol{\sigma_0}$
Asset 4	10%	5%	2	1	[0-30%]

As we can see on the graph, the value of information curve is no more monotonic and we observe a bump around $\sigma_0 = 5\%$. To explain this shape, we see that when σ_0 ranges between 0 and 5%, both Sharpe ratios decrease. It is understandable since the variance of the terminal wealth is an increasing function of σ and σ_0. Nevertheless, the decrease is much more rapid for the non-learning strategy reaching approximately zero when the drift volatility merely exceeds 5%, approximately the value of the volatility of the asset. The difference in the rate of decrease explains the bump we observe. Then, the Sharpe ratio of the learning strategy keeps decreasing and progressively recover when the volatility of the drift approaches the double of the volatility of the asset, 10% . At this point, the learning effect really allows the strategy to exploit its advantage of updating the drift as the new information become available. As soon as $\sigma_0 > 2\sigma$, the signal coming from the volatility of the drift differentiates from the volatility of the asset and is captured by the learning strategy which allows the value of information to increase.

5.2.2 Sharpe ratio of the asset

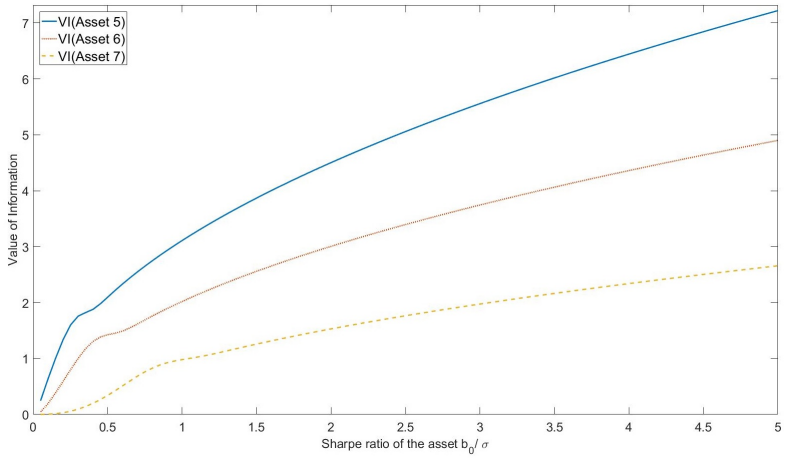


Figure 3: Value of information as a function of the Sharpe ratio of the asset with parameters as in Table 3.

Table 3: Parameter values used in Figure 3.

	$\mathbf{b_0}$	$\boldsymbol{\sigma}$	**Sharpe ratio**	**T**	$\boldsymbol{\sigma_0}$
Asset 5	5%	[1-100]%	[0-5]	1	75%
Asset 6	5%	[1-100]%	[0-5]	1	35%
Asset 7	5%	[1-100]%	[0-5]	1	10%

Figure 3 exhibits the value of information as a function of the Sharpe ratio of the asset, defined as b_0/σ, for a sample of three assets with parameters set in Table 3.

We see that the higher the Sharpe ratio of the asset the bigger the value of information for any level of σ_0. Obviously, the higher the volatility of the drift, the more value information has and the more necessary it is to update the strategy. We also notice that the slope of the curve is steeper when the Sharpe ratio of the asset is low, ranging roughly from 0 to 1. Intuitively, it shows that information has more value for assets with a low Sharpe ratio because assets with a high Sharpe ratio will perform well, no matter the learning. Another explanation is for a fix b_0, a low Sharpe ratio means a high volatility which increases the need for learning and consequently the value of information. Furthermore, we clearly see that the curve with the higher σ_0 dominates the lower one, which confirms the intuition and understanding of Figure 1.

5.2.3 Time

Figure 4 displays the value of information along the life of the investment for time $t \in [0, T]$. To obtain this graph, we use the set of parameters shown in Table 4. We simulate $N = 1,000,000$ optimal wealth trajectories $(X)^i_{i\in[\![1,N]\!]}$ and at each point in time we compute the empirical Sharpe ratio of the learning strategy:

$$\hat{Sh}^L_t = \frac{\sqrt{N-1}}{N} \frac{\sum_{i=1}^N (X^i_t - x_0)}{\sqrt{\sum_{i=1}^N (X^i_t)^2 - \left(\sum_{i=1}^N X^i_t\right)^2}}$$

and use Formula (5.5) for the non-learning strategy.

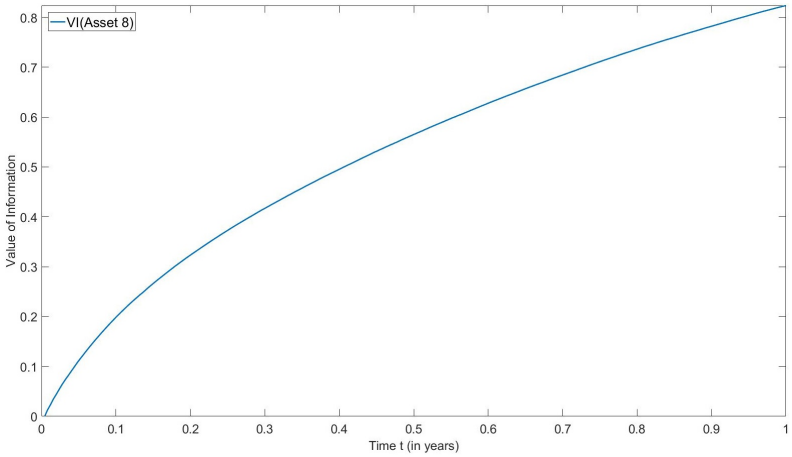


Figure 4: Value of information as a function of t with parameters as in Table 4.

Table 4: Parameter values used in Figure 4

	$\mathbf{b_0}$	$\boldsymbol{\sigma}$	$\boldsymbol{\sigma_0}$	**T**	$\boldsymbol{\vartheta}$	**Simulations**
Asset 8	5%	20%	40%	1	$(10\%)^2$	1,000,000

As Figure 4 shows, the value of information increases monotonically with time. Nonetheless, the speed of increase tends to slow as time goes by. The fact that the marginal gain on the value of information decreases with time is well known and analyzed in the recent article by Keppo et al. (2018) in the context of investment decisions and costs of data analytics.

5.2.4 Investment horizon

Figure 5 shows the value of information w.r.t. the investment time horizon T for a sample of three assets with parameters described in Table 5.

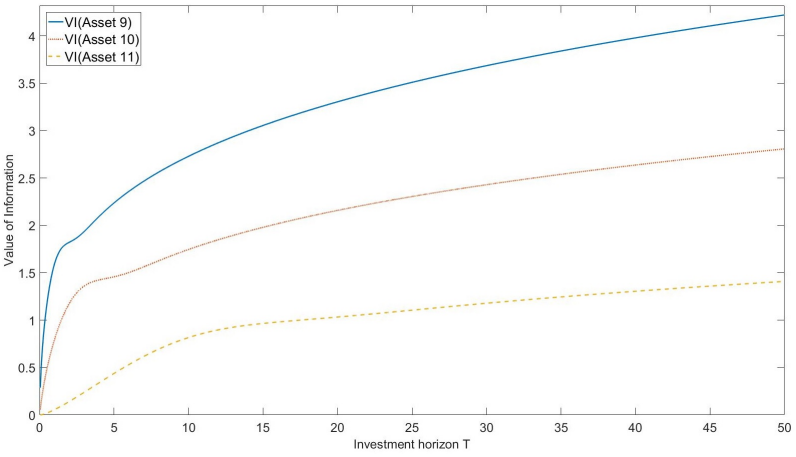


Figure 5: Value of information as a function of T with parameters as in Table 5.

Table 5: Parameter values used in Figure 5.

	$\mathbf{b_0}$	$\boldsymbol{\sigma}$	**Sharpe ratio**	**T**	$\boldsymbol{\sigma_0}$
Asset 9	5%	20%	25%	[0-50]	75%
Asset 10	5%	20%	25%	[0-50]	35%
Asset 11	5%	20%	25%	[0-50]	10%

Although the three curves are increasing no matter the value of the other parameters, the slope of these curves are mainly decreasing with the investment horizon. It suggests that the investor should consider that the marginal contribution to the value of information decreases with the investment horizon. It means that the spread between the optimal, Bayesian-Markowitz, and a sub-optimal, non-learning, strategy tightens for long investment horizons. It makes sense since for long horizons, at some point, the posterior distribution will not move so much since the values of the parameters of the distribution will be marginally affected by more learning. In addition, we remark that the bigger the drift volatility parameter, the higher the level of the associated value of information curve. This confirms the analysis of Figure 1 and 3.

6 Conclusion

In this paper, we have solved the Bayesian-Markowitz problem when the unknown drift is assumed to follow a probability distribution corresponding to a prior belief. The investor then updates her information on the drift from a predictive distribution based on observed data coming from assets prices. We have turned the non-standard problem into a standard one to exhibit the HJB equation associated to the standardized problem and apply dynamic programming techniques. To illustrate our theoretical results, we have computed the key diffusion coefficient ψ in the case of the multidimensional discrete law and the Gaussian law, and provided the full solution to the problem in the multidimensional Gaussian case. We have described a way of measuring the performance of investment strategies, the Sharpe ratio of terminal wealth, and used it to assess the value of information. To exhibit the value added of implementing a Bayesian learning approach in our problem, we have illustrated the value of information sensitivity to various key parameters and concluded that learning brings value to the optimal strategy that solves the Markowitz problem in a framework of drift uncertainty modeled, by a prior distribution with a positive definite covariance matrix, and a constant known volatility.

7 Appendices

A Proof of Lemma 2.3

Proof. We use similar arguments as in Ismail and Pham (2017).

Fix $\vartheta > 0$, for any $\epsilon > 0$, there exists an ϵ-optimal control for $U_0(\vartheta)$, $\alpha^\epsilon \in \mathcal{A}$ such that

$$U_0(\vartheta) \leq \mathbb{E}[X_T^{\alpha^\epsilon}] + \epsilon \text{ and } \mathrm{Var}(X_T^{\alpha^\epsilon}) \leq \vartheta.$$

Then for any $\lambda > 0$,

$$V_0(\lambda) \leq \lambda \mathrm{Var}(X_T^{\alpha^\epsilon}) - \mathbb{E}[X_T^{\alpha^\epsilon}] \leq \lambda\vartheta - U_0(\vartheta) + \epsilon.$$

Since ϵ is arbitrary and the above relation holds for any ϑ we deduce,

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

Conversely, fix $\lambda > 0$ and $\epsilon > 0$ and consider the ϵ-optimal control $\alpha^{*,\lambda,\epsilon} \in \mathcal{A}$ for V_0 such that

$$V_0(\lambda) \geq \lambda \mathrm{Var}(X_T^{\alpha^{*,\lambda,\epsilon}}) - \mathbb{E}[X_T^{\alpha^{*,\lambda,\epsilon}}] - \epsilon.$$

If we define

$$\vartheta^{\lambda,\epsilon} := \mathrm{Var}(X_T^{\alpha^{*,\lambda,\epsilon}}),$$

by definition of $U_0(\vartheta^{\lambda,\epsilon})$ the inequality $\mathbb{E}[X_T^{\alpha^{*,\lambda,\epsilon}}] \leq U_0(\vartheta^{\lambda,\epsilon})$ holds true.
Hence,

$$V_0(\lambda) \geq \lambda\vartheta^{\lambda,\epsilon} - \mathbb{E}[X_T^{\alpha^{*,\lambda,\epsilon}}] - \epsilon \geq \lambda\vartheta^{\lambda,\epsilon} - U_0(\vartheta^{\lambda,\epsilon}) - \epsilon$$

which holds for arbitrary $\epsilon > 0$. This shows the first duality relation, i.e. V_0 is the Fenchel Legendre transform of U_0, and ϑ^λ attains the infimum, namely

$$V_0(\lambda) = \inf_{\vartheta>0}[\lambda\vartheta - U_0(\vartheta)] = \lambda\vartheta^\lambda - U_0(\vartheta^\lambda).$$

We now prove the second duality relation. Fix $\vartheta > 0$ and for any $\lambda > 0$ consider the optimal control $\hat{\alpha}$ for U_0, we then have

$$\lambda\vartheta - U_0(\vartheta) \geq \lambda \mathrm{Var}(X_T^{\hat{\alpha}}) - \mathbb{E}[X_T^{\hat{\alpha}}] \geq V_0(\lambda).$$

The previous relation holds for any $\lambda > 0$, thus

$$U_0(\vartheta) \leq \inf_{\lambda>0}[\lambda\vartheta - V_0(\lambda)]. \tag{A.1}$$

Now fix $\vartheta > 0$ and consider $\alpha^{*,\lambda}$ the optimal control for V_0, we have

$$V_0(\lambda) = \lambda \mathrm{Var}(X_T^{\alpha^{*,\lambda}}) - \mathbb{E}[X_T^{\alpha^{*,\lambda}}].$$

Choose $\lambda(\vartheta)$ such that $\mathrm{Var}(X_T^{\alpha^{*,\lambda(\vartheta)}}) = \vartheta$. We then have

$$\begin{aligned} V_0(\lambda(\vartheta)) &= \lambda(\vartheta)\mathrm{Var}(X_T^{\alpha^{*,\lambda(\vartheta)}}) - \mathbb{E}[X_T^{\alpha^{*,\lambda(\vartheta)}}] \\ &= \lambda(\vartheta)\vartheta - \mathbb{E}[X_T^{\alpha^{*,\lambda(\vartheta)}}] \\ &\geq \lambda(\vartheta)\vartheta - U_0(\vartheta) \end{aligned}$$

by definition of U_0. Together with A.1, this shows the second duality relation; U_0 is the Fenchel-Legendre transform of V_0 and $\lambda(\vartheta)$ attains the infimum in this transform

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

Finally, for $\lambda > 0$, if $\alpha^{*,\lambda} \in \mathcal{A}$ is the optimal control for V_0 i.e. for problem (2.4) then

$$V_0(\lambda) = \lambda \mathrm{Var}(X_T^{\alpha^{*,\lambda}}) - \mathbb{E}[X_T^{\alpha^{*,\lambda}}] = \lambda\vartheta^\lambda - U_0(\vartheta^\lambda)$$

with $\vartheta^\lambda := \mathrm{Var}(X_T^{\alpha^{*,\lambda}})$. Moreover, we have

$$\begin{aligned} U_0(\vartheta) = \inf_{\lambda>0}[\lambda\vartheta - V_0(\lambda)] &= \lambda(\vartheta)\vartheta - V_0(\lambda(\vartheta)) \\ &= \lambda(\vartheta)\vartheta - \lambda(\vartheta)\vartheta^{\lambda(\vartheta)} + \mathbb{E}[X_T^{\alpha^{*,\lambda(\vartheta)}}] \\ &= \mathbb{E}[X_T^{\hat{\alpha}^\vartheta}] \end{aligned}$$

since $\vartheta^{\lambda(\vartheta)} = \vartheta$, by definition of $\lambda(\vartheta)$. This means that $\hat{\alpha}^\vartheta$ is an optimal control for U_0 i.e. problem (2.3). □

B Proof of Lemma 2.4

Proof. We use a similar approach as in Zhou and Li (2000).
Since $\forall \alpha \in \mathcal{A}$, we can write

$$\mathbb{E}[X_T^\alpha]^2 = -\inf_{\gamma\in\mathbb{R}} \left[\gamma^2 - 2\gamma\mathbb{E}[X_T^\alpha]\right]$$

where the infimum is achieved for $\gamma^* = \mathbb{E}[X_T^\alpha]$, we then have

$$\begin{aligned}
V_0(\lambda) &= \inf_{\alpha\in\mathcal{A}} \left[\lambda \mathrm{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha]\right] \\
&= \inf_{\alpha\in\mathcal{A}} \left[\lambda\big(\mathbb{E}[(X_T^\alpha)^2] - \mathbb{E}[X_T^\alpha]^2\big) - \mathbb{E}[X_T^\alpha]\right] \\
&= \inf_{\alpha\in\mathcal{A}} \left[\lambda\left(\mathbb{E}[(X_T^\alpha)^2] + \inf_{\gamma\in\mathbb{R}}\left[\gamma^2 - 2\gamma\mathbb{E}[X_T^\alpha]\right]\right) - \mathbb{E}[X_T^\alpha]\right] \\
&= \inf_{\alpha\in\mathcal{A}} \left[\inf_{\gamma\in\mathbb{R}}\left[\lambda\left(\mathbb{E}[(X_T^\alpha)^2] + \gamma^2 - 2\gamma\mathbb{E}[X_T^\alpha]\right) - \mathbb{E}[X_T^\alpha]\right]\right] \\
&= \inf_{\gamma\in\mathbb{R}} \left[\inf_{\alpha\in\mathcal{A}}\left[\lambda\mathbb{E}[(X_T^\alpha)^2] - (1+2\lambda\gamma)\mathbb{E}[X_T^\alpha]\right] + \lambda\gamma^2\right] \\
&= \inf_{\gamma\in\mathbb{R}} \left[\tilde{V}_0(\lambda,\gamma) + \lambda\gamma^2\right],
\end{aligned}$$

where $\tilde{V}_0(\lambda,\gamma) := \inf_{\alpha\in\mathcal{A}}\left[\lambda\mathbb{E}[(X_T^\alpha)^2] - (1+2\lambda\gamma)\mathbb{E}[X_T^\alpha]\right]$ as in (2.6), which proves (2.5).

The function $\tilde{V}_0$ is clearly linear in γ, so $\tilde{V}_0(\lambda,\gamma) + \lambda\gamma^2$ is strictly convex in γ and the infimum $V_0(\lambda)$ exists for some unique $\gamma^*(\lambda) := \mathrm{argmin}_{\gamma\in\mathbb{R}}[\tilde{V}_0(\lambda,\gamma^*) + \lambda(\gamma^*)^2]$. Take now this $\gamma^*(\lambda)$, $\tilde{\alpha}^{\lambda,\gamma}$ an optimal control for $\tilde{V}_0(\lambda,\gamma)$ and $\alpha^{*,\lambda} = \tilde{\alpha}^{\lambda,\gamma^*(\lambda)}$.
Then, dropping the dependence in λ of $\gamma^*(\lambda)$ to alleviate notations, we have

$$\operatorname*{argmin}_{\gamma\in\mathbb{R}}[\tilde{V}_0(\lambda,\gamma^*) + \lambda(\gamma^*)^2] \quad = \operatorname*{argmin}_{\gamma\in\mathbb{R}}\left[\lambda\mathbb{E}[(X_T^{\alpha^{*,\lambda}})^2] - (1+2\lambda\gamma^*)\mathbb{E}[X_T^{\alpha^{*,\lambda}}] + \lambda(\gamma^*)^2\right]$$

and find

$$\gamma^* = \mathbb{E}[X_T^{\alpha^{*,\lambda}}],$$

simply by differentiating the strictly convex function $\gamma^* \mapsto \lambda\mathbb{E}[(X_T^{\alpha^{*,\lambda}})^2] - (1+2\lambda\gamma^*)\mathbb{E}[X_T^{\alpha^{*,\lambda}}] + \lambda(\gamma^*)^2$.
Consequently, we have

$$\begin{aligned}
V_0(\lambda) &= \tilde{V}_0(\lambda,\gamma^*) + \lambda(\gamma^*)^2 \\
&= \lambda\mathbb{E}[(X_T^{\alpha^{*,\lambda}})^2] - (1+2\lambda\gamma^*)\mathbb{E}[X_T^{\alpha^{*,\lambda}}] + \lambda(\gamma^*)^2 \\
&= \lambda\left(\mathbb{E}[(X_T^{\alpha^{*,\lambda}})^2] - 2\gamma^*\mathbb{E}[X_T^{\alpha^{*,\lambda}}] + (\gamma^*)^2\right) - \mathbb{E}[X_T^{\alpha^{*,\lambda}}] \\
&= \lambda\mathrm{Var}(X_T^{\alpha^{*,\lambda}}) - \mathbb{E}[X_T^{\alpha^{*,\lambda}}]
\end{aligned}$$

which shows that $\alpha^{*,\lambda}$ is an optimal control for V_0. □

C Proof of theorem 4.1

Proof. Following the standard dynamic programming approach as in Pham (2009), we derive the HJB equation for the standard control problem (4.1):

$$\begin{cases} \partial_t v^{\lambda,\gamma} + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal \mathcal{D}_b^2 v^{\lambda,\gamma}\right) + \inf\limits_{\alpha\in\mathcal{A}}\left\{\alpha_t^\intercal b\partial_x v^{\lambda,\gamma} + \frac{1}{2}|\sigma^\intercal\alpha_t|^2\partial_{xx}^2 v^{\lambda,\gamma} + \alpha_t^\intercal\sigma\psi^\intercal\partial_{xb}^2 v^{\lambda,\gamma}\right\} = 0, \\ v^{\lambda,\gamma}(T,x,b) = \lambda x^2 - (1+2\lambda\gamma)x, \end{cases} \tag{C.1}$$

where the matrix function $\psi(t,b)$ is noted ψ to alleviate notations.

Assuming for now that $\partial_{xx}^2 v^{\lambda,\gamma} \geq 0$, we find

$$\underset{\alpha\in\mathcal{A}}{\mathrm{argmin}}\left\{\alpha_t^\intercal b\partial_x v^{\lambda,\gamma} + \frac{1}{2}|\sigma^\intercal\alpha_t|^2\partial_{xx}^2 v^{\lambda,\gamma} + \alpha_t^\intercal\sigma\psi^\intercal\partial_{xb}^2 v^{\lambda,\gamma}\right\} = -\Sigma^{-1}b\frac{\partial_x v^{\lambda,\gamma}}{\partial_{xx}^2 v^{\lambda,\gamma}} - (\psi\sigma^{-1})^\intercal\frac{\partial_{xb}^2 v^{\lambda,\gamma}}{\partial_{xx}^2 v^{\lambda,\gamma}} \tag{C.2}$$

which turns the standard control problem (C.1) into the same following problem

$$\begin{cases} \partial_t v^{\lambda,\gamma} + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal \mathcal{D}_b^2 v^{\lambda,\gamma}\right) - \frac{1}{2\partial_{xx}^2 v^{\lambda,\gamma}}|\partial_x v^{\lambda,\gamma}\sigma^{-1}b + \psi^\intercal\partial_{xb}^2 v^{\lambda,\gamma}|^2 = 0 \\ v^{\lambda,\gamma}(T,x,b) = \lambda x^2 - (1+2\lambda\gamma)x. \end{cases} \tag{C.3}$$

We look for a solution in the ansatz form $v^{\lambda,\gamma}(t,x,b) = K(t,b)x^2 + \Gamma(t,b)x + \chi(t,b)$ with $K \geq 0$ which ensures $\partial_{xx}^2 v^{\lambda,\gamma} \geq 0$ and (C.2). Formally we derive,

$$\begin{aligned} \partial_t v^{\lambda,\gamma} &= \partial_t K x^2 + \partial_t\Gamma x + \partial_t\chi, & \partial_x v^{\lambda,\gamma} &= 2Kx + \Gamma, \\ \partial_{xx}^2 v^{\lambda,\gamma} &= 2K, & \nabla_b v^{\lambda,\gamma} &= \nabla_b K x^2 + \nabla_b\Gamma x + \nabla_b\chi, \\ \mathcal{D}_b^2 v^{\lambda,\gamma} &= \mathcal{D}_b^2 K x^2 + \mathcal{D}_b^2\Gamma x + \mathcal{D}_b^2\chi, & \partial_{xb}^2 v^{\lambda,\gamma} &= 2\nabla_b K x + \nabla_b\Gamma. \end{aligned}$$

Plugging the previous partial derivatives into (C.3), the problem becomes:

$$\begin{cases} 0 = \left(\partial_t K + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal\mathcal{D}_b^2 K\right) - \left(2(\psi\sigma^{-1}b)^\intercal\nabla_b K + \frac{|\psi^\intercal\nabla_b K|^2}{K} + |\sigma^{-1}b|^2 K\right)\right)x^2 \\ + \left(\partial_t\Gamma + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal\mathcal{D}_b^2\Gamma\right) - \left(\left((\psi\sigma^{-1}b)^\intercal + \frac{(\psi\psi^\intercal\nabla_b K)^\intercal}{K}\right)\nabla_b\Gamma + \left(|\sigma^{-1}b|^2 + \left(\psi\sigma^{-1}b\right)^\intercal\frac{\nabla_b K}{K}\right)\Gamma\right)\right)x \\ + \partial_t\chi + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal\mathcal{D}_b^2\chi\right) - \frac{1}{4K}|\sigma^{-1}b\Gamma + \psi^\intercal\nabla_b\Gamma|^2, \\ v(T,x,b) = \lambda x^2 - (1+2\lambda\gamma)x, \end{cases}$$

and by identification we obtain the following system of PDEs:

$$\begin{cases} 0 = \partial_t K + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal\mathcal{D}_b^2 K\right) - \frac{1}{K}|\sigma^{-1}bK + \psi^\intercal\nabla_b K|^2, \\ 0 = \partial_t\Gamma + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal\mathcal{D}_b^2\Gamma\right) - \left(\left((\psi\sigma^{-1}b)^\intercal + \frac{(\psi\psi^\intercal\nabla_b K)^\intercal}{K}\right)\nabla_b\Gamma + \left(|\sigma^{-1}b|^2 + \left(\psi\sigma^{-1}b\right)^\intercal\frac{\nabla_b K}{K}\right)\Gamma\right), \\ 0 = \partial_t\chi + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal\mathcal{D}_b^2\chi\right) - \frac{1}{4K}|\sigma^{-1}b\Gamma + \psi^\intercal\nabla_b\Gamma|^2, \end{cases}$$

with terminal conditions

$$\begin{cases} K(T,b) = \lambda, \\ \Gamma(T,b) = -(1+2\lambda\gamma), \\ \chi(T,b) = 0. \end{cases}$$

We now introduce the functions $K = \lambda\tilde{K}$, $\Gamma = -(1+2\lambda\gamma)\tilde{\Gamma}$ and $\chi = (1+2\lambda\gamma)^2\tilde{\chi}/\lambda$, so that the previous system becomes

$$\begin{cases} 0 = \partial_t \tilde{K} + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal \mathcal{D}_b^2 \tilde{K}\right) - \frac{1}{\tilde{K}}|\sigma^{-1}b\tilde{K} + \psi^\intercal \nabla_b \tilde{K}|^2, \\ 0 = \partial_t \tilde{\Gamma} + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal \mathcal{D}_b^2 \tilde{\Gamma}\right) - \left(\left(\left(\psi\sigma^{-1}b\right)^\intercal + \frac{(\psi\psi^\intercal \nabla_b \tilde{K})^\intercal}{\tilde{K}}\right)\nabla_b\tilde{\Gamma} + \left(|\sigma^{-1}b|^2 + (\psi\sigma^{-1}b)^\intercal \frac{\nabla_b \tilde{K}}{\tilde{K}}\right)\tilde{\Gamma}\right), \\ 0 = \partial_t \tilde{\chi} + \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal \mathcal{D}_b^2 \tilde{\chi}\right) - \frac{1}{4\tilde{K}}|\sigma^{-1}b\tilde{\Gamma} + \psi^\intercal \nabla_b \tilde{\Gamma}|^2, \end{cases}$$

with terminal conditions

$$\begin{cases} \tilde{K}(T,b) = 1, \\ \tilde{\Gamma}(T,b) = 1, \\ \tilde{\chi}(T,b) = 0. \end{cases}$$

If there exists a unique solution of the equation for $\tilde{K}$, then $\tilde{\Gamma} = \tilde{K}$ is the unique solution that verifies the linear PDE for $\tilde{\Gamma}$ and $\tilde{\chi} = (\tilde{K}-1)/4$ is the unique solution of the linear PDE for $\tilde{\chi}$.

Since we look for positive $\tilde{K}$, we introduce $\tilde{K}(t,b) := e^{-R(t,b)}$ and rewrite the related PDE:

$$\begin{cases} 0 = \tilde{K}\left(-\partial_t R - \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal \mathcal{D}_b^2 R - \psi\psi^\intercal \nabla_b R \nabla_b R^\intercal\right) + 2\left(\psi\sigma^{-1}b\right)^\intercal \nabla_b R - |\psi^\intercal(\nabla_b R)|^2 - |\sigma^{-1}b|^2\right), \\ 0 = R(T,b). \end{cases}$$

From $\mathrm{tr}\left(\psi\psi^\intercal \nabla_b R \nabla_b R^\intercal\right) = |\psi^\intercal \nabla_b R|^2$ we finally obtain the expression of the semi-linear PDE in (4.2):

$$\begin{cases} 0 = -\partial_t R - \frac{1}{2}\mathrm{tr}\left(\psi\psi^\intercal \mathcal{D}_b^2 R\right) + 2\left(\psi\sigma^{-1}b\right)^\intercal \nabla_b R - \frac{1}{2}|\psi^\intercal(\nabla_b R)|^2 - |\sigma^{-1}b|^2, \\ 0 = R(T,b). \end{cases}$$

We obtain the following value function dependent upon (λ,γ):

$$\begin{aligned} v^{\lambda,\gamma}(t,x,b) =& \lambda\tilde{K}(t,b)x^2 - (1+2\lambda\gamma)\tilde{K}(t,b)x + \frac{(1+2\lambda\gamma)^2}{\lambda}\frac{\tilde{K}(t,b)-1}{4} \\ =& e^{-R(t,b)}\left(\lambda x^2 - (1+2\lambda\gamma)x + \frac{(1+2\lambda\gamma)^2}{4\lambda}\right) - \frac{(1+2\lambda\gamma)^2}{4\lambda}, \end{aligned} \tag{C.4}$$

and from Eq. (C.2), the optimal feedback control $\tilde{\alpha}^{\lambda,\gamma}$ is:

$$\tilde{\alpha}^{\lambda,\gamma} = \tilde{a}^{\lambda,\gamma}(t, X_t^{\tilde{\alpha}^{\lambda,\gamma}}, b),$$

with

$$\tilde{a}^{\lambda,\gamma}(t,x,b) = \left(\frac{1}{2\lambda} + \gamma - x\right)\left(\Sigma^{-1}b - (\psi\sigma^{-1}).^\intercal \nabla_b R(t,b)\right).$$

Moreover, from $\tilde{V}_0(\lambda,\gamma) = v^{\lambda,\gamma}(0,x_0,b_0)$ we compute

$$\begin{aligned}
\gamma^*(\lambda) =& \underset{\gamma\in\mathbb{R}}{\text{argmin}} \left[\tilde{V}_0(\lambda,\gamma) + \lambda\gamma^2\right] \\
=& \underset{\gamma\in\mathbb{R}}{\text{argmin}} \left[e^{-R(0,b_0)}\left(\lambda x_0^2 - (1+2\lambda\gamma)x_0 + \frac{(1+2\lambda\gamma)^2}{4\lambda}\right) - \frac{(1+2\lambda\gamma)^2}{4\lambda} + \lambda\gamma^2\right] \\
=& \underset{\gamma\in\mathbb{R}}{\text{argmin}} \left[e^{-R(0,b_0)}\left(\lambda x_0^2 - x_0 + \frac{1}{4\lambda}\right) - \frac{1}{4\lambda} + \left((1-2\lambda x_0)e^{-R(0,b_0)} - 1\right)\gamma + \lambda e^{-R(0,b_0)}\gamma^2\right] \\
=& x_0 + \frac{1}{2\lambda}\left(e^{R(0,b_0)} - 1\right).
\end{aligned}$$

Remembering (C.4), the value function at time 0 with the optimal γ is computed as followed:

$$\begin{aligned}
v^{\lambda,\gamma^*}(0,x_0,b_0) =& e^{-R(0,b_0)}\left(\lambda x_0^2 - (2\lambda x_0 + e^{R(0,b_0)}x_0)\right) + \left(\lambda x_0^2 + \frac{e^{2R(0,b_0)}}{4\lambda} + x_0 e^{R(0,b_0)}\right)\left(e^{-R(0,b_0)} - 1\right) \\
=& -\lambda x_0^2 + \frac{e^{R(0,b_0)}}{4\lambda}\left(1 - e^{R(0,b_0)}\right) - x_0 e^{R(0,b_0)},
\end{aligned}$$

and using Eq. (2.5), we find the expression in (4.13):

$$\begin{aligned}
V_0(\lambda) &= \tilde{V}_0(\lambda,\gamma^*) + \lambda(\gamma^*)^2 = v^{\lambda,\gamma^*}(0,x_0,b_0) + \lambda(\gamma^*)^2 \\
&= \frac{e^{R(0,b_0)}}{4\lambda}\left(1 - e^{R(0,b_0)}\right) - x_0 e^{R(0,b_0)} + \lambda x_0^2 + \frac{e^{2R(0,b_0)}}{4\lambda} + \frac{1}{4\lambda} + x_0 e^{R(0,b_0)} - x_0 - \frac{e^{R(0,b_0)}}{2\lambda} \\
&= -\frac{1}{4\lambda}\left(e^{R(0,b_0)} - 1\right) - x_0.
\end{aligned}$$

From Lemma (2.4), we find the optimal terminal wealth in (4.7)

$$\mathbb{E}[X_T^{\alpha^*,\lambda}] = \gamma^*(\lambda) = x_0 + \frac{1}{2\lambda}\left(e^{R(0,b_0)} - 1\right),$$

obtained with the optimal feedback control:

$$\alpha_t^{*,\lambda} = \tilde{a}^{\lambda,\gamma^*(\lambda)}(t, X_t^{\alpha^*,\lambda}, \hat{B}_t) = a_0^{Bayes,\lambda}(t, X_t^{\alpha^*,\lambda}, \hat{B}_t),$$

with

$$a_0^{Bayes,\lambda}(t,x,b) = \left(x_0 - x + \frac{e^{R(0,b_0)}}{2\lambda}\right)\left(\Sigma^{-1}b - (\psi\sigma^{-1})^\intercal \nabla_b R(t,b)\right). \tag{C.5}$$

The Lagrange multiplier $\lambda(\vartheta)$ which makes the variance of the optimal wealth equals to the variance constraint ϑ is calculated as follows. We know that,

$$V_0(\lambda(\vartheta)) = -\frac{1}{4\lambda(\vartheta)}\left(e^{R(0,b_0)} - 1\right) - x_0, \tag{C.6}$$

and from Eq. (2.4), we explicitly compute

$$\begin{aligned}
V_0(\lambda(\vartheta)) &= \lambda(\vartheta)\text{Var}(X_T^{\hat{\alpha}^\vartheta}) - \mathbb{E}[X_T^{\hat{\alpha}^\vartheta}] \\
&= \lambda(\vartheta)\vartheta - \mathbb{E}[X_T^{\hat{\alpha}^\vartheta}] \\
&= \lambda(\vartheta)\vartheta - x_0 - \frac{1}{2\lambda(\vartheta)}\left(e^{R(0,b_0)} - 1\right).
\end{aligned} \tag{C.7}$$

From Eq. (C.6) and (C.7) we obtain (4.8), namely

$$\lambda(\vartheta) = \sqrt{\frac{1}{4\vartheta}\left(e^{R(0,b_0)} - 1\right)}. \tag{C.8}$$

From the correspondance of the controls established in Lemma 2.3, we know that

$$\hat{\alpha}^{\vartheta} = \alpha^{*,\lambda(\vartheta)}.$$

We then compute the optimal controls of the Bayesian-Markowitz problem U_0 as

$$\hat{\alpha}^{\vartheta} = a_0^{Bayes,\lambda(\vartheta)}(t, X_t^{\hat{\alpha}^{\vartheta}}, \hat{B}_t),$$

where, from Eq. (C.5) with $\lambda = \lambda(\vartheta)$ as in (C.8),

$$a_0^{Bayes,\lambda(\vartheta)}(t,x,b) \;=\; \left(x_0 - x + e^{R(0,b_0)}\sqrt{\frac{\vartheta}{e^{R(0,b_0)}-1}}\right)\left(\Sigma^{-1}b - (\psi\sigma^{-1})^{\intercal}\nabla_b R(t,b)\right).$$

Finally from (C.7) and (C.8), we deduce the optimal performance of the Bayesian-Markowitz problem:

$$\begin{aligned} U_0(\vartheta) \;&=\; \lambda(\vartheta)\vartheta - V_0(\lambda(\vartheta)) \\ &=\; x_0 + \frac{1}{2\lambda(\vartheta)}\left(e^{R(0,b_0)} - 1\right) \;\;=\; x_0 + \sqrt{\vartheta(e^{R(0,b_0)} - 1)}. \end{aligned}$$

□

D Proofs of theorem 4.3

Without loss of generality, we consider in this demonstration the function $\tilde{R} := -R$. The value of the positive constant C may change from line to line and the matrix function $\psi(t,b)$ will be simply noted ψ to alleviate notations.

We define the nonlinear function $F : \mathcal{R} \times \mathbb{R}^n \to \mathbb{R}$ by:

$$F(t,b,p) \;=\; \frac{1}{2}|\psi^{\intercal}p|^2 + 2(\psi\sigma^{-1}b)^{\intercal}p + |\sigma^{-1}b|^2,$$

and we introduce the function $L_t : \mathcal{B}_t \times \mathbb{R}^n \to \mathbb{R}$ by

$$L_t(b,q) \;:=\; \max_{p\in\mathbb{R}^n}\left[-q^{\intercal}\psi^{\intercal}p - F(t,b,p)\right].$$

We notice that F is quadratic hence convex in p, so it is easy to see that for fixed $(t,b) \in \mathcal{R}$ the function $p \mapsto -q^{\intercal}\psi^{\intercal}p - F(t,b,p)$ reaches a maximum $p^* = -(\psi^{\intercal})^{-1}(q + 2\sigma^{-1}b)$. This gives us the explicit form of the function L_t which depends on t only through $\mathcal{B}_t$:

$$\begin{aligned} L_t(b,q) &= -q^{\intercal}\psi^{\intercal}p^* - F(t,b,p^*) \\ &= \frac{1}{2}|q|^2 + 2(\sigma^{-1}b)^{\intercal}q + |\sigma^{-1}b|^2. \end{aligned}$$

Conversely, the function L_t is convex in q so for fixed $b \in \mathcal{B}_t$ the function $q \mapsto -q^{\intercal}\psi^{\intercal}p - L_t(b,q)$ reaches its maximum for $q^* = -\left(\psi^{\intercal}p + 2\sigma^{-1}b\right)$. It shows that

$$\begin{aligned} \max_{q\in\mathbb{R}^n}\left[-q^{\intercal}\psi^{\intercal}p - L_t(b,q)\right] \;&=\; -(q^*)^{\intercal}\psi^{\intercal}p - L_t(b,q^*) \\ &=\; \frac{1}{2}|\psi^{\intercal}p|^2 + 2(\psi\sigma^{-1}b)^{\intercal}p + |\sigma^{-1}b|^2 \\ &=\; F(t,b,p), \end{aligned}$$

and it establishes the duality relation between the functions F and L. Let us know consider the truncated function $F^k : \mathcal{R} \times \mathbb{R}^n \to \mathbb{R}$ defined for each $k \in \mathbb{N}$ by

$$F^k(t,b,p) \;=\; \max_{q \in \mathcal{A}_k} \big[-q^\intercal \psi^\intercal p - L_t(b,q) \big].$$

We observe from the explicit form of L_t that $L_t \in C^1(\mathcal{B}_t \times \mathbb{R}^n)$ and L, $\nabla_b L$ satisfy a polynomial growth condition in b. Namely, we see that the following estimates hold true for $(b,q) \in \mathcal{B}_t \times \mathbb{R}^n$:

$$\begin{aligned} |L_t(b,q)| \;&\leq\; \frac{1}{2}|q|^2 + 2|\sigma^{-1}||b||q| + |\sigma^{-1}|^2|b|^2 \\ &\leq\; \frac{1}{2}|q|^2 + |\sigma^{-1}|(|q|^2 + |b|^2) + |\sigma^{-1}|^2|b|^2 \\ &\leq\; C(|q|^2 + |b|^2) \end{aligned}$$

and

$$\begin{aligned} |\nabla_b L_t(b,q)| \;&\leq\; 2(|(\sigma^{-1})||q| + |\Sigma^{-1}||b|) \\ &\leq\; C(|q| + |b|). \end{aligned}$$

So, $|L|$ and $|\nabla_b L|$ are of polynomial growth in b uniformly in $|q|$ when $|q| \leq k$. By classical theory (see Theorem 4.3 p 163 in Fleming and Soner (2006)), the previous estimates and the assumptions of the theorem tell us that there exists a unique quadratically growing smooth solution

$$\tilde{R}^k \in C^{1,2}\left([0,T) \times \mathcal{B}_t\right) \cap C(\mathcal{R}),$$

to the truncated semi-linear PDE,

$$-\partial_t \tilde{R}^k - \frac{1}{2}\mathrm{tr}(\psi\psi^\intercal \mathcal{D}_b^2 \tilde{R}^k) + F^k(t,b,\nabla_b \tilde{R}^k) \;=\; 0,$$

with terminal condition $\tilde{R}^k(T,.) = 0$.

We then know from the Feynman-Kac formula and standard arguments that $\tilde{R}^k$ can be represented as the solution of the stochastic control problem

$$\tilde{R}^k(t,b) = \inf_{q \in \mathcal{A}_k} \mathbb{E}\left[\int_t^\intercal L_t(\tilde{B}_s, q_s) ds | \tilde{B}_t = b \right] \tag{D.1}$$

where $\mathcal{A}_k$ is the compact set $\mathcal{A}_k = \{q \in \mathbb{R}^n : |q| \leq k\}, k > 0$, and the dynamic of $\tilde{B}$ is

$$d\tilde{B}_s = \psi(s, \tilde{B}_s)\left(q_s ds + dW_s\right). \tag{D.2}$$

Moreover, an optimal control for (D.1) is Markovian and given by

$$q_k^*(t,b) = \operatorname*{argmin}_{q \in \mathcal{A}_k} \left\{ q^\intercal \psi^\intercal \nabla_b \tilde{R}^k(t,b) + L_t(b,q) \right\}.$$

We then deduce that,

$$\tilde{R}^k(t,b) = \mathbb{E}\left[\int_t^\intercal L_s\left(\tilde{B}_s^*, q_k^*\left(s, \tilde{B}_s^*\right)\right) ds | \tilde{B}_t^* = b \right]$$

where $\tilde{B}_s^*$ solves the stochastic differential equation (SDE) (D.2) with controls $q_k^*\left(s, \tilde{B}_s^*\right)$. From the theorem of differentiation under the expectation and the integral sign we know that

$$\nabla_b \tilde{R}^k(t,b) = \mathbb{E}\left[\int_t^\intercal \nabla_b L_s\left(\tilde{B}_s^*, q_k^*\left(s, \tilde{B}_s^*\right)\right) ds | \tilde{B}_t^* = b\right].$$

To prove the linear growth condition of $\nabla_b \tilde{R}^k$, we will need the following inequality:

$$|\nabla_b L_t(b,q)|^2 \leq C_1 L_t(b,q) + C_2|b|^2,$$

with $C_1 > 0$ and $C_2 \geq 2C_1|\sigma^{-1}b|^2 > 0$.
Having in mind that the function L_t can be written as $L_t(b,q) = \frac{1}{2}|q + 2\sigma^{-1}|^2 - |\sigma^{-1}b|^2$, the previous inequality comes from:

$$\begin{aligned}
|\nabla_b L_t(b,q)|^2 &= |2(\sigma^{-1})^\intercal q + 2\Sigma^{-1}b|^2 \\
&\leq 4|\sigma^{-1}|^2|q + 2\sigma^{-1}b - \sigma^{-1}b|^2 \\
&\leq 8|\sigma^{-1}|^2\left(\frac{1}{2}|q + 2\sigma^{-1}b|^2 + |\sigma^{-1}b|^2\right) \\
&\leq C_1 L_t(b,q) + 2C_1|\sigma^{-1}|^2|b|^2 \\
&\leq C_1 L_t(b,q) + C_2|b|^2.
\end{aligned}$$

Then by the Cauchy-Schwartz inequality and the fact that $L_t(b,0) \leq C|b|^2$ for some positive constant C independent of k and t, we have

$$\begin{aligned}
|\nabla_b \tilde{R}^k(t,b)| &\leq \mathbb{E}\left[\int_t^T \left|\nabla_b L_s\left(\tilde{B}_s^*, q_k^*\left(s, \tilde{B}_s^*\right)\right)\right| ds | \tilde{B}_t^* = b\right] \\
&\leq C\mathbb{E}\left[\int_t^T \left|\nabla_b L_s\left(\tilde{B}_s^*, q_k^*\left(s, \tilde{B}_s^*\right)\right)\right|^2 ds | \tilde{B}_t^* = b\right]^{\frac{1}{2}} \\
&\leq C\left(C_1\mathbb{E}\left[\int_t^T L_s\left(\tilde{B}_s^*, q_k^*\left(s, \tilde{B}_s^*\right)\right) ds | \tilde{B}_t^* = b\right] + C_2\mathbb{E}\left[\int_t^T |\tilde{B}_s|^2 ds | \tilde{B}_t = b\right]\right)^{\frac{1}{2}} \\
&\leq C\left(C_1\mathbb{E}\left[\int_t^T L_s\left(\tilde{B}_s, 0\right) ds | \tilde{B}_t = b\right] + C_2\mathbb{E}\left[\int_t^T |\tilde{B}_s|^2 ds | \tilde{B}_t = b\right]\right)^{\frac{1}{2}} \\
&\leq C\mathbb{E}\left[\int_t^T |\tilde{B}_s|^2 ds | \tilde{B}_t = b\right]^{\frac{1}{2}} \\
&\leq C(1 + |b|).
\end{aligned}$$

We used that when $q := 0$, $\tilde{B}_s = \tilde{B}_t + \int_t^s \psi(u, \tilde{B}_u) dW_u$ and thanks to Ito's formula,

$$|\tilde{B}_s|^2 = |\tilde{B}_t|^2 + 2\tilde{B}_t^\intercal \int_t^s \psi(u, \tilde{B}_u) dW_u + \int_t^s \operatorname{tr}\left(\psi\psi^\intercal(u, \tilde{B}_u)\right) du.$$

Moreover,

$$\begin{aligned}
\mathbb{E}\left[\int_t^\intercal |\tilde{B}_s|^2 ds | \tilde{B}_t = b\right]^{\frac{1}{2}} &= \mathbb{E}\left[\int_t^\intercal \left(|\tilde{B}_t|^2 + 2\tilde{B}_t^\intercal \int_t^s \psi(u, \tilde{B}_u) dW_u + \int_t^s \operatorname{tr}\left(\psi\psi^\intercal(u, \tilde{B}_u)\right) du\right) ds | \tilde{B}_t = b\right]^{\frac{1}{2}} \\
&\leq \left((T-t)|b|^2 + n|\psi|^2(T-t)\right)^{\frac{1}{2}} \\
&\leq C(1 + |b|^2)^{\frac{1}{2}} \\
&\leq C\left(1 + |b|\right)
\end{aligned}$$

Consequently, since the function $q \to -q^\intercal \psi^\intercal \nabla_b R^k(t,b) - L_t(b,q)$ attains its maximum on $\mathbb{R}^n$ for

$$\tilde{q}_k(t,b) = -\left(\psi^\intercal \nabla_b R^k(t,b) + 2\sigma^{-1} b\right),$$

and knowing that by assumption $|\psi| < \infty$, it is easy to see that

$$|\tilde{q}_k(t,b)| \leq C(1+|b|).$$

As a consequence, for an arbitrarily large constant $\hat{C} > 0$, there exists a positive constant C independent of k such that,

$$|\tilde{q}_k(t,b)| \leq C, \quad t \in [0,T], \quad b \in \mathcal{B}_t \cap \{|b| \leq \tilde{C}\}.$$

Hence, for $k \geq C$, we have

$$\begin{aligned} F^k(t,b,\nabla_b \tilde{R}^k) &= \max_{q \in \mathcal{A}_k} \big[-q^\intercal \psi^\intercal \nabla_b \tilde{R}^k - L_t(b,q) \big], \\ &= \max_{q \in \mathbb{R}^n} \big[-q^\intercal \psi^\intercal \nabla_b \tilde{R}^k - L_t(b,q) \big], \\ &= F(t,b,\nabla_b \tilde{R}^k). \end{aligned}$$

for all $(t,b) \in [0,T] \times \mathcal{B}_t \cap \{|b| \leq \tilde{C}\}$. Letting $\tilde{C}$ goes to infinity implies that for k sufficiently large, $R^k = -\tilde{R}^k$ is a smooth solution satisfying a quadratic growth condition to (4.2) - (4.4).

E Proof of Lemma 4.5

Proof. To solve the system in the multidimensional Gaussian case, we solve the Riccati equation for M. We look for a symmetric solution M such that $GM = (GM)^\intercal$ which solves the following ODE

$$-M'(t) - 2M(t)^\intercal G(t)^\intercal \Sigma G(t) M(t) + 2G(t)M(t) + 2M(t)^\intercal G(t)^\intercal - \Sigma^{-1} = 0.$$

A particular solution to the ODE for M is $\hat{M}(t) = G^{-1}(t)\Sigma^{-1}$ and it is easy to see that $\hat{M}$ is symmetric since $\Sigma G = G^\intercal \Sigma$, and $G\hat{M} = (G\hat{M})^\intercal$.

Now we look for a function N such that $M = \hat{M} + N$. Note that N should be symmetric with $GN = (GN)^\intercal$. Plugging M into the ODE yields

$$-N'(t) - 2N^\intercal G(t)^\intercal \Sigma G(t) N(t) = 0.$$

Now, we change variable $N = \Theta^{-1}$ in the previous ODE to find

$$\Theta'(t) = 2G(t)^\intercal \Sigma G(t).$$

Noticing that $\Theta(T) = -\Sigma G(T) = -(\Sigma_0^{-1} + \Sigma^{-1}T)^{-1}$, we obtain after integration

$$\Theta(t) = -\left((\Sigma_0^{-1} + \Sigma^{-1}T)^{-1} + \int_t^\intercal 2G(s)^\intercal \Sigma G(s) ds\right).$$

Finally, writing $G(t)$ as $(\Sigma + \Sigma_0 t)^{-1}\Sigma_0$, the solution to the original ODE is

$$M(t) = (\hat{M} + \Theta^{-1})(t) = \Sigma_0^{-1} + \Sigma^{-1}t - \left[(\Sigma_0^{-1} + \Sigma^{-1}T)^{-1} + 2\int_t^{\intercal} \Sigma_0 \left(\Sigma + \Sigma_0 s\right)^{-1} \Sigma \left(\Sigma + \Sigma_0 s\right)^{-1} \Sigma_0 ds\right]^{-1}.$$

To obtain $U(t)$, we simply integrate and with some simplifications we obtain,

$$U(t) = \int_t^T \operatorname{tr}\left(\Sigma_0 \left(\Sigma + \Sigma_0 s\right)^{-1} - \left[\left((\Sigma_0^{-1} + \Sigma^{-1}T)^{-1} + 2\int_s^T G(u)^{\intercal}\Sigma G(u) du\right)\left(G(s)^{\intercal}\Sigma G(s)\right)^{-1}\right]^{-1}\right) ds.$$

It is easy to verify that $M(T) = U(T) = 0$.

The next step is to check that N satisfies the conditions we have imposed to derive the solution M. The condition $N = N^{\intercal}$ is satisfied since it is easy to see that $\Theta = \Theta^{\intercal}$. It is straightforward to see that the second condition, $GN = (GN)^{\intercal}$ is equivalent to $\Theta\Sigma_0^{-1}\Sigma = \Sigma\Sigma_0^{-1}\Theta$, once G is written as $(\Sigma + \Sigma_0 t)^{-1}\Sigma_0$. Noticing the symmetry of $G\Sigma_0^{-1}$, we develop the left side of the equality,

$$\begin{aligned}
\Theta\Sigma_0^{-1}\Sigma &= -\left((\Sigma_0^{-1} + \Sigma^{-1}T)^{-1} + \int_t^{\intercal} 2G(s)^{\intercal}\Sigma G(s) ds\right)\Sigma_0^{-1}\Sigma \\
&= -\left((\Sigma_0^{-1} + \Sigma^{-1}T)^{-1}\Sigma_0^{-1}\Sigma + \int_t^{\intercal} 2G(s)^{\intercal}\Sigma G(s)\Sigma_0^{-1}\Sigma ds\right) \\
&= -\left(\Sigma\left[(\Sigma\Sigma_0^{-1} + T)\Sigma_0\right]^{-1}\Sigma + \int_t^{\intercal} 2\Sigma G(s)\Sigma_0^{-1}G(s)^{\intercal}\Sigma ds\right) \\
&= -\left(\Sigma\Sigma_0^{-1}\left(\Sigma_0^{-1} + \Sigma^{-1}T\right)^{-1} + \int_t^{\intercal} 2\Sigma\Sigma_0^{-1}G^{\intercal}(s)\Sigma G(s) ds\right) \\
&= \Sigma\Sigma_0^{-1}\Theta.
\end{aligned}$$

□

For the one-dimensional case n = 1, we check that the functions:

$$M(t) = \frac{\sigma^2 + \sigma_0^2 t}{\sigma^2\left(\sigma^2 + \sigma_0^2(2T - t)\right)}(T - t) \quad \text{and} \quad U(t) = \log\left(\frac{\sigma^2 + \sigma_0^2 T}{\sqrt{(\sigma^2 + \sigma_0^2 t)(\sigma^2 + \sigma_0^2(2T - t))}}\right),$$

satisfy the following unidimensional ODE system:

$$\begin{cases} M'(t) = -\dfrac{2\sigma^2\sigma_0^4}{(\sigma^2 + \sigma_0^2 t)^2}M(t)^2 + \dfrac{4\sigma_0^2}{\sigma^2 + \sigma_0^2 t}M(t) - \sigma^{-2}, \\ U'(t) = -\dfrac{\sigma^2\sigma_0^4}{\sigma^2 + \sigma_0^2 t)^2}M(t), \end{cases}$$

with M(T) = U(T) = 0.
We easily see that M and U satisfy the terminal conditions. We first check that M satisfies the first ODE of the system. We derive the function M w.r.t. t,

$$M'(t) = \frac{\sigma^{-2}}{(\sigma^2 + \sigma_0^2(2T - t))^2}\left(\sigma_0^2(-2\sigma^2 t) + \sigma_0^4(2T^2 - 4Tt + t^2) - \sigma^4\right),$$

and notice that the right side of the equality for M yields:

$$
\begin{aligned}
&-\frac{2\sigma^2\sigma_0^4}{(\sigma^2+\sigma_0^2 t)^2}M(t)^2+\frac{4\sigma_0^2}{\sigma^2+\sigma_0^2 t}M(t)-\sigma^{-2}\\
&=\frac{\sigma^{-2}}{(\sigma^2+\sigma_0^2(2T-t))^2}\left(-2\sigma_0^4(T-t)^2+4\omega^2(T-t)(\sigma^2+\sigma_0^2(2T-t))-(\sigma^2+\sigma_0^2(2T-t))^2\right)\\
&=\frac{\sigma^{-2}}{(\sigma^2+\sigma_0^2(2T-t))^2}\left(\sigma_0^2(-2\sigma^2 t)+\sigma_0^4(2T^2-4Tt+t^2)-\sigma^4\right)\\
&=M'(t).
\end{aligned}
$$

We then verify the ODE for U by deriving the function U w.r.t. t:

$$
\begin{aligned}
U'(t)&=(\sigma^2+\sigma_0^2 T)\frac{\omega^2(2\sigma_0^2(T-t))}{2\left((\sigma^2+\sigma_0^2 t)(\sigma^2+\sigma_0^2(2T-t))\right)^{\frac{3}{2}}}\frac{\sqrt{(\sigma^2+\sigma_0^2 t)(\sigma^2+\sigma_0^2(2T-t))}}{\sigma^2+\sigma_0^2 T}\\
&=-\frac{\sigma_0^4(T-t)}{(\sigma^2+\sigma_0^2 t)(\sigma^2+\sigma_0^2(2T-t))},
\end{aligned}
$$

and from the right part of the ODE, we find the equality:

$$
\begin{aligned}
-\frac{\sigma^2\sigma_0^4}{(\sigma^2+\sigma_0^2 t)^2}M(t)&=-\frac{\sigma_0^4}{(\sigma^2+\sigma_0^2 t)^2}\frac{\sigma^2+\sigma_0^2 t}{\sigma^2+\sigma_0^2(2T-t)}(T-t)\\
&=-\frac{\sigma_0^4(T-t)}{(\sigma^2+\sigma_0^2 t)(\sigma^2+\sigma_0^2(2T-t))}\\
&=U'(t).
\end{aligned}
$$

References

Aguilar, O. and M. West (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics 18*(3), 338–357.

Avramov, D. and G. Zhou (2010). Bayesian portfolio analysis. *Annu. Rev. Financ. Econ. 2*(1), 25–47.

Bain, A. and D. Crisan (2009). *Fundamentals of stochastic filtering*. Number 60 in Stochastic Modelling and Applied Probability. Springer-Verlag New York.

Barry, C. B. (1974). Portfolio analysis under uncertain means, variances, and covariances. *The Journal of Finance 29*(2), 515–522.

Bauder, D., T. Bodnar, N. Parolya, and W. Schmid (2018, March). Bayesian mean-variance analysis: Optimal portfolio selection under parameter uncertainty. *ArXiv e-prints*.

Benth, F. E. and K. H. Karlsen (2005). A pde representation of the density of the minimal entropy martingale measure in stochastic volatility markets. *Stochastics An International Journal of probability and Stochastic Processes 77*(2), 109–137.

Best, M. J. and R. R. Grauer (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The review of financial studies 4*(2), 315–342.

Black, F. and R. Litterman (1992). Global portfolio optimization. *Financial analysts journal 48*(5), 28–43.

Bodnar, T., S. Mazur, and Y. Okhrin (2017). Bayesian estimation of the global minimum variance portfolio. *European Journal of Operational Research 256*(1), 292–307.

Cvitanić, J., A. Lazrak, L. Martellini, and F. Zapatero (2006). Dynamic portfolio choice with parameter uncertainty and the economic value of analysts? recommendations. *The Review of Financial Studies 19*(4), 1113–1156.

Ekstrom, E. and J. Vaicenavicius (2016). Optimal liquidation of an asset under drift uncertainty. *SIAM Journal on Financial Mathematics 7*(1), 357–381.

Fleming, W. H. and H. M. Soner (2006). *Controlled Markov processes and viscosity solutions*, Volume 25. Springer Science & Business Media.

Frost, P. A. and J. E. Savarino (1986). An empirical bayes approach to efficient portfolio selection. *Journal of Financial and Quantitative Analysis 21*(3), 293–305.

Guéant, O. and J. Pu (2017, 8). Portfolio choice, portfolio liquidation, and portfolio transition under drift uncertainty. *ArXiv e-prints*.

Ismail, A. and H. Pham (2017). Robust markowitz mean-variance portfolio selection under ambiguous covariance matrix.

Karatzas, I., J. P. Lehoczky, and S. E. Shreve (1987). Optimal portfolio and consumption decisions for a "small investor" on a finite horizon. *SIAM journal on control and optimization 25*(6), 1557–1586.

Karatzas, I. and X. Zhao (2001). *Bayesian Adaptative Portfolio Optimization*. Handbook of Mathematical Finance, Optimization, Pricing, Interest Rates, and Risk Management.

Keppo, J., H. M. Tan, and C. Zhou (2018). Investment decisions and falling cost of data analytics.

Klein, R. W. and V. S. Bawa (1976). The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics 3*(3), 215–231.

Krylov, N. V. (1987). *Nonlinear elliptic and parabolic equations of the second order*, Volume 7. Springer.

Ladyzhenskaia, O. A., V. A. Solonnikov, and N. N. Ural'ceva (1968). *Linear and quasi-linear equations of parabolic type*, Volume 23. American Mathematical Soc.

Lakner, P. (1995). Utility maximization with partial information. *Stochastic processes and their applications 56*(2), 247–273.

Lakner, P. (1998). Optimal trading strategy for an investor: the case of partial information. *Stochastic Processes and their Applications 76*(1), 77–97.

Markowitz, H. (1952, 3). Portfolio Selection. *The Journal of Finance 7*(1), 77–91.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, 247–257.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of financial economics 8*(4), 323–361.

Pham, H. (2002). Smooth solutions to optimal investment models with stochastic volatilities and portfolio constraints. *Applied Mathematics & Optimization 46*(1).

Pham, H. (2008). Portfolio optimization under partial observation: theoretical and numerical aspects.

Pham, H. (2009). *Continuous-time Stochastic Control and Optimization with Financial Applications* (1st ed.). Springer Publishing Company, Incorporated.

Rogers, L. C. G. (2001). The relaxed investor and parameter uncertainty. *Finance and Stochastics 5*(2), 131–154.

Samuelson, P. A. (1969, 8). Lifetime Portfolio Selection by Dynamic Stochastic Programming. *The Review of Economics and Statistics 51*(3), 239–246.

Zhou, X. Y. and D. Li (2000, 1). Continuous-Time Mean-Variance Portfolio Selection: A Stochastic LQ Framework. *Applied Mathematics & Optimization 42*(1), 19–33.