

Adaptive Robust Control Under Model Uncertainty

Tomasz R Bielecki, Tao Chen, Igor Cialenco, Areski Cousin, Monique Jeanblanc

▶ **To cite this version:**

Tomasz R Bielecki, Tao Chen, Igor Cialenco, Areski Cousin, Monique Jeanblanc. Adaptive Robust Control Under Model Uncertainty. SIAM Journal on Control and Optimization, 2019, 57 (2), pp.925-946. 10.1137/17M1137917 . hal-01582574

HAL Id: hal-01582574
https://hal.science/hal-01582574

Submitted on 6 Sep 2017

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.

Adaptive Robust Control Under Model Uncertainty

Tomasz R. Bielecki [a], Tao Chen, [a] Igor Cialenco, [a] Areski Cousin [b], Monique Jeanblanc [c]

First Circulated: June 06, 2017

Abstract: In this paper we propose a new methodology for solving an uncertain stochastic Markovian control problem in discrete time. We call the proposed methodology the adaptive robust control. We demonstrate that the uncertain control problem under consideration can be solved in terms of associated adaptive robust Bellman equation. The success of our approach is to the great extend owed to the recursive methodology for construction of relevant confidence regions. We illustrate our methodology by considering an optimal portfolio allocation problem, and we compare results obtained using the adaptive robust control method with some other existing methods.

Keywords: adaptive robust control, model uncertainty, stochastic control, dynamic programming, recursive confidence regions, Markov control problem, portfolio allocation.
MSC2010: 93E20, 93E35, 49L20, 60J05

1 Introduction

In this paper we propose a new methodology for solving an uncertain stochastic Markovian control problem in discrete time, which is then applied to an optimal portfolio selection problem. Accordingly, we only consider terminal optimization criterion.

The uncertainty in the problem comes from the fact that the (true) law of the underlying stochastic process is not known. What is known, we assume, is the family of potential probability laws that the true law belongs to. Thus, we are dealing here with a stochastic control problem subject to Knightian uncertainty.

Such problems have been extensively studied in the literature, using different approaches, some of them are briefly described in Section 2.

The classical approach to the problem goes back to several authors. The multiple-priors, or the maxmin approach, of [GS89] is probably one of the first ones and one of the best-known in the economics literature. In the context of our terminal criterion, it essentially amounts to a robust control problem (a game problem, in fact) of the form

$$\sup_{\varphi\in\mathcal{A}}\inf_{\mathbb{Q}\in\mathcal{M}}\mathbb{E}_{\mathbb{Q}}\,\mathcal{L}(X,\varphi,T),$$

where φ is the control process, $\mathcal{A}$ is the family of admissible controls, $\mathcal{M}$ is a family of possible underlying probabilistic models (or priors), $\mathbb{E}_{\mathbb{Q}}$ denotes expectation under the prior $\mathbb{Q}$, X is the underlying process, T is the finite optimization horizon, and where $\mathcal{L}$ is an optimization criterion. The family $\mathcal{M}$ represents the Knightian uncertainty.

[a] Department of Applied Mathematics, Illinois Institute of Technology
Emails: tbielecki@iit.edu (T. R. Bielecki), tchen29@iit.edu (T. Chen) and cialenco@iit.edu (I. Cialenco)
URLs: http://math.iit.edu/~bielecki and http://math.iit.edu/~igor

[b] Institut de Sciences Financière et d'Assurance, Université Lyon 1, Lyon, France
Email: areski.cousin@univ-lyon1.fr, URL: http://www.acousin.net/

[c] Univ Evry, LaMME UMR CNRS 807, Université Paris-Saclay, 91025, Evry, France
Email: monique.jeanblanc@univ-evry.fr, URL: http://www.maths.univ-evry.fr/pages_perso/jeanblanc

The above maxmin formulation has been further modified to the effect of

$$\sup_{\varphi\in\mathcal{A}}\inf_{Q\in\mathcal{M}}\mathbb{E}_{\mathbb{Q}}\left(\mathcal{L}(X,\varphi,T)-c(\mathbb{Q})\right),$$

where c is a penalty function. We refer to, e.g., [HSTW06], [Ski03], [MMR06] and [BMS07], and references therein, for discussion and various studies of this problem.

In our approach we do not use the penalty term. Instead, we apply a learning algorithm that is meant to reduce the uncertainty about the true probabilistic structure underlying the evolution of the process X. This leads us to consider what we call *adaptive robust control problem*. We stress that our problem and approach should not be confused with the problems and approaches presented in [IS12] and [BWH12].

A very interesting study of an uncertain control problem in continuous time, that involves learning, has been done in [KOZ14].

The paper is organized as follows. In Section 2 we briefly review some of the existing methodologies of solving stochastic control problems subject to model uncertainty, starting with robust control method, and continuing with strong robust method, model free robust method, Bayesian adaptive control method, and adaptive control method. Also here we introduce the underlying idea of the proposed method, called the adaptive robust control, and its relationship with the existing ones. Section 3 is dedicated to the adaptive robust control methodology. We begin with setting up the model and in Section 3.1 we formulate in strict terms the stochastic control problem. The solution of the adaptive robust problem is discussed in Section 3.2. Also in this section, we derive the associated Bellman equation and prove the Bellman principle of optimality for the considered adaptive robust problem. Finally, in Section 4 we consider an illustrative example, namely, the classical dynamic optimal allocation problem when the investor is deciding at each time on investing in a risky asset and a risk-free banking account by maximizing the expected utility of the terminal wealth. Also here, we give a comparative analysis between the proposed method and some of the existing classical methods.

2 Stochastic Control Subject to Model Uncertainty

Let $(\Omega,\mathscr{F})$ be a measurable space, and $T\in\mathbb{N}$ be a fixed time horizon. Let $\mathcal{T}=\{0,1,2,\ldots,T\}$, $\mathcal{T}'=\{0,1,2,\ldots,T-1\}$, and $\mathbf{\Theta}\subset\mathbb{R}^d$ be a non-empty set, which will play the role of the known parameter space throughout.[1]

On the space $(\Omega,\mathscr{F})$ we consider a random process $X=\{X_t,\ t\in\mathcal{T}\}$ taking values in some measurable space. We postulate that this process is observed, and we denote by $\mathbb{F}=(\mathscr{F}_t,t\in\mathcal{T})$ its natural filtration. The true law of X is unknown and assumed to be generated by a probability measure belonging to a parameterized family of probability distributions on $(\Omega,\mathscr{F})$, say $\mathbf{P}(\mathbf{\Theta})=\{\mathbb{P}_\theta,\theta\in\mathbf{\Theta}\}$. We will write $\mathbb{E}_{\mathbb{P}}$ to denote the expectation corresponding to a probability measure $\mathbb{P}$ on $(\Omega,\mathscr{F})$, and, for simplicity, we denote by $\mathbb{E}_\theta$ the expectation operator corresponding to the probability $\mathbb{P}_\theta$.

By $\mathbb{P}_{\theta^*}$ we denote the measure generating the true law of X, so that $\theta^*\in\mathbf{\Theta}$ is the (unknown) true parameter. Since $\mathbf{\Theta}$ is assumed to be known, the model uncertainty discussed in this paper occurs only if $\mathbf{\Theta}\neq\{\theta^*\}$, which we assume to be the case.

[1] In general, the parameter space may be infinite dimensional, consisting for example of dynamic factors, such as deterministic functions of time or hidden Markov chains. In this study, for simplicity, we chose the parameter space to be a subset of $\mathbb{R}^d$. In most applications, in order to avoid problems with constrained estimation, the parameter space is taken to be equal to the maximal relevant subset of $\mathbb{R}^d$.

The methodology proposed in this paper is motivated by the following generic optimization problem: we consider a family, say $\mathcal{A}$, of $\mathbb{F}$–adapted processes $\varphi = \{\varphi_t,\ t \in \mathcal{T}\}$ defined on $(\Omega, \mathscr{F})$, that take values in a measurable space. We refer to the elements of $\mathcal{A}$ as to admissible control processes. Additionally, we consider a functional of X and φ, which we denote by L. The stochastic control problem at hand is

$$\inf_{\varphi \in \mathcal{A}} \mathbb{E}_{\theta^*} \left(L(X, \varphi) \right). \tag{2.1}$$

However, stated as such, the problem can not be dealt with directly, since the value of θ^* is unknown. Because of this we refer to such problem as to an *uncertain stochastic control problem.* The question is then how to handle the stochastic control problem (2.1) subject to this type of model uncertainty.

The classical approaches to solving this uncertain stochastic control problem are:

- *to solve the robust control problem*

$$\inf_{\varphi \in \mathcal{A}} \sup_{\theta \in \mathbf{\Theta}} \mathbb{E}_{\theta} \left(L(X, \varphi) \right). \tag{2.2}$$

  We refer to, e.g., [HSTW06], [HS08], [BB95], for more information regarding robust control problems.

- *to solve the strong robust control problem*

$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}_{\nu}^{\varphi, \mathbf{\Theta}^K}} \mathbb{E}_{\mathbb{Q}} \left(L(X, \varphi) \right), \tag{2.3}$$

  where $\mathbf{\Theta}^K$ is the set of strategies chosen by a Knightian adversary (the nature) and $\mathcal{Q}_{\nu}^{\varphi, \mathbf{\Theta}^K}$ a set of probabilities, depending on the strategy φ and a given law ν on $\mathbf{\Theta}$. See Section 3 for a formal description of this problem, and see, e.g., [Sir14] and [BCP16] for related work.

- *to solve the model free robust control problem*

$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}} \left(L(X, \varphi) \right), \tag{2.4}$$

  where $\mathcal{P}$ is given family of probability measures on $(\Omega, \mathscr{F})$.

- *to solve a Bayesian adaptive control problem,* where it is assumed that the (unknown) parameter θ is random, modeled as a random variable Θ taking values in $\mathbf{\Theta}$ and having a prior distribution denoted by ν_0. In this framework, the uncertain control problem is solved via the following optimization problem

$$\inf_{\varphi \in \mathcal{A}} \int_{\mathbf{\Theta}} \mathbb{E}_{\theta} \left(L(X, \varphi) \right) \nu_0(d\theta).$$

  We refer to, e.g., [KV15].

- *to solve an adaptive control problem,* namely, first for each $\theta \in \mathbf{\Theta}$ solve

$$\inf_{\varphi \in \mathcal{A}} \mathbb{E}_{\theta} \left(L(X, \varphi) \right), \tag{2.5}$$

  and denote by φ^{θ} a corresponding optimal control (assumed to exist). Then, at each time $t \in \mathcal{T}'$, compute a point estimate $\widehat{\theta}_t$ of θ^*, using a chosen, $\mathscr{F}_t$ measurable estimator $\widehat{\Theta}_t$. Finally, apply at time t the control value $\varphi_t^{\widehat{\theta}_t}$. We refer to, e.g., [KV15], [CG91].

Several comments are now in order:

1. Regarding the solution of the robust control problem, [LSS06] observe that

   > *If the true model*[a] *is the worst one, then this solution will be nice and dandy. However, if the true model is the best one or something close to it, this solution could be very bad (that is, the solution need not be robust to model error at all!).*

   [a] True model in [LSS06] corresponds to θ^* in our notation.

   The message is that using the robust control framework may produce undesirable results.

2. It can be shown that

$$\inf_{\varphi\in\mathcal{A}}\sup_{\theta\in\mathbf{\Theta}}\mathbb{E}_\theta\left(L(S,\varphi)\right)) = \inf_{\varphi\in\mathcal{A}}\sup_{\nu_0\in\mathcal{P}(\mathbf{\Theta})}\int_{\mathbf{\Theta}}\mathbb{E}_\theta\left(L(S,\varphi)\right)\nu_0(d\theta). \tag{2.6}$$

   Thus, for any given prior distribution ν_0

$$\inf_{\varphi\in\mathcal{A}}\sup_{\theta\in\mathbf{\Theta}}\mathbb{E}_\theta\left(L(S,\varphi)\right) \geq \inf_{\varphi\in\mathcal{A}}\int_{\mathbf{\Theta}}\mathbb{E}_\theta\left(L(S,\varphi)\right)\nu_0(d\theta).$$

   The adaptive Bayesian problem appears to be less conservative. Thus, in principle, solving the adaptive Bayesian control problem for a given prior distribution may lead to a better solution of the uncertain stochastic control problem than solving the robust control problem.

3. It is sometimes suggested that the robust control problem does not involve learning about the unknown parameter θ^*, which in fact is the case, but that the adaptive Bayesian control problem involves "learning" about θ^*. The reason for this latter claim is that in the adaptive Bayesian control approach, in accordance with the Bayesian statistics, the unknown parameter is considered to be a random variable, say Θ with prior distribution ν_0. This random variable is then considered to be an unobserved state variable, and consequently the adaptive Bayesian control problem is regarded as a control problem with partial (incomplete) observation of the entire state vector. The typical way to solve a control problem with partial observation is by means of transforming it to the corresponding separated control problem. The separated problem is a problem with full observation, which is achieved by introducing additional state variable, and what in the Bayesian statistics is known as the posterior distribution of Θ. The "learning" is attributed to the use of the posterior distribution of Θ in the separated problem. However, the information that is used in the separated problem is exactly the same that in the original problem, no learning is really involved. This is further documented by the equality (2.6).

4. We refer to Remark 3.2 for a discussion regarding distinction between the strong robust control problem 2.3 and the adaptive robust control problem 2.7 that is stated below.

5. As said before, the model uncertainty discussed in this paper occurs if $\mathbf{\Theta} \neq \{\theta^*\}$. The classical robust control problem (2.2) does not involve any reduction of uncertainty about θ^*, as the parameter space is not "updated" with incoming information about the signal process X. Analogous remark applies to problems (2.3) and (2.4): there is no reduction of uncertainty about the underlying stochastic dynamics involved there.

   Clearly, incorporating "learning" into the robust control paradigm appears like a good idea. In fact, in [AHS03] the authors state

> *We see three important extensions to our current investigation. Like builders of rational expectations models, we have side-stepped the issue of how decision-makers select an approximating model. Following the literature on robust control, we envision this approximating model to be analytically tractable, yet to be regarded by the decision maker as not providing a correct model of the evolution of the state vector. The misspecifications we have in mind are small in a statistical sense but can otherwise be quite diverse. Just as we have not formally modelled how agents learned the approximating model, neither have we formally justified why they do not bother to learn about potentially complicated misspecifications of that model.* ***Incorporating forms of learning would be an important extension of our work.***[a]
>
> [a] The boldface emphasis is ours.

In the present work we follow up on the suggestion of Anderson, Hansen and Sargent stated above, and we propose a new methodology, which we call **adaptive robust control methodology**, and which is meant to incorporate learning about θ^* into the robust control paradigm.

This methodology amounts to solving the following problem

$$\inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}}} \mathbb{E}_{\mathbb{Q}}\left(L(S,\varphi)\right), \tag{2.7}$$

where $\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}}$ is a family of probability measures on some canonical space related to the process X, chosen in a way that allows for appropriate dynamic reduction of uncertainty about θ^*. Specifically, we chose the family $\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}}$ in terms of confidence regions for the parameter θ^* (see details in Section 3). Thus, the adaptive robust control methodology incorporates updating controller's knowledge about the parameter space – a form of learning, directed towards reducing uncertainty about θ^* using incoming information about the signal process S. Problem (2.7) is derived from problem (2.3); we refer to Section 3 for derivation of both problems.

In this paper we will compare the robust control methodology, the adaptive control methodology and the adaptive robust control methodology in the context of a specific optimal portfolio selection problem that is considered in finance. This will be done in Section 4.

3 Adaptive Robust Control Methodology

This section is the key section of the paper. We will first make precise the control problem that we are studying and then we will proceed to the presentation of our adaptive robust control methodology.

Let $\{(\Omega,\mathscr{F},\mathbb{P}_\theta),\ \theta\in\mathbf{\Theta}\subset\mathbb{R}^d\}$ be a family of probability spaces, and let $T\in\mathbb{N}$ be a fixed maturity time. In addition, we let $A\subset\mathbb{R}^k$ be a finite[2] set and

$$S : \mathbb{R}^n\times A\times\mathbb{R}^m\to\mathbb{R}^n$$

be a measurable mapping. Finally, let

$$\ell : \mathbb{R}^n\to\mathbb{R}$$

be a measurable function.

[2] A will represent the set of control values, and we assume it is finite for simplicity, in order to avoid technical issues regarding existence of measurable selectors.

We consider an underlying discrete time controlled dynamical system with state process X taking values in $\mathbb{R}^n$ and control process φ taking values in A. Specifically, we let

$$X_{t+1} = S(X_t, \varphi_t, Z_{t+1}), \quad t \in \mathcal{T}', \quad X_0 = x_0 \in \mathbb{R}^n, \tag{3.1}$$

where $Z = \{Z_t,\ t \in \mathcal{T}'\}$ is an $\mathbb{R}^m$-valued random sequence, which is $\mathbb{F}$-adapted and i.i.d. under each measure $\mathbb{P}_\theta$.[3] The true, but unknown law of Z corresponds to measure $\mathbb{P}_{\theta^*}$. A control process φ is admissible, if it is $\mathbb{F}$-adapted. We denote by $\mathcal{A}$ the set of all admissible controls.

Using the notation of Section 2 we set $L(X, \varphi) = \ell(X_T)$, so that the problem (2.1) becomes now

$$\inf_{\varphi \in \mathcal{A}} \mathbb{E}_{\theta^*} \ell(X_T). \tag{3.2}$$

3.1 Formulation of the adaptive robust control problem

In what follows, we will be making use of a recursive construction of confidence regions for the unknown parameter θ^* in our model. We refer to [BCC16] for a general study of recursive constructions of confidence regions for time homogeneous Markov chains, and to Section 4 for details of a specific recursive construction corresponding to the optimal portfolio selection problem. Here, we just postulate that the recursive algorithm for building confidence regions uses an $\mathbb{R}^d$-valued and observed process, say $(C_t,\ t \in \mathcal{T}')$, satisfying the following abstract dynamics

$$C_{t+1} = R(t, C_t, Z_{t+1}), \quad t \in \mathcal{T}',\ C_0 = c_0, \tag{3.3}$$

where $R(t, c, z)$ is a deterministic measurable function. Note that, given our assumptions about process Z, the process C is $\mathbb{F}$-adapted. This is one of the key features of our model. Usually C_t is taken to be a consistent estimator of θ^*.

Now, we fix a confidence level $\alpha \in (0, 1)$, and for each time $t \in \mathcal{T}'$, we assume that an (1-α)-confidence region, say $\mathbf{\Theta}_t$, for θ^*, can be represented as

$$\mathbf{\Theta}_t = \tau(t, C_t), \tag{3.4}$$

where, for each $t \in \mathcal{T}'$,

$$\tau(t, \cdot) : \mathbb{R}^d \to 2^{\mathbf{\Theta}}$$

is a deterministic set valued function.[4] Note that in view of (3.3) the construction of confidence regions given in (3.4) is indeed recursive. In our construction of confidence regions, the mapping $\tau(t, \cdot)$ will be a measurable set valued function, with compact values. The important property of the recursive confidence regions constructed in Section 4 is that $\lim_{t\to\infty} \mathbf{\Theta}_t = \{\theta^*\}$, where the convergence is understood $\mathbb{P}_{\theta^*}$ almost surely, and the limit is in the Hausdorff metric. This is not always the case though in general. In [BCC16] is shown that the convergence holds in probability, for the model set-up studied there. The sequence $\mathbf{\Theta}_t,\ t \in \mathcal{T}'$ represents learning about θ^* based on the observation of the history $H_t,\ t \in \mathcal{T}$ (cf. (3.6) below). We introduce the augmented state process $Y_t = (X_t, C_t),\ t \in \mathcal{T}$, and the augmented state space

$$E_Y = \mathbb{R}^n \times \mathbb{R}^d.$$

We denote by $\mathcal{E}_Y$ the collection of Borel measurable sets in E_Y. The process Y has the following dynamics,

$$Y_{t+1} = \mathbf{G}(t, Y_t, \varphi_t, Z_{t+1}),\ t \in \mathcal{T}',$$

[3] The assumption that the sequence Z is i.i.d. under each measure $\mathbb{P}_\theta$ is made in order to simplify our study.

[4] As usual, $2^{\mathbf{\Theta}}$ denotes the set of all subsets of $\mathbf{\Theta}$.

where $\mathbf{G}$ is the mapping

$$\mathbf{G} : \mathcal{T}' \times E_Y \times A \times \mathbb{R}^m \to E_Y$$

defined as

$$\mathbf{G}(t, y, a, z) = \big(S(x, a, z), R(t, c, z)\big), \tag{3.5}$$

where $y = (x, c) \in E_Y$.

For future reference, we define the corresponding histories

$$H_t = ((X_0, C_0), (X_1, C_1), \ldots, (X_t, C_t)),\ t \in \mathcal{T}, \tag{3.6}$$

so that

$$H_t \in \mathbf{H}_t = \underbrace{E_Y \times E_Y \times \ldots \times E_Y}_{t+1 \text{ times}}. \tag{3.7}$$

Clearly, for any admissible control process φ, the random variable H_t is $\mathscr{F}_t$-measurable. We denote by

$$h_t = (y_0, y_1, \ldots, y_t) = (x_0, c_0, x_1, c_1, \ldots, x_t, c_t) \tag{3.8}$$

a realization of H_t. Note that $h_0 = y_0$.

Remark 3.1. A control process $\varphi = (\varphi_t,\ t \in \mathcal{T}')$ is called history dependent control process if (with a slight abuse of notation)

$$\varphi_t = \varphi_t(H_t),$$

where (on the right hand side) $\varphi_t : \mathbf{H}_t \to A$, is a measurable mapping. Note that any admissible control process φ is such that φ_t is a function of $X_0, \ldots, X_t$. So, any admissible control process is history dependent. On the other hand, given our above set up, any history dependent control process is $\mathbb{F}$–adapted, and thus, it is admissible. From now on, we identify the set $\mathcal{A}$ of admissible strategies with the set of history dependent strategies.

For the future reference, for any admissible control process φ and for any $t \in \mathcal{T}'$, we denote by $\varphi^t = (\varphi_k,\ k = t, \ldots, T-1)$ the "t-tail" of φ; in particular, $\varphi^0 = \varphi$. Accordingly, we denote by $\mathcal{A}^t$ the collection of t-tails φ^t; in particular, $\mathcal{A}^0 = \mathcal{A}$.

Let $\psi_t : \mathbf{H}_t \to \mathbf{\Theta}$ be a Borel measurable mapping (Knightian selector), and let us denote by $\psi = (\psi_t,\ t \in \mathcal{T}')$ the sequence of such mappings, and by $\psi^t = (\psi_s,\ s = t, \ldots, T-1)$ the t-tail of the sequence ψ. The set of all sequences ψ, and respectively ψ^t, will be denoted by $\mathbf{\Psi}_K$ and $\mathbf{\Psi}_K^t$, respectively. Similarly, we consider the measurable selectors $\psi_t(\cdot) : \mathbf{H}_t \to \mathbf{\Theta}_t$, and correspondingly define the set of all sequences of such selectors by $\mathbf{\Psi}$, and the set of t-tails by $\mathbf{\Psi}^t$. Clearly, $\psi^t \in \mathbf{\Psi}^t$ if and only if $\psi^t \in \mathbf{\Psi}_K^t$ and $\psi_s(h_s) \in \tau(s, c_s),\ s = t, \ldots, T-1$.

Next, for each $(t, y, a, \theta) \in \mathcal{T}' \times E_Y \times A \times \mathbf{\Theta}$, we define a probability measure on $\mathcal{E}_Y$:

$$Q(B \mid t, y, a, \theta) = \mathbb{P}_\theta(Z_{t+1} \in \{z : \mathbf{G}(t, y, a, z) \in B\}) = \mathbb{P}_\theta\left(\mathbf{G}(t, y, a, Z_{t+1}) \in B\right),\ B \in \mathcal{E}_Y. \tag{3.9}$$

We assume that for each B the function $Q(B \mid t, y, a, \theta)$ of t, y, a, θ is measurable. This assumption will be satisfied in the context of the optimal portfolio problem discussed in Section 4.

Finally, using Ionescu-Tulcea theorem, for every control process $\varphi \in \mathcal{A}$ and for every initial probability distribution ν on E_Y, we define the family $\mathcal{Q}_\nu^{\varphi, \mathbf{\Psi}} = \{\mathbb{Q}_\nu^{\varphi,\psi},\ \boldsymbol{\psi} \in \mathbf{\Psi}\}$ of probability measures on the canonical space E_Y^{T+1}, with $\mathbb{Q}_\nu^{\varphi,\psi}$ given as follows

$$\mathbb{Q}_\nu^{\varphi,\psi}(B_0, B_1, \ldots, B_T) = \int_{B_0} \int_{B_1} \cdots \int_{B_T} \prod_{t=1}^{T} Q(dy_t | t-1, y_{t-1}, \varphi_{t-1}(h_{t-1}), \psi_{t-1}(h_{t-1})) \nu(dy_0) \tag{3.10}$$

Analogously we define the set $\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}_K} = \{\mathbb{Q}_\nu^{\varphi,\psi^K},\ \psi^K \in \mathbf{\Psi}_K\}$.

The *strong robust control* problem is then given as:

$$\inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}_K}} \mathbb{E}_{\mathbb{Q}}\ell(X_T). \tag{3.11}$$

The corresponding *adaptive robust control* problem is:

$$\inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}}} \mathbb{E}_{\mathbb{Q}}\ell(X_T). \tag{3.12}$$

Remark 3.2. The strong robust control problem is essentially a game problem between the hedger and his/her Knightian adversary – the nature, who may keep changing the dynamics of the underlying stochastic system over time. In this game, the nature is not restricted in its choices of model dynamics, except for the requirement that $\psi_t^K(H_t) \in \mathbf{\Theta}$, and each choice is potentially based on the entire history H_t up to time t. On the other hand, the adaptive robust control problem is a game problem between the hedger and his/her Knightian adversary – the nature, who, as in the case of strong robust control problem, may keep changing the dynamics of the underlying stochastic system over time. However, in this game, the nature is restricted in its choices of model dynamics to the effect that $\psi_t(H_t) \in \tau(t, C_t)$.

Note that if the parameter θ^* is known, then, using the above notations and the canonical construction[5], the control problem reduces to

$$\inf_{\varphi\in\mathcal{A}} \mathbb{E}_{\mathbb{Q}^*}\ell(X_T), \tag{3.13}$$

where, formally, the probability $\mathbb{Q}^*$ is given as in (3.10) with $\tau(t,c) = \{\theta^*\}$ for all t and c. It certainly holds that

$$\inf_{\varphi\in\mathcal{A}} \mathbb{E}_{\mathbb{Q}^*}\ell(X_T) \le \inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}}} \mathbb{E}_{\mathbb{Q}}\ell(X_T) \le \inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}_K}} \mathbb{E}_{\mathbb{Q}}\ell(X_T). \tag{3.14}$$

It also holds that

$$\inf_{\varphi\in\mathcal{A}} \sup_{\theta\in\mathbf{\Theta}} \mathbb{E}_\theta\ell(X_T) \le \inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}_K}} \mathbb{E}_{\mathbb{Q}}\ell(X_T). \tag{3.15}$$

Remark 3.3. We conjecture that

$$\inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}}} \mathbb{E}^{\mathbb{Q}}\ell(X_T) \le \inf_{\varphi\in\mathcal{A}} \sup_{\theta\in\mathbf{\Theta}} \mathbb{E}_\theta\ell(X_T).$$

However, at this time, we do not know how to prove this conjecture, and whether it is true in general.

3.2 Solution of the adaptive robust control problem

In accordance with our original set-up, in what follows we assume that $\nu(dx_0) = \delta_{h_0}(dx_0)$ (Dirac measure), and we use the notation $\mathbb{Q}_{h_0}^{\varphi,\psi}$ and $\mathcal{Q}_{h_0}^{\varphi,\mathbf{\Psi}}$ in place of $\mathbb{Q}_\nu^{\varphi,\psi}$ and $\mathcal{Q}_\nu^{\varphi,\mathbf{\Psi}}$, so that the problem (3.12) becomes

$$\inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_{h_0}^{\varphi,\mathbf{\Psi}}} \mathbb{E}_{\mathbb{Q}}\ell(X_T). \tag{3.16}$$

[5] Which, clearly, is not needed in this case.

For each $t \in \mathcal{T}'$, we then define a probability measure on the concatenated canonical space as follows

$$\mathbb{Q}_{h_t}^{\varphi^t,\psi^t}(B_{t+1},\dots,B_T) = \int_{B_{t+1}} \cdots \int_{B_T} \prod_{u=t+1}^{T} Q(dy_u \mid u-1, y_{u-1}, \varphi_{u-1}(h_{u-1}), \psi_{u-1}(h_{u-1})).$$

Accordingly, we put $\mathcal{Q}_{h_t}^{\varphi^t,\mathbf{\Psi}^t} = \{\mathbb{Q}_{h_t}^{\varphi^t,\psi^t},\ \psi^t \in \mathbf{\Psi}^t\}$. Finally, we define the functions U_t and U_t^* as follows: for $\varphi^t \in \mathcal{A}^t$ and $h_t \in H_t$

$$U_t(\varphi^t, h_t) = \sup_{\mathbb{Q}\in\mathcal{Q}_{h_t}^{\varphi^t,\mathbf{\Psi}^t}} \mathbb{E}_{\mathbb{Q}}\ell(X_T),\ t \in \mathcal{T}', \tag{3.17}$$

$$U_t^*(h_t) = \inf_{\varphi^t\in\mathcal{A}^t} U_t(\varphi^t, h_t),\ t \in \mathcal{T}', \tag{3.18}$$

$$U_T^*(h_T) = \ell(x_T). \tag{3.19}$$

Note in particular that

$$U_0^*(y_0) = U_0^*(h_0) = \inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_{h_0}^{\varphi,\mathbf{\Psi}}} \mathbb{E}_{\mathbb{Q}}\ell(X_T).$$

We call U_t^* the *adaptive robust Bellman* functions.

3.2.1 Adaptive robust Bellman equation

Here we will show that a solution to the optimal problem (3.16) can be given in terms of the adaptive robust Bellman equation associated to it.

Towards this end we will need to solve for functions $W_t,\ t \in \mathcal{T}$, the following adaptive robust Bellman equations (recall that $y = (x, c)$)

$$\begin{aligned} W_T(y) &= \ell(x), \quad y \in E_Y, \\ W_t(y) &= \inf_{a\in A} \sup_{\theta\in\tau(t,c)} \int_{E_Y} W_{t+1}(y')Q(dy' \mid t, y, a, \theta), \quad y \in E_Y,\ t = T-1,\dots,0, \end{aligned} \tag{3.20}$$

and to compute the related optimal selectors $\varphi_t^*,\ t \in \mathcal{T}'$.

In Lemma 3.4 below, under some additional technical assumptions, we will show that the optimal selectors in (3.20) exist; namely, for any $t \in \mathcal{T}'$, and any $y = (x, c) \in E_Y$, there exists a measurable mapping $\varphi_t^* : E_Y \to A$, such that

$$W_t(y) = \sup_{\theta\in\tau(t,c)} \int_{E_Y} W_{t+1}(y')Q(dy' \mid t, y, \varphi_t^*(y), \theta).$$

In order to proceed, to simplify the argument, we will assume that under measure $\mathbb{P}_\theta$, for each $t \in \mathcal{T}$, the random variable Z_t has a density with respect to the Lebesgue measure, say $f_Z(z;\theta),\ z \in \mathbb{R}^m$. We will also assume that the set A of available actions is finite. In this case, the problem (3.20) becomes

$$\begin{aligned} W_T(y) &= \ell(x), \quad y \in E_Y, \\ W_t(y) &= \min_{a\in A} \sup_{\theta\in\tau(t,c)} \int_{\mathbb{R}^m} W_{t+1}(\mathbf{G}(t, y, a, z))f_Z(z;\theta)dz, \quad y \in E_Y,\ t = T-1,\dots,0. \end{aligned} \tag{3.21}$$

where $T(t, y, a, z)$ is given in (3.5). Additionally, we take the standing assumptions that

(i) for any a and z, the function $S(\cdot, a, z)$ is continuous.

(ii) For each z, the function $f_Z(z;\cdot)$ is continuous in θ.

(iii) ℓ is continuous and bounded.[6]

(iv) For each $t \in \mathcal{T}'$, the function $R(t,\cdot,\cdot)$ is continuous.

Then, we have the following result.

Lemma 3.4. *The functions W_t, $t = T, T-1, \dots, 0$, are upper semi-continuous (u.s.c.), and the optimal selectors φ_t^*, $t = T, T-1, \dots, 0$, in (3.20) exist.*

Proof. The function W_T is continuous. Since $\mathbf{G}(T-1,\cdot,\cdot,z)$ is continuous, then, $W_T(\mathbf{G}(T-1,\cdot,\cdot,z))$ is continuous. Consequently, the function

$$w_{T-1}(y,a,\theta) = \int_{\mathbb{R}} W_T(\mathbf{G}(T-1,y,a,z)) f_Z(z;\theta) dz$$

is continuous, and thus u.s.c.

Next we will apply [BS78, Proposition 7.33] by taking (in the notations of [BS78])

$$\begin{aligned} \mathrm{X} &= E_Y \times A = \mathbb{R}^n \times \mathbb{R}^d \times A, \quad \mathrm{x} = (y,a), \\ \mathrm{Y} &= \mathbf{\Theta}, \quad \mathrm{y} = \theta, \\ \mathrm{D} &= \bigcup_{(y,a)\in E_Y \times A} \{(y,a)\} \times \tau(T-1,c), \\ f(\mathrm{x},\mathrm{y}) &= -w_{T-1}(y,a,\theta). \end{aligned}$$

Recall that in view of the prior assumptions, Y is compact. Clearly X is metrizable. From the above, f is lower semi-continuous (l.s.c). Also note that the cross section $\mathrm{D}_{\mathrm{x}} = \mathrm{D}_{(y,a)} = \{\theta \in \mathbf{\Theta} : (y,a,\theta) \in \mathrm{D}\}$ is given by $\mathrm{D}_{(y,a)}(t) = \tau(t,c)$. Hence, by [BS78, Proposition 7.33], the function

$$\widetilde{w}_{T-1}(y,a) = \inf_{\theta \in \tau(T-1,c)} (-w_{T-1}(y,a,\theta)), \quad (y,a) \in E_Y \times A$$

is l.s.c.. Consequently, the function $W_{T-1} = \inf_{a\in A}(-\widetilde{w}_{T-1}(y,a))$ is u.s.c., and there exists an optimal selector φ_{T-1}^*. The rest of the proof follows in the analogous way. □

The following proposition is the key result in this section.

Proposition 3.5. *For any $h_t \in H_t$, and $t \in \mathcal{T}$, we have*

$$U_t^*(h_t) = W_t(y_t). \tag{3.22}$$

Moreover, the policy φ^* constructed from the selectors in (3.20) is robust-optimal, that is

$$U_t^*(h_t) = U_t(\varphi_t^*, h_t), \quad t \in \mathcal{T}'. \tag{3.23}$$

[6] We assume that ℓ is bounded to guarantee the finiteness of the expectations in (3.16), as well as finiteness of functions W_t. Clearly this assumption can be relaxed, as it will be seen in the examples from Section 4.

Proof. We proceed similarly as in the proof of [Iye05, Theorem 2.1], and via backward induction in $t = T, T-1, \ldots, 1, 0$.

Take $t = T$. Clearly, $U_T^*(h_T) = W_T(y_T)$. For $t = T-1$ we have

$$\begin{aligned}
U_{T-1}^*(h_{T-1}) &= \inf_{\varphi^{T-1}=\varphi_{T-1}\in\mathcal{A}^{T-1}} \sup_{\mathbb{Q}\in\mathcal{Q}_{h_{T-1}}^{\varphi^{T-1},\mathbf{\Psi}^{T-1}}} \mathbb{E}^{\mathbb{Q}}\ell(X_T) \\
&= \inf_{\varphi^{T-1}=\varphi_{T-1}\in\mathcal{A}^{T-1}} \sup_{\theta\in\tau(T-1,c_{T-1})} \int_{E_Y} U_T^*(h_{T-1},y) Q(dy \mid T-1, y_{T-1}, \varphi_{T-1}(h_{T-1}), \theta) \\
&= \inf_{\varphi^{T-1}=\varphi_{T-1}\in\mathcal{A}^{T-1}} \sup_{\theta\in\tau(T-1,c_{T-1})} \int_{E_Y} W_T(y)\, Q(dy \mid T-1, y_{T-1}, \varphi_{T-1}(h_{T-1}), \theta) \\
&= \inf_{a\in A} \sup_{\theta\in\tau(T-1,c_{T-1})} \int_{E_Y} W_T(y)\, Q(dy \mid T-1, y_{T-1}, a, \theta) = W_{T-1}(y_{T-1}).
\end{aligned}$$

For $t = T-1, \ldots, 1, 0$ we have by induction

$$\begin{aligned}
U_t^*(h_t) &= \inf_{\varphi^t\in\mathcal{A}^t} \sup_{\mathbb{Q}\in\mathcal{Q}_{h_t}^{\varphi^t,\mathbf{\Psi}^t}} \mathbb{E}^{\mathbb{Q}}\ell(X_T) \\
&= \inf_{\varphi^t=(\varphi_t,\varphi^{t+1})\in\mathcal{A}^t} \sup_{\theta\in\tau(t,c_t)} \int_{E_Y} \sup_{\widehat{\mathbb{Q}}\in\mathcal{Q}_{h_t,y}^{\varphi^{t+1},\mathbf{\Psi}^{t+1}}} \mathbb{E}^{\widehat{\mathbb{Q}}}\ell(X_T) Q(dy \mid t, y_t, \varphi_t(h_t), \theta) \\
&\geq \inf_{\varphi^t=(\varphi_t,\varphi^{t+1})\in\mathcal{A}^t} \sup_{\theta\in\tau(c_t,t)} \int_{E_Y} U_{t+1}^*(h_t,y)\, Q(dy \mid t, y_t, \varphi_t(h_t), \theta) \\
&= \inf_{a\in A} \sup_{\theta\in\tau(t,c_t)} \int_{E_Y} U_{t+1}^*(h_t,y)\, Q(dy \mid t, y_t, a, \theta) \\
&= \inf_{a\in A} \sup_{\theta\in\tau(t,c_t)} \int_{E_Y} W_{t+1}(y)\, Q(dy \mid t, y_t, a, \theta) = W_t(y_t).
\end{aligned}$$

Now, fix $\epsilon > 0$, and let $\varphi^{t+1,\epsilon}$ denote an ϵ-optimal control process starting at time $t+1$, so that

$$U_{t+1}(\varphi^{t+1,\epsilon}, h_{t+1}) \leq U_{t+1}^*(h_{t+1}) + \epsilon.$$

Then we have

$$\begin{aligned}
U_t^*(h_t) &= \inf_{\varphi^t\in\mathcal{A}^t} \sup_{\mathbb{Q}\in\mathcal{Q}_{h_t}^{\varphi^t,\mathbf{\Psi}^t}} \mathbb{E}^{\mathbb{Q}}\ell(X_T) \\
&= \inf_{\varphi^t=(\varphi_t,\varphi^{t+1})\in\mathcal{A}^t} \sup_{\theta\in\tau(t,c_t)} \int_{E_Y} \sup_{\widehat{\mathbb{Q}}\in\mathcal{Q}_{h_t,y}^{\varphi^{t+1},\mathbf{\Psi}^{t+1}}} \mathbb{E}^{\widehat{\mathbb{Q}}}\ell(X_T) Q(dy \mid t, y_t, \varphi_t(h_t), \theta) \\
&\leq \inf_{\varphi^t=(\varphi_t,\varphi^{t+1})\in\mathcal{A}^t} \sup_{\theta\in\tau(t,c_t)} \int_{E_Y} \sup_{\widehat{\mathbb{Q}}\in\mathcal{Q}_{h_t,y}^{\varphi^{t+1,\epsilon},\mathbf{\Psi}^{t+1}}} \mathbb{E}^{\widehat{\mathbb{Q}}}\ell(X_T) Q(dy \mid t, y_t, \varphi_t(h_t), \theta) \\
&\leq \inf_{a\in A} \sup_{\theta\in\tau(t,c_t)} \int_{E_Y} U_{t+1}^*(h_t,y)\, Q(dy \mid t, y_t, a; \theta) + \epsilon \\
&= \inf_{a\in A} \sup_{\theta\in\tau(t,c_t)} \int_{E_Y} W_{t+1}(y)\, Q(dy \mid t, y_t, a; \theta) + \epsilon = W_t(y_t) + \epsilon.
\end{aligned}$$

Since ϵ was arbitrary, the proof of (3.22) is done. Equality (3.23) now follows easily. □

4 Example: Dynamic Optimal Portfolio Selection

In this section we will present an example that illustrates the adaptive robust control methodology.

We follow here the set up of [BGSCS05] in the formulation of our dynamic optimal portfolio selection. As such, we consider the classical dynamic optimal asset allocation problem, or dynamic optimal portfolio selection, when an investor is deciding at time t on investing in a risky asset and a risk-free banking account by maximizing the expected utility $u(V_T)$ of the terminal wealth, with u being a given utility function. The underlying market model is subject to the type of uncertainty that has been introduced in Section 2.

Consider a risk-free asset B with a constant interest rate $r = (B_{t+1} - B_t)/B_t$, and a risky asset with the corresponding return from time t to $t+1$ denoted by r^S_{t+1}. We assume that the return process r^S, is observed. The dynamics of the wealth process produced by a self-financing trading strategy is given by

$$V_{t+1} = V_t(1 + r + \varphi_t(r^S_{t+1} - r)), \quad t \in \mathcal{T}', \tag{4.1}$$

with the initial wealth $V_0 = v_0$, and where φ_t denotes the proportion of the portfolio wealth invested in the risky asset from time t to $t+1$. We assume that the process φ takes finitely many values, say $a_i,\ i = 1, \dots, N$ where $a_i \in [0, 1]$.

We further assume that $r^S_t + 1$, $t = 1, \dots, T$, is an i.i.d. sequence of log-normal distributed random variables.[7] Namely,

$$r^S_t = e^{Z_t} - 1,$$

where Z_t is an i.i.d. sequence of Gaussian random variables with mean μ and variance σ^2. Equivalently, we put $Z_t = \mu + \sigma\varepsilon_t$, where $\varepsilon_t,\ t \in \mathcal{T}'$ are i.i.d. standard Gaussian random variables. Note that under the above model assumptions, the wealth process remains positive a.s..

Using the notations from Section 3, here we have that $X_t = V_t$, and setting $x = v$ we get

$$S(v, a, z) = v(1 + r + a(e^z - 1 - r)), \quad \ell(v) = -u(v), \quad A = \{a_i,\ i = 1, \dots, N\}.$$

The model uncertainty comes from the unknown parameters μ and/or σ. We will discuss two cases: Case 1 - unknown mean μ and known standard deviation σ, and Case II - both μ and σ are unknown.

Case I. Assume that σ is known, and the model uncertainty comes only from the unknown parameter μ^*. Thus, using the notations from Section 3, we have that $\theta^* = \mu^*$, $\theta = \mu$, and we take $C_t = \widehat{\mu}_t$, $\mathbf{\Theta} = [\underline{\mu}, \overline{\mu}] \subset \mathbb{R}$, where $\widehat{\mu}$ is an estimator of μ, given the observations Z, that takes values in $\mathbf{\Theta}$. For the detailed discussion on the construction of such estimators we refer to [BCC16]. For this example, it is enough to take $\widehat{\mu}$ the Maximum Likelihood Estimator (MLE), which is the sample mean in this case, projected appropriately on $\mathbf{\Theta}$. Formally, the recursion construction of $\widehat{\mu}$ is defined as follows:

$$\begin{aligned} \widetilde{\mu}_{t+1} &= \frac{t}{t+1}\widehat{\mu}_t + \frac{1}{t+1}Z_{t+1}, \\ \widehat{\mu}_{t+1} &= \pi(\widetilde{\mu}_{t+1}), \qquad t \in \mathcal{T}', \end{aligned} \tag{4.2}$$

with $\widehat{\mu}_0 = c_0$, and where π is the projection to the closest point in $\mathbf{\Theta}$, i.e. $\pi(\mu) = \mu$ if $\mu \in [\underline{\mu}, \overline{\mu}]$, $\pi(\mu) = \underline{\mu}$ if $\mu < \underline{\mu}$, and $\pi(\mu) = \overline{\mu}$ if $\mu > \overline{\mu}$. We take as the initial guess c_0 any point in $\mathbf{\Theta}$. It is immediate to verify that

$$R(t, c, z) = \pi\left(\frac{t}{t+1}c + \frac{1}{t+1}z\right)$$

[7] This is the same as to assume that the excess log-returns are normally distributed.

is continuous in c and z. Putting the above together we get that the function $\mathbf{G}$ corresponding to (3.5) is given by

$$\mathbf{G}(t,v,c,a,z) = \left(v(1+r+a(e^z-1-r)), \frac{t}{t+1}c + \frac{1}{t+1}z\right).$$

Now, we note that the $(1-\alpha)$-confidence region for μ^* at time t is given as[8]

$$\mathbf{\Theta}_t = \tau(t, \widehat{\mu}_t),$$

where

$$\tau(t,c) = \left[c - \frac{\sigma}{\sqrt{t}}q_{\alpha/2}, c + \frac{\sigma}{\sqrt{t}}q_{\alpha/2}\right],$$

and where q_α denotes the α-quantile of a standard normal distribution. With these at hand we define the kernel Q according to (3.9), and the set of probability measures $\mathcal{Q}_{h_0}^{\varphi,\mathbf{\Psi}}$ on canonical space by (3.10).

Formally, the investor's problem[9], in our notations and setup, is formulated as follows

$$\inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_{h_0}^{\varphi,\mathbf{\Psi}}} \mathbb{E}_{\mathbb{Q}}[-u(V_T)] \tag{4.3}$$

where $\mathcal{A}$ is the set of self-financing trading strategies.

The corresponding adaptive robust Bellman equation becomes

$$\begin{cases} W_T(v,c) = -u(v), \\ W_t(v,c) = \inf_{a\in A} \sup_{\mu\in\tau_\alpha(t,c)} \mathbb{E}\left[W_{t+1}\left(\mathbf{G}(t,v,c,a,\mu+\sigma\varepsilon_{t+1})\right)\right], \end{cases} \tag{4.4}$$

where the expectation $\mathbb{E}$ is with respect to a standard one dimensional Gaussian distribution. In view of Proposition 3.5, the function $W_t(v,c)$ satisfies

$$W_t(v,c) = \inf_{\varphi^t\in\mathcal{A}^t} \sup_{\mathbb{Q}\in\mathcal{Q}_{h_t'}^{\varphi^t,\mathbf{\Psi}^t}} \mathbb{E}_{\mathbb{Q}}[-u(V_T)],$$

where $h_t' = (h_{t-1}, v, c)$ for any $h_{t-1} = (v_0, c_0, \ldots, v_{t-1}, c_{t-1}) \in \mathbf{H}_{t-1}$. So, in particular,

$$W_0(v_0,c_0) = \inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}_{h_0}^{\varphi,\mathbf{\Psi}}} \mathbb{E}_{\mathbb{Q}}[-u(V_T)],$$

and the optimal selector $\varphi_t^*, t\in\mathcal{T}$, from (4.4) solves the original investor's allocation problem.

To further reduce the computational complexity, we consider a CRRA utility of the form $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$, for $x>0$, and some $\gamma\neq 1$. Note that the function $\ell = -u$ is not bounded and thus it does not satisfy condition (iii). However, as already mentioned, condition (iii) is just a sufficient conditions, and one can check directly that in this particular example of CRRA utility, and under above assumptions on the dynamics of X, all results from Section 3.2.1 hold true.

In this case we have

$$u(V_T) = u(V_t)\left[\prod_{s=t}^{T-1}(1+r+\varphi_s(e^{Z_{s+1}}-1-r))\right]^{1-\gamma}, \quad t\in\mathcal{T}'.$$

[8] We take this interval to be closed, as we want it to be compact.

[9] The original investor's problem is $\sup_{\varphi\in\mathcal{A}}\inf_{\mathbb{Q}\in\mathcal{Q}_{h_0}^{\varphi,\mathbf{\Psi}}}\mathbb{E}_{\mathbb{Q}}[u(V_T)]$, where $\mathcal{A}$ is the set of self-financing trading strategies

Note that

$$W_t(v,c) = -u(v)\cdot \begin{cases} \sup_{\varphi^t\in\mathcal{A}^t} \inf_{\mathbb{Q}\in\mathcal{Q}^{\varphi^t,\Psi^t}_{h'_t}} \mathbb{E}_{\mathbb{Q}}\left[\left(\prod_{s=t}^{T-1}(1+r+\varphi_s(e^{Z_{s+1}}-1-r))\right)^{1-\gamma}\right], & \gamma<1,\\ \inf_{\varphi^t\in\mathcal{A}^t} \sup_{\mathbb{Q}\in\mathcal{Q}^{\varphi^t,\Psi^t}_{h'_t}} \mathbb{E}_{\mathbb{Q}}\left[\left(\prod_{s=t}^{T-1}(1+r+\varphi_s(e^{Z_{s+1}}-1-r))\right)^{1-\gamma}\right], & \gamma>1. \end{cases}$$

Next, following the ideas presented in [BGSCS05], we will prove that for $t\in\mathcal{T}$ and any $c\in[a,b]$ the ratio $W_t(v,c)/v^{1-\gamma}$ does not depend on v, and that the functions $\widetilde{W}_t$ defined as $\widetilde{W}_t(c) = W_t(v,c)/v^{1-\gamma}$ satisfy the following backward recursion

$$\begin{cases} \widetilde{W}_T(c) = \frac{1}{1-\gamma},\\ \widetilde{W}_t(c) = \inf_{a\in A}\sup_{\mu\in\tau_\alpha(t,c)} \mathbb{E}\Big[(1+r+a(e^{\mu+\sigma\varepsilon_{t+1}}-1-r))^{1-\gamma} \\ \qquad\qquad \widetilde{W}_{t+1}(\frac{t}{t+1}c+\frac{1}{t+1}(\mu+\sigma\varepsilon_{t+1}))\Big], \quad t\in\mathcal{T}'. \end{cases} \tag{4.5}$$

We will show this by backward induction in t. First, the equality $\widetilde{W}_T(c) = \frac{1}{1-\gamma}$ is obvious. Next, we fix $t\in\mathcal{T}'$, and we assume that $W_{t+1}(v,c)/v^{1-\gamma}$ does not depend on v. Thus, using (4.4) we obtain

$$\frac{W_t(v,c)}{v^{1-\gamma}} = \inf_{a\in A}\sup_{\mu\in\tau_\alpha(t,c)} \mathbb{E}\left[(1+r+a(e^{\mu+\sigma\varepsilon_{t+1}}-1-r))^{1-\gamma}\widetilde{W}_{t+1}\left(\pi(t,v,c,a,\mu+\sigma\varepsilon_{t+1})\right)\right]$$

does not depend on v since $\widetilde{W}_{t+1}$ does not depend on its first argument.

We finish this example with several remarks regarding the numerical implementations aspects of this problem. By considering CRRA utility functions, and with the help of the above substitution, we reduced the original recursion problem (4.4) to (4.5) which has a lower dimension, and which consequently significantly reduces the computational complexity of the problem.

In the next section we compare the strategies (and the corresponding wealth process) obtained by the adaptive robust method, strong robust methods, adaptive control method, as well as by considering the case of no model uncertainty when the true model is known.

Assuming that the true model is known, the trading strategies are computed by simply solving the optimization problem (3.2), with its corresponding Bellman equation

$$\begin{cases} \widetilde{W}_T = \frac{1}{1-\gamma},\\ \widetilde{W}_t = \inf_{a\in A}\mathbb{E}\left[(1+r+a(e^{\mu^*+\sigma\varepsilon_{t+1}}-1-r))^{1-\gamma}\widetilde{W}_{t+1}\right],\ t\in\mathcal{T}'. \end{cases} \tag{4.6}$$

Similar to the derivation of (4.5), one can show that the Bellman equation for the robust control problem takes the form

$$\begin{cases} \widetilde{W}_T = \frac{1}{1-\gamma},\\ \widetilde{W}_t = \inf_{a\in A}\sup_{\mu\in\mathbf{\Theta}}\mathbb{E}\left[(1+r+a(e^{\mu+\sigma\varepsilon_{t+1}}-1-r))^{1-\gamma}\widetilde{W}_{t+1}\right],\ t\in\mathcal{T}'. \end{cases} \tag{4.7}$$

Note that the Bellman equations (4.6) and (4.7) are recursive scalar sequences that can be computed numerically efficiently with no state space discretization required. The adaptive control strategies are obtained by solving, at each time iteration t, a Bellman equation similar to (4.6), but by iterating backward up to time t and where μ^* is replaced by its estimated value $\widehat{\mu}_t$.

To solve the corresponding adaptive control problem, we first perform the optimization phase. Namely, we solve the Bellman equations (4.6) with μ^* replaced by μ, and for all $\mu\in\mathbf{\Theta}$. The

optimal selector is denoted by φ_t^{μ}, $t \in \mathcal{T}'$. Next, we do the adaptation phase. For every $t \in \{0,1,2,\dots,T-1\}$, we compute the pointwise estimate $\widehat{\mu}_t$ of θ^*, and apply the certainty equivalent control $\varphi_t = \varphi_t^{\widehat{\mu}_t}$. For more details see, for instance, [KV15], [CG91].

Case II. Assume that both μ and σ are the unknown parameters, and thus in the notations of Section 3, we have $\theta^* = (\mu^*, (\sigma^*)^2)$, $\theta = (\mu, \sigma^2)$, $\mathbf{\Theta} = [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}^2, \overline{\sigma}^2] \subset \mathbb{R} \times \mathbb{R}_+$, for some fixed $\underline{\mu}, \overline{\mu} \in \mathbb{R}$ and $\underline{\sigma}^2, \overline{\sigma}^2 \in \mathbb{R}_+$. Similar to the Case I, we take the MLEs for μ^* and $(\sigma^*)^2$, namely the sample mean and respectively the sample variance, projected appropriately to the rectangle $\mathbf{\Theta}$. It is shown in [BCC16] that the following recursions hold true

$$\begin{aligned}
\widetilde{\mu}_{t+1} &= \frac{t}{t+1}\widehat{\mu}_t + \frac{1}{t+1}Z_{t+1},\\
\widetilde{\sigma}^2_{t+1} &= \frac{t}{t+1}\widehat{\sigma}^2_t + \frac{t}{(t+1)^2}(\widehat{\mu}_t - Z_{t+1})^2,\\
(\widehat{\mu}_{t+1}, \widehat{\sigma}^2_{t+1}) &= \pi(\widetilde{\mu}_{t+1}, \widetilde{\sigma}^2_{t+1}), \quad t = 1, \dots, T-1,
\end{aligned}$$

with some initial guess $\widehat{\mu}_0 = c'_0$, and $\widehat{\sigma}^2_0 = c''_0$, and where π is the projection[10] defined similarly as in (4.2). Consequently, we put $C_t = (C'_t, C''_t) = (\widehat{\mu}_t, \widehat{\sigma}^2_t), t \in \mathcal{T}$, and respectively we have

$$R(t,c,z) = \pi\left(\frac{t}{t+1}c' + \frac{1}{t+1}z, \frac{t}{t+1}c'' + \frac{t}{(t+1)^2}(c'-z)^2\right),$$

with $c = (c', c'')$. Thus, in this case, we take

$$\mathbf{G}(t,v,c,a,z) = \left(v(1+r+a(e^z-1-r)), \frac{t}{t+1}c' + \frac{1}{t+1}z, \frac{t}{t+1}c'' + \frac{t}{(t+1)^2}(c'-z)^2\right).$$

It is also shown in [BCC16] that here the $(1-\alpha)$-confidence region for $(\mu^*, (\sigma^*)^2)$ at time t is an ellipsoid given by

$$\mathbf{\Theta}_t = \tau(t, \widehat{\mu}_t, \widehat{\sigma}^2_t), \quad \tau(t,c) = \left\{c = (c', c'') \in \mathbb{R}^2 \ : \ \frac{t}{c''}(c'-\mu)^2 + \frac{t}{2(c'')^2}(c''-\sigma^2)^2 \le \kappa\right\},$$

where κ is the $(1-\alpha)$ quantile of the χ^2 distribution with two degrees of freedom. Given all the above, the adaptive robust Bellman equations are derived by analogy to (4.4)-(4.5). Namely, $\widetilde{W}_T(c) = \frac{1}{1-\gamma}$ and, for any $t \in \mathcal{T}'$,

$$\begin{aligned}
\widetilde{W}_t(c) = \sup_{a \in A} \inf_{(\mu,\sigma^2) \in \tau(t,c)} \mathbb{E}\Big[&(1+r+a(\mu+\sigma\varepsilon_{t+1}))^{1-\gamma} \\
&\times \widetilde{W}_{t+1}\left(\frac{t}{t+1}c' + \frac{1}{t+1}(e^{\mu+\sigma\varepsilon_{t+1}} - 1 - r), \frac{t}{t+1}c'' + \frac{t}{(t+1)^2}(c' - (\mu+\sigma\varepsilon_{t+1}))^2\right)\Big].
\end{aligned} \tag{4.8}$$

The Bellman equations for the true model, and strong robust method are computed similarly to (4.6) and (4.7).

4.1 Numerical Studies

In this section, we compute the terminal wealth generated by the optimal adaptive robust controls for both cases that are discussed in Section 4. In addition, for these two cases, we compute the

[10] We refer to [BCC16] for precise definition of the projection π, but essentially it is defined as the closest point in the set $\mathbf{\Theta}$.

terminal wealth by assuming that the true parameters μ^* and $(\sigma^*)^2$ are known, and then using respective optimal controls; we call this the true model control specification. We also compute the terminal wealth obtained from using the optimal adaptive controls. Finally, we compute the terminal wealth using optimal robust controls, which, as it turns out, are the same in Case I as the optimal strong robust controls. We perform an analysis of the results and we compare the four methods that are considered.

In the process, we first numerically solve the respective Bellman equation for each of the four considered methods. This is done by backward induction, as usual.

Note that expectation operator showing in the Bellman equation is just an expectation over the standard normal distribution. For all considered control methods and for all simulation conducted, we approximate the standard normal distribution by a 10-points optimal quantizer (see, e.g., [PP03]).

The Bellman equations for the true model control specification and for the adaptive control are essentially the same, that is (4.6), except that, as already said above, in the adaptive control implementation, we use the certainty equivalent approach: at time t the Bellman equation is solved for the time-t point estimate of the parameter μ^* (in Case I), or for the time-t point estimates of the parameters μ^* and $(\sigma^*)^2$ (in Case II).

In the classic robust control application the Bellman equation (4.7) is solved as is normally done in the dynamic min-max game problem, in both Case I and Case II.

In the adaptive robust control application the Bellman equation (4.5) is a recursion on a real-valued cost-to-go function $\widetilde{W}$. In Case I, this recursive equation is numerically solved by discretizing the state space associated with state variable $\widehat{\mu}_t$. In our illustrative example, this discretization of the state space has been done by simulating sample paths of the state process $\widehat{\mu}_\cdot$ up to horizon time T. This simulation has been made under the true model[11]. At each time t, the state space grid has been defined as the collection of sample values taken by this process at time t. Analogous procedure has been applied in Case II with regard to state variables $\widehat{\mu}_t$ and $\widehat{\sigma}^2_t$.

Case I. In order to implement adaptive robust control method for solving the optimal allocation problem, we start by constructing a grid in both time and space. The grid consists of a number of simulated paths of $\hat{\mu}$. Then, we solve equations (4.5) at all grid points for the optimal trading strategies.

As stated above, the implementation of the adaptive control method includes two steps. First, for each μ in the uncertain set $\mathbf{\Theta} = [\underline{\mu}, \overline{\mu}]$, we solve the following Bellman equations:

$$\begin{cases} \widetilde{W}_T = -\frac{1}{1-\gamma}, \\ \widetilde{W}_t = \inf_{a\in A} \mathbb{E}\left[\widetilde{W}_{t+1}(1 + r + a(e^{\mu+\sigma\varepsilon_{t+1}} - 1 - r))^{1-\gamma}\right], \ t \in \mathcal{T}'. \end{cases}$$

It is clear that at each time t, the optimal control is parameterized by μ. With this in mind, at each time $t \in \mathcal{T}'$, we choose the control value with $\hat{\mu}_t$ replacing μ in the formula for the optimal control.

Under classical robust control method, the investor's problem becomes

$$\inf_{\varphi\in\mathcal{A}} \sup_{\theta\in\mathbf{\Theta}} \mathbb{E}_\theta[-u(V_T)] = -v_0^{1-\gamma} \sup_{\varphi\in\mathcal{A}} \inf_{\mu\in[\underline{\mu},\overline{\mu}]} \mathbb{E}\left[\frac{1}{1-\gamma}\left(\prod_{t=0}^{T-1}\left(1 + r + \varphi_t(e^{\mu+\sigma\varepsilon_{t+1}} - 1 - r)\right)\right)^{1-\gamma}\right]. \tag{4.9}$$

[11] Of course, this cannot be done if the control method is applied on genuinely observed market data since the market model is (by nature) not known. One possible solution would be first to estimate the model parameters based on a past history of Z and second to generate sample paths of the state process $\widehat{\mu}_\cdot$ according to this estimated model.

The inner infimum problem in (4.9) is attained at $\mu = \underline{\mu}$. Accordingly, the robust control problem becomes

$$\inf_{\varphi\in\mathcal{A}}\sup_{\theta\in\mathbf{\Theta}}\mathbb{E}_\theta[-u(V_T)] = \frac{v_0^{1-\gamma}}{1-\gamma}\inf_{\varphi\in\mathcal{A}}\mathbb{E}\left[-\left(\prod_{t=0}^{T-1}\left(1+r+\varphi_t(e^{\underline{\mu}+\sigma\varepsilon_{t+1}}-1-r)\right)\right)^{1-\gamma}\right].$$

The corresponding Bellman equation becomes

$$\begin{cases}\widetilde{W}_T = -\frac{1}{1-\gamma},\\ \widetilde{W}_t = \inf_{a\in A}\mathbb{E}\left[\widetilde{W}_{t+1}(1+r+a(e^{\underline{\mu}+\sigma\varepsilon_{t+1}}-1-r))^{1-\gamma}\right],\ t\in\mathcal{T}'.\end{cases} \tag{4.10}$$

We compute the robust optimal strategy by solving equation (4.10) backwards.

It can be shown that for this allocation problem, the strong robust control problem (3.11) is also solved via the Bellman equation (4.10). Hence, in this case, strong robust control method and robust control method provide the same result.

For numerical study we choose the parameter set as $\mathbf{\Theta} = [-1, 1]$, and we consider a set of time horizons $T = 0.1, 0.2, \ldots, 0.9, 1$. The other parameters are chosen as follows

$$V_0 = 100, \quad r = 0.02, \quad \alpha = 0.05, \quad \gamma = 1.5, \quad \sigma = 0.3, \quad \mu^* = 0.06, \quad \widehat{\mu}_0 = 0.02.$$

For every T, we compute the terminal wealth V_T generated by application of the optimal strategies corresponding to four control methods mentioned above: adaptive robust, classical robust, adaptive and the optimal control (assuming the true parameters are known in the latter case). In each method we use 1000 simulated paths of the risky asset and $300T$ rebalancing time steps. Finally, we use the acceptability index Gain-to-Loss Ratio (GLR)

$$\mathrm{GLR}(V) = \begin{cases}\frac{\mathbb{E}_{\theta^*}[e^{-rT}V_T - V_0]}{\mathbb{E}_{\theta^*}[(e^{-rT}V_T - V_0)^-]}, & \mathbb{E}_{\theta^*}[e^{-rT}V_T - V_0] > 0,\\ 0, & \text{otherwise},\end{cases}$$

and 95% Value-at-Risk, $\mathrm{V@R}(V_T) = \inf\{v \in \mathbb{R} : \mathbb{P}_{\theta^*}(V_T + v < 0) \le 95\%\}$, to compare the performance of every method.

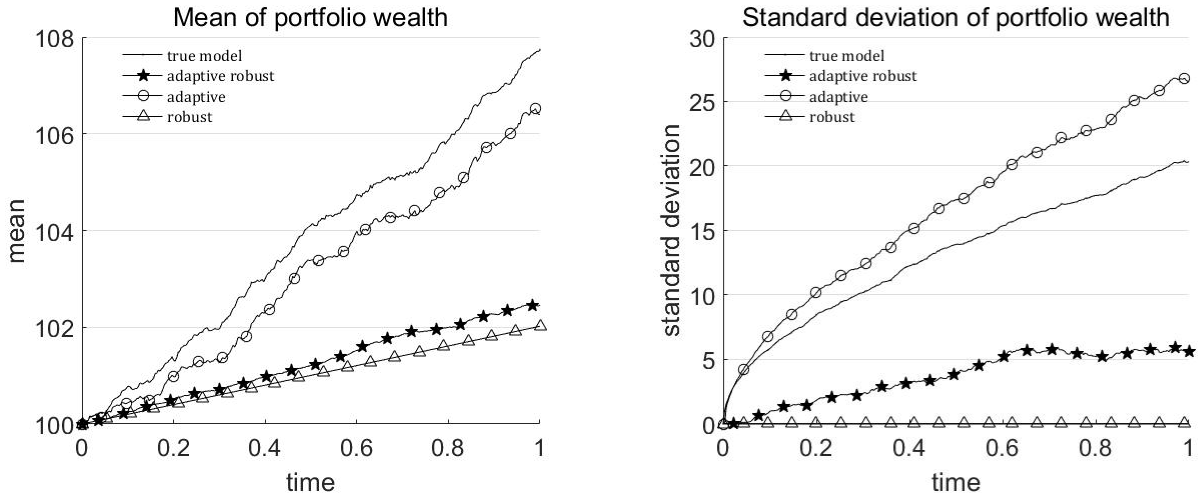


Figure 1: *Time-series of portfolio wealth means, and standard deviations. Unknown mean.*

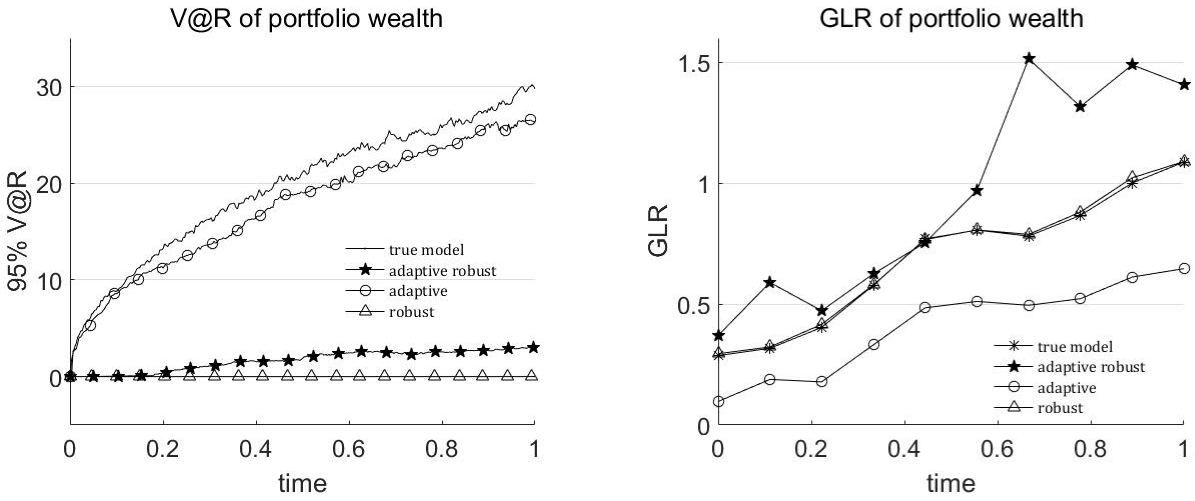


Figure 2: *Time-series of portfolio* V@R *and* GLR*. Unknown mean.*

According to Figure 2, it is apparent that adaptive robust method has the best performance among the considered methods. GLR in case of adaptive robust control is higher than in case of classical robust control and than in case of adaptive control for all the terminal times except for $T = 0.1$. Moreover, even though the adaptive control produces the highest mean terminal wealth (cf. Figure 1), the adaptive control is nevertheless the most risky method in the sense that the corresponding terminal wealth has the highest standard deviation and value at risk. The reason behind such phenomenon is, arguably, that adaptive control method uses the point estimator while solving the optimization problem, so it can be overly aggressive and offers no protection against the estimation error which always exists.

Optimal portfolio wealth corresponding to classical robust method is the lowest among all the four approaches analyzed. This is not surprising since classical robust control is designed to deal with the worst case scenario. Therefore, as illustrated by Figure 3, optimal holdings in the risky asset given by this method are always 0, which means that an investor following the classical robust method puts all the money in the banking account and, thus, fails to benefit from the price rise of the risky asset.

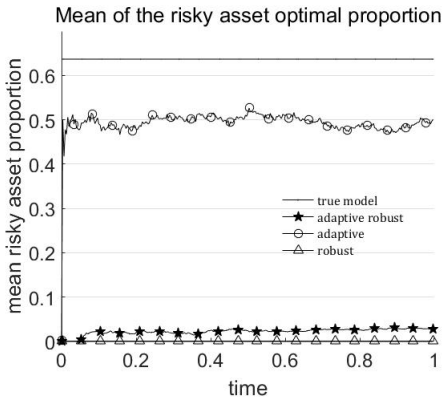


Figure 3: *Time-series of optimal strategies means. Unknown mean.*

Adaptive robust control method is meant to find the right balance between being aggressive

and conservative. As shown in Figures 1 and 2, adaptive robust produces higher terminal wealth than classical robust, and it bears lower risk than adaptive control. The robustness feature of the adaptive robust control method succeeds in controlling the risk stemming from the model uncertainty. Moreover, the learning feature of this method prevents it from being too conservative.
Case II. Here, the adaptive robust control method, the classical robust control method and the adaptive control method need to account uncertainty regarding the true parameter $\theta^* = (\mu^*, (\sigma^*)^2)$.

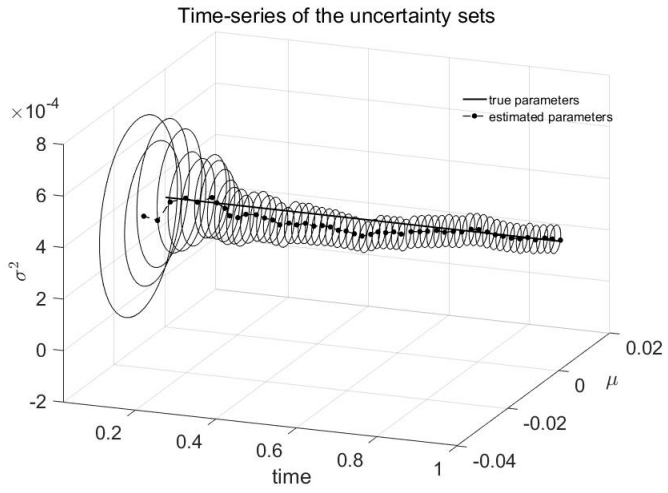


Figure 4: *Times-series of the confidence regions $\tau(t, \widehat{\mu}_t, \widehat{\sigma}_t^2)$ for one particular path of $(\widehat{\mu}_t, \widehat{\sigma}_t^2)$ at confidence level $\alpha = 10\%$.*

We choose the parameter set as $\mathbf{\Theta} = [-1, 1] \times [0, 0.5]$. As in Case I we consider a set of time horizons $T = 0.1, 0.2, \ldots, 0.9, 1$, and $300T$ time iterations (or rebalancing dates) evenly distributed over the time horizon T. Construction of the discretized state space (required in the adaptive control method) and application of the computed optimal strategies are made over 500 sample paths of the true model. The other parameters are chosen as follows

$$V_0 = 100, \quad r = 0.02, \quad \alpha = 0.05, \quad \gamma = 20, \quad \mu^* = 0.08, \quad \sigma^* = 0.30, \quad \widehat{\mu}_0 = 0.1, \quad \widehat{\sigma}_0 = 0.40.$$

Figure 4 shows the evolution of uncertainty sets for a particular sample path of the true model. We can show that the uncertainty on the true parameter (μ^*, σ^*) quickly reduces in size as realized excess returns are observed through time. Moreover, we can notice that for $\alpha = 10\%$, the ellipsoid regions contains the true parameter for nearly all time steps.

As in Case I we compared performance of adaptive robust control, adaptive control, robust control and optimal control, by simulating paths generated by the true model. Figure 5 shows how mean and standard deviation of optimal portfolio wealth values evolve through time. It also displays time-evolution of unexpected loss 95%-VaR and Gain-to-Loss Ratio (GLR). In this 2-dimensional case, the conclusion are in line with the previous case I where only μ were assumed unknown. The adaptive robust strategy outperforms adaptive control and robust control strategies in terms of GLR. Adaptive control strategy gives the highest portfolio wealth mean but at the cost of relatively high standard deviation. The unexpected loss 95%-VaR series is also not in favor of the adaptive control approach.

Finally, we want to mention that optimal holdings in the risky asset in this case exhibit similar behaviour as in Case I.

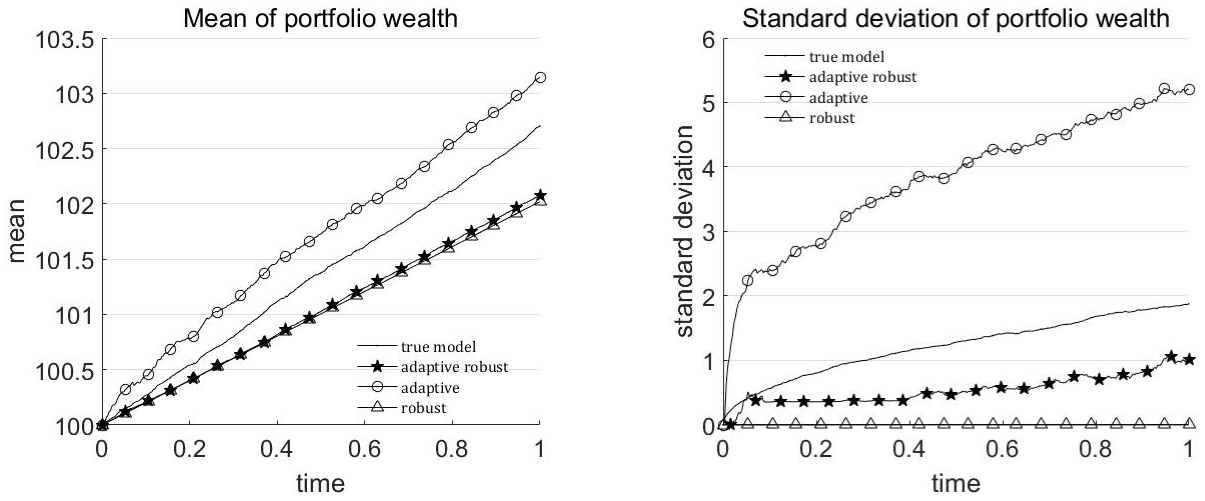


Figure 5: *Time-series of portfolio wealth means, and standard deviations. Unknown mean and variance.*

Acknowledgments

Part of the research was performed while Igor Cialenco was visiting the Institute for Pure and Applied Mathematics (IPAM), which is supported by the National Science Foundation. The research of Monique Jeanblanc is supported by 'Chaire Markets in transition', French Banking Federation and ILB, Labex ANR 11-LABX-0019.

References

[AHS03] E. W. Anderson, L. P. Hansen, and T. J. Sargent. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. *Journal of the European Economic Association*, 1(1):68–123, 2003.

[BB95] T. Başar and P. Bernhard. *H^∞-optimal control and related minimax design problems.* Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, second edition, 1995. A dynamic game approach.

[BCC16] T. R. Bielecki, I. Cialenco, and T. Chen. Recursive construction of confidence regions. 2016.

[BCP16] E. Bayraktar, A. Cosso, and H. Pham. Robust feedback switching control: Dynamic programming and viscosity solutions. *SIAM Journal on Control and Optimization*, 54(5):2594–2628, 2016.

[BGSCS05] M. W. Brandt, A. Goyal, P. Santa-Clara, and J. R. Stroud. A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *The Review of Financial Studies*, 18(3):831, 2005.

[BMS07] G. Bordigoni, A. Matoussi, and M. Schweizer. A stochastic control approach to a robust utility maximization problem. In *Stochastic analysis and applications*, volume 2 of *Abel Symp.*, pages 125–151. Springer, Berlin, 2007.

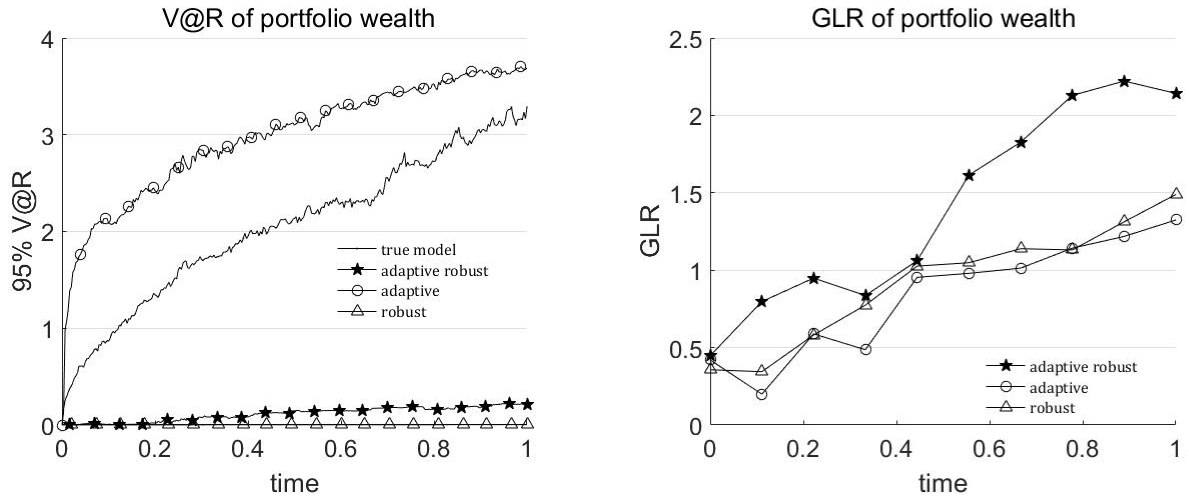


Figure 6: *Time-series of portfolio* V@R *and* GLR*. Unknown mean and variance.*

[BS78] D. P. Bertsekas and S. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Academic Press, 1978.

[BWH12] L. F. Bertuccelli, A. Wu, and J. P. How. Robust adaptive Markov decision processes: planning with model uncertainty. *IEEE Control Syst. Mag.*, 32(5):96–109, 2012.

[CG91] H. F. Chen and L. Guo. *Identification and stochastic adaptive control*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1991.

[GS89] I. Gilboa and D. Schmeidler. Maxmin expected utility with nonunique prior. *J. Math. Econom.*, 18(2):141–153, 1989.

[HS08] P. L. Hansen and T. J. Sargent. *Robustness*. Princeton University Press, 2008.

[HSTW06] L. P. Hansen, T. J. Sargent, G. Turmuhambetova, and N. Williams. Robust control and model misspecification. *J. Econom. Theory*, 128(1):45–90, 2006.

[IS12] P. Ioannou and J. Sun. *Robust Adaptive Control*. Dover Books on Electrical Engineering, 2012.

[Iye05] G. N. Iyengar. Robust Dynamic Programming. *Mathematics of Operations Research*, 30(2):257–280, 2005.

[KOZ14] S. Kallblad, J. Obłój, and T. Zariphopoulou. Time–consistent investment under model uncertainty: the robust forward criteria. *Forthcoming in Finance and Stochastis*, 2014.

[KV15] P. R. Kumar and P. Varaiya. *Stochastic systems: estimation, identification and adaptive control*, volume 75 of *Classics in applied mathematics*. SIAM, 2015.

[LSS06] Andrew E. B. Lim, J. George Shanthikumar, and Z. J. Max Shen. Model Uncertainty, Robust Optimization and Learning. In *Tutorials in Operations Research*, pages 66–94. INFORMS, 2006.

[MMR06] F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498, November 2006.

[PP03] G. Pagès and J. Printems. Optimal quadratic quantization for numerics: the Gaussian case. *Monte Carlo Methods and Applications*, 9:135–166, 2003.

[Sir14] M. Sirbu. A note on the strong formulation of stochastic control problems with model uncertainty. *Electronic Communications in Probability*, 19, 2014.

[Ski03] C. Skiadas. Robust control and recursive utility. *Finance Stoch.*, 7(4):475–489, 2003.