

# Estimating Private Beliefs of Bayesian Agents Based on Observed Decisions

Robert Mattila\*, Inês Lourenço\*, Cristian R. Rojas\*, Vikram Krishnamurthy† and Bo Wahlberg\*

**Abstract**—We consider sequential stochastic decision problems in which, at each time instant, an agent optimizes its local utility by solving a stochastic program and, subsequently, announces its decision to the world. Given this action, we study the problem of estimating the agent’s *private belief* (i.e., its posterior distribution over the set of states of nature based on its private observations). We demonstrate that it is possible to determine the set of private beliefs that are consistent with public data by leveraging techniques from inverse optimization. We further give a number of useful characterizations of this set; for example, tight bounds by solving a set of linear programs (under concave utility). As an illustrative example, we consider estimating the private belief of an investor in regime-switching portfolio allocation. Finally, our theoretical results are illustrated and evaluated in numerical simulations.

**Index Terms**—Estimation; Optimization; Stochastic optimal control

## I. INTRODUCTION

THE classical model for decision making under uncertainty assumes that a Bayesian agent chooses that which provides it with the highest expected utility [1]–[3]. A key component that determines the agent’s decision is its *private belief*: the Bayesian agent employs private information to compute its posterior distribution on the set of states of nature. Additionally, decisions are often sequential and, potentially, distributed between several distinct agents. For example, in social learning (e.g., [4], [5]) agents use noisy observations together with decisions by previous agents to estimate the underlying state of nature (and subsequently, maximize their local expected utilities).

In this paper, we address the question:

*Can an agent’s private belief be estimated given measurements of the agent’s decisions; and if so, how?*

Algorithms for estimating the private belief of an agent are useful since they provide the basis of predicting its future behavior.

For example, consider the behavior of a rational agent in portfolio allocation in regime-switching markets (e.g., [6]–[13]). Recall that a Markowitz-type investor aims to find a

\*This work supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP), the Swedish Research Council under contract 2016-06079 and the U.S. Air Force Office of Scientific Research under grant FA9550-18-1-0007. \*The authors are with the Division of Decision and Control Systems, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden. E-mails: {rmattila, ineslo, crro, bo}@kth.se †The author is with the Department of Electrical Engineering, Cornell Tech, New York, USA. E-mail: vikramk@cornell.edu

portfolio that maximizes its risk-adjusted return. In order to compute this return, a number of random factors have to be estimated (e.g., current market conditions and regime). A crucial difference between an unfortunate and a successful investor is the private information they have access to, and in turn, how accurate their private beliefs describe the true state of nature.

Another area where the present work has implications is Bayesian sequential social learning that is widely studied in economics and social networks [4], [5]: estimating an agent’s private belief and revealing it to all other agents raises important questions of privacy.

In summary, the main contributions of this paper are:

- We determine the set of private beliefs that are consistent with observed data using techniques from inverse optimization (e.g., [14]–[17]) in Theorem 1;
- Given a prior estimate of the private belief, we provide the closest belief (in  $\ell_2$ -sense) that is consistent with public data in Corollary 1;
- In Corollary 2, we derive bounds on the set of private beliefs that are consistent with observed data. Moreover, we give a constructive answer to the question of when an agent’s private belief can be uniquely reconstructed in Lemma 1;
- We specialize our results in a case-study on regime-switching Markowitz portfolio allocation;
- Finally, our theoretical results are illustrated, and their sensitivity is evaluated, in numerical simulations.

The rest of the paper is organized as follows. Below, we give a brief overview of related work. Section II formalizes the sequential decision-making framework. The private belief estimation problems are posed in Section III. Our main results – characterizations of the set of private beliefs that are consistent with public data – are derived in Section IV. Finally, we discuss an example in portfolio selection and provide numerical experiments in Sections V and VI, respectively.

### A. Related work

The theoretical results in this paper draw heavily on general results in inverse optimization (e.g., [14]–[17]). In inverse optimization, the formal goal is to find a minimal perturbation to the coefficients in the objective function of an optimization problem so as to make a given feasible point optimal. An alternative interpretation is that one wishes to reconstruct, or estimate, the parametric objective function that was used to generate an observed solution. The works [14]–[17] treat very general structures (e.g., linear, convex separable, convex). In

the present work, we narrow the scope and study a particular structure of the objective function; that of an expected utility. In this setting, since our parameters represent probabilities, we can derive useful bounds and interpretations.

The inverse optimal control problem – under what cost criteria is a given control law optimal – was studied in the 1960’s by Kalman [18], and is still an active area of research (e.g., [19]). Related disciplines with inverse problems include inverse reinforcement learning in machine learning [20] and revealed preferences in microeconomic theory [21]. The work [22] is complementary to the present, in the sense that the authors estimate the utility function given the decisions and noisy posterior (a form of revealed preferences).

The present work has strong connections to recently proposed inverse filtering problems [23], [24]. In such problems, a sequence of private beliefs is given and the aim is to infer sensor specifications of the filtering system. In [23], [24] it was, however, left as future work how to reconstruct these sequences of private beliefs in the more realistic setting when only actions based on them are observed – which is the topic of the present paper.

## II. DECISION MAKER’S MODEL

In this section, we outline our notation and the decision-making framework. We first consider an agent making a single decision, and then (potentially multiple agents making) sequential decisions. Finally, we detail how the private belief is updated in various settings.

### A. Notation

All vectors are column vectors, unless transposed, and  $\mathbf{1}$  denotes the vector of ones. The element  $i$  of a vector  $v$  is  $[v]_i$ . Inequalities ( $\leq, \geq$ ) between vectors are interpreted elementwise. Lastly,  $\Pr[\cdot]$  denotes the probability of event  $\cdot$ , and  $p(\cdot)$  the probability density.

### B. A Single Choice Under Uncertainty

Denote the underlying state of nature (i.e., all factors external to the decision-making agent) as the random variable  $x \in \mathcal{X}$ , where the state-space  $\mathcal{X}$  can be finite or a subset of Euclidean space. Associated with each state  $x$  is a cost function  $c(x, u)$ , where  $u \in \mathcal{U}$  is the decision variable and  $\mathcal{U}$  is the decision set which, again, can be either finite or a subset of Euclidean space. The agent has access to some private information  $\mathcal{I}$ . A rational agent makes its decision  $u^* \in \mathcal{U}$  based on optimizing its expected cost [1]–[3]:

$$\begin{aligned} u^* \in \arg \min_{u \in \mathcal{U}} \mathbb{E}_x \{ c(x, u) \mid \mathcal{I} \} \\ \text{s.t. } u \in \mathcal{C}, \end{aligned} \quad (1)$$

where  $\mathcal{C} \subset \mathcal{U}$  is the feasible set. The conditional expectation is computed with respect to the agent’s private information (sigma-algebra)  $\mathcal{I}$ . We denote the corresponding conditional distribution as the agent’s *private belief*<sup>1</sup>:

$$\pi(x') = p(x = x' \mid \mathcal{I}) \text{ or } [\pi]_{x'} = \Pr[x = x' \mid \mathcal{I}]. \quad (2)$$

<sup>1</sup>In consistency with the terminology employed in social learning [4], [5] – however, these are closely related to *subjective probabilities* in decision theory (see, e.g., [1]–[3] for further discussions).

Depending on whether  $\mathcal{X}$  is continuous or discrete, respectively,  $\pi(x')$  is either a conditional probability density over continuous  $\mathcal{X}$ , or  $\pi$  is a conditional probability mass vector over finite  $\mathcal{X}$ . Note that, when written out explicitly, for continuous  $\mathcal{X}$ , the optimization problem (1) is

$$\begin{aligned} \min_{u \in \mathcal{U}} \quad & \int_{\mathcal{X}} c(x', u) \pi(x') dx' \\ \text{s.t. } & u \in \mathcal{C}. \end{aligned} \quad (3)$$

### C. Sequential Decisions Under Uncertainty

In a more general framework, the agent does not just make a single decision but acts sequentially multiple times (equivalently, there can be multiple agents that sequentially make the decisions – see below). The state of nature  $x_k \in \mathcal{X}$  is not necessarily static, but can evolve as a random process, and we therefore add a time-subscript  $k$ . Naturally, as time progresses, more information is made available. We model this with a time-dependent information set  $\mathcal{I}_k$ . The sequential decision-making process is as follows:

- 1) New information is made available and the private information is updated to  $\mathcal{I}_k$ .
- 2) The agent uses its private information to update its private belief:

$$\pi_k(x) = p(x_k = x \mid \mathcal{I}_k) \text{ or } [\pi_k]_x = \Pr[x_k = x \mid \mathcal{I}_k], \quad (4)$$

depending on whether  $\mathcal{X}$  is continuous or discrete.

- 3) The agent solves the optimization problem:

$$\begin{aligned} \min_{u_k \in \mathcal{U}} \quad & \mathbb{E}_{x_k} \{ c(x_k, u_k) \mid \mathcal{I}_k \} \\ \text{s.t. } & u_k \in \mathcal{C}. \end{aligned} \quad (5)$$

- 4) An optimal decision  $u_k^*$  – from the set of minimizers of (5) – is made and publicly announced.
- 5) Time  $k$  is increased by one, and we return to step 1.

### D. Updating the Private Belief

Depending on the number of agents making the decisions, and the structure of the information set, there are multiple ways in which the private belief can be updated:

- In the single-agent setting, if  $\mathcal{I}_k$  is a set of observations  $\mathcal{I}_k = \{y_1, \dots, y_k\}$  then optimal filtering techniques can be used by the agent. In particular, if the observation process generating the  $y_k$ :s is linear and Gaussian, then the Kalman filter is optimal (e.g., [25]); and if the state of nature  $x_k$  is discrete, then the *hidden Markov model* (HMM) filter is optimal (e.g., [5], [26]). However, more generally, the private belief is synthesized from expert opinion and subjective beliefs; consider, for example, a doctor judging the likelihoods of various diseases, or the abstract information available to an investor (e.g., rumours and gossip).
- A common modeling framework for the multi-agent settings is social learning [4], [5]. In it, each agent acts once in a predetermined sequential order indexed by  $k = 1, 2, \dots$ . More specifically, at time  $k$ , agent  $k$  receives a private signal (observation)  $s_k$ . The information available

to agent  $k$  is the actions by previous agents (referred to as the *public belief*) along with the agent's private signal  $s_k: \mathcal{I}_k = \{u_1^*, \dots, u_{k-1}^*, s_k\}$ . Standard Bayesian filtering techniques are commonly assumed to be employed by the agent to compute its private belief  $\pi_k$  (of course, depending on the characteristics of the private signal) – see [4], [5]. Following the scheme above, in step 4), the agent broadcasts its decision  $u_k^*$  to all subsequent agents, and it becomes agent  $k+1$ 's turn to act.

### III. FORMULATION OF BELIEF ESTIMATION

In this section, we provide formal statements of the private belief estimation problem. We start with a general formulation (Problem 1), and then restrict the scope to give a more specific formulation (Problem 2).

#### A. General Formulation

Suppose an agent's cost function  $c(x, u)$  as well as its constraints  $\mathcal{C}$  are known; for example, when the costs are related to utilities in a game with public rules, or when they can be reconstructed (using, e.g., revealed preferences [21], [27]) or be estimated (see the case-study in Section V). In social learning, it is typical to assume that all agents share the same preferences [4], [5]. We refer to the costs and constraints, together with the announced actions, as *public data*.

The *private* information  $\mathcal{I}_k$  and the corresponding *private* belief  $\pi_k$  are, typically, not known – *nor are they supposed to be*. The problems we pose below aim to estimate the private belief of the agent.

**Problem 1** (Estimation of Private Beliefs). *Decisions are made according to the procedure specified in Section II-C. The cost functions  $\{c(x, u)\}_{x \in \mathcal{X}}$ , constraints  $\mathcal{C}$  and the decisions  $\{u_k^*\}_k$  of the agent(s) are known. Characterize the private beliefs  $\{\pi_k\}_k$  that are consistent<sup>2</sup> with the public data.*

A solution to this problem provides a basis for predicting the future behavior of an agent. Moreover, estimating an agent's private belief and revealing this belief to all other agents in the framework of social learning raises important questions of privacy.

#### B. Regime-Switching Models

In order to restrict the scope of the paper, we consider a discrete state-space  $\mathcal{X} = \{1, \dots, X\}$  for the state of nature  $x_k$ , and a continuous decision variable  $u_k \in \mathcal{U} \subset \mathbb{R}^U$ . Problems with discrete  $\mathcal{X}$  are known as *regime-switching* problems. In this setting, the decision problem (5) becomes

$$\begin{aligned} \min_{u_k \in \mathbb{R}^U} \quad & \sum_{i=1}^X [\pi_k]_i c(i, u_k) \\ \text{s.t.} \quad & u_k \in \mathcal{C}, \end{aligned} \quad (6)$$

when written out explicitly, where  $\pi_k \in [0, 1]^X$  such that  $\mathbb{1}^T \pi_k = 1$ . For this particular model structure, we can be more precise than in the general formulation (Problem 1):

<sup>2</sup>By *consistent*, we mean that a rational agent would have made its decision  $u_k^*$  under any such private belief.

**Problem 2** (Estimation of Private Beliefs in Regime-Switching Models). *Decisions are made according to the procedure specified in Section II-C. The state-space  $\mathcal{X} = \{1, \dots, X\}$  is discrete and the decision set  $\mathcal{U} \subset \mathbb{R}^U$  is continuous. The cost functions  $\{c(x, u)\}_{x \in \mathcal{X}}$ , constraints  $\mathcal{C}$  and a decision  $u_k^*$  of an agent are known.*

- a) *Determine the set  $\Pi_k$  of private beliefs that are consistent with the public data.*

*Since the set  $\Pi_k$  is, potentially, of a complicated structure, we would like to obtain useful characterizations of it:*

- b) *Provide bounds  $\pi_k \in \mathbb{R}^X$  and  $\bar{\pi}_k \in \mathbb{R}^X$  on the private belief, such that*

$$\pi_k \leq \pi_k \leq \bar{\pi}_k. \quad (7)$$

- c) *Given a prior estimate  $\tilde{\pi}_k \in \mathbb{R}^X$  of the private belief  $\pi_k$ , find the closest private belief  $\hat{\pi}_k$  (in the sense that  $\|\tilde{\pi}_k - \hat{\pi}_k\|_2$  is minimized) that is consistent with the public data (i.e., such that  $\hat{\pi}_k \in \Pi_k$ ).*

Of course, the bounds  $\pi_k = 0 \leq \pi_k \leq \mathbb{1} = \bar{\pi}_k$  solves Problem 2b, however, these are clearly not very informative. Below, we obtain non-trivial bounds.

### IV. ALGORITHMS TO ESTIMATE PRIVATE BELIEFS

This section addresses Problems 2a-c that were formulated above by leveraging results from inverse optimization (e.g., [14]–[17]). Our main result Theorem 1 (stated below) determines the set of consistent private beliefs (i.e., those under which the agent would make decision  $u_k^*$ ) and therefore provides a solution to Problem 2a. From this theorem, solutions to the other two subproblems are derived as corollaries.

#### A. Main Result

The key idea behind our main result is that the *Karush–Kuhn–Tucker* (KKT) conditions (e.g., [28], [29]) for decision problem (6) are necessary and sufficient when:

**Assumption 1.** *For fixed  $x$ , the function  $c(x, u)$  is convex and differentiable in  $u$ . The constraints  $\mathcal{C}$  are affine:*

$$\mathcal{C} = \{u \in \mathbb{R}^U : Au = b, u \geq 0\}, \quad (8)$$

for some  $A \in \mathbb{R}^{N \times U}$  and  $b \in \mathbb{R}^N$ .

Under Assumption 1, we have the following:

**Theorem 1** (Solution to Problem 2a). *Consider the setup in Problem 2 under Assumption 1. The agent that made decision  $u_k^*$  at time  $k$  could have had a private belief  $\pi_k \in \mathbb{R}^X$  if and only if this  $\pi_k$  lies in the set  $\Pi_k$ , specified by:*

$$\Pi_k = \left\{ \pi \in \mathbb{R}^X : \begin{array}{l} \exists \lambda \in \mathbb{R}^U, \nu \in \mathbb{R}^N \text{ s.t.} \\ \pi^T \mathbb{1} = 1, \pi \geq 0, \lambda \geq 0, \\ [\lambda]_i = 0 \text{ if } [u_k^*]_i \neq 0 \\ \quad \text{for } i = 1, \dots, U, \\ \sum_{i=1}^X [\pi]_i \nabla_u c(i, u_k^*) - \lambda + A^T \nu = 0 \end{array} \right\}. \quad (9)$$

*Proof.* The idea is to derive the KKT conditions and consider, not the decision variable  $u_k$ , but instead the private belief  $\pi_k$  as an unknown variable. This idea is adapted from [15], which deals with general conic programs.

The Lagrangian of the agent's decision problem (6) is

$$\mathcal{L}(u_k, \lambda, \nu) = \sum_{i=1}^X [\pi_k]_i c(i, u_k) - \lambda^T u_k + \nu^T (A u_k - b), \quad (10)$$

where  $\lambda \in \mathbb{R}_{\geq 0}^U$  and  $\nu \in \mathbb{R}^N$  are Lagrange multipliers. The corresponding KKT conditions are that any  $u_k^* \in \mathbb{R}^U$  and pair  $(\lambda, \nu)$  that are primal and dual optimal, respectively, must satisfy:

$$u_k^* \geq 0, \quad (11)$$

$$A u_k^* - b = 0, \quad (12)$$

$$\lambda \geq 0, \quad (13)$$

$$[\lambda]_i [u_k^*]_i = 0 \text{ for } i = 1, \dots, U, \quad (14)$$

$$\sum_{i=1}^X [\pi_k]_i \nabla_u c(i, u_k^*) - \lambda + A^T \nu = 0. \quad (15)$$

Since the cost function is convex and the constraints in problem (6) are defined by affine functions (under Assumption 1), constraint qualification (e.g., [28], [29]) guarantees that these equations are in fact also sufficient for optimality.

Hence, a candidate private belief in the simplex  $\{\pi \in \mathbb{R}^X : \pi \geq 0, \pi^T \mathbf{1} = 1\}$  would make the observed decision  $u_k^*$  optimal in (6) if and only if corresponding  $\nu$  and  $\lambda$  exist (i.e., that all together satisfy equations (11)-(15)).  $\square$

**Remark 1.** Our characterization of  $\Pi_k$  in (9), omits (11) and (12) since if the agent generated the decision  $u_k^*$ , clearly, it has to be (primal) feasible. Also, note that the simplex as well as equations (14) and (15) are affine in the private belief and the dual variables, so that the set  $\Pi_k$  is defined entirely by affine functions.

**Remark 2.** Assumption 1 can be relaxed to incorporate general conic constraints without significantly changing the results or the derivations – see, e.g., [15] for details.

### B. Characterizations of $\Pi_k$

We now derive useful characterizations of the set  $\Pi_k$ . In particular, we provide solutions to Problems 2b and 2c.

**Corollary 1** (Solution to Problem 2b). *Given a prior estimate  $\tilde{\pi}_k$  of the private belief, the closest estimate  $\hat{\pi}_k$  that is consistent with public data is given by the convex optimization problem:*

$$\begin{aligned} \min_{\hat{\pi}_k \in \mathbb{R}^X} \quad & \|\hat{\pi}_k - \tilde{\pi}_k\|_2 \\ \text{s.t.} \quad & \hat{\pi}_k \in \Pi_k, \end{aligned} \quad (16)$$

where the set  $\Pi_k$  is detailed in equation (9).

*Proof.* Since the cost function is convex and the constraint set is defined by affine functions, the optimization problem is convex – see also [15, Lemma 1].  $\square$

The second useful characterization is lower-and-upper bounding the agent's private belief componentwise via *linear programs* (LPs):

**Corollary 2** (Solution to Problem 2c). *The agent's private belief is bounded as*

$$\underline{\pi}_k \leq \pi_k \leq \bar{\pi}_k, \quad (17)$$

where the lower bound  $\underline{\pi}_k$  is obtained from the LP

$$[\underline{\pi}_k]_i = \min_{\pi \in \mathbb{R}^X} [\pi]_i \text{ s.t. } \pi \in \Pi_k, \quad (18)$$

for  $i = 1, \dots, X$ , and the upper bound  $\bar{\pi}_k$  by the LP

$$[\bar{\pi}_k]_i = - \min_{\pi \in \mathbb{R}^X} \{-[\pi]_i\} \text{ s.t. } \pi \in \Pi_k, \quad (19)$$

for  $i = 1, \dots, X$ .

*Proof.* Clearly, (18) and (19) define an axis-aligned bounding box of  $\Pi_k$  (e.g., [30]). Since the set  $\Pi_k$  is defined by affine functions and the objective functions are linear, the optimization problems are LPs.  $\square$

### C. Uniqueness

A natural question to ask is: when can the private belief of an agent be reconstructed uniquely? The next lemma provides a constructive answer to this question.

**Lemma 1.** *Consider the setup in Problem 2 under Assumption 1. The private belief  $\pi_k$  of the agent acting at time  $k$  can be uniquely reconstructed if and only if  $\underline{\pi}_k = \bar{\pi}_k$  (computed as in Corollary 2). Moreover, in this case, the bounds all coincide:  $\underline{\pi}_k = \pi_k = \bar{\pi}_k$ .*

*Proof.* First, if the agent's belief  $\pi_k$  is unique (in the sense that it is the only belief that would generate the observed decision  $u_k^*$ ), then this is the only value of  $\pi_k$  for which there exist  $\lambda$  and  $\nu$  such that the KKT conditions (13)-(15) hold. This means that problems (18) and (19) have only a single point (the agent's  $\pi_k$ ) in their feasible sets. Clearly, then, the minimizers must coincide so that  $\underline{\pi}_k = \bar{\pi}_k$ , and thus  $\pi_k = \underline{\pi}_k = \bar{\pi}_k$ .

Second, if  $\underline{\pi}_k = \bar{\pi}_k$ , then (by rewriting (19) as an equivalent maximization problem)

$$\min_{\pi \in \Pi_k} [\pi]_i = \max_{\pi \in \Pi_k} [\pi]_i, \quad (20)$$

for all  $i = 1, \dots, X$ . This implies that the set  $\Pi_k$  is a single point. Now, the agent's private belief  $\pi_k$  is, of course, consistent, so  $\pi_k \in \Pi_k$ . In turn, we have that

$$[\underline{\pi}_k]_i = \min_{\pi \in \{\pi_k\}} [\pi]_i = [\pi_k]_i = \max_{\pi \in \{\pi_k\}} [\pi]_i = [\bar{\pi}_k]_i, \quad (21)$$

for all  $i = 1, \dots, X$ .  $\square$

## V. CASE-STUDY: ESTIMATING THE BELIEF OF A RISK-VERSE INVESTOR BY OBSERVING PORTFOLIO ALLOCATION

This section considers the following inverse estimation problem involving portfolio allocation. Given the utility function and observed portfolio allocation of a risk-averse investor, we estimate the private belief of the investor. This establishes a basis for predicting how the investor operates by observing his or her behavior.

### A. Background

In portfolio allocation, there have been substantial advances in *regime-switching market models* (e.g., [6]–[13]) that take into account that markets (randomly) switch between different states; for example, bull (“optimistic”) and bear (“pessimistic”). These works, in particular [11]–[13], deal with the question: *Given  $U$  assets, how should a fixed amount of capital be invested so as to maximize the risk-adjusted return under switching market conditions?*

Denote the portfolio allocation vector by  $u_k \in \mathbb{R}^U$ , where a fraction  $[u_k]_i$  of the total capital will be invested in asset  $i$ . Usually, one requires that  $\mathbb{1}^T u_k = 1$  (that the full capital is exposed to the market), that  $u_k \geq 0$  (it is only allowed to buy, not sell, assets), and that investments are held for one full time-period. In regime-switching market models, it is assumed that there exists a set of market conditions  $\mathcal{X} = \{1, \dots, X\}$  for the market state  $x_k \in \mathcal{X}$ . Each market state  $x_k$  results in a different mean vector  $\mu_{x_k} \in \mathbb{R}^U$ , and a corresponding covariance matrix  $\Sigma_{x_k} \in \mathbb{R}^{U \times U}$ , for the different assets. For a given risk aversion parameter  $\gamma \in \mathbb{R}_{\geq 0}$  – which quantifies how the investor trades potential return against risk – a regime-switching mean-variance portfolio allocation problem is on the form:

$$\begin{aligned} \min_{u_k \in \mathbb{R}^U} \quad & \mathbb{E}_{x_k} \{ \gamma u_k^T \Sigma_{x_k} u_k - \mu_{x_k}^T u_k \mid \mathcal{I}_k \} \\ \text{s.t.} \quad & \mathbb{1}^T u_k = 1, \quad u_k \geq 0, \end{aligned} \quad (22)$$

whose solution provides the investor with the portfolio giving the optimal risk-adjusted return for period  $k$ . Here,  $\mathcal{I}_k$  is the investor’s *private* information that is employed to compute the posterior distribution of the current market state  $x_k$ .

### B. Estimating the Investor’s Belief

Public stock data allows everyone to form estimates of the expected returns and covariances under different market conditions, meaning that, in practice, the cost functions of a Markowitz-type investor can be approximated (for robustness, see Section VI-B). However, clearly, the success of an investor is closely related to how well he or she can estimate the current market conditions. This estimation depends on *private information*; for example, rumours or privileged information. Reconstructing an investor’s private belief could allow for, e.g., change detection; which could indicate insider trading, and/or reverse engineering trading strategies.

To demonstrate an application of the results in the previous section, assume that we have an initial estimate of the market conditions  $\tilde{\pi}_k$  – this could, for example, be our own estimate, or the investor’s private belief that was reconstructed in the previous time instant  $\hat{\pi}_{k-1}$ .<sup>3</sup> The solution to Problem 2c is then:

**Lemma 2.** *Suppose an investor solves the regime-switching portfolio allocation problem (22), and that we know the expected returns and covariances  $\{\mu_x, \Sigma_x\}_{x=1}^X$ , the risk-aversion parameter  $\gamma$  along with the portfolio allocation  $u_k^*$ .*

<sup>3</sup>Note that  $\hat{\pi}_k$  is only of importance in the case of non-uniqueness.

*The estimate  $\hat{\pi}_k$  of the investor’s private belief  $\pi_k$  that is consistent with public data and that is closest to our prior estimate  $\tilde{\pi}_k$  is the solution of the following convex optimization problem:*

$$\begin{aligned} \min_{\hat{\pi}_k \in \mathbb{R}^X, \lambda \in \mathbb{R}^U, \nu \in \mathbb{R}} \quad & \|\hat{\pi}_k - \tilde{\pi}_k\|_2 \\ \text{s.t.} \quad & \hat{\pi}_k \geq 0, \quad \hat{\pi}_k^T \mathbb{1} = 1, \quad \lambda \geq 0, \\ & [\lambda]_i = 0 \text{ if } [u_k^*]_i \neq 0 \text{ for } i = 1, \dots, U, \\ & \sum_{i=1}^X [\hat{\pi}_k]_i (2\gamma \Sigma_i u_k^* - \mu_i) = \lambda - \mathbb{1} \nu. \end{aligned} \quad (23)$$

*Proof.* Use Corollary 1 (together with Theorem 1), and compute the gradient of the cost function in (22).  $\square$

## VI. NUMERICAL RESULTS

We conduct several numerical experiments to illustrate our theoretical results. In order to visualize the results, we consider synthetic three-regime portfolio allocation problems, as in (22), with  $X = 3$  and  $U = 3$ .

### A. Illustration of the set $\Pi_k$ and the bounds $\underline{\pi}_k, \bar{\pi}_k$

We start with a case where the set of consistent private beliefs  $\Pi_k$  is not a singleton. We computed this set using Theorem 1, and it is depicted as the blue region (line) in the left plot of Fig. 1. As expected, the actual private belief  $\pi_k$  (marked with a green circle) lies inside this set. The shaded red region corresponds to the bounds  $\underline{\pi}_k$  and  $\bar{\pi}_k$ . As expected (and guaranteed by Corollary 2), the actual private belief and the set  $\Pi_k$  lie within the bounds.

Intuitively, the private belief  $\pi_k$  parametrizes the position of the minimizer (assume it is unique) of the unconstrained version of the decision problem. In the example in Fig. 1, some inequality constraint is active which, roughly, projects the unconstrained minimizer onto it. As can be seen in the figure, a whole set of unconstrained minimizers, corresponding to the blue line of private beliefs, is being truncated into the same constrained decision. Since all private beliefs in this region will generate the same (observed) constrained decision, we can not do a unique reconstruction without access to the investor-agent’s associated Lagrangian multipliers (which would give us an idea of how much the constraint is violated by its corresponding unconstrained minimizer).

### B. Sensitivity

The cost functions  $c(x, u)$  incurred by an agent are not always known with certainty (e.g., Section V). To explore the sensitivity of our results, we added random perturbations to the cost functions when estimating the private belief. To illustrate the results, we consider an example where the set  $\Pi_k$  is singleton (i.e., it contains only the actual private belief  $\pi_k$ ). The results of 40 simulations are displayed in the center and right plots in Fig. 1. Random zero-mean Gaussian elementwise perturbations of standard deviations 5% and 10% were added to both the means  $\mu_x$  and covariances  $\Sigma_x$  (see Section V). Clearly, the results are robust to small perturbations since all the estimated private beliefs (blue dots in the figures) lie close to the actual private belief (green circle).

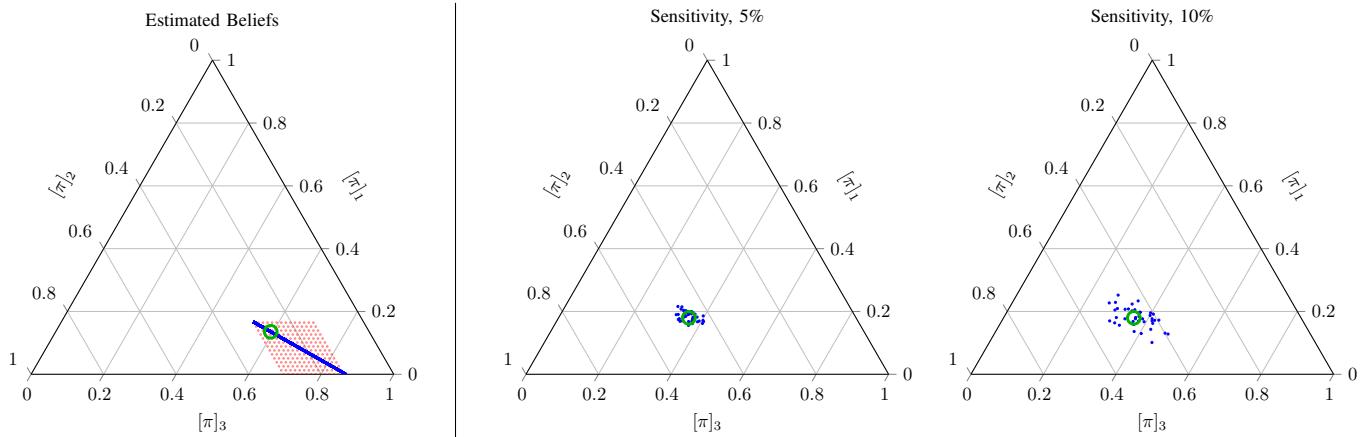


Figure 1: The actual private belief  $\pi_k$  is marked with a green circle (○) in all figures. *Left:* The blue region (—) corresponds to  $\Pi_k$  – the set of beliefs consistent with public data (defined in Theorem 1). The red region (●) displays the bounds  $\underline{\pi}_k$  and  $\bar{\pi}_k$ . *Center and Right:* The blue dots (•) mark the beliefs that are estimated when perturbed cost functions are used. Note that a different example (one in which the set  $\Pi_k$  is singleton) is used compared to that in the left figure.

## VII. CONCLUSION

In this paper, we formulated and proposed algorithms for estimating the private belief of an agent in sequential stochastic decision problems. We exploited results from inverse optimization to characterize the set of private beliefs consistent with public data. Our results have implications in, for example, social learning [4], [5] where they would allow other agent's to gain insight into any agent's private belief – raising important questions of privacy. Finally, we illustrated the results in numerical simulations and demonstrated that they are robust to small uncertainties in the agent's cost functions.

In future work, it is worthwhile deriving perturbation bounds (using, e.g, [31]) on the sets observed in Section VI-B, and connecting the present work with inverse filtering [23], [24].

## REFERENCES

- [1] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic theory*, vol. 1. Oxford university press New York, 1995.
- [2] D. G. Luenberger, *Microeconomic theory*. McGraw-Hill College, 1995.
- [3] M. J. Machina, "Choice under uncertainty: Problems solved and unsolved," *Journal of Economic Perspectives*, vol. 1, no. 1, pp. 121–154, 1987.
- [4] C. Chamley, *Rational herds: Economic models of social learning*. Cambridge University Press, 2004.
- [5] V. Krishnamurthy, *Partially Observed Markov Decision Processes*. Cambridge University Press, 2016.
- [6] V. Krishnamurthy and T. Rydén, "Consistent estimation of linear and non-linear autoregressive models with Markov regime," *Journal of time series analysis*, vol. 19, no. 3, pp. 291–307, 1998.
- [7] T. Rydén, T. Teräsvirta, and S. Åsbrink, "Stylized facts of daily return series and the hidden Markov model," *Journal of Applied Econometrics*, vol. 13, no. 3, pp. 217–244, 1998.
- [8] J. Bulla, "Hidden Markov models with t components. Increased persistence and other aspects," *Quantitative Finance*, vol. 11, no. 3, pp. 459–475, 2011.
- [9] J. G. Dias, J. K. Vermunt, and S. Ramos, "Clustering financial time series: New insights from an extended hidden Markov model," *European Journal of Operational Research*, vol. 243, no. 3, pp. 852–864, 2015.
- [10] E. Kole and D. van Dijk, "How to identify and forecast bull and bear markets?," *Journal of Applied Econometrics*, vol. 32, pp. 120–139, Jan. 2017.
- [11] G. G. Yin and X. Y. Zhou, "Markowitz's mean-variance portfolio selection with regime switching: from discrete-time models to their continuous-time limits," *IEEE Transactions on Automatic Control*, vol. 49, no. 3, pp. 349–360, 2004.
- [12] R. J. Elliott, T. K. Siu, and A. Badescu, "On mean-variance portfolio selection under a hidden Markovian regime-switching model," *Economic Modelling*, vol. 27, no. 3, pp. 678–686, 2010.
- [13] P. Nystrup, H. Madsen, and E. Lindström, "Dynamic portfolio optimization across hidden market regimes," *Quantitative Finance*, vol. 18, pp. 83–95, Jan. 2018.
- [14] R. K. Ahuja and J. B. Orlin, "Inverse optimization," *Operations Research*, vol. 49, no. 5, pp. 771–783, 2001.
- [15] G. Iyengar and W. Kang, "Inverse conic programming with applications," *Operations Research Letters*, vol. 33, pp. 319 – 330, 2005.
- [16] J. Zhang and C. Xu, "Inverse optimization for linearly constrained convex separable programming problems," *European Journal of Operational Research*, vol. 200, no. 3, pp. 671–679, 2010.
- [17] A. Keshavarz, Y. Wang, and S. Boyd, "Imputing a convex objective function," in *IEEE International Symposium on Intelligent Control*, pp. 613–619, Sep. 2011.
- [18] R. E. Kalman, "When is a linear control system optimal," *Journal of Basic Engineering*, vol. 86, no. 1, pp. 51–60, 1964.
- [19] M. C. Priess, R. Conway, J. Choi, J. M. Popovich, and C. Radcliffe, "Solutions to the inverse LQR problem with application to biological systems analysis," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 2, pp. 770–777, 2015.
- [20] A. Ng, "Algorithms for inverse reinforcement learning," in *Proceedings of the International Conference on Machine Learning*, 2000.
- [21] H. R. Varian, "Revealed preference," *Samuelsonian economics and the twenty-first century*, pp. 99–115, 2006.
- [22] A. Caplin and M. Dean, "Revealed preference, rational inattention, and costly information acquisition," *The American Economic Review*, vol. 105, no. 7, pp. 2183–2203, 2015.
- [23] R. Mattila, C. R. Rojas, V. Krishnamurthy, and B. Wahlberg, "Inverse filtering for hidden Markov models," in *Advances in Neural Information Processing Systems*, pp. 4207–4216, 2017.
- [24] R. Mattila, C. R. Rojas, V. Krishnamurthy, and B. Wahlberg, "Inverse filtering for linear Gaussian state-space models," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 5556–5561, 2018.
- [25] B. Anderson and J. Moore, *Optimal Filtering*. Prentice-Hall, 1979.
- [26] O. Cappé, E. Moulines, and T. Rydén, *Inference in Hidden Markov Models*. Springer, 2005.
- [27] H. R. Varian, *Microeconomic analysis*. Norton, 3rd ed., 1992.
- [28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [29] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.
- [30] C. Ericson, *Real-time collision detection*. CRC Press, 2004.
- [31] J. W. Daniel, "Stability of the solution of definite quadratic programs," *Mathematical Programming*, vol. 5, no. 1, pp. 41–53, 1973.